

BMJ Open

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 (<http://bmjopen.bmj.com>).

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

BMJ Open

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

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

	INFECTIOUS DISEASES, Epidemiology < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts

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

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 matthew.freeman@emory.edu; 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 pre-analysis is published on the Open Science Framework platform (<https://osf.io/4rkn6/>). 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.

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]

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

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,[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 quasi-experimental designs must be applied.[20–22]

47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

1
2
3 water treatment such as boiling, chlorinating or filtering water, with studies predominantly
4 conducted in rural settings.[18] Results of these trials have been mixed, since household-based
5 approaches have various limitations, including low uptake and inconsistent use,[23,24] post-
6 treatment contamination,[25–28] and a poor record of sustained use.[29,30] Household water
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

1
2
3 repeated enteric pathogen infections in the first two years of life—with or without symptomatic
4 diarrhea—are associated with serious morbidities, including gut impairment, growth shortfalls,
5
6
7 and cognitive deficits by ages 7-9 years.[53–59] Such outcomes can have profound impact on
8
9
10 health, development and well-being of individuals, communities, and entire countries.[60,61]
11
12 Host-level gastrointestinal conditions affected by environmental determinants, such as gut
13
14 microbiome composition, may also help explain the long-term sequelae of enteric infections.
15
16 While there is evidence of differences in gut microbiome composition across different cultures,
17
18 regions, and populations,[62–67] and environmental conditions,[68] to date specific WASH
19
20 determinants of these differences, such as access to piped water, have not been evaluated using
21
22 explicit counterfactuals. Thus measures of gut microbial conditions provide objective outcomes to
23
24 more accurately measure the effect of WASH interventions[69,70] and capture long-term
25
26 sequelae,[71] resulting in a more complete understanding of the health impacts.[72]
27
28
29
30
31
32

33 1.3. Overview of study

34
35 In the PAASIM study (*Pesquisa Sobre o Acesso à Água e a Saúde Infantil em Moçambique -*
36
37 *Research on Access to Water and Children's Health in Mozambique*), we address a series of
38
39 questions about the impact of community-level water system improvements on acute and chronic
40
41 health outcomes in children in a low-income urban area of Mozambique. This matched-control
42
43 cohort study follows mother-child dyads from late pregnancy through children 12 months of age,
44
45 examining the impact of living in an area with an improved water network and/or having a
46
47 household water connection on a variety of aspects of access to drinking water, microbes in the
48
49 child gut (including both pathogens and other resident gut microbes), and ultimately downstream
50
51 health outcomes (including diarrhea prevalence and growth) (**Figure 1**). This study will provide
52
53
54
55
56
57
58
59
60

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

1
2
3 from other groups, in particular the Dutch government, through infrastructure upgrades as well as
4 emergency response funds following Cyclone Idai in 2019.[76] Improvements in the city of Beira
5
6 include rehabilitation of water treatment facilities, replacing existing pipe mains that are failing,
7
8 reticulation of water supply to new areas previously without water service, improving service in
9
10 areas with poor coverage or low water pressure, and subsidizing water connection fees for the
11
12 poor.
13
14
15

16 17 18 19 **2.2 Study design**

20
21
22 Several aspects of a city-wide water supply improvement project pose challenges to
23
24 implementing a rigorous epidemiological study, particularly as people living in neighborhoods
25
26 with piped water often differ in myriad ways from people living in neighborhoods without piped
27
28 water. Water improvements to communities are often based on the needs or demographics of the
29
30 community, government or donor priorities, and engineering considerations. Provisions of water
31
32 supply to new areas previously without water service—or dramatic improvements in access and
33
34 availability—represent a fundamental development that changes the livability and sometimes the
35
36 makeup of the community. These issues lead to difficulty in finding a comparable control group
37
38 for epidemiological comparison. Furthermore, rollouts of water interventions often happen in
39
40 continuous phases over time, and these changes might coincide with other events or community
41
42 improvements that similarly impact health, making it difficult for that community to serve as its
43
44 own control in a pre post design.
45
46
47
48

49
50 By using a prospective matched cohort in this unique context of an ongoing natural experiment,
51
52 we are able to overcome many of these difficulties. The prospective nature of our study allows
53
54 better control of confounding through matching, restriction, and rigorous and thoughtful collection
55
56
57
58
59
60

1
2
3 of potential control variables. We specifically focus on one region of the city (**Figure 2**), where
4 some neighborhoods received water system improvements focused on preventing water losses
5 through replacement of the distribution system pipe system in dense, low-income settings; other
6 neighboring areas with similar demographic characteristics did not receive these improvements.
7
8 Neighborhood-level matching took place in the context of a natural experiment, where the delayed
9 rollout across the city allows us to find and compare intervention and control neighborhoods that
10 are similar in many ways, before the rollout eventually reached all potential control areas.
11
12

13
14
15
16
17
18
19 The Water Loss Reduction Project represents a subset of the improvements being carried out
20 by FIPAG, with co-funding from the Dutch government and the World Bank WASIS-II project.
21
22 The improvements are designed to reduce illegal connections, thereby increasing the water
23 pressure and quality and increasing the system's capacity for household connections. These areas
24 also received some benefits related to improvements to the water intake and distribution systems.
25
26 FIPAG undertook a campaign to offer new connections to households. These improvements were
27 completed in some informal settlements in these low income areas of the city in 2019, with other
28 adjacent neighborhoods with similar density and socioeconomic profile slated for completion in
29 future years but not within the timeframe of our study. These specific distribution system upgrades
30 therefore represent a unique opportunity to examine the impacts of community-scale water
31 improvements with neighboring communities who did not receive the intervention serving as
32 control areas for comparison. The neighborhoods under study are also in the lowest income—and
33 therefore highest need areas of the city.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 We will perform analyses that take into account the four factorial possible household types,
50 based on mother-child dyad living in a sub-neighborhood with or without the improved water
51 network and with or without a household connection (**Figure 3**). Our primary analyses focus on
52
53
54
55
56
57
58
59
60

1
2
3 assessing A) the *total network effect*, by comparing subjects living in sub-neighborhoods with the
4 improved water network (Household Types 1 and 2) to those living in sub-neighborhoods without
5 these improvements (Household Types 3 and 4); and B) the *direct household connection effect*, by
6 comparing subjects with household water connections on their premises (Household Types 1 and
7 3) to those without a connection (Household Types 2 and 4). Depending on the results of these two
8 primary analyses, secondary analyses may evaluate the other comparisons depicted in **Figure 3**.
9

10
11
12
13
14
15
16
17 The reasons for some neighborhoods receiving the improvements and others not were a result
18 of resource constraints. We conducted a population-based survey (described below) that allowed
19 us to both restrict and match study sub-neighborhoods, thereby creating a statistically appropriate
20 counterfactual for strong internal validity. The evaluation of a real-world intervention delivered in
21 an informal urban setting provides strong external validity for estimating the effects of similar
22 interventions in other low- and middle-income country urban sites. Our study design allows us to
23 isolate the effects of both overall water supply infrastructural improvements as well as the presence
24 of a household water connection. The presence of control areas not receiving upgrades adjacent to
25 intervention areas that are matched on socioeconomic and density variables is unique to this study
26 location. We collect data at multiple timepoints for each study household, allowing us to examine
27 variability in each of the measures taken from each household, rather than at a single point in time,
28 and also allowing for longitudinal analyses of the households and the individual enrolled subjects.
29 We also employ rigorous measures of exposure and novel and objective outcome measures,
30 including gut microbiome composition and molecular detection of enteric pathogens.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 **2.3. Patient & Public involvement**

52
53
54
55
56
57
58
59
60

1
2
3 The executive secretary of AURA, IP was directly involved in the formation of the research
4 questions, and FIPAG personnel were also engaged from the initiation of the project in helping
5 develop the study design. Our team also received input from other public agency stakeholders
6 during workshops that were held prior to initiation of the study. Study subjects and members of
7 the general public were not involved in the study design. We provide regular updates with data
8 summaries to public agency stakeholders, and plan to disseminate the main results to all study
9 participants and also through public presentations for stakeholders in both Beira and Maputo.
10
11
12
13
14
15
16
17
18
19
20
21

22 **2.4. Sub-neighborhood selection**

23
24 Sub-neighborhood eligibility, selection, and matching of intervention and control sub-
25 neighborhoods occurred through a two-step process:
26
27

28 1) *Intervention designation*: Our study is a natural experiment, where the investigators had no
29 control over the selection or timing of the intervention implementation. The study flow diagram is
30 shown in **Figure 4**. We worked with FIPAG to determine which neighborhoods in Beira were to
31 receive water distribution system upgrades prior to initiation of enrollment (2020) and before the
32 end of the study (2022). FIPAG provided maps and timelines for construction works related to the
33 upgrades, and the specific areas participating in the water loss reduction project. We also worked
34 with FIPAG and through satellite imagery to identify similarly dense low income areas in Beira
35 that were not slated to receive water network upgrades. A total of 17 potential neighborhoods were
36 considered for inclusion in the study, and neighborhoods were divided into 80 sub-neighborhoods,
37 delineated along natural boundaries such as roads or waterways. “Intervention” sub-neighborhoods
38 include areas with the upgraded water distribution system. “Control” sub-neighborhoods include
39 areas not receiving these improvements during the time period of the study. Within both
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 intervention and control sub-neighborhoods, some households have a connection to the water
4 system and others do not. We excluded nine sub-neighborhoods that were likely partially
5 contaminated by proximity to the intervention or that were scheduled to receive the interventions
6 within the timetable of our project; some control sub-neighborhoods are slated to receive the
7 intervention after completion of our study.
8
9

10
11
12
13
14
15 2) *Matching & Restriction*: We followed the suggestion of Arnold et al. (2010) to use baseline
16 (preintervention) data at the community level to match intervention to control communities when
17 randomization is not possible.[22] To characterize subneighborhoods for further matching and
18 restriction, we performed a population-based community survey in November-December, 2020 of
19 approximately 1,700 households; this provided approximately a 5% proportional sample of our
20 potential study sub-neighborhoods. We used a random grid sampling approach to estimate
21 household density, using Google Earth satellite imagery, where a grid was placed over an area,
22 and a random selection of squares were selected and counted independently in duplicate, and the
23 number of houses per unit was extrapolated across unsampled squares. The survey contained
24 modules regarding household demographics, water access and practices, sanitation access and
25 practices, household assets and wealth indicators, as well as questions related to COVID-19. A
26 socioeconomic status (SES) score was constructed using the 'simple poverty scorecard'[77]
27 developed specifically for Mozambique, and scores were aggregated at the sub-neighborhood
28 level, and categorized into tertiles.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46
47 We matched intervention sub-neighborhoods to control sub-neighborhoods, using coarsened
48 exact matching,[78,79] with intervention sub-neighborhoods being matched to control sub-
49 neighborhoods within the same tertile of both SES and population density. Four neighborhoods
50 (encompassing nine sub-neighborhoods) were found to be outliers in terms of their sub-
51
52
53
54
55
56
57
58
59
60

1
2
3 neighborhood-level SES or sanitation, and were excluded from the study sampling frame.
4
5 Ultimately we designated 36 intervention sub-neighborhoods, with an estimated 16,800
6
7 households, and 26 control sub-neighborhoods, with an estimated 9,500 households.
8
9

10 11 12 **2.5. Participant recruitment, eligibility, and retention**

13
14 We recruit pregnant women at the last trimester of pregnancy and follow the infant-mother
15
16 dyads until the child is 12 months old (**Figure 5**). We selected the first 12 months of life because
17
18 it is a critical development window,[80–82] it is a time when children are most at risk of acute and
19
20 chronic effects of enteropathogen infection,[83] and it is a short enough period of time to avoid
21
22 changes in water access that might occur. We recruit mothers at the end of their pregnancy so we
23
24 can collect data on household risk factors (including drinking water quality) during the gestational
25
26 period. Active recruitment occurs through identification of pregnant women in the 2020
27
28 population-based survey, lists of pregnant women visiting local health centers for pre-natal care,
29
30 and study staff visiting under-enrolled sub-neighborhoods throughout the recruitment period.
31
32 Based on Ministry of Health data for Sofala Province (where Beira is located), virtually all mothers
33
34 attend pre-natal clinical visits.[84] Passive strategies include referrals of pregnant women by study
35
36 participants and community leaders. We aim to have complete data on a total of 548 infant-mother
37
38 dyads, approximately evenly divided between the intervention and control groups. We will
39
40 continue to enroll dyads into both arms until we reach a minimum of 274 dyads with complete
41
42 data in each arm, to ensure temporal balance throughout the duration of the study period.
43
44
45
46
47
48

49 During an initial pre-birth visit, pregnant women are assessed for study eligibility: 1) 18 years
50
51 or older, 2) in third trimester of pregnancy, 3) resides in enrolled study cluster, 4) not planning to
52
53 move within the next 12 months, 5) carrying a singleton birth, and 6) consents to take part in the
54
55
56
57
58
59
60

1
2
3 study. We will re-assess study eligibility at each follow up visit and record if enrolled participants
4
5 have been lost to follow up.
6
7
8
9

10 **2.6. Data collection**

11
12 A local data collection firm (WE Consult) performs the in-country coordination of participant
13 enrollment, data collection, and sample collection. Enumerators conduct household visits before
14 birth for consent, eligibility, and conditions. At months 3, 6, 9, and 12 we deploy survey
15 instruments to collect data on key indicators through structured observations, reports from
16 respondents, and objective measurements (**Table S1**). We assess a number of variables related to
17 drinking water, including aspects of water quality, water access, water availability, water security,
18 water consumption, and participant satisfaction with water. Brief active surveillance calls also take
19 place monthly by phone to gather information on prenatal and perinatal environmental exposures
20 and illnesses (**Figure 5**). We ask the caregiver to report diarrhea and blood in the stool (dysentery)
21 of the index child in the previous week at the 3, 6, 9, and 12 month surveys and during active
22 surveillance calls; due to concerns about reporting biases, we also include negative control
23 outcomes.[85] At each post-birth visit we measure child: 1) length, weight, and head
24 circumference, and calculate length-for-age and weight-for-age Z-scores. Prevalence of stunting
25 and underweight are defined as two standard deviations below median of the reference
26 population.[86] All data are collected on electronic tablets using Open Data Kit (ODK) Collect,
27 an open-source program which allows offline data collection on a mobile device.[87] Additional
28 details are provided in the Supplementary Material.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54 **2.7. Sample collection, processing, and analysis**

55
56
57
58
59
60

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 384-well 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

1
2
3 antibody responses to dozens of pathogens from a single blood spot.[92] Our first measure will
4 occur at 6 months, to avoid detection of maternal antibodies that wane over the first 3-6 months of
5 life.[93]
6
7
8
9

10 11 12 *2.7.3. Drinking water*

13
14 *Sample collection:* We collect 100-mL household drinking water samples from source and
15 stored water at all household visits. To complement the household sampling, we collect samples
16 from a selection of 45 public standpipes located within the study area and 55 additional public
17 standpipes located elsewhere in the city of Beira. At public standpipes we also measure water
18 pressure by measuring time to fill a fixed volume (1L or 5L, depending on the pressure). Samples
19 are processed for fecal indicator bacteria within six hours of collection using Colilert-18 reagent
20 and the Quanti-Tray/2000 MPN method (IDEXX Laboratories, Westbrook, ME, USA), as well as
21 for free and total chlorine levels and additional physiochemical parameters (pH, conductivity, and
22 turbidity). Large volume samples will be collected from a subset of 50 households (1 L, processed
23 by membrane filtration) and 25 public standpipes (50 L, processed by dead end ultrafiltration[94])
24 in two different seasons, and tested for enteropathogens using the TAC assay.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

43 **2.8. Outcomes**

44
45 Our primary outcomes include: any bacteria or protozoa infection at age 12 months after birth;
46 individual pathogens or pathogen groups; child gut microbiome composition; and household
47 source water quality. While we measure viral pathogens using the TAC assay, they will be
48 excluded from the combined enteropathogen prevalence primary outcome measure, because
49 waterborne transmission is unlikely to dominate for these viral pathogens.[95–98] In addition to
50
51
52
53
54
55
56
57
58
59
60

1
2
3 the aforementioned reasons related to child development and infection risk, measuring pathogens
4
5 at 12 months will give us the greatest power to detect a difference, given higher levels of infection
6
7 at that age than in younger children. We will measure gut microbiome using 16S rRNA gene
8
9 amplicon sequencing in the full sample at 12 months and in a random subset of 200 children with
10
11 complete data at 3, 6, and 9 months, evenly distributed between intervention and control groups;
12
13 dyads eligible for sub-set sampling will include those with complete stool sample collection and
14
15 unchanged intervention exposure conditions. The 12-month samples will allow us to compare all
16
17 study children at a common time, when all children are consuming drinking water and once the
18
19 gut microbiome has become relatively established;[99] the longitudinal samples will allow for
20
21 comparison of development of the microbiome over time between the two groups. Microbiome
22
23 outcomes include alpha and beta diversity metrics, and identification of enriched taxonomic
24
25 groups. We also include household source water quality as a primary exposure outcome, as
26
27 understanding whether exposure to microbial contaminants is altered is considered a critical aspect
28
29 of evaluation of WASH projects.[70,100]
30
31
32
33
34

35 Additional non-primary outcomes include pathogen count, pathogen community similarity
36
37 (measured using Jaccard similarity index), diarrhea, child growth, and prior enteropathogen
38
39 infection (measured using serology on dried blood spot samples). We will measure additional
40
41 water quality exposure measures, as well as measures of exposure to the improved water system,
42
43 such as fidelity of the intervention (e.g., improvements to water quantity and coverage of
44
45 household taps) and receipt of the intervention by community members (e.g., reductions in water
46
47 insecurity, increased water use). These fidelity and uptake measures will be collected at all time
48
49 points through direct observation and respondent report. Available minimal detectable effect sizes
50
51 are summarized in **Table 1** and calculations are further detailed in the Supplementary Material.
52
53
54
55
56
57
58
59
60

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.

	Time point	Minimum detectable effect (Risk Ratio)	Anticipated control group prevalence ^a
Primary outcomes			
Patent enteric pathogen infection			
Prevalence of any bacterial or protozoan pathogen	12 mo ^b	0.74	70%
Prevalence of any bacterial pathogen	12 mo ^b	0.69	61%
Prevalence of any protozoan pathogen	12 mo ^b	0.50	32%
Any co-infection (bacterial, protozoan, or viral pathogens)	12 mo ^b	0.60	45%
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 (Same variables as patent enteric pathogen infection)	6, 9, 12 mo	--	--
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		

^aAnticipated control group prevalence based on control group prevalence for 10-14 month olds in the MapSan trial ([31] and J. Knee, pers. Comm.).

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

^cA 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.

^dThis is a conservative estimate as it does not account for weekly active surveillance.

^eDefined 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.

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 sub-neighborhood 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

1
2
3 population density through matching, and will additionally control for household sanitation,
4 mother's education-level, child sex, and any other variables that are imbalanced and are
5 conceivably potential confounders. For previous enteropathogen exposure evaluated using
6 serological measures we hypothesize that those in the intervention group will show delays in
7 pathogen acquisition.
8
9

10
11
12
13
14
15 To assess the impact of the intervention on microbiome outcomes, we will evaluate alpha
16 diversity (Chao1 species richness estimator, Pielou's J evenness estimator, and the Shannon
17 diversity index[102]) using the same modeling approach as described above for continuous
18 outcomes. Linear discriminate effect size (LEfSe) analyses will be used to evaluate specific 16S
19 rRNA gene-based Operational Taxonomic Units (OTUs) that differ between individuals in
20 intervention versus control groups, and will include effect size corrections[103]. We will examine
21 the impact of intervention group, controlling for other covariates, on community similarity using
22 Adonis permutation models,[104] based on weighted UniFrac and Bray-Curtis distances, and
23 evaluate and visualize differences using PCA and/or NMDS plots. We hypothesize that we will be
24 able to observe detectable differences in gut microbiome composition in children living in
25 intervention versus control sub-neighborhoods and we will report these differences at the
26 individual OTU and bacterial family levels.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

45 2.9.2. *Direct Household Connection Effect*

46
47 To assess the effect of having a water connection at the household or compound, we will use
48 models similar to those described above, but accounting for a household network connection. We
49 will also assess the interaction between the household and neighborhood network variables, which
50 will allow us to contrast and estimate indirect, direct, and total effects, as shown in **Figure 3**. We
51
52
53
54
55
56
57
58
59
60

1
2
3 hypothesize that participants with both improved water networks in their sub-neighborhoods and
4
5 household water connections will most benefit from the interventions in terms of our primary and
6
7 non-primary health outcomes and exposure outcomes of interest.
8
9

10 11 12 2.9.3. *Additional analyses* 13

14
15 For select primary outcomes, we will assess if there is effect modification by a third variable,
16
17 such as follow-up round/age, participant sex, and household sanitation access. We will use
18
19 interaction terms to identify potential interactions, and will present stratified results (e.g.,
20
21 separately by sex) if interactions are detected. The intervention status of subneighborhoods was
22
23 set at baseline, but if control subneighborhood(s) receive the intervention after the study has
24
25 started, we will perform sensitivity analyses dropping and/or recategorizing subneighborhood(s)
26
27 that crossed over.
28
29

30
31 There are several analyses where we do not use the matched-design and intervention variable
32
33 in our analyses. For example, we will assess associations between various water measures on
34
35 health, without regard to the intervention designation. We will also examine changes in the gut
36
37 microbiomes of children over time. Additional analyses will be described and documented in OSF.
38
39

40 41 42 2.10. **Sample size and power calculations** 43

44
45 Our minimal sample size of 548 households—half in intervention and half in matched control
46
47 sub-neighborhoods—was powered for our primary outcome of prevalence of any non-viral
48
49 pathogen. Utilizing data from the MapSan trial for children 10-14 months of age (J. Knee, *pers*
50
51 *comm*) we used a control group prevalence of 70% for any non-viral pathogen, and estimated the
52
53 ability to detect a relative risk of 0.74, $\alpha=0.05$, and power=80% using a two-sided test for
54
55
56
57
58
59

1
2
3 significance.[31] We estimated a sub-neighborhood-level interclass correlation (ICC) of 0.05 (a
4 moderate estimate) among our 62 designated sub-neighborhoods. We will also report on the final
5 ICC and other assumptions of this power analysis at the end of the study. Estimates of minimum
6 detectable effect sizes based on control prevalence of the outcome of interest (**Table S2**) show we
7 may be adequately powered to detect a difference in some individual pathogens if those pathogens
8 have high prevalence and/or if they are strongly associated with the water supply improvement
9 intervention (e.g., waterborne pathogens). We target planned recruitment at 900 pregnant women
10 in the third trimester, to account for incomplete data and loss to follow-up. We used sub-
11 neighborhood enrollment targets proportionate to our density estimates to achieve balance across
12 intervention and control sub-neighborhoods.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

29 **2.11. Blinding**

30
31 All lab personnel and field enumerators are blinded to the intervention status of the samples
32 and households. Participants cannot be blinded to their household-level water exposure status or
33 cluster-level exposure status, although participants may or may not know about water
34 improvements in their particular neighborhood. A primary analyst external to the core data
35 management team is blinded to the group assignments until the data cleaning and primary analysis
36 are completed. Details of these procedures are included in the Supplementary Material. Unblinding
37 will occur only after primary outcome models are developed and compared between two
38 independent analysts. Analyses examining the impact of the intervention on non-primary outcomes
39 or exposures of interest will not be unblinded until after analyses that examine the impact of the
40 intervention on our primary outcomes have been completed. Purely observational analyses that do
41 not require information on intervention group may be completed before unblinding occurs.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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). 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.

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. PMID: 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/International Monetary Fund (IMF). Global Monitoring 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-africa-region.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

- 1
2
3 13 GLAAS 2018/2019 country survey. World Health Organization 2018.
4 <https://www.who.int/publications/m/item/glaas-2018-2019-country-survey> (accessed 1 Oct
5 2021).
6
7
8 14 UN-Water. UN-Water GLAAS 2019: National systems to support drinking-water,
9 sanitation and hygiene - Global status report 2019. 2019.
10 [https://www.unwater.org/publications/un-water-glaas-2019-national-systems-to-support-](https://www.unwater.org/publications/un-water-glaas-2019-national-systems-to-support-drinking-water-sanitation-and-hygiene-global-status-report-2019/)
11 [drinking-water-sanitation-and-hygiene-global-status-report-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).
12
13 15 Ercumen A, Arnold BF, Kumpel E, *et al.* Upgrading a piped water supply from intermittent
14 to continuous delivery and association with waterborne illness: A matched cohort study in
15 urban India. *PLOS Med* 2015;**12**:e1001892. doi:10.1371/journal.pmed.1001892
16
17 16 Pickering AJ, Crider Y, Sultana S, *et al.* Effect of in-line drinking water chlorination at the
18 point of collection on child diarrhoea in urban Bangladesh: a double-blind, cluster-
19 randomised controlled trial. *Lancet Glob Health* 2019;**7**:e1247–56. doi:10.1016/S2214-
20 109X(19)30315-8
21
22 17 Reese H, Routray P, Torondel B, *et al.* Assessing longer-term effectiveness of a combined
23 household-level piped water and sanitation intervention on child diarrhoea, acute
24 respiratory infection, soil-transmitted helminth infection and nutritional status: a matched
25 cohort study in rural Odisha, India. *Int J Epidemiol* 2019;**48**:1757–67.
26 doi:10.1093/ije/dyz157
27
28 18 Clasen TF, Alexander KT, Sinclair D, *et al.* Interventions to improve water quality for
29 preventing diarrhoea. *Cochrane Database Syst Rev* Published Online First: 20 October
30 2015. doi:10.1002/14651858.CD004794.pub3
31
32 19 Wolf J, Pruss-Ustun A, Cumming O, *et al.* Assessing the impact of drinking water and
33 sanitation on diarrhoeal disease in low- and middle-income settings: systematic review and
34 meta-regression. *Trop Med Int Health* 2014;**19**:928–42. doi:10.1111/tmi.12331
35
36 20 Sanson-Fisher RW, Bonevski B, Green LW, *et al.* Limitations of the randomized controlled
37 trial in evaluating population-based health interventions. *Am J Prev Med* 2007;**33**:155–61.
38 doi:10.1016/j.amepre.2007.04.007
39
40 21 Spears D, Ban R, Cumming O. Trials and Tribulations: The Rise and Fall of the RCT in the
41 WASH Sector. In: *Randomized Control Trials in the Field of Development*. Oxford: :
42 Oxford University Press 2020. doi:10.1093/oso/9780198865360.003.0008
43
44 22 Arnold BF, Khush RS, Ramaswamy P, *et al.* Causal inference methods to study
45 nonrandomized, preexisting development interventions. *Proc Natl Acad Sci*
46 2010;**107**:22605–10. doi:10.1073/pnas.1008944107
47
48 23 Boisson S, Schmidt WP, Berhanu T, *et al.* Randomized controlled trial in rural Ethiopia to
49 assess a portable water treatment device. *Environ Sci Technol* 2009;**43**:5934–9.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 24 Luoto J, Mahmud M, Albert J, *et al.* Learning to dislike safe water products: results from a
4 randomized controlled trial of the effects of direct and peer experience on willingness to
5 pay. *Environ Sci Technol* 2012;**46**:6244–51. doi:10.1021/es2027967
6
7
8 25 Levy K, Nelson KL, Hubbard A, *et al.* Following the water: a controlled study of drinking
9 water storage in northern coastal Ecuador. *Environ Health Perspect* 2008;**116**:1533–40.
10 doi:10.1289/ehp.11296
11
12 26 Wright J, Gundry S, Conroy R. Household drinking water in developing countries: a
13 systematic review of microbiological contamination between source and point-of-use. *Trop*
14 *Med Int Health* 2004;**9**:106–17.
15
16 27 Levy K, Anderson L, Robb KA, *et al.* Household effectiveness vs. laboratory efficacy of
17 point-of-use chlorination. *Water Res* 2014;**54**:69–77. doi:10.1016/j.watres.2014.01.037
18
19 28 Stauber CE, Kominek B, Liang KR, *et al.* Evaluation of the impact of the plastic BioSand
20 filter on health and drinking water quality in rural Tamale, Ghana. *Int J Environ Res Public*
21 *Health* 2012;**9**:3806–23. doi:10.3390/ijerph9113806
22
23 29 Brown J, Clasen T. High adherence is necessary to realize health gains from water quality
24 interventions. *PLoS One* 2012;**7**:e36735. doi:10.1371/journal.pone.0036735
25
26 30 Enger KS, Nelson KL, Rose JB, *et al.* The joint effects of efficacy and compliance: a study
27 of household water treatment effectiveness against childhood diarrhea. *Water Res*
28 2013;**47**:1181–90. doi:10.1016/j.watres.2012.11.034
29
30 31 Knee J, Sumner T, Adriano Z, *et al.* Effects of an urban sanitation intervention on
31 childhood enteric infection and diarrhea in Maputo, Mozambique: A controlled before-and-
32 after trial. *eLife* 2021;**10**:e62278. doi:10.7554/eLife.62278
33
34 32 Gough EK, Moulton LH, Mutasa K, *et al.* Effects of improved water, sanitation, and
35 hygiene and improved complementary feeding on environmental enteric dysfunction in
36 children in rural Zimbabwe: A cluster-randomized controlled trial. *PLoS Negl Trop Dis*
37 2020;**14**:e0007963. doi:10.1371/journal.pntd.0007963
38
39 33 Null C, Stewart CP, Pickering AJ, *et al.* Effects of water quality, sanitation, handwashing,
40 and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-
41 randomised controlled trial. *Lancet Glob Health* 2018;**6**:e316–29. doi:10.1016/S2214-
42 109X(18)30005-6
43
44 34 Luby SP, Rahman M, Arnold BF, *et al.* Effects of water quality, sanitation, handwashing,
45 and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster
46 randomised controlled trial. *Lancet Glob Health* 2018;**6**:e302–15. doi:10.1016/S2214-
47 109X(17)30490-4
48
49 35 Post A, Ray I. Hybrid Modes of Urban Water Delivery in Low- and Middle-Income
50 Countries. *Oxf. Res. Encycl. Environ. Sci.* 2020.
51 doi:10.1093/acrefore/9780199389414.013.679
52
53
54
55
56
57
58
59
60

- 1
2
3 36 World Health Organization (WHO) and the United Nations Children's Fund (UNICEF).
4 Progress on household drinking water, sanitation and hygiene 2000-2020: five years into
5 the SDGs. Geneva: : World Health Organization (WHO) and the United Nations Children's
6 Fund (UNICEF) 2021.
7
8
9 37 Walker CL, Rudan I, Liu L, *et al*. Global burden of childhood pneumonia and diarrhoea.
10 *Lancet* 2013;**381**:1405–16. doi:10.1016/S0140-6736(13)60222-6
11
12 38 Liu L, Johnson HL, Cousens S, *et al*. Global, regional, and national causes of child
13 mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet*
14 2012;**379**:2151–61. doi:10.1016/S0140-6736(12)60560-1
15
16 39 Arnold BF, Galiani S, Ram PK, *et al*. Optimal recall period for caregiver-reported illness in
17 risk factor and intervention studies: a multicountry study. *Am J Epidemiol* 2013;**177**:361–
18 70. doi:10.1093/aje/kws281
19
20 40 Baqui AH, Black RE, Yunus M, *et al*. Methodological issues in diarrhoeal disease
21 epidemiology. Definition of diarrhoea episodes. *Int J Epidemiol* 1991;**20**:1057–63.
22
23 41 Wright J, Gundry SW, Conroy RM. A review of changing episode definitions and their
24 effects on estimates of diarrhoeal morbidity. *J Health Popul Nutr* 2007;**25**:448–55.
25
26 42 Wright JA, Gundry SW, Conroy R, *et al*. Defining episodes of diarrhoea: results from a
27 three-country study in Sub-Saharan Africa. *J Health Popul Nutr* 2006;**24**:8–16.
28
29 43 Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in
30 Guatemala: determining the optimal length of recall. *Epidemiol Infect* 2010;**138**:264–9.
31 doi:10.1017/S0950268809990422
32
33 44 Trostle A, Yépez-Montufar J, Corozo-Ángulo B, *et al*. Diarrheal illnesses on the
34 Ecuadorian coast: socio-environmental changes and concepts of health. *Cad Saúde Pública*
35 *Rep Public Health* 2010;**26**:1334–44.
36
37 45 Blum D, Feachem RG. Measuring the impact of water supply and sanitation investments on
38 diarrhoeal diseases: problems of methodology. *Int J Epidemiol* 1983;**12**:357–65.
39
40 46 Lee G, Cama V, Gilman RH, *et al*. Comparison of two types of epidemiological surveys
41 aimed at collecting daily clinical symptoms in community-based longitudinal studies. *Ann*
42 *Epidemiol* 2010;**20**:151–8. doi:10.1016/j.annepidem.2009.10.004
43
44 47 Schmidt WP, Arnold BF, Boisson S, *et al*. Epidemiological methods in diarrhoea studies--
45 an update. *Int J Epidemiol* 2011;**40**:1678–92.
46
47 48 Thapar N, Sanderson IR. Diarrhoea in children: an interface between developing and
48 developed countries. *Lancet* 2004;**363**:641–53. doi:10.1016/S0140-6736(04)15599-2
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 49 Holcomb DA, Knee J, Sumner T, *et al*. Human fecal contamination of water, soil, and
4 surfaces in households sharing poor-quality sanitation facilities in Maputo, Mozambique.
5 *Int J Hyg Environ Health* 2020;**226**:113496. doi:10.1016/j.ijheh.2020.113496
6
7
8 50 Chard AN, Baker KK, Tsai K, *et al*. Associations between soil-transmitted helminthiasis
9 and viral, bacterial, and protozoal enteroinfections: a cross-sectional study in rural Laos.
10 *Parasit Vectors* 2019;**12**:216. doi:10.1186/s13071-019-3471-2
11
12 51 Brown J, Cumming O, Bartram J, *et al*. A controlled, before-and-after trial of an urban
13 sanitation intervention to reduce enteric infections in children: research protocol for the
14 Maputo Sanitation (MapSan) study, Mozambique. *BMJ Open* 2015;**5**:e008215–e008215.
15 doi:10.1136/bmjopen-2015-008215
16
17 52 Liu J, Kabir F, Manneh J, *et al*. Development and assessment of molecular diagnostic tests
18 for 15 enteropathogens causing childhood diarrhoea: a multicentre study. *Lancet Infect Dis*
19 2014;**14**:716–24. doi:10.1016/S1473-3099(14)70808-4
20
21 53 Lunn PG, Northrop-Clewes CA, Downes RM. Intestinal permeability, mucosal injury, and
22 growth faltering in Gambian infants. *Lancet* 1991;**338**:907–10.
23
24 54 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet*
25 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8
26
27 55 Guerrant RL, Oria RB, Moore SR, *et al*. Malnutrition as an enteric infectious disease with
28 long-term effects on child development. *Nutr Rev* 2008;**66**:487–505. doi:10.1111/j.1753-
29 4887.2008.00082.x
30
31 56 Campbell DI, Elia M, Lunn PG. Growth faltering in rural Gambian infants is associated
32 with impaired small intestinal barrier function, leading to endotoxemia and systemic
33 inflammation. *J Nutr* 2003;**133**:1332–8.
34
35 57 Goto R, Mascie-Taylor CG, Lunn PG. Impact of intestinal permeability, inflammation
36 status and parasitic infections on infant growth faltering in rural Bangladesh. *Br J Nutr*
37 2009;**101**:1509–16. doi:10.1017/S0007114508083554
38
39 58 Lunn PG. Growth retardation and stunting of children in developing countries. *Br J Nutr*
40 2002;**88**:109–10. doi:10.1079/BJNBJN2002652
41
42 59 Guerrant RL, DeBoer MD, Moore SR, *et al*. The impoverished gut--a triple burden of
43 diarrhoea, stunting and chronic disease. *Nat Rev Gastroenterol Hepatol* 2013;**10**:220–9.
44 doi:10.1038/nrgastro.2012.239
45
46 60 Eppig C, Fincher CL, Thornhill R. Parasite prevalence and the worldwide distribution of
47 cognitive ability. *Proc Biol Sci* 2010;**277**:3801–8. doi:10.1098/rspb.2010.0973
48
49 61 Wierzba TF, Muhib F. Exploring the broader consequences of diarrhoeal diseases on child
50 health. *Lancet Glob Health* 2018;**6**:e230–1. doi:10.1016/S2214-109X(18)30047-0
51
52
53
54
55
56
57
58
59
60

- 1
2
3 62 Yatsunenkeno T, Rey FE, Manary MJ, *et al.* Human gut microbiome viewed across age and
4 geography. *Nature* 2012;**486**:222–7. doi:10.1038/nature11053
5
6 63 Schnorr SL. The diverse microbiome of the hunter-gatherer. *Nature* 2015;**518**:S14-5.
7 doi:10.1038/518S14a
8
9 64 Schnorr SL, Candela M, Rampelli S, *et al.* Gut microbiome of the Hadza hunter-gatherers.
10 *Nat Commun* 2014;**5**:3654. doi:10.1038/ncomms4654
11
12 65 Clemente JC, Pehrsson EC, Blaser MJ, *et al.* The microbiome of uncontacted Amerindians.
13 *Sci Adv* 2015;**1**:e1500183. doi:10.1126/sciadv.1500183
14
15 66 Martinez I, Stegen JC, Maldonado-Gomez MX, *et al.* The gut microbiota of rural papua
16 new guineans: composition, diversity patterns, and ecological processes. *Cell Rep*
17 2015;**11**:527–38. doi:10.1016/j.celrep.2015.03.049
18
19 67 Obregon-Tito AJ, Tito RY, Metcalf J, *et al.* Subsistence strategies in traditional societies
20 distinguish gut microbiomes. *Nat Commun* 2015;**6**:6505. doi:10.1038/ncomms7505
21
22 68 Gacesa R, Kurilshikov A, Vich Vila A, *et al.* Environmental factors shaping the gut
23 microbiome in a Dutch population. *Nature* 2022;**604**:732–9. doi:10.1038/s41586-022-
24 04567-7
25
26 69 Clasen T, Boisson S. Assessing the Health Impact of Water Quality Interventions in Low-
27 Income Settings: Concerns Associated with Blinded Trials and the Need for Objective
28 Outcomes. *Env Health Perspect* 2016;**124**:886–9. doi:10.1289/ehp.1510532
29
30 70 Goddard FGB, Ban R, Barr DB, *et al.* Measuring Environmental Exposure to Enteric
31 Pathogens in Low-Income Settings: Review and Recommendations of an Interdisciplinary
32 Working Group. *Environ Sci Technol* 2020;**54**:11673–91. doi:10.1021/acs.est.0c02421
33
34 71 Prendergast A, Kelly P. Enteropathies in the developing world: neglected effects on global
35 health. *Am J Trop Med Hyg* 2012;**86**:756–63. doi:10.4269/ajtmh.2012.11-0743
36
37 72 Brown J, Cumming O. Stool-Based Pathogen Detection Offers Advantages as an Outcome
38 Measure for Water, Sanitation, and Hygiene Trials. *Am J Trop Med Hyg* 2020;**102**.
39 doi:10.4269/ajtmh.19-0639
40
41 73 INE. Resultados definitivos Censo 2017. IV Recenseamento Geral da Populacao e
42 Habitacao de Mocambique. Maputo: : INE 2017. Www.Ine.Gov.Mzww.Ine.Gov.Mz
43
44 74 Water & Sanitation for the Urban Poor. Supporting the provision of universal water
45 services in Beira and Dondo. London: : Water & Sanitation for the Urban Poor 2022.
46
47 75 Development Projects : Water Services & Institutional Support II. World Bank.
48 <https://projects.worldbank.org/en/projects-operations/project-detail/P149377> (accessed 1
49 Oct 2021).
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 76 Netherlands M of FA of the K of the. Dutch Development results Mozambique.
4 <https://www.dutchdevelopmentresults.nl/2019/countries/mozambique#water> (accessed 20
5 Apr 2022).
6
7
8 77 Schreiner M. Simple Poverty Scorecard® Poverty-Assessment Tool Mozambique. ;:127.
9
10 78 Blackwell M, Iacus S, King G, *et al.* Cem: Coarsened Exact Matching in Stata. *Stata J*
11 2009;**9**:524–46. doi:10.1177/1536867X0900900402
12
13 79 Porro G, King G, Iacus S. CEM: Software for Coarsened Exact Matching. *J Stat Softw*
14 2009;**30**. doi:10.18637/jss.v030.i09
15
16 80 Mbuya MN, Humphrey JH. Preventing environmental enteric dysfunction through
17 improved water, sanitation and hygiene: an opportunity for stunting reduction in developing
18 countries. *Matern Child Nutr* 2016;**12 Suppl 1**:106–20. doi:10.1111/mcn.12220
19
20 81 Cumming O, Cairncross S. Can water, sanitation and hygiene help eliminate stunting?
21 Current evidence and policy implications. *Matern Child Nutr* 2016;**12 Suppl 1**:91–105.
22 doi:10.1111/mcn.12258
23
24 82 Oria RB, Murray-Kolb LE, Scharf RJ, *et al.* Early-life enteric infections: relation between
25 chronic systemic inflammation and poor cognition in children. *Nutr Rev* 2016;**74**:374–86.
26 doi:10.1093/nutrit/nuw008
27
28 83 Platts-Mills JA, Babji S, Bodhidatta L, *et al.* Pathogen-specific burdens of community
29 diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob*
30 *Health* 2015;**3**:e564–75. doi:10.1016/S2214-109X(15)00151-5
31
32 84 Direcção Provincial de Saúde. Balanço do Plano Economico e Social. Beira: : Governo de
33 Sofala, Direcção Provincial de Saúde, República de Moçambique 2021.
34
35 85 Arnold BF, Ercumen A. Negative Control Outcomes. *JAMA* 2016;**316**:2597–8.
36 doi:10.1001/jama.2016.17700
37
38 86 UNICEF. Nutrition: definitions of the indicators.
39 http://www.unicef.org/infobycountry/stats_popup2.html (accessed 5 Feb 2015).
40
41 87 Hartung C, Lerer A, Anokwa Y, *et al.* Open data kit: tools to build information services for
42 developing regions. In: *Proceedings of the 4th ACM/IEEE International Conference on*
43 *Information and Communication Technologies and Development - ICTD '10*. London,
44 United Kingdom: : ACM Press 2010. 1–12. doi:10.1145/2369220.2369236
45
46 88 Bosch F, Palmeirim MS, Ali SM, *et al.* Diagnosis of soil-transmitted helminths using the
47 Kato-Katz technique: What is the influence of stirring, storage time and storage temperature
48 on stool sample egg counts? *PLoS Negl Trop Dis* 2021;**15**:e0009032.
49 doi:10.1371/journal.pntd.0009032
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 89 Liu J, Gratz J, Amour C, *et al.* A laboratory-developed TaqMan Array Card for
4 simultaneous detection of 19 enteropathogens. *J Clin Microbiol* 2013;**51**:472–80.
5 doi:10.1128/JCM.02658-12
6
7
8 90 Bolyen E, Rideout JR, Dillon MR, *et al.* Reproducible, interactive, scalable and extensible
9 microbiome data science using QIIME 2. *Nat Biotechnol* 2019;**37**:852–7.
10 doi:10.1038/s41587-019-0209-9
11
12 91 Centers for Disease Control and Prevention Division of Parasitic Diseases and Malaria.
13 Blood Collection and Dried Blood Spot Preparation. Atlanta, GA: : Centers for Disease
14 Control and Prevention 2019.
15
16 92 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for
17 neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800.
18 doi:10.1016/j.ijpara.2012.07.002
19
20 93 Zinkernagel RM. Maternal antibodies, childhood infections, and autoimmune diseases. *N*
21 *Engl J Med* 2001;**345**:1331–5. doi:10.1056/NEJMra012493
22
23 94 Smith CM, Hill VR. Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes
24 from water. *Appl Environ Microbiol* 2009;**75**:5284–9. doi:10.1128/AEM.00456-09
25
26 95 Strina A, Rodrigues LC, Cairncross S, *et al.* Factors associated with rotavirus diarrhoea in
27 children living in a socially diverse urban centre in Brazil. *Trans R Soc Trop Med Hyg*
28 2012;**106**:445–51. doi:10.1016/j.trstmh.2012.02.001
29
30 96 Chen S-Y, Chiu C-H. Worldwide molecular epidemiology of norovirus infection. *Paediatr*
31 *Int Child Health* 2012;**32**:128–31. doi:10.1179/2046905512Y.0000000031
32
33 97 da Silva Poló T, Peiró JR, Mendes LCN, *et al.* Human norovirus infection in Latin
34 America. *J Clin Virol* 2016;**78**:111–9. doi:10.1016/j.jcv.2016.03.016
35
36 98 Nasser AM. Prevalence and fate of hepatitis a virus in water. *Crit Rev Environ Sci Technol*
37 1994;**24**:281–323. doi:10.1080/10643389409388470
38
39 99 Derrien M, Alvarez A-S, de Vos WM. The Gut Microbiota in the First Decade of Life.
40 *Trends Microbiol* 2019;**27**:997–1010. doi:10.1016/j.tim.2019.08.001
41
42 100 Cumming O, Arnold BF, Ban R, *et al.* The implications of three major new trials for the
43 effect of water, sanitation and hygiene on childhood diarrhea and stunting: a consensus
44 statement. *BMC Med* 2019;**17**:173. doi:10.1186/s12916-019-1410-x
45
46 101 Iacus SM, King G, Porro G. Matching for Causal Inference Without Balance Checking.
47 Rochester, NY: : Social Science Research Network 2008. doi:10.2139/ssrn.1152391
48
49 102 Magurran AE. *Measuring biological diversity*. Malden, Ma.: : Blackwell Pub. 2004.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 103 Segata N, Izard J, Waldron L, *et al.* Metagenomic biomarker discovery and explanation.
4 *Genome Biol* 2011;**12**:R60. doi:10.1186/gb-2011-12-6-r60
5
6
7 104 Oksanen J, Blanchet FG, Kindt R, *et al.* Package ‘vegan’: Community Ecology Package.
8 2016;:Ordination methods, diversity analysis and other functions for community and
9 vegetation ecologists.<https://cran.r-project.org>, <https://github.com/vegandevs/vegan>
10 (accessed 5 Feb 2017).
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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.

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



Figure 1. PAASIM theory of change. Mono Image.

307x40mm (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 Supérieur (ABES). Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

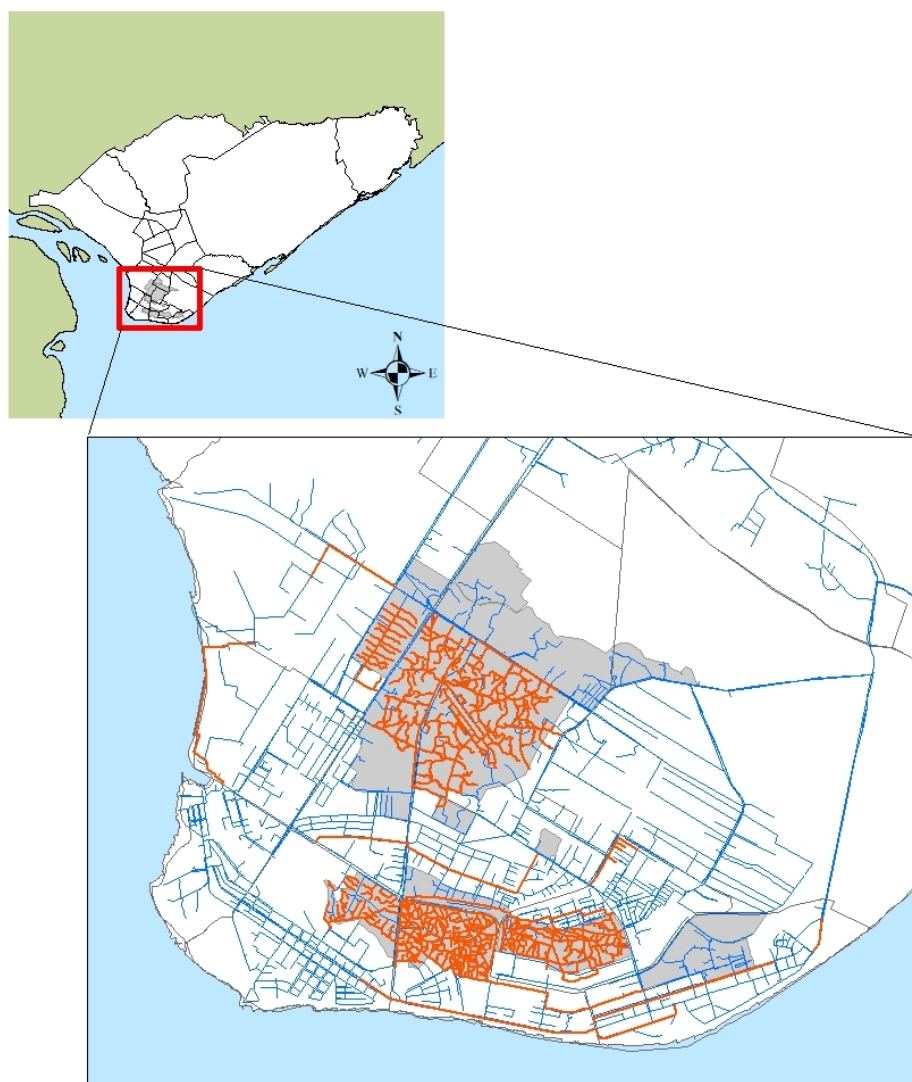


Figure 2. Map of PAASIM study site in Beira, Mozambique. Colour Image. Map of Beira, Mozambique, with enlargement highlighting study site. Red lines indicate the new distribution system water network. Blue lines indicate other parts of the water network. Gray shaded areas indicate neighborhoods enrolled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

215x279mm (96 x 96 DPI)

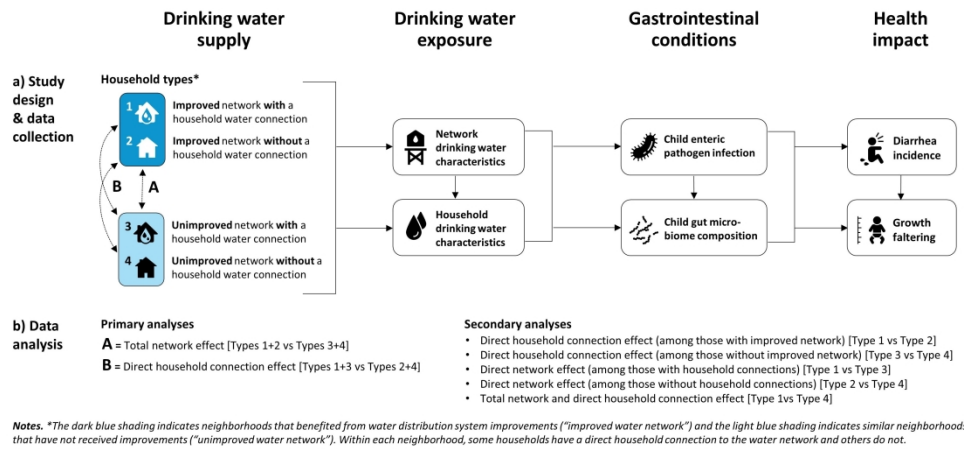


Figure 3. PAASIM study summary diagram. Colour Image. The diagram reflects the summarized a) study design and data collection, and b) data analysis approaches to isolate the effects of both overall water supply infrastructural improvements as well as the presence of a household water connection.

320x151mm (300 x 300 DPI)

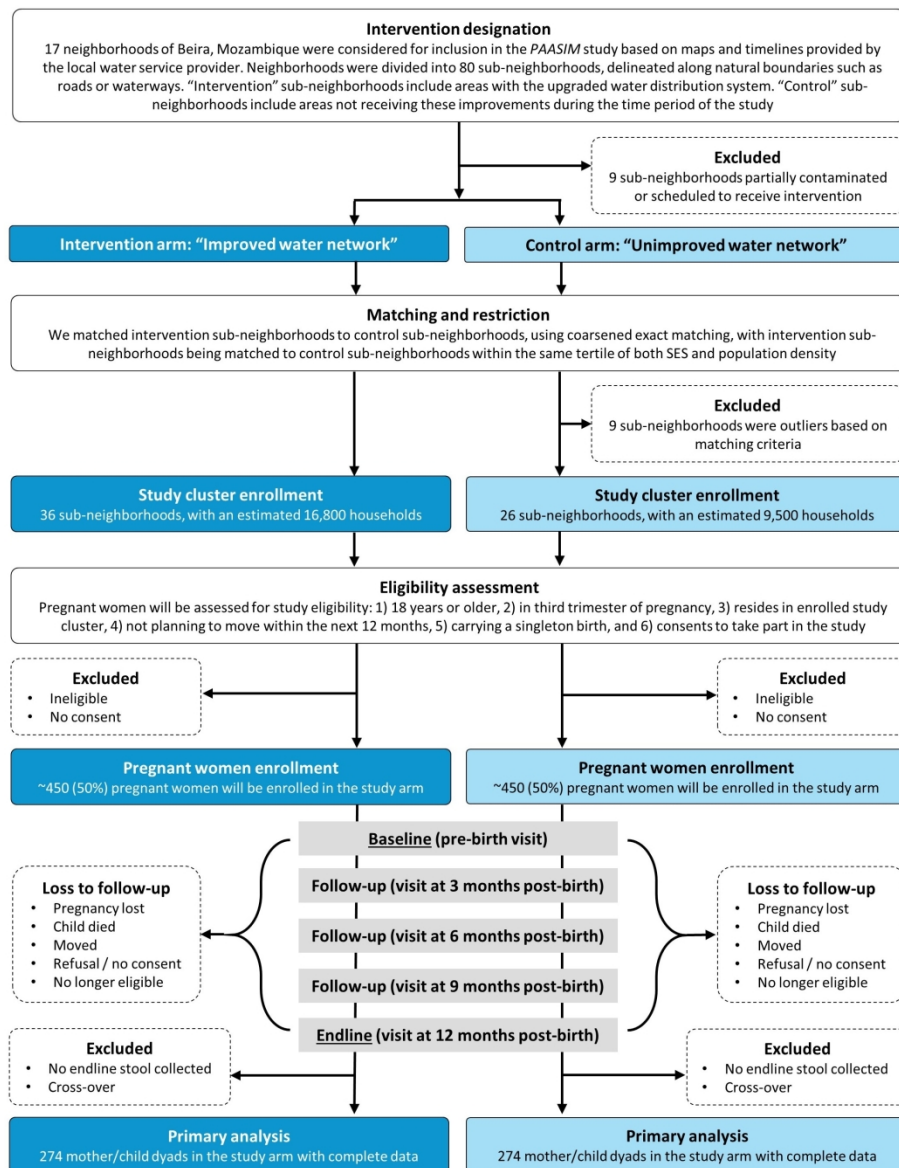


Figure 4. Colour Image. PAASIM study flow diagram.

193x247mm (300 x 300 DPI)

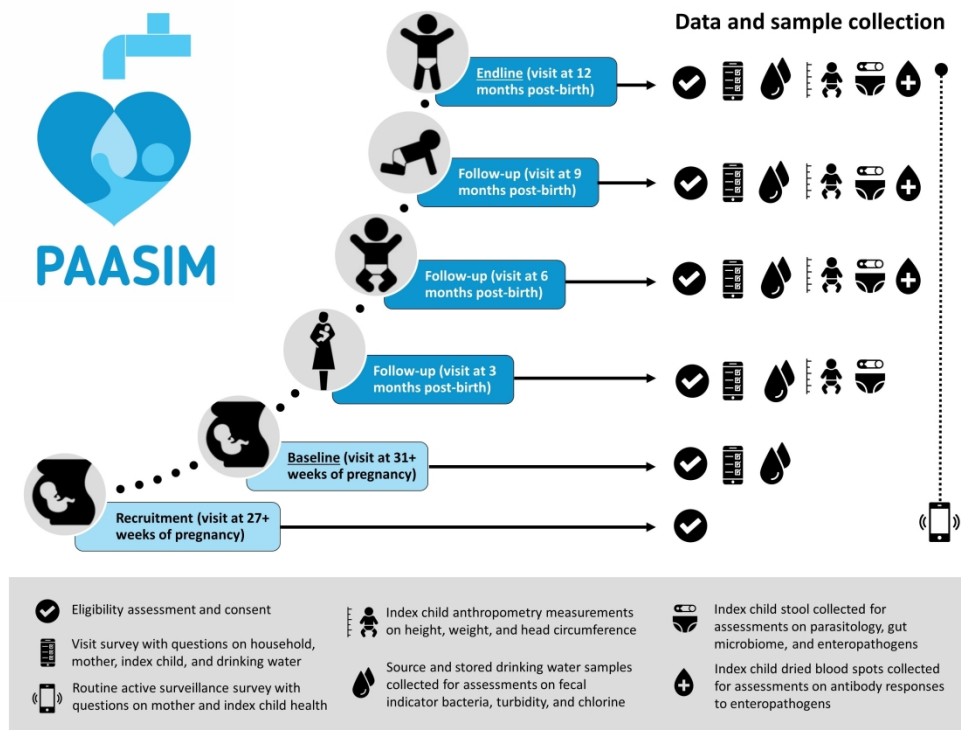


Figure 5. Data and sample collection timeline for outcomes in the PAASIM study. Colour Image. Data and sample collection of infant-mother dyads enrolled into the study will be used address a series of questions about the impact of community-level water system improvements on acute and chronic health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

292x223mm (300 x 300 DPI)

Supplementary Material

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*

1 *halotolerans* and *Allobacillus halotolerans*. Bioinformatic analyses will be completed using the QIIME2
2 software platform and other bioinformatics tools¹⁷.
3
4

5 2.5.2. Dried blood spots

6 *Sample collection:* A trained nurse collects up to six dried spots of capillary blood of the index child at
7 6, 9, and 12 month visits on Tropbio Filter Paper Blood Collection Disks (Cellabs, Sydney, Australia), using
8 a 2mm lancet.
9

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

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

24 2.5.3. Drinking water

25 *Sample collection:* We collect 100-mL household drinking water samples from source and stored water
26 at all household visits. We select the drinking water source sample by asking the mother what the source of
27 water is for the household that would be given to the study child to drink, or to mix formula. If the source
28 water is not in the household, the enumerator walks with the mother to the source (e.g., a neighbor's house
29 or public standpipe) to collect a sample and GPS location. We select the stored water sample by asking the
30 mother for any water in the household that is used for drinking purposes; stored water is water stored in a
31 jerry can, bucket, or other container in the household for later consumption.
32

33 To complement the household sampling, we also collect samples from a selection of 45 public
34 standpipes located within the study area and 55 additional public standpipes located elsewhere in the city
35 of Beira. At public standpipes we also measure water pressure by measuring time to fill a fixed volume (1L
36 or 5L, depending on the pressure).
37

38 *Sample processing and analysis:* All samples are placed on cold packs after collection for transport to
39 a lab in Beira. Samples are processed for fecal indicator bacteria within six hours of collection using
40 Colilert-18 reagent and the Quanti-Tray/2000 MPN method (IDEXX Laboratories, Westbrook, ME, USA).
41 Free and total chlorine levels are measured using a DR300 Pocket Colorimeter (Hach Company, Loveland,
42 CO, USA) and DPD powder pillows (Hach Company)²⁴. Additional physiochemical parameters (pH,
43 conductivity, and turbidity) are measured for public standpipe water samples using a Pocket Pro+ Multi 2
44 Tester and TL2300 turbidity meter (Hach Company).
45
46

47 **Details of Blinding**

48 At enrollment, households are assigned a unique identifier independent of intervention status. The core data
49 management team conducts quality assurance using geocoded data to ensure group exposure status aligns
50 with cluster-level designations of the exposure (i.e., intervention vs control areas). All geocoded data and
51 group exposure indicators are removed prior to an external analyst performing the data cleaning and primary
52 analysis. Data are cleaned, including decisions on missing data, outliers, and variable categorizations,
53 before the analyst receives any group exposure. The analyst performs the primary analyses making
54 comparisons between undefined group exposures. Once the analysis models are finalized, these rehearsal
55 results are input into tables to create table shells for the final analyses. The primary analyst then receives a
56
57
58
59
60

1
2 masking key, with true cluster-level designations of the exposure, and reruns the code with the appropriate
3 group exposure, and re-inputs the final results into the table shells.
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Table S1. Household survey data collection in the PAASIM Study

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Household-level							
General characteristics	Socio-economic status	X	*	*	*	X	Respondent reported Observation of households
	Demographics (household ownership, household members, children under 5 years of age, primary wage earner)	X	*	*	*	X	Respondent reported
Nutrition	Food insecurity	X	X	X	X	X	Respondent reported
	Types of food consumed by household members yesterday						Respondent reported
Sanitation	Access to improved sanitation	X	X	X	X	X	Observation of household latrines
	Location of sanitation facility	X	X	X	X	X	Observation of household compounds
	Sharing of sanitation facility	X	X	X	X	X	Respondent reported
	Sanitation facility characteristics (serviceable, drop hole cover, smooth and cleanable floor)	X	X	X	X	X	Observation of household latrines
	Presence of human feces in the household compound	X	X	X	X	X	Observation of household compounds
	Trash disposal	X	X	X	X	X	Respondent reported
Handwashing	Access to handwashing facility with soap and water	X	X	X	X	X	Observation of household compounds
Flooding	Flooding of household compound	X	X	X	X	X	Respondent reported
Animals	Presence of animals in the household or compound (chickens, ducks/turkey, dogs, cats, pigs, sheep, goats, rabbits, donkeys)	X	X	X	X	X	Respondent reported
	Presence of animal feces in the household compound	X	X	X	X	X	Observation of household compounds
	Child(ren) contact with animals	X	X	X	X	X	Respondent reported
Health	Deworming history (pre-school age children, school age children, and mother)		X	X	X	X	Respondent reported
Moving	Moved to another household		X	X	X	X	Respondent reported Observation of household compounds

Notes: *Asked if study household has moved to another location within the study area

36/bmjopen-2022-067341 on 2 March 2023. Downloaded from <http://bmjopen.bmj.com/> on June 11, 2025 at Agence Bibliographique de l'Enseignement Supérieur de Québec. All rights reserved. No reuse allowed without permission.

36/bmjopen-2022-06/241 on 2 March 2023. Downloaded from <http://bmjopen.bmj.com/> on June 11, 2025 at Agence Bibliographique de l'Enseignement Supérieur (A.B.E.S). Related to text and data mining, AI training, and similar technologies.

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Mother-level							
General characteristics	Demographics (age, religion, language, education, employment, marital status)	X				X	Respondent reported
Mental and physical health	WHO 5 well-being index	X	X	X	X	X	Respondent reported
	Medical diagnosis history (diabetes, high blood pressure, heart disease, urinary tract infection, kidney disease, sexually-transmitted disease, cancer, malaria)	X					Respondent reported
	Vaginal health practices	X					Respondent reported
Pregnancy health	Pregnancy history	X					Respondent reported
	Expected delivery date	X					Respondent reported
	Pre-natal care	X					Respondent reported
	Medical diagnosis during pregnancy (gestational diabetes, high blood pressure, placenta previa, COVID-19, Dengue, Zika, Chikungunya, Malaria)	X					Respondent reported
	Medications, vitamins, or supplements during pregnancy	X					Respondent reported
Birth care	Delivery location		X				Respondent reported
	Cesarean section		X				Respondent reported
	Post-natal care		X				Respondent reported
Breastfeeding	Breastfeeding intentions	X					Respondent reported
	Breastfeeding practices		X	X	X	X	Respondent reported
Sanitation	Primary place of defecation in last week	X	X	X	X	X	Respondent reported
	Exclusive use of sanitation facility in last week	X	X	X	X	X	Respondent reported
Travel	Estimated time spent away from home in the last 12 months					X	Respondent reported

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Child-level							
Birth	Birth outcome		X				Respondent reported
	Birthdate		X				Observation of child health card
	Birth weight		X				Respondent reported
	Sex		X				Observation of child health card
Health	Mortality		X	X	X	X	Respondent reported
	Vaccination history (DTP, Rotavirus, Polio, MMR)		X	X	X	X	Observation of child health card
	Medication history		X	X	X	X	Respondent reported
	Deworming history					X	Respondent reported
	Illness (diarrhea, dysentery, nasal congestion, fever, vomiting, physical injury) in the last week		X	X	X	X	Respondent reported
	Medical diagnosis (colic, ear infection, anemia, respiratory illness, malaria, asthma) and treatment history		X	X	X	X	Respondent reported
Anthropometry	Head circumference		X	X	X	X	Objective measurement
	Body length		X	X	X	X	Objective measurement
	Body weight		X	X	X	X	Objective measurement
Nutrition	Types of liquid food consumed in the last week		X	X	X	X	Respondent reported
	Types of solid food consumed in the last week			X	X	X	Respondent reported
Sanitation	Diaper wearing		X	X	X	X	Respondent reported
	Disposal of child stools		X	X	X	X	Respondent reported
Water consumption	Child consumption of drinking water source		X	X	X	X	Respondent reported
	Drinking water treatment for child consumption		X	X	X	X	Respondent reported
Travel	Estimated time spent away from home in the last 12 months					X	Respondent reported

36/bmjopen-2022-067341
 Downloaded from <http://bmjopen.bmj.com/> on June 15, 2023 at Agency Bibliographique de l'Université de Montréal
 For peer review only. All rights reserved. No reuse allowed without permission.

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Drinking water-level							
Water access	Municipality water connection and status	X	X	X	X	X	Observation of water meter
	Type of drinking water source	X	X	X	X	X	Respondent reported Observation of water source
	Alternate source for non-drinking water	X	X	X	X	X	Respondent reported
	Alternate source for drinking water outages	X	X	X	X	X	Respondent reported
	Distance to water main	X	X	X	X	X	GPS location
	Location of drinking water source	X	X	X	X	X	Respondent reported
	Distance to drinking water source	X	X	X	X	X	GPS location
	Time to collect water (minutes for round trip and trips per week)	X	X	X	X	X	Respondent reported
Water availability	Water availability (hours per day and days per week)	X	X	X	X	X	Respondent reported
	Household water storage (large and small containers)	X	X	X	X	X	Respondent reported
	Piped water flow rate (liters per minute)	X	X	X	X	X	Objective measurement
	Water availability at handwashing facility	X	X	X	X	X	Observation of handwashing facility
	Water availability at sanitation facility	X	X	X	X	X	Observation of sanitation facility
Drinking water quality	Treatment of stored water	X	X	X	X	X	Respondent reported
	Drinking water free/total chlorine (source and stored water)	X	X	X	X	X	Objective measurement
Water security	Sharing of water connection	X	X	X	X	X	Respondent reported
	Household water insecurity experiences	X	X	X	X	X	Respondent reported
	Sufficient quantity of drinking water	X	X	X	X	X	Respondent reported
User satisfaction with water	Satisfied with water service	X	X	X	X	X	Respondent reported
	Satisfied with water availability	X	X	X	X	X	Respondent reported
	Satisfied with water pressure	X	X	X	X	X	Respondent reported
	Satisfied with water color and appearance	X	X	X	X	X	Respondent reported
	Satisfied with water taste and smell	X	X	X	X	X	Respondent reported
	Satisfied with water affordability	X	X	X	X	X	Respondent reported
Water consumption	Monthly water expenses	X	X	X	X	X	Respondent reported
	Water usage (liters per day)	X	X	X	X	X	Respondent reported Observation of water meter

Table S2: Targets Assayed by the TaqMan Array Card (TAC) for the PAASIM study

Target	Gene of interest	Forward (5' to 3')	Reverse (5' to 3')	Probe (5' to 3')*
Bacterial Pathogens				
Enteroaggregative <i>E. coli</i> (EAEC)	aaiC	ATTGTCCTCAGGCATTTTAC	ACGACACCCCTGATAAACAA	TAGTGCATACTCATCATTTAAG
	aatA	CTGGCGAAAGACTGTATCAT	TTTTGCTTCATAAGCCGATAGA	TGGTTCTCATCTATTACAGACAGC
Diarrheagenic <i>E. coli</i> (DAEC)	afaB	GTCTCCCTGAATGTACAGCTTTCA	CMCTCTGCCACTCCACCTT	TCAAGCTGTTTGTTCGTC
Shiga toxin-producing <i>E. coli</i> (STEC)	stx1	ACTTCTCGACTGCAAAGACGTATG	ACAAATTATCCCCTGWGCCACTATC	CTCTGCAATAGGTACTCCA
	stx2	CCACATCGGTGTCTGTTATTAACC	GGTCAAACCGCGCTGATAG	TTGCTGTGGATATACGAGG
Enteropathogenic <i>E. coli</i> (EPEC)	eae	CATTGATCAGGATTTTTCTGGTGATA	CTCATGCGGAAATAGCCGTTA	ATACTGGCGAGACTATTTCAA
	bfpA	TGGTGCTTGCGCTTGCT	CGTTGCGCTCATTACTTCTG	CAGTCTGCGTCTGATTCCAA
Enterotoxigenic <i>E. coli</i> (ETEC)	LT	TTCCCACCGGATCACCAA	CAACCTTGTGGTGCATGATGA	CTTGGAGAGAAGAACCCT
	STh	GCTAAACCAGCAGGGTCTTCAAAA	CCCGGTACAAGCAGGATTACAACA	TGGTCCTGAAAGCATGAA
	STp	TGAATCACTTGACTCTTCAAAA	GGCAGGATTACAACAAAGTT	TGAACAACACATTTTACTGCT
<i>Shigella</i> spp./ Enteroinvasive <i>E. coli</i> (EIEC)	ipaH	CCTTTTCCGCGTTCCTTGA	CGGAATCCGGAGGTATTGC	CGCCTTCCGATACCGTCTCTGCA
<i>E. coli</i> O157	rbdE	TTTCACACTTATTGGATGGTCTCAA	CGATGAGTTTATCTGCAAGGTGAT	CTCTCTTCTCTGCGGTCCT
<i>Campylobacter jejuni</i> / <i>C. coli</i>	cadF	CTGCTAAACCATAGAAATAAAATTTCTCAC	CTTTGAAGGTAATTTAGATATGGATAATCG	CATTTTGACGATTTTTGGCTTGA
<i>Salmonella</i> spp.	ttr	CTCACCAGGAGATTACAACATGG	AGCTCAGACAAAAGTGACCATC	CACCGACGGCGAGACCGACTTT
<i>Vibrio cholerae</i>	hylA	ATCGTCAGTTTGGAGCCAGT	TCGATGCGTTAAACACGAAG	ACCGATGCGATTGCCCAA
<i>Clostridium difficile</i>	tcdB	GGTATTACCTAATGCTCCAATAG	TTTGTGCCATCATTTTCTAAGC	CCTGGTGTCCATCCTGTTTC
Viral Pathogens				
Adenovirus	fiber	AACTTTCTCTTAATAGACGCC	AGGGGGCTAGAAAACAAA	CTGACACGGGCACTCT
Astrovirus	capsid	CAGTTGCTTGCTGCGTTCA	CTTGCTAGCCATCACACTTCT	CACAGAAGAGCAACTCCATCGC
Hepatitis G	5' UTR	CGGCCAAAAGGTGGT GGA TG	CGACGAGCCTGACGTCGGG	AGGTCCCTCTGGCGCTTGTGGCGAG
Norovirus GI	ORF1-ORF2	CGTGCGATGCGATTCCATGA	CTTAGACGCCATCATCATTTAC	TGGACAGGAGATCGC
Norovirus GII	ORF1-ORF2	CAAGAACCTATGTTTAGATGGATGAG	TCGACGCCATCTTCATTACA	TGGGAGGGCGATCGCAATCT
Rotavirus	NSP3	ACCATCTWCACRTRACCCTCTATGAG	GGTCACATAACGCCCTATAGC	AGTTAAAAGCTAACACTGTCAA
Sapovirus	RdRp	GAYCASGCTCTCGCYACCTAC	CCCTCCATYTCAAACACTA; TTGGCCCTCGCCACCTAC	CCRCCTATRAACCA
SARS-CoV-2	N1	GACCCCAAATCAGCGAAAT	TCTGGTTACTGCCAGTTGAATCTG	ACCCCGCATTACGTTTGGTGGACC

Target	Gene of interest	Forward (5' to 3')	Reverse (5' to 3')	Probe (5' to 3')*
Protozoan pathogens				
<i>Cryptosporidium spp.</i>	18S rRNA	GGGTTGTATTTATTAGATAAAGAACCA	AGGCCAATACCCTACCGTCT	TGACATATCATTCAAGTTTCTGAC
<i>Giardia duodenalis</i>	18S rRNA	GACGGCTCAGGACAACGGTT	TTGCCAGCGGTGTCCG	CCCGCGGGCTCCCTGCTAG
<i>Entamoeba histolytica</i>	18S rRNA	ATTGTCGTGGCATCCTAACTCA	GCGGACGGCTCATTATAACA	TCATTGAATGAATTGGCCATT
<i>Cyclospora cayetanensis</i>	18S rRNA	AAAAGCTCGTAGTTGGATTTCTG	AACACCAACGCACGCAGC	AAGGCCGGATGACCACGA
Helminthic pathogens				
<i>Ascaris lumbricoides</i>	ITS1	GCCACATAGTAAATTGCACACAAAT	GCCTTTCTAACAAGCCCAACAT	TTGGCGGACAATTGCATGCGAT
<i>Trichuris trichiura</i>	18S rRNA	TTGAAACGACTTGCTCATCAACTT	CTGATTCTCCGTTAACCGTTGTC	CGATGGTACGCTACGTGCTTACCATGG
<i>Ancylostoma duodenale</i>	ITS2	GAATGACAGCAAACCTCGTTGTTG	ATACTAGCCACTGCCGAAACGT	ATCGTTTACCGACTTTAG
<i>Necator americanus</i>	ITS2	CTGTTTGTGCAACGGTACTTGC	ATAACAGCGTGACATGTTGC	CTGTACTACGCATTGTATAC
Antimicrobial resistance genes				
intl1		GATCGGTGCAATGCGTGT	GCCTTGATGTTACCCGAGAG	ATTCTGGCCGTGGTTCTGGGTTTT
mcr-1		GATCGCTGTCGTGCTCTTTG	ACCGCGCCATGATTAATAG	CGATGCTACTGATCACCACG
SHV		TCCATGATGAGCACCTTTAAA	TCCTGCTGGCGATAGTGGAT	TGCCGGTGACGAACAGCTGGAG
TEM		GCATCTTACGGATGGCATGA	GTCCTCCGATCGTTGTCAGAA	CAGTGCTGCCATAACCATGAGTGA
CTX-M1		CCGTCACGCTGTRTTAGGA	AATGCCACMCCCAGYCKKCC	CAGCAAAAACCTGCCGRATT
CTX-M8-M25		ATRACACSTTCCGGCTCGAT	GCTAAYGGCGTGGTGGTATC	TCAACACCGCGATCCCCG
CTX-M2-M74		GCGCAGACCCTGAAAAAYCT	TGYGCSCGCTGRGTTTCC	ACSCTGGGYAAAGCGC
CTX-M9		GCTTTATGCGCAGACGARTG	ATCACCGCGATAAAGCACCT	TCGATACCRMAGATAATACGC
KPC		GGCCGCCGTGCAATAC	GCCGCCAACTCCTTCA	TGATAACGCCGCCGCAATTTGT
NDM		ATATCACCGTTGGGATCGAC	TAGTGCTCAGTGTCGGCATC	AAGGACAGCAAGGCCAAGTCG
VIM		TSTACCCRTCCAATGGTCTC	AGAAGKCCRCTGTGTTTTT	TGTCCGTGATGGYGATGAGTTG
Controls				
16S		TGCAAGTCGAACGAAGCACTTTA	GCAGGTTACCCACGCGTTAC	CGCCACTCAGTCACAAA
human mtDNA		CAATGAATCTGAGGAGGCTAC	CGTGAAGAATAGGAGGTG	ACCTCACACGATTCTTTACCTTCACT
BHV		GAGCAAAGCCCCGCCGAAGGA	TACGAACAGCAGCACGGGCGG	GAACCTGCCACGCGCTGAAAC
BRSV		GCAATGCTGCAGGACTAGGTATAAT	ACACTGTAATTGATGACCCCATCT	ACCAAGACTTGTATGATGCTGCCAAAGCA

*All probes have FAM on the 5' end and MGB on the 3' end

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	K1	K2	M1	M2	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	.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	.05
.05	.8	8	10	36	26	.4262	.25	.1066	.05
.05	.8	8	10	36	26	.4804	.3	.1441	.05
.05	.8	8	10	36	26	.5257	.35	.184	.05
.05	.8	8	10	36	26	.5648	.4	.2259	.05
.05	.8	8	10	36	26	.5994	.45	.2698	.05
.05	.8	8	10	36	26	.6308	.5	.3154	.05
.05	.8	8	10	36	26	.6596	.55	.3628	.05
.05	.8	8	10	36	26	.6867	.6	.412	.05
.05	.8	8	10	36	26	.7126	.65	.4632	.05
.05	.8	8	10	36	26	.7376	.7	.5163	.05
.05	.8	8	10	36	26	.7624	.75	.5718	.05
.05	.8	8	10	36	26	.7875	.8	.63	.05
.05	.8	8	10	36	26	.5398	.5	.2699	.1
.05	.8	8	10	36	26	.5737	.55	.3155	.1
.05	.8	8	10	36	26	.6056	.6	.3634	.1
.05	.8	8	10	36	26	.6362	.65	.4135	.1
.05	.8	8	10	36	26	.6660	.7	.4662	.1
.05	.8	8	10	36	26	.6955	.75	.5216	.1
.05	.8	8	10	36	26	.7254	.8	.5804	.1
.05	.8	8	10	36	26	.7567	.85	.6432	.1
.05	.8	8	10	36	26	.7907	.9	.7117	.1

References Cited

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>
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.
4. 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>
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.
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.
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>

BMJ Open

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

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

Heading:	
Secondary Subject Heading:	Gastroenterology and hepatology, Epidemiology
Keywords:	Gastrointestinal infections < GASTROENTEROLOGY, Public health < INFECTIOUS DISEASES, Epidemiology < INFECTIOUS DISEASES



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

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 Cumbe,⁴ Bacelar Muneme,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 matthew.freeman@emory.edu; 1518 Clifton Road NE, Atlanta, GA 30322

Word count: 5,680

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 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.

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.

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- 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

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

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

1 The predominance of studies in the WASH sector have focused on household- and compound-
2 level interventions because they lend themselves more readily to RCT methodology. Outside of
3 the few aforementioned studies of community-wide infrastructure improvements, evaluations of
4 the health impact of water quality improvements in low-income settings often focus on household
5 water treatment such as boiling, chlorinating or filtering water, with studies predominantly
6 conducted in rural settings.[19] Results of these trials have been mixed, since household-based
7 approaches have various limitations, including low uptake and inconsistent use,[24,25] post-
8 treatment contamination,[26–29] and a poor record of sustained use.[30,31] Household water
9 treatment interventions do not increase water quantity and availability and typical household
10 WASH interventions are likely insufficient to prevent growth faltering in most cases.[32–35]

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

17 **1.2. Objective measures of gut health are needed**

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

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

19 1.3. Overview of study

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

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

11 **2. METHODS & ANALYSIS**

12 **2.1. Study setting and description of the intervention**

13 Our study site is the coastal city of Beira, the second largest city in Mozambique (population
14 ~530,000),[74] which serves as a gateway for both the central interior portion of the country and
15 a trade corridor to neighboring land-locked nations. The center of Beira is bordered by unplanned,
16 informal settlements inhabited by over 300,000 low-income residents.[75] A 2018 survey in Beira
17 by Water & Sanitation for the Urban Poor (WSUP) showed that of 5,643 respondents, only 28%
18 had a household water connection, and among those households without a connection, 83% used
19 their neighbor's tap as their main source of water (unpublished data, courtesy of WSUP).
20 Therefore, improvement of water supply and delivery infrastructure is a priority.

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

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

11 2.2 Study design

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

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

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

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

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 **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
18 shown in **Figure 4**. We worked with FIPAG to determine which neighborhoods in Beira were to
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

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

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

1 We matched intervention sub-neighborhoods to control sub-neighborhoods, using coarsened
2 exact matching,[79,80] with intervention sub-neighborhoods being matched to control sub-
3 neighborhoods within the same tertile of both SES and population density. Four neighborhoods
4 (encompassing nine sub-neighborhoods) were found to be outliers in terms of their sub-
5 neighborhood-level SES or sanitation, and were excluded from the study sampling frame.
6 Ultimately we designated 36 intervention sub-neighborhoods, with an estimated 16,800
7 households, and 26 control sub-neighborhoods, with an estimated 9,500 households.

8 9 **2.5. Participant recruitment, eligibility, and retention**

10 We recruit pregnant women at the last trimester of pregnancy and follow the infant-mother
11 dyads until the child is 12 months old (**Figure 5**). We selected the first 12 months of life because
12 it is a critical development window,[81–83] it is a time when children are most at risk of acute and
13 chronic effects of enteropathogen infection,[84] and it is a short enough period of time to avoid
14 changes in water access that might occur. We recruit mothers at the end of their pregnancy so we
15 can collect data on household risk factors (including drinking water quality) during the gestational
16 period. Active recruitment occurs through identification of pregnant women in the 2020
17 population-based survey, lists of pregnant women visiting local health centers for pre-natal care,
18 and study staff visiting under-enrolled sub-neighborhoods throughout the recruitment period.
19 Based on Ministry of Health data for Sofala Province (where Beira is located), virtually all mothers
20 attend pre-natal clinical visits.[85] Passive strategies include referrals of pregnant women by study
21 participants and community leaders. We aim to have complete data on a total of 548 infant-mother
22 dyads, approximately evenly divided between the intervention and control groups. We will
23 continue to enroll dyads into both arms until we reach a minimum of 274 dyads with complete

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

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

7 8 **2.6. Data collection**

9 A local data collection firm (WE Consult) performs the in-country coordination of participant
10 enrollment, data collection, and sample collection. Enumerators conduct household visits before
11 birth for consent, eligibility, and conditions. At months 3, 6, 9, and 12 we deploy survey
12 instruments to collect data on key indicators through structured observations, reports from
13 respondents, and objective measurements (**Table S1**). We assess a number of variables related to
14 drinking water, including aspects of water quality, water access, water availability, water security,
15 water consumption, and participant satisfaction with water. Enumerators also conduct brief active
16 surveillance calls on a monthly-basis by phone with caregivers to gather supplemental information
17 on prenatal and perinatal environmental exposures and illnesses, on child illness symptoms, and
18 intake of medicines, vitamins, breastfeeding, and introduction of complementary foods (**Figure**
19 **5**). To facilitate communication with the study team, participants receive a 150 MZN phone credit
20 at each visit. Aside from these phone credits, there is no financial incentive provided to participants
21 to partake in the study, per Mozambican guidelines for human subjects research. We ask the
22 caregiver to report diarrhea and blood in the stool (dysentery) of the index child in the previous
23 week at the 3, 6, 9, and 12 month surveys and during active surveillance calls; due to concerns

1 about reporting biases, we also include negative control outcomes.[86] At each post-birth visit we
2 measure child: 1) length, weight, and head circumference, and calculate length-for-age and weight-
3 for-age Z-scores. Prevalence of stunting and underweight are defined as two standard deviations
4 below median of the reference population.[87] All data are collected on electronic tablets using
5 Open Data Kit (ODK) Collect, an open-source program which allows offline data collection on a
6 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, with
10 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.

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

1 gene amplicon to characterize gut microbiome community structure and composition.
2 Bioinformatic analyses will be completed using the QIIME2 software platform.[91]

4 2.7.2. Dried blood spots

5 *Sample collection:* A trained nurse from INS collects up to six dried spots of capillary blood
6 of the index child at 6, 9, and 12 month visits on Tropic Filter Paper Blood Collection Disks
7 (Cellabs, Sydney, Australia), using a 2mm lancet. Samples are stored at -20 °C at INS facilities in
8 Beira and shipped at ambient temperature .[92] We will use the Luminex platform to carry out
9 high throughput, multiplex antibody assays that enable the simultaneous measurement of
10 quantitative antibody responses to dozens of pathogens from a single blood spot.[93] Our first
11 measure will occur at 6 months, to avoid detection of maternal antibodies that wane over the first
12 3-6 months of life.[94]

14 2.7.3. Drinking water

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

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

4 **2.8. Outcomes**

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

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

3 Additional non-primary outcomes include pathogen count, pathogen community similarity
4 (measured using Jaccard similarity index), diarrhea, child growth, and prior enteropathogen
5 infection (measured using serology on dried blood spot samples). We will measure additional
6 water quality exposure measures, as well as measures of exposure to the improved water system,
7 such as fidelity of the intervention (e.g., improvements to water quantity and coverage of
8 household taps) and receipt of the intervention by community members (e.g., reductions in water
9 insecurity, increased water use). These fidelity and uptake measures will be collected at all time
10 points through direct observation and respondent report. Available minimal detectable effect sizes
11 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.

	Time point	Minimum detectable effect (Risk Ratio)	Anticipated control group prevalence ^a
Primary outcomes			
Patent enteric pathogen infection			
Prevalence of any bacterial or protozoan pathogen	12 mo ^b	0.74	70%
Prevalence of any bacterial pathogen	12 mo ^b	0.69	61%
Prevalence of any protozoan pathogen	12 mo ^b	0.50	32%
Any co-infection (bacterial, protozoan, or viral pathogens)	12 mo ^b	0.60	45%
Gut microbiome composition			
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			
<i>(Same variables as gut microbiome composition for 12 mo)</i>			
Prior enteric pathogen exposure	3, 6, 9 mo ^c		
<i>(Same variables as patent enteric pathogen infection)</i>			
Diarrhea 1-week period prevalence (caregiver report)	6, 9, 12 mo	--	--
	Weekly	0.26	14.4%
Anthropometric Measurements			
Length-for-age Z-score	3, 6, 9, 12 mo	--	--
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		

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

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

^cA 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.

^dThis is a conservative estimate as it does not account for weekly active surveillance.

^eDefined 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 sub-neighborhood 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

1
2
3 each given time point. For these analyses, we will control for sub-neighborhood-level SES and
4
5 population density through matching, and will additionally control for household sanitation,
6
7 mother's education-level, child sex, and any other variables that are imbalanced and are
8
9 conceivably potential confounders. For previous enteropathogen exposure evaluated using
10
11 serological measures we hypothesize that those in the intervention group will show delays in
12
13 pathogen acquisition.
14
15

16
17 To assess the impact of the intervention on microbiome outcomes, we will evaluate alpha
18
19 diversity (Chao1 species richness estimator, Pielou's J evenness estimator, and the Shannon
20
21 diversity index [105]) using the same modeling approach as described above for continuous
22
23 outcomes. Linear discriminate effect size (LEfSe) analyses will be used to evaluate specific 16S
24
25 rRNA gene-based Operational Taxonomic Units (OTUs) that differ between individuals in
26
27 intervention versus control groups, and will include effect size corrections [106]. We will examine
28
29 the impact of intervention group, controlling for other covariates, on community similarity using
30
31 Adonis permutation models,[107] based on weighted UniFrac and Bray-Curtis distances, and
32
33 evaluate and visualize differences using PCA and/or NMDS plots. We hypothesize that we will be
34
35 able to observe detectable differences in gut microbiome composition in children living in
36
37 intervention versus control sub-neighborhoods and we will report these differences at the
38
39 individual OTU and bacterial family levels.
40
41
42
43
44
45
46

47 *2.9.2. Direct Household Connection Effect*

48
49 To assess the effect of having a water connection at the household or compound, we will use
50
51 models similar to those described above, but accounting for a household network connection. We
52
53 will also assess the interaction between the household and neighborhood network variables, which
54
55
56
57
58
59
60

1
2
3 will allow us to contrast and estimate indirect, direct, and total effects, as shown in **Figure 3**. We
4
5 hypothesize that participants with both improved water networks in their sub-neighborhoods and
6
7 household water connections will most benefit from the interventions in terms of our primary and
8
9 non-primary health outcomes and exposure outcomes of interest.
10
11
12
13
14

15 2.9.3. Additional analyses

16
17 For select primary outcomes, we will assess if there is effect modification by a third variable,
18
19 such as follow-up round/age, participant sex, and household sanitation access. We will use
20
21 interaction terms to identify potential interactions, and will present stratified results (e.g.,
22
23 separately by sex) if interactions are detected. The intervention status of subneighborhoods was
24
25 set at baseline, but if control subneighborhood(s) receive the intervention after the study has
26
27 started, we will perform sensitivity analyses dropping and/or recategorizing subneighborhood(s)
28
29 that crossed over.
30
31
32

33
34 There are several analyses where we do not use the matched-design and intervention variable
35
36 in our analyses. For example, we will assess associations between various water measures on
37
38 health, without regard to the intervention designation. We will also examine changes in the gut
39
40 microbiomes of children over time. Additional analyses will be described and documented in OSF.
41
42
43
44

45 2.10. Sample size and power calculations

46
47 Our minimal sample size of 548 households—half in intervention and half in matched control
48
49 sub-neighborhoods—was powered for our primary outcome of prevalence of any non-viral
50
51 pathogen. Utilizing data from the MapSan trial for children 10-14 months of age (J. Knee, *pers*
52
53 *comm*) we used a control group prevalence of 70% for any non-viral pathogen, and estimated the
54
55
56
57
58
59
60

1
2
3 ability to detect a relative risk of 0.74, $\alpha=0.05$, and power=80% using a two-sided test for
4 significance.[32] We estimated a sub-neighborhood-level interclass correlation (ICC) of 0.05 (a
5 moderate estimate) among our 62 designated sub-neighborhoods. We will also report on the final
6
7
8
9
10 ICC and other assumptions of this power analysis at the end of the study. Estimates of minimum
11
12 detectable effect sizes based on control prevalence of the outcome of interest (**Table S2**) show we
13
14 may be adequately powered to detect a difference in some individual pathogens if those pathogens
15
16 have high prevalence and/or if they are strongly associated with the water supply improvement
17
18 intervention (e.g., waterborne pathogens). We target planned recruitment at 900 pregnant women
19
20 in the third trimester, to account for incomplete data and loss to follow-up. We used sub-
21
22 neighborhood enrollment targets proportionate to our density estimates to achieve balance across
23
24 intervention and control sub-neighborhoods.
25
26
27
28
29
30

31 **2.11. Blinding**

32
33 All lab personnel and field enumerators are blinded to the intervention status of the samples
34
35 and households. Participants cannot be blinded to their household-level water exposure status or
36
37 cluster-level exposure status, although participants may or may not know about water
38
39 improvements in their particular neighborhood. A primary analyst external to the core data
40
41 management team is blinded to the group assignments until the data cleaning and primary analysis
42
43 are completed. Details of these procedures are included in the Supplementary Material. Unblinding
44
45 will occur only after primary outcome models are developed and compared between two
46
47 independent analysts. Analyses examining the impact of the intervention on non-primary outcomes
48
49 or exposures of interest will not be unblinded until after analyses that examine the impact of the
50
51 intervention on our primary outcomes have been completed. Purely observational analyses that do
52
53
54
55
56
57
58
59
60

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 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.

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/International Monetary Fund (IMF). Global Monitoring 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

- 13 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).
- 14 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
- 15 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).
- 16 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
- 17 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
- 18 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
- 19 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
- 20 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
- 21 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
- 22 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

- 1
2
3 23 Arnold BF, Khush RS, Ramaswamy P, *et al.* Causal inference methods to study
4 nonrandomized, preexisting development interventions. *Proc Natl Acad Sci*
5 2010;**107**:22605–10. doi:10.1073/pnas.1008944107
6
7
8 24 Boisson S, Schmidt WP, Berhanu T, *et al.* Randomized controlled trial in rural Ethiopia to
9 assess a portable water treatment device. *Environ Sci Technol* 2009;**43**:5934–9.
10
11 25 Luoto J, Mahmud M, Albert J, *et al.* Learning to dislike safe water products: results from a
12 randomized controlled trial of the effects of direct and peer experience on willingness to
13 pay. *Environ Sci Technol* 2012;**46**:6244–51. doi:10.1021/es2027967
14
15 26 Levy K, Nelson KL, Hubbard A, *et al.* Following the water: a controlled study of drinking
16 water storage in northern coastal Ecuador. *Environ Health Perspect* 2008;**116**:1533–40.
17 doi:10.1289/ehp.11296
18
19 27 Wright J, Gundry S, Conroy R. Household drinking water in developing countries: a
20 systematic review of microbiological contamination between source and point-of-use. *Trop*
21 *Med Int Health* 2004;**9**:106–17.
22
23 28 Levy K, Anderson L, Robb KA, *et al.* Household effectiveness vs. laboratory efficacy of
24 point-of-use chlorination. *Water Res* 2014;**54**:69–77. doi:10.1016/j.watres.2014.01.037
25
26 29 Stauber CE, Kominek B, Liang KR, *et al.* Evaluation of the impact of the plastic BioSand
27 filter on health and drinking water quality in rural Tamale, Ghana. *Int J Environ Res Public*
28 *Health* 2012;**9**:3806–23. doi:10.3390/ijerph9113806
29
30 30 Brown J, Clasen T. High adherence is necessary to realize health gains from water quality
31 interventions. *PLoS One* 2012;**7**:e36735. doi:10.1371/journal.pone.0036735
32
33 31 Enger KS, Nelson KL, Rose JB, *et al.* The joint effects of efficacy and compliance: a study
34 of household water treatment effectiveness against childhood diarrhea. *Water Res*
35 2013;**47**:1181–90. doi:10.1016/j.watres.2012.11.034
36
37 32 Knee J, Sumner T, Adriano Z, *et al.* Effects of an urban sanitation intervention on
38 childhood enteric infection and diarrhea in Maputo, Mozambique: A controlled before-and-
39 after trial. *eLife* 2021;**10**:e62278. doi:10.7554/eLife.62278
40
41 33 Gough EK, Moulton LH, Mutasa K, *et al.* Effects of improved water, sanitation, and
42 hygiene and improved complementary feeding on environmental enteric dysfunction in
43 children in rural Zimbabwe: A cluster-randomized controlled trial. *PLoS Negl Trop Dis*
44 2020;**14**:e0007963. doi:10.1371/journal.pntd.0007963
45
46 34 Null C, Stewart CP, Pickering AJ, *et al.* Effects of water quality, sanitation, handwashing,
47 and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-
48 randomised controlled trial. *Lancet Glob Health* 2018;**6**:e316–29. doi:10.1016/S2214-
49 109X(18)30005-6
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 35 Luby SP, Rahman M, Arnold BF, *et al.* Effects of water quality, sanitation, handwashing,
4 and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster
5 randomised controlled trial. *Lancet Glob Health* 2018;**6**:e302–15. doi:10.1016/S2214-
6 109X(17)30490-4
7
8
9 36 Post A, Ray I. Hybrid Modes of Urban Water Delivery in Low- and Middle-Income
10 Countries. *Oxf. Res. Encycl. Environ. Sci.* 2020.
11 doi:10.1093/acrefore/9780199389414.013.679
12
13 37 World Health Organization (WHO) and the United Nations Children’s Fund (UNICEF).
14 Progress on household drinking water, sanitation and hygiene 2000-2020: five years into
15 the SDGs. Geneva: : World Health Organization (WHO) and the United Nations Children’s
16 Fund (UNICEF) 2021.
17
18 38 Walker CL, Rudan I, Liu L, *et al.* Global burden of childhood pneumonia and diarrhoea.
19 *Lancet* 2013;**381**:1405–16. doi:10.1016/S0140-6736(13)60222-6
20
21 39 Liu L, Johnson HL, Cousens S, *et al.* Global, regional, and national causes of child
22 mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet*
23 2012;**379**:2151–61. doi:10.1016/S0140-6736(12)60560-1
24
25 40 Arnold BF, Galiani S, Ram PK, *et al.* Optimal recall period for caregiver-reported illness in
26 risk factor and intervention studies: a multicountry study. *Am J Epidemiol* 2013;**177**:361–
27 70. doi:10.1093/aje/kws281
28
29 41 Baqui AH, Black RE, Yunus M, *et al.* Methodological issues in diarrhoeal disease
30 epidemiology. Definition of diarrhoea episodes. *Int J Epidemiol* 1991;**20**:1057–63.
31
32 42 Wright J, Gundry SW, Conroy RM. A review of changing episode definitions and their
33 effects on estimates of diarrhoeal morbidity. *J Health Popul Nutr* 2007;**25**:448–55.
34
35 43 Wright JA, Gundry SW, Conroy R, *et al.* Defining episodes of diarrhoea: results from a
36 three-country study in Sub-Saharan Africa. *J Health Popul Nutr* 2006;**24**:8–16.
37
38 44 Zafar SN, Luby SP, Mendoza C. Recall errors in a weekly survey of diarrhoea in
39 Guatemala: determining the optimal length of recall. *Epidemiol Infect* 2010;**138**:264–9.
40 doi:10.1017/S0950268809990422
41
42 45 Trostle A, Yépez-Montufar J, Corozo-Ángulo B, *et al.* Diarrheal illnesses on the
43 Ecuadorian coast: socio-environmental changes and concepts of health. *Cad Saúde Pública*
44 *Rep Public Health* 2010;**26**:1334–44.
45
46 46 Blum D, Feachem RG. Measuring the impact of water supply and sanitation investments on
47 diarrhoeal diseases: problems of methodology. *Int J Epidemiol* 1983;**12**:357–65.
48
49 47 Lee G, Cama V, Gilman RH, *et al.* Comparison of two types of epidemiological surveys
50 aimed at collecting daily clinical symptoms in community-based longitudinal studies. *Ann*
51 *Epidemiol* 2010;**20**:151–8. doi:10.1016/j.annepidem.2009.10.004
52
53
54
55
56
57
58
59
60

- 1
2
3 48 Schmidt WP, Arnold BF, Boisson S, *et al.* Epidemiological methods in diarrhoea studies--
4 an update. *Int J Epidemiol* 2011;**40**:1678–92.
5
6
7 49 Thapar N, Sanderson IR. Diarrhoea in children: an interface between developing and
8 developed countries. *Lancet* 2004;**363**:641–53. doi:10.1016/S0140-6736(04)15599-2
9
10 50 Holcomb DA, Knee J, Sumner T, *et al.* Human fecal contamination of water, soil, and
11 surfaces in households sharing poor-quality sanitation facilities in Maputo, Mozambique.
12 *Int J Hyg Environ Health* 2020;**226**:113496. doi:10.1016/j.ijheh.2020.113496
13
14 51 Chard AN, Baker KK, Tsai K, *et al.* Associations between soil-transmitted helminthiasis
15 and viral, bacterial, and protozoal enteroinfections: a cross-sectional study in rural Laos.
16 *Parasit Vectors* 2019;**12**:216. doi:10.1186/s13071-019-3471-2
17
18 52 Brown J, Cumming O, Bartram J, *et al.* A controlled, before-and-after trial of an urban
19 sanitation intervention to reduce enteric infections in children: research protocol for the
20 Maputo Sanitation (MapSan) study, Mozambique. *BMJ Open* 2015;**5**:e008215–e008215.
21 doi:10.1136/bmjopen-2015-008215
22
23 53 Liu J, Kabir F, Manneh J, *et al.* Development and assessment of molecular diagnostic tests
24 for 15 enteropathogens causing childhood diarrhoea: a multicentre study. *Lancet Infect Dis*
25 2014;**14**:716–24. doi:10.1016/S1473-3099(14)70808-4
26
27 54 Lunn PG, Northrop-Clewes CA, Downes RM. Intestinal permeability, mucosal injury, and
28 growth faltering in Gambian infants. *Lancet* 1991;**338**:907–10.
29
30 55 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet*
31 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8
32
33 56 Guerrant RL, Oria RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with
34 long-term effects on child development. *Nutr Rev* 2008;**66**:487–505. doi:10.1111/j.1753-
35 4887.2008.00082.x
36
37 57 Campbell DI, Elia M, Lunn PG. Growth faltering in rural Gambian infants is associated
38 with impaired small intestinal barrier function, leading to endotoxemia and systemic
39 inflammation. *J Nutr* 2003;**133**:1332–8.
40
41 58 Goto R, Mascie-Taylor CG, Lunn PG. Impact of intestinal permeability, inflammation
42 status and parasitic infections on infant growth faltering in rural Bangladesh. *Br J Nutr*
43 2009;**101**:1509–16. doi:10.1017/S0007114508083554
44
45 59 Lunn PG. Growth retardation and stunting of children in developing countries. *Br J Nutr*
46 2002;**88**:109–10. doi:10.1079/BJNBJN2002652
47
48 60 Guerrant RL, DeBoer MD, Moore SR, *et al.* The impoverished gut--a triple burden of
49 diarrhoea, stunting and chronic disease. *Nat Rev Gastroenterol Hepatol* 2013;**10**:220–9.
50 doi:10.1038/nrgastro.2012.239
51
52
53
54
55
56
57
58
59
60

- 1
2
3 61 Eppig C, Fincher CL, Thornhill R. Parasite prevalence and the worldwide distribution of
4 cognitive ability. *Proc Biol Sci* 2010;**277**:3801–8. doi:10.1098/rspb.2010.0973
5
6
7 62 Wierzbza TF, Muhib F. Exploring the broader consequences of diarrhoeal diseases on child
8 health. *Lancet Glob Health* 2018;**6**:e230–1. doi:10.1016/S2214-109X(18)30047-0
9
10 63 Yatsunencko T, Rey FE, Manary MJ, *et al.* Human gut microbiome viewed across age and
11 geography. *Nature* 2012;**486**:222–7. doi:10.1038/nature11053
12
13 64 Schnorr SL. The diverse microbiome of the hunter-gatherer. *Nature* 2015;**518**:S14-5.
14 doi:10.1038/518S14a
15
16 65 Schnorr SL, Candela M, Rampelli S, *et al.* Gut microbiome of the Hadza hunter-gatherers.
17 *Nat Commun* 2014;**5**:3654. doi:10.1038/ncomms4654
18
19 66 Clemente JC, Pehrsson EC, Blaser MJ, *et al.* The microbiome of uncontacted Amerindians.
20 *Sci Adv* 2015;**1**:e1500183. doi:10.1126/sciadv.1500183
21
22 67 Martinez I, Stegen JC, Maldonado-Gomez MX, *et al.* The gut microbiota of rural papua
23 new guineans: composition, diversity patterns, and ecological processes. *Cell Rep*
24 2015;**11**:527–38. doi:10.1016/j.celrep.2015.03.049
25
26 68 Obregon-Tito AJ, Tito RY, Metcalf J, *et al.* Subsistence strategies in traditional societies
27 distinguish gut microbiomes. *Nat Commun* 2015;**6**:6505. doi:10.1038/ncomms7505
28
29 69 Gacesa R, Kurilshikov A, Vich Vila A, *et al.* Environmental factors shaping the gut
30 microbiome in a Dutch population. *Nature* 2022;**604**:732–9. doi:10.1038/s41586-022-
31 04567-7
32
33 70 Clasen T, Boisson S. Assessing the Health Impact of Water Quality Interventions in Low-
34 Income Settings: Concerns Associated with Blinded Trials and the Need for Objective
35 Outcomes. *Env Health Perspect* 2016;**124**:886–9. doi:10.1289/ehp.1510532
36
37 71 Goddard FGB, Ban R, Barr DB, *et al.* Measuring Environmental Exposure to Enteric
38 Pathogens in Low-Income Settings: Review and Recommendations of an Interdisciplinary
39 Working Group. *Environ Sci Technol* 2020;**54**:11673–91. doi:10.1021/acs.est.0c02421
40
41 72 Prendergast A, Kelly P. Enteropathies in the developing world: neglected effects on global
42 health. *Am J Trop Med Hyg* 2012;**86**:756–63. doi:10.4269/ajtmh.2012.11-0743
43
44 73 Brown J, Cumming O. Stool-Based Pathogen Detection Offers Advantages as an Outcome
45 Measure for Water, Sanitation, and Hygiene Trials. *Am J Trop Med Hyg* 2020;**102**.
46 doi:10.4269/ajtmh.19-0639
47
48 74 INE. Resultados definitivos Censo 2017. IV Recenseamento Geral da Populacao e
49 Habitacao de Mocambique. Maputo: : INE 2017. Www.Ine.Gov.Mzwww.Ine.Gov.Mz
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 75 Water & Sanitation for the Urban Poor. Supporting the provision of universal water
4 services in Beira and Dondo. London: : Water & Sanitation for the Urban Poor 2022.
5
6
7 76 Development Projects : Water Services & Institutional Support II. World Bank.
8 <https://projects.worldbank.org/en/projects-operations/project-detail/P149377> (accessed 1
9 Oct 2021).
10
11 77 Netherlands M of FA of the K of the. Dutch Development results Mozambique.
12 <https://www.dutchdevelopmentresults.nl/2019/countries/mozambique#water> (accessed 20
13 Apr 2022).
14
15 78 Schreiner M. Simple Poverty Scorecard® Poverty-Assessment Tool Mozambique. ;:127.
16
17 79 Blackwell M, Iacus S, King G, *et al.* Cem: Coarsened Exact Matching in Stata. *Stata J*
18 2009;**9**:524–46. doi:10.1177/1536867X0900900402
19
20 80 Porro G, King G, Iacus S. CEM: Software for Coarsened Exact Matching. *J Stat Softw*
21 2009;**30**. doi:10.18637/jss.v030.i09
22
23 81 Mbuya MN, Humphrey JH. Preventing environmental enteric dysfunction through
24 improved water, sanitation and hygiene: an opportunity for stunting reduction in developing
25 countries. *Matern Child Nutr* 2016;**12 Suppl 1**:106–20. doi:10.1111/mcn.12220
26
27 82 Cumming O, Cairncross S. Can water, sanitation and hygiene help eliminate stunting?
28 Current evidence and policy implications. *Matern Child Nutr* 2016;**12 Suppl 1**:91–105.
29 doi:10.1111/mcn.12258
30
31 83 Oria RB, Murray-Kolb LE, Scharf RJ, *et al.* Early-life enteric infections: relation between
32 chronic systemic inflammation and poor cognition in children. *Nutr Rev* 2016;**74**:374–86.
33 doi:10.1093/nutrit/nuw008
34
35 84 Platts-Mills JA, Babji S, Bodhidatta L, *et al.* Pathogen-specific burdens of community
36 diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob*
37 *Health* 2015;**3**:e564–75. doi:10.1016/S2214-109X(15)00151-5
38
39 85 Direcção Provincial de Saúde. Balanço do Plano Economico e Social. Beira: : Governo de
40 Sofala, Direcção Provincial de Saúde, República de Moçambique 2021.
41
42 86 Arnold BF, Ercumen A. Negative Control Outcomes. *JAMA* 2016;**316**:2597–8.
43 doi:10.1001/jama.2016.17700
44
45 87 UNICEF. Nutrition: definitions of the indicators.
46 http://www.unicef.org/infobycountry/stats_popup2.html (accessed 5 Feb 2015).
47
48 88 Hartung C, Lerer A, Anokwa Y, *et al.* Open data kit: tools to build information services for
49 developing regions. In: *Proceedings of the 4th ACM/IEEE International Conference on*
50 *Information and Communication Technologies and Development - ICTD '10*. London,
51 United Kingdom: : ACM Press 2010. 1–12. doi:10.1145/2369220.2369236
52
53
54
55
56
57
58
59
60

- 1
2
3 89 Bosch F, Palmeirim MS, Ali SM, *et al.* Diagnosis of soil-transmitted helminths using the
4 Kato-Katz technique: What is the influence of stirring, storage time and storage temperature
5 on stool sample egg counts? *PLoS Negl Trop Dis* 2021;**15**:e0009032.
6 doi:10.1371/journal.pntd.0009032
7
8
9 90 Liu J, Gratz J, Amour C, *et al.* A laboratory-developed TaqMan Array Card for
10 simultaneous detection of 19 enteropathogens. *J Clin Microbiol* 2013;**51**:472–80.
11 doi:10.1128/JCM.02658-12
12
13 91 Bolyen E, Rideout JR, Dillon MR, *et al.* Reproducible, interactive, scalable and extensible
14 microbiome data science using QIIME 2. *Nat Biotechnol* 2019;**37**:852–7.
15 doi:10.1038/s41587-019-0209-9
16
17 92 Centers for Disease Control and Prevention Division of Parasitic Diseases and Malaria.
18 Blood Collection and Dried Blood Spot Preparation. Atlanta, GA: : Centers for Disease
19 Control and Prevention 2019.
20
21 93 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for
22 neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800.
23 doi:10.1016/j.ijpara.2012.07.002
24
25 94 Zinkernagel RM. Maternal antibodies, childhood infections, and autoimmune diseases. *N*
26 *Engl J Med* 2001;**345**:1331–5. doi:10.1056/NEJMra012493
27
28 95 Smith CM, Hill VR. Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes
29 from water. *Appl Environ Microbiol* 2009;**75**:5284–9. doi:10.1128/AEM.00456-09
30
31 96 Vos T, Lim SS, Abbafati C, *et al.* Global burden of 369 diseases and injuries in 204
32 countries and territories, 1990–2019: a systematic analysis for the Global Burden of
33 Disease Study 2019. *The Lancet* 2020;**396**:1204–22. doi:10.1016/S0140-6736(20)30925-9
34
35 97 Kotloff KL, Nasrin D, Blackwelder WC, *et al.* The incidence, aetiology, and adverse
36 clinical consequences of less severe diarrhoeal episodes among infants and children
37 residing in low-income and middle-income countries: a 12-month case-control study as a
38 follow-on to the Global Enteric Multicenter Study (GEMS). *Lancet Glob Health*
39 2019;**7**:e568–84. doi:10.1016/S2214-109X(19)30076-2
40
41 98 Strina A, Rodrigues LC, Cairncross S, *et al.* Factors associated with rotavirus diarrhoea in
42 children living in a socially diverse urban centre in Brazil. *Trans R Soc Trop Med Hyg*
43 2012;**106**:445–51. doi:10.1016/j.trstmh.2012.02.001
44
45 99 Chen S-Y, Chiu C-H. Worldwide molecular epidemiology of norovirus infection. *Paediatr*
46 *Int Child Health* 2012;**32**:128–31. doi:10.1179/2046905512Y.0000000031
47
48 100 da Silva Poló T, Peiró JR, Mendes LCN, *et al.* Human norovirus infection in Latin
49 America. *J Clin Virol* 2016;**78**:111–9. doi:10.1016/j.jcv.2016.03.016
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 101 Nasser AM. Prevalence and fate of hepatitis a virus in water. *Crit Rev Environ Sci Technol*
4 1994;**24**:281–323. doi:10.1080/10643389409388470
5
6
7 102 Derrien M, Alvarez A-S, de Vos WM. The Gut Microbiota in the First Decade of Life.
8 *Trends Microbiol* 2019;**27**:997–1010. doi:10.1016/j.tim.2019.08.001
9
10 103 Cumming O, Arnold BF, Ban R, *et al*. The implications of three major new trials for the
11 effect of water, sanitation and hygiene on childhood diarrhea and stunting: a consensus
12 statement. *BMC Med* 2019;**17**:173. doi:10.1186/s12916-019-1410-x
13
14 104 Iacus SM, King G, Porro G. Matching for Causal Inference Without Balance Checking.
15 Rochester, NY: : Social Science Research Network 2008. doi:10.2139/ssrn.1152391
16
17 105 Magurran AE. *Measuring biological diversity*. Malden, Ma.: : Blackwell Pub. 2004.
18
19 106 Segata N, Izard J, Waldron L, *et al*. Metagenomic biomarker discovery and explanation.
20 *Genome Biol* 2011;**12**:R60. doi:10.1186/gb-2011-12-6-r60
21
22 107 Oksanen J, Blanchet FG, Kindt R, *et al*. Package ‘vegan’: Community Ecology Package.
23 2016;:Ordination methods, diversity analysis and other functions for community and
24 vegetation ecologists. <https://cran.r-project.org>, <https://github.com/vegandevs/vegan>
25 (accessed 5 Feb 2017).
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

FIGURE 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.

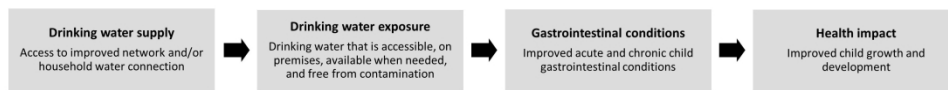


Figure 1. PAASIM theory of change. Mono Image.

307x40mm (300 x 300 DPI)

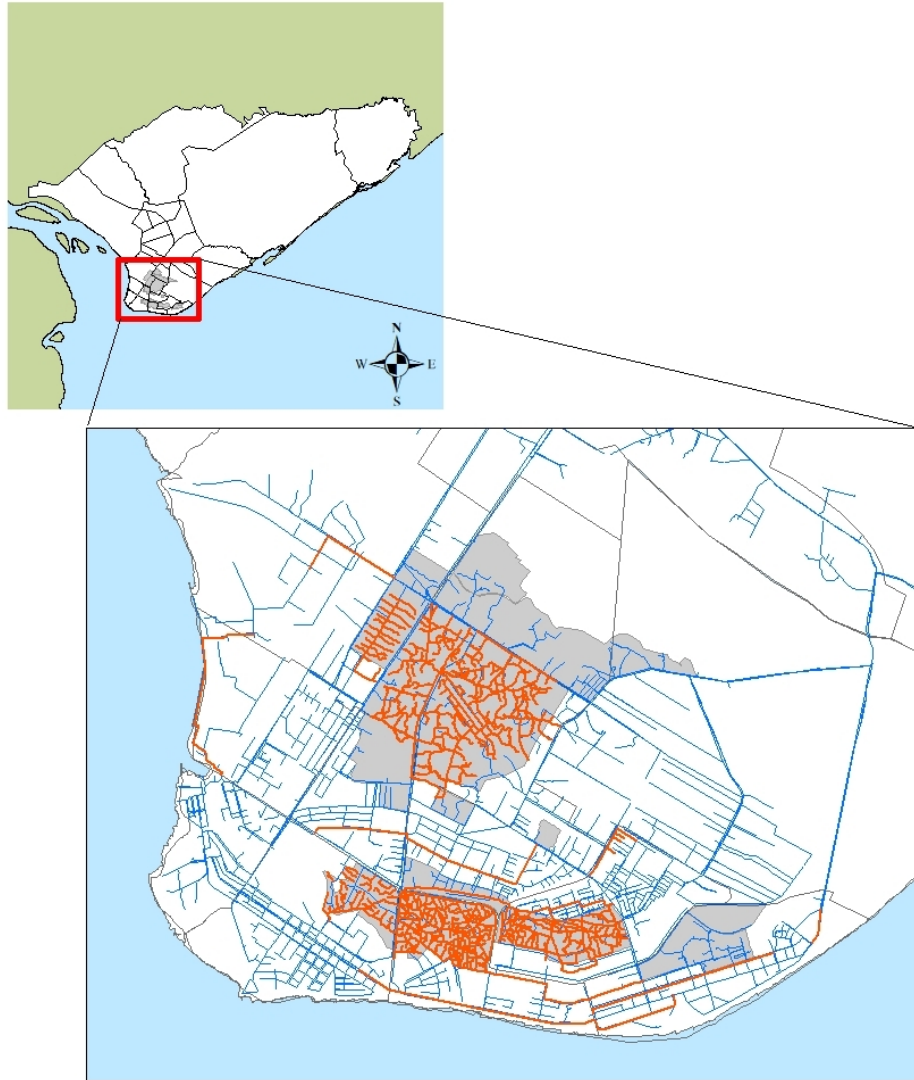


Figure 2. Map of PAASIM study site in Beira, Mozambique. Colour Image. Map of Beira, Mozambique, with enlargement highlighting study site. Red lines indicate the new distribution system water network. Blue lines indicate other parts of the water network. Gray shaded areas indicate neighborhoods enrolled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

215x279mm (96 x 96 DPI)

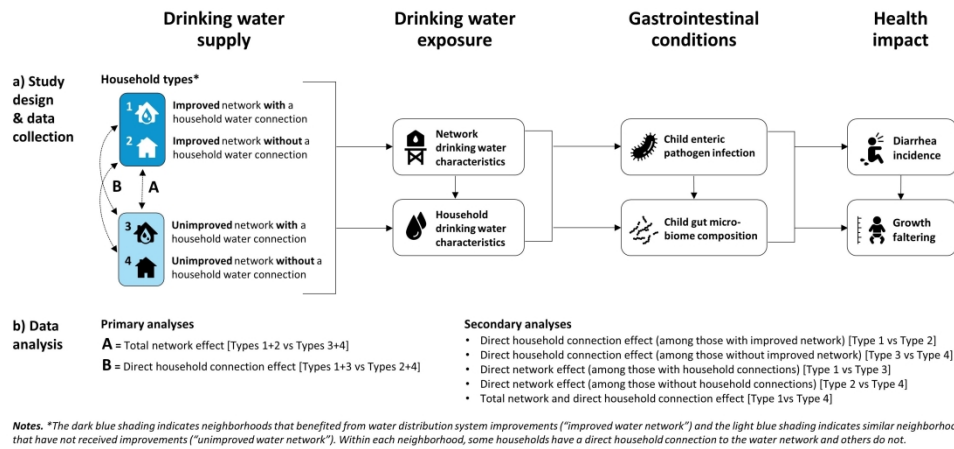


Figure 3. PAASIM study summary diagram. Colour Image. The diagram reflects the summarized a) study design and data collection, and b) data analysis approaches to isolate the effects of both overall water supply infrastructural improvements as well as the presence of a household water connection.

320x151mm (300 x 300 DPI)

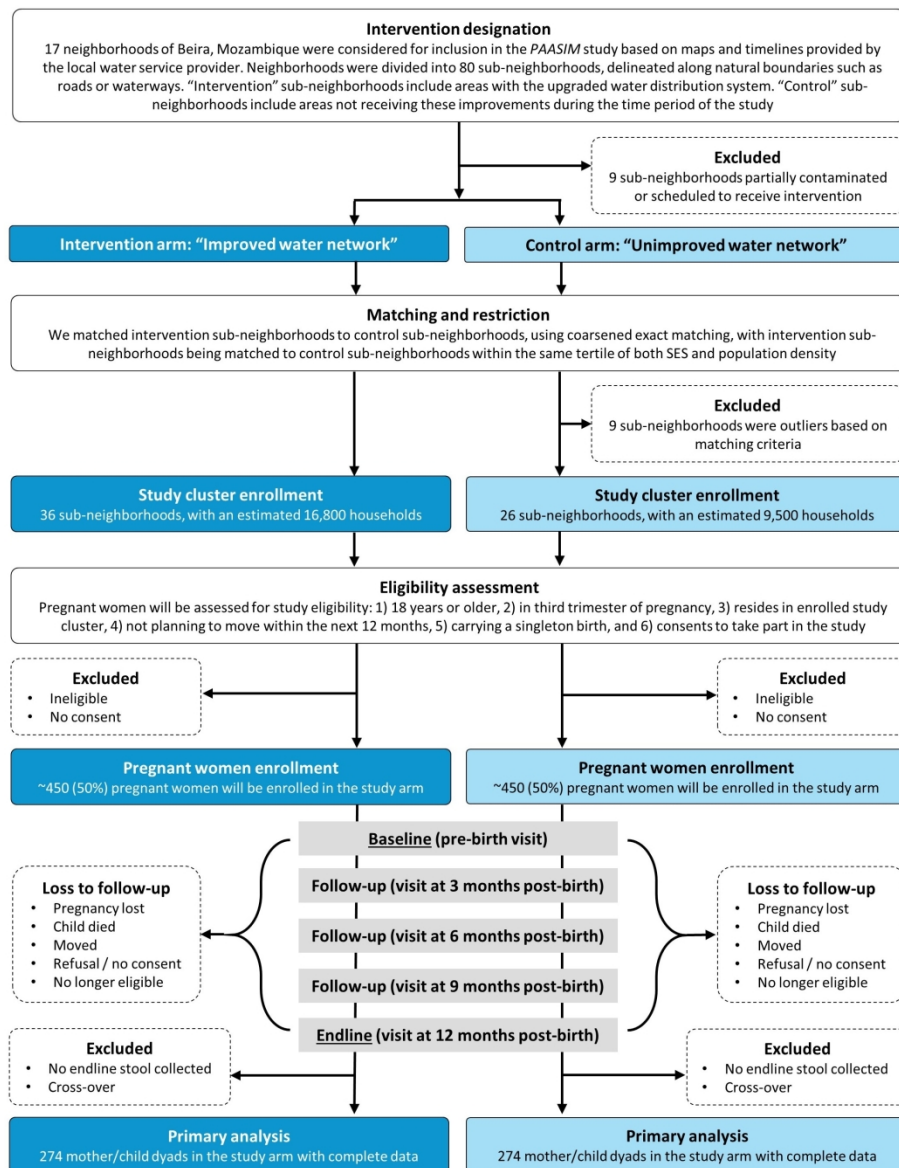


Figure 4. Colour Image. PAASIM study flow diagram.

193x247mm (300 x 300 DPI)

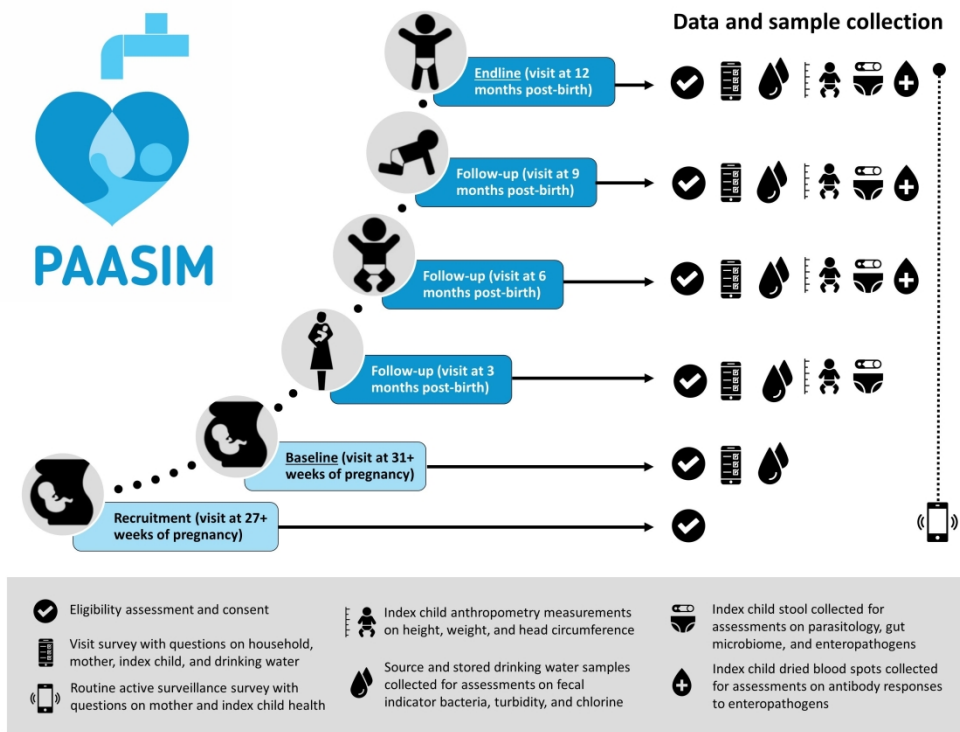


Figure 5. Data and sample collection timeline for outcomes in the PAASIM study. Colour Image. Data and sample collection of infant-mother dyads enrolled into the study will be used address a series of questions about the impact of community-level water system improvements on acute and chronic health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

292x223mm (300 x 300 DPI)

Supplementary Material

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

1
2 as two standard deviations below median of the reference population.⁸ We measure weight using digital
3 baby weighing scales with 5g gradations (ADE, Model# M112600, Hamburg, Germany), and we measure
4 height using a baby length measuring board with 1mm gradations (ADE, Model# MZ10040).⁹
5 Measurements are repeated twice and recorded into the survey. The enumerator is prompted to conduct a
6 third measurement if there are differences of >1 cm in length or head circumference or if weight is off by
7 >0.5kg in the two repeated measurements. Each child's LAZ and WAZ will be calculated using the WHO
8 Child Growth Standards for the reference population^{10,11} and WHO Anthropometric macros.¹²

10 Details of sample collection, processing, and analysis

11 2.5.1. Stool

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

18
19 *Sample processing:* Lab personnel transfer 1g fecal material from the diaper into a DNA/RNA Shield
20 Fecal Collection Tube (Zymo Research, Irvine, CA, USA); we aim to collect three aliquots in separate
21 collection tubes. The DNA/RNA Shield stabilization buffer lyses the cells and renders the sample DNA and
22 RNA stable for 30 days at ambient temperature.¹³ Samples are stored and shipped at -20 °C. Nucleic acids
23 are extracted from fecal samples using the QIAamp 96 Virus QIAcube HT Kit on a QIAcube HT (Qiagen
24 Sciences Inc., Germantown, MD, USA), and stored at -80°C until further processing.

25
26 *Analysis - parasites:* Immediately after the first three aliquots are placed in the collection tubes, two
27 additional aliquots of fecal material are taken from the diaper to prepare a slide for Kato-Katz analysis of
28 parasite ova¹⁴. If sufficient material is available for a second slide analysis is carried out in duplicate.
29 Samples are analyzed for hookworm (*e.g., Necator americanus, Ancylostoma duodenale*) immediately
30 following slide preparation, and for *Ascaris* spp., *Schistosoma mansoni*, *Trichuris trichiura*, *Taenia*,
31 *Enterobius vermicularis*, and *Strongyloides stercoralis* after overnight incubation at room temperature.
32 Eligible participants are referred for deworming medicine at the 12-month visit, after returning results of
33 the parasitological exam to study subjects in collaboration with Instituto Nacional de Saude (INS) staff in
34 Beira.

35
36 *Analysis - enteric pathogens:* Extracted nucleic acids are analyzed using the TaqMan Array card (TAC,
37 ThermoFisher Scientific, Waltham, MA, USA) assay, which allows quantification by real-time PCR via a
38 384-well microfluidic card for simultaneous detection of multiple viral, bacterial, and parasitic enteric
39 pathogen targets as well as antimicrobial resistance genes. Immediately prior to nucleic acid extraction,
40 samples used in downstream TAC assays are seeded with the Inforce 3 Bovine Vaccine (Zoetis, Parsippany-
41 Troy Hills, NJ, USA)¹⁵ containing Bovine Herpesvirus 1 (BHV) and Bovine Respiratory Syncytial Virus
42 (BRSV) as extrinsic controls, to monitor extraction and amplification efficiency. Pathogens will be linked
43 to the diarrheal disease episode based on relative cycle threshold values from the TaqMan results.¹⁶ The
44 specific targets we will test for are shown in **Table S2**. The batch of TaqMan Array Cards will be QA/QC
45 using positive control plasmids. Initially, a standard curve for quantification will be run in duplicate, and
46 the limit of detection and limit of quantification will be determined by running cards at a low concentration
47 (concentration determined based on the standard curve tests) until a 95% positivity rate out of at least ten
48 assays is obtained. Additionally, when processing samples, standard curves will be run in singlicate once
49 per month with the positive control plasmid. A negative control will be run per card.

50
51 *Analysis - Gut microbiome composition:* We will characterize the gut microbial community structure
52 and composition by sequencing of the V4 region of the 16S ribosomal RNA (rRNA) gene amplicon.
53 Immediately prior to nucleic acid extraction, samples used in downstream 16s assays will be seeded with
54 the ZymoBIOMICS Spike-in Control I (High Microbial Load) (Zymo Research) containing *Imtechella*
55
56
57

1 *halotolerans* and *Allobacillus halotolerans*. Bioinformatic analyses will be completed using the QIIME2
2 software platform and other bioinformatics tools¹⁷.
3
4

5 2.5.2. Dried blood spots

6 *Sample collection:* A trained nurse collects up to six dried spots of capillary blood of the index child at
7 6, 9, and 12 month visits on Tropbio Filter Paper Blood Collection Disks (Cellabs, Sydney, Australia), using
8 a 2mm lancet.
9

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

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

24 2.5.3. Drinking water

25 *Sample collection:* We collect 100-mL household drinking water samples from source and stored water
26 at all household visits. We select the drinking water source sample by asking the mother what the source of
27 water is for the household that would be given to the study child to drink, or to mix formula. If the source
28 water is not in the household, the enumerator walks with the mother to the source (e.g., a neighbor's house
29 or public standpipe) to collect a sample and GPS location. We select the stored water sample by asking the
30 mother for any water in the household that is used for drinking purposes; stored water is water stored in a
31 jerry can, bucket, or other container in the household for later consumption.
32

33 To complement the household sampling, we also collect samples from a selection of 45 public
34 standpipes located within the study area and 55 additional public standpipes located elsewhere in the city
35 of Beira. At public standpipes we also measure water pressure by measuring time to fill a fixed volume (1L
36 or 5L, depending on the pressure).
37

38 *Sample processing and analysis:* All samples are placed on cold packs after collection for transport to
39 a lab in Beira. Samples are processed for fecal indicator bacteria within six hours of collection using
40 Colilert-18 reagent and the Quanti-Tray/2000 MPN method (IDEXX Laboratories, Westbrook, ME, USA).
41 Free and total chlorine levels are measured using a DR300 Pocket Colorimeter (Hach Company, Loveland,
42 CO, USA) and DPD powder pillows (Hach Company)²⁴. Additional physiochemical parameters (pH,
43 conductivity, and turbidity) are measured for public standpipe water samples using a Pocket Pro+ Multi 2
44 Tester and TL2300 turbidity meter (Hach Company).
45
46

47 **Details of Blinding**

48 At enrollment, households are assigned a unique identifier independent of intervention status. The core data
49 management team conducts quality assurance using geocoded data to ensure group exposure status aligns
50 with cluster-level designations of the exposure (i.e., intervention vs control areas). All geocoded data and
51 group exposure indicators are removed prior to an external analyst performing the data cleaning and primary
52 analysis. Data are cleaned, including decisions on missing data, outliers, and variable categorizations,
53 before the analyst receives any group exposure. The analyst performs the primary analyses making
54 comparisons between undefined group exposures. Once the analysis models are finalized, these rehearsal
55 results are input into tables to create table shells for the final analyses. The primary analyst then receives a
56
57
58
59
60

1
2 masking key, with true cluster-level designations of the exposure, and reruns the code with the appropriate
3 group exposure, and re-inputs the final results into the table shells.
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Table S1. Household survey data collection in the PAASIM Study

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Household-level							
General characteristics	Socio-economic status	X	*	*	*	X	Respondent reported Observation of households
	Demographics (household ownership, household members, children under 5 years of age, primary wage earner)	X	*	*	*	X	Respondent reported
Nutrition	Food insecurity	X	X	X	X	X	Respondent reported
	Types of food consumed by household members yesterday						Respondent reported
Sanitation	Access to improved sanitation	X	X	X	X	X	Observation of household latrines
	Location of sanitation facility	X	X	X	X	X	Observation of household compounds
	Sharing of sanitation facility	X	X	X	X	X	Respondent reported
	Sanitation facility characteristics (serviceable, drop hole cover, smooth and cleanable floor)	X	X	X	X	X	Observation of household latrines
	Presence of human feces in the household compound	X	X	X	X	X	Observation of household compounds
	Trash disposal	X	X	X	X	X	Respondent reported
Handwashing	Access to handwashing facility with soap and water	X	X	X	X	X	Observation of household compounds
Flooding	Flooding of household compound	X	X	X	X	X	Respondent reported
Animals	Presence of animals in the household or compound (chickens, ducks/turkey, dogs, cats, pigs, sheep, goats, rabbits, donkeys)	X	X	X	X	X	Respondent reported
	Presence of animal feces in the household compound	X	X	X	X	X	Observation of household compounds
	Child(ren) contact with animals	X	X	X	X	X	Respondent reported
Health	Deworming history (pre-school age children, school age children, and mother)		X	X	X	X	Respondent reported
Moving	Moved to another household		X	X	X	X	Respondent reported Observation of household compounds

Notes: *Asked if study household has moved to another location within the study area

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Mother-level							
General characteristics	Demographics (age, religion, language, education, employment, marital status)	X				X	Respondent reported
Mental and physical health	WHO 5 well-being index	X	X	X	X	X	Respondent reported
	Medical diagnosis history (diabetes, high blood pressure, heart disease, urinary tract infection, kidney disease, sexually-transmitted disease, cancer, malaria)	X					Respondent reported
	Vaginal health practices	X					Respondent reported
Pregnancy health	Pregnancy history	X					Respondent reported
	Expected delivery date	X					Respondent reported
	Pre-natal care	X					Respondent reported
	Medical diagnosis during pregnancy (gestational diabetes, high blood pressure, placenta previa, COVID-19, Dengue, Zika, Chikungunya, Malaria)	X					Respondent reported
	Medications, vitamins, or supplements during pregnancy	X					Respondent reported
Birth care	Delivery location		X				Respondent reported
	Cesarean section		X				Respondent reported
	Post-natal care		X				Respondent reported
Breastfeeding	Breastfeeding intentions	X					Respondent reported
	Breastfeeding practices		X	X	X	X	Respondent reported
Sanitation	Primary place of defecation in last week	X	X	X	X	X	Respondent reported
	Exclusive use of sanitation facility in last week	X	X	X	X	X	Respondent reported
Travel	Estimated time spent away from home in the last 12 months					X	Respondent reported

36/bmjopen-2022-067341 on 2 March 2023. Downloaded from <http://bmjopen.bmj.com/> on June 11, 2025 at Agence Bibliographique de l'Enseignement Supérieur (A.B.E.S.) Related to text and data mining, AI training, and similar technologies.

36/bmjopen-2022-067341
Downloaded from <http://bmjopen.bmj.com/> on June 15, 2023 at Agency Bibliographique de l'Enfance et de la Maternité
For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Child-level							
Birth	Birth outcome		X				Respondent reported
	Birthdate		X				Observation of child health card
	Birth weight		X				Respondent reported
	Sex		X				Observation of child health card
Health	Mortality		X	X	X	X	Respondent reported
	Vaccination history (DTP, Rotavirus, Polio, MMR)		X	X	X	X	Observation of child health card
	Medication history		X	X	X	X	Respondent reported
	Deworming history					X	Respondent reported
	Illness (diarrhea, dysentery, nasal congestion, fever, vomiting, physical injury) in the last week		X	X	X	X	Respondent reported
	Medical diagnosis (colic, ear infection, anemia, respiratory illness, malaria, asthma) and treatment history		X	X	X	X	Respondent reported
Anthropometry	Head circumference		X	X	X	X	Objective measurement
	Body length		X	X	X	X	Objective measurement
	Body weight		X	X	X	X	Objective measurement
Nutrition	Types of liquid food consumed in the last week		X	X	X	X	Respondent reported
	Types of solid food consumed in the last week			X	X	X	Respondent reported
Sanitation	Diaper wearing		X	X	X	X	Respondent reported
	Disposal of child stools		X	X	X	X	Respondent reported
Water consumption	Child consumption of drinking water source		X	X	X	X	Respondent reported
	Drinking water treatment for child consumption		X	X	X	X	Respondent reported
Travel	Estimated time spent away from home in the last 12 months					X	Respondent reported

Category	Indicator	Time point					Type of data
		Pre-birth	3mo	6mo	9mo	12mo	
Drinking water-level							
Water access	Municipality water connection and status	X	X	X	X	X	Observation of water meter
	Type of drinking water source	X	X	X	X	X	Respondent reported Observation of water source
	Alternate source for non-drinking water	X	X	X	X	X	Respondent reported
	Alternate source for drinking water outages	X	X	X	X	X	Respondent reported
	Distance to water main	X	X	X	X	X	GPS location
	Location of drinking water source	X	X	X	X	X	Respondent reported
	Distance to drinking water source	X	X	X	X	X	GPS location
	Time to collect water (minutes for round trip and trips per week)	X	X	X	X	X	Respondent reported
Water availability	Water availability (hours per day and days per week)	X	X	X	X	X	Respondent reported
	Household water storage (large and small containers)	X	X	X	X	X	Respondent reported
	Piped water flow rate (liters per minute)	X	X	X	X	X	Objective measurement
	Water availability at handwashing facility	X	X	X	X	X	Observation of handwashing facility
	Water availability at sanitation facility	X	X	X	X	X	Observation of sanitation facility
Drinking water quality	Treatment of stored water	X	X	X	X	X	Respondent reported
	Drinking water free/total chlorine (source and stored water)	X	X	X	X	X	Objective measurement
Water security	Sharing of water connection	X	X	X	X	X	Respondent reported
	Household water insecurity experiences	X	X	X	X	X	Respondent reported
	Sufficient quantity of drinking water	X	X	X	X	X	Respondent reported
User satisfaction with water	Satisfied with water service	X	X	X	X	X	Respondent reported
	Satisfied with water availability	X	X	X	X	X	Respondent reported
	Satisfied with water pressure	X	X	X	X	X	Respondent reported
	Satisfied with water color and appearance	X	X	X	X	X	Respondent reported
	Satisfied with water taste and smell	X	X	X	X	X	Respondent reported
	Satisfied with water affordability	X	X	X	X	X	Respondent reported
Water consumption	Monthly water expenses	X	X	X	X	X	Respondent reported
	Water usage (liters per day)	X	X	X	X	X	Respondent reported Observation of water meter

36/bmjopen-2022-067291 on 3 March 2023 at 12:31 GMT. Downloaded from <http://bmjopen.bmj.com/> on June 15, 2024 at 7:09:56 AM by guest. All rights reserved. No reuse allowed without permission. See <http://creativecommons.org/licenses/by/4.0/> for details.

Table S2: Targets Assayed by the TaqMan Array Card (TAC) for the PAASIM study

Target	Gene of interest	Forward (5' to 3')	Reverse (5' to 3')	Probe (5' to 3')*
Bacterial Pathogens				
Enteroaggregative <i>E. coli</i> (EAEC)	aaiC	ATTGTCCTCAGGCATTTTAC	ACGACACCCCTGATAAACAA	TAGTGCATACTCATCATTTAAG
	aatA	CTGGCGAAAGACTGTATCAT	TTTTGCTTCATAAGCCGATAGA	TGGTTCATCTATTACAGACAGC
Diarrheagenic <i>E. coli</i> (DAEC)	afaB	GTCTCCCTGAATGTACAGCTTTCA	CMCTCTGCCACTCCACCTT	TCAAGCTGTTTGTTCGTC
Shiga toxin-producing <i>E. coli</i> (STEC)	stx1	ACTTCTCGACTGCAAAGACGTATG	ACAAATTATCCCCTGWGCCACTATC	CTCTGCAATAGGTACTCCA
	stx2	CCACATCGGTGTCTGTTATTAACC	GGTCAAACCGCGCTGATAG	TTGCTGTGGATATACGAGG
Enteropathogenic <i>E. coli</i> (EPEC)	eae	CATTGATCAGGATTTTTCTGGTGATA	CTCATGCGGAAATAGCCGTTA	ATACTGGCGAGACTATTTCAA
	bfpA	TGGTGCTTGCGCTTGCT	CGTTGCGCTCATTACTTCTG	CAGTCTGCGTCTGATTCCAA
	LT	TTCCCACCGGATCACCAA	CAACCTTGTGGTGCATGATGA	CTTGGAGAGAAGAACCCT
Enterotoxigenic <i>E. coli</i> (ETEC)	STh	GCTAAACCAGCAGGGTCTTCAAAA	CCCGGTACAAGCAGGATTACAACA	TGGTCCTGAAAGCATGAA
	STp	TGAATCACTTGACTCTTCAAAA	GGCAGGATTACAACAAAGTT	TGAACAACACATTTTACTGCT
<i>Shigella</i> spp./ Enteroinvasive <i>E. coli</i> (EIEC)	ipaH	CCTTTTCCGCGTTCCTTGA	CGGAATCCGGAGGTATTGC	CGCCTTCCGATACCGTCTCTGCA
<i>E. coli</i> O157	rbdE	TTTCACACTTATTGGATGGTCTCAA	CGATGAGTTTATCTGCAAGGTGAT	CTCTCTTTCCTCTGCGGTCT
<i>Campylobacter jejuni</i> / <i>C. coli</i>	cadF	CTGCTAAACCATAGAAATAAAATTTCTCAC	CTTTGAAGGTAATTTAGATATGGATAATCG	CATTTTGACGATTTTTGGCTTGA
<i>Salmonella</i> spp.	ttr	CTCACCAGGAGATTACAACATGG	AGCTCAGACAAAAGTGACCATC	CACCGACGGCGAGACCGACTTT
<i>Vibrio cholerae</i>	hylA	ATCGTCAGTTTGGAGCCAGT	TCGATGCGTTAAACACGAAG	ACCGATGCGATTGCCCAA
<i>Clostridium difficile</i>	tcdB	GGTATTACCTAATGCTCCAATAG	TTTGTGCCATCATTTTCTAAGC	CCTGGTGTCCATCCTGTTTC
Viral Pathogens				
Adenovirus	fiber	AACTTTCTCTCTTAATAGACGCC	AGGGGGCTAGAAAACAAAA	CTGACACGGGCACTCT
Astrovirus	capsid	CAGTTGCTTGCTGCGTTCA	CTTGCTAGCCATCACACTTCT	CACAGAAGAGCAACTCCATCGC
Hepatitis G	5' UTR	CGGCCAAAAGGTGGT GGA TG	CGACGAGCCTGACGTCGGG	AGGTCCCTCTGGCGCTTGTGGCGAG
Norovirus GI	ORF1-ORF2	CGTGCGATGCGATTCCATGA	CTTAGACGCCATCATCATTTAC	TGGACAGGAGATCGC
Norovirus GII	ORF1-ORF2	CAAGAACCTATGTTTAGATGGATGAG	TCGACGCCATCTTCATTACA	TGGGAGGGCGATCGCAATCT
Rotavirus	NSP3	ACCATCTWCACRTRACCCTCTATGAG	GGTCACATAACGCCCTATAGC	AGTTAAAAGCTAACACTGTCAAA
Sapovirus	RdRp	GAYCASGCTCTCGCYACCTAC	CCCTCCATYTCAAACACTA; TTGGCCCTCGCCACCTAC	CCRCCTATRACCA
SARS-CoV-2	N1	GACCCCAAATCAGCGAAAT	TCTGGTTACTGCCAGTTGAATCTG	ACCCCGCATTACGTTTGGTGGACC

36/bmjopen-2022-067341 on 2 March 2023. Downloaded from <http://bmjopen.bmj.com/> on June 11, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES).
 Enregistrement Supérieur (ABES).
 All rights reserved. No reuse allowed without permission. All training, and similar technologies.

Target	Gene of interest	Forward (5' to 3')	Reverse (5' to 3')	Probe (5' to 3')*
Protozoan pathogens				
<i>Cryptosporidium spp.</i>	18S rRNA	GGGTTGTATTTATTAGATAAAGAACCA	AGGCCAATACCCTACCGTCT	TGACATATCATTCAAGTTTCTGAC
<i>Giardia duodenalis</i>	18S rRNA	GACGGCTCAGGACAACGGTT	TTGCCAGCGGTGTCCG	CCCGCGGGCTCCCTGCTAG
<i>Entamoeba histolytica</i>	18S rRNA	ATTGTCGTGGCATCCTAACTCA	GCGGACGGCTCATTATAACA	TCATTGAATGAATTGGCCATT
<i>Cyclospora cayetanensis</i>	18S rRNA	AAAAGCTCGTAGTTGGATTTCTG	AACACCAACGCACGCAGC	AAGGCCGGATGACCACGA
Helminthic pathogens				
<i>Ascaris lumbricoides</i>	ITS1	GCCACATAGTAAATTGCACACAAAT	GCCTTTCTAACAAGCCCAACAT	TTGGCGGACAATTGCATGCGAT
<i>Trichuris trichiura</i>	18S rRNA	TTGAAACGACTTGCTCATCAACTT	CTGATTCTCCGTTAACCGTTGTC	CGATGGTACGCTACGTGCTTACCATGG
<i>Ancylostoma duodenale</i>	ITS2	GAATGACAGCAAACCTCGTTGTTG	ATACTAGCCACTGCCGAAACGT	ATCGTTTACCGACTTTAG
<i>Necator americanus</i>	ITS2	CTGTTTGTGCAACGGTACTTGC	ATAACAGCGTGACATGTTGC	CTGTACTACGCATTGTATAC
Antimicrobial resistance genes				
intl1		GATCGGTGCAATGCGTGT	GCCTTGATGTTACCCGAGAG	ATTCCTGGCCGTGGTTCTGGGTTTT
mcr-1		GATCGCTGTCGTGCTCTTTG	ACCGCGCCATGATTAATAG	CGATGCTACTGATCACCACG
SHV		TCCATGATGAGCACCTTTAAA	TCCTGCTGGCGATAGTGGAT	TGCCGGTGACGAACAGCTGGAG
TEM		GCATCTTACGGATGGCATGA	GTCCTCCGATCGTTGTCAGAA	CAGTGTGCCATAACCATGAGTGA
CTX-M1		CCGTCACGCTGTRTTAGGA	AATGCCACMCCCAGYCKKCC	CAGCAAAAACCTGCCGRATT
CTX-M8-M25		ATRACACSTTCCGGCTCGAT	GCTAAYGGCGTGGTGGTATC	TCAACACCGCGATCCCCG
CTX-M2-M74		GCGCAGACCCTGAAAAAYCT	TGYGCSCGCTGRGTTTCC	ACSCTGGGYAAAGCGC
CTX-M9		GCTTTATGCGCAGACGARTG	ATCACCGCGATAAAGCACCT	TCGATACCRMAGATAATACGC
KPC		GGCCGCCGTGCAATAC	GCCGCCAACTCCTTCA	TGATAACGCCGCCGCAATTTGT
NDM		ATATCACCGTTGGGATCGAC	TAGTGCTCAGTGTCGGCATC	AAGGACAGCAAGGCCAAGTCG
VIM		TSTACCCRTCCAATGGTCTC	AGAAGKCCRCTGTGTTTTT	TGTCCGTGATGGYGATGAGTTG
Controls				
16S		TGCAAGTCGAACGAAGCACTTTA	GCAGGTTACCCACGCGTTAC	CGCCACTCAGTCACAAA
human mtDNA		CAATGAATCTGAGGAGGCTAC	CGTGAAGAATAGGAGGTG	ACCTCACAGATTCTTTACCTTCACT
BHV		GAGCAAAGCCCCGCCGAAGGA	TACGAACAGCAGCACGGGCGG	GAACCTGCCACGCGCTGAAAC
BRSV		GCAATGCTGCAGGACTAGTGATAAT	ACACTGTAATTGATGACCCCATCT	ACCAAGACTTGTATGATGCTGCCAAAGCA

*All probes have FAM on the 5' end and MGB on the 3' end

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	K1	K2	M1	M2	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	.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	.05
.05	.8	8	10	36	26	.4262	.25	.1066	.05
.05	.8	8	10	36	26	.4804	.3	.1441	.05
.05	.8	8	10	36	26	.5257	.35	.184	.05
.05	.8	8	10	36	26	.5648	.4	.2259	.05
.05	.8	8	10	36	26	.5994	.45	.2698	.05
.05	.8	8	10	36	26	.6308	.5	.3154	.05
.05	.8	8	10	36	26	.6596	.55	.3628	.05
.05	.8	8	10	36	26	.6867	.6	.412	.05
.05	.8	8	10	36	26	.7126	.65	.4632	.05
.05	.8	8	10	36	26	.7376	.7	.5163	.05
.05	.8	8	10	36	26	.7624	.75	.5718	.05
.05	.8	8	10	36	26	.7875	.8	.63	.05
.05	.8	8	10	36	26	.5398	.5	.2699	.1
.05	.8	8	10	36	26	.5737	.55	.3155	.1
.05	.8	8	10	36	26	.6056	.6	.3634	.1
.05	.8	8	10	36	26	.6362	.65	.4135	.1
.05	.8	8	10	36	26	.6660	.7	.4662	.1
.05	.8	8	10	36	26	.6955	.75	.5216	.1
.05	.8	8	10	36	26	.7254	.8	.5804	.1
.05	.8	8	10	36	26	.7567	.85	.6432	.1
.05	.8	8	10	36	26	.7907	.9	.7117	.1

References Cited

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>
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.
4. 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>
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.
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.
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>