

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<u>http://bmjopen.bmj.com</u>).

If you have any questions on BMJ Open's open peer review process please email <u>info.bmjopen@bmj.com</u>

Study Design and Rationale for the PAASIM Project, a Matched Cohort Study on Urban Water Supply Improvements and Infant Enteric Pathogen Infection, Gut Microbiome Development, and Health in Mozambique

Journal:	BMJ Open
Manuscript ID	bmjopen-2022-067341
Article Type:	Protocol
Date Submitted by the Author:	09-Aug-2022
Complete List of Authors:	Levy, Karen; University of Washington School of Public Health, Department of Environmental and Occupational Health Sciences; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Garn, Joshua; University of Nevada Reno, Division of Biostatistics, Epidemiology and Environmental Health, School of Public Health, Adriano, Zaida; WE Consult Muneme, Bacelar; WE Consult Fagnant-Sperati, Christine; University of Washington School of Public Health, Department of Environmental & Occupational Health Sciences Hubbard, Sydney; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Júnior, Antonio ; WE Consult Manuel, João ; Ministry of Health of Mozambique, Beira Operations Research Center, National Health Institute (INS) Mangamela, Magalhães ; Autoridade Reguladora de Água, Instituto Público McGunegill, Sandy; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Miller-Petrie, Molly; University of Washington School of Public Health, Department of Environmental & Occupational Health Sciences Snyder, Jedidiah ; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Victor, Courtney; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Waller , Lance A. ; Emory University, Department of Biostatistics and Bioinformatics, Rollins School of Public Health Waller , Lance A. ; Emory University, Department of Biostatistics and Bioinformatics, Rollins School of Public Health Brown, Joe; University of North Carolina at Chapel Hill Gillings School of Global Public Health, Environmental Science and Engineering Nala, Rassul; Instituto Nacional de Saúde, Ministry of Health Freeman, Matthew; Emory University, Gangarosa Department of Environmental Health, Environmental Science and Engineering Nala, Rassul; Instituto Nacional de Saúde, Ministry of Health
Keywords:	Gastrointestinal infections < GASTROENTEROLOGY, Public health <

SCHOLARONE™ Manuscripts
SCHOLARONE™ Manuscripts
SCHOLARONE " Manuscripts

Study Design and Rationale for the PAASIM Project, a Matched Cohort Study on Urban Water Supply Improvements and Infant Enteric Pathogen Infection, Gut Microbiome Development, and Health in Mozambique

Short Title: Study Design and Rationale for the PAASIM Matched Cohort Study

Authors

Karen Levy,^{1,2*} Joshua V. Garn,³ Zaida Adriano,⁴ Bacelar de Barros,⁴ Christine S. Fagnant-Sperati,¹ Sydney Hubbard,² Antonio Júnior,⁴ João Luís Manuel,⁵ Magalhães Mangamela,⁶ Sandy McGunegill,² Molly K. Miller-Petrie,¹ Jedidiah S. Snyder,² Courtney Victor,² Lance Waller,⁷ Konstantinos T. Konstantinidis,⁸ Thomas Clasen,² Joe Brown,⁹ Rassul Nalá,¹⁰ Matthew C. Freeman^{2*}

1. Department of Environmental & Occupational Health Sciences, School of Public Health, University of Washington, Seattle, WA, USA.

2. Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA.

3. Division of Biostatistics, Epidemiology and Environmental Health, School of Public Health, University of Nevada—Reno, Reno, NV, USA.

4. WE Consult, Maputo, Mozambique.

5. Beira Operations Research Center, National Health Institute (INS), Ministry of Health, Beira, Mozambique.

6. Autoridade Reguladora de Água, Instituto Público (AURA, IP), Maputo, Mozambique

7. Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA.

8. School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA.

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

10. Instituto Nacional de Saúde (INS), Ministry of Health, Marracuene, Mozambique.

*Corresponding authors. Authors contributed equally.

Karen Levy klevyx@uw.edu; 3980 15th Avenue NE, Seattle, WA 98195

Matthew C. Freeman <u>matthew.freeman@emory.edu</u>; 1518 Clifton Road NE, Atlanta, GA 30322

Word count: 5,452

ABSTRACT

Introduction: Despite clear linkages between provision of clean water and improvements in child health, limited information exists about the health impacts of large water infrastructure improvements in low-income settings. Billions of dollars are spent annually to improve urban water supply, and rigorous evaluation of these improvements, especially targeting informal settlements, is critical to guide policy and investment strategies. Objective measures of infection and exposure to pathogens, and measures of gut function, are needed to understand the effectiveness and impact of water supply improvements.

Methods and analysis: In the PAASIM study, we examine the impact of water system improvements on acute and chronic health outcomes in children in a low-income urban area of Beira, Mozambique, comprising 62 sub-neighborhoods and ~26,300 households. This prospective matched cohort study follows 548 mother-child dyads from late pregnancy through 12 months of age. Primary outcomes include measures of enteric pathogen infections, gut microbiome composition, and source drinking water microbiological quality, measured at the child's 12 month visit. Additional outcomes include diarrhea prevalence, child growth, previous enteric pathogen exposure, child mortality, and various measures of water access and quality. Our analyses will compare a) subjects living in sub-neighborhoods with the improved water to those living in sub-neighborhoods without these improvements; and b) subjects with household water connections on their premises to those without such a connection. This study will provide critical information to understand how to optimize investments for improving child health, filling the information gap about the impact of piped water provision to low-income urban households, using novel gastrointestinal disease outcomes.

Ethics and dissemination: The study was approved by the Emory University Institutional Review Board and the National Bio-Ethics Committee for Health in Mozambique. The preanalysis is published on the Open Science Framework platform (<u>https://osf.io/4rkn6/</u>). Results will be shared with relevant stakeholders locally, and through publications.

STRENGTHS AND LIMITATIONS OF THE STUDY

- This matched cohort study of an urban water supply improvement project will provide critical information about the health impacts of providing piped water and household connections to low-income households.
- We employ rigorous measures of exposure and novel and objective outcome measures, • including gut microbiome composition and molecular detection of enteric pathogens.
- The study design allows for examination of both neighborhood and household-level • effects of water supply improvements.
- As a natural experiment, we are unable to randomize the intervention, leading to potential residual confounding.
- We are unable to examine the impacts of all aspects of the city-wide water improvement project, due to lack of comparable populations, and instead focus only on the low income neighborhoods.

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES)

data mining, AI training, and similar technologies

Protected by copyright, including for uses related to text and

1. INTRODUCTION

Large-scale provision of disinfected, treated drinking water is considered one of the greatest public health achievements of the 20th century[1] and played an important role in improving child health in high-income countries.[2] In low-income countries with high burdens of infectious diseases, inadequate water, sanitation, and hygiene (WASH) conditions are strongly associated with poor child health outcomes, including diarrheal diseases, responsible for >400,000 deaths of children <5 annually,[3,4], and linear growth faltering, underlying 15-17% of mortality of children <5.[5] In Mozambique 27% of stunting is attributed to unimproved water and sanitation.[6]

1.1. Robust studies of the health impacts of community water supply are needed

Rapid urbanization is occurring globally, with urban areas expected to account for 96% of the additional 1.4 billion human population by 2030[7] and 68% of the global population expected to live in urban areas by 2050.[8] While sub-Saharan Africa is still predominantly rural, by 2050 the continent is projected to be 56% urban.[9] To cope with urban growth, expanded infrastructure and services in cities and peri-urban areas will be essential.[7] In 2014 alone, over US\$4.4 billion was committed to WASH in sub-Saharan Africa, with the majority going to improve water supply.[10] Implementation challenges in lower-income settings—such as intermittent service and pathogen intrusion in the distribution system due to pipe breaks, pressure drops, or illegal connections—limit the potential for engineered systems to provide a continuous supply of treated drinking water directly to homes, in adequate quantities to improve hygiene.[11,12] Given the considerable investment in providing piped services to low-income communities, rigorous evaluation of community-scale water provision is critical, to understand the real-world effectiveness and health impact of such systems in low-income contexts.[10,13,14]

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Despite the clear biological link between safe water and child health and development, limited information exists about the health impacts of large water infrastructure improvements in lowincome settings. A small number of studies have evaluated upgrades from intermittent to continuous water delivery in urban areas, [15] or localized improvements to water quality at shared water points.[16] Other studies have evaluated sanitation interventions, without examining drinking water or combined water and sanitation interventions.[17] A recent review of interventions to improve water quality globally found no studies evaluating reliable piped-in water supplies delivered to households and specifically called for rigorous research to assess the health impact of reticulated water supply systems.[18] The review concluded that "there is currently insufficient evidence to know if source-based improvements such as protected wells, communal tap stands, or chlorination/filtration of community sources consistently reduce diarrhoea." A World Health Organization (WHO) review of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings concurs, stating that "evidence from well-conducted intervention studies assessing exclusive use of adequate access and supply of safe water...is still very limited."[19] One reason for this limited evidence is that community-scale interventions are difficult to study using randomized control trial (RCT) methodology – the gold standard for causal inference. It is often infeasible to randomize intervention groups due to policy, planning, and engineering considerations, and lack of adequate comparison groups. As such, alternative quasiexperimental designs must be applied.[20–22]

The predominance of studies in the WASH sector have focused on household- and compoundlevel interventions because they lend themselves more readily to RCT methodology. Outside of the few aforementioned studies of community-wide infrastructure improvements, evaluations of the health impact of water quality improvements in low-income settings often focus on household

water treatment such as boiling, chlorinating or filtering water, with studies predominantly conducted in rural settings.[18] Results of these trials have been mixed, since household-based approaches have various limitations, including low uptake and inconsistent use,[23,24] post-treatment contamination,[25–28] and a poor record of sustained use.[29,30] Household water treatment interventions do not increase water quantity and availability and typical household WASH interventions are likely insufficient to prevent growth faltering in most cases.[31–34]

It is crucial to assess the impact of community-scale infrastructure improvements,[21] as this is an area that is particularly relevant to inform aid agencies, development banks, and other policy makers.[35] The area of most rapid growth in water access is via piped water supply connections, not household water treatment, and larger infrastructure interventions are also critical to achieving the scale of water supply improvements necessary to make impactful changes.[36]

1.2. Objective measures of gut health are needed

A vast majority of WASH studies use as their primary health outcome caregiver-reported diarrhea, primarily because acute diarrheal illness is responsible for ~10-12% of all deaths in children <5.[37,38] However, diarrhea is an unreliable outcome due to courtesy, social desirability and recall bias,[39] local definitions of diarrhea,[40–44] other self-reporting issues,[39,45–47] and the multiple potential etiologies of diarrhea symptoms.[48] Such biases are especially problematic where interventions cannot be blinded as is mainly the case for water interventions. Shedding of enteropathogens, organisms that cause acute gastrointestinal illness, provides an unambiguous indicator of current infection, and increasingly is being used in the WASH field.[31,49,50] Advances in diagnostic techniques make it feasible to test for a wide variety of enteric pathogens simultaneously.[51,52] It is useful to understand enteric pathogen infections because chronic and

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

repeated enteric pathogen infections in the first two years of life—with or without symptomatic diarrhea—are associated with serious morbidities, including gut impairment, growth shortfalls, and cognitive deficits by ages 7-9 years.[53–59] Such outcomes can have profound impact on health, development and well-being of individuals, communities, and entire countries.[60,61] Host-level gastrointestinal conditions affected by environmental determinants, such as gut microbiome composition, may also help explain the long-term sequelae of enteric infections. While there is evidence of differences in gut microbiome composition across different cultures, regions, and populations,[62–67] and environmental conditions,[68] to date specific WASH determinants of these differences, such as access to piped water, have not been evaluated using explicit counterfactuals. Thus measures of gut microbial conditions provide objective outcomes to more accurately measure the effect of WASH interventions[69,70] and capture long-term sequelae,[71] resulting in a more complete understanding of the health impacts.[72]

1.3. Overview of study

In the PAASIM study (*Pesquisa Sobre o Acesso à Água e a Saúde Infantil em Moçambique -Research on Access to Water and Children's Health in Mozambique*), we address a series of questions about the impact of community-level water system improvements on acute and chronic health outcomes in children in a low-income urban area of Mozambique. This matched-control cohort study follows mother-child dyads from late pregnancy through children 12 months of age, examining the impact of living in an area with an improved water network and/or having a household water connection on a variety of aspects of access to drinking water, microbes in the child gut (including both pathogens and other resident gut microbes), and ultimately downstream health outcomes (including diarrhea prevalence and growth) (**Figure 1**). This study will provide Page 9 of 53

BMJ Open

critical information for agencies who seek to understand how to optimize investments for improving child health, helping to fill the information gap about the impact of providing piped water to urban, low-income households by isolating the effects of major community-level water supply improvements on novel gastrointestinal disease outcomes.

2. METHODS & ANALYSIS

2.1. Study setting and description of the intervention

Our study site is the coastal city of Beira, the second largest city in Mozambique (population ~530,000),[73] which serves as a gateway for both the central interior portion of the country and a trade corridor to neighboring land-locked nations. The center of Beira is bordered by unplanned, informal settlements inhabited by over 300,000 low-income residents who often disproportionately feel the impact of severe weather due to lack of infrastructure.[74] A 2018 survey in Beira by Water & Sanitation for the Urban Poor (WSUP) showed that of 5,643 respondents, only 28% had a household water connection, and among those households without a connection, 83% used their neighbor's tap as their main source of water (unpublished data, courtesy of WSUP). Therefore, improvement of water supply and delivery infrastructure is a priority.

The World Bank funded the Water Service & Institutional Support (WASIS-II) Project in 2016 to address the low access to improved water supply in Mozambique,[75] investing \$140 million with the Mozambican public institutions FIPAG (responsible for the public and private investment program in urban water supply systems that serves as the water utility in Beira) and AURA, IP (the water regulatory authority responsible for the economic regulation and consumer protection of service provision). In addition, improvements in Beira are being augmented by investments

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

from other groups, in particular the Dutch government, through infrastructure upgrades as well as emergency response funds following Cyclone Idai in 2019.[76] Improvements in the city of Beira include rehabilitation of water treatment facilities, replacing existing pipe mains that are failing, reticulation of water supply to new areas previously without water service, improving service in areas with poor coverage or low water pressure, and subsidizing water connection fees for the poor.

2.2 Study design

Several aspects of a city-wide water supply improvement project pose challenges to implementing a rigorous epidemiological study, particularly as people living in neighborhoods with piped water often differ in myriad ways from people living in neighborhoods without piped water. Water improvements to communities are often based on the needs or demographics of the community, government or donor priorities, and engineering considerations. Provisions of water supply to new areas previously without water service—or dramatic improvements in access and availability—represent a fundamental development that changes the livability and sometimes the makeup of the community. These issues lead to difficulty in finding a comparable control group for epidemiological comparison. Furthermore, rollouts of water interventions often happen in continuous phases over time, and these changes might coincide with other events or community improvements that similarly impact health, making it difficult for that community to serve as its own control in a pre post design.

By using a prospective matched cohort in this unique context of an ongoing natural experiment, we are able to overcome many of these difficulties. The prospective nature of our study allows better control of confounding through matching, restriction, and rigorous and thoughtful collection Page 11 of 53

BMJ Open

of potential control variables. We specifically focus on one region of the city (**Figure 2**), where some neighborhoods received water system improvements focused on preventing water losses through replacement of the distribution system pipe system in dense, low-income settings; other neighboring areas with similar demographic characteristics did not receive these improvements. Neighborhood-level matching took place in the context of a natural experiment, where the delayed rollout across the city allows us to find and compare intervention and control neighborhoods that are similar in many ways, before the rollout eventually reached all potential control areas.

The Water Loss Reduction Project represents a subset of the improvements being carried out by FIPAG, with co-funding from the Dutch government and the World Bank WASIS-II project. The improvements are designed to reduce illegal connections, thereby increasing the water pressure and quality and increasing the system's capacity for household connections. These areas also received some benefits related to improvements to the water intake and distribution systems. FIPAG undertook a campaign to offer new connections to households. These improvements were completed in some informal settlements in these low income areas of the city in 2019, with other adjacent neighborhoods with similar density and socioeconomic profile slated for completion in future years but not within the timeframe of our study. These specific distribution system upgrades therefore represent a unique opportunity to examine the impacts of community-scale water improvements with neighboring communities who did not receive the intervention serving as control areas for comparison. The neighborhoods under study are also in the lowest income—and therefore highest need areas of the city.

We will perform analyses that take into account the four factorial possible household types, based on mother-child dyad living in a sub-neighborhood with or without the improved water network and with or without a household connection (**Figure 3**). Our primary analyses focus on

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

assessing A) the *total network effect*, by comparing subjects living in sub-neighborhoods with the improved water network (Household Types 1 and 2) to those living in sub-neighborhoods without these improvements (Household Types 3 and 4); and B) the *direct household connection effect*, by comparing subjects with household water connections on their premises (Household Types 1 and 3) to those without a connection (Household Types 2 and 4). Depending on the results of these two primary analyses, secondary analyses may evaluate the other comparisons depicted in **Figure 3**.

The reasons for some neighborhoods receiving the improvements and others not were a result of resource constraints. We conducted a population-based survey (described below) that allowed us to both restrict and match study sub-neighborhoods, thereby creating a statistically appropriate counterfactual for strong internal validity. The evaluation of a real-world intervention delivered in an informal urban setting provides strong external validity for estimating the effects of similar interventions in other low- and middle-income country urban sites. Our study design allows us to isolate the effects of both overall water supply infrastructural improvements as well as the presence of a household water connection. The presence of control areas not receiving upgrades adjacent to intervention areas that are matched on socioeconomic and density variables is unique to this study location. We collect data at multiple timepoints for each study household, allowing us to examine variability in each of the measures taken from each household, rather than at a single point in time, and also allowing for longitudinal analyses of the households and the individual enrolled subjects. We also employ rigorous measures of exposure and novel and objective outcome measures, including gut microbiome composition and molecular detection of enteric pathogens.

2.3. Patient & Public involvement

The executive secretary of AURA, IP was directly involved in the formation of the research questions, and FIPAG personnel were also engaged from the initiation of the project in helping develop the study design. Our team also received input from other public agency stakeholders during workshops that were held prior to initiation of the study. Study subjects and members of the general public were not involved in the study design. We provide regular updates with data summaries to public agency stakeholders, and plan to disseminate the main results to all study participants and also through public presentations for stakeholders in both Beira and Maputo.

2.4. Sub-neighborhood selection

Sub-neighborhood eligibility, selection, and matching of intervention and control subneighborhoods occurred through a two-step process: BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

1) Intervention designation: Our study is a natural experiment, where the investigators had no control over the selection or timing of the intervention implementation. The study flow diagram is shown in **Figure 4**. We worked with FIPAG to determine which neighborhoods in Beira were to receive water distribution system upgrades prior to initiation of enrollment (2020) and before the end of the study (2022). FIPAG provided maps and timelines for construction works related to the upgrades, and the specific areas participating in the water loss reduction project. We also worked with FIPAG and through satellite imagery to identify similarly dense low income areas in Beira that were not slated to receive water network upgrades. A total of 17 potential neighborhoods were considered for inclusion in the study, and neighborhoods were divided into 80 sub-neighborhoods, delineated along natural boundaries such as roads or waterways. "Intervention" sub-neighborhoods include areas not receiving these improvements during the time period of the study. Within both

intervention after completion of our study.

intervention and control sub-neighborhoods, some households have a connection to the water system and others do not. We excluded nine sub-neighborhoods that were likely partially contaminated by proximity to the intervention or that were scheduled to receive the interventions within the timetable of our project; some control sub-neighborhoods are slated to receive the

2) Matching & Restriction: We followed the suggestion of Arnold et al. (2010) to use baseline (preintervention) data at the community level to match intervention to control communities when randomization is not possible.[22] To characterize subneighborhoods for further matching and restriction, we performed a population-based community survey in November-December, 2020 of approximately 1,700 households; this provided approximately a 5% proportional sample of our potential study sub-neighborhoods. We used a random grid sampling approach to estimate household density, using Google Earth satellite imagery, where a grid was placed over an area, and a random selection of squares were selected and counted independently in duplicate, and the number of houses per unit was extrapolated across unsampled squares. The survey contained modules regarding household demographics, water access and practices, sanitation access and practices, household assets and wealth indicators, as well as questions related to COVID-19. A socioeconomic status (SES) score was constructed using the 'simple poverty scorecard'[77] developed specifically for Mozambique, and scores were aggregated at the sub-neighborhood level, and categorized into tertiles.

We matched intervention sub-neighborhoods to control sub-neighborhoods, using coarsened exact matching [78,79] with intervention sub-neighborhoods being matched to control subneighborhoods within the same tertile of both SES and population density. Four neighborhoods (encompassing nine sub-neighborhoods) were found to be outliers in terms of their subPage 15 of 53

BMJ Open

neighborhood-level SES or sanitation, and were excluded from the study sampling frame. Ultimately we designated 36 intervention sub-neighborhoods, with an estimated 16,800 households, and 26 control sub-neighborhoods, with an estimated 9,500 households.

2.5. Participant recruitment, eligibility, and retention

We recruit pregnant women at the last trimester of pregnancy and follow the infant-mother dyads until the child is 12 months old (Figure 5). We selected the first 12 months of life because it is a critical development window, [80-82] it is a time when children are most at risk of acute and chronic effects of enteropathogen infection, [83] and it is a short enough period of time to avoid changes in water access that might occur. We recruit mothers at the end of their pregnancy so we can collect data on household risk factors (including drinking water quality) during the gestational period. Active recruitment occurs through identification of pregnant women in the 2020 population-based survey, lists of pregnant women visiting local health centers for pre-natal care, and study staff visiting under-enrolled sub-neighborhoods throughout the recruitment period. Based on Ministry of Health data for Sofala Province (where Beira is located), virtually all mothers attend pre-natal clinical visits.[84] Passive strategies include referrals of pregnant women by study participants and community leaders. We aim to have complete data on a total of 548 infant-mother dyads, approximately evenly divided between the intervention and control groups. We will continue to enroll dyads into both arms until we reach a minimum of 274 dyads with complete data in each arm, to ensure temporal balance throughout the duration of the study period.

During an initial pre-birth visit, pregnant women are assessed for study eligibility: 1) 18 years or older, 2) in third trimester of pregnancy, 3) resides in enrolled study cluster, 4) not planning to move within the next 12 months, 5) carrying a singleton birth, and 6) consents to take part in the

study. We will re-assess study eligibility at each follow up visit and record if enrolled participants have been lost to follow up.

2.6. Data collection

A local data collection firm (WE Consult) performs the in-country coordination of participant enrollment, data collection, and sample collection. Enumerators conduct household visits before birth for consent, eligibility, and conditions. At months 3, 6, 9, and 12 we deploy survey instruments to collect data on key indicators through structured observations, reports from respondents, and objective measurements (Table S1). We assess a number of variables related to drinking water, including aspects of water quality, water access, water availability, water security, water consumption, and participant satisfaction with water. Brief active surveillance calls also take place monthly by phone to gather information on prenatal and perinatal environmental exposures and illnesses (Figure 5). We ask the caregiver to report diarrhea and blood in the stool (dysentery) of the index child in the previous week at the 3, 6, 9, and 12 month surveys and during active surveillance calls; due to concerns about reporting biases, we also include negative control outcomes.[85] At each post-birth visit we measure child: 1) length, weight, and head circumference, and calculate length-for-age and weight-for-age Z-scores. Prevalence of stunting and underweight are defined as two standard deviations below median of the reference population.[86] All data are collected on electronic tablets using Open Data Kit (ODK) Collect, an open-source program which allows offline data collection on a mobile device.[87] Additional details are provided in the Supplementary Material.

2.7. Sample collection, processing, and analysis

BMJ Open

We briefly describe sample collection and downstream processing and analysis here, with additional details provided in the Supplementary Material.

2.7.1. Stool

Stool of the index child is collected at months 3, 6, 9, and 12. Three aliquots are placed in temperature stable lysis buffer collection tubes, and two additional aliquots are used to prepare a slide for Kato-Katz analysis of parasite ova[88]. Eligible participants are referred for deworming medicine at the 12-month visit, after returning results of the parasitological exam to study subjects in collaboration with Instituto Nacional de Saúde (INS) staff in Beira.

Extracted nucleic acids are analyzed: (1) using the TaqMan Array card (TAC, ThermoFisher Scientific, Waltham, MA, USA) assay, which allows quantification by real-time PCR via a 384well microfluidic card for simultaneous detection of multiple viral, bacterial, and parasitic enteric pathogen targets as well as antimicrobial resistance genes,[89] customized for our targets of interest (**Table S2**); and (2) by sequencing of the V4 region of the 16S ribosomal RNA (rRNA) gene amplicon to characterize gut microbiome community structure and composition. Bioinformatic analyses will be completed using the QIIME2 software platform.[90]

2.7.2. Dried blood spots

Sample collection: A trained nurse or laboratory technician collects up to six dried spots of capillary blood of the index child at 6, 9, and 12 month visits on Tropbio Filter Paper Blood Collection Disks (Cellabs, Sydney, Australia), using a 2mm lancet. Samples are stored at -20 °C and shipped at ambient temperature.[91] We will use the Luminex platform to carry out high throughput, multiplex antibody assays that enable the simultaneous measurement of quantitative

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

antibody responses to dozens of pathogens from a single blood spot.[92] Our first measure will occur at 6 months, to avoid detection of maternal antibodies that wane over the first 3-6 months of life.[93]

2.7.3. Drinking water

Sample collection: We collect 100-mL household drinking water samples from source and stored water at all household visits. To complement the household sampling, we collect samples from a selection of 45 public standpipes located within the study area and 55 additional public standpipes located elsewhere in the city of Beira. At public standpipes we also measure water pressure by measuring time to fill a fixed volume (1L or 5L, depending on the pressure). Samples are processed for fecal indicator bacteria within six hours of collection using Colliert-18 reagent and the Quanti-Tray/2000 MPN method (IDEXX Laboratories, Westbrook, ME, USA), as well as for free and total chlorine levels and additional physiochemical parameters (pH, conductivity, and turbidity). Large volume samples will be collected from a subset of 50 households (1 L, processed by membrane filtration) and 25 public standpipes (50 L, processed by dead end ultrafiltration[94]) in two different seasons, and tested for enteropathogens using the TAC assay.

2.8. Outcomes

Our primary outcomes include: any bacteria or protozoa infection at age 12 months after birth; individual pathogens or pathogen groups; child gut microbiome composition; and household source water quality. While we measure viral pathogens using the TAC assay, they will be excluded from the combined enteropathogen prevalence primary outcome measure, because waterborne transmission is unlikely to dominate for these viral pathogens.[95–98] In addition to

Page 19 of 53

BMJ Open

the aforementioned reasons related to child development and infection risk, measuring pathogens at 12 months will give us the greatest power to detect a difference, given higher levels of infection at that age than in younger children. We will measure gut microbiome using 16S rRNA gene amplicon sequencing in the full sample at 12 months and in a random subset of 200 children with complete data at 3, 6, and 9 months, evenly distributed between intervention and control groups; dyads eligible for sub-set sampling will include those with complete stool sample collection and unchanged intervention exposure conditions. The 12-month samples will allow us to compare all study children at a common time, when all children are consuming drinking water and once the gut microbiome has become relatively established;[99] the longitudinal samples will allow for comparison of development of the microbiome over time between the two groups. Microbiome outcomes include alpha and beta diversity metrics, and identification of enriched taxonomic groups. We also include household source water quality as a primary exposure outcome, as understanding whether exposure to microbial contaminants is altered is considered a critical aspect of evaluation of WASH projects.[70,100] BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Additional non-primary outcomes include pathogen count, pathogen community similarity (measured using Jaccard similarity index), diarrhea, child growth, and prior enteropathogen infection (measured using serology on dried blood spot samples). We will measure additional water quality exposure measures, as well as measures of exposure to the improved water system, such as fidelity of the intervention (e.g., improvements to water quantity and coverage of household taps) and receipt of the intervention by community members (e.g., reductions in water insecurity, increased water use). These fidelity and uptake measures will be collected at all time points through direct observation and respondent report. Available minimal detectable effect sizes are summarized in **Table 1** and calculations are further detailed in the Supplementary Material.

Table 1. Primary and non-primary health outcomes and exposure outcomes for the PAASIM Study. Patent enteric pathogen infection in stool is measured via TaqMan Array Card (TAC) assay; Stool microbiome composition is measured via 16s rRNA amplicon sequencing; prior enteric pathogen exposure is measured via serological assays of dried blood spots. Drinking water quality measured by IDEXX as *E. coli* most probable number/100mL. See text for further details. Calculations and additional values for the minimum detectable effect are described in the Supplementary Material.

Primary outcomes	Time point	Minimum detectable effect (Risk Ratio)	Anticipated cont group prevalenc
Patent enteric nathogen infection		(Hisk Ratio)	
Prevalence of any bacterial or protozoan nathogen	12 mo ^b	0 74	70%
Prevalence of any bacterial nathogen	12 mo^{b}	0.69	61%
Prevalence of any protozoan nathogen	12 mo^{b}	0.50	32%
Any co-infection (bacterial protozoan or viral pathogens)	12 mo^{b}	0.50	45%
Any co-infection (bacterial, protozoan, or vital pathogens)	12 110	0.00	4570
Gut microbiome composition	12 mo ^c		
Alpha diversity	12 1110		NA
Beta diversity			NA
Enriched taxa			NA
			1111
Household source drinking water quality	12 mod		NA
Trousenerd source armining water quality	12 110		1111
Additional health outcomes			
Patent enteric pathogen infection			
Pathogen count	12 mo ^b		
Pathogen community similarity	12 mo^{b}		
Individual nathogens	12 mo ^b	n/a-0.49	2-31%
Any virus	12 mo^{b}	0 46	28%
	12 1110	0.10	2070
Gut microbiome composition	3. 6. 9 mo ^c		
(Same variables as gut microbiome composition for 12 mo)	.,.,.		
Prior enteric pathogen exposure	6, 9, 12 mo		
(Same variables as patent enteric pathogen infection)			
Diarrhea 1-week period prevalence (caregiver report)	Weekly	0.26	14.4%
Anthropometric Measurements	3, 6, 9, 12 mo		
Length-for-age Z-score			
Weight-for-age Z-score		-	
Stunting prevalence ^e		0.49	31%
Underweight prevalence ^e		0.22	13%
All-cause mortality (while enrolled in study)	Continuous		
Additional Exposure outcomes			
Primary drinking water source	3, 6, 9, 12 mo		
Drinking water quality (source)	3, 6, 9 mo ^f		
Drinking water quality (stored)	3, 6, 9, 12 mo		
Water access	3, 6, 9, 12 mo		
Water availability	3, 6, 9, 12 mo		
Water security	3, 6, 9, 12 mo		
Water consumption	3, 6, 9, 12 mo		
User satisfaction with water	3, 6, 9, 12 mo		

^b Samples also collected at 3, 6, and 9 months of age may also be analyzed, depending on results of primary analysis at 12 months.

⁶ A subset of n=200 samples will be analyzed for gut microbiome composition in children at 3, 6, and 9 months of age. All 12 month samples will be analyzed.

Samples will be analyzed.
^d This is a conservative estimate as it does not account for weekly active surveillance.

^c Defined as two standard deviations below median of the reference population.

^f Samples will be analyzed at 3, 6, 9, and 12 mo, but 12 month samples are the primary outcome of interest.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1

2

BMJ Open

2.9. Analysis plan

The pre-analysis plan for this study is published on the Open Science Framework platform (https://osf.io/4rkn6/).

2.9.1. Total network effect

To assess the impact of the intervention on our primary enteric pathogen infection outcomes and water quality exposure outcome (**Table 1**), we will use an intention-to-treat (ITT) analysis approach to compare children living in intervention versus control sub-neighborhoods, without regard to uptake/use of the intervention (i.e., direct household connection on the premises). We will use multivariable log-linear binomial regression models, as pathogen infection is a binary variable, and will use generalized estimating equations (GEE) to account for clustering at the subneighborhood level. We group matched on sub-neighborhood-level SES and population density, using weighting to account for unequal numbers between the intervention and control areas within each matching stratum.[101] We will additionally control for household- and individual-level confounders, including household SES, household sanitation, mother's education-level, and child sex. We may adjust for additional variables if there are found to be imbalances in potential confounders in our baseline assessment. We hypothesize that the intervention will lead to reductions in enteric pathogens among children and microbial water contamination of source water.

For additional outcomes and exposure variables of interest (**Table 1**), we will use a similar modeling approach, using log-linear binomial regression models for binary outcomes, linear regression models for continuous outcomes, and Poisson (or negative binomial) models for count outcomes. For outcomes measured at multiple time points, we will present results separately for each given time point. For these analyses, we will control for sub-neighborhood-level SES and

population density through matching, and will additionally control for household sanitation, mother's education-level, child sex, and any other variables that are imbalanced and are conceivably potential confounders. For previous enteropathogen exposure evaluated using serological measures we hypothesize that those in the intervention group will show delays in pathogen acquisition. To assess the impact of the intervention on microbiome outcomes, we will evaluate alpha diversity (Chao1 species richness estimator, Pielou's *J* evenness estimator, and the Shannon

diversity (Chao1 species richness estimator, Pielou's *J* evenness estimator, and the Shannon diversity index[102]) using the same modeling approach as described above for continuous outcomes. Linear discriminate effect size (LEfSe) analyses will be used to evaluate specific 16S rRNA gene-based Operational Taxonomic Units (OTUs) that differ between individuals in intervention versus control groups, and will include effect size corrections[103]. We will examine the impact of intervention group, controlling for other covariates, on community similarity using Adonis permutation models,[104] based on weighted UniFrac and Bray-Curtis distances, and evaluate and visualize differences using PCA and/or NMDS plots. We hypothesize that we will be able to observe detectable differences in gut microbiome composition in children living in intervention versus control sub-neighborhoods and we will report these differences at the individual OTU and bacterial family levels.

2.9.2. Direct Household Connection Effect

To assess the effect of having a water connection at the household or compound, we will use models similar to those described above, but accounting for a household network connection. We will also assess the interaction between the household and neighborhood network variables, which will allow us to contrast and estimate indirect, direct, and total effects, as shown in **Figure 3**. We

BMJ Open

hypothesize that participants with both improved water networks in their sub-neighborhoods and household water connections will most benefit from the interventions in terms of our primary and non-primary health outcomes and exposure outcomes of interest.

2.9.3. Additional analyses

For select primary outcomes, we will assess if there is effect modification by a third variable, such as follow-up round/age, participant sex, and household sanitation access. We will use interaction terms to identify potential interactions, and will present stratified results (e.g., separately by sex) if interactions are detected. The intervention status of subneighborhoods was set at baseline, but if control subneighborhood(s) receive the intervention after the study has started, we will perform sensitivity analyses dropping and/or recategorizing subneighborhood(s) that crossed over.

There are several analyses where we do not use the matched-design and intervention variable in our analyses. For example, we will assess associations between various water measures on health, without regard to the intervention designation. We will also examine changes in the gut microbiomes of children over time. Additional analyses will be described and documented in OSF.

2.10. Sample size and power calculations

Our minimal sample size of 548 households-half in intervention and half in matched control sub-neighborhoods-was powered for our primary outcome of prevalence of any non-viral pathogen. Utilizing data from the MapSan trial for children 10-14 months of age (J. Knee, *pers comm*) we used a control group prevalence of 70% for any non-viral pathogen, and estimated the ability to detect a relative risk of 0.74, alpha=0.05, and power=80% using a two-sided test for

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

significance.[31] We estimated a sub-neighborhood-level interclass correlation (ICC) of 0.05 (a moderate estimate) among our 62 designated sub-neighborhoods. We will also report on the final ICC and other assumptions of this power analysis at the end of the study. Estimates of minimum detectable effect sizes based on control prevalence of the outcome of interest (**Table S2**) show we may be adequately powered to detect a difference in some individual pathogens if those pathogens have high prevalence and/or if they are strongly associated with the water supply improvement intervention (e.g., waterborne pathogens). We target planned recruitment at 900 pregnant women in the third trimester, to account for incomplete data and loss to follow-up. We used sub-neighborhood enrollment targets proportionate to our density estimates to achieve balance across intervention and control sub-neighborhoods.

2.11. Blinding

All lab personnel and field enumerators are blinded to the intervention status of the samples and households. Participants cannot be blinded to their household-level water exposure status or cluster-level exposure status, although participants may or may not know about water improvements in their particular neighborhood. A primary analyst external to the core data management team is blinded to the group assignments until the data cleaning and primary analysis are completed. Details of these procedures are included in the Supplementary Material. Unblinding will occur only after primary outcome models are developed and compared between two independent analysts. Analyses examining the impact of the intervention on non-primary outcomes or exposures of interest will not be unblinded until after analyses that examine the impact of the intervention on our primary outcomes have been completed. Purely observational analyses that do not require information on intervention group may be completed before unblinding occurs.

<u>3. ETHICS & DISSEMINATION</u>

The study protocol, informed consent forms, and data collection tools were approved by 1) Mozambique National Bio-Ethics Committee for Health (IRB00002657) and (2) Emory University's Institutional Review Board (IRB00098584). As this study is a natural experiment that the investigators do not control, we do not have a data monitoring committee or any interim stopping guidelines. Enrollment for this study began during the COVID-19 pandemic, and precautions were taken to secure the safety of study staff and participants based on guidance from INS, Emory University, and the University of Washington.

Any changes to this published protocol will be noted in OSF, and, where relevant, in future publications. De-identified data sufficient to replicate study findings will be publicly available on OSF upon completion and publication of the study results. A report will also be prepared and shared with the municipality and health authorities in Beira, and other relevant stakeholders. All microbial DNA sequence data will be made available through the SRA database of NCBI upon validation and/or publication of the corresponding manuscript.

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

REFERENCES

- 1 Centers for Disease Control and Prevention. Achievements in Public Health, 1900-1999: Control of Infectious Diseases. *MMWR Morb Mortal Wkly Rep* 1999;**48**:621–9.
- 2 Cutler D, Miller G. The role of public health improvements in health advances: the twentieth-century United States. *Demography* 2005;**42**:1–22.
- 3 Troeger C, Blacker BF, Khalil IA, *et al.* Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* 2018;18:1211-1228. PMCID: PMC6202444. doi:10.1016/S1473-3099(18)30362-1
- 4 Pruss-Ustun A, Bartram J, Clasen T, *et al.* Burden of disease from inadequate water, sanitation and hygiene in low- and middle-income settings: a retrospective analysis of data from 145 countries. *Trop Med Int Health* 2014;**19**:894–905. doi:10.1111/tmi.12329
- 5 Black RE, Victora CG, Walker SP, *et al.* Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 2013;**382**:427–51. doi:10.1016/S0140-6736(13)60937-X
- 6 Danaei G, Andrews KG, Sudfeld CR, *et al.* Risk Factors for Childhood Stunting in 137 Developing Countries: A Comparative Risk Assessment Analysis at Global, Regional, and Country Levels. *PLoS Med* 2016;**13**:e1002164. doi:10.1371/journal.pmed.1002164
- 7 World Bank/Interational Monitary Fund (IMF). Global Montioring Report 2013: Rural-Urban Dynamics and the Millennium Development Goals. Washington, D.C.: World Bank/IMF 2013.
- 8 United Nations Department of Economic and Social Affairs. *World Urbanization Prospects: The 2018 Revision.* United Nations 2019. doi:10.18356/b9e995fe-en
- 9 United Nations-HABITAT. *The State of African Cities: Re-imagining sustainable urban transitions*. 2014. http://unhabitat.org/books/state-of-african- cities-2014-re-imagining-sustainable-urban-transitions/
- 10 United Nations. Investing in Water and Sanitation: Increasing Access, Reducing Inequalities. GLAAS 2014 Findings-- Special report for Africa. UN Water & World Health Organization 2015. http://www.who.int/water_sanitation_health/glaas/glaas2014-africaregion.pdf
- 11 Bain R, Cronk R, Wright J, *et al.* Fecal contamination of drinking-water in low- and middle-income countries: a systematic review and meta-analysis. *PLOS Med* 2014;**11**:e1001644. doi:10.1371/journal.pmed.1001644
- 12 Hunter PR, Zmirou-Navier D, Hartemann P. Estimating the impact on health of poor reliability of drinking water interventions in developing countries. *Sci Total Environ* 2009;**407**:2621–4. doi:10.1016/j.scitotenv.2009.01.018

- 2021). 109X(19)30315-8 doi:10.1093/ije/dyz157 2015. doi:10.1002/14651858.CD004794.pub3 doi:10.1016/j.amepre.2007.04.007
 - GLAAS 2018/2019 country survey. World Health Organization 2018. https://www.who.int/publications/m/item/glaas-2018-2019-country-survey (accessed 1 Oct
 - UN-Water. UN-Water GLAAS 2019: National systems to support drinking-water, sanitation and hygiene - Global status report 2019. 2019. https://www.unwater.org/publications/un-water-glaas-2019-national-systems-to-supportdrinking-water-sanitation-and-hygiene-global-status-report-2019/ (accessed 1 Oct 2021).
 - Ercumen A, Arnold BF, Kumpel E, et al. Upgrading a piped water supply from intermittent to continuous delivery and association with waterborne illness: A matched cohort study in urban India. PLOS Med 2015;12:e1001892. doi:10.1371/journal.pmed.1001892
 - Pickering AJ, Crider Y, Sultana S, et al. Effect of in-line drinking water chlorination at the point of collection on child diarrhoea in urban Bangladesh: a double-blind, clusterrandomised controlled trial. Lancet Glob Health 2019;7:e1247-56. doi:10.1016/S2214-
 - Reese H, Routray P, Torondel B, et al. Assessing longer-term effectiveness of a combined household-level piped water and sanitation intervention on child diarrhoea, acute respiratory infection, soil-transmitted helminth infection and nutritional status: a matched cohort study in rural Odisha, India. Int J Epidemiol 2019;48:1757-67.
 - Clasen TF, Alexander KT, Sinclair D, et al. Interventions to improve water quality for preventing diarrhoea. Cochrane Database Syst Rev Published Online First: 20 October
 - Wolf J, Pruss-Ustun A, Cumming O, et al. Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and meta-regression. Trop Med Int Health 2014;19:928-42. doi:10.1111/tmi.12331
 - Sanson-Fisher RW, Bonevski B, Green LW, et al. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med 2007;33:155-61.
 - Spears D, Ban R, Cumming O. Trials and Tribulations: The Rise and Fall of the RCT in the WASH Sector. In: Randomized Control Trials in the Field of Development. Oxford: : Oxford University Press 2020. doi:10.1093/oso/9780198865360.003.0008
 - Arnold BF, Khush RS, Ramaswamy P, et al. Causal inference methods to study nonrandomized, preexisting development interventions. Proc Natl Acad Sci 2010;107:22605-10. doi:10.1073/pnas.1008944107
 - Boisson S, Schmidt WP, Berhanu T, et al. Randomized controlled trial in rural Ethiopia to assess a portable water treatment device. Environ Sci Technol 2009;43:5934–9.

- 24 Luoto J, Mahmud M, Albert J, *et al.* Learning to dislike safe water products: results from a randomized controlled trial of the effects of direct and peer experience on willingness to pay. *Environ Sci Technol* 2012;**46**:6244–51. doi:10.1021/es2027967
 - 25 Levy K, Nelson KL, Hubbard A, *et al.* Following the water: a controlled study of drinking water storage in northern coastal Ecuador. *Environ Health Perspect* 2008;**116**:1533–40. doi:10.1289/ehp.11296
 - 26 Wright J, Gundry S, Conroy R. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. *Trop Med Int Health* 2004;**9**:106–17.
 - 27 Levy K, Anderson L, Robb KA, *et al.* Household effectiveness vs. laboratory efficacy of point-of-use chlorination. *Water Res* 2014;**54**:69–77. doi:10.1016/j.watres.2014.01.037
 - 28 Stauber CE, Kominek B, Liang KR, *et al.* Evaluation of the impact of the plastic BioSand filter on health and drinking water quality in rural Tamale, Ghana. *Int J Environ Res Public Health* 2012;**9**:3806–23. doi:10.3390/ijerph9113806
- 29 Brown J, Clasen T. High adherence is necessary to realize health gains from water quality interventions. *PLoS One* 2012;7:e36735. doi:10.1371/journal.pone.0036735
- 30 Enger KS, Nelson KL, Rose JB, *et al.* The joint effects of efficacy and compliance: a study of household water treatment effectiveness against childhood diarrhea. *Water Res* 2013;47:1181–90. doi:10.1016/j.watres.2012.11.034
- 31 Knee J, Sumner T, Adriano Z, *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. doi:10.7554/eLife.62278
- 32 Gough EK, Moulton LH, Mutasa K, *et al.* Effects of improved water, sanitation, and hygiene and improved complementary feeding on environmental enteric dysfunction in children in rural Zimbabwe: A cluster-randomized controlled trial. *PLoS Negl Trop Dis* 2020;**14**:e0007963. doi:10.1371/journal.pntd.0007963
- 33 Null C, Stewart CP, Pickering AJ, 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. doi:10.1016/S2214-109X(18)30005-6
- 34 Luby SP, Rahman M, Arnold BF, *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. doi:10.1016/S2214-109X(17)30490-4
- Post A, Ray I. Hybrid Modes of Urban Water Delivery in Low- and Middle-Income Countries. Oxf. Res. Encycl. Environ. Sci. 2020. doi:10.1093/acrefore/9780199389414.013.679

2		
3 4 5 6 7	36	World Health Organization (WHO) and the United Nations Children's Fund (UNICEF). Progress on household drinking water, sanitation and hygiene 2000-2020: five years into the SDGs. Geneva: : World Health Organization (WHO) and the United Nations Children's Fund (UNICEF) 2021.
8 9 10 11	37	Walker CL, Rudan I, Liu L, <i>et al.</i> Global burden of childhood pneumonia and diarrhoea. <i>Lancet</i> 2013; 381 :1405–16. doi:10.1016/S0140-6736(13)60222-6
12 13 14 15	38	Liu L, Johnson HL, Cousens S, <i>et al.</i> Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. <i>Lancet</i> 2012; 379 :2151–61. doi:10.1016/S0140-6736(12)60560-1
16 17 18 19	39	Arnold BF, Galiani S, Ram PK, <i>et al.</i> Optimal recall period for caregiver-reported illness in risk factor and intervention studies: a multicountry study. <i>Am J Epidemiol</i> 2013; 177 :361–70. doi:10.1093/aje/kws281
20 21 22 23	40	Baqui AH, Black RE, Yunus M, <i>et al.</i> Methodological issues in diarrhoeal disease epidemiology. Definition of diarrhoea episodes. <i>Int J Epidemiol</i> 1991; 20 :1057–63.
24 25 26	41	Wright J, Gundry SW, Conroy RM. A review of changing episode definitions and their effects on estimates of diarrhoeal morbidity. <i>J Health Popul Nutr</i> 2007; 25 :448–55.
27 28 29 30	42	Wright JA, Gundry SW, Conroy R, <i>et al.</i> Defining episodes of diarrhoea: results from a three-country study in Sub-Saharan Africa. <i>J Health Popul Nutr</i> 2006; 24 :8–16.
30 31 32 33 34	43	Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in Guatemala: determining the optimal length of recall. <i>Epidemiol Infect</i> 2010; 138 :264–9. doi:10.1017/S0950268809990422
35 36 37 38	44	Trostle A, Yépez-Montufar J, Corozo-Ángulo B, <i>et al.</i> Diarrheal illnesses on the Ecuadorian coast: socio-environmental changes and concepts of health. <i>Cad Saúde Pública Rep Public Health</i> 2010; 26 :1334–44.
39 40 41 42	45	Blum D, Feachem RG. Measuring the impact of water supply and sanitation investments on diarrhoeal diseases: problems of methodology. <i>Int J Epidemiol</i> 1983; 12 :357–65.
43 44 45 46	46	Lee G, Cama V, Gilman RH, <i>et al.</i> Comparison of two types of epidemiological surveys aimed at collecting daily clinical symptoms in community-based longitudinal studies. <i>Ann Epidemiol</i> 2010; 20 :151–8. doi:10.1016/j.annepidem.2009.10.004
47 48 49	47	Schmidt WP, Arnold BF, Boisson S, <i>et al.</i> Epidemiological methods in diarrhoea studies an update. <i>Int J Epidemiol</i> 2011; 40 :1678–92.
50 51 52 53 54 55 56 57 58	48	Thapar N, Sanderson IR. Diarrhoea in children: an interface between developing and developed countries. <i>Lancet</i> 2004; 363 :641–53. doi:10.1016/S0140-6736(04)15599-2

- 49 Holcomb DA, Knee J, Sumner T, *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. doi:10.1016/j.ijheh.2020.113496
- 50 Chard AN, Baker KK, Tsai K, *et al.* Associations between soil-transmitted helminthiasis and viral, bacterial, and protozoal enteroinfections: a cross-sectional study in rural Laos. *Parasit Vectors* 2019;**12**:216. doi:10.1186/s13071-019-3471-2
- 51 Brown J, Cumming O, Bartram J, *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. doi:10.1136/bmjopen-2015-008215
- 52 Liu J, Kabir F, Manneh J, *et al.* Development and assessment of molecular diagnostic tests for 15 enteropathogens causing childhood diarrhoea: a multicentre study. *Lancet Infect Dis* 2014;**14**:716–24. doi:10.1016/S1473-3099(14)70808-4
- 53 Lunn PG, Northrop-Clewes CA, Downes RM. Intestinal permeability, mucosal injury, and growth faltering in Gambian infants. *Lancet* 1991;**338**:907–10.
- 54 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8
- 55 Guerrant RL, Oria RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with long-term effects on child development. *Nutr Rev* 2008;**66**:487–505. doi:10.1111/j.1753-4887.2008.00082.x
- 56 Campbell DI, Elia M, Lunn PG. Growth faltering in rural Gambian infants is associated with impaired small intestinal barrier function, leading to endotoxemia and systemic inflammation. *J Nutr* 2003;**133**:1332–8.
- 57 Goto R, Mascie-Taylor CG, Lunn PG. Impact of intestinal permeability, inflammation status and parasitic infections on infant growth faltering in rural Bangladesh. *Br J Nutr* 2009;**101**:1509–16. doi:10.1017/S0007114508083554
- 58 Lunn PG. Growth retardation and stunting of children in developing countries. *Br J Nutr* 2002;**88**:109–10. doi:10.1079/BJNBJN2002652
- 59 Guerrant RL, DeBoer MD, Moore SR, *et al.* The impoverished gut--a triple burden of diarrhoea, stunting and chronic disease. *Nat Rev Gastroenterol Hepatol* 2013;**10**:220–9. doi:10.1038/nrgastro.2012.239
- 60 Eppig C, Fincher CL, Thornhill R. Parasite prevalence and the worldwide distribution of cognitive ability. *Proc Biol Sci* 2010;**277**:3801–8. doi:10.1098/rspb.2010.0973
- 61 Wierzba TF, Muhib F. Exploring the broader consequences of diarrhoeal diseases on child health. *Lancet Glob Health* 2018;6:e230–1. doi:10.1016/S2214-109X(18)30047-0

BMJ Open

(\mathbf{a})	
62	Yatsunenko T, Rey FE, Manary MJ, <i>et al.</i> Human gut microbiome viewed across age and geography. <i>Nature</i> 2012; 486 :222–7. doi:10.1038/nature11053
63	Schnorr SL. The diverse microbiome of the hunter-gatherer. <i>Nature</i> 2015; 518 :S14-5. doi:10.1038/518S14a
64	Schnorr SL, Candela M, Rampelli S, <i>et al</i> . Gut microbiome of the Hadza hunter-gatherers. <i>Nat Commun</i> 2014; 5 :3654. doi:10.1038/ncomms4654
65	Clemente JC, Pehrsson EC, Blaser MJ, <i>et al.</i> The microbiome of uncontacted Amerindians. <i>Sci Adv</i> 2015;1:e1500183. doi:10.1126/sciadv.1500183
66	Martinez I, Stegen JC, Maldonado-Gomez MX, <i>et al.</i> The gut microbiota of rural papua new guineans: composition, diversity patterns, and ecological processes. <i>Cell Rep</i> 2015; 11 :527–38. doi:10.1016/j.celrep.2015.03.049
67	Obregon-Tito AJ, Tito RY, Metcalf J, <i>et al.</i> Subsistence strategies in traditional societies distinguish gut microbiomes. <i>Nat Commun</i> 2015; 6 :6505. doi:10.1038/ncomms7505
68	Gacesa R, Kurilshikov A, Vich Vila A, <i>et al.</i> Environmental factors shaping the gut microbiome in a Dutch population. <i>Nature</i> 2022; 604 :732–9. doi:10.1038/s41586-022-04567-7
69	Clasen T, Boisson S. Assessing the Health Impact of Water Quality Interventions in Low- Income Settings: Concerns Associated with Blinded Trials and the Need for Objective Outcomes. <i>Env Health Perspect</i> 2016; 124 :886–9. doi:10.1289/ehp.1510532
70	Goddard FGB, Ban R, Barr DB, <i>et al.</i> Measuring Environmental Exposure to Enteric Pathogens in Low-Income Settings: Review and Recommendations of an Interdisciplinary Working Group. <i>Environ Sci Technol</i> 2020; 54 :11673–91. doi:10.1021/acs.est.0c02421
71	Prendergast A, Kelly P. Enteropathies in the developing world: neglected effects on global health. <i>Am J Trop Med Hyg</i> 2012; 86 :756–63. doi:10.4269/ajtmh.2012.11-0743
72	Brown J, Cumming O. Stool-Based Pathogen Detection Offers Advantages as an Outcome Measure for Water, Sanitation, and Hygiene Trials. <i>Am J Trop Med Hyg</i> 2020; 102 . doi:10.4269/ajtmh.19-0639
73	INE. Resultados definitivos Censo 2017. IV Recenseamento Geral da Populacao e Habitacao de Mocambique. Maputo: : INE 2017. Www.Ine.Gov.Mzww.Ine.Gov.Mz
74	Water & Sanitation for the Urban Poor. Supporting the provision of universal water services in Beira and Dondo. London: : Water & Sanitation for the Urban Poor 2022.
75	Development Projects : Water Services & Institutional Support II. World Bank. https://projects.worldbank.org/en/projects-operations/project-detail/P149377 (accessed 1 Oct 2021).

- 76 Netherlands M of FA of the K of the. Dutch Development results Mozambique. https://www.dutchdevelopmentresults.nl/2019/countries/mozambique#water (accessed 20 Apr 2022).
- 77 Schreiner M. Simple Poverty Scorecard® Poverty-Assessment Tool Mozambique. ;:127.
- 78 Blackwell M, Iacus S, King G, *et al.* Cem: Coarsened Exact Matching in Stata. *Stata J* 2009;**9**:524–46. doi:10.1177/1536867X0900900402
- 79 Porro G, King G, Iacus S. CEM: Software for Coarsened Exact Matching. *J Stat Softw* 2009;**30**. doi:10.18637/jss.v030.i09
- 80 Mbuya MN, Humphrey JH. Preventing environmental enteric dysfunction through improved water, sanitation and hygiene: an opportunity for stunting reduction in developing countries. *Matern Child Nutr* 2016;**12 Suppl 1**:106–20. doi:10.1111/mcn.12220
- 81 Cumming O, Cairncross S. Can water, sanitation and hygiene help eliminate stunting? Current evidence and policy implications. *Matern Child Nutr* 2016;**12 Suppl 1**:91–105. doi:10.1111/mcn.12258
- 82 Oria RB, Murray-Kolb LE, Scharf RJ, *et al.* Early-life enteric infections: relation between chronic systemic inflammation and poor cognition in children. *Nutr Rev* 2016;**74**:374–86. doi:10.1093/nutrit/nuw008
- 83 Platts-Mills JA, Babji S, Bodhidatta L, *et al.* Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Health* 2015;**3**:e564–75. doi:10.1016/S2214-109X(15)00151-5
- 84 Direcção Provincial de Saúde. Balanço do Plano Economico e Social. Beira: : Governo de Sofala, Direcção Provincial de Saúde, República de Moçambique 2021.
- 85 Arnold BF, Ercumen A. Negative Control Outcomes. *JAMA* 2016;**316**:2597–8. doi:10.1001/jama.2016.17700
- 86 UNICEF. Nutrition: definitions of the indicators. http://www.unicef.org/infobycountry/stats_popup2.html (accessed 5 Feb 2015).
- Hartung C, Lerer A, Anokwa Y, et al. Open data kit: tools to build information services for developing regions. In: Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development ICTD '10. London, United Kingdom: ACM Press 2010. 1–12. doi:10.1145/2369220.2369236
- 88 Bosch F, Palmeirim MS, Ali SM, *et al.* Diagnosis of soil-transmitted helminths using the Kato-Katz technique: What is the influence of stirring, storage time and storage temperature on stool sample egg counts? *PLoS Negl Trop Dis* 2021;15:e0009032. doi:10.1371/journal.pntd.0009032

BMJ Open

89	Liu J, Gratz J, Amour C, <i>et al.</i> A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. <i>J Clin Microbiol</i> 2013; 51 :472–80. doi:10.1128/JCM.02658-12
90	Bolyen E, Rideout JR, Dillon MR, <i>et al.</i> Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. <i>Nat Biotechnol</i> 2019; 37 :852–7. doi:10.1038/s41587-019-0209-9
91	Centers for Disease Control and Prevention Division of Parasitic Diseases and Malaria. Blood Collection and Dried Blood Spot Preparation. Atlanta, GA: : Centers for Disease Control and Prevention 2019.
02	Lammie PJ, Moss DM, Brook Goodhew E, <i>et al.</i> Development of a new platform for neglected tropical disease surveillance. <i>Int J Parasitol</i> 2012; 42 :797–800. doi:10.1016/j.ijpara.2012.07.002
93	Zinkernagel RM. Maternal antibodies, childhood infections, and autoimmune diseases. <i>N Engl J Med</i> 2001; 345 :1331–5. doi:10.1056/NEJMra012493
94	Smith CM, Hill VR. Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes from water. <i>Appl Environ Microbiol</i> 2009;75:5284–9. doi:10.1128/AEM.00456-09
95	Strina A, Rodrigues LC, Cairncross S, <i>et al.</i> Factors associated with rotavirus diarrhoea in children living in a socially diverse urban centre in Brazil. <i>Trans R Soc Trop Med Hyg</i> 2012; 106 :445–51. doi:10.1016/j.trstmh.2012.02.001
96	Chen S-Y, Chiu C-H. Worldwide molecular epidemiology of norovirus infection. <i>Paediatr Int Child Health</i> 2012; 32 :128–31. doi:10.1179/2046905512Y.0000000031
97	da Silva Poló T, Peiró JR, Mendes LCN, <i>et al.</i> Human norovirus infection in Latin America. <i>J Clin Virol</i> 2016; 78 :111–9. doi:10.1016/j.jcv.2016.03.016
98	Nasser AM. Prevalence and fate of hepatitis a virus in water. <i>Crit Rev Environ Sci Technol</i> 1994; 24 :281–323. doi:10.1080/10643389409388470
99	Derrien M, Alvarez A-S, de Vos WM. The Gut Microbiota in the First Decade of Life. <i>Trends Microbiol</i> 2019; 27 :997–1010. doi:10.1016/j.tim.2019.08.001
100	Cumming O, Arnold BF, Ban R, <i>et al.</i> The implications of three major new trials for the effect of water, sanitation and hygiene on childhood diarrhea and stunting: a consensus statement. <i>BMC Med</i> 2019; 17 :173. doi:10.1186/s12916-019-1410-x
101	Iacus SM, King G, Porro G. Matching for Causal Inference Without Balance Checking.
101	Rochester, NY: : Social Science Research Network 2008. doi:10.2139/ssrn.1152391

- **BMJ** Open
- 103 Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011;12:R60. doi:10.1186/gb-2011-12-6-r60
- 104 Oksanen J, Blanchet FG, Kindt R, et al. Package 'vegan': Community Ecology Package. 2016;:Ordination methods, diversity analysis and other functions for community and vegetation ecologists.https://cran.r-project.org, https://github.com/vegandevs/vegan (accessed 5 Feb 2017).

L. IIBL Adt R, *et al.* Pa , diversity analysis , ps://cran.r-project.org, /
AUTHORS' CONTRIBUTIONS

MM and MCF conceived of the overarching idea of evaluating the intervention; KL and MCF conceived of the specific study and secured funding; JVG and MCF designed the analysis plan, with input from LW; ZAC, SH, SM, JSS, and MKMP designed protocols for recruitment of participants and oversaw collection of field data, with guidance from RN and JLM; JSS and JVG oversaw data management, with help from SH, MKMP, SM, and CSFS; CV and CSFS designed specimen management and laboratory protocols; TC, JB, and LW advised on study design, epidemiological approaches and research methods; RN, MM, and JLM provided oversight on relevant scientific questions in Mozambique; RN, JSS, CV, KL, MCF, ZAC, SM, and MKMP managed human subjects protocol submissions; KK and KL oversaw microbiome analysis approach; RN oversaw parasitology analysis approach; KL and JB oversaw enteric pathogen analysis approach; KL and JVG wrote the manuscript, with input from all authors.

FUNDING STATEMENT

This work was supported by National Institute of Allergy and Infectious Diseases (NIAID) Grant #R01AI130163, National Institute of Environmental Health Sciences (NIEHS) Grant #5T32ES12870), and a contract from Autoridade Reguladora de Água, Instituto Público (AURA, IP), Mozambique. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare no conflicts of interests.

FIGURE CAPTIONS

Figure 1. PAASIM theory of change.

No caption needed

Figure 2. Map of PAASIM study site in Beira, Mozambique.

Map of Beira, Mozambique, with enlargement highlighting study site. Red lines indicate the new distribution system water network. Blue lines indicate other parts of the water network. Gray shaded areas indicate neighborhoods enrolled in the study.

Figure 3. PAASIM study summary diagram.

The diagram reflects the summarized a) study design and data collection, and b) data analysis approaches to isolate the effects of both overall water supply infrastructural improvements as well as the presence of a household water connection.

Figure 4. PAASIM study flow diagram.

No caption needed

Figure 5. Data and sample collection timeline for outcomes in the PAASIM study.

Data and sample collection of infant-mother dyads enrolled into the study will be used address a series of questions about the impact of community-level water system improvements on acute and chronic health outcomes in children in a low-income urban area of Mozambique.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

1	
2	
3	
4	
5	
6	
7	
8	Drinking water supply Drinking water exposure Gastrointestinal conditions Health impact Access to improved network and/or Drinking water that is accessible, on Improved acute and chronic child Improved child growth and
9	household water connection and free from contamination premises, available when needed, and free from contamination gastrointestinal conditions development
10	Τ
11	Figure 1. DAACIM the survey of charges. Many June 19
12	Figure 1. PAASIM theory of change. Mono Image.
13	307x40mm (300 x 300 DPI)
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
20	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
40	
48	
49	
50	•
51	
52	
53	
54	
55	
56	
57	
58	
59	For poor rovious only http://bmionon.hmi.com/site/ahout/swidelines.yhtml
60	For peer review only - http://binjopen.binj.com/site/about/guidelines.xntml



Figure 2. Map of PAASIM study site in Beira, Mozambique. Colour Image. Map of Beira, Mozambique, with enlargement highlighting study site. Red lines indicate the new distribution system water network. Blue lines indicate other parts of the water network. Gray shaded areas indicate neighborhoods enrolled in the study.

215x279mm (96 x 96 DPI)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



Figure 3. PAASIM study summary diagram. Colour Image. The diagram reflects the summarized a) study design and data collection, and b) data analysis approaches to isolate the effects of both overall water supply infrastructural improvements as well as the presence of a household water connection.

320x151mm (300 x 300 DPI)

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES)

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies



For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml





193x247mm (300 x 300 DPI)



Figure 5. Data and sample collection timeline for outcomes in the PAASIM study. Colour Image. Data and sample collection of infant-mother dyads enrolled into the study will be used address a series of questions about the impact of community-level water system improvements on acute and chronic health outcomes in children in a low-income urban area of Mozambique.

292x223mm (300 x 300 DPI)

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Details of data collection

A data collection firm (WeConsult) performs the in-country coordination of participant enrollment, data collection, and sample collection. Enumerators are recruited from Beira; all are fluent in Portuguese (spoken by >80% of the study population)¹ and more than half speak the local language. Enumerators conduct household visits for consent and eligibility pre-birth, and at months 3, 6, 9, and 12 post-birth; active surveillance calls take place monthly (**Figure 4**). All data are collected on electronic tablets using Open Data Kit (ODK) Collect, an open source program which allows offline data collection on a mobile device.²

2.4.1. Household visit

Survey instruments consist of several modules aimed at collecting data on key indicators through structured observations, reports from respondents, and objective measurements (Table S1). The survey includes household questions (e.g., SES, demographics, nutrition and food insecurity, flooding, animals, and sanitation and handwashing facilities), questions about the mothers (e.g., demographics, medical and physical health, pregnancy and birth care, and breastfeeding practices), questions about the index child (e.g., birth outcome, nutrition, health and recent illness, anthropometry, and vaccination history), and household questions about drinking water. We also collect GPS coordinates at each household visit.

2.4.2. Active surveillance calls

A brief monthly active surveillance call is conducted by phone to gather information on prenatal and perinatal environmental exposures and illnesses, on child illness symptoms, and intake of medicines, vitamins, breastfeeding, and introduction of complementary foods. The first active surveillance pre-birth call is targeted to occur one month after enrollment and calls continue on a monthly basis between visits. To facilitate communication with the study team participants receive a 150 MZN phone credit at each visit.

2.4.3. Drinking water

Drinking water characteristics, which are on the causal pathway between the intervention and our study outcomes (**Figure 1**), will be used to provide evidence of biological plausibility of the intervention's effects on these downstream outcomes. We will characterize water access using the Joint Monitoring Program definition of "safely managed," i.e., an improved water source that is accessible on premises (located within the dwelling, yard or plot), available when needed (sufficient water available for at least 12 hours per day), and free from contamination (no *E. coli* detected in a 100mL sample).^{3,4} We also assess a number of additional variables related to drinking water, including aspects of water quality, water access, water availability, water security, water consumption, and participant satisfaction with water. These variables are evaluated throughout the study at the household level and also at the community level through objective, observed, and participant-reported measures.

2.4.4. Diarrhea

At months 3, 6, 9, and 12, we ask the caregiver to report diarrhea and blood in the stool (dysentery) of the index child in the previous week. We use the case definition of diarrhea as 3 or more loose stools in a 24 hour period.⁵ Due to concerns about reporting biases, we also include a negative control outcome⁶; caregivers report on accidents that resulted in physical injury in the previous week. In addition, we note objective characteristics of the stool samples collected, including observed blood and mucus and an infant stool form scale describing consistency (4-point scale), amount (4-point scale), and color (6 categories).⁷

2.4.5. Anthropometry

At months 3, 6, 9, and 12, we measure child: 1) length, weight, and head circumference. From these measures we will calculate Z-scores (LAZ and WAZ), and prevalence of stunting and underweight, defined

Page 43 of 53

BMJ Open

as two standard deviations below median of the reference population.⁸ We measure weight using digital baby weighing scales with 5g gradations (ADE, Model# M112600, Hamburg, Germany), and we measure height using a baby length measuring board with 1mm gradations (ADE, Model# MZ10040).⁹ Measurements are repeated twice and recorded into the survey. The enumerator is prompted to conduct a third measurement if there are differences of >1 cm in length or head circumference or if weight is off by >0.5kg in the two repeated measurements. Each child's LAZ and WAZ will be calculated using the WHO Child Growth Standards for the reference population^{10,11} and WHO Anthropometric macros.¹²

Details of sample collection, processing, and analysis

2.5.1. Stool

Sample collection: Stool of the index child is collected at months 3, 6, 9, and 12 using a diaper, which is provided to the primary caregiver before a household visit. If a fecal sample is not provided during the initial home visit, we leave the primary caregiver with additional diapers and two more attempts are made to collect the sample within 4 hours of production. If needed, we provide a cooler and cold-pack, and collect the sample within 7 hours of production.

Sample processing: Lab personnel transfer 1g fecal material from the diaper into a DNA/RNA Shield Fecal Collection Tube (Zymo Research, Irvine, CA, USA); we aim to collect three aliquots in separate collection tubes. The DNA/RNA Shield stabilization buffer lyses the cells and renders the sample DNA and RNA stable for 30 days at ambient temperature.¹³ Samples are stored and shipped at -20 °C. Nucleic acids are extracted from fecal samples using the QIAamp 96 Virus QIAcube HT Kit on a QIAcube HT (Qiagen Sciences Inc., Germantown, MD, USA), and stored at -80°C until further processing.

Analysis - parasites: Immediately after the first three aliquots are placed in the collection tubes, two additional aliquots of fecal material are taken from the diaper to prepare a slide for Kato-Katz analysis of parasite ova¹⁴. If sufficient material is available for a second slide analysis is carried out in duplicate. Samples are analyzed for hookworm (*e.g., Necator americanus, Ancylostoma duodenale*) immediately following slide preparation, and for *Ascaris* spp., *Schistosoma mansoni, Trichuris trichiura, Taenia, Enterobius vermicularis*, and *Strongyloides stercoralis* after overnight incubation at room temperature. Eligible participants are referred for deworming medicine at the 12-month visit, after returning results of the parasitological exam to study subjects in collaboration with Instituto Nacional de Saude (INS) staff in Beira.

Analysis - enteric pathogens: Extracted nucleic acids are analyzed using the TaqMan Array card (TAC, ThermoFisher Scientific, Waltham, MA, USA) assay, which allows quantification by real-time PCR via a 384-well microfluidic card for simultaneous detection of multiple viral, bacterial, and parasitic enteric pathogen targets as well as antimicrobial resistance genes. Immediately prior to nucleic acid extraction, samples used in downstream TAC assays are seeded with the Inforce 3 Bovine Vaccine (Zoetis, Parsippany-Troy Hills, NJ, USA)¹⁵ containing Bovine Herpesvirus 1 (BHV) and Bovine Respiratory Syncytial Virus (BRSV) as extrinsic controls, to monitor extraction and amplification efficiency. Pathogens will be linked to the diarrheal disease episode based on relative cycle threshold values from the TaqMan results.¹⁶ The specific targets we will test for are shown in **Table S2**. The batch of TaqMan Array Cards will be QA/QC using positive control plasmids. Initially, a standard curve for quantification will be run in duplicate, and the limit of detection and limit of quantification will be determined by running cards at a low concentration (concentration determined based on the standard curve tests) until a 95% positivity rate out of at least ten assays is obtained. Additionally, when processing samples, standard curves will be run in singlicate once per month with the positive control plasmid. A negative control will be run per card.

Analysis - Gut microbiome composition: We will characterize the gut microbial community structure and composition by sequencing of the V4 region of the 16S ribosomal RNA (rRNA) gene amplicon. Immediately prior to nucleic acid extraction, samples used in downstream 16s assays will be seeded with the ZymoBIOMICS Spike-in Control I (High Microbial Load) (Zymo Research) containing *Imtechella*

and data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

halotolerans and *Allobacillus halotolerans*. Bioinformatic analyses will be completed using the QIIME2 software platform and other bioinformatics tools¹⁷.

2.5.2. Dried blood spots

 Sample collection: A trained nurse collects up to six dried spots of capillary blood of the index child at 6, 9, and 12 month visits on Tropbio Filter Paper Blood Collection Disks (Cellabs, Sydney, Australia), using a 2mm lancet.

Sample processing: Samples are allowed to dry overnight, then three aliquots of two spots each are placed in a Ziploc bag with silica desiccant. Dried blood spots can be stored at ambient temperatures for up to 100 days, even in tropical climates^{18,19} but samples are stored at -20 °C and shipped at ambient temperature.²⁰

Analysis – antibodies: We will use the Luminex platform to carry out high throughput, multiplex antibody assays that enable the simultaneous measurement of quantitative antibody responses to dozens of pathogens from a single blood spot.²¹ Bead coupling of antigens will occur at the U.S. Centers for Disease Control and Prevention (CDC), and CDC collaborators will also provide support in determining appropriate antigen cut points. Our panel includes a subset of enteropathogens that have targets on the TAC assay, including *Giardia, Cryptosporidium, Entamoeba histolytica*, norovirus, *Campylobacter*, enterotoxigenic *E. coli* and *V. cholerae*, following previous similar studies.²² Our first measure will occur at 6 months, to avoid detection of maternal antibodies that wane over the first 3-6 months of life.²³

2.5.3. Drinking water

Sample collection: We collect 100-mL household drinking water samples from source and stored water at all household visits. We select the drinking water source sample by asking the mother what the source of water is for the household that would be given to the study child to drink, or to mix formula. If the source water is not in the household, the enumerator walks with the mother to the source (e.g., a neighbor's house or public standpipe) to collect a sample and GPS location. We select the stored water sample by asking the mother for any water in the household that is used for drinking purposes; stored water is water stored in a jerry can, bucket, or other container in the household for later consumption.

To complement the household sampling, we also collect samples from a selection of 45 public standpipes located within the study area and 55 additional public standpipes located elsewhere in the city of Beira. At public standpipes we also measure water pressure by measuring time to fill a fixed volume (1L or 5L, depending on the pressure).

Sample processing and analysis: All samples are placed on cold packs after collection for transport to a lab in Beira. Samples are processed for fecal indicator bacteria within six hours of collection using Colilert-18 reagent and the Quanti-Tray/2000 MPN method (IDEXX Laboratories, Westbrook, ME, USA). Free and total chlorine levels are measured using a DR300 Pocket Colorimeter (Hach Company, Loveland, CO, USA) and DPD powder pillows (Hach Company)²⁴. Additional physiochemical parameters (pH, conductivity, and turbidity) are measured for public standpipe water samples using a Pocket Pro+ Multi 2 Tester and TL2300 turbidity meter (Hach Company).

Details of Blinding

At enrollment, households are assigned a unique identifier independent of intervention status. The core data management team conducts quality assurance using geocoded data to ensure group exposure status aligns with cluster-level designations of the exposure (i.e., intervention vs control areas). All geocoded data and group exposure indicators are removed prior to an external analyst performing the data cleaning and primary analysis. Data are cleaned, including decisions on missing data, outliers, and variable categorizations, before the analyst receives any group exposure. The analyst performs the primary analyses making comparisons between undefined group exposures. Once the analysis models are finalized, these rehearsal results are input into tables to create table shells for the final analyses. The primary analyst then receives a

Table S1. Household survey data collection in the *PAASIM* Study

able S1. Household	survey data collection in the PAASIM Study						pen-2022 :opyright	
Category	Indicator	Time poir	nt				Type of da	â
		Pre-birth	3mo	6mo	9mo	12mo	ludi	
Household-level		V	*	*	*	V		wawawtad
Jeneral	Socio-economic status	~				X	Obconvetic	oreported of households
characteristics	Domographics (household ownership	V	*	*	*	v		reported
	beussheld members, shildren under E voors	~				X		reported
	of ago, primary wago parnor)						023 elat	
Nutrition	Food incourity	v	v	v	v	v		roported
NUTITION	Types of food consumed by bousehold	^	^	^	^	^	Responding	reported
	members vesterday							reporteu
Sanitation	Access to improved sanitation	x	x	x	x	x	Ohsen and the	of household latrines
Janitation	Location of sanitation facility	×	X	x	x	×		of household compound
	Sharing of sanitation facility	×	X	x	x	×	Respondent	reported
	Sanitation facility characteristics	X	X	X	X	×	Ohsen	of household latrines
	(serviceable drop hole cover smooth and	A	~	Λ	Λ	Λ		
	cleanable floor)						, A J	
	Presence of human feces in the household	X	X	Х	Х	Х	Observertion	of household compound
	compound		/ĸ				nin b	·
	Trash disposal	Х	X	Х	Х	Х	Respondent	reported
Handwashing	Access to handwashing facility with soap	X	Х	X	X	X	Observation	of household compound
	and water						sin	·
Flooding	Flooding of household compound	Х	Х	Х	Х	X	Respondent	reported
Animals	Presence of animals in the household or	Х	Х	Х	Х	X	Respondent	reported
	compound (chickens, ducks/turkey, dogs,						, 11 chn	, , , , , , , , , , , , , , , , , , ,
	cats, pigs, sheep, goats, rabbits, donkeys)						, 20 10	
	Presence of animal feces in the household	Х	Х	Х	Х	Х	Observation	of household compoun
	compound						s. at /	
	Child(ren) contact with animals	Х	Х	Х	Х	Х	Responde	reported
Health	Deworming history (pre-school age children,		Х	Х	Х	Х	Responder	reported
	school age children, and mother)						5	
Moving	Moved to another household		Х	Х	Х	Х	Responder	reported
0							Ohservatio	of household compound

Page 47 of s	53		BMJ Oper	ı				136/bmjopen-
2 3 4								2022-061 right, ind
5	Category	Indicator	Time point	nt 3mo	600	9mo	12mo	Type of daga
6 7	Mother-level			51110		51110	121110	
8	General	Demographics (age, religion, language,	Х				х	Responders reported
9	characteristics	education, employment, marital status)						arc
10	Mental and	WHO 5 well-being index	Х	Х	Х	Х	Х	Respondent reported
11	physical health	Medical diagnosis history (diabetes, high	Х					Responded reported
12		blood pressure, heart disease, urinary tract						ed the D
13		infection, kidney disease, sexually-						o tt
15		transmitted disease, cancer, malaria)						
16		Vaginal health practices	Х					Responed reported
17		Pregnancy history	Х					Responented
18	Pregnancy health	Expected delivery date	Х					Responder reported
19		Pre-natal care	X					Respond Responded
20		Medical diagnosis during pregnancy	Х					Respondent reported
21		(gestational diabetes, high blood pressure,						
23		placenta previa, COVID-19, Dengue, Zika,						tra
24		Chikungunya, Malaria)						ini ni
25		Medications, vitamins, or supplements	Х	C				Responded reported
26		during pregnancy						anc c
27	Birth care	Delivery location		Х				Responged reported
28		Cesarean section		Х			5	Respondent reported
30		Post-natal care		Х				Respondent reported
31	Breastfeeding	Breastfeeding intentions	Х					Respondent reported
32	-	Breastfeeding practices		Х	Х	Х	Х 📂	Respondent reported
33	Sanitation	Primary place of defecation in last week	Х	Х	Х	Х	Х	Responded reported
34		Exclusive use of sanitation facility in last	Х	Х	Х	Х	Х	Respondent reported
35		week						Ag
30 - 37	Travel	Estimated time spent away from home in					Х	Responder reported
38		the last 12 months						E C C C C C C C C C C C C C C C C C C C
39		-						6
40								og
41								apt
42								niqu
43								Ie c
44 15		For peer review only - http://bu	miopen.bmi	com/si	te/aboi	ut/quide	elines.xh	tml —
46			y					
47								

						∍n-20
Category	Indicator	Time point				Type of data
		Pre-birth 3mo	6mo	9mo	12mo	067
Child-level	1					1ud
Birth	Birth outcome	Х				Respozdeng reported
	Birthdate	Х				Observation of child health card
	Birth weight	Х				Respon हिंतु से reported
	Sex	Х				Observ အမွဲထဲမှာ of child health card
Health	Mortality 🔨	Х	Х	Х	Х	Responed reported
	Vaccination history (DTP, Rotavirus, Polio,	Х	Х	Х	Х	Observစ်မြို့တို of child health card
	MMR)					to
	Medication history	Х	Х	Х	Х	Responହିୁସି ହାନ୍ତି reported
	Deworming history				Х	Respondent reported
	Illness (diarrhea, dysentery, nasal	Х	Х	Х	Х	Respond in the reported
	congestion, fever, vomiting, physical injury)					r (An ata
	in the last week					
	Medical diagnosis (colic, ear infection,	X	Х	Х	Х	Responed reported
	anemia, respiratory illness, malaria, asthma)					(g, /bm
	and treatment history					
Anthropometry	Head circumference	X	Х	Х	Х	Object ye reasurement
	Body length	X	X	Х	Х	Object ve reasurement
	Body weight	X	Х	X	Х	Object ye 📅 easurement
Nutrition	Types of liquid food consumed in the last	Х	X	X	Х	Respondent reported
	week					ini on
	Types of solid food consumed in the last		Х	x	X	Responderter reported
	week					ne
Sanitation	Diaper wearing	Х	Х	Х	x	Respondent reported
	Disposal of child stools	Х	Х	Х	х	Resported
Water consumption	Child consumption of drinking water source	Х	Х	Х	Х	Responder reported
	Drinking water treatment for child	Х	Х	Х	Х	Respondens reported
	consumption					gen
Travel	Estimated time spent away from home in				Х	Responder & reported
	the last 12 months					Bib

Page 49	of	53
---------	----	----

Category	Indicator	Time poi	nt				Type of days
		Pre-birth	3mo	6mo	9mo	12mo	ght.
Drinking water-level		•					ine 61
Water access	Municipality water connection and status	Х	Х	Х	Х	Х	Observ ation of water meter
	Type of drinking water source	Х	Х	Х	Х	Х	Responder reported
							Observation of water source
	Alternate source for non-drinking water	Х	Х	Х	Х	Х	Responder a reported
	Alternate source for drinking water outages	Х	Х	Х	Х	Х	Responded for the reported
	Distance to water main	Х	Х	Х	Х	Х	GPS lo မွှားမျှံမှုက်
	Location of drinking water source	Х	Х	Х	Х	Х	Respoin ସ୍ଥିୟ ମିଧି reported
	Distance to drinking water source	Х	Х	Х	Х	Х	GPS loge
	Time to collect water (minutes for round trip	Х	Х	Х	Х	х	Respo
	and trips per week)						t an
Water availability	Water availability (hours per day and days	Х	Х	Х	Х	Х	Responder reported
	per week)						lata
	Household water storage (large and small	Х	Х	Х	Х	Х	Responding reported
	containers)	6					nin S)
	Piped water flow rate (liters per minute)	X	Х	Х	Х	Х	Objective reasurement
	Water availability at handwashing facility	X	Х	Х	Х	Х	Observation of handwashing facilit
	Water availability at sanitation facility	X	X	Х	Х	Х	Observertion of sanitation facility
Drinking water	Treatment of stored water	Х	X	X	Х	Х	Resported
quality	Drinking water free/total chorine (source	Х	Х	Х	Х	Х	Objectige 🙀 easurement
	and stored water)						
Water security	Sharing of water connection	Х	Х	Х	Х	Х	Respondeng reported
	Household water insecurity experiences	Х	Х	Х	Х	X	Responderter reported
	Sufficient quantity of drinking water	Х	Х	Х	Х	X	Respondent reported
User satisfaction	Satisfied with water service	Х	Х	Х	Х	Х	Respondent reported
with water	Satisfied with water availability	Х	Х	Х	Х	Х	Resported
	Satisfied with water pressure	Х	Х	Х	Х	Х	Responder reported
	Satisfied with water color and appearance	Х	Х	Х	Х	Х	Respondent reported
	Satisfied with water taste and smell	Х	Х	Х	Х	Х	Responder reported
	Satisfied with water affordability	Х	Х	Х	Х	Х	Respondent reported
Water consumption	Monthly water expenses	Х	Х	Х	Х	Х	Responder F reported
	Water usage (liters per day)	Х	Х	Х	Х	Х	Responder reported
							Observation of water meter

Page 50 of 53

1	
2	
3	
4	
5	
6	
7	
, 0	
0	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
-1/	

Target	Assayed Dy Ine	E TaqMan Array Card (TAC) for 1	the PAASIM study		D D D Broho (E' to 2')*
laiget	Gene of interest	Bacterial F	Pathogens	ludi :	
Enteroaggregative F	aaiC	ATTGTCCTCAGGCATTTCAC	ΑΓΓΑΛΟΥΓΕΛΟΥ	ngi	
coli (EAEC)	aatA			ior i	
Diarrheagenic <i>E. coli</i> (DAEC)	afaB	GTCTCCCTGAATGTACAGCTTTCA	CMCTCTGCCACTCCACCTT	Ense uses r	TCAAGCTGTTTGTTCGTC
Shiga toxin-producing	stx1	ACTTCTCGACTGCAAAGACGTATG	ACAAATTATCCCCTGWGCCACTATC	elat	CTCTGCAATAGGTACTCCA
E. coli (STEC)	stx2	CCACATCGGTGTCTGTTATTAACC	GGTCAAAACGCGCCTGATAG	em(- TTGCTGTGGATATACGAGG
Enteropathogenic <i>E.</i>	eae	CATTGATCAGGATTTTTCTGGTGATA	CTCATGCGGAAATAGCCGTTA	tot	ATACTGGCGAGACTATTTCAA
coli (EPEC)	bfpA	TGGTGCTTGCGCTTGCT	CGTTGCGCTCATTACTTCTG	ext	
	LT	TTCCCACCGGATCACCAA	CAACCTTGTGGTGCATGATGA	and	CTTGGAGAGAAGAACCCT
Enterotoxigenic <i>E. coli</i>	STh	GCTAAACCAGCAGGGTCTTCAAAA	CCCGGTACAAGCAGGATTACAACA	d da	
	STp	TGAATCACTTGACTCTTCAAAA	GGCAGGATTACAACAAAGTT	ta A	TGAACAACACATTTTACTGCT
Shigella spp./ Enteroinvasive E. coli (FIEC)	ipaH	CCTTTTCCGCGTTCCTTGA	CGGAATCCGGAGGTATTGC	BES) . mining	CGCCTTTCCGATACCGTCTCTGCA
E. coli 0157	rbdE		CGATGAGTTTATCTGCAAGGTGAT	≥ .	стстстттсстстдсддтсст
Campylobacter jejuni/ C. coli	cadF	СТӨСТАААССАТАБАААТААААТТТСТСАС	CTTTGAAGGTAATTTAGATATGGATAATCO	trainir	CATTTTGACGATTTTTGGCTTGA
Salmonella spp.	ttr	CTCACCAGGAGATTACAACATGG	AGCTCAGACCAAAAGTGACCATC	, ĝ	CACCGACGGCGAGACCGACTTT
Vibrio cholerae	hylA	ATCGTCAGTTTGGAGCCAGT	TCGATGCGTTAAACACGAAG	and	ACCGATGCGATTGCCCAA
Clostridium difficile	tcdB	GGTATTACCTAATGCTCCAAATAG	TTTGTGCCATCATTTTCTAAGC	Isii	CCTGGTGTCCATCCTGTTTC
		Viral Pa	thogens	nila	<u> </u>
Adenovirus	fiber	AACTTTCTCTCTTAATAGACGCC	AGGGGGCTAGAAAACAAAA	r te	CTGACACGGGCACTCT
Astrovirus	capsid	CAGTTGCTTGCTGCGTTCA	CTTGCTAGCCATCACACTTCT	chi	
Hepatitis G	5' UTR	CGGCCAAAAGGTGGT GGA TG	CGACGAGCCTGACGTCGGG		
Norovirus GI	ORF1-ORF2	CGCTGGATGCGATTCCATGA	CTTAGACGCCATCATCATTTAC	ogie	TGGACAGGAGATCGC
Norovirus GII	ORF1-ORF2	CAAGAACCTATGTTTAGATGGATGAG	TCGACGCCATCTTCATTCACA	ŝ	¥ TGGGAGGGCGATCGCAATCT
Rotavirus	NSP3	ACCATCTWCACRTRACCCTCTATGAG	GGTCACATAACGCCCCTATAGC	ר שר	AGTTAAAAGCTAACACTGTCAAA
Sapovirus	RdRp	GAYCASGCTCTCGCYACCTAC	CCCTCCATYTCAAACACTA; TTGGCCCTCGCCACCTAC		CCRCCTATRAACCA
SARS-CoV-2	N1	GACCCCAAAATCAGCGAAAT	TCTGGTTACTGCCAGTTGAATCTG	2	ACCCCGCATTACGTTTGGTGGACC

f 53			ВМЈ Ор	en	cted by	36/bm
					сору	open-
	Targat	Cono of interact	Forward (E' to 2')		right,	Ň 00 20 20 20 20 20 20 20 20 20 20 20 20
	Target	Gene of interest	Forward (5 to 3)	Reverse (5 to 5)	inc	
	Cruntocnoridium con				lud	<u>3</u> 44
	Cryptospondium spp.	18S rRNA	GGGTIGTATITATTAGATAAAGAACCA	AGGCCAATACCCTACCGTCT	ing	g IGACATATCATICAAGTITCIGAC
	Giaraia auoaenalis	18S rRNA	GACGGCTCAGGACAACGGTT	TTGCCAGCGGTGTCCG	for	
	Entamoeba histolytica Cuclospora	18S rRNA	ATTGTCGTGGCATCCTAACTCA	GCGGACGGCTCATTATAACA	ы	
	cavetanensis	18S rRNA	AAAAGCTCGTAGTTGGATTTCTG	AACACCAACGCACGCAGC	nse es i	AAGGCCGGATGACCACGA
-			Helminth	ic pathogens	rela	202
-	Ascaris lumbricoides	ITS1	GCCACATAGTAAATTGCACACAAAT	GCCTTTCTAACAAGCCCAACAT	ted	⇔ ☐ TTGGCGGACAATTGCATGCGAT
	Trichuris trichiura	18S rRNA	TTGAAACGACTTGCTCATCAACTT	CTGATTCTCCGTTAACCGTTGTC	tot	
	Ancylostoma duodenale	ITS2	GAATGACAGCAAACTCGTTGTTG	ATACTAGCCACTGCCGAAACGT	fexi	
	Necator americanus	ITS2	CTGTTTGTCGAACGGTACTTGC	ATAACAGCGTGCACATGTTGC	per t an	
-			Antimicrobia	resistance genes	<u>ā</u> d	d f
-	intl1		GATCGGTCGAATGCGTGT	GCCTTGATGTTACCCGAGAG	ata (A	
	mcr-1		GATCGCTGTCGTGCTCTTTG		<u>n</u> B	
	SHV			TCCTGCTGGCGATAGTGGAT	ning	
	TEM		GCATCTTACGGATGGCATGA	GICCICCGATCGITGICAGAA	ġ.	
	CTX-M1		CCGTCACGCTGTTRTTAGGA		r tr	
	CTX-M8-M25		ATRACACSTTCCGGCTCGAT	GCTAAYGGCGTGGTGGTATC	aini	
	CTX-M2-M74			TGYGCSCGCTGBGTTTCC	ing	
	CTX-M9		GCTTTATGCGCAGACGARTG	ATCACCGCGATAAAGCACCT	an	
	КРС		GGUGGUGTGCAATAC	GCCGCCCAACTCCTTCA	d s	
	NDM		ATATCACCGTTGGGATCGAC	TAGTGCTCAGTGTCGGCATC	imi	
	VIM				lar t	
-			Co	Introls	- EC	
-	165		ΤΟΓΑΔΟΤΓΟΔΑΓΟΔΑΘΓΑΓΤΤΤΔ	GCAGGTTACCCACGCGTTAC	onr	
	human mtDNA				log	
	BHV		GAGCAAAGCCCCGCCGAAGGA		ies.	
	BRSV					

Table S3: Details of Power Calculations

The table below details our minimum detectible effect for individual and groups of pathogens given the control group prevalence (p1) ranging from 10% to 80%. It generates the minimum detectible effect as a risk ratio (delta) and resulting minimum detectible prevalence in the comparison group.

Our assumptions are the number of sub-neighborhoods in the intervention and comparison (M1/M2), average number of households enrolled per sub-neighborhood in intervention and control (K1/K2), standard alpha (0.05) and power (.80). We estimate the minimum detectible effect for three values of for the intra-class correlation coefficient, representing low, moderate and high clustering ($\rho = .01, 0.05, 0.1$). Our power calculations rely on the moderate estimates for ICC. We account for clustering at the sub-neighborhood level.

We apply the following code using STATA v16:

power twoproportions (.2(.05).8), cluster effect(ratio) m1(36) m2(26) k1(8) k2(10) power(.8) alpha(.05) direction(lower) rho(.01/.5/.1)

alpha	power	К1	к2	M1	м2	delta	p1	p2	rhc
.05	.8	8	10	36	26	.5120	.2	.1024	.01
.05	.8	8	10	36	26	.5676	.25	.1419	.01
.05	.8	8	10	36	26	.6114	.3	.1834	.01
.05	.8	8	10	36	26	.6476	.35	.2266	.01
.05	.8	8	10	36	26	.6785	.4	.2714	.01
.05	.8	8	10	36	26	.7058	.45	.3176	.01
.05	.8	8	10	36	26	.7302	.5	.3651	.01
.05	.8	8	10	36	26	.7527	.55	.414	.01
.05	.8	8	10	36	26	.7736	.6	.4642	.01
.05	.8	8	10	36	26	.7935	.65	.5158	.01
.05	.8	8	10	36	26	.8127	.7	.5689	.01
.05	.8	8	10	36	26	.8317	.75	.6238	.01
.05	.8	8	10	36	26	.8509	.8	.6807	.01
.05	.8	8	10	 36	26	.3585	.2	.07169	.0!
.05	. 8	8	10	36	26	.4262	.25	.1066	. 0.
.05	.8	8	10	36	26	.4804	.3	.1441	.0
.05	.8	8	10	36	26	.5257	.35	.184	.0
.05	.8	8	10	36	26	.5648	.4	.2259	.0
.05	.8	8	10	36	26	.5994	.45	.2698	.0.
.05	. 8	8	10	36	26	.6308	.5	.3154	. 0.
.05	. 8	8	10	36	26	.6596	.55	.3628	.0
.05	. 8	8	10	36	26	.6867	.6	.412	.0
.05	. 8	8	10	36	26	.7126	.65	.4632	.0
.05	. 8	8	10	36	2.6	.7376	. 7	.5163	. 0
.05	. 8	8	10	36	2.6	.7624	. 75	.5718	. 0
.05	.8	8	10	36	26	.7875	.8	.63	.0
.05	.8	8	10	36	26	.5398	.5	.2699	
.05	. 8	8	10	36	26	.5737	.55	.3155	
.05	. 8	8	10	36	26	.6056	.6	.3634	
.05	.8	8	10	36	26	.6362	.65	.4135	
.05	. 8	8	10	36	26	.6660	.7	.4662	_
.05	. 8	8	10	36	26	.6955	.75	.5216	-
.05	.8	- 8	10	36	26	.7254	.8	.5804	
.05	.8	- 8	10	36	26	.7567	.85	.6432	
05	8	8	10	36	26	7907	9	7117	•

Protected by copyright, including for uses related to text and data mining, Al training, and similar techn	Enseignement Superieur (ABES) .	MJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11,
ologies.		2025 at Agence Bibliographique de

1.	United Nations. The Mozambique third population and housing census [Internet]. 2007. Available from:
	http://unstats.un.org/unsd/demographic/sources/census/country_impl.htm
2.	Hartung C, Lerer A, Anokwa Y, Tseng C, Brunette W, Borriello G. Open data kit: tools to build information services for developing regions. In: Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development - ICTD '10 [Internet]. London, United Kingdom: ACM Press; 2010 [cited 2022 Apr 20]. p. 1–12. Available from: http://dl.acm.org/citation.cfm?doid=2369220.2369236
•	United Nations Children's Fund and World Health Organization. Integrating Water Quality Testing into Household Surveys: Thematic report on drinking water. New York: UNICEF and WHO; 2020.
•	WHO UNICEF Joint Monitoring Programme. JMP methodology: 2017 update & SDG baselines [Internet]. Geneva: WHO & UNICEF; 2018 [cited 2020 Mar 10]. Available from: https://washdata.org/sites/default/files/documents/reports/2018-04/JMP-2017-update-methodology.pdf
j.	Levine GA, Walson JL, Atlas HE, Lamberti LM, Pavlinac PB. Defining Pediatric Diarrhea in Low-Resource Settings. J Pediatr Infect Dis Soc. 2017 Sep 1;6(3):289–93.
5.	Arnold BF, Ercumen A. Negative Control Outcomes. JAMA. 2016 Dec 27;316(24):2597–8.
' .	Bekkali N, Hamers SL, Reitsma JB, Van Toledo L, Benninga MA. Infant Stool Form Scale: Development and Results. J Pediatr. 2009 Apr 1;154(4):521-526.e1.
;.	UNICEF. Nutrition: definitions of the indicators [Internet]. [cited 2015 Feb 5]. Available from: http://www.unicef.org/infobycountry/stats_popup2.html
€.	Kotloff KL, Blackwelder WC, Nasrin D, Nataro JP, Farag TH, van Eijk A, et al. The Global Enteric Multicenter Study (GEMS) of diarrheal disease in infants and young children in developing countries: epidemiologic and clinical methods of the case/control study. Clin Infect Dis. 2012 Dec;55 Suppl 4:S232-45.
10.	De Onis M. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight- for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva, Switzerland: World Health Organization; 2006.
11.	World Health Organization. WHO child growth standards, SAS macros. Geneva, Switzerland: World Health Organization; 2010.
12.	World Health Organization. Child growth standards: WHO Anthro (version 3.2.2, January 2011) and macros [Internet]. 2016 [cited 2017 Feb 5]. Available from: http://www.who.int/childgrowth/software/en/

- 13. Kazantseva J, Malv E, Kaleda A, Kallastu A, Meikas A. Optimisation of sample storage and DNA extraction for human gut microbiota studies. BMC Microbiol. 2021 May 29;21(1):158.
- Bosch F, Palmeirim MS, Ali SM, Ame SM, Hattendorf J, Keiser J. Diagnosis of soiltransmitted helminths using the Kato-Katz technique: What is the influence of stirring, storage time and storage temperature on stool sample egg counts? PLoS Negl Trop Dis. 2021 Jan 22;15(1):e0009032.
- 15. Bivins A, Bibby K. Wastewater Surveillance during Mass COVID-19 Vaccination on a College Campus. Environ Sci Technol Lett. 2021 Sep 14;8(9):792–8.
- 16. Liu J, Gratz J, Amour C, Kibiki G, Becker S, Janaki L, et al. A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. J Clin Microbiol. 2013 Feb;51(2):472–80.
- 17. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019 Aug;37(8):852–7.
- 18. Marques BL, Brandao CU, Silva EF, Marques VA, Villela-Nogueira CA, Do OK, et al. Dried blood spot samples: optimization of commercial EIAs for hepatitis C antibody detection and stability under different storage conditions. J Med Virol. 2012 Oct;84(10):1600–7.
- 19. Parker SP, Cubitt WD. The use of the dried blood spot sample in epidemiological studies. J Clin Pathol. 1999 Sep;52(9):633–9.
- 20. Centers for Disease Control and Prevention Division of Parasitic Diseases and Malaria. Blood Collection and Dried Blood Spot Preparation. Atlanta, GA: Centers for Disease Control and Prevention; 2019. Report No.: ECL.DR.C.032.
- Lammie PJ, Moss DM, Brook Goodhew E, Hamlin K, Krolewiecki A, West SK, et al. Development of a new platform for neglected tropical disease surveillance. Int J Parasitol. 2012 Aug;42(9):797–800.
- Zambrano LD, Priest JW, Ivan E, Rusine J, Nagel C, Kirby M, et al. Use of Serologic Responses against Enteropathogens to Assess the Impact of a Point-of-Use Water Filter: A Randomized Controlled Trial in Western Province, Rwanda. Am J Trop Med Hyg. 2017 Sep;97(3):876–87.
- 23. Zinkernagel RM. Maternal antibodies, childhood infections, and autoimmune diseases. N Engl J Med. 2001 Nov;345(18):1331–5.
- 4500-CI CHLORINE (RESIDUAL). In: Standard Methods For the Examination of Water and Wastewater [Internet]. American Public Health Association; 2018 [cited 2021 Oct 11]. (Standard Methods for the Examination of Water and Wastewater). Available from: https://www.standardmethods.org/doi/abs/10.2105/SMWW.2882.078

Study Design and Rationale for the PAASIM Project, a Matched Cohort Study on Urban Water Supply Improvements and Infant Enteric Pathogen Infection, Gut Microbiome Development, and Health in Mozambique

Journal:	BMJ Open
Manuscript ID	bmjopen-2022-067341.R1
Article Type:	Protocol
Date Submitted by the Author:	30-Nov-2022
Complete List of Authors:	Levy, Karen; University of Washington School of Public Health, Department of Environmental and Occupational Health Sciences; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Garn, Joshua; University of Nevada Reno, Division of Biostatistics, Epidemiology and Environmental Health, School of Public Health, Adriano, Zaida; WE Consult Muneme, Bacelar; WE Consult Fagnant-Sperati, Christine; University of Washington School of Public Health, Department of Environmental & Occupational Health Sciences Hubbard, Sydney; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Júnior, Antonio ; WE Consult Manuel, João ; Ministry of Health of Mozambique, Beira Operations Research Center, National Health Institute (INS) Mangamela, Magalhães ; Autoridade Reguladora de Água, Instituto Público McGunegill, Sandy; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Miller-Petrie, Molly; University of Washington School of Public Health, Department of Environmental & Occupational Health Sciences Snyder, Jedidiah ; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Willer , Lance A. ; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health Waller , Lance A. ; Emory University, Department of Biostatistics and Bioinformatics, Rollins School of Public Health Konstantinidis, Konstantinos; Georgia Institute of Technology, School of Civil & Environmental Engineering Clasen, Thomas; Emory University, Department of Environmental Health Sciences, Rollins School of Public Health Brown, Joe; University of North Carolina at Chapel Hill Gillings School of Global Public Health, Environmental Science and Engineering Nala, Rassul; Instituto Nacional de Saúde, Ministry of Health Freeman, Matthew; Emory University, Gangarosa Department of Environmental Health, Rollins School of Public Health
Primary Subject	Public health

1 2	
3	Heading:
5	Secondary Subject Heading:
6 7 8	Keywords:
9 10	
11	
13	
14 15	
16 17	
18 19	
20 21	
22	
24	
25 26	
27 28	
29 30	
31 32	
33	
35	
36 37	
38 39	
40 41	
42	
44	
45 46	
47 48	
49 50	
51	
53	
54 55	
56 57	
58 59	
60	For peer reviev



2		
3	1	Study Design and Rationale for the PAASIM Project a Matched Cohort Study on Urban
4	2	Water Sunnly Improvements and Infant Enterin Dathagen Infaction. Cut Microbiome
5	2	Water Supply Improvements and Imant Enteric Fathogen Infection, Gut Microbiome
6	3	Development, and Health in Wozambique
7	4	
8	5	Short Title: Study Design and Rationale for the PAASIM Matched Cohort Study
9	6	
10	7	Authors
11	8	Karen Levy, ^{1,2*} Joshua V. Garn, ³ Zaida Adriano Cumbe, ⁴ Bacelar Muneme, ⁴ Christine S.
12	9	Fagnant-Sperati, ¹ Sydney Hubbard, ² Antonio Júnior, ⁴ João Luís Manuel, ⁵ Magalhães
1 <i>3</i>	10	Mangamela, ⁶ Sandy McGunegill, ² Molly K. Miller-Petrie, ¹ Jedidiah S. Snyder, ² Courtney
15	11	Victor, ² Lance Waller, ⁷ Konstantinos T. Konstantinidis, ⁸ Thomas Clasen, ² Joe Brown, ⁹ Rassul
16	12	Nalá. ¹⁰ Matthew C. Freeman ^{2*}
17	13	
18	14	1 Department of Environmental & Occupational Health Sciences, School of Public Health
19	15	University of Washington Seattle WA USA
20	15	2 Congarage Department of Environmental Health Polling School of Dublic Health Emery
21	10	2. Galigatosa Department of Environmental Health, Konnis School of Fublic Health, Enlory
22	1/ 10	University, Atlanta, GA, USA.
23	18	3. Division of Biostatistics, Epidemiology and Environmental Health, School of Public Health,
24	19	University of Nevada—Reno, Reno, NV, USA.
25 26	20	4. WE Consult, Maputo, Mozambique.
20 27	21	5. Beira Operations Research Center, National Health Institute (INS), Ministry of Health, Beira,
27	22	Mozambique.
20	23	6. Autoridade Reguladora de Água, Instituto Público (AURA, IP), Maputo, Mozambique
30	24	7. Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory
31	25	University, Atlanta, GA, USA.
32	26	8. School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta,
33	27	GA USA
34	28	9 Department of Environmental Sciences and Engineering Gillings School of Public Health
35	29	University of North Carolina at Chanel Hill Chanel Hill North Carolina USA
36	30	10 Instituto Nacional de Saúde (INS) Ministry of Health Marracuene Mozambique
37	31	10. Instituto Macional de Sadde (1145), Winnstry of Meanin, Martaedene, Mozamorque.
38	22	
39 40	32 22	*Corresponding outhors, Authors contributed equally
40	22	Vortesponding authors. Authors contributed equally.
42	34 25	Karen Levy <u>klevyx(a)uw.edu</u> , 3980 ISin Avenue NE, Seattle, WA 98195
43	35	Mattnew C. Freeman <u>mattnew.freeman@emory.edu</u> ; 1518 Clifton Road NE, Atlanta, GA
44	36	30322
45	37	
46	38	Word count: 5,680
47	39	
48	40	
49 50	41	
5U 51		
51 52		
52 53		
54		
55		
56		
57		
58		

ABSTRACT

Introduction: Despite clear linkages between provision of clean water and improvements in child health, limited information exists about the health impacts of large water infrastructure improvements in low-income settings. Billions of dollars are spent annually to improve urban water supply, and rigorous evaluation of these improvements, especially targeting informal settlements, is critical to guide policy and investment strategies. Objective measures of infection and exposure to pathogens, and measures of gut function, are needed to understand the effectiveness and impact of water supply improvements.

Methods and analysis: In the PAASIM study, we examine the impact of water system improvements on acute and chronic health outcomes in children in a low-income urban area of Beira, Mozambique, comprising 62 sub-neighborhoods and ~26,300 households. This prospective matched cohort study follows 548 mother-child dyads from late pregnancy through 12 months of age. Primary outcomes include measures of enteric pathogen infections, gut microbiome composition, and source drinking water microbiological quality, measured at the child's 12 month visit. Additional outcomes include diarrhea prevalence, child growth, previous enteric pathogen exposure, child mortality, and various measures of water access and quality. Our analyses will compare a) subjects living in sub-neighborhoods with the improved water to those living in sub-neighborhoods without these improvements; and b) subjects with household water connections on their premises to those without such a connection. This study will provide critical information to understand how to optimize investments for improving child health, filling the information gap about the impact of piped water provision to low-income urban households,

using novel gastrointestinal disease outcomes.

Ethics and dissemination: The study was approved by the Emory University Institutional Review Board and the National Bio-Ethics Committee for Health in Mozambigue. The pre-analysis plan is published on the Open Science Framework platform (https://osf.io/4rkn6/). Results will be shared with relevant stakeholders locally, and through publications.

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

 STRENGTHS AND LIMITATIONS OF THE STUDY This matched cohort study of an urban water supply improvement project will p critical information about the health impacts of providing piped water and hous connections to low-income households. We employ rigorous measures of exposure and novel and objective outcome ma including gut microbiome composition and molecular detection of enteric pathon The study design allows for examination of both neighborhood and household-b effects of water supply improvements. 	
 STRENGTHS AND LIMITATIONS OF THE STUDY This matched cohort study of an urban water supply improvement project will p critical information about the health impacts of providing piped water and hous connections to low-income households. We employ rigorous measures of exposure and novel and objective outcome ma including gut microbiome composition and molecular detection of enteric pathon The study design allows for examination of both neighborhood and household-b effects of water supply improvements. 	
 This matched cohort study of an urban water supply improvement project will p critical information about the health impacts of providing piped water and hous connections to low-income households. We employ rigorous measures of exposure and novel and objective outcome main including gut microbiome composition and molecular detection of enteric pathon The study design allows for examination of both neighborhood and household-leffects of water supply improvements. 	
 critical information about the health impacts of providing piped water and hous connections to low-income households. We employ rigorous measures of exposure and novel and objective outcome measures of exposure and molecular detection of enteric pathon including gut microbiome composition and molecular detection of enteric pathon The study design allows for examination of both neighborhood and household-leffects of water supply improvements. 	provide
 We employ rigorous measures of exposure and novel and objective outcome measures of exposure and molecular detection of enteric pathology of the study design allows for examination of both neighborhood and household-leffects of water supply improvements. 	ehold
 6 including gut microbiome composition and molecular detection of enteric patho 7 The study design allows for examination of both neighborhood and household-l 8 effects of water supply improvements. 	easures
 The study design allows for examination of both neighborhood and household-l effects of water supply improvements.)gens
8 effects of water supply improvements.	level
9 • As a natural experiment, we are unable to randomize the intervention, leading to	o potential
10 residual confounding.	1
• We are unable to examine the impacts of all aspects of the city-wide water impl	rovement
12 project, due to lack of comparable populations, and instead focus only on the lo	w income
13 neighborhoods.	
14	

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

<u>1. INTRODUCTION</u>

Large-scale provision of disinfected, treated drinking water is considered one of the greatest public health achievements of the 20th century[1] and played an important role in improving child health in high-income countries.[2] In low- and middle-income countries (LMICs) with high burdens of infectious diseases, inadequate water, sanitation, and hygiene (WASH) conditions are strongly associated with poor child health outcomes, including diarrheal diseases, which are responsible for >500,000 deaths of children <5 annually.[3,4] Repeated enteropathogen infections, regardless of symptoms, can lead to stunting and linear growth faltering, a comorbidity that impacts 4.7% of children in LMICs and is responsible for a 4.8-times increase in mortality.[5] In Mozambique, 27% of stunting is attributed to unimproved water and sanitation.[6]

1.1. Robust studies of the health impacts of community water supply are needed

Rapid urbanization is occurring globally, with urban areas expected to account for 96% of the additional 1.4 billion human population by 2030[7] and 68% of the global population is expected to live in urban areas by 2050.[8] While sub-Saharan Africa is still predominantly rural, by 2050 the continent is projected to be 56% urban.[9] To cope with urban growth, expanded infrastructure and services in cities and peri-urban areas will be essential.[7] In 2017 alone, an estimated US\$3.95 billion was distributed in official development assistance to WASH in Africa (47% to the Eastern and Southern African region, specifically), with the majority going to improve water supply and sanitation systems.[10] Implementation challenges in lower-income settings—such as intermittent service and pathogen intrusion in the distribution system due to pipe breaks, pressure drops, or illegal connections—limit the potential for engineered systems to provide a continuous supply of treated drinking water directly to homes, in adequate quantities to improve hygiene.[11,12] Given

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

the considerable investment in providing piped services to low-income communities, rigorous
evaluation of community-scale water provision is critical, to understand the real-world
effectiveness and health impact of such systems in low-income contexts.[13–15]

Despite the clear biological link between safe water and child health and development, limited information exists about the health impacts of large water infrastructure improvements in low-income settings. A small number of studies have evaluated upgrades from intermittent to continuous water delivery in urban areas, [16] or localized improvements to water quality at shared water points.[17] Other studies have evaluated sanitation interventions, without examining drinking water or combined water and sanitation interventions.[18] A recent review of interventions to improve water quality globally found no studies evaluating reliable piped-in water supplies delivered to households and specifically called for rigorous research to assess the health impact of reticulated water supply systems.[19] The review concluded that "there is currently insufficient evidence to know if source-based improvements such as protected wells, communal tap stands, or chlorination/filtration of community sources consistently reduce diarrhoea."([19] p2) A World Health Organization (WHO) review of drinking water on diarrhoeal disease in low- and middle-income settings concurs, stating that "Although evidence on the effects of providing drinking water of higher quality supplied on premises...on health is scarce, data on the effects of supplying safely managed WASH services in LMICs are completely missing."([20] p56) One reason for this limited evidence is that community-scale interventions are difficult to study using randomized control trial (RCT) methodology – the gold standard for causal inference. It is often infeasible to randomize intervention groups due to policy, planning, and engineering considerations, and lack of adequate comparison groups. As such, alternative quasi-experimental designs must be applied.[21–23]

Page 7 of 54

BMJ Open

The predominance of studies in the WASH sector have focused on household- and compound-level interventions because they lend themselves more readily to RCT methodology. Outside of the few aforementioned studies of community-wide infrastructure improvements, evaluations of the health impact of water quality improvements in low-income settings often focus on household water treatment such as boiling, chlorinating or filtering water, with studies predominantly conducted in rural settings.[19] Results of these trials have been mixed, since household-based approaches have various limitations, including low uptake and inconsistent use. [24.25] post-treatment contamination, [26–29] and a poor record of sustained use. [30,31] Household water treatment interventions do not increase water quantity and availability and typical household WASH interventions are likely insufficient to prevent growth faltering in most cases.[32–35]

It is crucial to assess the impact of community-scale infrastructure improvements,[22] as this is an area that is particularly relevant to inform local and national policy-makers, aid agencies, and development banks.[36] The area of most rapid growth in water access is via piped water supply connections, not household water treatment, and larger infrastructure interventions are also critical to achieving the scale of water supply improvements necessary to make impactful changes.[37]

- - **1.2. Objective measures of gut health are needed**

A vast majority of WASH studies use as their primary health outcome caregiver-reported diarrhea, primarily because acute diarrheal illness is responsible for ~10-12% of all deaths in children <5.[38,39] However, self-reported diarrhea is an unreliable outcome due to courtesy, social desirability and recall bias,[40] local definitions of diarrhea,[41–45] other self-reporting issues,[40,46–48] and the multiple potential etiologies of diarrhea symptoms.[49] Such biases are especially problematic where interventions cannot be blinded as is mainly the case for water

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

interventions. Shedding of enteropathogens, organisms that cause acute gastrointestinal illness, provides an unambiguous indicator of current infection, and increasingly is being used in the WASH field.[32,50,51] Advances in diagnostic techniques make it feasible to test for a wide variety of enteric pathogens simultaneously.[52,53] It is useful to understand enteric pathogen infections because chronic and repeated enteric pathogen infections in the first two years of life— with or without symptomatic diarrhea—are associated with serious morbidities, including gut impairment, growth shortfalls, and cognitive deficits by ages 7-9 years.[54-60] Such outcomes can have profound impact on health, development and well-being of individuals, communities, and entire countries.[61,62] Host-level gastrointestinal conditions affected by environmental determinants, such as gut microbiome composition, may also help explain the long-term sequelae of enteric infections. While there is evidence of differences in gut microbiome composition across different cultures, regions, and populations, [63–68] and environmental conditions, [69] to date specific WASH determinants of these differences, such as access to piped water, have not been evaluated using explicit counterfactuals. Thus measures of gut microbial conditions may provide objective outcomes to more accurately measure the effect of WASH interventions[70,71] and capture long-term sequelae, [72] resulting in a more complete understanding of the health impacts.[73]

1.3. Overview of study

In the PAASIM study (*Pesquisa Sobre o Acesso à Água e a Saúde Infantil em Moçambique -Research on Access to Water and Children's Health in Mozambique*), we address a series of questions about the impact of community-level water system improvements on acute and chronic health outcomes in children in a low-income urban area of Mozambique. This matched-control

BMJ Open

cohort study follows mother-child dyads from late pregnancy through children 12 months of age, examining the impact of living in an area with an improved water network and/or having a household water connection on a variety of aspects of access to drinking water, microbes in the child gut (including both pathogens and other resident gut microbes), and ultimately downstream health outcomes (including diarrhea prevalence and growth) (Figure 1). This study will provide critical information for agencies who seek to understand how to optimize investments for improving child health, helping to fill the information gap about the impact of providing piped water to urban, low-income households by isolating the effects of major community-level water supply improvements on novel gastrointestinal disease outcomes.

11 2. METHODS & ANALYSIS

2.1. Study setting and description of the intervention

Our study site is the coastal city of Beira, the second largest city in Mozambique (population \sim 530,000),[74] which serves as a gateway for both the central interior portion of the country and a trade corridor to neighboring land-locked nations. The center of Beira is bordered by unplanned, informal settlements inhabited by over 300,000 low-income residents. [75] A 2018 survey in Beira by Water & Sanitation for the Urban Poor (WSUP) showed that of 5,643 respondents, only 28% had a household water connection, and among those households without a connection, 83% used their neighbor's tap as their main source of water (unpublished data, courtesy of WSUP). Therefore, improvement of water supply and delivery infrastructure is a priority.

The World Bank funded the Water Service & Institutional Support (WASIS-II) Project in 2016 to address the low access to improved water supply in Mozambique,[76] investing \$140 million with the Mozambican public institutions FIPAG (responsible for the public and private investment

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

program in urban water supply systems that serves as the water utility in Beira) and AURA, IP (the water regulatory authority responsible for the economic regulation and consumer protection of service provision). In addition, improvements in Beira are being augmented by investments from other groups, in particular the Dutch government, through infrastructure upgrades as well as emergency response funds following Cyclone Idai in 2019.[77] Improvements in the city of Beira include rehabilitation of water treatment facilities, replacing existing pipe mains that are failing, reticulation of water supply to new areas previously without water service, improving service in areas with poor coverage or low water pressure, and subsidizing water connection fees for the poor.

11 2.2 Study design

Several aspects of a city-wide water supply improvement project pose challenges to implementing a rigorous epidemiological study, particularly as people living in neighborhoods with piped water often differ in myriad ways from people living in neighborhoods without piped water. Water improvements to communities are often based on the needs or demographics of the community, government or donor priorities, and engineering considerations. Provisions of water supply to new areas previously without water service—or dramatic improvements in access and availability—represent a fundamental development that changes the livability and sometimes the makeup of the community.[2] These issues lead to difficulty in finding a comparable control group for epidemiological comparison. Furthermore, rollouts of water interventions often happen in continuous phases over time, and these changes might coincide with other events or community improvements that similarly impact health, making it difficult for that community to serve as its own control in a pre post design.

Page 11 of 54

BMJ Open

By using a prospective matched cohort in this unique context of an ongoing natural experiment, we are able to overcome many of these difficulties. The prospective nature of our study allows better control of confounding through matching, restriction, and rigorous and thoughtful collection of potential control variables. We specifically focus on one region of the city (Figure 2), where some neighborhoods received water system improvements focused on preventing water losses through replacement of the distribution system pipe system in dense, low-income settings; other neighboring areas with similar demographic characteristics did not receive these improvements. Neighborhood-level matching took place in the context of a natural experiment, where the delayed rollout across the city allows us to find and compare intervention and control neighborhoods that are similar in many ways, before the rollout eventually reached all potential control areas.

The Water Loss Reduction Project represents a subset of the improvements being carried out by FIPAG, with co-funding from the Dutch government and the World Bank WASIS-II project. The improvements are designed to reduce illegal connections, thereby increasing the water pressure and quality and increasing the system's capacity for household connections. These areas also received some benefits related to improvements to the water intake and distribution systems. FIPAG undertook a campaign to offer new connections to households. These improvements were completed in some informal settlements in these low income areas of the city in 2019, with other adjacent neighborhoods with similar density and socioeconomic profile slated for completion in future years but not within the timeframe of our study. These specific distribution system upgrades therefore represent a unique opportunity to examine the impacts of community-scale water improvements with neighboring communities who did not receive the intervention serving as control areas for comparison. The neighborhoods under study are also in the lowest income-and therefore highest need areas of the city.

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de I Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

1	We will perform analyses that take into account the four factorial possible household types,
2	based on mother-child dyad living in a sub-neighborhood with or without the improved water
3	network and with or without a household connection (Figure 3). Our primary analyses focus on
4	assessing A) the total network effect, by comparing subjects living in sub-neighborhoods with the
5	improved water network (Household Types 1 and 2) to those living in sub-neighborhoods without
6	these improvements (Household Types 3 and 4); and B) the direct household connection effect, by
7	comparing subjects with household water connections on their premises (Household Types 1 and
8	3) to those without a connection (Household Types 2 and 4). Depending on the results of these two
9	primary analyses, secondary analyses may evaluate the other comparisons depicted in Figure 3.
10	The reasons for some neighborhoods receiving the improvements and others not were a result
11	of resource constraints. According to FIPAG, decisions on the order of improvements in different
12	neighborhoods were guided by resource constraints as well as engineering logistics. We conducted
13	a population-based survey (described below) that allowed us to both restrict and match study sub-
14	neighborhoods, thereby creating a statistically appropriate counterfactual for strong internal
15	validity. The evaluation of a real-world intervention delivered in an informal urban setting
16	provides strong external validity for estimating the effects of similar interventions in other low-
17	and middle-income country urban sites. Our study design allows us to isolate the effects of both
18	overall water supply infrastructural improvements as well as the presence of a household water
19	connection. The presence of control areas not receiving upgrades adjacent to intervention areas
20	that are matched on socioeconomic and density variables is unique to this study location. We
21	collect data at multiple timepoints for each study household, allowing us to examine variability in
22	each of the measures taken from each household, rather than at a single point in time, and also
23	allowing for longitudinal analyses of the households and the individual enrolled subjects. We also

1	
2	
3	
4	
5	
5	
0	
/	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
10	
עו 20	
2U 21	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
20 21	
21	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
10	
44 15	
45	
40	
4/	
48	
49	
50	
51	
52	
53	
54	
55	
55	
50	
5/ 50	
58	
59	

employ rigorous measures of exposure and novel and objective outcome measures, including gut
 microbiome composition and molecular detection of enteric pathogens.

3

4

2.3. Patient & Public involvement

5 The executive secretary of AURA, IP was directly involved in the formation of the research 6 questions, and FIPAG personnel were also engaged from the initiation of the project in helping 7 develop the study design. Our team also received input from other public agency stakeholders 8 during workshops that were held prior to initiation of the study. Study subjects and members of 9 the general public were not involved in the study design. We provide regular updates with data 10 summaries to public agency stakeholders, and plan to disseminate the main results to all study 11 participants and also through public presentations for stakeholders in both Beira and Maputo.

13

12

2.4. Sub-neighborhood selection

14 Sub-neighborhood eligibility, selection, and matching of intervention and control sub-15 neighborhoods occurred through a two-step process:

16 1) Intervention designation: Our study is a natural experiment, where the investigators had no 17 control over the selection or timing of the intervention implementation. The study flow diagram is shown in Figure 4. We worked with FIPAG to determine which neighborhoods in Beira were to 18 19 receive water distribution system upgrades prior to initiation of enrollment (2020) and before the 20 end of the study (2023). FIPAG provided maps and timelines for construction works related to the 21 upgrades, and the specific areas participating in the water loss reduction project. We also worked 22 with FIPAG and through satellite imagery to identify similarly dense low-income areas in Beira 23 that were not slated to receive water network upgrades. A total of 17 potential neighborhoods were

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

considered for inclusion in the study, and neighborhoods were divided into 80 sub-neighborhoods, delineated along natural boundaries such as roads or waterways. "Intervention" sub-neighborhoods include areas with the upgraded water distribution system. "Control" sub-neighborhoods include areas not receiving these improvements during the time period of the study. Within both intervention and control sub-neighborhoods, some households have a connection to the water system and others do not. We excluded nine control sub-neighborhoods that were in close proximity to intervention sub-neighborhoods or that were scheduled to receive the interventions within the timetable of our project; some control sub-neighborhoods are slated to receive the intervention after completion of our study.

2) Matching & Restriction: We followed the suggestion of Arnold et al. (2010) to use baseline (preintervention) data at the community level to match intervention to control communities when randomization is not possible.[23] To characterize subneighborhoods for further matching and restriction, we performed a population-based community survey in November-December, 2020 of approximately 1,700 households; this provided approximately a 5% proportional sample of our potential study sub-neighborhoods. We used a random grid sampling approach to estimate household density, using Google Earth satellite imagery, where a grid was placed over an area, and a random selection of squares were selected and counted independently in duplicate, and the number of houses per unit was extrapolated across unsampled squares. The survey contained modules regarding household demographics, water access and practices, sanitation access and practices, household assets and wealth indicators, as well as questions related to COVID-19. A socioeconomic status (SES) score was constructed using the 'simple poverty scorecard'[78] developed specifically for Mozambique, and scores were aggregated at the sub-neighborhood level, and categorized into tertiles.

BMJ Open

We matched intervention sub-neighborhoods to control sub-neighborhoods, using coarsened exact matching [79.80] with intervention sub-neighborhoods being matched to control sub-neighborhoods within the same tertile of both SES and population density. Four neighborhoods (encompassing nine sub-neighborhoods) were found to be outliers in terms of their sub-neighborhood-level SES or sanitation, and were excluded from the study sampling frame. Ultimately we designated 36 intervention sub-neighborhoods, with an estimated 16,800 households, and 26 control sub-neighborhoods, with an estimated 9,500 households.

2.5. Participant recruitment, eligibility, and retention

We recruit pregnant women at the last trimester of pregnancy and follow the infant-mother dyads until the child is 12 months old (Figure 5). We selected the first 12 months of life because it is a critical development window, [81–83] it is a time when children are most at risk of acute and chronic effects of enteropathogen infection, [84] and it is a short enough period of time to avoid changes in water access that might occur. We recruit mothers at the end of their pregnancy so we can collect data on household risk factors (including drinking water quality) during the gestational period. Active recruitment occurs through identification of pregnant women in the 2020 population-based survey, lists of pregnant women visiting local health centers for pre-natal care, and study staff visiting under-enrolled sub-neighborhoods throughout the recruitment period. Based on Ministry of Health data for Sofala Province (where Beira is located), virtually all mothers attend pre-natal clinical visits.[85] Passive strategies include referrals of pregnant women by study participants and community leaders. We aim to have complete data on a total of 548 infant-mother dyads, approximately evenly divided between the intervention and control groups. We will continue to enroll dyads into both arms until we reach a minimum of 274 dyads with complete
BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

data in each arm, to ensure temporal balance throughout the duration of the study period.

During an initial pre-birth visit, pregnant women are assessed for study eligibility: 1) 18 years or older, 2) in third trimester of pregnancy, 3) resides in enrolled study cluster, 4) not planning to move within the next 12 months, 5) carrying a singleton birth, and 6) consents to take part in the study. We will re-assess study eligibility at each follow up visit and record if enrolled participants have been lost to follow up.

2.6. Data collection

A local data collection firm (WE Consult) performs the in-country coordination of participant enrollment, data collection, and sample collection. Enumerators conduct household visits before birth for consent, eligibility, and conditions. At months 3, 6, 9, and 12 we deploy survey instruments to collect data on key indicators through structured observations, reports from respondents, and objective measurements (Table S1). We assess a number of variables related to drinking water, including aspects of water quality, water access, water availability, water security, water consumption, and participant satisfaction with water. Enumerators also conduct brief active surveillance calls on a monthly-basis by phone with caregivers to gather supplemental information on prenatal and perinatal environmental exposures and illnesses, on child illness symptoms, and intake of medicines, vitamins, breastfeeding, and introduction of complementary foods (Figure 5). To facilitate communication with the study team, participants receive a 150 MZN phone credit at each visit. Aside from these phone credits, there is no financial incentive provided to participants to partake in the study, per Mozambican guidelines for human subjects research. We ask the caregiver to report diarrhea and blood in the stool (dysentery) of the index child in the previous week at the 3, 6, 9, and 12 month surveys and during active surveillance calls; due to concerns

BMJ Open

about reporting biases, we also include negative control outcomes.[86] At each post-birth visit we
measure child: 1) length, weight, and head circumference, and calculate length-for-age and weightfor-age Z-scores. Prevalence of stunting and underweight are defined as two standard deviations
below median of the reference population.[87] All data are collected on electronic tablets using
Open Data Kit (ODK) Collect, an open-source program which allows offline data collection on a
mobile device.[88] Additional details are provided in the Supplementary Material.

8 2.7. Sample collection, processing, and analysis

9 We briefly describe sample collection and downstream processing and analysis here, with10 additional details provided in the Supplementary Material.

12 2.7.1. Stool

13 Stool of the index child is collected at months 3, 6, 9, and 12. Three aliquots are placed in 14 temperature stable lysis buffer collection tubes, and two additional aliquots are used to prepare a 15 slide for Kato-Katz analysis of parasite ova[89]. Eligible participants are referred for deworming 16 medicine at the 12-month visit, after returning results of the parasitological exam to study subjects 17 in collaboration with Instituto Nacional de Saúde (INS) staff in Beira.

Extracted nucleic acids are analyzed: (1) using the TaqMan Array card (TAC, ThermoFisher Scientific, Waltham, MA, USA) assay, which allows quantification by real-time PCR via a 384well microfluidic card for simultaneous detection of multiple viral, bacterial, and parasitic enteric pathogen targets as well as antimicrobial resistance genes,[90] customized for our targets of interest (**Table S2**); and (2) by sequencing of the V4 region of the 16S ribosomal RNA (rRNA)

gene amplicon to characterize gut microbiome community structure and composition.

Bioinformatic analyses will be completed using the QIIME2 software platform.[91] 2.7.2. Dried blood spots Sample collection: A trained nurse from INS collects up to six dried spots of capillary blood of the index child at 6, 9, and 12 month visits on Tropbio Filter Paper Blood Collection Disks (Cellabs, Sydney, Australia), using a 2mm lancet. Samples are stored at -20 °C at INS facilities in Beira and shipped at ambient temperature [92] We will use the Luminex platform to carry out high throughput, multiplex antibody assays that enable the simultaneous measurement of quantitative antibody responses to dozens of pathogens from a single blood spot.[93] Our first measure will occur at 6 months, to avoid detection of maternal antibodies that wane over the first CLIC.

3-6 months of life.[94]

2.7.3. Drinking water

Sample collection: We collect 100-mL household drinking water samples from source and stored water at all household visits. To complement the household sampling, we collect samples from a selection of 45 public standpipes located within the study area and 55 additional public standpipes located elsewhere in the city of Beira. At public standpipes we also measure water pressure by measuring time to fill a fixed volume (1L or 5L, depending on the pressure). Samples are processed for fecal indicator bacteria within six hours of collection using Collert-18 reagent and the Quanti-Tray/2000 MPN method (IDEXX Laboratories, Westbrook, ME, USA), as well as for free and total chlorine levels and additional physiochemical parameters (pH, conductivity, and turbidity). Large volume samples will be collected from a subset of 50 households (1 L, processed

BMJ Open

by membrane filtration) and 25 public standpipes (50 L, processed by dead end ultrafiltration[95])
 in two different seasons, and tested for enteropathogens using the TAC assay.

2.8. Outcomes

Our primary outcomes include: any bacteria or protozoa infection at age 12 months after birth; individual pathogens or pathogen groups; child gut microbiome composition; and household source water quality. We include viral, protozoal, and bacterial pathogens responsible for the vast majority of enteric pathogen infections and global disease burden. [96,97] While we measure viral pathogens using the TAC assay, they will be excluded from the combined enteropathogen prevalence primary outcome measure, because waterborne transmission is unlikely to dominate for these viral pathogens.[98–101] In addition to the aforementioned reasons related to child development and infection risk, measuring pathogens at 12 months will give us the greatest power to detect a difference, given higher levels of infection at that age than in younger children. We will measure gut microbiome using 16S rRNA gene amplicon sequencing in the full sample at 12 months and in a random subset of 200 children with complete data at 3, 6, and 9 months, evenly distributed between intervention and control groups; dyads eligible for sub-set sampling will include those with complete stool sample collection and unchanged intervention exposure conditions. The 12-month samples will allow us to compare all study children at a common time, when all children are consuming drinking water and once the gut microbiome has become relatively established: [102] the longitudinal samples will allow for comparison of development of the microbiome over time between the two groups. Microbiome outcomes include alpha and beta diversity metrics, and identification of enriched taxonomic groups. We also include household source water quality as a primary exposure outcome, as understanding whether exposure to

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

microbial contaminants is altered is considered a critical aspect of evaluation of WASH projects.[71,103]

Additional non-primary outcomes include pathogen count, pathogen community similarity (measured using Jaccard similarity index), diarrhea, child growth, and prior enteropathogen infection (measured using serology on dried blood spot samples). We will measure additional water quality exposure measures, as well as measures of exposure to the improved water system, such as fidelity of the intervention (e.g., improvements to water quantity and coverage of household taps) and receipt of the intervention by community members (e.g., reductions in water insecurity, increased water use). These fidelity and uptake measures will be collected at all time points through direct observation and respondent report. Available minimal detectable effect sizes are summarized in **Table 1** and calculations are further detailed in the Supplementary Material.

2

3

4

60

BMJ Open

Table 1. Primary and non-primary health outcomes and exposure outcomes for the PAASIM Study. Patent enteric pathogen infection in stool is measured via TaqMan Array Card (TAC) assay; Stool microbiome composition is measured via 16s rRNA amplicon sequencing; prior enteric pathogen exposure is measured via serological assays of dried blood spots. Drinking water quality measured by IDEXX as *E. coli* most probable number/100mL. See text for further details. Calculations and additional values for the minimum detectable effect are described in the Supplementary Material.

Primary outcomes	Time point	Minimum detectable effect (Pisk Patio)	Anticipated con group prevalen
Patant antaria nathagan infaction		(KISK Katio)	
Prevalence of any basterial or protozoon nothogon	12 mah	0.74	700/
Prevalence of any bacterial nothercen	12 mob	0.74	/0%
Prevalence of any bacterial pathogen	12 mob	0.09	01%
Any co-infection (bacterial protozoan or viral pathogens)	12 mo ^b	0.50	<u> </u>
Any co-infection (bacterial, protozoan, or viral pathogens)	12 110	0.00	4370
Gut microbiome composition	12 mo ^c		
Alpha diversity			NA
Beta diversity			NA
Enriched taxa			NA
		-	
Household source drinking water quality	12 mo ^d		NA
Additional health outcomes			
Patent enteric pathogen infection			
Pathogen count	12 mo ^b		
Pathogen community similarity	12 mo ^b		
Individual pathogens	12 mo ^b	n/a-0.49	2-31%
Any virus	12 mo ^b	0.46	28%
Gut microbiome composition (Same variables as gut microbiome composition for 12 mo)	3, 6, 9 mo ^c		
Prior enteric pathogen exposure	6, 9, 12 mo		
(Same variables as patent enteric pathogen infection) Diarrhea 1-week period prevalence (caregiver report)	Weekly	0.26	14.4%
Anthronometric Measurements	3 6 0 12 mg		
Length-for-age Z-score	5, 0, 9, 12 110		
Weight-for-age Z-score			
Stunting prevalence ^e		0 49	31%
Underweight prevalence ^e		0.22	13%
All-cause mortality (while enrolled in study)	Continuous		
Additional Exposure outcomes		1	1
Primary drinking water source	3, 6, 9, 12 mo		
Drinking water quality (source)	3, 6, 9 mo ^t		
Drinking water quality (stored)	3, 6, 9, 12 mo		
Water access	3, 6, 9, 12 mo		
Water availability	3, 6, 9, 12 mo		
Water security	3, 6, 9, 12 mo		
Water consumption	3, 6, 9, 12 mo		
User satisfaction with water	3, 6, 9, 12 mo		

^aAnticipated control group prevalence based on control group prevalence for 10-14 month olds in the MapSan trial ([32] and J.
 Knee, pers. Comm.).

^b Samples also collected at 3, 6, and 9 months of age may also be analyzed, depending on results of primary analysis at 12 months.

^e A subset of n=200 samples will be analyzed for gut microbiome composition in children at 3, 6, and 9 months of age. All 12 month
 samples will be analyzed.

^d *This is a conservative estimate as it does not account for weekly active surveillance.*

^e *Defined as two standard deviations below median of the reference population.*

^fSamples will be analyzed at 3, 6, 9, and 12 mo, but 12 month samples are the primary outcome of interest.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

2.9. Analysis plan

The pre-analysis plan for this study is published on the Open Science Framework platform (https://osf.io/4rkn6/).

2.9.1. Total network effect

To assess the impact of the intervention on our primary enteric pathogen infection outcomes and water quality exposure outcome (Table 1), we will use an intention-to-treat (ITT) analysis approach to compare children living in intervention versus control sub-neighborhoods, without regard to uptake/use of the intervention (i.e., direct household connection on the premises). This answers the relevant policy question "what if an improvement is delivered to an area?" We will use multivariable log-linear binomial regression models, as pathogen infection is a binary variable, and will use generalized estimating equations (GEE) to account for clustering at the subneighborhood level. We group matched on sub-neighborhood-level SES and population density, using weighting to account for unequal numbers between the intervention and control areas within each matching stratum.[104] We will additionally control for household- and individual-level confounders, including household SES, household sanitation, mother's education-level, and child sex. We may adjust for additional variables if there are found to be imbalances in potential confounders in our baseline assessment. We hypothesize that the intervention will lead to reductions in enteric pathogens among children and microbial water contamination of source water.

For additional outcomes and exposure variables of interest (**Table 1**), we will use a similar modeling approach, using log-linear binomial regression models for binary outcomes, linear regression models for continuous outcomes, and Poisson (or negative binomial) models for count outcomes. For outcomes measured at multiple time points, we will present results separately for

Page 23 of 54

BMJ Open

each given time point. For these analyses, we will control for sub-neighborhood-level SES and population density through matching, and will additionally control for household sanitation, mother's education-level, child sex, and any other variables that are imbalanced and are conceivably potential confounders. For previous enteropathogen exposure evaluated using serological measures we hypothesize that those in the intervention group will show delays in pathogen acquisition.

To assess the impact of the intervention on microbiome outcomes, we will evaluate alpha diversity (Chao1 species richness estimator, Pielou's *J* evenness estimator, and the Shannon diversity index [105]) using the same modeling approach as described above for continuous outcomes. Linear discriminate effect size (LEfSe) analyses will be used to evaluate specific 16S rRNA gene-based Operational Taxonomic Units (OTUs) that differ between individuals in intervention versus control groups, and will include effect size corrections [106]. We will examine the impact of intervention group, controlling for other covariates, on community similarity using Adonis permutation models,[107] based on weighted UniFrac and Bray-Curtis distances, and evaluate and visualize differences using PCA and/or NMDS plots. We hypothesize that we will be able to observe detectable differences in gut microbiome composition in children living in intervention versus control sub-neighborhoods and we will report these differences at the individual OTU and bacterial family levels.

2.9.2. Direct Household Connection Effect

To assess the effect of having a water connection at the household or compound, we will use models similar to those described above, but accounting for a household network connection. We will also assess the interaction between the household and neighborhood network variables, which

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

will allow us to contrast and estimate indirect, direct, and total effects, as shown in **Figure 3**. We hypothesize that participants with both improved water networks in their sub-neighborhoods and household water connections will most benefit from the interventions in terms of our primary and non-primary health outcomes and exposure outcomes of interest. *2.9.3. Additional analyses*For select primary outcomes, we will assess if there is effect modification by a third variable,

such as follow-up round/age, participant sex, and household sanitation access. We will use interaction terms to identify potential interactions, and will present stratified results (e.g., separately by sex) if interactions are detected. The intervention status of subneighborhoods was set at baseline, but if control subneighborhood(s) receive the intervention after the study has started, we will perform sensitivity analyses dropping and/or recategorizing subneighborhood(s) that crossed over.

There are several analyses where we do not use the matched-design and intervention variable in our analyses. For example, we will assess associations between various water measures on health, without regard to the intervention designation. We will also examine changes in the gut microbiomes of children over time. Additional analyses will be described and documented in OSF.

2.10. Sample size and power calculations

Our minimal sample size of 548 households-half in intervention and half in matched control sub-neighborhoods-was powered for our primary outcome of prevalence of any non-viral pathogen. Utilizing data from the MapSan trial for children 10-14 months of age (J. Knee, *pers comm*) we used a control group prevalence of 70% for any non-viral pathogen, and estimated the

ability to detect a relative risk of 0.74, alpha=0.05, and power=80% using a two-sided test for significance.[32] We estimated a sub-neighborhood-level interclass correlation (ICC) of 0.05 (a moderate estimate) among our 62 designated sub-neighborhoods. We will also report on the final ICC and other assumptions of this power analysis at the end of the study. Estimates of minimum detectable effect sizes based on control prevalence of the outcome of interest (**Table S2**) show we may be adequately powered to detect a difference in some individual pathogens if those pathogens have high prevalence and/or if they are strongly associated with the water supply improvement intervention (e.g., waterborne pathogens). We target planned recruitment at 900 pregnant women in the third trimester, to account for incomplete data and loss to follow-up. We used sub-neighborhood enrollment targets proportionate to our density estimates to achieve balance across intervention and control sub-neighborhoods.

2.11. Blinding

All lab personnel and field enumerators are blinded to the intervention status of the samples and households. Participants cannot be blinded to their household-level water exposure status or cluster-level exposure status, although participants may or may not know about water improvements in their particular neighborhood. A primary analyst external to the core data management team is blinded to the group assignments until the data cleaning and primary analysis are completed. Details of these procedures are included in the Supplementary Material. Unblinding will occur only after primary outcome models are developed and compared between two independent analysts. Analyses examining the impact of the intervention on non-primary outcomes or exposures of interest will not be unblinded until after analyses that examine the impact of the intervention on our primary outcomes have been completed. Purely observational analyses that do

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

not require information on intervention group may be completed before unblinding occurs.

3. ETHICS & DISSEMINATION

The study protocol, informed consent forms, and data collection tools were approved by 1) Mozambique National Bio-Ethics Committee for Health (IRB00002657) and (2) Emory University's Institutional Review Board (IRB00098584). Prior to enrollment, study staff fully explained and carried out the consent process and documented the procedure. Subjects provided written consent with a signature. In the case of illiteracy of the subject, study staff verbally summarized the material with the subject, and the participants were required to provide written consent by marking the document with a thumbprint. As this study is a natural experiment that the investigators do not control, we do not have a data monitoring committee or any interim stopping guidelines. Enrollment for this study began during the COVID-19 pandemic, and precautions were taken to secure the safety of study staff and participants based on guidance from INS, Emory University, and the University of Washington.

Any changes to this published protocol will be noted in OSF, and, where relevant, in future publications. De-identified data sufficient to replicate study findings will be publicly available on OSF upon completion and publication of the study results. A report will also be prepared and shared with the municipality and health authorities in Beira, and other relevant stakeholders. All microbial DNA sequence data will be made available through the SRA database of NCBI upon validation and/or publication of the corresponding manuscript.

AUTHORS' CONTRIBUTIONS

MM and MCF conceived of the overarching idea of evaluating the intervention; KL and MCF conceived of the specific study and secured funding; JVG and MCF designed the analysis plan, with input from LW; ZAC, AJ, BM, SH, SM, JSS, and MKMP designed protocols for recruitment of participants and oversaw collection of field data, with guidance from RN and JLM; JSS and JVG oversaw data management, with help from SH, MKMP, SM, and CSFS; CV and CSFS designed specimen management and laboratory protocols; TC, JB, and LW advised on study design, epidemiological approaches and research methods; RN, MM, and JLM provided oversight on relevant scientific questions in Mozambique; RN, JSS, CV, KL, MCF, ZAC, SM, and MKMP managed human subjects protocol submissions; KK and KL oversaw microbiome analysis approach; RN oversaw parasitology analysis approach; KL and JVG wrote the manuscript, with input from all authors.

FUNDING STATEMENT

This work was supported by National Institute of Allergy and Infectious Diseases (NIAID) Grant #R01AI130163, National Institute of Environmental Health Sciences (NIEHS) Grant #5T32ES12870), and a contract from Autoridade Reguladora de Água, Instituto Público (AURA, IP), Mozambique. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare no conflicts of interests.

REFERENCES

- 1 Centers for Disease Control and Prevention. Achievements in Public Health, 1900-1999: Control of Infectious Diseases. *MMWR Morb Mortal Wkly Rep* 1999;**48**:621–9.
- 2 Cutler D, Miller G. The role of public health improvements in health advances: the twentieth-century United States. *Demography* 2005;**42**:1–22.
- 3 Pruss-Ustun A, Bartram J, Clasen T, *et al.* Burden of disease from inadequate water, sanitation and hygiene in low- and middle-income settings: a retrospective analysis of data from 145 countries. *Trop Med Int Health* 2014;**19**:894–905. doi:10.1111/tmi.12329
- 4 Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Under-5 Mortality and Adult Mortality 1950-2019. Seattle, United States of America: : Institute for Health Metrics and Evaluation (IHME) 2020.
- 5 Victora CG, Christian P, Vidaletti LP, *et al.* Revisiting maternal and child undernutrition in low-income and middle-income countries: variable progress towards an unfinished agenda. *The Lancet* 2021;**397**:1388–99. doi:10.1016/S0140-6736(21)00394-9
- 6 Danaei G, Andrews KG, Sudfeld CR, *et al.* Risk Factors for Childhood Stunting in 137 Developing Countries: A Comparative Risk Assessment Analysis at Global, Regional, and Country Levels. *PLoS Med* 2016;**13**:e1002164. doi:10.1371/journal.pmed.1002164
- 7 World Bank/Interational Monitary Fund (IMF). Global Montioring Report 2013: Rural-Urban Dynamics and the Millennium Development Goals. Washington, D.C.: World Bank/IMF 2013.
- 8 United Nations Department of Economic and Social Affairs. *World Urbanization Prospects: The 2018 Revision*. United Nations 2019. doi:10.18356/b9e995fe-en
- 9 United Nations-HABITAT. *The State of African Cities: Re-imagining sustainable urban transitions*. 2014. http://unhabitat.org/books/state-of-african- cities-2014-re-imagining-sustainable-urban-transitions/
- 10 Jones O, Mansour G, Burr P. THE STATE OF WASH FINANCING IN EASTERN AND SOUTHERN AFRICA Regional Level Assessment. Nairobi: : UNICEF Eastern and Southern Africa Regional Office 2019.
- 11 Bain R, Cronk R, Wright J, *et al.* Fecal contamination of drinking-water in low- and middle-income countries: a systematic review and meta-analysis. *PLOS Med* 2014;**11**:e1001644. doi:10.1371/journal.pmed.1001644
- 12 Hunter PR, Zmirou-Navier D, Hartemann P. Estimating the impact on health of poor reliability of drinking water interventions in developing countries. *Sci Total Environ* 2009;**407**:2621–4. doi:10.1016/j.scitotenv.2009.01.018

GLAAS 2018/2019 country survey. World Health Organization 2018. https://www.who.int/publications/m/item/glaas-2018-2019-country-survey (accessed 1 Oct 2021). United Nations. Investing in Water and Sanitation: Increasing Access, Reducing Inequalities. GLAAS 2014 Findings-- Special report for Africa. UN Water & World Health Organization 2015. http://www.who.int/water sanitation health/glaas/glaas2014-africa-region.pdf UN-Water. UN-Water GLAAS 2019: National systems to support drinking-water, sanitation and hygiene - Global status report 2019. 2019. https://www.unwater.org/publications/un-water-glaas-2019-national-systems-to-support-drinking-water-sanitation-and-hygiene-global-status-report-2019/ (accessed 1 Oct 2021). Ercumen A, Arnold BF, Kumpel E, et al. Upgrading a piped water supply from intermittent to continuous delivery and association with waterborne illness: A matched cohort study in urban India. PLOS Med 2015;12:e1001892. doi:10.1371/journal.pmed.1001892 Pickering AJ, Crider Y, Sultana S, et al. Effect of in-line drinking water chlorination at the point of collection on child diarrhoea in urban Bangladesh: a double-blind, cluster-randomised controlled trial. Lancet Glob Health 2019;7:e1247-56. doi:10.1016/S2214-109X(19)30315-8 Reese H, Routray P, Torondel B, et al. Assessing longer-term effectiveness of a combined household-level piped water and sanitation intervention on child diarrhoea, acute respiratory infection, soil-transmitted helminth infection and nutritional status: a matched cohort study in rural Odisha, India. Int J Epidemiol 2019;48:1757-67. doi:10.1093/ije/dyz157 Clasen TF, Alexander KT, Sinclair D, et al. Interventions to improve water quality for preventing diarrhoea. Cochrane Database Syst Rev Published Online First: 20 October 2015. doi:10.1002/14651858.CD004794.pub3 Wolf J, Hubbard S, Brauer M, 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. doi:10.1016/S0140-6736(22)00937-0 Sanson-Fisher RW, Bonevski B, Green LW, et al. Limitations of the randomized controlled trial in evaluating population-based health interventions. Am J Prev Med 2007;33:155–61. doi:10.1016/j.amepre.2007.04.007 Spears D, Ban R, Cumming O. Trials and Tribulations: The Rise and Fall of the RCT in the WASH Sector. In: Randomized Control Trials in the Field of Development. Oxford: : Oxford University Press 2020. doi:10.1093/oso/9780198865360.003.0008 For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Arnold BF, Khush RS, Ramaswamy P, et al. Causal inference methods to study nonrandomized, preexisting development interventions. Proc Natl Acad Sci 2010;107:22605-10. doi:10.1073/pnas.1008944107 Boisson S, Schmidt WP, Berhanu T, et al. Randomized controlled trial in rural Ethiopia to assess a portable water treatment device. Environ Sci Technol 2009;43:5934-9. Luoto J, Mahmud M, Albert J, et al. Learning to dislike safe water products: results from a randomized controlled trial of the effects of direct and peer experience on willingness to pay. Environ Sci Technol 2012;46:6244-51. doi:10.1021/es2027967 Levy K, Nelson KL, Hubbard A, et al. Following the water: a controlled study of drinking water storage in northern coastal Ecuador. Environ Health Perspect 2008;116:1533-40. doi:10.1289/ehp.11296 Wright J, Gundry S, Conroy R. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Trop *Med Int Health* 2004;9:106–17. Levy K, Anderson L, Robb KA, et al. Household effectiveness vs. laboratory efficacy of point-of-use chlorination. Water Res 2014;54:69-77. doi:10.1016/j.watres.2014.01.037 Stauber CE, Kominek B, Liang KR, et al. Evaluation of the impact of the plastic BioSand filter on health and drinking water quality in rural Tamale, Ghana. Int J Environ Res Public Health 2012;9:3806-23. doi:10.3390/ijerph9113806 Brown J, Clasen T. High adherence is necessary to realize health gains from water quality interventions. PLoS One 2012;7:e36735. doi:10.1371/journal.pone.0036735 Enger KS, Nelson KL, Rose JB, et al. The joint effects of efficacy and compliance: a study of household water treatment effectiveness against childhood diarrhea. Water Res 2013;47:1181–90. doi:10.1016/j.watres.2012.11.034 Knee J, Sumner T, Adriano Z, et al. Effects of an urban sanitation intervention on childhood enteric infection and diarrhea in Maputo, Mozambique: A controlled before-andafter trial. eLife 2021;10:e62278. doi:10.7554/eLife.62278 Gough EK, Moulton LH, Mutasa K, et al. Effects of improved water, sanitation, and hygiene and improved complementary feeding on environmental enteric dysfunction in children in rural Zimbabwe: A cluster-randomized controlled trial. PLoS Negl Trop Dis 2020;14:e0007963. doi:10.1371/journal.pntd.0007963 Null C, Stewart CP, Pickering AJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a clusterrandomised controlled trial. Lancet Glob Health 2018;6:e316–29. doi:10.1016/S2214-109X(18)30005-6

35	Luby SP, Rahman M, Arnold BF, <i>et al.</i> Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. <i>Lancet Glob Health</i> 2018; 6 :e302–15. doi:10.1016/S2214-109X(17)30490-4
36	Post A, Ray I. Hybrid Modes of Urban Water Delivery in Low- and Middle-Income Countries. Oxf. Res. Encycl. Environ. Sci. 2020. doi:10.1093/acrefore/9780199389414.013.679
37	World Health Organization (WHO) and the United Nations Children's Fund (UNICEF). Progress on household drinking water, sanitation and hygiene 2000-2020: five years into the SDGs. Geneva: : World Health Organization (WHO) and the United Nations Children's Fund (UNICEF) 2021.
38	Walker CL, Rudan I, Liu L, <i>et al.</i> Global burden of childhood pneumonia and diarrhoea. <i>Lancet</i> 2013; 381 :1405–16. doi:10.1016/S0140-6736(13)60222-6
39	Liu L, Johnson HL, Cousens S, <i>et al.</i> Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. <i>Lancet</i> 2012; 379 :2151–61. doi:10.1016/S0140-6736(12)60560-1
40	Arnold BF, Galiani S, Ram PK, <i>et al.</i> Optimal recall period for caregiver-reported illness in risk factor and intervention studies: a multicountry study. <i>Am J Epidemiol</i> 2013; 177 :361–70. doi:10.1093/aje/kws281
41	Baqui AH, Black RE, Yunus M, <i>et al.</i> Methodological issues in diarrhoeal disease epidemiology. Definition of diarrhoea episodes. <i>Int J Epidemiol</i> 1991; 20 :1057–63.
42	Wright J, Gundry SW, Conroy RM. A review of changing episode definitions and their effects on estimates of diarrhoeal morbidity. <i>J Health Popul Nutr</i> 2007; 25 :448–55.
43	Wright JA, Gundry SW, Conroy R, <i>et al.</i> Defining episodes of diarrhoea: results from a three-country study in Sub-Saharan Africa. <i>J Health Popul Nutr</i> 2006; 24 :8–16.
44	Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in Guatemala: determining the optimal length of recall. <i>Epidemiol Infect</i> 2010; 138 :264–9. doi:10.1017/S0950268809990422
45	Trostle A, Yépez-Montufar J, Corozo-Ángulo B, <i>et al.</i> Diarrheal illnesses on the Ecuadorian coast: socio-environmental changes and concepts of health. <i>Cad Saúde Pública Rep Public Health</i> 2010; 26 :1334–44.
46	Blum D, Feachem RG. Measuring the impact of water supply and sanitation investments on diarrhoeal diseases: problems of methodology. <i>Int J Epidemiol</i> 1983; 12 :357–65.
47	Lee G, Cama V, Gilman RH, <i>et al.</i> Comparison of two types of epidemiological surveys aimed at collecting daily clinical symptoms in community-based longitudinal studies. <i>Ann Epidemiol</i> 2010; 20 :151–8. doi:10.1016/j.annepidem.2009.10.004

- 48 Schmidt WP, Arnold BF, Boisson S, *et al.* Epidemiological methods in diarrhoea studies-an update. *Int J Epidemiol* 2011;**40**:1678–92.
- 49 Thapar N, Sanderson IR. Diarrhoea in children: an interface between developing and developed countries. *Lancet* 2004;**363**:641–53. doi:10.1016/S0140-6736(04)15599-2
- 50 Holcomb DA, Knee J, Sumner T, *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. doi:10.1016/j.ijheh.2020.113496
- 51 Chard AN, Baker KK, Tsai K, *et al.* Associations between soil-transmitted helminthiasis and viral, bacterial, and protozoal enteroinfections: a cross-sectional study in rural Laos. *Parasit Vectors* 2019;**12**:216. doi:10.1186/s13071-019-3471-2
- 52 Brown J, Cumming O, Bartram J, *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. doi:10.1136/bmjopen-2015-008215
- 53 Liu J, Kabir F, Manneh J, et al. Development and assessment of molecular diagnostic tests for 15 enteropathogens causing childhood diarrhoea: a multicentre study. *Lancet Infect Dis* 2014;14:716–24. doi:10.1016/S1473-3099(14)70808-4
- 54 Lunn PG, Northrop-Clewes CA, Downes RM. Intestinal permeability, mucosal injury, and growth faltering in Gambian infants. *Lancet* 1991;**338**:907–10.
- 55 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8
- 56 Guerrant RL, Oria RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with long-term effects on child development. *Nutr Rev* 2008;**66**:487–505. doi:10.1111/j.1753-4887.2008.00082.x
- 57 Campbell DI, Elia M, Lunn PG. Growth faltering in rural Gambian infants is associated with impaired small intestinal barrier function, leading to endotoxemia and systemic inflammation. *J Nutr* 2003;**133**:1332–8.
- 58 Goto R, Mascie-Taylor CG, Lunn PG. Impact of intestinal permeability, inflammation status and parasitic infections on infant growth faltering in rural Bangladesh. *Br J Nutr* 2009;**101**:1509–16. doi:10.1017/S0007114508083554
- 59 Lunn PG. Growth retardation and stunting of children in developing countries. *Br J Nutr* 2002;**88**:109–10. doi:10.1079/BJNBJN2002652
- 60 Guerrant RL, DeBoer MD, Moore SR, *et al.* The impoverished gut--a triple burden of diarrhoea, stunting and chronic disease. *Nat Rev Gastroenterol Hepatol* 2013;**10**:220–9. doi:10.1038/nrgastro.2012.239

BMJ Open

61	Eppig C, Fincher CL, Thornhill R. Parasite prevalence and the worldwide distribution of cognitive ability. <i>Proc Biol Sci</i> 2010; 277 :3801–8. doi:10.1098/rspb.2010.0973
62	Wierzba TF, Muhib F. Exploring the broader consequences of diarrhoeal diseases on child health. <i>Lancet Glob Health</i> 2018; 6 :e230–1. doi:10.1016/S2214-109X(18)30047-0
63	Yatsunenko T, Rey FE, Manary MJ, <i>et al.</i> Human gut microbiome viewed across age and geography. <i>Nature</i> 2012; 486 :222–7. doi:10.1038/nature11053
64	Schnorr SL. The diverse microbiome of the hunter-gatherer. <i>Nature</i> 2015; 518 :S14-5. doi:10.1038/518S14a
65	Schnorr SL, Candela M, Rampelli S, <i>et al</i> . Gut microbiome of the Hadza hunter-gatherers. <i>Nat Commun</i> 2014; 5 :3654. doi:10.1038/ncomms4654
66	Clemente JC, Pehrsson EC, Blaser MJ, <i>et al.</i> The microbiome of uncontacted Amerindians. <i>Sci Adv</i> 2015;1:e1500183. doi:10.1126/sciadv.1500183
67	Martinez I, Stegen JC, Maldonado-Gomez MX, <i>et al.</i> The gut microbiota of rural papua new guineans: composition, diversity patterns, and ecological processes. <i>Cell Rep</i> 2015; 11 :527–38. doi:10.1016/j.celrep.2015.03.049
68	Obregon-Tito AJ, Tito RY, Metcalf J, <i>et al.</i> Subsistence strategies in traditional societies distinguish gut microbiomes. <i>Nat Commun</i> 2015; 6 :6505. doi:10.1038/ncomms7505
59	Gacesa R, Kurilshikov A, Vich Vila A, <i>et al.</i> Environmental factors shaping the gut microbiome in a Dutch population. <i>Nature</i> 2022; 604 :732–9. doi:10.1038/s41586-022-04567-7
70	Clasen T, Boisson S. Assessing the Health Impact of Water Quality Interventions in Low- Income Settings: Concerns Associated with Blinded Trials and the Need for Objective Outcomes. <i>Env Health Perspect</i> 2016; 124 :886–9. doi:10.1289/ehp.1510532
71	Goddard FGB, Ban R, Barr DB, <i>et al.</i> Measuring Environmental Exposure to Enteric Pathogens in Low-Income Settings: Review and Recommendations of an Interdisciplinary Working Group. <i>Environ Sci Technol</i> 2020; 54 :11673–91. doi:10.1021/acs.est.0c02421
72	Prendergast A, Kelly P. Enteropathies in the developing world: neglected effects on global health. <i>Am J Trop Med Hyg</i> 2012; 86 :756–63. doi:10.4269/ajtmh.2012.11-0743
73	Brown J, Cumming O. Stool-Based Pathogen Detection Offers Advantages as an Outcome Measure for Water, Sanitation, and Hygiene Trials. <i>Am J Trop Med Hyg</i> 2020; 102 . doi:10.4269/ajtmh.19-0639
71	INE. Resultados definitivos Censo 2017. IV Recenseamento Geral da Populacao e

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

- 75 Water & Sanitation for the Urban Poor. Supporting the provision of universal water services in Beira and Dondo. London: : Water & Sanitation for the Urban Poor 2022.
- 76 Development Projects : Water Services & Institutional Support II. World Bank. https://projects.worldbank.org/en/projects-operations/project-detail/P149377 (accessed 1 Oct 2021).
- 77 Netherlands M of FA of the K of the. Dutch Development results Mozambique. https://www.dutchdevelopmentresults.nl/2019/countries/mozambique#water (accessed 20 Apr 2022).
- 78 Schreiner M. Simple Poverty Scorecard® Poverty-Assessment Tool Mozambique. ;:127.
- 79 Blackwell M, Iacus S, King G, *et al.* Cem: Coarsened Exact Matching in Stata. *Stata J* 2009;**9**:524–46. doi:10.1177/1536867X0900900402
- 80 Porro G, King G, Iacus S. CEM: Software for Coarsened Exact Matching. *J Stat Softw* 2009;**30**. doi:10.18637/jss.v030.i09
- 81 Mbuya MN, Humphrey JH. Preventing environmental enteric dysfunction through improved water, sanitation and hygiene: an opportunity for stunting reduction in developing countries. *Matern Child Nutr* 2016;**12 Suppl 1**:106–20. doi:10.1111/mcn.12220
- 82 Cumming O, Cairneross S. Can water, sanitation and hygiene help eliminate stunting? Current evidence and policy implications. *Matern Child Nutr* 2016;**12 Suppl 1**:91–105. doi:10.1111/mcn.12258
- 83 Oria RB, Murray-Kolb LE, Scharf RJ, *et al.* Early-life enteric infections: relation between chronic systemic inflammation and poor cognition in children. *Nutr Rev* 2016;**74**:374–86. doi:10.1093/nutrit/nuw008
- 84 Platts-Mills JA, Babji S, Bodhidatta L, *et al.* Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Health* 2015;**3**:e564–75. doi:10.1016/S2214-109X(15)00151-5
- 85 Direcção Provincial de Saúde. Balanço do Plano Economico e Social. Beira: : Governo de Sofala, Direcção Provincial de Saúde, República de Moçambique 2021.
- 86 Arnold BF, Ercumen A. Negative Control Outcomes. *JAMA* 2016;**316**:2597–8. doi:10.1001/jama.2016.17700
- 87 UNICEF. Nutrition: definitions of the indicators. http://www.unicef.org/infobycountry/stats_popup2.html (accessed 5 Feb 2015).
- 88 Hartung C, Lerer A, Anokwa Y, et al. Open data kit: tools to build information services for developing regions. In: Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development - ICTD '10. London, United Kingdom: : ACM Press 2010. 1–12. doi:10.1145/2369220.2369236

BMJ Open

89	Bosch F, Palmeirim MS, Ali SM, <i>et al.</i> Diagnosis of soil-transmitted helminths using the Kato-Katz technique: What is the influence of stirring, storage time and storage temperature on stool sample egg counts? <i>PLoS Negl Trop Dis</i> 2021; 15 :e0009032. doi:10.1371/journal.pntd.0009032
90	Liu J, Gratz J, Amour C, <i>et al.</i> A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. <i>J Clin Microbiol</i> 2013; 51 :472–80. doi:10.1128/JCM.02658-12
91	Bolyen E, Rideout JR, Dillon MR, <i>et al.</i> Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. <i>Nat Biotechnol</i> 2019; 37 :852–7. doi:10.1038/s41587-019-0209-9
92	Centers for Disease Control and Prevention Division of Parasitic Diseases and Malaria. Blood Collection and Dried Blood Spot Preparation. Atlanta, GA: : Centers for Disease Control and Prevention 2019.
93	Lammie PJ, Moss DM, Brook Goodhew E, <i>et al.</i> Development of a new platform for neglected tropical disease surveillance. <i>Int J Parasitol</i> 2012; 42 :797–800. doi:10.1016/j.ijpara.2012.07.002
94	Zinkernagel RM. Maternal antibodies, childhood infections, and autoimmune diseases. <i>N Engl J Med</i> 2001; 345 :1331–5. doi:10.1056/NEJMra012493
95	Smith CM, Hill VR. Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes from water. <i>Appl Environ Microbiol</i> 2009; 75 :5284–9. doi:10.1128/AEM.00456-09
96	Vos T, Lim SS, Abbafati C, <i>et al.</i> Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. <i>The Lancet</i> 2020; 396 :1204–22. doi:10.1016/S0140-6736(20)30925-9
97	Kotloff KL, Nasrin D, Blackwelder WC, <i>et al.</i> The incidence, aetiology, and adverse clinical consequences of less severe diarrhoeal episodes among infants and children residing in low-income and middle-income countries: a 12-month case-control study as a follow-on to the Global Enteric Multicenter Study (GEMS). <i>Lancet Glob Health</i> 2019;7:e568–84. doi:10.1016/S2214-109X(19)30076-2
98	Strina A, Rodrigues LC, Cairncross S, <i>et al.</i> Factors associated with rotavirus diarrhoea in children living in a socially diverse urban centre in Brazil. <i>Trans R Soc Trop Med Hyg</i> 2012; 106 :445–51. doi:10.1016/j.trstmh.2012.02.001
99	Chen S-Y, Chiu C-H. Worldwide molecular epidemiology of norovirus infection. <i>Paediatr Int Child Health</i> 2012; 32 :128–31. doi:10.1179/2046905512Y.0000000031
100	da Silva Poló T, Peiró JR, Mendes LCN, <i>et al.</i> Human norovirus infection in Latin America. <i>J Clin Virol</i> 2016; 78 :111–9. doi:10.1016/j.jcv.2016.03.016

- 101 Nasser AM. Prevalence and fate of hepatitis a virus in water. *Crit Rev Environ Sci Technol* 1994;**24**:281–323. doi:10.1080/10643389409388470
- 102 Derrien M, Alvarez A-S, de Vos WM. The Gut Microbiota in the First Decade of Life. *Trends Microbiol* 2019;**27**:997–1010. doi:10.1016/j.tim.2019.08.001

- 103 Cumming O, Arnold BF, Ban R, *et al.* The implications of three major new trials for the effect of water, sanitation and hygiene on childhood diarrhea and stunting: a consensus statement. *BMC Med* 2019;**17**:173. doi:10.1186/s12916-019-1410-x
- 104 Iacus SM, King G, Porro G. Matching for Causal Inference Without Balance Checking. Rochester, NY: : Social Science Research Network 2008. doi:10.2139/ssrn.1152391
- 105 Magurran AE. Measuring biological diversity. Malden, Ma.: : Blackwell Pub. 2004.
- 106 Segata N, Izard J, Waldron L, *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;**12**:R60. doi:10.1186/gb-2011-12-6-r60
- 107 Oksanen J, Blanchet FG, Kindt R, *et al.* Package 'vegan': Community Ecology Package. 2016;:Ordination methods, diversity analysis and other functions for community and vegetation ecologists.https://cran.r-project.org, https://github.com/vegandevs/vegan (accessed 5 Feb 2017).

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

FIGURE CAPTIONS

Figure 1. PAASIM theory of change.

No caption needed

Figure 2. Map of PAASIM study site in Beira, Mozambique.

Map of Beira, Mozambique, with enlargement highlighting study site. Red lines indicate the new distribution system water network. Blue lines indicate other parts of the water network. Gray shaded areas indicate neighborhoods enrolled in the study.

Figure 3. PAASIM study summary diagram.

The diagram reflects the summarized a) study design and data collection, and b) data analysis approaches to isolate the effects of both overall water supply infrastructural improvements as well as the presence of a household water connection.

Figure 4. PAASIM study flow diagram.

No caption needed

Figure 5. Data and sample collection timeline for outcomes in the PAASIM study.

Data and sample collection of infant-mother dyads enrolled into the study will be used address a series of questions about the impact of community-level water system improvements on acute and chronic health outcomes in children in a low-income urban area of Mozambique.

low-inc.





215x279mm (96 x 96 DPI)

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES)

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies



Figure 3. PAASIM study summary diagram. Colour Image. The diagram reflects the summarized a) study design and data collection, and b) data analysis approaches to isolate the effects of both overall water supply infrastructural improvements as well as the presence of a household water connection.

320x151mm (300 x 300 DPI)

html





193x247mm (300 x 300 DPI)

BMJ Open: first published as 10.1136/bmjopen-2022-067341 on 2 March 2023. Downloaded from http://bmjopen.bmj.com/ on June 11, 2025 at Agence Bibliographique de Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.





Figure 5. Data and sample collection timeline for outcomes in the PAASIM study. Colour Image. Data and sample collection of infant-mother dyads enrolled into the study will be used address a series of questions about the impact of community-level water system improvements on acute and chronic health outcomes in children in a low-income urban area of Mozambique.

292x223mm (300 x 300 DPI)

Details of data collection

A data collection firm (WeConsult) performs the in-country coordination of participant enrollment, data collection, and sample collection. Enumerators are recruited from Beira; all are fluent in Portuguese (spoken by >80% of the study population)¹ and more than half speak the local language. Enumerators conduct household visits for consent and eligibility pre-birth, and at months 3, 6, 9, and 12 post-birth; active surveillance calls take place monthly (**Figure 4**). All data are collected on electronic tablets using Open Data Kit (ODK) Collect, an open source program which allows offline data collection on a mobile device.²

2.4.1. Household visit

Survey instruments consist of several modules aimed at collecting data on key indicators through structured observations, reports from respondents, and objective measurements (Table S1). The survey includes household questions (e.g., SES, demographics, nutrition and food insecurity, flooding, animals, and sanitation and handwashing facilities), questions about the mothers (e.g., demographics, medical and physical health, pregnancy and birth care, and breastfeeding practices), questions about the index child (e.g., birth outcome, nutrition, health and recent illness, anthropometry, and vaccination history), and household questions about drinking water. We also collect GPS coordinates at each household visit.

2.4.2. Active surveillance calls

A brief monthly active surveillance call is conducted by phone to gather information on prenatal and perinatal environmental exposures and illnesses, on child illness symptoms, and intake of medicines, vitamins, breastfeeding, and introduction of complementary foods. The first active surveillance pre-birth call is targeted to occur one month after enrollment and calls continue on a monthly basis between visits. To facilitate communication with the study team participants receive a 150 MZN phone credit at each visit.

2.4.3. Drinking water

Drinking water characteristics, which are on the causal pathway between the intervention and our study outcomes (**Figure 1**), will be used to provide evidence of biological plausibility of the intervention's effects on these downstream outcomes. We will characterize water access using the Joint Monitoring Program definition of "safely managed," i.e., an improved water source that is accessible on premises (located within the dwelling, yard or plot), available when needed (sufficient water available for at least 12 hours per day), and free from contamination (no *E. coli* detected in a 100mL sample).^{3,4} We also assess a number of additional variables related to drinking water, including aspects of water quality, water access, water availability, water security, water consumption, and participant satisfaction with water. These variables are evaluated throughout the study at the household level and also at the community level through objective, observed, and participant-reported measures.

2.4.4. Diarrhea

At months 3, 6, 9, and 12, we ask the caregiver to report diarrhea and blood in the stool (dysentery) of the index child in the previous week. We use the case definition of diarrhea as 3 or more loose stools in a 24 hour period.⁵ Due to concerns about reporting biases, we also include a negative control outcome⁶; caregivers report on accidents that resulted in physical injury in the previous week. In addition, we note objective characteristics of the stool samples collected, including observed blood and mucus and an infant stool form scale describing consistency (4-point scale), amount (4-point scale), and color (6 categories).⁷

2.4.5. Anthropometry

At months 3, 6, 9, and 12, we measure child: 1) length, weight, and head circumference. From these measures we will calculate Z-scores (LAZ and WAZ), and prevalence of stunting and underweight, defined

as two standard deviations below median of the reference population.⁸ We measure weight using digital baby weighing scales with 5g gradations (ADE, Model# M112600, Hamburg, Germany), and we measure height using a baby length measuring board with 1mm gradations (ADE, Model# MZ10040).⁹ Measurements are repeated twice and recorded into the survey. The enumerator is prompted to conduct a third measurement if there are differences of >1 cm in length or head circumference or if weight is off by >0.5kg in the two repeated measurements. Each child's LAZ and WAZ will be calculated using the WHO Child Growth Standards for the reference population^{10,11} and WHO Anthropometric macros.¹²

Details of sample collection, processing, and analysis

2.5.1. Stool

Sample collection: Stool of the index child is collected at months 3, 6, 9, and 12 using a diaper, which is provided to the primary caregiver before a household visit. If a fecal sample is not provided during the initial home visit, we leave the primary caregiver with additional diapers and two more attempts are made to collect the sample within 4 hours of production. If needed, we provide a cooler and cold-pack, and collect the sample within 7 hours of production.

Sample processing: Lab personnel transfer 1g fecal material from the diaper into a DNA/RNA Shield Fecal Collection Tube (Zymo Research, Irvine, CA, USA); we aim to collect three aliquots in separate collection tubes. The DNA/RNA Shield stabilization buffer lyses the cells and renders the sample DNA and RNA stable for 30 days at ambient temperature.¹³ Samples are stored and shipped at -20 °C. Nucleic acids are extracted from fecal samples using the QIAamp 96 Virus QIAcube HT Kit on a QIAcube HT (Qiagen Sciences Inc., Germantown, MD, USA), and stored at -80°C until further processing.

Analysis - parasites: Immediately after the first three aliquots are placed in the collection tubes, two additional aliquots of fecal material are taken from the diaper to prepare a slide for Kato-Katz analysis of parasite ova¹⁴. If sufficient material is available for a second slide analysis is carried out in duplicate. Samples are analyzed for hookworm (*e.g., Necator americanus, Ancylostoma duodenale*) immediately following slide preparation, and for *Ascaris* spp., *Schistosoma mansoni, Trichuris trichiura, Taenia, Enterobius vermicularis*, and *Strongyloides stercoralis* after overnight incubation at room temperature. Eligible participants are referred for deworming medicine at the 12-month visit, after returning results of the parasitological exam to study subjects in collaboration with Instituto Nacional de Saude (INS) staff in Beira.

Analysis - enteric pathogens: Extracted nucleic acids are analyzed using the TaqMan Array card (TAC, ThermoFisher Scientific, Waltham, MA, USA) assay, which allows quantification by real-time PCR via a 384-well microfluidic card for simultaneous detection of multiple viral, bacterial, and parasitic enteric pathogen targets as well as antimicrobial resistance genes. Immediately prior to nucleic acid extraction, samples used in downstream TAC assays are seeded with the Inforce 3 Bovine Vaccine (Zoetis, Parsippany-Troy Hills, NJ, USA)¹⁵ containing Bovine Herpesvirus 1 (BHV) and Bovine Respiratory Syncytial Virus (BRSV) as extrinsic controls, to monitor extraction and amplification efficiency. Pathogens will be linked to the diarrheal disease episode based on relative cycle threshold values from the TaqMan results.¹⁶ The specific targets we will test for are shown in **Table S2**. The batch of TaqMan Array Cards will be QA/QC using positive control plasmids. Initially, a standard curve for quantification will be run in duplicate, and the limit of detection and limit of quantification will be determined by running cards at a low concentration (concentration determined based on the standard curve tests) until a 95% positivity rate out of at least ten assays is obtained. Additionally, when processing samples, standard curves will be run in singlicate once per month with the positive control plasmid. A negative control will be run per card.

Analysis - Gut microbiome composition: We will characterize the gut microbial community structure and composition by sequencing of the V4 region of the 16S ribosomal RNA (rRNA) gene amplicon. Immediately prior to nucleic acid extraction, samples used in downstream 16s assays will be seeded with the ZymoBIOMICS Spike-in Control I (High Microbial Load) (Zymo Research) containing *Imtechella*

BMJ Open

halotolerans and *Allobacillus halotolerans*. Bioinformatic analyses will be completed using the QIIME2 software platform and other bioinformatics tools¹⁷.

2.5.2. Dried blood spots

Sample collection: A trained nurse collects up to six dried spots of capillary blood of the index child at 6, 9, and 12 month visits on Tropbio Filter Paper Blood Collection Disks (Cellabs, Sydney, Australia), using a 2mm lancet.

Sample processing: Samples are allowed to dry overnight, then three aliquots of two spots each are placed in a Ziploc bag with silica desiccant. Dried blood spots can be stored at ambient temperatures for up to 100 days, even in tropical climates^{18,19} but samples are stored at -20 °C and shipped at ambient temperature.²⁰

Analysis – antibodies: We will use the Luminex platform to carry out high throughput, multiplex antibody assays that enable the simultaneous measurement of quantitative antibody responses to dozens of pathogens from a single blood spot.²¹ Bead coupling of antigens will occur at the U.S. Centers for Disease Control and Prevention (CDC), and CDC collaborators will also provide support in determining appropriate antigen cut points. Our panel includes a subset of enteropathogens that have targets on the TAC assay, including *Giardia, Cryptosporidium, Entamoeba histolytica*, norovirus, *Campylobacter*, enterotoxigenic *E. coli* and *V. cholerae*, following previous similar studies.²² Our first measure will occur at 6 months, to avoid detection of maternal antibodies that wane over the first 3-6 months of life.²³

2.5.3. Drinking water

Sample collection: We collect 100-mL household drinking water samples from source and stored water at all household visits. We select the drinking water source sample by asking the mother what the source of water is for the household that would be given to the study child to drink, or to mix formula. If the source water is not in the household, the enumerator walks with the mother to the source (e.g., a neighbor's house or public standpipe) to collect a sample and GPS location. We select the stored water sample by asking the mother for any water in the household that is used for drinking purposes; stored water is water stored in a jerry can, bucket, or other container in the household for later consumption.

To complement the household sampling, we also collect samples from a selection of 45 public standpipes located within the study area and 55 additional public standpipes located elsewhere in the city of Beira. At public standpipes we also measure water pressure by measuring time to fill a fixed volume (1L or 5L, depending on the pressure).

Sample processing and analysis: All samples are placed on cold packs after collection for transport to a lab in Beira. Samples are processed for fecal indicator bacteria within six hours of collection using Colilert-18 reagent and the Quanti-Tray/2000 MPN method (IDEXX Laboratories, Westbrook, ME, USA). Free and total chlorine levels are measured using a DR300 Pocket Colorimeter (Hach Company, Loveland, CO, USA) and DPD powder pillows (Hach Company)²⁴. Additional physiochemical parameters (pH, conductivity, and turbidity) are measured for public standpipe water samples using a Pocket Pro+ Multi 2 Tester and TL2300 turbidity meter (Hach Company).

Details of Blinding

At enrollment, households are assigned a unique identifier independent of intervention status. The core data management team conducts quality assurance using geocoded data to ensure group exposure status aligns with cluster-level designations of the exposure (i.e., intervention vs control areas). All geocoded data and group exposure indicators are removed prior to an external analyst performing the data cleaning and primary analysis. Data are cleaned, including decisions on missing data, outliers, and variable categorizations, before the analyst receives any group exposure. The analyst performs the primary analyses making comparisons between undefined group exposures. Once the analysis models are finalized, these rehearsal results are input into tables to create table shells for the final analyses. The primary analyst then receives a

masking key, with true cluster-level designations of the exposure, and reruns the code with the appropriate group exposure, and re-inputs the final results into the table shells.

tor per terien ony

Catagony	Indicator	Time nei	ot				
Category	Indicator	Pre-birth	3mo	6mo	9mo	12mo	
Household-level							in o
General characteristics	Socio-economic status	Х	*	*	*	Х	Respondents reported Observetings of households
	Demographics (household ownership, household members, children under 5 years of age, primary wage earner)	Х	*	*	*	Х	Responding reported religing 2023
Nutrition	Food insecurity Types of food consumed by household members vesterday	Х	Х	Х	Х	Х	Responding reported Responding reported
Sanitation	Access to improved sanitation	Х	х	Х	Х	Х	Observersion of household latring
	Location of sanitation facility	X	X	X	X	X	Observertion of household comp
	Sharing of sanitation facility	Х	Х	Х	Х	Х	Respondent reported
	Sanitation facility characteristics (serviceable, drop hole cover, smooth and cleanable floor)	x	Х	Х	Х	Х	Obser 문화(1986) of household latring
	Presence of human feces in the household compound	x	X	X	Х	Х	Observertion of household comp
	Trash disposal	Х	X	Х	Х	Х	Respondent reported
Handwashing	Access to handwashing facility with soap and water	Х	Х	X	x	Х	Observatioရှိ of household comp
Flooding	Flooding of household compound	Х	Х	Х	Х	X	<u>Respondent</u> reported
Animals	Presence of animals in the household or compound (chickens, ducks/turkey, dogs, cats, pigs, sheep, goats, rabbits, donkeys)	Х	Х	Х	х	х	Respondent reported
	Presence of animal feces in the household compound	Х	Х	Х	Х	Х	Observation of household comp ဖို့ မှု
	Child(ren) contact with animals	Х	Х	Х	Х	Х	Responde
Health	Deworming history (pre-school age children, school age children, and mother)		Х	Х	х	Х	Responder <mark>g</mark> reported
Moving	Moved to another household		Х	Х	Х	Х	Responder reported

Mother-level General							I VDE O L data
Mother-level General		Pre-birth	3mo	6mo	9mo	12mo	udir udir
General							
characteristics	Demographics (age, religion, language, education, employment, marital status)	Х				Х	Responder reported
Mental and	WHO 5 well-being index	Х	Х	Х	Х	Х	Respondented
physical health	Medical diagnosis history (diabetes, high blood pressure, heart disease, urinary tract infection, kidney disease, sexually- transmitted disease, cancer, malaria)	Х					Responder Responder d to to the sup
	Vaginal health practices	Х					Respon ဦးဦးကြီ reported
	Pregnancy history	Х					Responed First reported
Pregnancy health	Expected delivery date	Х					Respond and reported
	Pre-natal care	Х					Respond reported
	(gestational diabetes, high blood pressure, placenta previa, COVID-19, Dengue, Zika, Chikungunya, Malaria) Medications, vitamins, or supplements	x	16				Al mjö Pl training Respondent reported
	during pregnancy						and
Birth care	Delivery location		Х				Resported
	Cesarean section		Х			6	Respondent reported
	Post-natal care		Х				Respondent reported
Breastfeeding	Breastfeeding intentions	Х					Respongdent reported
	Breastfeeding practices		Х	Х	Х	Χ 📂	Respondents reported
	Primary place of defecation in last week	Х	Х	Х	Х	Х	Respoinder reported
Sanitation		v	Х	Х	Х	Х	Respondert reported
Sanitation	Exclusive use of sanitation facility in last week	X					Ge

Cohorowy	Indiantos	Time nei					
Category	Indicator	Pre-birth	nt 3mo	6mo	9mo	12mo	iype on daya E S
Child-level							734
Birth	Birth outcome		Х				Respondent reported
	Birthdate		Х				Observation of child health c
	Birth weight		Х				Respondent reported
	Sex		Х				Observer of child health c
Health	Mortality 🔨		Х	Х	Х	Х	Responented
	Vaccination history (DTP, Rotavirus, Polio,		Х	Х	Х	Х	Observe do of child health c
	MMR)						I to
	Medication history		Х	Х	Х	Х	Respoint reported
	Deworming history					Х	Respondent reported
	Illness (diarrhea, dysentery, nasal		Х	Х	Х	Х	Respondent reported
	congestion, fever, vomiting, physical injury)						fron / Jata
	in the last week						
	Medical diagnosis (colic, ear infection,	4	Х	Х	Х	Х	Responder reported
	anemia, respiratory illness, malaria, asthma)						9, · //br
	and treatment history						Al t
Anthropometry	Head circumference		X	Х	Х	Х	Object ve reeasurement
	Body length		X	Х	Х	Х	Object de la construction de la
	Body weight		Х	X	Х	Х	Objective measurement
Nutrition	Types of liquid food consumed in the last		Х	Х	Х	Х	Respondent reported
	week						sim
	Types of solid food consumed in the last			Х	Х	X	Respondenter reported
	week						· tec
Sanitation	Diaper wearing		Х	Х	Х	Х	Respondent reported
	Disposal of child stools		Х	Х	Х	Х	Respondered reported
Water consumption	Child consumption of drinking water source		Х	Х	Х	Х	Respondent reported
	Drinking water treatment for child		Х	Х	Х	Х	Respondens reported
	consumption						ger
Travel	Estimated time spent away from home in					Х	Responder dr reported
	the last 12 months						Bit
							liog
							gra
							h

7 8

Category	Indicator	Time noi	nt				
category		Pre-birth	3mo	6mo	9mo	12mo	line dight
Drinking water-level							
Water access	Municipality water connection and status	Х	Х	Х	Х	Х	Observation of water meter
	Type of drinking water source	Х	Х	Х	Х	Х	Respondered reported
							Observation of water source
	Alternate source for non-drinking water	Х	Х	Х	Х	Х	Responder reported
	Alternate source for drinking water outages	Х	Х	Х	Х	Х	Responder reported
	Distance to water main	Х	Х	Х	Х	Х	GPS lo ေ
	Location of drinking water source	Х	Х	Х	Х	Х	Respon
	Distance to drinking water source	Х	Х	Х	Х	Х	GPS lo ଙ୍ଗଞ୍ଚିଷ୍ଡ୍ର
	Time to collect water (minutes for round trip	Х	Х	Х	Х	Х	Respon ସୁର୍ବି ଅନୁ reported
	and trips per week)						l and
Nater availability	Water availability (hours per day and days	Х	Х	Х	Х	Х	Responder reported
	per week)						ata Arron
	Household water storage (large and small	X	Х	Х	Х	Х	Responed Responed
	containers)	h					ning:
	Piped water flow rate (liters per minute)	X	Х	Х	Х	Х	Objective measurement
	Water availability at handwashing facility	X	X	Х	Х	Х	Observation of handwashing facility
	Water availability at sanitation facility	X	X	Х	Х	Х	Observertion facility
Drinking water	Treatment of stored water	Х	X	X	Х	Х	Resported
quality	Drinking water free/total chorine (source	Х	X	X	Х	Х	Objective measurement
	and stored water)						<u><u>s</u></u>
Nater security	Sharing of water connection	X	X	X	X	X	Responed reported
	Household water insecurity experiences	X	X	X	X	X	
	Sufficient quantity of drinking water	Х	Х	Х	Х	X	Responed reported
Jser satisfaction	Satisfied with water service	Х	Х	Х	Х	Х	Respondent reported
with water	Satisfied with water availability	Х	Х	Х	Х	Х	Resported
	Satisfied with water pressure	Х	Х	Х	Х	Х	Respondent reported
	Satisfied with water color and appearance	Х	X	X	X	X	Respondent reported
	Satisfied with water taste and smell	Х	Х	Х	Х	Х	Responderg reported
	Satisfied with water affordability	X	X	X	X	X	Respondent reported
Nater consumption	Monthly water expenses	Х	X	X	X	X	Responder
	Water usage (liters per day)	Х	Х	Х	Х	Х	Responder reported
							Observation of water meter

Target	Assayed by the	TaqMan Array Card (TAC) for th	1e PAASIM study		D D D D Probe (5' to 3')*
laiget	Gene of interest	Bacterial Pa	athogens		2
Entoroggragativo E	aaiC				
coli (EAEC)					
Diarrheagenic <i>E. coli</i> (DAEC)	afaB	GTCTCCCTGAATGTACAGCTTTCA	CMCTCTGCCACTCCACCTT	Ense	TCAAGCTGTTTGTTCGTC
Shiga toxin-producing	stx1	ACTTCTCGACTGCAAAGACGTATG	ACAAATTATCCCCTGWGCCACTATC	elan (CTCTGCAATAGGTACTCCA
E. coli (STEC)	stx2	CCACATCGGTGTCTGTTATTAACC	GGTCAAAACGCGCCTGATAG		TTGCTGTGGATATACGAGG
Enteropathogenic E.	eae	CATTGATCAGGATTTTTCTGGTGATA	CTCATGCGGAAATAGCCGTTA		ATACTGGCGAGACTATTTCAA
coli (EPEC)	bfpA	TGGTGCTTGCGCTTGCT	CGTTGCGCTCATTACTTCTG	Sup	CAGTCTGCGTCTGATTCCAA
	LT	TTCCCACCGGATCACCAA	CAACCTTGTGGTGCATGATGA	beri	CTTGGAGAGAAGAACCCT
Enterotoxigenic <i>E. coli</i>	STh	GCTAAACCAGCAGGGTCTTCAAAA	CCCGGTACAAGCAGGATTACAACA	4 e c 3 er =	TGGTCCTGAAAGCATGAA
(ETEC)	STp	TGAATCACTTGACTCTTCAAAA	GGCAGGATTACAACAAAGTT	₽ P	TGAACAACACATTTTACTGCT
<i>Shigella</i> spp./ Enteroinvasive <i>E. coli</i> (<i>EIEC</i>)	іраН	CCTTTTCCGCGTTCCTTGA	CGGAATCCGGAGGTATTGC	BES) .	CGCCTTTCCGATACCGTCTCTGCA
E. coli O157	rbdE	TTTCACACTTATTGGATGGTCTCAA	CGATGAGTTTATCTGCAAGGTGAT		CTCTCTTTCCTCTGCGGTCCT
Campylobacter jejuni/ C. coli	cadF	СТӨСТАААССАТАБАААТААААТТТСТСАС	CTTTGAAGGTAATTTAGATATGGATAATCG	irainir	CATTTTGACGATTTTTGGCTTGA
Salmonella spp.	ttr	CTCACCAGGAGATTACAACATGG	AGCTCAGACCAAAAGTGACCATC	5	CACCGACGGCGAGACCGACTTT
Vibrio cholerae	hylA	ATCGTCAGTTTGGAGCCAGT	TCGATGCGTTAAACACGAAG		ACCGATGCGATTGCCCAA
Clostridium difficile	tcdB	GGTATTACCTAATGCTCCAAATAG	TTTGTGCCATCATTTTCTAAGC		CCTGGTGTCCATCCTGTTTC
		Viral Path	nogens		3
Adenovirus	fiber	AACTTTCTCTCTTAATAGACGCC	AGGGGGCTAGAAAACAAAA		CTGACACGGGCACTCT
Astrovirus	capsid	CAGTTGCTTGCTGCGTTCA	CTTGCTAGCCATCACACTTCT		
Hepatitis G	5' UTR	CGGCCAAAAGGTGGT GGA TG	CGACGAGCCTGACGTCGGG		
Norovirus GI	ORF1-ORF2	CGCTGGATGCGATTCCATGA	CTTAGACGCCATCATCATTTAC		TGGACAGGAGATCGC
Norovirus GII	ORF1-ORF2	CAAGAACCTATGTTTAGATGGATGAG	TCGACGCCATCTTCATTCACA	ñ a	
Rotavirus	NSP3	ACCATCTWCACRTRACCCTCTATGAG	GGTCACATAACGCCCCTATAGC	- Ge	AGTTAAAAGCTAACACTGTCAAA
Sapovirus	RdRp	GAYCASGCTCTCGCYACCTAC	CCCTCCATYTCAAACACTA; TTGGCCCTCGCCACCTAC		CCRCCTATRAACCA
SARS-CoV-2	N1	GACCCCAAAATCAGCGAAAT	TCTGGTTACTGCCAGTTGAATCTG	5	ACCCCGCATTACGTTTGGTGGACC
		BMJ Op	BG Pag		
---	------------------	-----------------------------	---------------------------	-------------	---------------------------
-				copyright.	
Target	Gene of interest	Forward (5' to 3')	Reverse (5' to 3')	, inc	5' to 3')*
Cruptosporidium spp	400 514	Protozoa	n patnogens		
Cigrdia duodonalis	18S rRNA	GGGTTGTATTTATTAGATAAAGAACCA	AGGCCAATACCCTACCGTCT	ing	
Giaraia auodenaiis Satawa aka kistaktisa	18S rRNA	GACGGCTCAGGACAACGGTT	TTGCCAGCGGTGTCCG	for	
Entamoeba histolytica Cyclospora	18S rRNA	ATTGTCGTGGCATCCTAACTCA	GCGGACGGCTCATTATAACA	ы	
cayetanensis	18S rRNA	AAAAGCTCGTAGTTGGATTTCTG	AACACCAACGCACGCAGC	es r	AAGGCCGGATGACCACGA
·		Helminth	ic pathogens	elat	2022
Ascaris lumbricoides	ITS1	GCCACATAGTAAATTGCACACAAAT	GCCTTTCTAACAAGCCCAACAT	led	
Trichuris trichiura	18S rRNA	TTGAAACGACTTGCTCATCAACTT	CTGATTCTCCGTTAACCGTTGTC	ent to t	
Ancylostoma duodenale	ITS2	GAATGACAGCAAACTCGTTGTTG	ATACTAGCCACTGCCGAAACGT	Sul	ATCGTTTACCGACTTTAG
Necator americanus	ITS2	CTGTTTGTCGAACGGTACTTGC	ATAACAGCGTGCACATGTTGC	an	CTGTACTACGCATTGTATAC
		Antimicrobial	resistance genes	d da	
ntl1		GATCGGTCGAATGCGTGT	GCCTTGATGTTACCCGAGAG	(A) ata	ATTCCTGGCCGTGGTTCTGGGTTTT
mcr-1		GATCGCTGTCGTGCTCTTTG	ACCGCGCCCATGATTAATAG	mir	CGATGCTACTGATCACCACG
SHV		ТСССАТБАТБАБСАССТТТААА	TCCTGCTGGCGATAGTGGAT	s) ing	TGCCGGTGACGAACAGCTGGAG
TEM		GCATCTTACGGATGGCATGA	GTCCTCCGATCGTTGTCAGAA	j, A	
CTX-M1		CCGTCACGCTGTTRTTAGGA	AATGCCACMCCCAGYCKKCC	ltra	CAGCAAAAACTTGCCGRATT
CTX-M8-M25		ATRACACSTTCCGGCTCGAT	GCTAAYGGCGTGGTGGTATC	aini	
CTX-M2-M74		GCGCAGACCCTGAAAAAYCT	TGYGCSCGCTGRGTTTCC	ng,	ACSCTGGGYAAAGCGC
CTX-M9		GCTTTATGCGCAGACGARTG	ATCACCGCGATAAAGCACCT	an	TCGATACCRMAGATAATACGC
KPC		GGCCGCCGTGCAATAC	GCCGCCCAACTCCTTCA	d si	TGATAACGCCGCCGCCAATTTGT
NDM		ATATCACCGTTGGGATCGAC	TAGTGCTCAGTGTCGGCATC	mil	AAGGACAGCAAGGCCAAGTCG
VIM		TSTACCCRTCCAATGGTCTC	AGAAGKGCCRCTGTGTTTTT	ar to	TGTCCGTGATGGYGATGAGTTG
		Co	ntrols	ech	D
165		TGCAAGTCGAACGAAGCACTTTA	GCAGGTTACCCACGCGTTAC	nol	
numan mtDNA		CAATGAATCTGAGGAGGCTAC	CGTGCAAGAATAGGAGGTG	ogi	
3HV		GAGCAAAGCCCCGCCGAAGGA	TACGAACAGCAGCACGGGCGG	es.	
		GCAATGCTGCAGGACTAGGTATAAT	ACACTGTAATTGATGACCCCATTCT	ě	

Table S3: Details of Power Calculations

The table below details our minimum detectible effect for individual and groups of pathogens given the control group prevalence (p1) ranging from 10% to 80%. It generates the minimum detectible effect as a risk ratio (delta) and resulting minimum detectible prevalence in the comparison group.

Our assumptions are the number of sub-neighborhoods in the intervention and comparison (M1/M2), average number of households enrolled per sub-neighborhood in intervention and control (K1/K2), standard alpha (0.05) and power (.80). We estimate the minimum detectible effect for three values of for the intra-class correlation coefficient, representing low, moderate and high clustering ($\rho = .01, 0.05, 0.1$). Our power calculations rely on the moderate estimates for ICC. We account for clustering at the sub-neighborhood level.

We apply the following code using STATA v16:

power twoproportions (.2(.05).8), cluster effect(ratio) m1(36) m2(26) k1(8) k2(10) power(.8) alpha(.05) direction(lower) rho(.01/.5/.1)

alpha	power	К1	к2	M1	м2	delta	p1	p2	rho
.05	.8	8	10	36	26	.5120	.2	.1024	.01
.05	.8	8	10	36	26	.5676	.25	.1419	.01
.05	.8	8	10	36	26	.6114	.3	.1834	.01
.05	.8	8	10	36	26	.6476	.35	.2266	.01
.05	.8	8	10	36	26	.6785	.4	.2714	.01
.05	.8	8	10	36	26	.7058	.45	.3176	.01
.05	.8	8	10	36	26	.7302	.5	.3651	.01
.05	.8	8	10	36	26	.7527	.55	.414	.01
.05	.8	8	10	36	26	.7736	.6	.4642	.0
.05	.8	8	10	36	26	.7935	.65	.5158	.0
.05	.8	8	10	36	26	.8127	.7	.5689	.0
.05	.8	8	10	36	26	.8317	.75	.6238	.0
.05	.8	8	10	36	26	.8509	.8	.6807	.0
.05	.8	8	10	36	26	.3585	.2	.07169	.0
.05	.8	8	10	36	26	.4262	.25	.1066	.0
.05	.8	8	10	36	26	.4804	.3	.1441	.0
.05	.8	8	10	36	26	.5257	.35	.184	.0
.05	.8	8	10	36	26	.5648	.4	.2259	.0
.05	.8	8	10	36	26	.5994	.45	.2698	.0
.05	.8	8	10	36	26	.6308	.5	.3154	.0
.05	.8	8	10	36	26	.6596	.55	.3628	.0
.05	.8	8	10	36	26	.6867	.6	.412	.0
.05	.8	8	10	36	26	.7126	.65	.4632	.0
.05	.8	8	10	36	26	.7376	.7	.5163	.0
.05	.8	8	10	36	26	.7624	.75	.5718	.0
.05	.8	8	10	36	26	.7875	.8	.63	.0
.05	.8	8	10	 36	26	.5398	.5	.2699	
.05	.8	8	10	36	26	.5737	.55	.3155	
.05	.8	8	10	36	26	.6056	.6	.3634	
.05	.8	8	10	36	26	.6362	.65	.4135	
.05	.8	8	10	36	26	.6660	.7	.4662	
.05	.8	8	10	36	26	.6955	.75	.5216	
.05	.8	8	10	36	26	.7254	.8	.5804	
.05	.8	8	10	36	26	.7567	.85	.6432	
.05	. 8	8	10	36	2.6	.7907	9	.7117	

References Cited

 United Nations. The Mozambique third population and housing census [Internet]. 2007. Available from:

http://unstats.un.org/unsd/demographic/sources/census/country_impl.htm

- Hartung C, Lerer A, Anokwa Y, Tseng C, Brunette W, Borriello G. Open data kit: tools to build information services for developing regions. In: Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development - ICTD '10 [Internet]. London, United Kingdom: ACM Press; 2010 [cited 2022 Apr 20]. p. 1–12. Available from: http://dl.acm.org/citation.cfm?doid=2369220.2369236
- 3. United Nations Children's Fund and World Health Organization. Integrating Water Quality Testing into Household Surveys: Thematic report on drinking water. New York: UNICEF and WHO; 2020.
- WHO UNICEF Joint Monitoring Programme. JMP methodology: 2017 update & SDG baselines [Internet]. Geneva: WHO & UNICEF; 2018 [cited 2020 Mar 10]. Available from: https://washdata.org/sites/default/files/documents/reports/2018-04/JMP-2017-updatemethodology.pdf
- 5. Levine GA, Walson JL, Atlas HE, Lamberti LM, Pavlinac PB. Defining Pediatric Diarrhea in Low-Resource Settings. J Pediatr Infect Dis Soc. 2017 Sep 1;6(3):289–93.
- 6. Arnold BF, Ercumen A. Negative Control Outcomes. JAMA. 2016 Dec 27;316(24):2597–8.
- 7. Bekkali N, Hamers SL, Reitsma JB, Van Toledo L, Benninga MA. Infant Stool Form Scale: Development and Results. J Pediatr. 2009 Apr 1;154(4):521-526.e1.
- 8. UNICEF. Nutrition: definitions of the indicators [Internet]. [cited 2015 Feb 5]. Available from: http://www.unicef.org/infobycountry/stats_popup2.html
- 9. Kotloff KL, Blackwelder WC, Nasrin D, Nataro JP, Farag TH, van Eijk A, et al. The Global Enteric Multicenter Study (GEMS) of diarrheal disease in infants and young children in developing countries: epidemiologic and clinical methods of the case/control study. Clin Infect Dis. 2012 Dec;55 Suppl 4:S232-45.
- De Onis M. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva, Switzerland: World Health Organization; 2006.
- 11. World Health Organization. WHO child growth standards, SAS macros. Geneva, Switzerland: World Health Organization; 2010.
- World Health Organization. Child growth standards: WHO Anthro (version 3.2.2, January 2011) and macros [Internet]. 2016 [cited 2017 Feb 5]. Available from: http://www.who.int/childgrowth/software/en/

13.	Kazantseva J, Malv E, Kaleda A, Kallastu A, Meikas A. Optimisation of sample storage and DNA extraction for human gut microbiota studies. BMC Microbiol. 2021 May 29;21(1):158.
14.	Bosch F, Palmeirim MS, Ali SM, Ame SM, Hattendorf J, Keiser J. Diagnosis of soil- transmitted helminths using the Kato-Katz technique: What is the influence of stirring, storage time and storage temperature on stool sample egg counts? PLoS Negl Trop Dis. 2021 Jan 22;15(1):e0009032.
15.	Bivins A, Bibby K. Wastewater Surveillance during Mass COVID-19 Vaccination on a College Campus. Environ Sci Technol Lett. 2021 Sep 14;8(9):792–8.
16.	Liu J, Gratz J, Amour C, Kibiki G, Becker S, Janaki L, et al. A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. J Clin Microbiol. 2013 Feb;51(2):472–80.
17.	Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019 Aug;37(8):852–7.
18.	Marques BL, Brandao CU, Silva EF, Marques VA, Villela-Nogueira CA, Do OK, et al. Dried blood spot samples: optimization of commercial EIAs for hepatitis C antibody detection and stability under different storage conditions. J Med Virol. 2012 Oct;84(10):1600–7.
19.	Parker SP, Cubitt WD. The use of the dried blood spot sample in epidemiological studies. J Clin Pathol. 1999 Sep;52(9):633–9.
20.	Centers for Disease Control and Prevention Division of Parasitic Diseases and Malaria. Blood Collection and Dried Blood Spot Preparation. Atlanta, GA: Centers for Disease Control and Prevention; 2019. Report No.: ECL.DR.C.032.
21.	Lammie PJ, Moss DM, Brook Goodhew E, Hamlin K, Krolewiecki A, West SK, et al. Development of a new platform for neglected tropical disease surveillance. Int J Parasitol. 2012 Aug;42(9):797–800.
22.	Zambrano LD, Priest JW, Ivan E, Rusine J, Nagel C, Kirby M, et al. Use of Serologic Responses against Enteropathogens to Assess the Impact of a Point-of-Use Water Filter: A Randomized Controlled Trial in Western Province, Rwanda. Am J Trop Med Hyg. 2017 Sep;97(3):876–87.
23.	Zinkernagel RM. Maternal antibodies, childhood infections, and autoimmune diseases. N Engl J Med. 2001 Nov;345(18):1331–5.
24.	4500-Cl CHLORINE (RESIDUAL). In: Standard Methods For the Examination of Water and Wastewater [Internet]. American Public Health Association; 2018 [cited 2021 Oct 11]. (Standard Methods for the Examination of Water and Wastewater). Available from: https://www.standardmethods.org/doi/abs/10.2105/SMWW.2882.078