Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

David M. Berendes Date

Taking care of the backside: Aspects of the sanitation chain beyond the household toilet and their associations with fecal contamination in the public and private domains and enteric infection risk in children

By

David Berendes
Doctor of Philosophy

Environmental Health Sciences

Christine L. Moe
Advisor

Julie A. Clennon
Committee Member

Amy E. Kirby
Committee Member

Juan S. Leon
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Taking care of the backside: Aspects of the sanitation chain beyond the household toilet and their associations with fecal contamination in the public and private domains and enteric infection risk in children

By

David M. Berendes
M.S.P.H., Emory University, 2011
B.S., Duke University, 2009

Advisor: Christine L. Moe Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Environmental Health Sciences
2016

Abstract

Taking care of the backside: Aspects of the sanitation chain beyond the household toilet and their associations with fecal contamination in the public and private domains and enteric infection risk in children

By David M. Berendes

Enteric infections cause over 1.7 billion cases of diarrhea each year and other long-term morbidity, especially in children. Exposure to fecal contamination, and subsequent enteric infection risk, occurs in both public and private domains, especially in urban slums, where interactions between household-level sanitation, neighborhood-level fecal sludge management (FSM), and fecal contamination are poorly understood. This dissertation examined the interactions of household sanitation and FSM, and their spatial heterogeneity, in poor urban neighborhoods in Accra, Ghana and Vellore, India, and evaluated associations with fecal contamination in both domains and risk of enteric infection.

Household surveys described household and neighborhood structural and behavioral risk factors, which were tested for local spatial clustering and overlaid on drainage maps. Fecal indicator bacteria and selected enteric pathogens were quantified in environmental samples inside and outside the household. Stool samples collected from children in both cross-sectional and longitudinal studies were tested for a panel of enteropathogens. Associations between microbiological outcomes and household or neighborhood risk factors were examined by generalized linear models.

Cross-sectional study results in Accra indicated that sub-neighborhood clustering of household sanitation with good FSM was associated with lower levels of fecal indicator bacteria, but not human enteric viruses, in open drains.

Cross-sectional study results from Vellore suggested that poor FSM was associated with higher prevalence of pediatric enteric infection and higher concentrations of norovirus in drains. Further, associations between household sanitation and within-household fecal contamination varied by neighborhood coverage level and household hygiene practices.

The longitudinal assessment of enteric infection in 0-2 year olds suggested that aspects of neighborhood-level FSM and urban geography—clustering of flooding during a monsoon—affect enteric infection risk independent of exposure behaviors. Further, enteric viruses were major causes of diarrhea with atypical associations with "traditional" water, sanitation, and hygiene-associated risk factors, requiring new strategies for mitigation.

Overall, this dissertation underscores the importance of the public domain and urban geography in fecal contamination exposures and enteric infection risk. The spatial heterogeneity and FSM associated with household sanitation have community-level exposure and health implications, which must be measured and accounted for in future infrastructural and behavioral interventions.

Taking care of the backside: Aspects of the sanitation chain beyond the household toilet and their associations with fecal contamination in the public and private domains and enteric infection risk in children

By

David M. Berendes
M.S.P.H, Emory University, 2011
B.S., Duke University, 2009

Advisor: Christine L. Moe Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Environmental Health Sciences
2016

Table of Contents

Introduction

Enteric infections are a source of a massive, likely underestimated, burden of infection, especially among children. Worldwide, 1.7 billion cases of diarrhea, the most recognized symptom of enteric infection, are estimated each year, and 90% of mortality occurs in children under 5 years of age.[1,2] While diarrhea contributes to child morbidity and mortality in the short term, the longer-term effects of repeat enteric infections—both symptomatic and asymptomatic—may have an even larger impact than severe dehydration alone, including malnutrition, stunting, and poor cognitive and developmental outcomes.[2–7] Because the burden and effects of asymptomatic infection on longer-term outcomes have been rarely quantified until recently, the true burden of enteric infections is larger than that estimated by diarrhea alone.[6,8]

Enteric organisms and environmental transmission

Almost all enteric pathogens have primarily fecal-oral transmission pathways.[3,9] These pathogens can be broadly grouped into bacteria, protozoa, soil-transmitted helminths, and viruses.[9] Though shed in stool, and sometimes vomitus, from infected individuals, these pathogens vary in environmental transmission based on their characteristics, including persistence and median infectious dose (ID50), and environmental conditions. For example, soil-transmitted helminth (STH) transmission may be direct or indirect, as their life cycles are more complex than bacteria, involving ingestion or dermal routes of exposure, and development in both the human host and the environment.[10,11]

While bacteria remain a classic example of fecal-oral transmission through the environment, viruses and protozoa tend to have lower ID50s and longer environmental

persistence than bacteria, yielding more direct (less environmentally-mediated) pathways with a greater variety of potential fecal exposures.[9,12] For example, norovirus is positive-sense, single-stranded RNA virus that is a leading cause of gastroenteritis in high-income countries, and is quickly emerging as a large contributor to enteric infection burden worldwide.[13–16] Noroviruses are human-specific, persistent in the environment for months and possibly years, and cause acute vomiting, stomach cramps, fever, and diarrhea within 1-2 days of exposure, with symptoms lasting for generally 2-3 days.[12,13,17] Their low ID50 (modeled as low as 18 viral particles) and prolonged environmental persistence provide a variety of potential environmental transmission pathways, including contact with fomites and environmental surfaces, person to person, and through food.[12,18,19] Foodborne transmission of norovirus is an especially large contributor to gastroenteritis throughout the world, as norovirus was responsible for an estimated 684 million cases throughout the world in 2010.[20]

Adenoviruses are an example of an enteric virus with different epidemiology yet similar environmental transmission characteristics when compared with norovirus. Adenoviruses are double-stranded DNA viruses that are responsible for both respiratory and gastrointestinal illnesses in children and adults. Adenovirus infection is common in humans and, similar to norovirus, infected individuals may shed virus in high concentrations for months after symptomatic infection is resolved. Though some are animal-specific, adenoviruses that cause gastrointestinal illness and are human-specific (e.g. Adenoviruses subgroup F, 40 and 41) have been suggested as indicator organisms for human sewage contamination of water due to the frequency of detection.[21,22] In

addition to waterborne transmission, adenoviruses may also be transmitted through more direct pathways (e.g. hands or fomites).[21]

In contrast to viruses, bacteria, such as Enteroaggregative *E. coli* (EAEC), have different epidemiology and environmental characteristics. While *E. coli* are commonly present in human and animal feces, only a subset of *E. coli* is diarrheagenic. EAEC is one of six subgroups of *E. coli* that cause gastroenteritis and cause symptoms by adhering and aggregating on the intestinal wall, followed by toxin production and increased mucous secretions.[23,24] Though early studies found that EAEC is commonly associated with diarrhea, especially in young children, in both low- and high-income countries, recent work in a multi-site study of enteric infection in children under 2 showed EAEC was, depending on age, the second or third most frequently detected pathogen in both diarrheal (20-25%) and non-diarrheal (25-30%) stool specimens.[25] Environmental transmission of EAEC is generally thought to be fecal-oral and may be transmitted through contaminated food or water. However, reservoirs and specific transmission pathways have not yet been identified and zoonotic pathways are still being considered as well.[26]

Sanitation-related approaches to containing enteric organisms and fecal contamination

Historically, water, sanitation, and hygiene (WASH) interventions have been designed to interrupt exposure to environmental fecal contamination in various pathways.[27,28] Due to recent WASH sector focus in rural areas—where access is poorest—and the complexity associated with designing and implementing community level systems, recent interventions have targeted fecal exposures at the household level.[27–30] As depicted in the F-diagram (Figure I-1), these exposures include

contaminated food, water, hands, or flies.[27] For example, water treatment at the point of collection or point of use inactivates enteric pathogens before consumption, handwashing reduces the amount of feces on hands prior to contact with the mouth, and improved household sanitation contains excreta away from human contact, reducing environmental fecal contamination.[28,29]

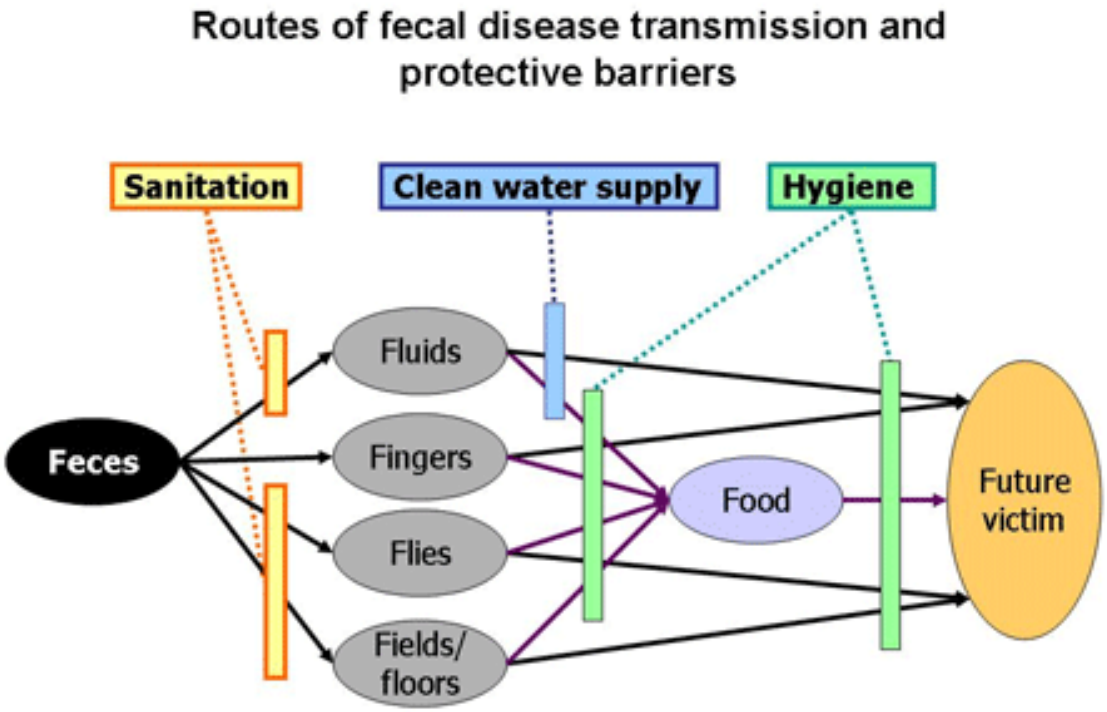


Figure I-1: The F-diagram. This diagram shows environmental pathways of exposure to fecal contamination and where water, sanitation, and hygiene can interrupt this transmission (reproduced from the World Bank website on June 24, 2016: http://water.worldbank.org/sites/water.worldbank.org/files/thumbnail/fdiagram.gif).

Types of household sanitation vary within and between regions in both the user interface and the downstream fecal sludge management (FSM). The user interface is often thought of as the toilet itself, while the FSM associated with a sanitation system may contain the excreta onsite or deliver it to other portions of the sanitation system,

sometimes even the environment. For example, flush toilets may have slabs or commodes, depending on the users' customs or habits, and may connect to a sewer system, septic tank, pit, or even the environment. Dry sanitation options include pit latrines (in this case, without a water seal) with or without slabs, composting toilets, buckets or hanging toilets, or open defecation (the absence of facilities). While these systems do not directly connect to a piped sewerage system, those that contain excreta onsite (pits, septic tanks, etc.) must still be managed. Generally, this management takes the form of removal of fecal sludge via emptying services (trucks), followed by discharge or reuse. This discharge is ideally to a facility where it is treated before entering the environment, though direct discharge (without treatment) is common.[29,31–33]

WASH, and especially sanitation, has interdependencies between the household- and community-levels, meaning that interventions or evaluations cannot be made at a single household alone while disregarding the larger community context. Evaluating a water, sanitation, or hygiene intervention in isolation can lead to erroneous conclusions about the effects of that intervention, as exposure pathways are interlinked and may respond in a threshold-saturation, rather than dose-response, manner.[34–38] Thus, instead of each additional WASH provision adding linear, quantifiable health benefits, there may be a threshold to be reached before health benefits are realized.[34,36–38] For example, provision of isolated, household sanitation may be ineffective if the community sanitation is poor, as an individual household's facility by itself may do little to reduce environmental concentrations of fecal contamination and subsequent fecal exposures.[35,39] Conversely, high community-level coverage of sanitation has been shown to have health benefits for households both with and without sanitation facilities.[36,37] In Guatemala,

Bateman and Smith showed that children with access to a flush toilet, but living in a low coverage sanitation cluster in urban areas, had 67% higher prevalence of stunting compared to those in a high coverage cluster. However, children without access to a toilet, but living in a high coverage cluster, had the same prevalence of stunting as those with access to a toilet.[36] Further, Root observed that, between two Zimbabwean communities—one with 68% household sanitation coverage and one with 0%, children living in the community with 68% sanitation coverage but with no access to sanitation themselves had lower risk of diarrhea when compared with children in the community with no sanitation.[37] These findings underscore the potential for community-level environmental improvements in fecal contamination that may be achieved by sanitation coverage and the importance of viewing sanitation interventions in the context of the community, and not simply the household alone.

Spatial Analysis and WASH-associated Infections

Collection of Global Positioning System (GPS) data to inform understanding of spatial heterogeneity in WASH, and especially sanitation, coverage at different scales is essential to understanding the context in which it operates.[39–41] Because effective sanitation is expected to reduce environmental fecal contamination through containment of excreta, it follows that sanitation must act in both the household and community environments—wherever exposures take place in a localized area.[27] Spatial data allows for better understanding of the levels of containment, fate, and transport of this fecal contamination, and subsequent transmission of infections, providing detail about residents' interactions with the local environment.[42–45] However, there have been limited studies to-date measuring spatial coverage and taking into account interactions between

household- and community-level factors associated with transmission of enteric infection.[35–38,46]

Spatial analysis of enteric infections has mostly focused on clustering of incidence of disease. Recent studies have sought to characterize spatial clusters of diarrhea and enteric infection, examining characteristics within and outside of clusters of high incidence and describing large-scale patterning of infection incidence.[47–51] Recent research in Ethiopia used SaTScan, a program for analyzing cluster of values of points in space, to examine spatial, temporal, and space-time clustering of child diarrhea at the district level.[49,50,52] The study found district-level high and low prevalences of child diarrhea and then fit multi-level models to explain these effects by examining WASH- and nutrition-associated behaviors.[49] Further, they used a similar approach to evaluate childhood stunting by examining individual and community-level exposures.[53]

Researchers in China recently used spatial autocorrelation and spatial regression methods to examine flood-related incidence of infectious diseases and reported significant associations between areas of river flooding and diarrhea.[54] Generally, spatial regression allows for quantification of spatial correlation within the model. Spatial variation in exposure-outcome relationships is either quantified with a lag term (spatial lag regression, utilizing a dependent variable) or controlled for via statistical approaches (spatial error regression).[55]

Few studies have examined clustering of sanitation infrastructure as an exposure. In rural Ecuador, Fuller et al. most recently showed associations between localized spatial coverage of sanitation—using a 500m radius around each household—and childhood stunting.[56] In this study, households within the radius were included in a sanitation

coverage level for "community sanitation," while those outside were not. Similarly, recent work on typhoid from low-income areas of Kolkata, India, has emphasized the role of neighbors' health practices on an individual household's infection risk, with risk of developing typhoid associated with the sum of exposures of households within 100m.[57]

Measurement of health outcomes

Much of the current understanding of the effects of household- and community-level sanitation on enteric infection has been limited by poorly measured and imprecise health outcomes. Reported diarrhea, one of the most frequent outcomes measured following sanitation interventions, is subject to recall, response, and observer bias, leading to potential overestimation of the effects of the intervention.[30,58] Specifically, respondents in the intervention group know they are part of the trial and are often less likely to report diarrhea when it has occurred after the intervention than those in the control group. As a short-term outcome, measurement of diarrhea also fails to encompass the spectrum of longer-term health outcomes associated with enteric infections, including stunting and cognitive disabilities.[3,6] Given the links between subclinical infections, environmental enteropathy (EE), and longer-term health outcomes, failure to examine associations between sanitation and asymptomatic infection may be an important omission in the current WASH literature.[4–6]

Current microbiological techniques can now efficiently detect both symptomatic and asymptomatic enteric infection from stool samples. Recent development of the multiplex assays, including the TaqMan Array Card system and the xTAG® system, allow for simultaneous PCR detection of multiple enteropathogens.[59–63] These and similar

single- and multiplex assays have been used in several recent and ongoing large cohort studies of WASH, enteric infection, EE, and other long-term outcomes to better understand the underlying prevalence and incidence of specific enteric infections.[64–67]

These data not only provide more sensitive outcome measures, but also—when combined with clinical, environmental, and survey data—improved accuracy in understanding the type of infection associated with diarrhea and determining pathway- and pathogen-specific risk factors.[66] Because enteric pathogens have varying abilities to cause symptoms in human hosts, previous understanding of environmental determinants of enteric infection from studies of diarrhea may be incomplete, or specific to a certain set of predominantly diarrheagenic organisms.[66] Given the multitude of organisms causing enteric infection and the complexity of their environmental transmission, further research—even to confirm existing transmission pathways—is useful.[9,12] This work complements that of quantitative microbial risk assessments (QMRAs), which have been important in determining pathogen-specific risk from WASH-associated pathways, especially in the urban environment.[68–73]

Measurement of exposure outcomes

In addition to health studies, exposure assessments are also useful to understand the role of sanitation in reducing enteric infections, as effective sanitation is expected to reduce the amount of excreta in the environment.[27] Measurement of fecal contamination in environmental samples is less invasive than collecting stool samples, and provides a more proximal, objective measure of the effectiveness of local household or neighborhood sanitation compared to self-reported diarrhea.[30,58,74] Sample processing can be done onsite in many study areas, as presence and concentrations of fecal indicator

organisms can be measured through inexpensive, sensitive lab tests that yield results within 24 hours.[74,75] Further, pathogens can be measured through both conventional and quantitative PCR (qPCR), an increasingly common technology throughout the world.[66,76]

However, measurement of fecal contamination varies between studies in the choice of environmental sample/organism tested and the sensitivity of the assay. Within the household, sanitation studies have sampled from stored water, children's and caregiver's hands, household surfaces, utensils, and sentinel objects to measure fecal contamination levels.[77–84] Though each of these types of samples may provide a proxy for within-household fecal contamination and exposure, limitations exist with each. For example, samples of hand contamination—the most physically proximal measure of hand-associated exposures—have been shown to exhibit high temporal variation, even with non-sanitation activities.[79] Further, correlation between microbial concentrations in different types of samples at a single household has not been well-documented, limiting comparability between studies.

Outside the household, concentrations of fecal contamination in drains has been well-studied at the neighborhood level, but rarely at sub-neighborhood scales to test variation with local household sanitation coverage.[69,70,85–90] For most studies, both the spatial and temporal scale of study have been large: spatially, on the order of neighborhood or city catchment areas and, for time, on the order of years. Studies have shown that residential catchments have lower fecal contamination loads in drains compared with industrial areas, but are more subject to spikes in contamination after rainfall from "first flush" runoff from surfaces.[87] Further, concentrations of fecal contamination in these areas has been shown to have diurnal patterning over the course of

a day, peaking from 12:00-15:00.[88] In addition to measurement of fecal coliforms and *E. coli*, research has shown enteric viruses persist in urban sewerage, drainage, and affected rivers as well.[89,90] Though few studies have focused on how sanitation presence or coverage affects levels of fecal contamination in urban drains, a recent QMRA in Thailand showed infection risks from *E. coli* and *Salmonella* spp. in urban canals varied with the dumping of human excreta from a nearby, poorly-functioning, treatment facility.[73]

Measurement of fecal contamination in soil has been limited in the literature, with a primary focus on children's play areas, parks, and other public exposure domains. Studies of these types of public places have focused on parasites, including STHs.[91–94] Though parasite eggs were often prevalent in these areas and sanitation was frequently referenced as an infection control strategy, the WASH literature has paid little attention to these studies or soil-based exposures in general. More recently, understanding of potential risks associated with geophagy, especially in areas where domestic animal husbandry is common, has refocused attention on characterizing fecal contamination in soil.[95] Though a recent QMRA in Kampala, Uganda has estimated exposure from soil contamination as a less important pathway when compared with urban drains, other recent studies focused on fecal contamination in soil as it relates to the presence or absence of sanitation nearby.[70,80] Notably, at the household level in rural Tanzania, one of the few studies focusing on fecal contamination in soil observed that it was associated with human activity in the area more-so than with the presence or type of household sanitation.[80]

Beyond the choice of the environmental sample to measure, the choice of organism also varies in measurements of fecal contamination from environmental samples. Fecal indicator bacteria (FIB), including *E. coli* and fecal coliforms, have long been target organisms for drinking water quality, despite little to moderate correlation with health outcomes; however, study of variation in environmental concentrations of FIB and correlations with health outcomes has been limited.[74,96] FIB concentration measurement is easy, relying on detection of β-D glucoronidase and β-D galactosidase for identification of viable colonies, in contrast to previous PCR-based methods.[74,75,97] However, FIB-based conclusions about levels of human-specific fecal contamination are limited, as their environmental concentrations may reflect growth or die-off in the environment in addition to the initial concentration discharged.[87] Specifically, fecal coliforms include organisms that are not specific to feces, while *E. coli* is present in the gastrointestinal tract of all warm-blooded animals.[98] This colonization of the gut means that *E. coli* is shed with all excreta that has passed through the gastrointestinal tract, and thus detection of *E. coli* in environmental media represent previous animal and human fecal contamination.[74] Techniques like membrane filtration detect markers of current viability and not pathogenicity or organism-specific details, limiting inferences about etiology of public health risk.[98]

In addition to measurements of FIB in environmental samples, recent studies have added PCR-based detection of pathogens in order to quantify human-specific fecal contamination.[70,80,82,84,85] The development of PCR, and especially qPCR methods, within the last 25 years has allowed for specific and sensitive detection and quantification of target organisms from environmental samples within and outside the household.[76] qPCR

uses nucleic acid primers to amplify a specific target sequence of DNA, followed by attachment of fluorescent probes to the amplified DNA or intercalation of fluorescent stain to the double stranded DNA and subsequent amplification and quantification.[76] However, in isolation, PCR detection does not provide information about the infectivity of the organism, though combinations of methods are possible to further predict infectivity.[99,100] Also, the concentration of nucleic acid detected may under- or overestimate organism concentrations based on many factors, including the environmental matrix and limit of detection of the assay and the efficiency of recovery of the target organism during sample processing.[76,100]

Though detection of pathogens, rather than FIB, in environmental samples using PCR provides a more specific measure for pathogen exposure, both PCR and enzyme-based FIB identification vary in sensitivity with the environmental media measured.[101] PCR detection limits for pathogens may vary by orders of magnitude with the conditions of the assay, including the sample matrix and presence of inhibitors, especially in environmental samples.[101] Soil, fecal sludge, and stool samples all have the potential for complete or partial inhibition due to the presence of various substances that inhibit the enzymes used in the assay.[101] These issues with environmental media, when combined with the low infectious doses and environmental persistence of some fecal pathogens like viruses, limit the resolution of and ability to draw conclusions about exposure levels from detection of environmental concentrations alone.[12,17,18,76] However, similar issues of competitive growth of other organisms, like *Aeromonas* spp., and inconsistency in detection with varying pH, alkalinity, hardness, iron, and conductivity of water have been demonstrated in FIB detection.[75] Nonetheless, sampling and testing of fecal indicator

bacteria and pathogens in environmental media is being examined more widely in the literature and is a necessary component of understanding risk from fecal material in private and public environments.[79,80]

Sanitation and Risk of Exposure to Fecal Contamination in the Urban Environment

Despite 82% coverage of improved sanitation in urban areas worldwide, urban sanitation is a growing issue in low-income countries.[29] From 1990 to 2015, improved urban sanitation coverage in the poorest countries only rose from 37% to 47%, and only from 39% to 40% in Sub-Saharan Africa.[29] This poor coverage, especially in low-income urban neighborhoods and informal settlements, must be addressed as the world's population distribution shifts to being predominantly urban.[102,103] By 2009, over 50% of the world's population was urban, and by 2050, the urban population is estimated to almost double from 3.3 billion to 6.3 billion. This influx adds to the 600 million urban residents already without sanitation.[103] The combination of high population density, poor sanitation, and high frequency of person-to-person contact provides an environment for elevated transmission and risk of infectious diseases.[103–107]

In many low-income, urban areas, including informal settlements, there is insufficient space to build individual improved household toilets, forcing household members to rely on public or shared sanitation.[108–111] Though they may be the only feasible option in many urban areas, the microbiological safety of these toilets has been a point of contention within international organizations.[112] Public toilets have been shown to vary greatly in their cleanliness, ability to safely contain excreta, and availability of handwashing facilities, among other issues.[113] Because public toilets may be shared by a

large, undefined number of households, efforts have been made to classify them separately from shared household sanitation, where a defined set of households share the facility.[109] Despite a proposal to classify sanitation facilities shared by at most 5 households or 30 people (whichever is less) as improved sanitation facilities, the Joint Monitoring Program (JMP) of the World Health Organization (WHO) and United Nations Children's Fund (UNICEF) classifies only individual household toilets as improved sanitation.[29,114] Improved sanitation, as defined by the JMP, constitutes facilities that safely separate and contain feces from the user.[29] Recent studies, including some using Demographic and Health Surveys (DHS) data, have indicated that use of shared household sanitation facilities is associated with elevated prevalence and odds of diarrheal disease among children, compared with those using individual household toilets.[112,115] However, identification of the mechanism by which shared sanitation influences health risk has been difficult. Studies of fecal contamination on children's hands, within the latrine itself, and in stored water have shown no consistent differences between households using shared facilities and those using individual household toilets.[78,116]

At the household-level, management of fecal sludge after pit latrines or septic tanks fill is a problem. The few households with onsite sanitation in urban areas often do not have room to build new pits after they fill, forcing users to either connect their facilities to sewerage or open drains, abandon their pits in an unsafe manner, or rely on emptying services, which can be unreliable and potentially unsafe.[31,32,108] Connection to closed sewerage has been shown to reduce diarrhea incidence; however, this need for FSM is frequently overlooked in non-sewered settings, requiring labor-intensive and

often expensive emptying of the pit or household tank under the toilet and subsequent trucking, digestion, and treatment of the fecal sludge.[31,32,41,117] Safe FSM is challenging in crowded, urban environments because emptying services cannot reach latrines to directly exhaust them and/or must commute long distances to safely dispose of the sludge. These challenges can result in over-accumulation of fecal sludge in pits, with the pit becoming unusable, in addition to the potential for urban environmental contamination with excreta.[32,41]

At the community- or city-level, FSM conditions are even worse. Recent estimates from the Water and Sanitation Program (WSP) of the World Bank suggest that less than half of fecal sludge is "safely-managed" (contained from human contact throughout the entire sanitation chain) in a selection of major cities in low-income countries.[31,32] Key findings mirror those at the household-level, including insufficient household-level sludge containment to prevent further urban contamination, given that few latrines are emptied regularly. Additionally, illegal dumping is frequent among those emptying pits, and transport of the sludge beyond the neighborhood, as well as treatment and reuse of fecal sludge, is limited.[31]

Open drains are a frequent fate for household excreta, whether through direct connection or illegal dumping, and may contaminate the local environment with this untreated fecal sludge.[31,118] Open drains often take the place of sewerage systems, which are designed specifically to contain excreta in a closed (piped) system but are expensive and labor-intensive to install.[117] Consequently, levels of fecal contamination, including pathogens, in open drains approximate those of pure sewage.[73,85–89,119] Further, studies have indicated that up to 35% of households with facilities previously classified as

improved may actually discharge into drains and the environment, with excreta remaining untreated.[118]

The role of open drains in the spread of fecal contamination and transmission of enteric infections has not been extensively studied in the literature, possibly because they are not designed to transport sludge specifically within and between urban neighborhoods. In one of the few studies including open drainage alone as a community-level intervention, Moraes et al. found a reduction in the incidence of diarrheal disease and prevalence of intestinal helminths in children in neighborhoods in Salvador, Brazil with drainage only compared with the control group, which received no intervention. However, in this same study, children in a third group of neighborhoods with both drainage and sewerage interventions had even larger reductions above those with drainage alone.[120,121]

In contrast, QMRAs in Accra, Ghana and Kampala, Uganda examining multiple pathways of fecal exposure in urban settings identified exposure to open drains as the pathway with the highest contribution to disease burden in children, among other sanitation-related pathways.[69,70] Further studies, including another QMRA in Thailand, showed open drains, and downstream canals and rivers receiving fecal sludge, to have high fecal contamination and to be a consistent high risk pathway for various types of infections.[73,89,119] Thus, their net effect on risk of enteric infection is unclear and may be specific to the context, geography, and population distribution of the area.

Combined with their high concentrations of fecal contamination, the prevalence of open drains in urban areas adds to risk of enteric infection from flood water. The poorest urban areas are often the most flood-prone, and risk from contact with floodwater

is difficult to mitigate as it generally requires expensive, community-level changes to the physical environment and infrastructure.[122,123] Even in high-income countries, studies have indicated increased risk from bacterial, viral, and protozoal agents associated with flood water contact, especially for children.[71,72] In low-income settings, QMRA has shown that flooding of open drains is associated with elevated bacterial and viral infections, contributing to 5% of disability-adjusted life years (DALYs) per year.[69,70]

Children's exposure to fecal contamination in urban environments varies and includes both WASH pathways like drinking water, as well as environmental pathways, such as exposure to drain water or flood water, that become more prominent because of density.[68–70] Frequent ground contact and the possibility of geophagy implies that soil-based exposures may also pose a risk of exposure to fecal contamination, important because of the large volume of food traffic in public areas.[95,124] Additionally, exposures through food, including through contaminated produce from urban farms and transmission of parasites through food, has been shown to be an important contributor to enteric infection in children.[125–128] Finally, children in urban environments may be exposed to frequent person-to-person or person-object-person transmission of fecal contamination and enteric infection. Secondary transmission of infection and transfer of fecal contamination via hands or fomites can occur both within the household—from parents, older siblings, contaminated surfaces—and outside the household in places with high potential for cross-contamination, like public toilets.[19,129,130] Thus, it is important to consider exposure to fecal contamination both within and outside the household in order to more appropriately assess competing enteric infection risk pathways in urban communities.

Recent approaches to quantifying health risks from poor urban sanitation

In addition to traditional, regression-based approaches to estimating associations between exposures, like urban household toilets and drains, and enteric infection or fecal contamination outcomes, QMRA has been used to estimate the contributions of competing risk pathways to the burden of enteric infection in children. Using this approach, risks of infection with specific organisms via individual exposure pathways are estimated from primary or secondary data on exposure frequencies and environmental concentrations and combined with dose-response models for organism-specific infection.[131]

Two QMRAs, one in Accra, Ghana and one in Kampala, Uganda, have examined both water and sanitation-related exposures in poor urban environments.[69,70] In Accra, Labite et al. modeled various bacterial, viral, and parasitic infection risks from exposures associated with the water supply and sanitation pathways and concentrations in primary (concentrations of bacterial in drains, sea water, and beach sand) and secondary data.[69] In Kampala, Katukiza et al. estimated risks of viral and bacterial infections associated with water and sanitation-related pathways using more detailed environmental concentration data, including testing of both viruses and bacteria in water sources, drains, and gray water, and exposure assessment.[70,85] Both concluded that risks from open drains were associated with the greatest infection burden in children: across pathogens, exposure to open drains in Accra were estimated to contribute to 64% of infections, while those in Kampala were estimated to contribute to 39% of infections (with another 24% contributed by tertiary drains—smaller drains transporting gray water specifically).[69,70]

Other QMRAs have focused on drinking water-specific and flood-specific exposures in low- and high-income settings.[68,71,72] In Accra, Ghana, Machdar et al. compared drinking water-specific risks of bacterial, viral, and parasitic infections using measurement of *E. coli* and *Ascaris* spp. in different drinking water sources and estimates of exposure from surveys, finding that household stored water was the major pathway of contamination and that pathogen-specific burdens were higher than WHO reference levels.[68] In the Netherlands, both ten Veldhuis et al. and de Man et al. examined urban floodwater for microbe-specific risks.[71,72] Ten Veldhuis et al. conducted a screening-level assessment, measuring concentrations of *E. coli*, *Enterococci*, and *Campylobacter* spp. in floods and concluding that further research was justified based on high estimated disease burdens.[71] De Man et al. measured *Campylobacter* spp. (61% prevalence in flood water), *Giardia* spp. (35%), *Cryptosporidium* spp. (30%), noroviruses (29%), and enteroviruses (35%) and estimated exposures by surveys, finding that combined sewers contributed to 33% and storm sewers contributed to 23% of infection risk in children.[72] These QMRAs concluded that risk from enteric organisms in floodwater in these settings is significant enough to warrant more precise examination in future studies and should not be overlooked as an important contributor to overall disease burden. Overall, QMRAs provide a powerful tool to model exposures that are difficult to measure in the real world and inform new areas of study that have not, historically, been examined.

Another, higher-level approach to quantifying poor sanitation, particularly in urban areas, has been the development of fecal waste flow diagrams ("shit flows diagrams" or SFD)s.[31,32] SFDs are city-level assessments to be used as an advocacy tool for government action around poor FSM. They use secondary data collected about

household-level sanitation facilities and neighborhood-level FSM to estimate the proportion of excreta that is "safely-managed" (contained and managed along the entire sanitation chain through treatment) vs. the proportion that is unsafely discharged into the environment. Initially aimed at 12 sentinel cities of varying sizes and conditions around the world, SFDs are now being completed in multiple cities across the world and being managed centrally through online forums (sfd.susana.org). However, SFDs do not include information on exposure and specific geographic data (and therefore cannot be linked directly to human health) and are based on a large number of assumptions. Thus, they serve as a conservative, "worst-case" scenario of the current sanitation situation and are limited in their utility outside of the advocacy realm.

Microbial source tracking (MST) provides an alternative to conventional indicator bacteria as a measure of fecal contamination. MST uses molecular markers or phage-based methods identify the sources of particular microorganisms, and therefore fecal contamination, fecal contamination in the environment.[132] MST requires sub-species level identification to be precise; however, tracking strains have been identified for multiple hosts and it has been used in some studies in low-income settings to describe human and animal contributions to environmental fecal contamination.[79,80,132,133] Early studies using MST, like one in rural Tanzania, have also highlighted the significant human contribution to soil contamination in household areas.[80] Despite the lab and technical capacity needed, MST remains a promising method of understanding sources of environmental fecal contamination that may help inform strategies to better control this contamination.

Study cities

Study sites for this dissertation consisted of 6 neighborhoods in Accra, Ghana and Vellore, India. While the neighborhoods are described in further detail in the following chapters themselves, this section provides a high-level overview of population and sanitation in each city.

Accra, Ghana: Accra is the capital of Ghana, located on the southern coast of the country. Based on 2010 census data, Accra, Ghana is a city of over 1.6 million people, with the Accra Metropolitan Area (AMA) spanning an area of about 140 square kilometers.[134] Slightly less than half of the population (43%) is under the age of 15. An estimated 42% of households use a public toilet facility and over one-third of households (36%) share sanitation facilities with another household, making Accra one of the major cities in the world served by shared sanitation.[109,113] Though only about 2% of households report not having a sanitation facility, this figure may be underestimated due to the high use of public toilets.[134]

Vellore, India: Vellore is part of the Tamil Nadu state and is located approximately 130km inland of Chennai, the state's capital. Vellore has an approximate population of 500,000 in about 20 square kilometers of urban (metropolitan) area.[135,136] According to the 2011 Indian census, an estimated 71% of households in urban, metropolitan Vellore had a sanitation facility. Most (86%) of the households without a facility onsite report open defecating (25% of the population overall), while the rest (4% of the population overall) use public toilets.[137]

Rationale and Aims of the Dissertation

Low-income urban environments provide a unique combination of fecal exposures for children, both in the private and public domain, requiring both household- and community-level interventions.[104,138] In order to identify successful sanitation solutions to reduce human exposure to fecal contamination, research is needed to understand the relative contributions of these exposures to health outcomes. Evaluation of interventions must take the neighborhood context into account when assessing overall effectiveness.[39,46] Further, there is a need to examine spatial heterogeneity of sanitation facilities and fecal contamination and to use fecal contamination measures that are more proximal to household or neighborhood sanitation to examine how sanitation affects potential fecal exposures in urban areas. This dissertation fills these gaps in understanding by examining associations between the type and spatial distribution of household sanitation within poor urban environments and proximal (fecal contamination) and distal (pediatric enteric infection) outcomes. It uses mixed methods in study sites in Accra, Ghana and Vellore, India to assess the following aims:

Aim 1: To examine the household and neighborhood-level factors affecting soil and drain contamination with fecal pathogens and indicator organisms in the public domain.

Aim 2: To examine the neighborhood-level factors affecting household contamination with fecal pathogens and indicator organisms, controlling for household-level sanitation practices.

Aim 3: To examine the neighborhood-level factors affecting children's risk of enteric infection, controlling for household-level sanitation practices, and compare children's exposure pathways within the urban environment.

Urban Sanitation Coverage and Environmental Fecal Contamination: Links Between the Household and Public Environments

David M. Berendes,[1,2*] Amy E. Kirby,[2] Julie A. Clennon,[2,3] Chantal Agbemebiese,[4] Joseph A. Ampofo,[5] George E. Armah,[4] Kelly K. Baker,[2†] Pengbo Liu,[2†] Heather E. Reese,[2†] Katharine A. Robb,[2] Habib Yakubu,[2] Christine L. Moe[2]

Email: dberen2@emory.edu

[1]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

[2]Center for Global Safe Water, Sanitation, and Hygiene, Emory University, Atlanta, GA, USA

[3]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA

[4]Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Accra, Ghana

[5]Council for Scientific and Industrial Research Water Research Institute, Accra, Ghana

[†]Current addresses: Department of Occupational and Environmental Health, University of Iowa College of Public Health, Iowa City, IA, USA (K. K. Baker), Poxvirus and Rabies Branch, Division of High Consequence Pathogens and Pathology, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention (P. Liu); Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA (H. E. Reese)

Abstract

Fecal exposures within the public domain contribute significantly to enteric disease risk. This study examined associations between sanitation and fecal contamination in urban, public environments. Soil (n = 72) and open drain (n = 90) samples were tested for *E. coli* and enteric viruses in four low-income neighborhoods in Accra, Ghana. Household sanitation was categorized by onsite fecal sludge containment ("contained" vs. "uncontained") and degree of sharing ("minimally-shared" categorized as ≤5 households or ≤30 people, compared to all others), per previous Joint Monitoring Program guidelines or proposals, using data from 793 household surveys. Associations between fecal contamination and household sanitation (local coverage and spatial clusters) within 50 and 100m were examined, controlling for season, neighborhood, and population density. Compared to samples from other locations in study neighborhoods, *E. coli* concentrations in drain samples within 50m of contained household sanitation clusters were significantly lower (by >3 log-units), and concentrations in drains within 50 or 100m of minimally-shared household sanitation clusters were also significantly lower (by approximately 1 log-unit). Enteric virus detection in drains and *E. coli* concentrations in soil were not significantly associated with household sanitation coverage or clustering. These findings suggest clustering of contained household sanitation, even if shared, affects levels of fecal contamination within the immediate public domain.

Introduction

An estimated 68% of the world's population lives with improved sanitation, as classified by the Joint Monitoring Program (JMP) for the Millennium Development Goals (MDGs). Improved sanitation consists of systems that ensure the safe containment of excreta, separating it from human contact and thereby minimizing health risk.[29] For the 32% of the world's population without improved sanitation, this health risk can stem from contact with environmental fecal contamination within the household or in the public domain, especially in urban areas.[27,69,70] Recent work has indicated that systems-level approaches to the entire sanitation chain need to be considered to contain excreta in the public domain, a goal included in the proposed Sustainable Development Goals (SDGs).[29,31,46,103]

Improved sanitation infrastructure and services in urban areas have not kept up with population growth in low-income countries. From 1990 to 2015, urban sanitation coverage only increased from 79% to 82% worldwide, and 37% to 47% in the world's poorest countries.[29] Meanwhile, over half of the world's population now lives in urban areas, and that figure is expected to increase to two-thirds of the world's population by 2050, including large urban growth in low and lower-middle income countries.[139] Growth in low-income urban neighborhoods is expected to parallel this overall urban growth, with the population doubling from 2001 to 2030, adding to the 600 million urban dwellers without access to sanitation.[103,138,140]

Solutions to the problems of urban sanitation are necessarily complex because exposures to fecal contamination occur both inside and outside the household.[39] While numerous studies have linked poor urban sanitation with increased diarrheal disease,

urban sanitation interventions have had mixed results on health impact.[103–107,141,142] This variation in effectiveness may be because interventions have generally targeted the household only.[39,142,143] In rural settings, sanitation interventions have historically required high levels of coverage at the community level to achieve reductions in diarrheal disease, implying reductions in both private and public domain contamination. However, the spatial heterogeneity of sanitation coverage has rarely been measured in either rural or urban settings.[27,36–38] In urban areas, multiple quantitative microbial risk assessments (QMRAs) have identified public domain exposures, including open drains, as high risk for children.[69,70] These public domain exposures may result from poor containment of excreta (poor fecal sludge management, "FSM") in unsewered, onsite household sanitation.[31,32] While sewerage has been the most-studied and effective FSM intervention, there has been little study of the effects of improved onsite sanitation, including improved pit latrines and ventilated improved pit (VIP) latrines with good pit-emptying services, on environmental fecal contamination in the public domain.[117]

There has been mixed evidence about the effectiveness of shared sanitation, classified just below "improved" on the sanitation ladder, at reducing fecal exposures. Shared sanitation constitutes any facility shared by two or more households that would otherwise, by design, be considered improved.[29] When compared to individual improved facilities, shared sanitation facilities have been associated with significantly elevated prevalence of pediatric diarrhea and other adverse health outcomes. However, the causal mechanism through which shared sanitation affects these health outcomes is unclear and frequently confounded by factors like socio-economic conditions.[115,144,145] Both urban and rural studies have shown no consistent differences between shared and unshared

improved sanitation facilities when measuring fecal contamination within toilets, in stored water, or on children's hands.[78,116] The effect of shared sanitation on fecal contamination in public, urban environments, where it may be the only feasible and sustainable sanitation option, is being explored by several ongoing studies.[67,108,146]

Given the interconnectedness of public and private domains in cities, there is a need to understand under what conditions sanitation facilities can reduce levels of fecal contamination in the public, urban environment. This study seeks to examine whether the type and spatial heterogeneity of sanitation facilities are associated with fecal contamination, as measured by *E. coli* concentrations, and enteric virus levels in soil and drain water in the public domains of four low-income, urban neighborhoods. *E. coli*, as a measure of human and animal-specific fecal contamination, represents concentrations of excreta in the environment that are expected to be contained—and therefore reduced—by sanitation.[27,98] In the context of the new SDGs, this examination of the type and density of onsite excreta containment will contribute to understanding the conditions under which sanitation coverage can lead to community-level benefits in dense, urban environments.

Methods

Study Site

This study used data collected from the SaniPath Study in four low-income neighborhoods in Accra, Ghana between September 2011 and March 2013 in collaboration with the Water Research Institute of the Center for Scientific and Industrial Research Institute, Ghana (WRI), The Noguchi Memorial Institute for Medical Research of the University of Ghana (NMIMR), and the Training, Research, and Networking for Development (TREND) Group. The SaniPath Study was conducted to quantify the

relative contributions of various household- and neighborhood-level risks of exposure to fecal contamination through multiple environmental pathways.

Accra has two rainy seasons, March – July and September – October, each year. Soil and drain samples were collected from March – December 2012, and household surveys were conducted from August – September 2012. Public toilet surveys were conducted from March – September 2012. Though all were low-income areas, the four study neighborhoods (Alajo, Bukom, Old Fadama, and Shiabu) were selected for variation in types of settlements, location, flooding, and household and public sanitation coverage.[86,113] Further details on neighborhood selection and characteristics are described elsewhere.[86,113]

Ethics

All study protocols were approved by the Institutional Review Board (IRB) at Emory University and the NMIMR IRB, University of Ghana.

Environmental Sampling and Processing

Samples of soil in public places and water in open public drains were collected in each neighborhood. As directed by community leaders and local field staff, locations for sample collection were purposefully selected where children were observed to have been playing or had contact with drains. Global Positioning System (GPS) coordinates were collected at each location at the time of sampling using a Garmin eTrex Venture HC device (Garmin Ltd., Olathe, KS, USA). Samples were tested for *E. coli*, adenovirus, and genogroup I and II norovirus (GI and GII norovirus). *E. coli* was chosen as a general indicator of fecal contamination, while adenovirus and norovirus were chosen because of

their high infection burden in recent studies of West African children.[16,96,147,148] Samples were tested for *E. coli* by membrane filtration using BBL MI agar (Becton Dickinson, Franklin Lakes, New Jersey, USA) following United States Environmental Protection Agency (USEPA) method 1604.[149] For virus analyses, DNA and RNA extraction utilized the MP Bio FastSoil DNA kit (MP Bio, Santa Ana, CA, USA) and Qiagen viral RNA mini kit (Qiagen Sciences, Germantown, Maryland, USA). Samples were tested for adenovirus and GI and GII norovirus by quantitative PCR using published methods.[150,151] Quantitative PCR utilized the QuantiFast Pathogen PCR kit with Internal Control and the OneStep RT-PCR kit (Qiagen Sciences, Germantown, Maryland, USA). The QuantiFast Pathogen PCR was used as a screening PCR for target viruses and assay inhibition. Positive samples and samples with inhibitors were quantified with the OneStep RT-PCR kit. Further details about sample collection and processing can be found in the supporting information (SI).

Household Surveys

Within study neighborhoods, households were defined as a person or group of people sharing cooking or living areas. Compounds consisted of a group of households sharing the same structure. Surveys were conducted in 200 households per neighborhood, selected by dividing the neighborhood into segments, randomly choosing a starting household within each segment, and conducting systematic sampling, as previously described by Peprah et al.[113] GPS coordinates were collected at each household. The target respondent for the survey was the primary caregiver of the youngest child, generally the female head of household. The number of people living in the household and compound and ownership of animals was recorded in the survey. Enumerators

categorized the type of household sanitation facilities present by observation. Facilities were classified into "contained" (ventilated improved pit (VIP) or Kumasi ventilated improved pit (KVIP) latrine, pour-flush/flush toilets into a septic/sewage system, or traditional pit latrines with slabs) and "uncontained" categories (bucket/pan latrines, other latrines, no facility present) based on JMP structural guidelines.[29,113] There were too few study households with uncontained, onsite facilities present to separate this group from households without onsite sanitation. "Improved" or "unimproved" sanitation categories were not used because most facilities were shared by at least two households.[29] Facilities were classified as "minimally-shared" according to two definitions proposed in the 2013 Update to the JMP guidelines and considered in previous studies: 1) 5 or fewer households sharing the facility, or 2) 30 people or fewer people sharing a facility.[109,144,152] These definitions were not mutually exclusive. Households not categorized into a given definition were classified into the comparison group, which therefore consisted of both households without sanitation and households with sanitation shared by more than 5 households or 30 people (too few households in this category were present to create a separate comparison group for analysis).

Public Toilet Surveys

Surveys and observations at public toilets have been described previously.[113] GPS points of all public toilets in the study area were collected during transect walks with a community leader, though only a subset of public toilets were observed. Public toilets were divided into "contained" and "uncontained" categories, as described for household sanitation facilities.

Analyses

Population density surrounding sampling locations was based on the 2010 Ghana Census data (Ghana Statistical Service, Accra, Ghana and estimated by Weeks et al.[153]). Samples were assigned the population density value of the enumeration area in which they were located (example: Figure 1).

Presence and type of household sanitation were evaluated for most-likely local clustering within neighborhoods using the GPS coordinates of study households and Kulldorff's Bernoulli spatial scan in SaTScan version 9.4.[52] Kulldorff's Bernoulli spatial scan evaluates binary outcomes in point data distributed in space to assess the degree of nonrandom clustering of '0' or '1' values. A spatial cluster of "high" coverage of contained sanitation, for example, would be a cluster of households with a significantly higher proportion of contained sanitation facilities, compared to the total proportion of households with contained sanitation facilities in that neighborhood. A cluster of "low" coverage of contained sanitation would be a cluster of households with a significantly lower proportion of contained sanitation facilities compared to the overall neighborhood proportion (example: Figure 1). An α of 0.05 was used to determine the significance of most-likely clusters.

Sanitation surrounding a sampling location was evaluated in three ways: (1) by the presence/absence of a public toilet within 50 or 100m; (2) by the "local household sanitation coverage" calculated among all study households within 50 or 100m (e.g. the number of study households within 50m of the sample location with contained household sanitation divided by the total number of study households within 50m of the sample location); and (3) by the presence of a spatial cluster of high or low coverage of

household sanitation within 50 or 100m. For (3), a sample was classified as "within a cluster" if one or more households within the given radius were part of a most-likely cluster, as calculated by Kulldorff's Bernoulli spatial scan. For (2) and (3), each type of household sanitation (any household sanitation (any facility present), contained household sanitation, and minimally-shared household sanitation) was evaluated. Radii of 50m and 100m from a sampling location were chosen to represent a realistic scale for environment-household sanitation interactions in low-income urban neighborhoods, based on boundaries from previous urban sanitation assessments.[154–156]

Statistical analyses were conducted in R version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria) using standard packages and the "logistf" package for penalized likelihood estimation.[157] Linear regression was used to assess continuous outcomes (*E. coli* concentrations), while logistic regression was used to assess binary outcomes (presence/absence of enteric viruses). Penalized likelihood estimation, referred to as Firth approximation, was used when separation was observed in logistic regression models.[158] Because the goal of regression modeling was to compare sub-neighborhood effects of sanitation on public domain fecal contamination, regression models for sanitation were adjusted for season, neighborhood, and population density in the sampling location's enumeration area (given the absence of a census of households in study neighborhoods). Other covariates were assessed for confounding using a 10% change in the estimate of the main effect (sanitation) as the cutoff for inclusion, as described in Kleinbaum and Klein.[159] Effect modification of sanitation by population density was included in the model if significant at the 0.10 level. An α of 0.05 was used for all other tests of significance.

Results

Neighborhood demographics and coverage/clustering of household sanitation

Population density, prevalence of animal ownership, and coverage of public toilets and household sanitation were compared across neighborhoods using chi-square tests. Household sanitation was further compared by contained/uncontained status, minimally-shared status, and spatial clustering (example: Figure 1), as defined in the methods. Neighborhoods differed significantly in all demographic and sanitation attributes (Table 1). Public toilets observed in all study neighborhoods, except Old Fadama (where the most public toilets were located), were classified as contained. Because of the small proportion of public toilets observed and the absence of sub-neighborhood variation in contained/uncontained status, subsequent analyses across neighborhoods were restricted to the presence/absence of a public toilet. Animal ownership was low to moderate (14-35%) and varied by neighborhood. Chickens (16.6% of households), cats (6.9%), goats (4.2%), and dogs (4.0%) were the most common animals owned. Household sanitation coverage was highest in Alajo (58.5%) and lowest in Old Fadama (1.5%); however, clusters of high coverage of any household sanitation and contained household sanitation were detected in Bukom and Shiabu. Significant clusters of high coverage of minimally-shared sanitation were detected in Alajo and Shiabu using either definition of minimally-shared sanitation, but detection in Bukom varied by the definition used. Overall, sanitation coverage varied between neighborhoods, with significant spatial heterogeneity at the sub-neighborhood level.

Variation in environmental fecal contamination between urban neighborhoods and by season

To understand differences in environmental fecal contamination between neighborhoods, concentrations of *E. coli* and detection of enteric viruses (adenovirus and GI and GII norovirus) in samples of soil in public areas and water in open drains were compared between study neighborhoods by analysis of variance (ANOVA) and chi-square tests of conditional independence, accounting for season. *E. coli* concentrations in soil were moderately-high, varied by 0.2-1.5 $\log_{10}$CFU/g across neighborhoods and were normally-distributed (data not shown). Adenovirus, GI norovirus, and GII norovirus were detected in 3 (4.4%), 1 (1.4%), and 0 (0%) soil samples, respectively, and thus were excluded from further analyses.

E. coli concentrations in drain samples were high in all neighborhoods, varying by 0.3-0.8 $\log_{10}$CFU/100mL (Table 2a) and were normally-distributed (data not shown). Detection of GII norovirus, but not other enteric viruses, in drains varied significantly by neighborhood (Table 2b). Overall, though fecal contamination in soil and some drain samples varied significantly between neighborhoods, there was no single neighborhood where fecal contamination—across types of samples and organisms tested—was consistently higher than in other neighborhoods.

To understand seasonal variation in environmental fecal contamination, *E. coli* concentrations and enteric virus detection were compared between rainy and dry seasons using linear and logistic regression, respectively, controlling for neighborhood. GII norovirus was significantly less likely to be detected in drain samples in the rainy season, compared to the dry season (odds ratio: 0.19, 95% confidence interval: 0.03-0.78);

however, detection of other enteric viruses in drains and *E. coli* concentrations in soil or drains did not differ significantly by season (SI Table S1).

Variation in environmental fecal contamination by animal ownership and population density

Associations between local animal ownership and environmental fecal contamination were examined by linear regression models for *E. coli* concentrations in soil or drain samples, controlling for season and neighborhood. Similar to local household sanitation prevalence, prevalence of household animal ownership was quantified as the number of study households within 50 or 100m of a sampling location owning an animal divided by the total number of study households within that radius. Soil and drain *E. coli* concentrations were not significantly associated with the prevalence of animal ownership within the given radii (SI Table S1).

Associations between population density and environmental fecal contamination were evaluated by linear and logistic regression, controlling for season and neighborhood. None of the fecal contamination outcomes were significantly associated with population density (SI Table S1, data not shown for enteric viruses).

Variation in E. coli contamination in soil by neighborhood sanitation

Associations between neighborhood sanitation and *E. coli* concentrations in soil were assessed by linear regression, controlling for neighborhood, season, and population density. Sanitation was assessed by presence of a public toilet and by local household sanitation coverage within 50 and 100m of sample locations. Because there was little variation in exposure to sunlight, and because the presence of feces (within 3m) or a

toilet/open defecation area (within 30m) were not meaningful confounders (data not shown), these variables were excluded from the models. *E. coli* concentrations in soil samples were not significantly associated with the presence of a public toilet or coverage of household sanitation, regardless of type of containment or number of households or people sharing the facility, within 50 or 100m (SI Table S2).

To understand whether associations between household sanitation and soil *E. coli* concentrations differed when sanitation was clustered within a neighborhood, spatial clusters of household sanitation coverage were substituted into the previous linear regression models evaluating local household sanitation coverage. Soil *E. coli* concentrations were moderately— though not significantly—higher within 50 or 100m of spatial clusters of high coverage of household sanitation, regardless of type, compared to the rest of the neighborhood (Table 3).

Variation in E. coli contamination in drain water by neighborhood sanitation

Associations between neighborhood sanitation and *E. coli* concentrations in drains were assessed using linear regression, controlling for season, neighborhood, and population density. *E. coli* concentrations in drain water samples did not vary significantly with the presence of a public toilet or the local household sanitation coverage within 50 or 100m of the sample location (SI Table S3). However, *E. coli* concentrations in drain water samples within 50 or 100m of clusters of high coverage of any household sanitation, contained household sanitation, or minimally-shared household sanitation were lower, and in most cases significantly lower, than in the rest of the study area (Table 4). Specifically, *E. coli* concentrations in drain water samples within 50m of clusters of high coverage of contained household sanitation were significantly lower than

in the rest of the study area ($p = 0.008$). Further, those within 50m of clusters of high coverage of minimally-shared household sanitation (when defined by the number of people per toilet) and within 100m of clusters of high coverage of minimally-shared household sanitation, using either definition, were also significantly lower than in the rest of the study area ($p = 0.015$ and $p = 0.005$ for minimally-shared sanitation cluster (using "persons per toilet" definition) within 50m and 100m of drain sample locations, respectively; $p = 0.016$ for minimally-shared sanitation cluster (using "households per toilet" definition) within 100m of drain sample locations). *E. coli* concentrations were significantly higher within 50m ($p = 0.010$ using "households per toilet" definition; $p = 0.024$ using "persons per toilet" definition) or 100m ($p = 0.012$ using "households per toilet" definition) of clusters of low coverage of contained or minimally-shared household sanitation.

Variation in enteric virus detection in drain water by neighborhood sanitation

Associations between neighborhood sanitation and adenovirus or norovirus detection in drains were assessed using logistic regression, controlling for season, neighborhood, and population density. Generally, viral detection in drain water samples did not vary significantly with the presence of public toilets or local household sanitation coverage within 50 or 100m (SI Table S4). Only adenovirus was significantly less likely to be detected in drain samples with increasing local coverage of any household sanitation. Viral detection within 50 or 100m of clusters of high coverage of household sanitation was not significantly different from the rest of the neighborhood, regardless of the type of sanitation clustered (SI Table S5). Further, no consistent trends in viral detection in high or low clusters of sanitation coverage were observed.

Discussion

This study examined whether levels of fecal contamination in soils and drains in the urban, public environment were associated with local sanitation coverage and other neighborhood characteristics. While *E. coli* concentrations in drains did not vary significantly with the local coverage of household sanitation, *E. coli* concentrations in samples of drain water collected within 50 or 100m of spatial clusters of high coverage of contained or minimally-shared household sanitation were significantly lower than concentrations in sample from drains that were not in clusters. There were no significant differences between *E. coli* concentrations in drains within and outside of clusters of high coverage of any household sanitation. *E. coli* concentrations in soil did not vary significantly with local coverage or clustering of household sanitation, regardless of type. There was no association between the detection of enteric viruses in drains and local coverage or clusters of household sanitation.

This study is one of the first to examine associations between sanitation coverage and fecal contamination in the public environment, considering the effects of the type of sanitation and spatial heterogeneity in coverage. Other studies that have examined associations between household sanitation, including shared sanitation, and fecal contamination have focused on the user's immediate environment—within the toilet or household—and observed few differences between shared and individual household toilets.[78,116] Studies of fecal contamination in the public environment have examined drains at larger scales, including entire urban areas, and have universally reported high levels of fecal contamination, and especially *E. coli*, in drains. [70,85,87,88,90,160–163] Generally, however, *E. coli* concentrations in these studies have not been as high as

observed in this study, including previous assessments of urban irrigation water from drains in Accra.[88,90,160,162,163] Though human-specific fecal contamination in urban drains has been observed, previous analyses have not examined variation in concentration with the type and spatial heterogeneity of sanitation facilities within the catchment area.[90,162]

Associations between lower *E. coli* concentrations in drains and clustering of high coverage of contained household sanitation may reflect a combination of functional containment of human excreta at the household and sufficient local levels of sanitation coverage to yield community-level benefits.[36–38] Although *E. coli* indicates fecal contamination from both humans and animals, *E. coli* concentrations in drains did not vary with household animal ownership, suggesting that humans may be the primary contributor of excreta in these neighborhoods.[164] Containment of human excreta is the primary role of sanitation, thus *E. coli* concentrations from human sources are expected to vary with sanitation coverage.[27] However, this relationship may be moderated by the type of household sanitation. *E. coli* concentrations in drains did not vary with the presence of clusters of high coverage of any household sanitation. Only household sanitation that contained excreta onsite (i.e. good FSM) was associated with significantly lower *E. coli* concentrations in nearby drains. If functioning properly, contained household sanitation facilities should retain excreta onsite, keeping it away from locations where human contact is possible, including contact with drains.[29,32] Household sanitation facilities that were classified as "uncontained" in this study either failed to contain excreta along the entire sanitation chain or were absent entirely, increasing the potential for environmental contamination.[29,32] Generally, open drains are a common location for uncontained household sanitation facilities in urban areas of low-income

countries to discharge excreta, most of which remains untreated, presenting a high risk fecal exposure pathway.[32,69,70,86,118]

Neighborhood-level coverage or clustering of contained sanitation may also affect *E. coli* concentrations in drains, though in a threshold—rather than dose-response—manner. In this study, evaluation of local household sanitation coverage surrounding a drain sampling site showed no significant association with *E. coli* concentrations, as would have been expected if there was a dose-response relationship. In contrast, clustering of household sanitation, representing localized areas with significantly higher coverage than the rest of the neighborhood, was associated with lower *E. coli* concentrations in drains. Though the effects of high localized sanitation coverage on *E. coli* or other fecal contamination have not been examined previously, studies have observed their effects on diarrheal outcomes. Reductions in diarrheal disease have been observed in households both with and without sanitation facilities in communities attaining high overall sanitation coverage.[34,36–38] Coverage levels within the two "high" clusters of contained household sanitation in this study were lower (44 and 68%) than those observed in previous studies that reported health impacts (generally greater than 75-80%), consistent with the idea of environmental fecal contamination levels an intermediate outcome for diarrhea.[27] These findings suggest that a threshold, rather than a dose-response, model may exist for contained household sanitation specifically.

Associations between clusters of high coverage of household sanitation and lower *E. coli* concentrations in drain water continued to be observed despite the fact that most household toilets (57-63%) in these neighborhoods were shared by multiple households. Clusters of minimally-shared sanitation had significantly lower *E. coli* concentrations in

drain water when compared with the rest of the neighborhood. Unfortunately, few study households near drain sampling locations either 1) shared facilities with more than 30 people or 5 households, or 2) had individual improved household facilities, limiting our ability to assess comparisons with these specific groups. Shared sanitation is common in Accra, and studies of shared sanitation, when compared to individual household toilets, have shown elevated prevalence and odds of diarrhea, resulting in its exclusion from the improved sanitation category.[29,109,112,113,115] While research into fecal contamination associated with shared sanitation has generally focused on within-toilet maintenance, our results suggest that containment of excreta did not vary with the number of users; however, relationships between shared sanitation and FSM effectiveness should be further explored.

While higher coverage of contained household sanitation was expected to decrease *E. coli* concentrations across the entire public domain, local coverage and clustering of household sanitation were not associated with *E. coli* concentrations in soil, suggesting soil contamination may be influenced by other factors. Previous sanitation research in rural households, rather than urban areas, has suggested that fecal contamination in soil varies with human foot traffic in the area, regardless of the type or coverage of local sanitation.[80] This evidence suggests that fecal contamination in soil in our study may have been more influenced by the high foot traffic and mixing of people than the sanitation in surrounding households.

Though *E. coli* concentrations in drain water varied with the type and coverage of local household sanitation, adenovirus and norovirus detection did not. While this finding may indicate differences between *E. coli* concentrations and enteric virus occurrence in

sewage, it may also reflect limitations in laboratory techniques and sampling. Enteric viruses, especially human adenovirus and GII norovirus, have been frequently detected in urban drain water in both high- and low-income countries, including in drain water used for irrigation in Accra.[17,21,22,89,90,165,166] Further, detection of adenovirus has been used to indicate and track human sewage contamination in rural and aquatic environments, areas subject to less regular human activity than open urban drains.[22,166] While the longer persistence of enteric viruses in the environment, when compared with that of *E. coli*, make them useful fecal indicators, it may also make detection of smaller scale variation with sanitation coverage in urban areas more difficult, especially when limited to presence/absence data.[17,21,89] Use of presence/absence data with elevated lower limits of detection (LODs) in environmental media may have introduced false negative results into the data, limiting the precision of our analyses.[76,101] Though the viral results were inconclusive when compared with the *E. coli* results, variation in enteric virus concentrations in the environment should be evaluated further in the context of sanitation studies. Measurement of enteric viruses may provide a more stable, long term, and human-specific indicator of fecal contamination than fecal indicator bacteria.[21]

Detection of fecal contamination in environmental samples had limitations beyond the sensitivity of PCR data following environmental sample processing. Soil moisture content, an important covariate for soil contamination, was not measured; however, exposure to sunlight at the time of collection (yes or no) was used as a proxy.[80] Though providing a more specific outcome than other measures of fecal contamination, virus detection by PCR does not indicate infectivity, restricting inference about public health risk. Further, *E. coli* may be detected year-round in soils and waters of tropical

climates, thus we cannot definitively identify humans as the sources of *E. coli* in soil or drain water.[167] Small sample size was also a limitation; however, despite purposive sampling, selection of sampling locations was independent of local sanitation coverage or spatial distribution.

Measurement of household and neighborhood sanitation had strengths and limitations as well. Observation of household sanitation limited response bias; however, household sanitation was only assessed for a sample, and not a census, of households in study neighborhoods. Bias in evaluation of spatial clustering was minimized, though, by the use of systematic sampling and the choice of spatial scan statistic. Systematic sampling estimated the underlying household distribution in space, while the spatial scan statistic is robust to uneven population distributions.[168] Further, use of spatially-explicit census data improved modeling of sub-neighborhood spatial heterogeneity.

Though we observed significantly lower *E. coli* concentrations in spatial clusters of high coverage of contained household sanitation, concentrations in all sampled drains remained high enough to pose significant health risks to children upon contact.[86] Future studies should evaluate environmental fecal contamination at each point in the sanitation chain—within and outside the household—and consider absolute, context-independent risks from poor FSM in order to develop effective strategies for containment. Future studies should also consider the relationships between fecal indicator bacteria concentrations and enteric virus detection in environmental risk pathways to further understand whether variation in environmentally-persistent viruses can be a suitable indicator for FSM effectiveness.

Given the importance of exposures to fecal contamination in the public environment on pediatric diarrhea,[16,25,69,70,147,148,169] these findings provide new evidence that localized, high coverage of contained household sanitation, even when shared, is associated with reduced environmental fecal contamination in a high risk pathway outside the household. This finding underscores the importance of household-level sanitation and FSM on environmental fecal contamination, and subsequent risk, at the community level.

Supporting information: Detail on soil and drain sample collection and processing and Tables S1-S5.

Acknowledgements: The authors would like to thank Dorothy Peprah for her work on the household survey in Accra, Nii Wellington and the survey team at TREND and lab team at WRI and NMIMR for their work on the SaniPath study, and Clair Null for her work on the SaniPath study and invaluable guidance in preparing this manuscript. Additionally, the authors would like to acknowledge the community leaders and household participants who took the time to be part of this study.

Tables and Figures

Table 1: Population density, animal ownership, and neighborhood sanitation, including spatial clustering[a]

	Alajo (n = 200)	Bukom (n = 199)	Old Fadama (n = 197)	Shiabu (n = 197)	Overall (n = 793)
Population density (per km^2)	21,475	75,927	50,835	21,594	42,458
Public Toilets[b]	7	7	19	13	46
Number of contained public toilets (%[c])	4 (100)	5 (100)	0 (0)	6 (100)	15 (75)
HHs reporting owning domestic animals	65 (32.5)	41 (20.6)	28 (14.2)	69 (35.0)	203 (25.6)
# of HHs with ≥ 1 Sanitation Facility (%)	117 (58.5)	53 (26.6)	3 (1.5)	92 (46.7)	265 (33.4)
# of HHs in High Coverage Cluster (%[d])	-	29 (79.3)	-	94 (70.2)	
# of HHs in Low Coverage Cluster (%[d])	10 (0.0)	66 (6.1)	-	42 (7.1)	
# of HHs with ≥ 1 Contained Sanitation Facility[e] (%)	111 (55.5)	12 (6.0)	3 (1.5)	83 (42.1)	209 (26.4)
# of HHs in High Coverage Cluster (%[d])	-	16 (43.8)	-	93 (67.7)	
# of HHs in Low Coverage Cluster (%[d])	-	-	-	55 (9.1)	
# of HHs with minimally-shared sanitation facility[f] ≤ 5 HHs per toilet in the compound (%)	79 (39.5)	11 (5.5)	3 (1.5)	58 (29.4)	151 (19.0)
# of HHs in High Coverage Cluster (%[d])	32 (81.3)	-	-	93 (45.2)	
# of HHs in Low Coverage Cluster (%[d])	-	-	-	56 (7.1)	
≤ 30 people per toilet in the compound (%)	70 (35.0)	27 (13.6)	3 (1.5)	68 (34.5)	168 (21.2)
# HHs in High Coverage Cluster (%[d])	32 (71.9)	24 (54.2)	-	98 (51.0)	
# HHs in Low Coverage Cluster (%[d])	-	-	-	45 (4.4)	

[a]Chi-square tests of independence for each attribute across neighborhoods were significant at the 0.05 level. "-" indicates no cluster was significant at the 0.10 level by Kulldorff's Bernoulli spatial scan; [b]Number of public toilets in each neighborhood; [c]Observations were conducted at 4 public toilets in Alajo, 5 in Bukom, 6 in Old Fadama (though one was unable to be observed), and 6 in Shiabu [113]; [d]Coverage of the attribute in the cluster: number of households (% of

households in cluster) with that attribute. All clusters presented were detected by Kulldorff's Bernoulli spatial scan as significant at $\alpha = 0.05$. [e]A contained sanitation facility was one that, in the absence of sharing, would have been considered "improved" per the current JMP guidelines (including ventilated improved pit (VIP) or Kumasi ventilated improved pit (KVIP) latrine, pour-flush/flush toilets into a septic/sewage system, or traditional pit latrines with slabs);[29] [f]As shown, minimally-shared sanitation facilities were categorized as those with ≤ 5 households or those with ≤ 30 people using them and compared to all other households in the study area (both those with a facility shared by > 5 households or > 30 people and those without a facility).

Table 2: E. coli *concentrations and enteric virus detection in public domain samples, by neighborhood*[a]

	Alajo		Bukom		Old Fadama		Shiabu		Overall	
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)
a) *E. coli* detection in public domain samples										
E. coli in soil (log_{10}CFU/g)[†]	22	1.8 (1.6)	13	2.0 (1.5)	23	3.3 (1.5)	14	2.2 (1.2)	72	2.4 (1.6)
E. coli in drains (log_{10}CFU/100mL)[†]	26	8.4 (0.8)	19	8.9 (1.0)	23	8.1 (0.9)	25	8.8 (1.3)	90	8.5 (1.1)
	N	Virus+ (%)	N	Virus+ (%)	N	Virus+ (%)	N	Virus+ (%)	N	Virus+ (%)
b) Viral detection in drain samples[b]										
Adenovirus	25	17 (68)	19	16 (84)	21	16 (76)	23	21 (91)	88	70 (80)
GI norovirus	25	4 (16)	19	5 (26)	19	6 (32)	21	6 (29)	84	21 (25)
GII norovirus[‡]	23	6 (26)	18	8 (44)	19	13 (68)	24	13 (54)	84	40 (48)

[a]Samples collected Mar. – Nov. 2012. [b]Virus were detected in less than 5% of soil samples, thus soil sample results are presented in the text only. [†]$p < 0.05$ for effect of neighborhood in two-way ANOVA controlling for season of sample collection (rainy vs. dry). [‡]$p < 0.05$ for Chi-square test of conditional independence of neighborhood (including season of sample collection (rainy vs. dry)).

Table 3: E. coli *contamination in soil in the public domain by sanitation coverage cluster*

	Within 50m of soil sample (n = 58)		Within 100m of soil sample (n = 67)	
Main effect of model[a]	β	SE(β)	β	SE(β)
Any HH sanitation				
High Coverage Cluster	1.48	1.54	-0.51	0.89
Low Coverage Cluster	-0.64	0.75	-0.42	0.63
Contained HH sanitation				
High Coverage Cluster	2.03	1.06	1.09	0.93
Low Coverage Cluster	-0.66	0.74	-0.58	0.62
Minimally-shared HH sanitation				
≤ 5 HHs/toilet				
High Coverage Cluster	1.21	0.69	0.46	0.61
Low Coverage Cluster	-1.48	1.54	-1.63	1.58
≤ 30 people/toilet				
High Coverage Cluster	0.84	0.79	0.35	0.74
Low Coverage Cluster	-1.48	1.54	-1.63	1.58

[a]All models are adjusted for neighborhood (Alajo as reference), population density around the location of the sample, and season of sample collection (rainy/dry).

Table 4: E. coli *contamination in public drains by sanitation coverage cluster*

	Within 50m of drain sample (n = 58)		Within 100m of drain sample (n = 72)	
Main effect of model[a]	β	SE(β)	β	SE(β)
Any HH sanitation				
High Coverage Cluster	-0.64	0.48	-0.63	0.42
Low Coverage Cluster	0.14	0.37	0.28	0.32
Contained HH sanitation				
High Coverage Cluster	-3.65**††	1.33	-0.26	0.40
Low Coverage Cluster	1.45†	0.56	4.06*†	1.58
Minimally-shared HH sanitation				
≤ 5 households/toilet				
High Coverage Cluster	-0.90	0.56	-0.99†	0.40
Low Coverage Cluster	1.46††	0.54	1.30†	0.50
≤ 30 people/toilet				
High Coverage Cluster	-1.05†	0.42	-1.04††	0.36
Low Coverage Cluster	1.51†	0.65	0.19	0.55

[a]All models are adjusted for neighborhood (Alajo as reference), population density around the location of the sample, and season of sample collection (rainy/dry). *Interaction of main effect and population density ($p<0.10$); **Significant interaction of main effect and population density ($p<0.05$). †$p < 0.05$; ††$p < 0.01$

Figure 1: Neighborhood sanitation coverage and sample sites, Shiabu, Accra, Ghana

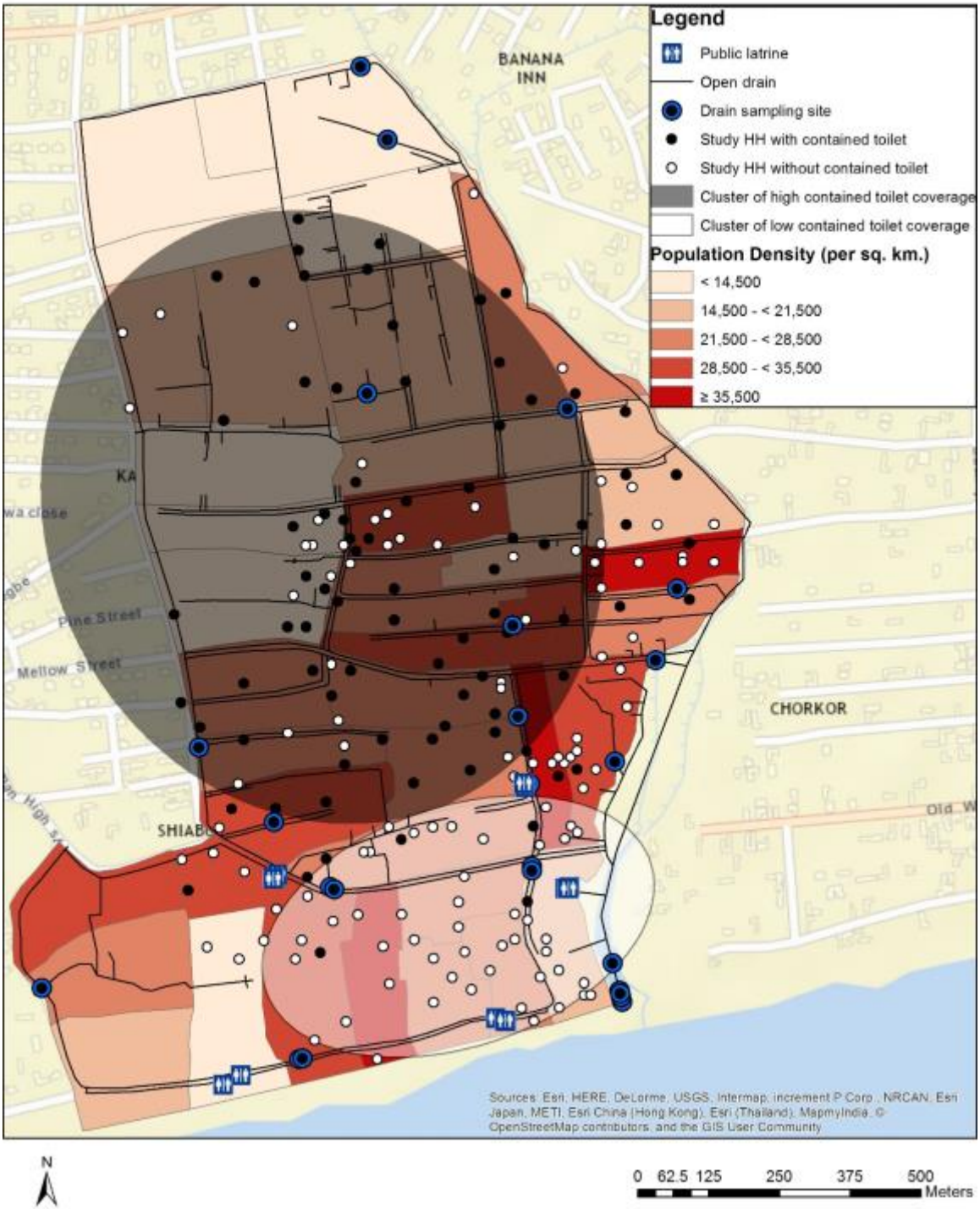


Figure 1: Neighborhood sanitation coverage and sample sites, Shiabu, Accra, Ghana. Drain sampling sites are illustrated using outlined circles. Households with a contained toilet are illustrated using black dots, while those without contained toilets (with uncontained toilets or no household sanitation facility present) are illustrated using white dots. Clusters of high (gray) and low (white) coverage of contained toilets are illustrated using ellipses.

Supporting Information (SI)**:**

Soil: Soil samples were collected using a sterile plastic scoop and sterile 250 mLWhirl-Pak bags (Nasco, Fort Atkinson, WI, USA). Seven separate samples totaling approximately 30g were collected within a 3m radius and combined into a single 250 mL Whirl-Pak bag. Samples were sealed, placed on ice in a cooler, and transported to WRI within 6 hours of collection. At the lab, the sample was weighed, mixed by rotation, and stored at 4°C until analysis. At the time of collection, the staff noted the date and whether the sample location was exposed to sunlight, within 3m of feces, and/or within 30m of a toilet or open defecation area.

Prior to membrane filtration, 10g of the composite sample were weighed into a sterile 50 mL conical tube with 20 mL of sterile phosphate-buffered saline (PBS). The sample was then vortexed for 30 seconds, adjusted to a pH of 9.0 by addition of 0.1N sodium hydroxide (NaOH), and shaken vigorously on a rotator or shaker for 30 minutes at room temperature. After 15 minutes of settling, 10 mL of the supernatant were aliquoted into a new sterile 50 mL conical tube, from which subsequent aliquots were taken for $1:10^0$, $1:10^1$, and $1:10^2$ dilutions for membrane filtration. Additionally, 1.5 mL of undiluted sample supernatant were aliquoted for PCR analysis at NMIMR. For PCR testing of GI and GII norovirus, 0.18g polyethelene glycol (PEG) was added, followed by centrifugation for 20 minutes at 6000 RPM and resuspension in sterile water, to concentrate the virus for RNA isolation.

Drain water: Samples of drain water were collected using a sterile bailer or a Sludge Nabber (Nasco, Fort Atkinson, WI, USA) with sterile 500 mL Whirl-Pak bags. The bailer or Sludge Nabber was submerged (or turned horizontally in shallow water) until full. Samples were deposited into the 500 mL Whirl-Pak bag until filled. Whirl-Pak bags were

sealed and labeled with the date, placed on ice in a cooler, and transported to WRI, where they were stored at 4°C until analysis.

Samples were diluted 1:10^5, 1:10^6, and 1:10^7 in sterile PBS prior to membrane filtration. Additionally, 1.5 mL of undiluted sample were aliquoted for PCR analysis at NMIMR. For PCR testing of GI and GII norovirus, 0.18g PEG were added, followed by centrifugation for 20 minutes at 6000 RPM and resuspension in sterile water, to concentrate the virus for RNA isolation.

All soil and drain samples were quantifiable above the lower limit of detection (1 CFU/100mL) for *E. coli* by membrane filtration. Samples evaluated by PCR were considered positive if both wells in OneStep analysis had cycle threshold (C_t) values ≤ 41 and within 5 C_t of one another. Estimated theoretical lower limits of detection for adenovirus, GI norovirus, and GII norovirus were 6.7 x 10^2, 3.3 x 10^4, and 6.7 x 10^3 genome copies, respectively, per gram for soil samples and 3.3 x 10^4, 1.7 x 10^6, and 3.3 x 10^5 genome copies per 100mL, respectively, for drain water samples.

Table S1: E. coli *contamination in soil and drains by season, population density, and local household animal ownership*

	E. coli in soil sample[a]		*E. coli* in drain sample[b]	
Main effect of model	β	SE(β)	β	SE(β)
Rainy season[c]	-0.43	0.60	0.16	0.33
Population density[d,e]	-1.29×10^{-6}	3.61×10^{-6}	-7.98×10^{-7}	2.01×10^{-6}
Prevalence of reported household animal ownership[d,f]				
Within 50m	-1.51×10^{-2}	6.72×10^{-2}	-2.61×10^{-2}	7.12×10^{-2}
Within 100m	0.08	9.50×10^{-2}	-5.29×10^{-2}	6.73×10^{-2}

[a]$\log_{10}$CFU/g; [b]$\log_{10}$CFU/100mL; [c]Adjusted for neighborhood; [d]Adjusted for season and neighborhood; [e]Per person per km^2; [f]Per 10% increase in prevalence;

Table S2: E. coli *contamination in soil in the public domain by local sanitation coverage*

	Within 50m of soil sample (n = 58)		Within 100m of soil sample (n = 67)	
Main effect of model[a]	β[b]	SE(β)	β[b]	SE(β)
Public toilet[c]	-0.19	0.46	6.33×10^{-2}	0.41
Any HH sanitation	5.39×10^{-3}	7.68×10^{-2}	-0.12	0.13
Contained HH sanitation	6.21×10^{-2}	8.38×10^{-2}	-4.51×10^{-2}	0.14
Minimally-shared HH sanitation				
≤ 5 HHs/toilet	0.16	8.34×10^{-2}	0.11	0.14
≤ 30 people/toilet	-6.18×10^{-2}	7.83×10^{-2}	1.50×10^{-3}	0.16

[a]All models are adjusted for neighborhood (Alajo as reference), population density around the location of the sample, and season of sample collection (rainy/dry).
[b]Estimate is for a 10% increase in sanitation coverage within the specified radius. [c]Presence or absence of public toilet within 50 or 100m.

Table S3: E. coli *contamination in public drains by local sanitation coverage*

	Within 50m of drain sample (n = 58)		Within 100m of drain sample (n = 72)	
Main effect of model[a]	β[b]	SE(β)	β[b]	SE(β)
Public toilet[c]	0.26	0.35	0.46	0.25
Any HH sanitation	-8.74×10^{-2}	6.80×10^{-2}	-7.77×10^{-3}	7.06×10^{-2}
Contained HH sanitation	-4.35×10^{-2}	7.92×10^{-2}	6.64×10^{-2}	7.11×10^{-2}
Minimally-shared HH sanitation				
≤ 5 HHs/toilet	-0.11	8.26×10^{-2}	-4.08×10^{-2}	7.82×10^{-2}
≤ 30 people/toilet	-0.13	7.81×10^{-2}	-2.27×10^{-2}	7.97×10^{-2}

[a]All models are adjusted for neighborhood (Alajo as reference), population density around the location of the sample, and season of sample collection (rainy/dry).
[b]Estimate is for a 10% increase in sanitation coverage within the specified radius. [c]Presence or absence of public toilet within 50 or 100m.

Table S4: Adenovirus, NoV GI, and NoV GII contamination in public drains by local sanitation coverage[a]

	Within 50m of drain sample (n = 58)			Within 100m of drain sample (n = 72)		
Main effect of model	Adenovirus OR (95% CI)[b]	GI norovirus OR (95% CI)[b]	GII norovirus OR (95% CI)[b]	Adenovirus OR (95% CI)[b]	GI norovirus OR (95% CI)[b]	GII norovirus OR (95% CI)[b]
Public toilet[c]	0.42	0.15	1.41	0.56	0.32	0.79
	(0.09, 2.02)	(0.00, 1.34)[d]	(0.28, 7.46)	(0.15, 2.12)	(0.04, 1.43)	(0.23, 2.65)
Any HH sanitation	0.65	0.95	0.91	0.77	1.06	0.80
	(0.44, 0.97)[†]	(0.65, 1.40)	(0.66, 1.25)	(0.53, 1.12)	(0.70, 1.61)	(0.56, 1.13)
Contained HH sanitation	0.76	1.19	0.77	0.99	1.01	0.83
	(0.50, 1.16)	(0.77, 1.84)	(0.52, 1.14)	(0.68, 1.44)	(0.69, 1.49)	(0.59, 1.19)
Minimally-shared HH sanitation						
≤ 5 HHs/toilet	0.82	1.22	0.98	0.92	1.24	0.80
	(0.52, 1.28)	(0.76, 1.94)	(0.67, 1.44)	(0.62, 1.36)	(0.80, 1.95)	(0.55, 1.17)
≤ 30 people/toilet	0.67	0.94	1.01	0.79	0.69	0.87
	(0.43, 1.07)	(0.62, 1.43)	(0.72, 1.43)	(0.53, 1.18)	(0.40, 1.19)	(0.59, 1.29)

[a]All models are adjusted for neighborhood (Alajo as reference), population density around the location of the sample, and season of sample collection (rainy/dry). [b]Estimate is for a 10% increase in sanitation coverage within the specified radius. [c]Presence or absence of public toilet within 50 or 100m. [d]estimated by Firth approximation

Table S5: Adenovirus, NoV GI, and NoV GII contamination in public drains by sanitation coverage cluster[a]

	50m vicinity of drain sample (n = 58)			100m vicinity of drain sample (n = 72)		
Main effect of model	Adenovirus OR (95% CI)	GI norovirus OR (95% CI)	GII norovirus OR (95% CI)	Adenovirus OR (95% CI)	GI norovirus OR (95% CI)	GII norovirus OR (95% CI)
Any HH sanitation						
High Coverage Cluster	0.60 (0.07, 7.10)[b]	0.16 (0.00, 1.94)[b]	0.27 (0.02, 2.72)	0.95 (0.12, 11.2)[b]	0.30 (0.01, 2.75)	0.44 (0.05, 3.19)
Low Coverage Cluster	2.94 (0.52, 30.9)[b]	0.43 (0.02, 3.60)	1.67 (0.32, 9.03)	2.35 (0.43, 19.2)	0.66 (0.08, 3.67)	1.85 (0.45, 8.01)
Contained HH sanitation[c]						
High Coverage Cluster	2.25 (0.17, 314)[b]	0.30 (0.01, 3.78)	1.76 (0.08, 6.23)	4.48 (0.40, 620)[b]	1.65 (0.22, 12.1)	0.70 (0.12, 3.90)
Low Coverage Cluster	0.36 (0.00, 6.36)[b]	3.10 (0.17, 105)	0.94 (0.08, 10.4)	0.29 (0.00, 4.71)[b]	5.81 (0.47, 165)	1.46 (0.17, 13.3)
Minimally-shared HH sanitation						
≤ 5 HHs/toilet						
High Coverage Cluster	0.48 (0.05, 6.17)[b]	1.19 (0.05, 13.6)	1.06 (0.08, 13.6)	0.47 (0.05, 4.62)	0.29 (0.01, 2.64)	0.43 (0.05, 2.83)
Low Coverage Cluster	0.24 (0.00, 4.10)[b]	4.39 (0.26, 145)	0.49 (0.04, 4.85)	0.28 (0.00, 4.66)[b]	2.20 (0.16, 34.5)	1.17 (0.13, 10.7)
≤ 30 people/toilet						
High Coverage Cluster	0.45 (0.07, 3.10)[b]	0.61 (0.03, 5.26)	1.30 (0.16, 9.97)	0.33 (0.05, 2.15)	0.39 (0.02, 3.03)	0.74 (0.10, 4.31)
Low Coverage Cluster	0.33 (0.02, 5.53)[b]	2.17 (0.07, 43.8)	0.41 (0.02, 5.93)	0.69 (0.04, 10.6)[b]	0.81 (0.03, 11.0)	2.18 (0.24, 24.0)

[a]All models are adjusted for neighborhood (Alajo as reference), population density around the location of the sample, and season of sample collection (rainy/dry); [b]estimated by Firth approximation; [c] A contained sanitation facility was one that, in the absence of sharing, would have been considered "improved" per the current JMP guidelines (including ventilated improved pit (VIP) or Kumasi ventilated improved pit (KVIP) latrine, pour-flush/flush toilets into a septic/sewage system, or traditional pit latrines with slabs)

The Influence of Household- and Community-Level Sanitation and Fecal Sludge Management on Urban Fecal Contamination in Households and Drains and Enteric Infection in Children

Running Head: Urban Sanitation and Fecal Contamination

David Berendes[1,2], Amy Kirby[2,3], Julie A. Clennon[2,4], Suraja Raj[2,3], Habib Yakubu[2,3], Juan Leon[2,3], Katharine Robb[2,3], Arun Kartikeyan[5], Priya Hemavathy[5], Annai Gunasekaran[5], Ben Ghale[5], J. Senthil Kumar[6], Venkata Raghava Mohan[6], Gagandeep Kang[5], Christine Moe[2,3]

[1]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [2]Center for Global Safe Water, Sanitation, and Hygiene, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [3]Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [4]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [5]Wellcome Research Laboratory, Christian Medical College, Vellore, India; [6]Department of Community Health, Christian Medical College, Vellore, India

Keywords: Sanitation, Urban, Fecal Contamination, Fecal Pathogens, Fecal Sludge Management

Word counts: 243 (abstract), 6,805 (text)

Number of figures and tables: 1 figure, 5 tables, 2 supplementary tables

Abstract:

Urban sanitation challenges policymakers to manage fecal sludge inside and outside the household. This study examined associations between household sanitation, fecal contamination, and enteric infection in two low-income neighborhoods in Vellore, India. Surveys and spatial analysis examined the presence and clustering of toilets and fecal sludge management (FSM) practices in 200 households. Fecal contamination was measured in environmental samples from 50 households and household drains. Enteric infection was assessed from stool specimens from children under five in these households. The two neighborhoods differed significantly in sanitation coverage (78% vs. 33%) and spatial clustering. Across neighborhoods, 49% of toilets discharged directly into open drains ("poor FSM"). Children in these households had 3.78 times higher prevalence of enteric infection when compared to children in other households, even those without toilets. Drains in poor FSM clusters had higher levels of enteric pathogens than elsewhere in the neighborhoods. Conversely, children in households with a toilet that contained excreta in a tank ("good FSM") had 55% lower prevalence of enteric infection compared to the rest of the study area. Household toilets in low sanitation coverage areas were associated with increased fecal contamination on floors where children played, while those in high coverage areas were associated with reductions in household fecal contamination and enteric infection in children when coupled with good household hygiene and FSM practices. Both sanitation coverage levels and FSM were associated with fecal contamination of the household and neighborhood and, subsequently, enteric infection prevalence in this urban setting.

Introduction

Poor water, sanitation, and hygiene are associated with multiple adverse health and developmental outcomes.[4,170] Following the Millennium Development Goals, the focus of the sanitation sector has been in rural areas, where an estimated 70% of those without access to improved sanitation live.[102] However, the need for sanitation solutions in poor urban neighborhoods and informal settlements is a growing concern, as the world's population distribution has shifted to being predominantly urban within the last decade.[103,138] By 2050, the urban population is estimated to almost double from its current estimate of 3.3 billion to 6.3 billion, providing a new challenge for sanitation implementers and policymakers.[103] Amongst the poorest and densest urban neighborhoods with little existing sanitation infrastructure, frequent person-to-person contact and poor environmental conditions facilitate transmission of fecal-oral infections, yielding frequent diarrhea in young children.[31,103–107,118]

Despite links between poor sanitation and health, evidence of effective sanitation in urban settings remains weak. Recent meta-analyses have identified few intervention data specific to urban neighborhoods and limited evidence of the positive effects of sanitation on diarrheal disease overall.[30,142] Among other limitations, the authors highlight bias in the measurement of the outcome (mostly self-reported diarrhea) and poor mechanistic evidence of changes in (more-proximal) exposure to fecal contamination with concurrent changes in sanitation coverage.[30] While several quantitative microbial risk assessments (QMRAs) have modeled environmental transmission pathways of exposure to fecal contamination and risk of enteric infection, few studies have measured associations between microbial indicators or pathogens in the

household or public environments and type or coverage of sanitation.[69,70] Because effective sanitation is expected to decrease enteric infection risk through safe containment of excreta, it is important to examine changes in environmental fecal contamination to better understand how sanitation affects enteric infection.[27]

Systems-level approaches to urban sanitation, where containment of excreta cannot be achieved by the presence of a household toilet alone, have not been well-examined.[138] When compared with rural settings, urban sanitation presents complex challenges, in particular the spillover of fecal contamination from private to public domains and vice-versa. This public-private domain interaction necessitates consideration of the entire sanitation chain in order to ensure safe containment, transport, treatment, and ultimately disposal or reuse of excreta.[32] Components of the sanitation chain start with the user interface (the household toilet), but also include transport (for example, sewerage or onsite containment followed by emptying and transport of excreta) and eventual treatment of the fecal sludge. All of these components are encompassed by the current focus on 'fecal sludge management' (FSM).[171] To date, associations between urban FSM downstream of the toilet, fecal contamination, and adverse health outcomes have only been evaluated for sewerage interventions, which was associated with reduced diarrhea incidence.[117] Because sewerage may not be feasible in some urban settings, and because many current sewer and open drain connections do not result in treatment of the excreta, it is important to examine the effects of other FSM models, including onsite containment, on health outcomes.[143]

In addition to the linear sanitation chain, studies also must consider the spatial heterogeneity of urban sanitation coverage. While the effect of sanitation on

environmental fecal contamination and enteric infection is often measured at the household-level, sanitation may have community-level benefits to those living within an area of high sanitation coverage, even if they themselves do not own a toilet.[36–38] Because of the interconnectedness of public and private urban environments, there is a growing need to examine how the concentrations of fecal contamination in the environment vary with the underlying spatial distribution and clustering of sanitation in the community.[46]

While urban management of fecal sludge in low-income countries is generally poor, the effects of FSM on fecal contamination and enteric infection within the urban environment have not been quantified.[31] There is a need to determine under what conditions toilets and FSM function to decrease fecal contamination and enteric infection risk in the urban environment. This study examines the effects of household toilets, their associated FSM, and their spatial clustering on fecal contamination within the household and the local urban environment, as well as their effects on pediatric enteric infection. By assessing proximal exposure outcomes and more distal infection outcomes in two different urban neighborhood environments, this work will contribute to understanding how the urban environment affects the success of sanitation interventions.

Materials and Methods

Data source

This study was conducted as a sub-study of a SaniPath Tool[172] deployment in two low-income urban neighborhoods in Vellore, India. Data was collected in February-March and September, 2014.

Study site

Two neighborhoods, Chinnallapuram (CAP) and Old Town (OT), in Vellore, India served as the study site. This study site was chosen because of the low income status, poor sanitation, and long-standing relationship with the Christian Medical College (CMC) and thus, spatial and demographic data were available from previous studies.[136,173] Further, the OT neighborhood is the study site for the Interactions of Malnutrition and Enteric Infections: Consequences for Child Health and Development (MAL-ED) study in Vellore.[136] Of the seven contiguous sub-neighborhood areas in Old Town selected for the MAL-ED study, five were selected for the SaniPath Tool deployment.

The CAP study area is a semi-urban neighborhood with a reported population density of 30,520 per km^2 over approximately a 0.41 km^2 area.[174] OT is an urban neighborhood with an estimated population density of 41,977 per km^2 over an area of 0.33 km^2.[136] The study area within Old Town was approximately 0.18 km^2. Vellore is subject to two monsoon seasons (a southwest monsoon from June to September and a northeast monsoon from October to December), with the remaining January to May period as a dry season.[136] Neighborhood sanitation consisted of household toilets, public toilets, and open defecation. Household toilets either discharged directly into open drains (poor FSM) or discharged into a tank under the house that contained the excreta (good FSM).

In each neighborhood, environmental samples and stool samples were collected from 25 households selected: 1) from previous sampling frames of CMC studies in those neighborhoods, and 2) based on the score from a hygiene survey developed by CMC and

implemented one month prior to SaniPath data collection.[175] Briefly, this survey assessed 18 general household hygiene characteristics and behaviors related to water collection, child/infant cleaning and feeding practices, and defecation, to create a household hygiene score. Scores less than or equal to 9 were classified as 'poor' hygiene, while scores greater than 9 were classified as 'good' hygiene. To ensure variation in general household-level hygiene practices, 13-14 'poor' hygiene and 11-12 'good' hygiene households were selected randomly in each neighborhood.

Ethical approval

Approval was obtained from the Emory University Institutional Review Board (IRB) and the Christian Medical College IRB. Informed consent was obtained prior to sample collection and survey administration at each household.

Environmental and stool sample collection, analysis, and processing

Environmental and stool samples were collected in March 2014. Five types of samples were collected at the household: hand rinse from children less than 5 years of age, rinse of a sentinel object, swab of household floors, 500 mL of drain water, and a stool sample from children under 5 years of age (see below for details). After collection, samples were stored on ice for up to 4 hours until arrival at the lab, where they were refrigerated at 4°C. Environmental samples were analyzed for *Escherichia coli* (*E. coli*) by membrane filtration, and Enteroaggregative *Escherichia coli* (EAEC) and Genogroup I and II norovirus by real-time, quantitative PCR and RT-PCR. *E. coli* was chosen as an indicator of fecal contamination, while EAEC and norovirus were chosen based on their high prevalence in the Vellore field site for the MAL-ED study.[25,96] Further, both are

predominantly human-specific infections.[23] *E. coli* concentrations were assessed within 6 hours of receipt by membrane filtration and plating on m-ColiBlue24® medium (Hach Company, Loveland, Colorado, USA) according to EPA method 1604.[149] EAEC and norovirus concentrations were assessed by quantitative real-time PCR (see below).

Stool samples were analyzed for enteropathogens using the MAL-ED study protocols, with the exception of *Campylobacter* spp., which was assessed by PCR.[25,65,176]

Hand rinses: Hand rinses were collected from the child under 5 years old in the household who was previously enrolled in other CMC studies. The child's right hand was inserted into a sterile 2 L Whirl-Pak bag containing 500 mL of sterile phosphate-buffered saline (PBS) solution. The staff massaged the fingers and palm for 30 seconds in the PBS, then the child removed their right hand and inserted the left hand, with the staff repeating the massage procedure. At the lab, the hand rinse was diluted $1:10^0$, $1:10^1$, and $1:10^2$ in sterile PBS prior to membrane filtration. Prior to PCR, 200 mL of the original hand rinse sample was precipitated with 12% polyethylene glycol 8000 (PEG), centrifuged for 20 minutes at 6000 RPM, and suspended in 5 mL sterile water, of which 1.5 mL was further concentrated by precipitation with 12% PEG 8000 prior to nucleic acid extraction.[177]

Sentinel objects: A child's toy or feeding spoon, volunteered by the mother, was used as the 'sentinel object' in the household. The object was inserted into a sterile 2 L Whirl-Pak bag containing 500 mL of sterile PBS, massaged from the outside of the bag for 1 minute, and subsequently removed and returned to the family. The rinses from sentinel objects were processed identically to the hand rinse samples for membrane filtration and PCR.

Household floor swabs: Composite household floor swabs were collected using EnviroMax Plus Sterile Environmental Swabs (Puritan Medical Products, Guilford, ME, USA). Once the child play area was identified for the field staff by the mother, the staff used a 20 cm framing square to outline a 25 cm^2 area in each of 4 corners and the center of the floor and swabbed back and forth across those sections of the floor. Two swabs were used to cover the entire area and subsequently combined into a single sample covering a total surface area of 125 cm^2. Each household floor swab was eluted in 7 mL of PBS solution in a sterile container, and the eluates from both swabs were combined for an approximate sample volume of 14 mL. From this volume, dilutions of $1:10^0$, $1:10^1$, and $1:10^2$ were made and membrane filtered. Nucleic acids were extracted from 1.5 mL of swab eluate following one round of PEG precipitation.

Drain water: Samples of drain water were collected from the drain directly in front of the household. A sterile bailer or stainless steel ladle was used to collect approximately 500 mL of drain water, which was deposited into a sterile 2 L Whirl-Pak bag (Nasco, Fort Atkinson, WI, USA), taking care not to disturb sediment on the bottom or nearby trash. Drain samples were diluted $1:10^1$, $1:10^2$, and $1:10^3$ in sterile PBS at the lab prior to membrane filtration. DNA and RNA were extracted from 1.5 mL samples of the original sample prior to analysis by PCR.

Because almost all drain water samples collected during the initial (February-March) sampling period had colony numbers above the countable range on the filter membrane, 10 of the original 25 households in each neighborhood were resampled spatially at random in September 2014, per the original sampling protocol. These samples were analyzed by membrane filtration after $1:10^4$-10^6 dilution.

Stool samples: Stool samples were collected from all children under 5 years old in the 50 study households. Stool samples were processed and analyzed according to the MAL-ED protocols for pathogen detection in stool.[65]

Quantitative Real-time PCR: Total nucleic acids were extracted using the Qiagen Xtractor system (Qiagen Sciences, Germantown, Maryland, USA), following manufacturer's instructions. EAEC was detected using primers and probes targeting the *aatA* gene.[59] GI norovirus was detected using genogroup-specific COG1 primers and RING1-TP probe, while GII norovirus was detected using the genogroup-specific COG2 primers and RING2-TP probe.[151] All samples were tested using the Qiagen QuantiFast Pathogen + IC kit (PCR for EAEC and RT-PCR for norovirus) for initial screening and assessment of potential PCR inhibition. Any samples that were positive (at least one well with a cycle threshold (C_t) value less than 45) or inhibited were quantified using the OneStep PCR (EAEC) or RT-PCR (norovirus) kit (Qiagen) and a standard curve. The standard curve for EAEC was generated from a plasmid containing the *aatA* gene. The standard curve for GI and GII norovirus was generated from in vitro transcribed RNA.[100] Positive and negative controls for EAEC or norovirus were included with every PCR run.

Samples tested for norovirus GI or GII using the OneStep kit and classified as positive (both wells had C_t values less than or equal to 45 and a difference of less than or equal to 4 between C_t values for duplicate wells) were quantified by a simple average of both wells. Due to inconsistencies with the standard curve, samples tested for EAEC using the OneStep kit were not quantified but instead were classified as positive or negative. Samples with no detectable EAEC or norovirus were assigned the value of the theoretical lower limit of detection for the assay (334 cell equivalents (CE) for EAEC or

genome equivalent copies (GEC) for norovirus GI and GII per 100mL (2.52 $\log_{10}$ CE or GEC/100mL)).

Survey data collection

Household surveys were conducted in 100 households in each study neighborhood, 25 of which were households with concurrent environmental and stool sample collection and 75 of which were divided across political subdivisions of the neighborhood and chosen spatially at random within those subdivisions. To be eligible for this study, households had to have a child under five years of age. A Global Positioning System (GPS) location was collected for each household using Garmin eTrex Venture HC devices (Garmin International Incorporated, Olathe, Kansas, USA) prior to administration of the survey. The target respondent for the survey was the person responsible for water, sanitation, hygiene, and food activities, generally the mother of the youngest child, or rarely, the grandmother. If the respondent was not available and the household was one of the households where environmental stool samples were to be collected, survey enumerators returned to the household at a later time, otherwise, the nearest available household was selected for survey. The household survey included questions about the household's population, sanitation, and fecal sludge management (FSM) practices, as well as the children's and adult's defecation practices.

Analyses

Values below the lower detection limit for membrane filtration were approximated on the log scale using the value of the lower limit of detection of 1 CFU/100mL and accounting for sample dilution. Concentrations per pair of hands,

sentinel object, and 125cm^2 of household floor were then back-calculated using the rinse volume for these samples. All microbial concentrations were $\log_{10}$-transformed prior to statistical analyses.

Aspatial statistical analyses were conducted in R version 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria) using base packages as well as the "logistf" package for penalized likelihood estimation and the "lme4" package for mixed-effects logistic regression.[157,178,179] Linear regression was used to assess continuous outcomes (*E. coli* concentrations in all environmental samples and norovirus GII concentrations in drain water), while logistic regression was used to assess binary outcomes (presence/absence of EAEC and norovirus GII in drain water, as well as stool pathogen detection). Penalized likelihood estimation, referred to as Firth approximation, was used when separation was observed in logistic regression models.[158] Aspatial model residuals were assessed for global autocorrelation (see below).

Binary (presence/absence) data from household surveys were evaluated for most-likely local clustering using Kulldorff's Bernoulli spatial scan in SaTScan version 9.4[52]. Kulldorff's Bernoulli spatial scan evaluates binary outcomes in point data distributed in space to assess the degree of nonrandom clustering of '0' or '1' values. Though p-values are presented, an α of 0.05 was used for all tests of significance in autocorrelation, clustering (including adjustment for multiple comparisons in Getis G* analyses), and regression modeling.

Results

Frequency and within-neighborhood spatial clustering of household sanitation

To compare sanitation coverage and spatial heterogeneity within and between study neighborhoods, we assessed the frequency and type of household sanitation and fecal sludge management (FSM) and their most-likely clustering in Chinnallapuram (CAP) and Old Town (OT) (Table 1 and Figure 1). In both neighborhoods, all toilets were pour-flush toilets (data not shown) and generally either: 1) discharged directly to an open drain (defined as "poor FSM") or 2) were connected to a tank under the household that contained the excreta (defined as "good FSM"). Compared to households in CAP, households in OT reported a significantly lower proportion of household toilets (33% vs. 78%). More household toilets in OT had poor FSM than in CAP (82% vs. 35%). Open defecation was higher amongst adults (68% vs. 18%) and children under 5 (80% vs. 40%) in OT than in CAP. Public toilet use varied between CAP and OT. While about the same percentages of households reported any use of a public toilet (59% in CAP and 54% in OT), significantly more households reported high use (more than 10 times per month) in OT compared with CAP (18% vs. 4%).

Significant spatial clusters of low toilet coverage and poor FSM were present in both CAP (Figure 1a) and OT (Figure 1b); however, significant clustering of high toilet coverage was only present in CAP (Table 1b).

Microbiological concentrations in environmental samples and enteric pathogen detection in stool

Levels of fecal contamination in the household were characterized by examining rinses of children's hands, rinses of sentinel objects, and swabs of household floors. *E. coli*, enteroaggregative *E. coli* (EAEC), and GI and GII norovirus levels were quantified in these household samples. *E. coli* was detected in 45/50 rinses of children's hands (geometric mean: 107.2 coliform-forming units (CFU) per pair of hands (standard deviation (SD): 11.7 CFU per pair of hands)), 32/49 rinses of sentinel objects (geometric mean: 13.2 CFU per 100mL (SD: 5.9 CFU per 100mL)), and 48/50 household swabs (geometric mean: 245.5 CFU per 125 cm^2 (SD: 9.8 CFU per 125 cm^2)). Distributions of sample *E. coli* concentrations were all approximately normal when log-transformed, though sentinel object rinses did exhibit a large proportion of left-censored values from non-detects (data not shown). EAEC was detected in 1/50 hand rinse samples, 0/50 sentinel objects rinses, and 1/50 floor swab. GI norovirus was not detected in any samples within the household. GII norovirus was detected in 1/50 hand rinse samples, 0/50 sentinel object rinses, and 0/50 floor swabs. These EAEC and GI and GII norovirus samples types were omitted from further analyses due to low levels of detection.

Levels of fecal contamination outside the household were characterized by examining drain samples, which were analyzed for *E. coli*, EAEC, and GI and GII norovirus. *E. coli* was detected in all 50 drain samples collected, with concentrations above the detection limit in 49/50 samples, thus 10 drain sample locations from each neighborhood were randomly chosen for re-sampling in September, 2014. Of these 20 samples, *E. coli* was detected all samples and was within a quantifiable range in 19/20

samples, with a geometric mean of 6.83 $\log_{10}$CFU per 100mL (SD: 0.56 $\log_{10}$CFU per 100mL) and an approximately normal distribution when log-transformed. However, because this drain sampling took place 6 months after initial data collection was completed, drain *E. coli* concentrations were excluded from further analyses. EAEC, GI norovirus, and GII norovirus were detected in 15/50, 1/50, and 19/50 drain samples, respectively. Mean concentrations of EAEC and GII norovirus were 2.67 $\log_{10}$cell equivalents (CE) per 100mL (SD: 0.41 $\log_{10}$CE per 100mL) and 3.43 $\log_{10}$genome equivalent copies (GEC) per 100mL (SD: 1.41 $\log_{10}$GEC per 100mL), respectively.

To assess differences in fecal contamination within households by neighborhood and household hygiene practices, logistic and linear regression models for *E. coli* data were constructed (Table 2). Overall, there was no significant variation in *E. coli* detection or concentrations in samples by neighborhood or hygiene status.

To determine the prevalence of enteric infections in children, stool specimens from all children under 5 years of age in study households where environmental samples were collected were assayed for viral, bacterial, protozoan, and parasitic enteric pathogens. Overall, one or more enteric pathogens was detected in 51/76 (67%) of children's stool. Astrovirus (7/76 children), *Campylobacter* spp. (32/76), *Entamoeba histolytica* (1/76), *Giardia* spp. (17/76), GII norovirus (5/76), and pathogenic *E. coli* (14/76) were detected in stool specimens.

Variation in enteric pathogen prevalence by neighborhood and household hygiene status was assessed using mixed-effects logistic regression (Table 3). The prevalence of any enteric infection, as well as infection with specific pathogens, did not vary significantly by neighborhood. Households with poor hygiene status had significantly

higher detection of *Campylobacter* spp. in stool compared to those with good hygiene status. Because no single pathogen was associated with more than half of infections across both neighborhoods, pathogen-specific analyses were not conducted and only the presence of any enteric infection (i.e. pooled pathogens) was assessed in further regression analyses.

The significance of most-likely clusters of pathogen-positive/negative-stool was assessed by Kulldorff's Bernoulli spatial scan (data not shown). No significant clusters of infection were detected.

Association between household- and cluster-level sanitation coverage, FSM practices, and within-household fecal contamination

The influence of household sanitation and FSM practices on within-household fecal contamination was examined using multivariate linear regression, with and without spatial lag. Household-level and cluster-level sanitation variables were considered and associations were compared within and between study neighborhoods. Because significant spatial lag was not observed in regression models and aspatial model results did not exhibit significant global autocorrelation, aspatial modeling results are presented (Tables S1-3). At the household level, the presence of a household toilet was associated with significantly lower *E. coli* concentrations on children's hands in households with "good" hygiene status (difference of 1.54 $\log_{10}$CFU/pair of hands, Table S1). Further, children in households with toilets containing excreta onsite (good FSM) and good hygiene had significantly lower *E. coli* on hands as well (difference of 1.19 $\log_{10}$CFU/pair of hands). *E. coli* concentrations on floors of households with a toilet and good hygiene status in CAP were significantly lower than those with poor hygiene

(difference of 2.06 log_{10}CFU/125 cm^2), while households in OT with a toilet had significantly higher *E. coli* concentrations on floors than those without a toilet (difference of 1.53 log_{10}CFU/125 cm^2, Table S2). *E. coli* concentrations in rinses of sentinel objects did not vary significantly with any of the sanitation variables tested (Table S3).

Association between demographics, neighborhood, and household- and cluster-level sanitation variables and fecal contamination outside the household

To examine variation in pathogen levels in the public domain with neighborhood, household hygiene, and sanitation, logistic regression models for EAEC and both logistic and linear regression models for GII norovirus in drains were constructed. An alpha of 0.05 was used for determining significance; however, associations with large effect sizes and nonsignificant p-values are included. Of note, though not significant, EAEC detection was lower in drains outside households with poor hygiene, compared to those with good hygiene statuses (OR: 0.30, 95% CI: 0.08, 1.02, p = 0.06), but did not vary by neighborhood. Controlling for both neighborhood and hygiene status, variation in EAEC detection did not approach significance with household-level sanitation variables (all p-values > 0.05). However, though not significant, EAEC was less likely to be detected outside households in low coverage clusters of household toilets (OR: 0.27, 95% CI: 0.05, 1.14, p = 0.09), though variation with other cluster-level sanitation variables did not approach significance.

GII norovirus detection and concentrations in drains did not vary significantly by neighborhood or hygiene status (Table 4a). At the household-level, the odds of detecting norovirus GII in drain samples and the concentrations of norovirus GII in drains were higher for drains adjacent to households with toilets compared to those without toilets,

though this finding was not significant (Table 4b). At the cluster-level, households within the cluster of high toilet coverage had significantly higher GII norovirus concentrations in drains (Table 4c).

Household- and cluster-level sanitation and enteric infection in children

Associations between household- and cluster-level sanitation and concurrent enteric infection in children (detection of any enteric pathogen in children's stool) were evaluated by mixed-effects logistic regression (Table 5). Children in households with a toilet with good FSM had 55% lower prevalence of infection compared to children in households with toilets with poor FSM or no toilet present, though the association was not significant. Conversely, children in households with a toilet with poor FSM had 3.78 times the prevalence of infection across both neighborhoods. Similar relationships were observed when comparing households with toilets and good or poor FSM practices to households without toilets. Prevalence of enteric infection did not vary significantly with cluster-level sanitation variables.

Discussion

This study examined the effects of household toilets, fecal sludge management (FSM), and spatial heterogeneity of sanitation coverage on fecal contamination within the household and the local urban environment and enteric infection in young children. The results suggest that FSM and neighborhood-level coverage and spatial clustering of sanitation have significant effects on household- and neighborhood-level fecal contamination and pediatric enteric infection prevalence. Enteric infection was least

prevalent among children in households with good FSM and most prevalent among children in households with poor FSM. Clusters of high sanitation coverage with poor FSM had higher concentrations of norovirus GII in drains adjacent to the house compared to the rest of the study area. In households in the neighborhood with low (~30%) toilet coverage, the presence of a toilet was associated with higher *E. coli* concentrations on household floors, while in clusters of high sanitation coverage, modest, though nonsignificant, decreases in *E. coli* concentrations and enteric infection were observed. Hygiene practices within the household were an important effect modifier of toilet presence on *E. coli* concentrations in both cases.

This study is one of the first to quantify the spatial heterogeneity in both household sanitation and reported FSM in an urban setting and assess its impact on household fecal contamination, as measured by *E. coli* concentrations, and pediatric enteric infection. Comparison of onsite excreta containment in tanks under the household (which must be desludged and transported away) to open drainage is new to the literature, which has previously focused on sewerage when considering FSM.[117,180,181] While some previous literature has shown lower incidence of pediatric diarrhea associated with urban drainage interventions, our findings indicate that toilets that fail to contain fecal sludge along the entire sanitation chain can be potential risk factors for fecal exposure and enteric infection.[120,180,182] Finally, associations between household- or community-level sanitation and fecal indicator bacteria or enteric pathogen concentrations within the household and local public domain in low-income, urban settings have rarely been observed. Studies collecting these data have generally modeled them in QMRAs, rather than evaluating them as outcomes alone.[68–70,73,85]

Onsite containment of feces, if properly desludged and transported away after filling, should yield lower levels of environmental contamination, and subsequent prevalence of enteric infection, when compared with open drainage. While drain concentrations of norovirus GII outside of households with good FSM were not significantly lower than those outside households with poor FSM, they may not be indicative of the levels of all pathogens and may not mirror disease transmission, as norovirus is known to have more person-to-person transmission than through open drain water.[15] Even if concentrations remain unchanged in drains, containment of fecal sludge onsite reduces the volume discharged into drains, lowering water levels in small drains immediately outside the household and reducing drain contact and transmission of fecal pathogens.[86] Concentrations of fecal indicator bacteria and pathogens in open drains approximate that of pure sewage, especially in low-income urban areas.[69,85,86,88,160] Many enteric pathogens, including those most detected in stool in this study, survive for long periods in the environment, especially environmental waters with large organic loads, thus prevention of contact with drains or re-routing of sewage is critical.[9,17] Child contact with open drains has been identified as a significant contributor to diarrhea disease burden in urban environments, including elevated attributable risk in settings with recent increases in sanitation coverage.[69,70,73,154] Containment of fecal sludge, mainly via sewerage, has been shown to interrupt pathogen transmission from open drains.[69,117] In the absence of this containment, combining increased coverage of toilets with high prevalence of open drains puts children at risk of exposure to high concentrations of pathogenic organisms.[89] Given the cost, planning, and operation and maintenance logistics involved with sewered systems, onsite containment of excreta may provide a

more realistic alternative for downstream management of fecal sludge that should be further evaluated.[183]

In areas of low sanitation coverage, increased fecal contamination in households with toilets may reflect sharing of sanitation, and subsequently poorer toilet maintenance and hygiene. Although our study did not collect information on households' sharing of toilets, the practice is not uncommon among households without sanitation in poor urban areas of India.[78] Household shared sanitation has previously been linked to increased diarrheal prevalence in children.[115,144] While household shared sanitation has not been directly associated with increased fecal contamination, measurement of fecal contamination in households has been inconsistent.[78,116] In rural communities and schools, studies have shown increased fecal contamination on hands associated with the presence and use of toilets, suggesting that poor hygiene practices modify this relationship.[79,83] Within study neighborhoods, the hygiene survey did not collect data on general toilet maintenance and no differences were observed in self-reported handwashing practices between different types of household sanitation, though this metric is subject to response bias.[184–186] However, only households with both a toilet and good hygiene status (representative of positive household WASH behaviors) had significant reductions in fecal contamination on children's hands and household floors. Further, higher concentrations of fecal contamination on children's hands were moderately associated with presence of enteric pathogens in their stool. These findings underscore the importance of improving hygiene practices at the same time as improving sanitation infrastructure to offset the potential for increased contact with fecal contamination at toilets.[83]

In contrast to increases in fecal contamination associated with household toilets in low sanitation coverage areas, clusters of high coverage of toilets were associated with decreased within-household fecal contamination and enteric infection. While decreases in fecal contamination were modest, we observed a 31% decrease in the prevalence of enteric infection in children in these households. When evaluated without regard to spatial clustering within the neighborhood, household toilet presence had no effect on these outcomes. Evidence of differential associations between sanitation and fecal contamination by local sanitation coverage level, including clustering, supports the premise of sanitation as a community-level good with a threshold to be attained before benefits can be observed. Under these threshold-saturation models, once reaching the threshold, localized areas of high sanitation coverage reduce community-level fecal contamination, conferring benefits to all households in that area, not only the ones with toilets.[27,34,36–38] However, increases in fecal contamination observed in drains in high sanitation coverage clusters with poor FSM suggest that community-level reductions in fecal contamination were not achieved. Combined with evidence that household hygiene may modify associations between toilets and fecal contamination, this finding suggests that both FSM and hygiene practices are important to attain community-level benefits from sanitation.

While this study included multiple FSM typologies and used microbiological measurements of environmental samples to evaluate the effectiveness of sanitation in urban household and neighborhood environments, it is limited in its assessment of hygiene practices as a potential household confounder related to fecal contamination. Classification of households into "good" and "poor" hygiene status was based on a 50%

score cutoff within the index, thus households near the cutoff likely did not vary significantly in their WASH practices, leading to potential misclassification. Further, classification of household hygiene status included questions about practices at the household toilet, thus households without a toilet were more likely to score lower on the hygiene scale, limiting our ability to separate conclusions about sanitation and hygiene practices.

While a census was not feasible, household selection approximated the spatial distribution of households by random selection within each of five divisions in the study area. Using this approach and estimating clusters by spatial scan provided a more accurate assessment of the spatial heterogeneity underlying the neighborhood, which is known to be significant with regard to socio-economic variables in low-income, urban settings.[187]

Even with limited sample sizes, several variables were non-significant, but had large effect sizes worth further investigation. However, assessment of fecal contamination in environmental samples and enteric infection in stool as exposure and infection outcome measures, respectively, improved objectivity. Because household exposure metrics for fecal contamination are inconsistent between studies and subject to high temporal variation, three sample types within the household (hand rinses, rinses of sentinel objects, and swabs of household floors where the child played) were measured, with relatively high correlation between *E. coli* concentrations in different types of samples from the same household.[77–79,84] Samples of urban drains in front of the household allowed for quantification of fecal contamination spread in the public domain. Enteric infection was assessed from stool specimens from children under 5 employing a

multi-pathogen panel used in other clinical assessments, avoiding issues of response bias associated with self-reported diarrhea and inability to detect asymptomatic infections.[30,65]

Future urban sanitation studies must move beyond the household toilet to consider household FSM within the neighborhood environment. Despite new efforts to diagnose FSM conditions, the effects of FSM typologies on household- and neighborhood-level fecal contamination and enteric infection is not well described in the literature.[31,32,188] Open drains persist as default sewerage options throughout low-income, urban settings. Given the interconnectedness of urban communities, future studies should quantify the effects of this hazard at the household-, neighborhood-, and city-scales.[31]

Overall, this study provides evidence of the importance of both fecal sludge management and the spatial heterogeneity of sanitation within neighborhoods when evaluating the effectiveness of urban sanitation. In order to reduce fecal contamination and improve health, good FSM must accompany increases in toilet coverage in order to remove and safely contain sewage from drains and other compartments of the urban public and private domain. Because isolated toilets in low coverage areas may not yield the same benefits as clusters of toilets in high coverage areas, sanitation must be considered at the community-level. As urban populations continue to grow, sanitation interventions must contain excreta in both the public and private domain in order to protect health and the environment.

Acknowledgements: The authors would like to thank the field and lab staff at CMC in Vellore, including Sheela Roy and field teams, for their effort on this study and all aspects of the SaniPath study in Vellore, India.

Financial Support: This study was funded by the Bill and Melinda Gates Foundation.

Disclosures: The funders had no role in the design or implementation of this study. The authors declare no competing interests.

Tables and Figures

Table 1: Reported frequency and clustering of household sanitation and fecal sludge management (FSM) in Chinnallapuram and Old Town

	Chinnallapuram (n = 100) Number (%)	**Old Town** (n = 100) Number (%)	**Overall** (n = 200) Number (%)	**p-value**[a]
a) Household-level				
Household toilet[b]	78 (78.0)	33 (33.0)[c]	111 (55.5)	<0.01
FSM: *Toilet excreta contained onsitc*[d]	37 (47.4)	3 (9.1)	40 (36.0)	<0.01
FSM: *Toilet discharges directly to drain*[d]	27 (34.6)	27 (81.8)	54 (48.6)	<0.01
FSM: *Other/don't know*[d]	14 (18.0)	2 (6.1)	16 (14.4)	0.18
Open defecation				
< 5 year olds	40 (40.0)	80 (80.0)	120 (60.0)	<0.01
Respondent (adult)	19 (19.0)	68 (68.0)	87 (43.5)	<0.01
Public toilet use (by respondent)				
None	41 (41.0)	46 (46.0)	87 (43.5)	0.57
Low (1-5 times per month)	51 (51.0)	31 (31.0)	82 (41.0)	0.01
Medium (6-10 times per month)	4 (4.0)	5 (5.0)	9 (4.5)	>0.99
High (>10 times per month)	4 (4.0)	18 (18.0)	22 (11.0)	<0.01
	Chinnallapuram	**Old Town**		
b) Most likely clusters[e]	Count (Cluster prevalence)[f]	Count (Cluster prevalence)[f]		
Household toilet				
High Coverage Cluster	43 (100.0)	-		
Low Coverage Cluster	40 (50.0)	27 (0.0)		
FSM: Toilet discharges directly to drain				
High Coverage Cluster	18 (77.8)	9 (100.0), 7 (100.0)		

[a]p-value for T-test of proportions between neighborhoods; [b]All toilets were pour-flush toilets; [c]Of the 33 households reporting having a toilet, 32 responded to the subsequent questions about FSM; [d]Percent in parentheses represents the percentage of all households with toilets; [e]No significant clusters of households with toilet excreta contained onsite were observed; [f]Each cluster presented had a prevalence of the attribute within it that was significantly different from the overall prevalence of the attribute in the neighborhood at 0.05 level, otherwise " – " is presented in section (b) of the table.

Table 2: Variation in detection and concentrations of E. coli in environmental samples within households with neighborhood and hygiene status

	Child hand rinse (n = 50)				**Sentinel object rinses** (n = 49)				**Household swabs** (n = 50)			
	E. coli detection	*E. coli* concentration			*E. coli* detection	*E. coli* concentration			*E. coli* detection	*E. coli* concentration		
	OR (95% CI)[a]	β[b]	SE(β)[c]	p-value	OR (95% CI)[a]	β[d]	SE(β)[c]	p-value	OR (95% CI)[a]	β[e]	SE(β)[c]	p-value
Neighb.[f]	0.22 (0.01, 1.62)	0.51	0.30	0.09	1.27 (0.39, 4.22)	0.15	0.22	0.51	1.00 (0.04, 26.3)	0.07	0.28	0.81
Poor hyg.[g]	0.18 (0.01, 1.36)	-0.07	0.31	0.83	0.99 (0.30, 3.28)	0.19	0.22	0.40	0.85 (0.03, 22.2)	-0.28	0.28	0.33
Neighb.[f]	0.19 (0.01, 1.48)	0.51	0.30	0.10	1.28 (0.39, 4.23)	0.16	0.22	0.48	0.99 (0.04, 26.1)	0.06	0.28	0.84
Poor hyg.[g]	0.16 (0.01, 1.24)	-0.05	0.30	0.88	1.01 (0.31, 3.34)	0.20	0.22	0.38	0.85 (0.03, 22.3)	-0.28	0.28	0.34

[a]Odds ratio for detection. Though p-values for ORs are omitted for reasons of space in the table, none approached significance at $\alpha = 0.05$. [b]Estimate is in $\log_{10}$coliform-forming units (CFU) per pair of hands. [c]Standard error. [d]Estimate is in $\log_{10}$CFU per 100mL. [e]Estimate is in $\log_{10}$CFU per 125 $cm.^2$ [f]Old Town neighborhood (reference is Chinnallapuram). [g]Hygiene status was divided into "poor" or "good" hygiene categories based on a 18-point scale (0-9 as "poor', 10-18 as "good") discussed in the methods and presented in Collinet-Adler et al.[175]

Table 3: Variation in detection of enteric pathogens in stool with neighborhood and hygiene status[a]

	Any enteric pathogen		***Campylobacter* spp.**		***Giardia* spp.**		**Pathogenic *E. coli*[b]**	
	PR (95% CI)[c]	p-value	PR (95% CI)[c]	p-value	PR (95% CI)[c]	p-value	PR (95% CI)[c]	p-value
Neighborhood: Old Town	1.32 (0.50, 3.49)	0.57	1.91 (0.71, 5.14)	0.18	1.56 (0.48, 6.46)	0.45	0.73 (0.14, 3.78)	0.64
Poor hygiene	1.97 (0.75, 5.62)	0.17	3.42 (1.30, 12.3)	0.02	1.69, (0.53, 7.27)	0.39	0.55 (0.07, 3.37)	0.34
Neighborhood: Old Town	1.37 (0.51, 3.72)	0.53	2.15 (0.80, 5.97)	0.13	1.60 (0.49, 6.78)	0.47	0.70 (0.11, 5.02)	0.56
Poor hygiene	2.01 (0.76, 5.77)	0.16	3.61 (1.37, 11.4)[†]	0.01	1.72 (0.54, 7.57)	0.54	0.53 (0.09, 4.10)	0.31

[a]N = 76 children from which stool specimens were collected (43 in Chinnallapuram, 33 in Old Town). Enteric pathogens detected in stool specimens included astrovirus, *Campylobacter* spp., *Entamoeba histolytica*, *Giardia* spp., GII norovirus, and pathogenic *E. coli*. A full list of organisms tested in stool specimens is presented in Houpt et al.[65]. Only pathogens detected in >20% of stool specimens were regressed against neighborhood and hygiene status. [b]*EAEC, EHEC, EPEC, ETEC*. [c]Prevalence ratio for detection of enteric pathogen in stool specimen and 95% confidence interval.

Table 4: Variation in GII norovirus detection and concentration in drain water

	GII norovirus detection		GII norovirus concentration		
	OR (95% CI)[a]	p-value	β[b]	SE(β)[c]	p-value
a) Demographics and hygiene					
Neighborhood: Old Town	0.42 (0.13, 1.34)	0.15	-0.38	0.39	0.35
Poor hygiene	0.92 (0.29, 2.91)	0.88	-0.05	0.40	0.90
Neighborhood: Old Town	0.42 (0.12, 1.33)	0.15	-0.38	0.40	0.35
Poor hygiene	0.88 (0.27, 2.86)	0.83	-0.07	0.41	0.87
b) Sanitation (household-level)					
Household toilet	4.73 (0.93, 28.9)	0.07	0.90	0.53	0.10
Toilet excreta contained onsite[d]	1.51 (0.37, 6.10)	0.56	0.41	0.49	0.41
Toilet discharges to drain[d]	1.14 (0.28, 4.74)	0.85	0.17	0.50	0.74
Toilet excreta contained onsite[e]	1.81 (0.38, 8.95)	0.45	0.61	0.55	0.28
Toilet discharges to drain[e]	1.50 (0.31, 7.62)	0.61	0.44	0.55	0.43
Open defecation (< 5 year old)	0.22 (0.03, 1.10)	0.09	-0.95	0.52	0.07
Open defecation (adult)	0.29 (0.04, 1.62)	0.18	-0.62	0.58	0.30
Any public toilet use (adult)	1.47 (0.44, 4.96)	0.53	0.33	0.42	0.43
High public toilet use (>10x/mo., adult)	2.38 (0.36, 17.0)	0.36	0.75	0.64	0.24
c) Sanitation (cluster-level)					
Cluster of high HH toilet coverage	3.32 (0.62, 21.2)	0.17	1.47	0.57	0.01
Cluster of low HH toilet coverage	1.05 (0.30, 3.68)	0.94	-0.51	0.43	0.24
Cluster of high coverage of poor FSM	2.29 (0.31, 17.8)	0.41	0.98	0.67	0.15

[a]Odds ratios for detection and 95% confidence intervals presented. Models are adjusted for neighborhood and hygiene status ("good" or "poor", as discussed previously). [b]Concentration differences are in $\log_{10}$genome equivalent copies/100mL; [c]Standard error. [d]Estimated relative to all other households, including those with toilets with other associated FSM and those without toilets. [e]Estimated relative to households without a toilet or those with "other" FSM practices.

Table 5: Any enteric pathogen detection in child stool[a] by household- and cluster-level attributes[b]

	Prevalence ratio (95% CI)	p-value
a) Household-level		
Household toilet	1.57 (0.45, 5.49)	0.48
Toilet excreta contained onsite[c]	0.45 (0.14, 1.43)	0.17
Toilet discharges to drain[c]	3.78 (1.01, 14.2)	0.05
Toilet excreta contained onsite[d]	0.71 (0.19, 2.64)	0.61
Toilet discharges to drain[d]	3.24 (0.76, 13.8)	0.11
Open defecation (< 5 year old)	0.38 (0.10, 1.50)	0.17
Open defecation (adult)	0.83 (0.21, 3.32)	0.79
Any public toilet use (adult)	1.50 (0.54, 4.20)	0.44
High public toilet use (>10 times per month, adult)	0.78 (0.16, 3.74)	0.76
b) Cluster-level[d]		
High cluster of household toilets	0.75 (0.17, 3.33)	0.71
Low cluster of household toilets	0.73 (0.26, 2.09)	0.56
High cluster of household toilets discharging to drain	2.55 (0.43, 15.1)	0.30

[a]Pathogens detected in stool included astrovirus, *Campylobacter* spp., *Entamoeba histolytica*, *Giardia* spp., GII norovirus, and pathogenic *E. coli*. [b]Models with prevalence ratios for detection of any enteric pathogen in stool specimens presented, adjusted for neighborhood and hygiene status ("good" or "poor", as discussed previously). [c]Estimated relative to all other households, including those with toilets with other associated FSM and those without a toilet. [d]Estimated relative to households without a toilet or those with "other" FSM practices.

Figure 1a: Sanitation coverage and clustering in Chinnallapuram

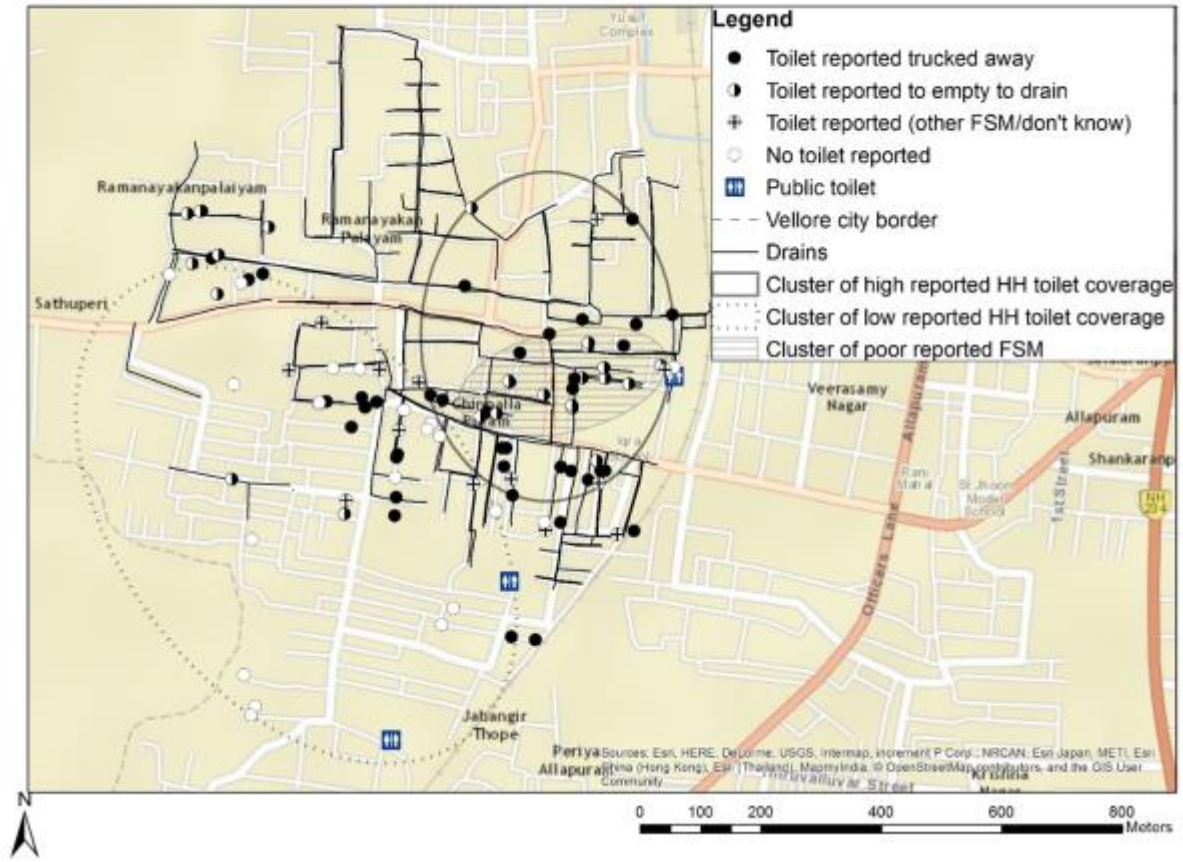


Figure 1a: Sanitation coverage and clustering in Chinnallapuram.

Figure 1b: Sanitation coverage and clustering in Old Town

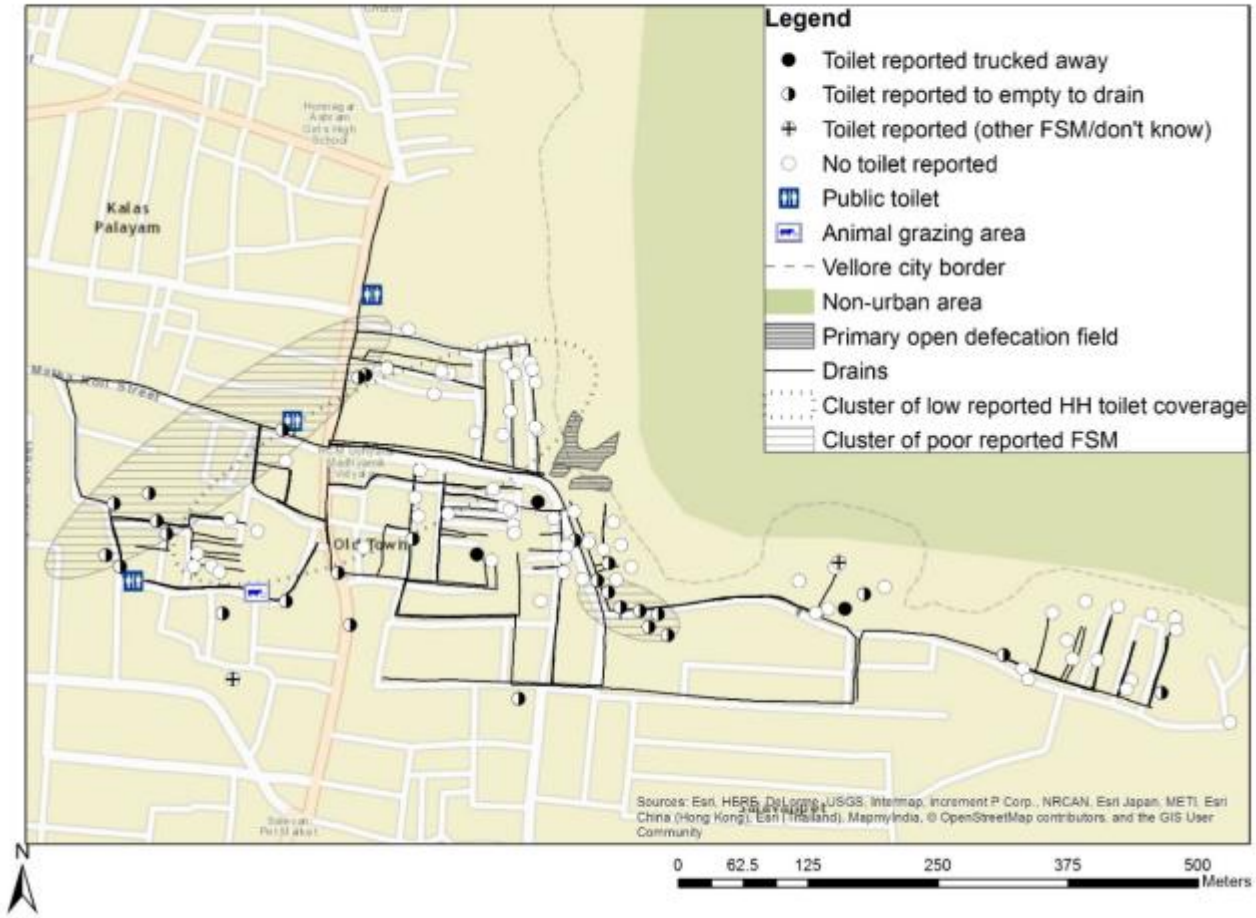


Figure 1b: Sanitation coverage and clustering in Old Town

Supplemental Information

Table S1: Analysis of E. coli concentrations in child hand rinses by household- and cluster-level sanitation and FSM[a]

	β[b]	Standard Error	p-value
a) Household-level			
Household toilet	-1.54	0.77	0.05
Neighborhood: Old Town	0.50	0.31	0.11
Poor hygiene	-1.63	0.77	0.04
Household toilet/Poor hygiene	2.03	0.85	0.02
Toilet contains excreta onsite[c]	-1.19	0.56	0.04
Toilet discharges to drain[c]	0.45	0.50	0.36
Neighborhood: Old Town	0.43	0.29	0.14
Poor hygiene	-0.29	0.47	0.54
Toilet contains excreta onsite/Poor hygiene	2.14	0.72	<0.01
Toilet discharges to drain/Poor hygiene	-0.87	0.80	0.28
Open defecation (< 5 year old)	0.52	0.39	0.19
Neighborhood: Old Town	0.39	0.31	0.21
Poor hygiene	-0.36	0.38	0.35
Open defecation (adult)	0.34	0.43	0.44
Neighborhood: Old Town	0.38	0.34	0.28
Poor hygiene	-0.26	0.41	0.53
Any public toilet use (adult)	-0.04	0.31	0.91
Neighborhood: Old Town	0.50	0.31	0.11
Poor hygiene	-0.04	0.30	0.89
High public toilet use (>10 times per month, adult)	-0.01	0.48	0.99
Neighborhood: Old Town	0.51	0.30	0.10
Poor hygiene	-0.05	0.31	0.89
b) Most likely clusters			
High cluster of household toilet coverage	-0.44	0.44	0.32
Neighborhood: Old Town	0.35	0.34	0.31
Poor hygiene	-0.08	0.30	0.78
Low cluster of household toilet coverage	-0.13	0.32	0.69
Neighborhood: Old Town	0.48	0.31	0.13
Poor hygiene	-0.02	0.31	0.96
High cluster of household toilets discharging to drain	0.05	0.51	0.93
Neighborhood: Old Town	0.51	0.30	0.10
Poor hygiene	-0.03	0.33	0.92

[a]Fecal sludge management. Multivariate models are presented for each sanitation variable, adjusting for neighborhood and hygiene status ("poor" or "good", as discussed previously), and are separated by a blank row in the table. In all models, interaction terms between the sanitation variable and neighborhood or hygiene were tested and are indicated with a "/" if significant at $\alpha = 0.10$. N = 50 samples. [b]Estimates are in $\log_{10}$coliform-forming units (CFU) per pair of hands. [c]Reference population is households without a toilet or those with "other" FSM practices.

Table S2: Analysis of E. coli concentrations in household swabs by household- and cluster-level sanitation and FSM[a]

	β[b]	Standard Error	p-value
a) Household-level			
Household toilet	-2.06	0.82	0.02
Neighborhood: Old Town	-0.79	0.45	0.09
Poor hygiene	-1.69	0.69	0.02
Household toilet/Old Town	1.53	0.57	0.01
Household toilet/Poor hygiene	2.19	0.77	0.01
Toilet contains excreta onsite[c]	-0.52	0.59	0.38
Toilet discharges to drain[c]	0.05	0.53	0.92
Neighborhood: Old Town	0.23	0.31	0.46
Poor hygiene	-0.67	0.50	0.19
Toilet contains excreta onsite/Poor hygiene	1.42	0.76	0.07
Toilet discharges to drain/Poor hygiene	0.75	0.85	0.39
Open defecation (< 5 year old)	0.24	0.37	0.52
Neighborhood: Old Town	0.00	0.30	0.99
Poor hygiene	-0.42	0.36	0.25
Open defecation (adult)	0.10	0.49	0.84
Neighborhood: Old Town	0.77	0.42	0.07
Poor hygiene	0.10	0.37	0.79
Open defecation (adult)/Neighborhood: Old Town	-1.15	0.59	0.06
Any public toilet use (adult)	0.41	0.29	0.16
Neighborhood: Old Town	0.14	0.28	0.64
Poor hygiene	-0.32	0.28	0.26
High public toilet use (>10 times per month, adult)	0.52	0.45	0.25
Neighborhood: Old Town	0.01	0.28	0.97
Poor hygiene	-0.35	0.29	0.23
b) Most likely clusters			
High cluster of household toilet coverage	-0.27	0.42	0.52
Neighborhood: Old Town	-0.04	0.32	0.89
Poor hygiene	-0.30	0.29	0.30
Low cluster of household toilet coverage	0.43	0.39	0.27
Neighborhood: Old Town	0.60	0.37	0.11
Poor hygiene	-0.17	0.28	0.54
Low cluster of household toilet coverage/Old Town	-1.38	0.56	0.02
High cluster of household toilets discharging to drain	0.42	0.48	0.39
Neighborhood: Old Town	0.03	0.28	0.93
Poor hygiene	-0.17	0.31	0.59

[a]Fecal sludge management. Multivariate models are presented for each sanitation variable, adjusting for neighborhood and hygiene status ("poor" or "good", as discussed previously), and are separated by a blank row in the table. In all models, interaction terms between the sanitation variable and neighborhood or hygiene were tested and are indicated with a "/" if significant at $\alpha = 0.10$. N = 50 samples. [b]Estimates are in $\log_{10}$coliform-forming units (CFU) per 125 cm^2.
[c]Reference population is households without a toilet or those with "other" FSM practices.

Table S3: Analysis of E. coli concentrations in sentinel object rinses by household- and cluster-level sanitation and FSM[a]

	β[b]	Standard Error	p-value
a) Household-level			
Household toilet	0.06	0.30	0.83
Neighborhood: Old Town	0.18	0.24	0.46
Poor hygiene	0.23	0.28	0.41
Toilet contains excreta onsite[c]	0.03	0.30	0.91
Toilet discharges to drain[c]	0.16	0.30	0.60
Neighborhood: Old Town	0.16	0.24	0.49
Poor hygiene	0.26	0.26	0.32
Open defecation (< 5 year old)	0.37	0.29	0.20
Neighborhood: Old Town	0.08	0.23	0.74
Poor hygiene	-0.03	0.28	0.93
Open defecation (adult)	-0.14	0.32	0.66
Neighborhood: Old Town	0.21	0.26	0.41
Poor hygiene	0.29	0.30	0.34
Any public toilet use (adult)	-0.14	0.23	0.54
Neighborhood: Old Town	0.13	0.23	0.56
Poor hygiene	0.22	0.23	0.34
High public toilet use (>10 times per month, adult)	-0.08	0.35	0.82
Neighborhood: Old Town	0.17	0.23	0.47
Poor hygiene	0.21	0.23	0.37
b) Most likely clusters			
High cluster of household toilet coverage	-0.08	0.33	0.80
Neighborhood: Old Town	0.13	0.25	0.61
Poor hygiene	0.19	0.23	0.40
Low cluster of household toilet coverage	-0.01	0.24	0.97
Neighborhood: Old Town	0.16	0.23	0.50
Poor hygiene	0.20	0.23	0.40
High cluster of household toilets discharging to drain	-0.14	0.37	0.71
Neighborhood: Old Town	0.17	0.23	0.46
Poor hygiene	0.16	0.24	0.51

[a]Fecal sludge management. Multivariate models are presented for each sanitation variable, adjusting for neighborhood and hygiene status ("poor" or "good", as discussed previously), and are separated by a blank row in the table. In all models, interaction terms between the sanitation variable and neighborhood or hygiene were tested and are indicated with a "/" if significant at $\alpha = 0.10$. N = 49 samples. [b]Estimates are in $\log_{10}$coliform-forming units (CFU) per 100mL.
[c]Reference population is households without a toilet or those with "other" FSM practices.

Risk Factors for Pediatric Enteric Infection in a Low-Income Urban Neighborhood: Examining the Contributions of the Household Environment, Neighborhood Geography, and Exposure Behaviors

David Berendes[1,2*], Juan Leon[2,3], Amy Kirby[2,3], Julie Clennon[2,4], Suraja Raj[2], Habib Yakubu[2], Katharine Robb[2], Arun Kartikeyan[5], Priya Hemavathy[5], Annai Gunasekaran[5], Ben Chirag Ghale[5], J. Senthil Kumar[6], Venkata Raghava Mohan[6], Gagandeep Kang[5], Christine Moe[2,3]

[1]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [2]Center for Global Safe Water, Sanitation, and Hygiene, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [3]Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [4]Department of Biostatistics, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [5]Wellcome Research Laboratory, Christian Medical College, Vellore, India; [6]Department of Community Health, Christian Medical College, Vellore, India

***** Corresponding author
Email: dberen2@emory.edu (DB)

Abstract

Fecal exposures contribute to pediatric enteric infection, and longer-term nutrition, growth, and cognitive impacts beyond diarrhea. In urban environments, exposure to fecal contamination may occur in both the household and public domains; however, the relative importance of the environmental and behavioral components of exposure pathways is poorly understood. This study examined associations between household and neighborhood environments, children's exposure behaviors, and enteric infection, evaluated as general infection and by type of pathogen to compare risk factors. As part of the MAL-ED study, 230 children in a low income, urban, Indian neighborhood provided stool specimens at 14-17 scheduled time points and during diarrheal episodes in the first two years of life that were analyzed for bacterial, parasitic (protozoa and helminths), and viral pathogens. Exposures were assessed for 100 of these children using locational data and interviews with caregivers about household, neighborhood, and behavioral aspects of pediatric exposure to fecal contamination. Study households had poor reported sanitation coverage (33%) and fecal sludge management (82% of household toilets discharged into open drains, which were ubiquitous in the neighborhood). Significant household factors, associated with 44-56% higher risk of pediatric enteric infection, included the presence of older siblings and open defecation by older siblings or adult caregivers. Presence of a household toilet was associated with significantly lower infection risk (OR: 0.73, 95% CI: 0.55-0.97). Neighborhood risk factors included residence in the drain flooding cluster during the northeast monsoon (OR: 2.39, 95% CI: 1.24-4.63), regardless of reported contact with flood water. Compared to general infection, viral infection differed most in risk factors, and included

frequent use of public toilets as a unique risk factor for GII norovirus infection (OR: 2.05, 95% CI: 1.09 – 3.86). Overall, enteric infection may have important infrastructural and geographic risk factors that must be addressed to improve the health of pediatric populations in low-income, urban settings.

Author Summary

Enteric infections, both symptomatic and asymptomatic, are important causes of morbidity and mortality in children through diarrhea, malnutrition, stunting, and poor cognitive outcomes. Though they may be transmitted through the environment by poor water, sanitation, and hygiene conditions, specific risk factors for this transmission in urban areas is not well understood. We examined the household and neighborhood environment and children's behaviors to understand risk factors for enteric infection during the first two years of life. Infections were separated by etiology, using analyses of stool specimens, to compare risk factors. Within the household, when examining 'any' enteric infection (pooled across pathogens), we found that the presence of older siblings and defecation practices of other family members were significant risk factors. Outside the household, aspects of neighborhood geography, like clustering of drain flooding, were significant risk factors, regardless of reported frequency of exposure. Viral infections were most unique in their risk factors, such as use of public toilets associated with norovirus infection risk. Because many of these health risks were affected by environmental neighborhood conditions, and not the child's exposure behaviors alone, reducing risk of enteric infection requires improvement of infrastructure and containment of excreta at the neighborhood level, in addition to behavior change.

Introduction

Despite an estimated 1.7 billion cases of diarrhea annually, most of which are in children, the impact of enteric infections worldwide is underestimated due to high, undetected rates of asymptomatic infection [1,25,66]. Even in the absence of diarrhea, these infections are detrimental to child health, growth, and cognition [189]. While water, sanitation, and hygiene (WASH) conditions are thought to influence enteric infection incidence, few studies have evaluated their impact on both symptomatic and asymptomatic infections [4].

Existing knowledge of associations between WASH-related exposures and transmission of enteric infection is limited by the outcome of interest, detection methods, and differences in study settings.[30,58,142,190] Most WASH studies have targeted reported diarrhea, an outcome that reflects mixed etiologic agents, ignores asymptomatic infection, and is subject to observer and respondent bias, limiting its generalizability to risk of enteric infection[30,58,65,66,191]. In addition to recent focus on WASH—and especially sanitation—in rural settings, there is an urgent need to understand the complexities of urban WASH.[102,103] Over half of the world's population lives in cities, and the urban population is expected to almost double by 2050, adding to the 600 million without access to basic sanitation [103,139].

Urban environments include exposure pathways to fecal contamination in both the household and public domains [69–73,103,130]. When functioning properly, household toilets and fecal sludge management (FSM) contain excreta at the household and neighborhood levels, along the entire sanitation chain, preventing public contact [29,32]. However, poor

handwashing behaviors after defecation, maintenance and cleanliness of the toilet, and FSM may also contaminate both environments (demonstrated in Aim 2) [32,80–83]. In the public domain, open defecation fields, open drains, and flood water all represent potential exposures to fecal contamination from poor FSM [31,32,69–72]. Public toilets may also be a point of exposure to fecal contamination if maintenance and containment are poor [19,109,113,116,192]. From household to city scales, uncooked vegetables and municipal water can also serve as vehicles of fecal contamination [125–127,193].

Within a neighborhood, urban density and geography dictate the management of fecal sludge, with implications for exposure in the public domain. Urban environments are often too dense to construct individual household toilets, forcing residents to either open defecate or use public toilets depending on convenience, distance, location, and other factors [113]. Further, the few household toilets present are often not connected to sewers, for logistical and financial reasons, and instead contain excreta onsite [32,117,183]. Spatial constraints prevent construction of new systems once the old ones are filled, necessitating transport of excreta away from households, by truck or connection to open drains [32,117,183]. The cost and logistics associated with excreta emptying and transport make connections to open drains very common, and most excreta remains untreated [69,70,118,183,194]. Children may easily have contact with drains when playing. Depending on the local geography, these drains may also routinely flood, leading to clustering of fecal exposures [122,156,195].

Pediatric exposure to environmental fecal contamination in urban settings may be direct (personal contact) or indirect (contact with other household members or fomites),

and varies by dose and frequency. Direct contact with fecal contamination in open drains, for example, can result in exposure to very high concentrations of fecal contamination from a single event [69,70,86]. Conversely, despite lower concentrations of fecal contamination, consumption of municipal water may yield the same cumulative dose through higher frequency of exposure and direct ingestion [68]. Contact with other household members or surfaces in the household or public environment yield indirect, frequent exposures with varying doses based on exposure behavior and hygiene practices [79,130,191,196].

Though enteric infections generally have fecal-oral transmission, they are caused by diverse etiologic agents—bacteria, helminths, protozoa, and viruses—that vary in their associations with WASH conditions [9,65]. Bacteria have historically represented the classical waterborne transmission of disease on which the WASH sector was founded [9]. Helminths have greater variation in their associations with WASH, however, as they have a more complex life cycle that requires development in both the human host and the environment, and may be transmitted through dermal contact or via ingestion [10,11,197]. Protozoa and enteric viruses survive for long periods outside the host and have low median infectious doses (ID50), facilitating transmission from person-to-person and through environmental surfaces [9,12,198,199].

Given the numerous environmental pathways for exposure to fecal contamination in urban settings and the diversity of etiologic agents causing enteric infection, there is a need to compare pediatric exposure pathways across the public and private domains for different groups of enteric pathogens. The goal of this study was to examine the

associations between the urban environment (including the household and neighborhood conditions and location), children's exposure behaviors, and enteric infection for different types of enteric pathogens in a cohort of under-two year olds in a low-income, urban neighborhood in Vellore, India. Examining the environmental and behavioral aspects of pediatric exposure pathways associated with enteric infection by etiologic agent will aid in understanding how fecal contamination in the public and private domains affects health risk for children in low-income, urban settings.

Methods

Data sources

This study was conducted using data from three sources in the Old Town neighborhood of Vellore, India: 1) The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and Consequences for Child Health and Development Project (MAL-ED study); 2) a SaniPath Tool deployment; and 3) surveys of public toilets [136,172,189]. In Vellore, the MAL-ED study was conducted by the Christian Medical College and Hospital, Vellore (CMC). Enrollment took place from March 2010 – February 2012, with data collection ending in February 2014. The SaniPath Tool deployment and public toilet surveys were conducted by Emory University in collaboration with CMC. Household surveys took place in February – March 2014, with the public toilet surveys in February 2015.

Study site

Yearly, Vellore has a dry season (January – May), a southwest monsoon (June – September), and a northeast monsoon (October – December) [136].The Old Town

neighborhood of Vellore, India, is a small, low-income, urban area with high population density (approximately 42,000/km^2), poor sanitation, and high burden of enteric disease (Aim 2)[136]. CMC has a longstanding relationship with the community, including mapping conducted during previous studies.

Ethical approval

Prior to subject recruitment for the MAL-ED study in Vellore, ethical approval was obtained from the Christian Medical College Institutional Review Board (IRB) [136]. Approval was obtained from the Emory University IRB and the Christian Medical College IRB prior to SaniPath Tool deployment and public toilet surveys. Informed consent was obtained prior to survey administration at the location of the interview.

Stool specimen collection and testing

Stool specimens were collected from children in accordance with MAL-ED study protocols [136,200]. Specimens were collected monthly over the child's first year of life, and then every 2-3 months over the next year (defined as "routine stool collection"). Caregivers were asked to submit specimens from any diarrheal events during the follow-up period (classified as "diarrheal stool collection"). All stool specimens were tested for a range of bacteria (*Salmonella, Shigella, Vibrio, Yersinia, Aeromonas, Plesiomonas,* diarrheagenic *E. coli* (Shiga toxin-producing *E. coli*, enterotoxigenic *E. coli*, enteropathogenic *E. coli*, enteroinvasive *E. coli*, and enteroaggregative *E. coli*), *Campylobacter*), protozoa (*Balantidium coli, Cryptosporidium, Chilomastix mesnili, Cyclospora, Entamoeba histolytica, Giardia lamblia, Endolimax nana, Iodamoeba butschlii, Isospora*), helminths (*Hymenolepis nana, Strongyloides stercoralis, Ascaris*

lumbricoides, *Taenia*, *Trichuris trichiura*, *Schistosoma*, *Enterobius vermicularis*, *Hymenolepis diminuta*, hookworm species), and viruses (genogroup I and II norovirus, astrovirus, rotavirus, adenovirus) by culture, microscopy, immunoassay, and PCR as described previously [25,65]. Pathogen detection was grouped into bacteria, parasites (protozoa and helminths), and viruses for analysis. Enteroaggregative *E. coli* (EAEC) and genogroup II norovirus were previously selected for analyses of environmental samples in households and open drains as primarily human-specific bacterial and viral pathogens, respectively, that had high prevalence in the MAL-ED study population in Vellore. Genogroup I norovirus was selected as a primarily human-specific pathogen with a low prevalence in the study population for comparison (Aim 2). EAEC and GI and GII norovirus were examined as specific infections in this study.

Transect walks and household survey data collection: SaniPath Tool

Transect walks with a community leader were conducted in the study neighborhood. Locations of potential fecal exposures in the public domain, including public toilets and animal grazing areas, were documented using Global Positioning System (GPS) points. Locations of other potential public fecal exposure locations, including open drains (both location and direction of flow) and the primary open defecation field, were recorded by CMC in previous studies [136].

Survey methods are provided in detail in Aim 2. Briefly, 100 households were surveyed within the neighborhood, 25 of which were selected based on results of a CMC hygiene survey completed 1 month prior to SaniPath data collection to ensure diversity of household hygiene practices [175]. The other 75 households were chosen randomly across the MAL-ED study areas within the neighborhood [136]. The survey assessed household

characteristics and sanitation (e.g. demographics, presence of a toilet, FSM), neighborhood characteristics near the household (e.g. flooding of the drain outside the house ("drain flooding") or flooding within the house itself ("house flooding")), defecation practices (e.g. open defecation, use of public toilets), and other exposure behaviors of the child (e.g. frequency of contact with drains). Contact with fecal exposures was divided into frequency categories and presented "per month" or "per week", depending on the exposure. The target respondent was the female head of household. GPS points were collected at the time of household survey.

Survey data collection: public toilet surveys

Surveys at public toilets were conducted with the current, onsite manager regarding toilet facilities, use, and maintenance, followed by observations of the public toilet location and its stalls. Three of four public toilets were open and functional at the time of the MAL-ED study and household surveys. Two of those three were open and functional at the time of survey.

Analyses

Map construction was completed in ArcMap version 10.1 (ESRI, Redlands, CA, USA). Kulldorff's Bernoulli and Normal spatial and space-time scans were used to evaluate most-likely clustering of any enteric infection, as well as by type of infection, in space and space-time using SaTScan version 9.4 [52]. Kulldorff's Bernoulli scan evaluates point data with binary values to assess the distribution of '0' and '1' values in space and space-time for non-random clustering. Kulldorff's Normal scan evaluates point data with continuous values in a similar manner. "Any enteric infection" was defined as the

detection of at least one pathogen in a child's stool at a given sampling event. Kulldorff's Bernoulli spatial scan was also used to evaluate most-likely clustering of reported drain and house flooding in space.

Aspatial analyses were conducted in R version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria) using standard packages and the "lme4" package for mixed effects models [179]. Mixed effects Poisson regression was used to estimate associations between pathogen detection and diarrheal stool incidence. Mixed effects logistic regression with a random effect for the child sampled was used to estimate bivariate associations between household and neighborhood environments or exposures, and the outcome: incidence of infection (detection of pathogens in children's stool). All variables significant at $p < 0.10$ in bivariate analyses were tested in multivariate models, including interactions terms, and the final model was selected using the Akaike Information Criterion (AIC) score with the goal of providing the best fit to the data [201]. Frequency categories for a child's reported exposures were modeled in two ways: any contact vs. no contact, and highest group of contact vs. all others. An α of 0.05 was used for all tests of significance. Residuals from final models were tested for spatial autocorrelation using the Global Moran's I in ArcMap 10.1.

Results

Household and neighborhood environments and exposure behaviors in the study neighborhood

Household and neighborhood environments and child exposure behaviors were quantified through household surveys with the adult caregiver (Table 1). Study children in most households were reported to have older siblings. Reported household sanitation

coverage was poor (33%), with poor fecal sludge management ("poor FSM," defined as a toilet discharging directly into an open drain) present in 82% of study households with a toilet. Only 3 households reported having a toilet with good FSM (defined as a toilet discharging into a tank containing excreta onsite), and only one reported the toilet filling at least once yearly, indicating regular exhausting of fecal sludge (data not shown). Drains were ubiquitous in the neighborhood, and over half (58%) of respondents reported drains flooding in front of the household. While more than 80% of respondents reported their children had monthly contact with drains and floodwater, frequent contact (>10 contacts per month) was more limited (15% for drain contact and 26% for flood contact). Open defecation was common across all age groups, and approximately half of all households reported use of public toilets on a monthly basis, though only 13% reported frequent use (>10 times per month). Weekly consumption of uncooked produce and daily consumption of municipal water were prevalent among study children. About one-third of families reported regularly treating drinking water in some way (boiling, chlorine, or other methods).

Potential fecal exposure points and neighborhood drains were mapped from GPS data and most-likely spatial clustering of reported flooding was evaluated by Kulldorff's Bernoulli spatial scan. Neighborhood drains generally flowed from east/northeast to west/southwest (Fig 1). Significant drain flooding and house flooding clusters (24 and 7 study households, respectively) were detected downstream of the open defecation field, one of the highest elevation points present. Within each cluster, flooding was reported in 100% of study households.

Public toilet maintenance and conditions

Observations and interviews with managers at the two functioning public toilets were used to evaluate their conditions, maintenance, and use (data not shown). Each respondent had begun working at his public toilet in 2014. Toilets had multiple, separate stalls for female and male users and were estimated to serve about 100 users per day. One manager reported rarely cleaning the toilet (less than once every two days), while the other reported cleaning the toilet 3-4 times per day. Visible excrement and a noticeable smell were observed within stalls of both toilets. Both managers relied on community leaders and government officials to coordinate exhausting of the fecal sludge in tanks under the toilet. Each tank had an outflow pipe discharging into open drains.

Distribution of pathogenic organisms in children's stool

Detection of enteropathogen-positive samples per child was compared by pathogen and type of stool collected (routine vs. diarrheal), and between children in SaniPath-specific and all MAL-ED study households, to assess the overall burden of infection and the representativeness of study households in SaniPath to the overall MAL-ED study. Chi-squared tests of proportions were used to compare percentages between groups. Approximately 67% of routine and 80% of diarrhea stool samples were positive for at least one pathogen (Table 2). Bacteria were detected in a higher proportion of samples than parasites or viruses across both types of stool collected. Detection of enteropathogens in stool from children in SaniPath households compared to children from all MAL-ED study households did not differ significantly.

Pathogens associated with diarrheal stool

Associations between pathogen detection and type of stool collected (routine vs. diarrheal) were evaluated using mixed effects Poisson regression (data not shown). Viruses, both in general and specifically GI and GII norovirus, were detected significantly more often in diarrheal stool than in routine stool collection (odds ratio (OR) for viral detection: 4.37, 95% confidence interval (CI): 3.39 – 5.63; GI norovirus OR: 3.13, 95% CI: 1.87 – 5.25; GII norovirus OR: 5.20, 95% CI: 3.82 – 7.09). Bacteria and parasites were detected more often in diarrheal stool compared to routine stool collection, but these differences were not significant (OR for bacterial detection: 1.17, 95% CI: 0.90 – 1.53; OR for parasitic detection: 1.21, 95% CI: 0.88 – 1.66). EAEC detection did not differ between type of stool specimen (OR: 0.90, 95% CI: 0.58 – 1.40). Because of high prevalence (Table 2) and significant association with diarrheal stool, GII norovirus was specifically evaluated in further analyses, in addition to pathogen groups.

Spatial and temporal clustering of enteric infections

To assess clustering of enteric infections in space and space-time, Kulldorff's Bernoulli and Normal spatial and space-time scans were conducted on enteric infection data from all MAL-ED study households in the SaniPath study area. Space-time clustering was assessed using data from 2011 and 2012, when the MAL-ED study population was highest. Each 12-month calendar year was analyzed separately to minimize dropout within the time period due to completion of follow-up. When evaluated in space only, presence of enteric infection at the first month's follow-up and the number of infections during the first year of follow-up were not significantly spatially clustered, regardless of whether infections were pooled (any enteric infection) or evaluated by type.

When evaluated in space and time, significant clustering of incidence of any enteric infection was not detected, but incidences of bacterial, parasitic, and viral infections were significantly clustered in 2011 or 2012 (Figures 2 and 3). Significant space-time clustering of bacterial infection incidence was observed at the end of the dry season (May, 2011, Fig 2) and during the southwest monsoon (July – September, 2012, Fig 3). Clustering during the southwest monsoon surrounded the open defecation field. In 2011, parasitic infection incidence was clustered during the monsoons, with significant clustering of low incidence during the dry season (Fig 2). Multiple, significant clusters of viral infection incidence were observed during the dry season in 2012 (Fig 3). Both bacterial and viral infections showed acute space-time clusters (within a single month period).

Seasonality of enteric infection risk was evaluated by mixed effects logistic regression on all stool samples, controlling for type of stool specimen. Risk of bacterial or parasitic infection did not differ significantly by season ($p > 0.05$, data not shown). Viral infection risk was significantly lower during the northeast monsoon (October – December) compared to the dry season (OR: 0.60, 95% CI: 0.38 – 0.95, $p = 0.04$). GII norovirus infection risk did not vary significantly by season ($p > 0.05$).

Effect of household and neighborhood environments on enteric infection risk, by pathogen type

Associations between household or neighborhood environments and risk of enteric infection were evaluated for any enteric infection, as well as by type of pathogen, using mixed effects logistic regression, controlling for season and type of stool specimen (Tables 3 and 4). Risk of infection with any pathogen ("any enteric infection") was

significantly higher for children in households with older siblings, but was significantly lower for those in households with toilets (Table 3a). Though overall, household toilets with poor FSM were associated with lower (though not significant, $p = 0.06$) risk of enteric infection (OR: 0.75, 95% CI: 0.55-1.01), this association varied by season. During the dry season, presence of a household toilet was associated with significantly lower risk of enteric infection (OR: 0.61, $p = 0.02$); however, during the northeast monsoon season, this association was no longer protective (OR: 1.10, p-value for interaction term: 0.06). Associations between the presence of household toilets with good FSM and pediatric enteric infection were not tested because few households with toilets with good FSM (3) were present.

Several risk factors were significant for specific types of infection, but not enteric infection overall. Risk of viral infections increased with increasing number of people in the child's household ($p = 0.03$). Risk of parasitic infection (Table 3) was higher for children in households with older siblings ($p = 0.01$). Risk of bacterial infection was 29% lower ($p < 0.01$) and risk of parasitic infection was 39% lower (though not significant, $p = 0.06$) for children in households with a toilet, but risk of viral infection did not differ by household toilet status.

When evaluated aspatially, neighborhood conditions and flooding were not significantly associated with enteric infection risk in study children (Table 3b).

Spatially, risk of enteric infection differed between children in drain flooding clusters and those in the rest of the neighborhood by season (Table 4). During the northeast monsoon, risk of any enteric infection was significantly higher for children living in the drain flooding cluster compared to those not in the cluster. Risk of bacterial

infection showed a similar association. During the southwest monsoon, risk of viral infection was significantly lower in this cluster compared to the rest of the neighborhood. Risk of enteric infection did not differ significantly between children inside and outside of house flooding clusters, regardless of season.

Effect of exposure behaviors on enteric infection risk, by pathogen type

Relationships between children's exposure behaviors and enteric infection risk were evaluated by mixed effects logistic regression, controlling for season and type of stool collected (Table 5). Children reporting any contact with floodwater, or those living in households with older siblings or adults who open defecated, had significantly higher risk of any enteric infection compared to the rest of the neighborhood. Children in households reporting treating their drinking water had significantly lower risk, compared to the rest of the neighborhood. Children reported to have any monthly contact with drain water had significantly higher risk of parasitic infection, while those reported to have any monthly contact with floodwater had significantly higher risk of bacterial infection. Risk of bacterial infections was significantly higher in households where others open defecated. Risk of viral infection was not associated with open defecation practices at the household. GII norovirus infection was significantly higher among children reported to have high (>10) monthly use of public toilets (OR: 2.05, 95% CI: 1.09 – 3.86). Children in households that reported treating their drinking water had lower risk of infection across all types of pathogens, with significantly lower bacterial and parasitic infection risks. Consumption of municipal water or produce were not significantly associated with enteric infection risk.

Multivariate modeling of enteric infection, by pathogen type

To determine the relative contributions of household and neighborhood conditions and exposure behaviors to pediatric enteric infection risk by pathogen, multivariate mixed effects logistic regression models were constructed, adjusting for season and type of stool collected (Table 6). Model residuals did not show evidence of spatial autocorrelation (all p-values for Global Moran's I test > 0.05). Bacterial infection risk in children was significantly associated with residence in the drain flooding cluster during the northeast monsoon, open defecation by the adult caregiver, and any reported monthly flood contact by the child (Table 6a). Presence of a household toilet was almost interchangeable with adult open defecation with regard to the fit of the model (OR for presence of a household toilet: 0.71, 95% CI: 0.55 – 0.91, model not shown). Increased risk of viral infection in children was observed with increasing size of the household population (Table 6b). Children in households in the cluster of drain flooding during the southwest monsoon and those in households that reported treating their drinking water had lower risk of viral infection, though drinking water treatment was not significant in the final model. In the final multivariate model for GII norovirus infection (data not shown), risk was significantly higher for children with frequent use of public toilets (OR: 1.94 95% CI: 1.05-3.59), but significantly lower when treatment of drinking water was reported (OR: 0.44, 95% CI: 0.23-0.85). Because the models failed to converge when testing multiple parameters, a final model for parasitic infection is not presented.

Discussion

The goal of this study was to examine the contributions of the household and neighborhood environments and exposure behaviors to pediatric enteric infection risk in a

low-income, urban setting, separating infection by pathogen group to compare risk factors. Household risk factors for any enteric infection, pooled across pathogens, included the presence of older siblings and open defecation practiced by older siblings or adult caregivers, underscoring the importance of other family members in the child's exposure to fecal contamination. Presence of a toilet in the household was a significant protective factor, though associations varied by season. Within the household, water treatment—a potential proxy for household hygiene practices—was associated with significantly lower risk of enteric infection. At the neighborhood-level, given the poor FSM, enteric infection risk was significantly higher in the cluster of reported drain flooding during the northeast monsoon. Consequently, behavioral risk factors included any contact with drain or floodwater. When separated by type of infection, the household and neighborhood-level risk factors associated with viral infections exhibited the largest differences from the pooled infection results. For example, risk of GII norovirus infection was significantly associated with high use of public toilets, a risk factor not observed for bacterial or parasitic infections. The variation in risk factors for pediatric enteric infection underscores the complexity of environmental transmission within urban areas and the importance of limiting fecal exposures in both the private and public domains.

This study is unique in examining urban environmental, infrastructural, and behavioral risk factors for enteric infection in a pediatric cohort, as well as comparing risk factors between bacterial, parasitic, and viral infections. While previous studies have enumerated risk factors for diarrheal disease in urban households and neighborhoods, few have used pathogen-specific approaches [91,117,120,121,130,154,180–182,202–205]. Previous examinations of pathogen-specific environmental risk factors by exposure pathway in

urban settings have been quantitative microbial risk assessments (QMRAs), which have not shown such enteropathogen-specific differences in risk pathways because of the extrapolation of *E. coli* measurements to other pathogens[69,70,73].

Though exposure occurs within the household, older family members are an important connection between young children and fecal contamination in the public domain, particularly where poor household sanitation coverage and high open defecation exist. Beyond contaminating the local environment, those who open defecate have more exposure to fecal contamination, including the feces of others and flies carrying enteric pathogens, than if they had used a household toilet [206]. Households with a toilet, in a similar study area, had significantly lower density of flies and incidence of diarrheal disease among young children, compared to those without a toilet [175]. Elsewhere, presence of an improved household toilet has been associated with lower levels of fecal contamination on hands [81]. Thus, these pathways may explain why the presence of a household toilet was associated with significantly lower risk of enteric infection, despite the study child likely being too young to use it during much of the follow-up period.

While most adult caregivers and older siblings practiced open defecation, siblings may provide more exposure to fecal contamination to the study child than the adults did. Older children who open defecate have been shown to do so closer to the home, display poorer hygiene practices after defecation, and have higher incidences of enteric infection than adults [129,207–210]. They may also have more contact with the child, as sibling care is common in Indian culture [202,210].

Seasonal variation in the association between household toilets discharging to open drains and risk of enteric infection reflects the negative impact of poor

neighborhood- and household-level FSM on the beneficial effects of functional household toilets. Overall, 82% of household toilets in this neighborhood discharged to the drain, contaminating the local environment with fecal pathogens (Aim 2). A previous, sub-neighborhood, cross-sectional study amongst a subset of this population and that of a nearby neighborhood with better household toilet coverage showed children in households connecting their toilets to drains had the highest prevalence of enteric infection, even when compared with households without a toilet (Aim 2). Further, the northeast monsoon was the period of highest flood-associated risk, as demonstrated in Table 4. Due to the direct connection to the neighborhood environment, toilets with poor FSM may have become nonfunctional during this flooding period. Thus, users may have elected to either 1) use the household toilet, in which case their feces was likely not removed from the household and immediate neighborhood environment due to the flooding; or 2) open defecate in a higher elevation area without flooding, potentially exposing them to more fecal contamination both at the open defecation site and through contact with flood water. Within the study neighborhood, significant spatial clusters of high levels of reported contact with drain and flood water were present and overlapped with clusters of reported drain and house flooding. However, in a nearby neighborhood with better FSM, no significant spatial clusters of high reported contact with drain or flood water were detected despite similar, significant clusters of reported drain and house flooding.[211]

Within the household, reductions in enteric infection risk associated with reported water treatment underscore the importance of exposure from drinking water, but may also be a proxy for household hygiene behaviors. Contaminated drinking water is an

important fecal exposure, and household water treatment has been associated with 33% reductions in risk of diarrheal disease, consistent with observations of reduced enteric infection in this study [142,212]. However, because risk was reduced across all pathogens, including some that are not primarily waterborne, reported water treatment may also have been representative of improved socioeconomic status and ability to maintain good household hygiene practices [12,15,191,209,213].

At the community-level, elevated enteric infection risk within the drain flooding cluster during the northeast monsoon demonstrated the effects of low sanitation coverage and poor FSM. Poor coverage of household toilets resulted in more frequent open defecation. While some open defecation was reported in drains near households, it was primarily concentrated at a high elevation field where fecal contamination could run off into drains following heavy rains (Aim 2)[214]. Further, most household toilets discharged directly into open drains. Previous work in this community has shown that drains receiving household toilet excreta directly have significantly higher levels of fecal contamination compared to the rest of the neighborhood, even households reported practicing open defecation (Aim 2). Even when flooded, the levels of fecal contamination in drains in the flooding cluster, which was downstream of both the open defecation field and many neighborhood drains, were likely high enough to cause infection with a single contact event [69–72,85–87].

Risk from drain and floodwater exposures in this study was not simply due to behaviors that were modifiable. Enteric infection risk was elevated in the drain flooding cluster, independent of reported contact, and did not show a dose-response relationship with drain or floodwater contact. Instead, children reported to have any contact with drain

or floodwater had significantly elevated risk of infection, which contrasts with previous studies in other urban settings that observed exposure risk driven by frequency of drain contact [69,70,86]. In the current context, exposure associated with these pathways must be managed at the neighborhood level through interventions to both improve FSM and reduce flooding. Simply reducing contact by providing fences or covers for open drains is unlikely to be effective in reducing risk from open drains [69,70].

When compared to risk factors for any enteric infection, those associated with viral infections were relatively unique because of the importance of direct contact with people and fomites, reflecting the biology and epidemiology of enteric viruses. Given their low infectious doses and high shedding in both symptomatic and asymptomatic individuals, virus transmission is frequently via person-to-person contact [12,15,18,215]. Large household populations in dense urban areas facilitate more person-to-person contact and transmission of infections or fecal exposures between family members within the household [103].

Although hand hygiene measures have been specifically implemented to interrupt transmission of viruses, cleaning of environmental surfaces, such as those in public toilets, to reduce exposure may also be important [19,191,216]. Viruses can survive and remain infectious for months in groundwater and days to weeks on environmental surfaces [12,17]. Environmental transmission of viruses through hands, fomites, or other surfaces can require very few particles (18 for norovirus), underscoring the importance of regular cleaning of public toilets and handwashing after defecation [18]. Visible feces in stalls at the public toilets suggested cleaning was not regularly practiced or not comprehensive when practiced. Further, handwashing basins and soap were absent from

all public toilets surveyed in this neighborhood, making users less likely to practice good hygiene after defecation [209,217–219]. Given the poor maintenance, lack of hygiene facilities, and high volume of users, public toilets may provide excellent sources of viral infection, even with infrequent use [17,19,85,89].

This study has some notable strengths and limitations in its measurement of outcomes and exposures. The use of stool specimens collected regularly to assess both symptomatic and asymptomatic enteric infections provides greater detail in understanding relationships with environmental risk factors than previous outcome measures [25,30,58,66,189]. However, detection of enteropathogens in diarrheal and non-diarrheal stool does not indicate current illness. Pathogens can continue to be shed in stool for weeks or months after illness resolves [12,215]. For exposure, behaviors may be overestimated because household surveys were conducted when the children were 2-5 years old, while infection outcomes were measured during the first two years of life [124]. Static assumptions had to be made about the household and neighborhood conditions. Despite the broad range of public domain exposures assessed, zoonotic exposures, a significant contributor to diarrheal diseases even in urban areas, were not measured [220]. Because observation was not logistically feasible and self-report had large potential for response bias, measures of household protective behaviors are limited to responses about water treatment [186]. Finally, sample size, and especially the small number of households with toilets with good FSM, limited our ability to make inferences about the joint effects of FSM and toilets on enteric infection risk.

Mitigation of health risk requires understanding the urban geography, fecal contamination levels, and frequency of contact associated with exposure to fecal

contamination, and is not possible through household WASH interventions alone. Future sanitation and hygiene interventions within the household must be complemented by efforts to reduce fecal contamination within the neighborhood, especially where neighborhood-level FSM is poor. The complexity of urban settings necessitates multiple approaches to reduce environmental fecal contamination exposures in the public domain [34,35,103,138]. Studies need to examine not only public domain interventions and health effects, but also how fecal exposures are influenced by behavioral, infrastructural, and geographic characteristics in order to identify specific intervention points [46]. Where possible, enteric infection outcomes encompassing both symptomatic and asymptomatic infections should be used to better understand pathogen-specific (or pathogen group-specific) environmental exposures and transmission.

Pediatric enteric infection has household- and neighborhood-level risk factors that include the defecation practices of other family members and open drain flooding. Though frequency of contact is a component of exposure, neighborhood-level clustering of drain flooding, combining urban geography with poor FSM, demonstrates the importance of neighborhood infrastructure in mitigating fecal exposures and enteric infection risk. Though viral infections may not be mitigated by traditional WASH strategies, they may have important environmental intervention points as well that are necessary to reduce their especially large burden associated with symptomatic enteric infection. Approaches accounting for the interconnectedness of the public and private domains are essential to designing effective interventions to improve the health of children in urban environments.

Acknowledgements

The authors would like to thank the field and lab staff at CMC in Vellore, including Sheela Roy, for their effort on this study and all aspects of the SaniPath study in Vellore, India.

Tables and Figures

Table 1: Reported household/neighborhood conditions and exposure behaviors[a]

Household conditions	Count (%) or Mean (SD)
Demographics	
5-12 year old present	62 (62.0)
Average number of people per household	6.4 (2.2)
Sanitation	
Household toilet	33 (33.0)
With poor FSM[b] (discharging to drain)	27 (81.8)
With good FSM[b] (excreta contained onsite)	3 (9.1)
Other/Don't know	3 (9.1)
Neighborhood conditions	
Open drain in front of household	96 (96.0)
Flooding	
Drain floods	57 (57.6)
House floods	23 (23.0)
Exposure behaviors[c]	
Drain contact	
Any	86 (86.0)
>10 times per month	15 (15.0)
Flood water contact	
Any	82 (82.0)
>10 times per month	26 (26.0)
Open defecation	
<5 year olds	80 (80.0)
5-12 year olds	45 (78.9)
Adult	68 (68.0)
Public toilet use	
Any	46 (46.0)
>10 times per month	13 (13.0)
Municipal water consumption	
Any	84 (84.0)
Daily	69 (69.0)
Drinking water treatment at the household	32 (32.0)
Raw produce consumption	
Any	70 (70.0)
Daily	17 (17.0)

[a]Data from household survey (n = 100 households); [b]Fecal sludge management (describing the containment of excreta along the entire sanitation chain, from toilet to treatment [32,221]); [c]Where an age group is not indicated, the exposure behavior represents that of the study child, as reported by the adult respondent

Table 2: Detection of pathogens in children's stool in SaniPath and all MAL-ED households from 2010-2014

	Percent of child's stool specimens collected that were positive[a]					
	Stool collected from children in SaniPath households (n = 100)[b]			Stool from children in all MAL-ED study households (n = 230)		
Single infections	Routine[c]	Diarrheal[d]	All stool collected	Routine[c]	Diarrheal[d]	All stool collected
Any pathogen[e]	67.2	82.6	69.2	67.6	79.4	69.2
Bacterial infection	60.6	64.5	60.9	59.9	61.0	59.8
EAEC	9.8	8.1	9.6	9.5	7.6	9.3
Parasitic infection[f]	16.1	18.3	16.3	17.9	21.3	18.3
Viral infection	7.9	39.4	12.5	8.6	35.9	12.8
GI norovirus	1.0	4.9	1.7	1.0	5.0	1.7
GII norovirus	1.5	19.5	4.3	2.3	17.6	4.7
Combined infections						
Viral + bacterial	5.6	26.4	8.6	6.4	23.7	9.0
Viral + parasite	1.4	6.2	2.2	1.8	6.6	2.6
Bacterial + parasite	11.6	12.2	11.7	12.2	13.4	12.3
Bacterial, viral, and parasite	1.2	4.8	1.8	1.6	4.8	2.2

[a]Calculated as the average, by child, of the proportion of stool specimens that were positive for a given pathogen; [b]SaniPath households consisted of a subset of all MAL-ED study households that were surveyed for demographics, exposure behaviors, and household and local neighborhood conditions. No significant differences were observed in the detection of pathogens (of any type, including combined infections) in stool from children in SaniPath households compared to children in all MAL-ED study households; [c]Routine stool was collected monthly over the first year of follow-up, then every 2-3 months during the second year of follow-up; [d]Diarrheal stool collected whenever a child had an incidence of diarrhea, as reported by the caregiver; [e]A full list of pathogens tested in stool specimens is available in Houpt et al. 2014; [f]Includes protozoa and soil-transmitted helminths

Table 3: Bivariate relationships between household and neighborhood conditions and pathogen detection in stool collected from children in SaniPath households, 2010-2014[a]

a) Household conditions	Bacteria OR (95% CI)	Parasite OR (95% CI)	Virus OR (95% CI)	Any pathogen OR (95% CI)
Demographics				
Num. people per household	1.01 (0.96, 1.07)	1.11 (1.00, 1.24)	1.09 (1.01, 1.17)†	1.03 (0.96, 1.09)
5-12 year old (YO) present	1.26 (1.00, 1.61)	1.91 (1.17, 3.12)††	1.02 (0.71, 1.46)	1.44 (1.10, 1.89)††
Sanitation				
Household toilet	0.71 (0.56, 0.91)††	0.61 (0.36, 1.02)	1.07 (0.74, 1.54)	0.73 (0.55, 0.97)†
b) Neighborhood conditions				
Drain present in front of HH	1.12 (0.60, 2.08)	3.91 (0.83, 18.5)	0.97 (0.38, 2.46)	1.32 (0.66, 2.63)
Flooding				
Drain flooding	1.24 (0.98, 1.57)	0.99 (0.61, 1.62)	0.81 (0.57, 1.15)	1.26 (0.96, 1.65)
House flooding	1.10 (0.83, 1.46)	1.01 (0.57, 1.81)	0.83 (0.54, 1.27)	1.04 (0.75, 1.44)

[a]All models adjusted for monsoon seasons (relative to dry season) and stool type (routine vs. diarrheal stool); [b]Toilet discharges to an open drain ; [c]Compared to households with other toilet/FSM combinations and households without toilets; [d]Toilet excreta is contained onsite; [e]Compared to households without toilets; †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$

Table 4: Bivariate relationships between spatial clustering of flooding in neighborhood, seasonality, and pathogen detection in stool collected from children in SaniPath households, 2010-2014[a]

	Bacteria OR (95% CI)	Virus OR (95% CI)	Any pathogen OR (95% CI)
Drain flooding cluster			
Year-round[b]	0.97 (0.73, 1.28)	0.86 (0.56, 1.32)	0.96 (0.70, 1.32)
Dry season (Jan. – May)	Ref.	Ref.	Ref.
Southwest monsoon (June – Sept.)	0.97 (0.56, 1.66)	0.27 (0.11, 0.68)††	0.86 (0.48, 1.53)
Northeast monsoon (Oct. – Dec.)	2.26 (1.23, 4.19)††	0.93 (0.32, 2.67)	2.39 (1.24, 4.63)††
House flooding cluster			
Year-round[b]	1.21 (0.74, 1.98)	1.20 (0.59, 2.46)	1.13 (0.65 (1.96)
Dry season (Jan. – May)	Ref.	Ref.	Ref.
Southwest monsoon (June – Sept.)	0.74 (0.28, 1.94)	0.18 (0.02, 1.63)	0.75 (0.28, 2.02)
Northeast monsoon (Oct. – Dec.)	1.02 (0.35, 2.96)	2.11 (0.47, 9.42)	2.19 (0.65, 7.38)

[a]All models adjusted stool type (routine vs. diarrheal stool). Parasite results not shown due to small number of events in cluster, preventing model convergence; [b]Pooled across seasons; †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$

Table 5: Bivariate relationships between household exposure behaviors and pathogen detection in stool collected from children in SaniPath households, 2010-2014[a]

	Bacteria OR (95% CI)	Parasite OR (95% CI)	Virus OR (95% CI)	Any pathogen OR (95% CI)
Drain contact				
Any	1.45 (0.99, 2.11)	2.41 (1.00, 5.81)†	1.26 (0.71, 2.26)	1.51 (0.99, 2.29)
>10 times per month	1.25 (0.88, 1.75)	1.73 (0.90, 3.33)	1.18 (0.74, 1.88)	1.33 (0.90, 1.97)
Flood water contact				
Any	1.47 (1.07, 2.00)†	1.35 (0.69, 2.64)	1.52 (0.94, 2.44)	1.44 (1.01, 2.04)†
>10 times per month	0.96 (0.68, 1.36)	1.56 (0.79, 3.06)	0.86 (0.52, 1.44)	0.93 (0.63, 1.36)
Open defecation				
<5 YO	1.33 (1.00, 1.77)	1.34 (0.73, 2.49)	1.05 (0.69, 1.61)	1.35 (0.97, 1.87)
5-12 YO	1.68 (1.18, 2.39)††	1.19 (0.58, 2.43)	0.99 (0.61, 1.60)	1.56 (1.05, 2.33)†
Adult	1.51 (1.19, 1.92)†††	1.41 (0.83, 2.40)	1.20 (0.83, 1.74)	1.48 (1.12, 1.95)††
Public toilet use				
Any	0.85 (0.67, 1.08)	1.30 (0.79, 2.13)	1.05 (0.73, 1.50)	0.93 (0.70, 1.23)
>10 times per month	0.92 (0.65, 1.30)	1.10 (0.54, 2.26)	1.46 (0.90, 2.35)	1.10 (0.73, 1.66)
Municipal water consumption				
Any	0.93 (0.67, 1.29)	0.98 (0.50, 1.91)	1.34 (0.81, 2.21)	0.93 (0.64, 1.36)
Daily	0.94 (0.72, 1.22)	0.87 (0.52, 1.47)	1.10 (0.75, 1.60)	0.87 (0.64, 1.17)
Drinking water treatment	0.73 (0.57, 0.94)†	0.56 (0.33, 0.94)†	0.69 (0.47, 1.02)	0.66 (0.50, 0.88)††
Produce consumption				
Any	0.95 (0.73, 1.24)	0.92 (0.54, 1.57)	1.05 (0.72, 1.55)	0.87 (0.65, 1.18)
Daily	1.07 (0.78, 1.47)	1.04 (0.55, 2.00)	1.28 (0.82, 1.99)	1.15 (0.80, 1.66)

[a]All models adjusted for monsoon seasons and stool type (routine vs. diarrheal stool); †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$

Table 6: Multivariate fecal exposure models by pathogen group for children in SaniPath households, 2010-2014[a]

a) Bacterial infection	β	SE	OR (95% CI)
Cluster of drain flooding	-0.25	0.16	0.78 (0.57, 1.06)
Cluster of drain flooding during northeast monsoon	0.85	0.29	2.34 (1.32, 4.15)††
Open defecation (adult)	0.42	0.13	1.53 (1.19, 1.97)††
Flood contact (any reported per month)	0.39	0.16	1.48 (1.09, 2.00)†
b) Viral infection			
Size of household population	0.09	0.04	1.09 (1.01, 1.18)†
Cluster of reported drain flooding	0.19	0.26	1.21 (0.73, 2.00)
Cluster of drain flooding during southwest monsoon	-1.27	0.45	0.28 (0.12, 0.68)††
Reported water treatment	-0.36	0.20	0.70 (0.47, 1.02)

[a]Using AIC for model selection, all models adjusted for season (dry season, southwest (June to September) monsoon, or northeast (October to December) monsoon [136] and type of stool specimen (routine vs. diarrheal stool) (not shown). A final multivariate model for parasitic infection is not presented due to a lack of model convergence preventing testing of all relevant covariates; [b]Toilet excreta contained onsite; †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$

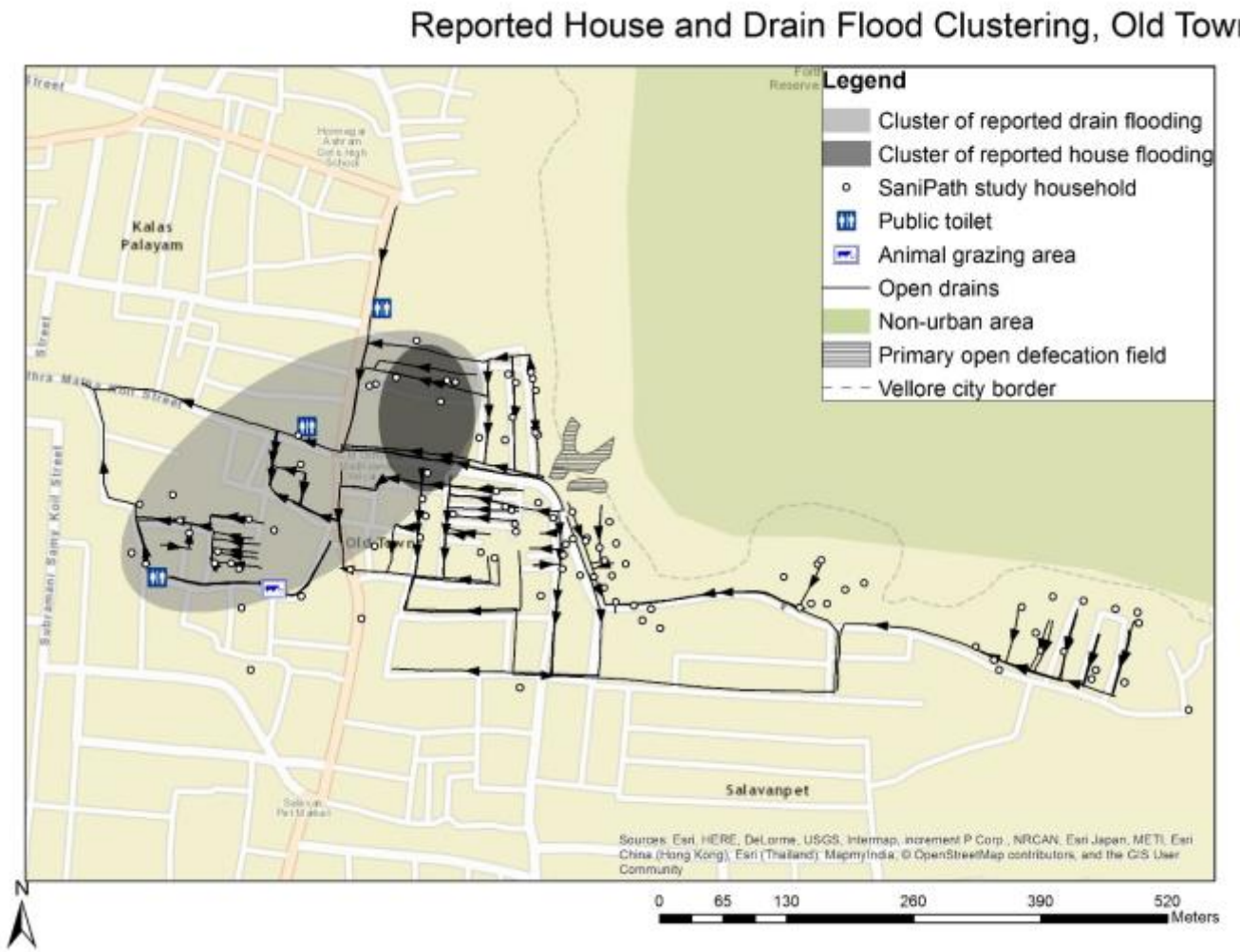


Figure 1: Reported drain and house flood clustering, Old Town. Significant clusters of reported drain and house flooding, determined by Kulldorff's Bernouilli spatial scan [52], are presented using light and dark gray ellipses, respectively. Each SaniPath study household (represented by white dots) within each of these clusters reported flooding. Black lines represent drains, with arrows indicating the direction of drain flow. Only drains within the Old Town neighborhood boundary are presented.

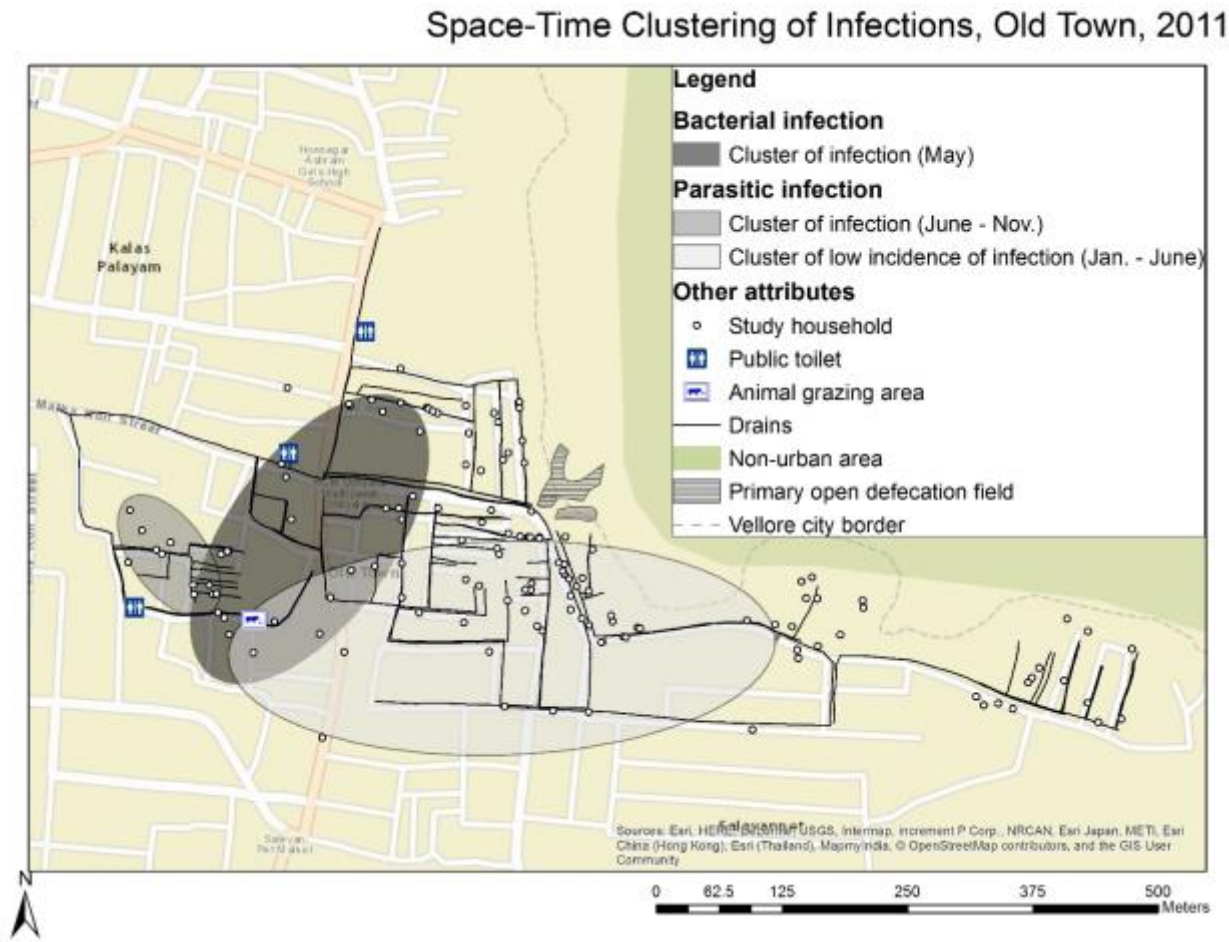


Figure 2: Space-time clustering of infections, Old Town, 2011. Significant space-time clusters of enteric infection, by type of infection (bacterial or parasitic), are presented using ellipses. The cluster of low incidence of infection is represented in space by the light gray ellipse, while clusters of high incidence of infection are represented in space by darker gray ellipses. The time period of the high or low space-time cluster is noted in the figure legend in parentheses. Study households are represented by white dots, and include every MAL-ED study household within the SaniPath-specific sub-neighborhood areas that had at least one stool specimen collected from the study child during 2011. Black lines represent drains, with arrows indicating the direction of drain flow. Only drains within the Old Town neighborhood boundary are presented.

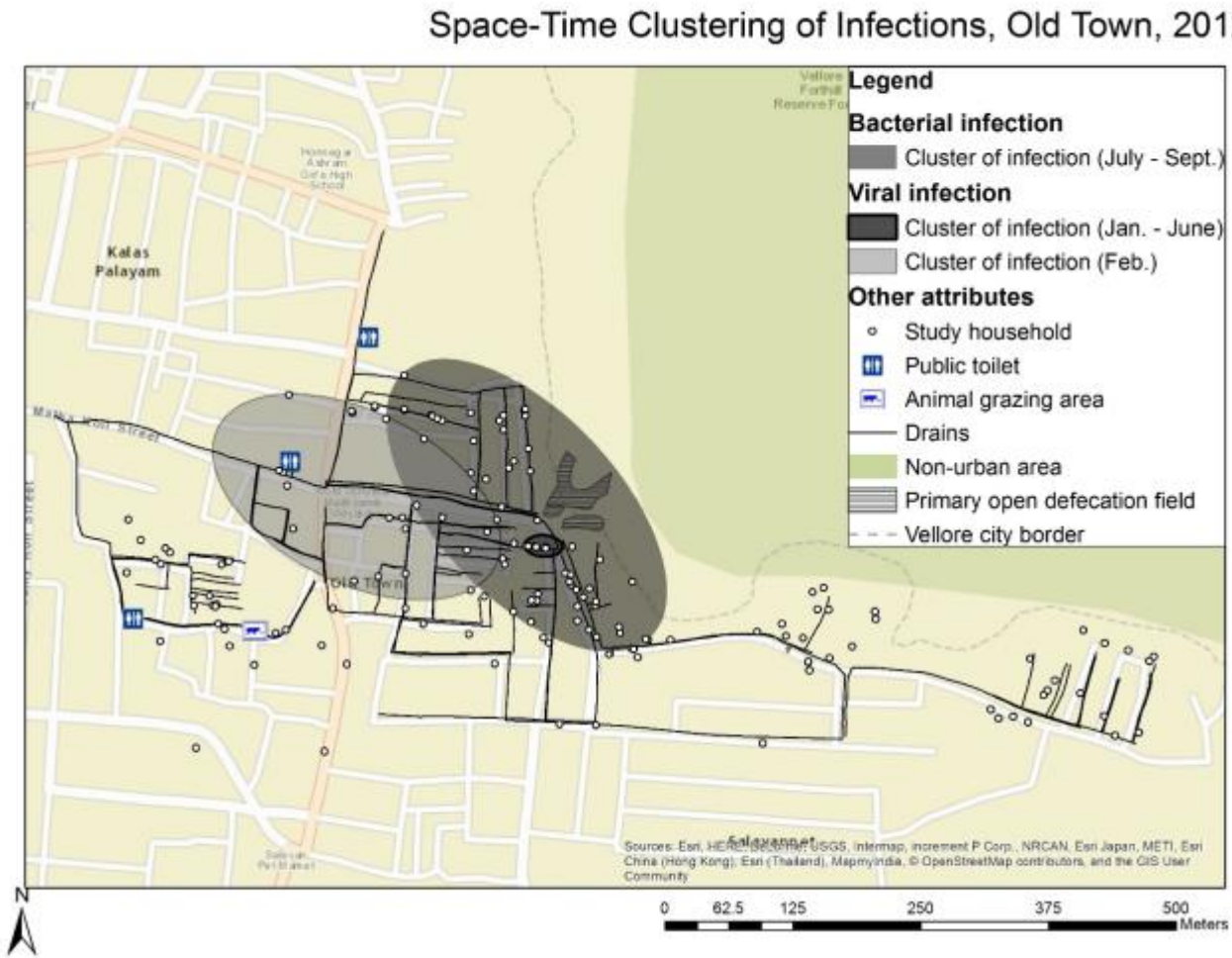


Figure 3: Space-time clustering of infections, Old Town, 2012. Significant space-time clusters of enteric infection, by type of infection (bacterial or viral), are presented using ellipses. Clusters of high incidence of infection are represented in space by gray ellipses, with the time period of the cluster noted in the figure legend in parentheses. Study households are represented by white dots, and include every MAL-ED study household within the SaniPath-specific sub-neighborhood areas that had at least one stool specimen collected from the study child during 2012. Black lines represent drains, with arrows indicating the direction of drain flow. Only drains within the Old Town neighborhood boundary are presented.

Conclusions

This dissertation used a mechanistic approach to understanding the effectiveness of sanitation systems in six low-income, urban neighborhoods in two countries by examining their associations with fecal contamination in household and community environments and with enteric infection in children. Aim 1 of the dissertation showed that clusters of significantly higher coverage of contained household sanitation (that is, household sanitation with good FSM) were associated with significantly lower concentrations of *E. coli* in drain water. This finding highlights the potential for the type and coverage of household sanitation to affect fecal contamination in the public domain.

Additionally, the findings from Accra suggest that household sanitation, including FSM, has a threshold—rather than a dose-response—effect on fecal contamination. While significant differences in fecal contamination inside and outside of clusters of high coverage of household facilities with good FSM were observed, no significant differences were observed with the level of local household sanitation coverage, a measure of dose-response. Previous studies have shown threshold relationships between household sanitation coverage and diarrhea disease, as well as stunting, in children. [36–38,56] Work from the 1990s and early 2000s suggested that at a 75-80% threshold of neighborhood coverage of household sanitation, even households without their own sanitation facility showed decreases in diarrhea incidence.[36,37] Although the more recent evidence from Fuller et al. (2016) uses a measure of improved sanitation within 500m and was not able to determine an exact threshold for sanitation coverage to have benefits on childhood stunting, the authors still underscore the importance of neighborhood-level coverage in influencing an individual child/household's outcome.[56] The results from Aim 1, evaluating an intermediate outcome (fecal contamination) on the pathway to diarrheal disease, enteric

infection, and stunting, make sense in the context of this previous research. The coverage levels observed in Accra (44-68%) were lower than the 75-80% suggested for health outcomes, as would be expected in the context of the F-diagram, where decreased fecal contamination should be observed before changes in health outcomes are observed.[27]

Results from Aim 2 showed that the spatial heterogeneity of sanitation and the downstream FSM associated with household toilets were significantly associated with both fecal contamination and enteric infection outcomes. Associations between household toilets and fecal contamination or enteric infection varied by coverage of sanitation and type of FSM, suggesting that both are important in ensuring effective containment of fecal sludge at the household and neighborhood level.

Finally, the third Aim showed associations between public domain exposures linked to poor FSM practices and enteric infection in children living in the immediate area. In some cases, these exposures were shown to be affected by the location of the household, and not the frequency of reported personal contact, suggesting that neighborhood geography and infrastructure is important in increasing risk. Further, risk factors varied between different groups of enteric infections, reflecting the variation in environmental transmission pathways and biological characteristics between pathogens and indicating new areas for WASH-related research focus.

The results of Aims 1 and 2 show differences in associations between GII norovirus presence in drains and FSM. Specifically, GII norovirus concentrations were significantly higher in areas of poor FSM in Vellore (Aim 2), while no significant associations were observed in Accra (Aim 1). This discrepancy likely reflects two key differences between the studies: the use of presence/absence vs. concentration data and differences in disposal practices. In the Vellore

study, though GII norovirus concentrations were significantly higher in areas of poor FSM, no significant differences in detection (presence/absence) of GII norovirus were observed in areas of poor FSM compared to the rest of the neighborhood, similar to the findings in Accra. This is likely because use of presence/absence data provides a coarser level of detail, thus differences may be more difficult to detect. Also, given its environmental persistence, differences in GII norovirus levels in the environment are likely above the limit of detection.[17] Further, sampling locations in open drains in Vellore were closer to households and their discharge pipes when compared with Accra, providing a more direct indication of nearby household FSM.

The research in this dissertation has underscored the importance of excreta containment, particularly in dense, urban settings, because of the variety of potential pathways of exposure to fecal contamination. Previous QMRAs have highlighted the importance of poor FSM through drain exposures.[69,70] In low-income neighborhoods of both Accra, Ghana and Kampala, Uganda, exposure to "gray water" in drains was the highest risk pathway, even when compared with drinking water, for young children.

Conversely, good FSM—primarily in the form of sewerage—has been shown to be effective in reducing diarrhea and enteric infections, as shown in a recent systematic review.[117] Notably, the results of this dissertation are supported by the observed effects of community-level sewerage interventions on diarrhea in Brazil (21-66% reductions).[120,180] However, the earlier study in Brazil by Moraes et al. (2003) also suggests that poor FSM (drainage) yields reductions in childhood diarrhea when compared to diarrhea in children in the control group (who had toilets but no designed drainage or sewerage), which contrasts with results presented in Aim 2 of this dissertation that suggest that poor FSM is associated with increased prevalence of enteric infections.[120] This differences between studies may result from differences in the definition of

urban drainage and the outcome measured. Specifically, drains in the "drainage group" in Brazil were covered, which may have reduced exposure when compared with the open drains observed in Vellore. The control group in the Brazilian study had high levels of disposal of feces on the ground or in open drains, by comparison. Further, the target outcome of the Brazilian study was reported diarrhea in children under 5 (compared to enteric infection measured from stool specimens from children in Vellore), which may have been subject to response bias associated with the intervention.[190]

Strengths and Limitations

This dissertation has some notable strengths and limitations in the methodological approach. The use of microbiological analysis of stool and environmental samples provides more objective outcome measures than self-reported diarrhea and incorporates quantification of both asymptomatic and symptomatic infection.[30,58,66] Comparison of pathogen-specific risk factors provides important health data to support established and identify new pathways of fecal exposure and disease from previous QMRAs.[69–73,119] Measurement of associations between sanitation coverage and fecal contamination in environmental samples provides direct measurement of the mechanism by which sanitation operates, reducing potential confounders or competing pathways.[27] As illustrated in the F-diagram, effective sanitation works to reduce incidence of enteric infection by reducing levels of fecal contamination in the environment. However, the inconsistency in environmental sampling in the sanitation literature, including variation in types of samples and organisms measured and their correlations with health outcomes makes contextualization of these results difficult.[74,78–81,84,96,116] Specifically, with regard to the effectiveness of sanitation, samples of hands, sentinel objects, soil, swabs, and in

stored water have been measured for *E. coli*, fecal coliforms (thermotolerant coliforms), *Enterococci*, *Bacteriodales*, and enteric viruses, all of which differ in infectious doses, environmental persistence, and source of the contamination (e.g. human only, human and animal contributions). Given this variety in methods, overall conclusions about the effects of sanitation on fecal contamination are difficult when evaluating *E. coli* and enteric viruses in samples within and outside the household in this dissertation. Further, some types of samples evaluated, such as hand rinses and soil, have been shown to vary with factors other than sanitation, including time since last handwashing event (for hand rinses), activity of the subject (for hand rinses), and amount of foot traffic in the area (for soil), that were not measured.[79–81]

While measurement of the concentration of fecal contamination in environmental samples provides a more proximal outcome with regard to the impact of sanitation on exposure than health outcomes, qPCR assays had relatively low sensitivity (high lower limits of detection) for pathogen detection, likely due to the type of environmental media tested.[76,101] In addition to DNA or RNA from the target organism, soil and drain water may have additional components, including humic acids or compounds and metals, that inhibit PCR processes.[101] Given that the lower limits of detection were generally higher than the ID50 (dose which, on average, results in 50% of the population becoming infected) of the pathogens and PCR assays are unable to indicate infectivity, inferences about public health risk from environmental concentrations is limited.[12,18,85]

In both study sites, despite the amount of microbiological data collected from environmental samples and stool specimens, sample size was limited. Analytical models for urban environmental fecal contamination did not have sufficient sample size to include all potential confounders, like presence of a public toilet or ownership of animals in the area, tested

in previous bivariate analyses. While assessment of these covariates through bivariate analyses is useful, the ability to control for them in multivariate models would have yielded stronger evidence. Further, the small sample size limited the power of bivariate and multivariate models to detect significant effects, restricting inferences made from the data.

The temporality and content of household surveys have strengths and limitations as well. Despite its necessity due to study logistics, assessment of exposures (household surveys) after outcomes (MAL-ED study data) was a major limitation of the analysis. While household and neighborhood characteristics were likely consistent over time, exposure behaviors of children measured at 3-5 years of age likely differed from exposure behaviors during the first two years of life, when children were followed for enteric infection.[124] Further, while this work was one of the few to assess FSM practices, in contrast to presence/absence of a toilet alone, management of fecal sludge may change over time, particularly for households with a toilet with onsite containment, and removal of fecal sludge from the household does not guarantee safe treatment.[31,32,118] Thus, measurement at a single time point may not have consistently represented the household's FSM status at all points during follow-up.

The absence of detailed data on hygiene-related behaviors and animal presence, particularly at the Vellore study site, may have limited inferences at the household as well. While the CMC hygiene survey provided an approximation of WASH, and specifically hygiene, behaviors at the household, all handwashing behaviors were self-reported, a method that is known to have response bias.[175,191,222] Consequently, there was little variation in reported handwashing behaviors within neighborhoods, limiting our precision in controlling for these practices (as proxies for household hygiene) when evaluating pediatric enteric infection risk. Further, while observation and mapping of animal grazing areas suggests domestic livestock

were common in Vellore, we were not able to measure animal ownership as in Accra. Presence of livestock in and around human living quarters may have been an important contributor to fecal contamination levels and enteric infection risk within the home.[220] Animal contact may be a source of exposure to fecal contamination and zoonotic enteric infections, particularly through direct contact or contamination of soil, thus the associations between direct (i.e. with drain water) or indirect (i.e. open defecation) measures of environmental contact may be overestimated.

The absence of detailed climate data is a limitation when quantifying environmental fecal contamination levels and enteric infection. While models were adjusted for season using the date of sample collection, relationships between enteric infection and climate may be more complex, as diarrheal diseases have been shown to vary with temperature, rainfall, and other factors.[223–226] While the relationship between environmental fecal contamination in urban settings and climate has not been well-studied, fecal contamination in drinking water sources and other tropical ecosystems has been shown to vary with various climatic factors as well.[167,227,228] Thus, collection of rainfall, temperature, and other climate data may provide more accurate adjustment for the effects of climate on fecal contamination and enteric infection in future studies.[224]

Finally, the examination of spatial heterogeneity in sanitation coverage and fecal exposures through cluster analysis of survey households was a strength of this work. Spatial heterogeneity in sanitation has been observed on multiple scales, and space-time clustering of enteric infections has been demonstrated in previous studies.[40,43,168,187,229–231] Specifically, clustering of rotavirus and cholera incidence have been shown to be clustered in space and time in urban, Indian settings, while heterogeneity in sanitation has been observed at national and city levels.[40,168,231] Though no studies have directly evaluated sanitation clustering and enteric

infection, Fuller et al. (2016) showed that the sanitation coverage within 500m of a household was related to stunting, acknowledging the importance of neighborhood-level sanitation coverage.[56] Failure to account for this clustering at the sub-neighborhood level could lead to misclassification in exposure-disease relationships.[232,233] Though spatial clustering was assessed from the locations of survey responses based on a sample, and not a census, of households, systematic sampling was used in both study locations to approximate the underlying distribution of the population. Further, the spatial scan statistic used is robust to variation in the underlying population distribution, and significance levels and parameters of the scan were adjusted to identify only the most significant clusters.[52,168]

Future Directions

Future studies need to identify feasible sanitation solutions that reduce environmental fecal contamination and, subsequently, enteric infection. These solutions require interventions in both the private and public domain supported by the work of local governments, community members, and the research community. The results of this dissertation indicate that focusing top-down interventions at the household alone will be insufficient to control fecal contamination, and subsequent enteric infection, within the community. Household infrastructural and behavioral interventions need to be complemented by community-level infrastructure to reduce risks in the public domain. The SFDs provide an indication of the initial problem, but must be followed by specific actions, as well as evaluation, of these actions to understand what is effective and ineffective. Evaluation and research need to contextualize sanitation within the household and community context where the work is being done, by using spatial coverage data for example, to understand public and private drivers of infection risk. This understanding should begin by

evaluation of levels of environmental fecal contamination, rather than self-reported health outcomes. Use of objective health outcomes that measure symptomatic and asymptomatic infection is also warranted to understand environmental causes and mechanisms of environmental enteropathy and subsequent longer-term sequelae. While cooperation between researchers, local governments, and the community is a challenge in any setting, initiation of a cycle of evidence-based policy-making, informed by research and followed by rigorous evaluation and learning, is essential to mitigating health risks and developing effective interventions.

Policy and Involvement of Local Governments and Communities

Because of the interconnectedness of fecal exposures in the private and public domains, WASH interventions must be targeted in both areas and cannot be seen as the responsibility of the individual alone. The results of this dissertation suggest that exposures in the public domain, including open drain flooding and public toilet use, are important contributors to pediatric enteric infection risk. However, these exposures could not be easily modified by changing an individual's behavior alone. For example, risk of enteric infection associated with floodwater was elevated in spatial clusters of reported drain flooding, regardless of frequency of reported floodwater contact. That is, location of residence (i.e. living in the spatial cluster of drain flooding) was associated with increased enteric infection risk, likely because children living in these clusters encountered floodwater everywhere they went and could not escape it. Therefore, community-level infrastructural interventions, and not behavioral interventions alone, may be needed to modify the neighborhood geography and mitigate flooding.[123] Further, while hygiene activities, like handwashing with soap after toilet use, should reduce the magnitude of exposure,

handwashing infrastructure was not present at any of the public toilets.[27] Recent research has underscored the importance of infrastructure interventions, with or without behavior change messaging, in changing handwashing behaviors, as messaging alone has been unsuccessful in promoting long-term changes and improvements in health.[218,219,234–236] Because of the importance of the public environment in mitigating behaviors, fecal exposures, and risk of enteric infection, communities and governments need to play a role and provide infrastructure, awareness, and education to promote healthy behaviors and reduce environmental fecal contamination levels in the public domain.

As discussed in the introduction, the importance of FSM has been highlighted through the work of the SFDs.[32] While SFDs provide an important advocacy tool to bring awareness to the amount of untreated fecal sludge unsafely discharged into urban communities, follow-up action is needed. Providing this "next step" after a municipal government is "triggered" by the SFD is an important role that researchers and development actors must work together to fill through affordable, feasible, yet safe downstream interventions. The research in this dissertation takes the understanding of the consequences of poor neighborhood-level FSM one step further by providing evidence of how poor FSM may contaminate the local environment and, indirectly and directly, affect human health.

In terms of targeting sanitation interventions, the results of this dissertation underscore the importance of household sanitation facility coverage, FSM improvements, and public domain improvements. Rather than choosing one over another, the results indicate that both sanitation coverage and FSM improvements are necessary to achieve reductions in fecal contamination and, subsequently, enteric infection incidence in children. Improvements in the public domain (for example, flood reduction) should be contextualized with FSM interventions. Doing this will

improve understanding of the community-level nature of this issue and the need for community-level, in addition to household-level, interventions. However, findings from the research in both Vellore and Accra also suggest that open drains in poor areas are never a good downstream FSM strategy, even in the interim. Open drains in both areas had high levels of both *E. coli* and enteric pathogens, and the data from Vellore showed that flooded drains were a significant risk factor for infection. Instead of constructing open drains, efforts to improve downstream management of fecal sludge should focus on better containment at the household and available, accessible, and affordable emptying and transport services.

Research and Evaluation

Evaluation of sanitation interventions must continue to measure not only the presence or absence of the toilet, but presence or absence of containment along the entire sanitation chain, particularly at the neighborhood and city scale.[32] While toilet provision combined with sewerage interventions have been shown to be effective in urban areas, sewerage is not a financially or logistically viable FSM option in many urban neighborhoods, and especially low-income areas.[32,117,120,121,183] The findings of higher prevalence of enteric infection in children households with toilets with poor FSM, even above those in households without toilets, suggests that poor FSM at the household may be worse than having no toilet at all. Densely populated urban areas need FSM interventions that remove fecal sludge from the residential environment, like trucking of contained fecal sludge to a treatment plant or transfer station, that are feasible and contain the excreta along the entire sanitation chain.[32,194]

As sanitation interventions continue to be implemented in urban settings, spatial data is an important resource for understanding the urban context. Results from this dissertation suggest

that differential effects of sanitation on household fecal contamination may be observed between high and low sanitation coverage areas, adjusting for household hygiene and WASH practices. Specifically, floors of households with a toilet had significantly higher levels of fecal contamination than those of households without a toilet in the neighborhood with low sanitation coverage, which may be due to sharing of toilets increasing contamination, as shown in a recent study from peri-urban Peru.[237] Because the effectiveness of sanitation may vary with the local sanitation coverage, it is important to describe and quantify the spatial heterogeneity of household toilet and FSM coverage within or between neighborhoods in order to identify where new interventions can be targeted and where existing sanitation, and especially FSM, interventions must be improved.[36–38] Further, in this dissertation, collection of spatial data from household survey responses and drain maps allowed identification of clusters with potentially high risk exposures. Understanding how components of the sanitation chain are distributed at the neighborhood- and city-level provides understanding of the interconnectedness of neighborhoods and their fecal sludge.

Before the effect of sanitation on health outcomes is evaluated, implementers and evaluators must measure and reduce environmental levels of fecal contamination in key areas. Further research is needed to understand more precisely how fecal contamination levels in different environments vary with sanitation coverage. Even studies to establish a baseline in high-income contexts would prove useful. Regardless, sanitation interventions should be measured by their ability to reduce levels of fecal contamination in the immediate, and even downstream, environment.[27] Results presented here suggest that high coverage of household sanitation that contains excreta onsite may reduce fecal contamination in the public domain. Understanding of this intermediate outcome is a key step—in the short term—in identifying why

some sanitation interventions have context-specific effectiveness and—in the longer term—in designing sanitation interventions that effectively reduce environmental fecal contamination.

Though only recently a focus in the larger research community, enteric infections and diarrhea caused by viruses need to be studied further in the WASH context. Results presented here indicate that viral infections were prevalent in children under two in Vellore and were predominantly associated with symptomatic (diarrheal) infection. Viral infections are prevalent, yet understudied (with the exception of rotavirus), in children in low-income contexts.[16,25,147,148,169] The results of this dissertation further suggest that levels of viruses in the environment may not respond to sanitation coverage in the same manner as levels of FIB. Because of their low infectious doses and long environmental persistence, viruses present a unique challenge to the WASH community.[12,18] Though vaccines have been developed for some viruses, like rotavirus, their efficacy in low-income contexts has not matched that demonstrated in high-income settings.[238–241] Though the cause is unknown, WASH-related environmental conditions may play a role in enabling or preventing repeat enteric infections, environmental enteropathy, and subsequent efficacy of this or other oral vaccines.[4,238,241] Thus, understanding the environmental distribution of viruses and their associations with sanitation coverage and other WASH interventions is an important part of reducing the overall burden of infection.

In addition to effects of WASH on vaccine efficacy, WASH interventions should be evaluated for their effects on environmental enteropathy (EE) and other longer term health outcomes.[4–6] Health outcomes associated with EE include malnutrition, stunting, and poor cognitive outcomes. Thus, the potential impact of WASH is much greater than reduced morbidity and mortality due to diarrhea alone.[4,6,7,242] Because EE is thought to be caused by repeat, generally subclinical enteric infections that cause morphological changes in the gut and

impair the absorption of nutrients, it is important for future studies to measure the effects of WASH on both symptomatic and asymptomatic infection.[7,242]

Contribution to the field

Despite the additional research needed on urban sanitation and FSM, this dissertation has provided a significant contribution to the field. Examination of associations between downstream, non-sewer-based FSM and objective health outcomes are some of the first in the field. Further, its examination of the associations between community-level coverage and fecal contamination reopens previous discussion (along with the work of Fuller et al.) on the importance of the sanitation in the vicinity of a household. The ability to deconstruct the role of sanitation into associations with fecal contamination and subsequently look at enteric infection provides an approach for future studies to better understand sanitation mechanistically, which is essential to future provision of effective interventions. As urban centers in low-income countries become more densely populated, it will be even more important for sanitation to function effectively on the household and community level, in combination with efforts to promote clean water and proper hygiene, to minimize health risks to those most vulnerable.

Appendix: Additional analyses and data

Aim 1 (Accra, Ghana)

Associations between drain size, construction, and detection of enteric viruses

Methods:

Variation in enteric virus presence/absence in drain water with increasing drain size and drain construction was tested by logistic regression. Drain size was divided into four categories: small (< 0.5m), medium (0.5 – 1m), large (1 – 3m), and extra-large (> 3 m). Drain construction was divided into ecological construction or formal construction.

Results:

To determine whether attributes of the drain itself were associated with enteric viruses in drain water, logistic regression models were fit for drain size and type of construction, controlling for neighborhood and season of sample (Table A1.1). Overall, no consistent trends in odds of detection of enteric viruses with increasing or decreasing drain size were detected. Detection of enteric viruses was 1.6-2.2 times more likely in ecologically-constructed drains when compared with formally-constructed drains, though these relationships were not statistically significant.

Discussion:

Drain size and construction were not significantly associated with detection of enteric viruses. Given that these enteric viruses are human-specific and, in Accra, smaller drains are more likely to be near to households than larger drains, one would expect the odds of detection to decrease with increasing size of drains.[86] None of the ORs for enteric virus detection displayed this trend consistently. Though detection of adenovirus was most likely in small drains, detection

of noroviruses was mixed across drain size. However, because noroviruses can survive for months to possibly years in soil and groundwater, one would expect to be more likely to detect noroviruses in ecologically-constructed drains when compared with those that had more formal construction.[12,17] Overall, drain construction showed little effect on detection of enteric viruses in this setting.

Table A1.1: Variation in enteric viruses in drain water with drain size and type of construction[a]

		Odds Ratio (95% Confidence Interval)		
		Adenovirus	GI norovirus	GII norovirus
Drain size:	> 3m	0.82 (0.08, 8.67)	2.70 (0.38, 25.5)	0.55 (0.09, 3.22)
	1-3m	0.60 (0.06, 4.71)	1.52 (0.20, 14.2)	1.06 (0.19, 6.17)
	0.5-1m	0.42 (0.06, 2.02)	2.28 (0.48, 16.8)	0.83 (0.20, 3.42)
	< 0.5m	Reference	Reference	Reference
Ecological construction		2.23 (0.43, 13.6)	1.78 (0.47, 7.19)	1.62 (0.48, 5.70)
Formal construction		Reference	Reference	Reference

[a]All models control for neighborhood (using Alajo as the reference) and season of sample collection (primary rainy season vs. all other).

Associations between sanitation variables and C_t values of enteric viruses in Accra, Ghana

Methods:

Because of inconsistent standard curves, analyses of associations between sanitation variables and enteric viruses in drain water were limited to presence/absence data. To approximate relationships between the amount (concentration) of enteric viruses present in a drain sample and selected sanitation variables, Ct values from qPCR were modeled by linear regression.

Results:

To estimate variation in approximate viral concentrations with the presence of public toilets and levels of local household sanitation coverage, C_t values were modeled by linear regression, controlling for neighborhood, season of sampling, and population density. Overall, none of the public toilet or local household sanitation coverage classifications showed consistent, significant associations at 50 and 100m distances from drain samples (Table A1.2). C_t values for adenovirus in drain water were significantly higher (indicating significantly lower concentration) within 100m of public toilets compared to outside of that distance. Further, GI norovirus C_t values were significantly higher (indicating again significantly lower concentration) with increasing coverage of minimally shared household sanitation (using the people per toilet definition) within 100m.

To estimate variation in approximate viral concentrations with the presence of cluster of household sanitation coverage, C_t values were modeled by linear regression, controlling for neighborhood, season of sampling, and population density (Table A1.3). Few significant associations were observed between household sanitation clusters and C_t values for enteric

viruses in drain water, though GII norovirus was the only organism to show consistent trends of 1) lower concentrations (higher C_t values) within clusters of high coverage of household sanitation variables and 2) higher concentrations (lower C_t values) within clusters of low coverage of household sanitation variables. The lone deviation from these trends was observed between GII norovirus C_t values within and outside of clusters of low coverage of minimally-shared household sanitation (using the people per toilet definition) within 50m, which was likely a result of sparse data and small sample size, as shown by the large standard error associated with the effect size.

Discussion:

Overall, enteric virus C_t values were not significantly associated with local household sanitation coverage or sanitation coverage clusters, though consistent trends were observed with GII norovirus, but not other viral outcomes. The absence of significance in enteric virus concentrations is likely influenced by a few factors, including the small sample size and the finite precision of the C_t values themselves when tested on a continuous scale with linear regression. Both likely weakened the power of these regression models to detect significant differences. However, the trends observed between GII norovirus concentrations and sanitation coverage clusters follow our initial expectations. Given that viral infections were not associated with household sanitation (as observed in Aim 3), we would expect that viral shedders should be evenly distributed between clusters and non-clusters of sanitation, but that the presence of clusters of high coverage of household sanitation, especially contained and even minimally shared sanitation, should limit the spread of GII norovirus into the environment.[27]

Table A1.2: Adenovirus, NoV GI, and NoV GII C_t values for public drain water by local sanitation prevalence in Accra, Ghana[a]

	Within 50m of drain sample (n = 58)			Within 100m of drain sample (n = 72)		
	Adenovirus	GI norovirus	GII norovirus	Adenovirus	GI norovirus	GII norovirus
Main effect of model	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]
Public toilet[d]	1.92 (1.80)	0.68 (0.98)	-1.04 (0.97)	3.69 (1.31)††	0.88 (0.69)	0.02 (0.76)
Any HH sanitation	1.35 (3.56)	1.72 (1.91)	0.34 (1.92)	-1.62 (3.78)	1.82 (1.90)	2.09 (2.10)
Contained HH sanitation	1.27 (4.10)	1.25 (2.21)	2.84 (2.17)	1.08 (3.84)	1.64 (1.93)	2.16 (2.13)
Minimally-shared HH sanitation						
≤ 5 HHs/toilet	-0.90 (6.52)	3.79 (3.48)	1.59 (3.51)	-5.32 (4.15)	1.84 (2.11)	2.88 (2.32)
≤ 30 people/toilet	1.49 (7.06)	5.62 (3.73)	-0.46 (3.80)	4.21 (4.25)	4.38 (2.10)†	2.35 (2.37)

[a]All models adusted for neighborhood, population density, and season of sample; [b]Estimate is the difference in C_t value; [c]Standard error; [d]Odds of detection in samples within the given distance (50 or 100m) of a public toilet compared to those not within the given distance; †$p < 0.05$; ††$p < 0.01$;

Table A1.3: Adenovirus, NoV GI, and NoV GII C_t values for public drain water by sanitation coverage cluster in Accra, Ghana[a]

	50m vicinity of drain sample (n = 58)			100m vicinity of drain sample (n = 72)		
	Adenovirus	GI norovirus	GII norovirus	Adenovirus	GI norovirus	GII norovirus
Main effect of model	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]	Estimate[b] (SE)[c]
Any HH sanitation						
High Coverage Cluster	2.49 (2.49)	1.28 (1.35)	0.94 (1.35)	2.26 (2.26)	1.40 (1.14)	1.49 (1.25)
Low Coverage Cluster	2.88 (1.88)	0.78 (1.03)	-0.84 (1.03)	2.85 (1.69)	0.27 (0.87)	-1.81 (1.20)*
Contained HH sanitation						
High Coverage Cluster	11.3 (6.97)**	1.34 (1.40)	0.40 (1.41)	-1.51 (2.15)	1.20 (1.08)	1.09 (1.19)
Low Coverage Cluster	1.60 (3.07)	-2.82 (1.61)	-0.87 (1.65)	14.5 (8.86)*	-2.31 (1.36)	-1.24 (1.52)
Minimally-shared HH sanitation						
≤ 5 HHs/toilet						
High Coverage Cluster	2.80 (2.92)	1.57 (1.58)	2.20 (1.56)	2.82 (2.23)	1.64 (1.12)	1.30 (1.25)
Low Coverage Cluster	2.50 (2.97)	10.5 (5.21)†**	-0.49 (1.61)	2.69 (2.82)	4.98 (4.54)*	-0.91 (1.58)
≤ 30 people/toilet						
High Coverage Cluster	1.36 (2.28)	1.24 (1.22)	0.75 (1.23)	1.92 (2.02)	1.18 (1.02)	0.96 (1.13)
Low Coverage Cluster	2.88 (3.50)	1.54 (1.89)	39.4 (14.1)†**	-0.11 (2.95)	1.76 (1.48)	-3.06 (1.60)

[a]All models adusted for neighborhood, population density, and season of sample; [b]Estimate is for the difference in C_t value; [c]Standard error; †$p < 0.05$; *Interaction of population density and main predictor (sanitation variable) significant at $\alpha = 0.10$; **Interaction of population density and main predictor (sanitation variable) significant at $\alpha = 0.05$;

Aim 2 (Vellore, India)

Christian Medical College Hygiene Survey

Scores on the Christian Medical College (CMC) hygiene survey at each household were based on 18 criteria:

1. Ownership of a container with the sole purpose of storing drinking water
2. Presence of a cover for the container in (1)
3. Presence of an object (e.g. cup, etc.) to dip into water to collect water from the container in (1)
4. Boiling of drinking water
5. Covering of food
6. Cleaning or washing of fruit before consumption
7. Cleaning of the mother's breast before the child feeds
8. Indicator for the family member feeding the child (mother, grandmother, aunt, sister, or other)
9. Hand cleanliness (observation of whether nails are cut and whether there is dirt under them)
10. Handwashing before feeding the child
11. Cleaning of feeding implement before child eats
12. Availability of water at the toilet
13. Handwashing after use of the toilet
14. Frequency of bathing (less than once daily, once daily, twice daily)
15. Occurrence of bathing after using the toilet
16. Location where those in the household defecate
17. Location where the child defecates
18. Whether or not the family pours drinking water into a separate container (for households using bottled water)

These scores were tabulated numerically and divided into "good" (10-18) and "poor" (0-9).

Correlation between fecal contamination levels in different types of environmental samples

To evaluate how fecal contamination in different environmental sample types was correlated within and outside the household, correlation analyses were conducted for *E. coli*

concentrations in environmental samples collected within the household and norovirus GII concentrations in drain samples (Table A2.1). No significant correlations were observed between *E. coli* concentrations in samples from CAP. However, the correlation between *E. coli* concentrations in household floor swabs and sentinel object rinses was borderline significant in CAP, significant in OT, and significant in both neighborhoods combined (r = 0.38 in each neighborhood and overall). The correlation between *E. coli* concentrations in swabs and hand rinses from children was significant within OT (r = 0.48) and in both neighborhoods combined (r = 0.33). Norovirus GII concentrations in drains were not significantly correlated with *E. coli* concentrations in any sample types.

Table A2.1: Correlations between sample types within and outside the household[a]

	N	Sentinel objects (*E. coli*) r (95% CI)	HH floor swabs (*E. coli*) r (95% CI)	Drains (NoV GII) r (95% CI)
A) Chinnallapuram				
Child's hands (*E. coli*)	25	0.24 (-0.17, 0.58)	0.07 (-0.34, 0.45)	-0.04 (-0.43, 0.36)
Sentinel objects (*E. coli*)	25		0.38 (-0.02, 0.67)	0.18 (-0.23, 0.54)
HH floor swabs (*E. coli*)	25			0.18 (-0.24, 0.53)
B) Old Town				
Child's hands (*E. coli*)	25	0.22 (-0.20, 0.57)	0.48 (0.11, 0.74)†	0.06 (-0.35, 0.44)
Sentinel objects (*E. coli*)	24		0.38 (-0.03, 0.68)	-0.18 (-0.55, 0.24)
HH floor swabs (*E. coli*)	25			-0.01 (-0.41, 0.38)
C) Both neighborhoods				
Rinses of children's hands (*E. coli*)	50	0.25 (-0.04, 0.49)	0.33 (0.06, 0.56)†	-0.02 (-0.30, 0.26)
Sentinel objects (*E. coli*)	49		0.38 (0.11, 0.60)††	-0.03 (-0.31, 0.25)
HH floor Swabs (*E. coli*)	50			0.06 (-0.22, 0.34)

†p < 0.05; ††p < 0.01. [a]Drain *E. coli* concentrations were excluded from analysis due to resampling of drains at a different time period than all other sample types.

Further spatial analysis in Vellore, India

Methods:

Mapping and spatial analyses of households and drains were conducted in ArcMap 10.1 (ESRI, Redlands, CA, USA). Household microorganism concentrations were evaluated for spatial global autocorrelation (if there are spatial trends of similar values or not of pathogen concentrations throughout the neighborhoods) using Global Moran's I statistic in the Point Pattern Analysis program (PPA, San Diego State University, San Diego, CA).[243] The Global Moran's I statistic determines the correlation between values that are near to each other, with 0 being perfect random distribution, -1 being perfect negative autocorrelation (values near each other being extremely negatively correlated, or unlike, each other), and 1 being perfect positive autocorrelation (values near each other being extremely correlated to each other). Local clustering (identifying locations of clustering for high or low concentration values at the sub-neighborhood level) was assessed using the local Getis-Ord G* statistic in PPA.[244] The local G* statistic compares local attribute values within a distance surrounding each household or sample location under consideration to the global average to determine localized high or low clusters. Focal clustering, clustering of high or low concentration values at the sub-neighborhood level around specific foci) was assessed by the Getis G statistic in PPA.[244] The Getis G statistic is identical to the Getis G* statistic in calculation except that it omits the attribute value of the focal point itself from the calculation. Focal clustering of point data was evaluated using Diggle's method[245] in ClusterSEER version 2.5.1 (BioMedware, Ann Arbor, MI, USA). Diggle's statistic compares a model of values moving away from a particular focal point to a model outlining the underlying distribution of values in the dataset to goodness of fit to the data. SpaceStat (BioMedware, Ann Arbor, MI, USA) was used for regression with a spatial lag, which is used to

account for spatial clustering within outcome data and quantify the explanatory power of this clustering.

Results:

The concentrations of the target enteric microbes in environmental samples were examined for spatial clustering trends on global and local scales (global autocorrelation and local clustering), as well as focal clustering (clustering around a point) around open defecation and animal grazing areas in the neighborhoods (data not shown). The significance of spatial clustering varied across sample types. *E. coli* concentrations from swabs of household floors were significantly non-clustered (negatively globally autocorrelated) at all scales, from 50-500m, in CAP. Localized clustering of similar high/low *E. coli* concentrations was observed for sentinel object samples within 150-190m of each other in CAP. In OT, high values of swab *E. coli* concentrations were significantly clustered within 120-130m of the animal grazing area, while low values were significantly clustered within 220m of the primary open defecation field (significant focal clustering). The locations of the animal grazing area and primary open defecation field are shown in Figure 1b (Aim 2). Global, local, and focal clustering of norovirus GII concentrations in drains was not observed.

Final multivariate fecal contamination models

Final aspatial regression models of *E. coli* concentrations, by sample type, were chosen by adjusted R^2 values for models containing household sanitation and FSM variables (Table A2.2). Fecal contamination on children's hands was significantly elevated in households with a toilet with good FSM, compared to those without a toilet (Table A2.2a). However, in the same model, households with a toilet with good FSM and good hygiene status (representative of other water,

sanitation, and hygiene (WASH) practices) had significantly lower fecal contamination on children's hands. The interaction between good hygiene status and household toilet presence was associated with a significant reduction in fecal contamination on household floor swabs as well (Table A2.2b). Notably, the FSM associated with the toilet was not in the model for *E. coli* concentrations on swabs from household floors. As observed in bivariate analyses, households with a toilet in Old Town had significantly higher *E. coli* concentrations on swabs of household floors compared to those without a toilet in the neighborhood. No models with sanitation and FSM at the household or cluster level significantly fit the fecal contamination data from sentinel objects, thus no model is presented.

Table A2.2: Multivariate regression for within-household fecal contamination in child hand rinse and swabs across neighborhoods

	Parameter attributes			Model attributes		
	β	SE(β)[a]	p-value	F	p-value	Adjusted R^2
a) *E. coli* concentrations in hand rinses (log_{10}CFU/pair of hands, n = 50)				3.81	0.004	0.26
Toilet with excreta contained onsite	0.95	0.46	0.044			
Toilet discharging to drain	-0.42	0.61	0.492			
Good hygiene status	0.29	0.47	0.545			
Interaction: Toilet with excreta contained onsite & Good hygiene status	-2.14	0.72	0.005			
Interaction: Toilet discharging to drain & Good hygiene status	0.87	0.80	0.281			
Neighborhood: Old Town	0.43	0.29	0.143			
b) *E. coli* concentrations in household floor swabs (log_{10}CFU/125 cm^2, n = 50)				3.16	0.016	0.18
Household toilet	0.13	0.47	0.781			
Good hygiene status	1.69	0.69	0.018			
Interaction: Household toilet & Good hygiene status	-2.20	0.77	0.007			
Neighborhood: Old Town	-0.79	0.45	0.090			
Interaction: Household toilet & Old Town	1.53	0.57	0.010			

[a]Standard error.

Fecal contamination associations with enteric infection

To examine associations between within-household fecal contamination and enteric infection in children, mixed-effects logistic regression models were constructed, examining prevalence of enteric infection (with any pathogen) as the outcome and *E. coli* concentrations in environmental samples as the predictor (Table A2.3). Overall, no significant associations between *E. coli* concentrations in environmental samples and prevalence of enteric infection were observed. In both neighborhoods combined, however, detection of an enteric pathogen in a child's stool was higher in children whose hands had higher concentrations of *E. coli* when sampled, though this association was not significant.

Table A2.3: Any enteric pathogen detection in child stool by E. coli concentrations in environmental samples [a,b]

	Chinnallapuram PR for infection (95% CI)	**Old Town** PR for infection (95% CI)	**Both neighborhoods**[c] PR for infection (95% CI)
E. coli on hands[d]	1.70 (0.63, 4.62)	1.44 (0.65, 3.18)	1.41 (0.80, 2.49)
E. coli on sentinel objects	2.02 (0.64, 6.38)	0.84 (0.43, 1.64)	1.04 (0.60, 1.81)
E. coli on swabs	1.74 (0.67, 4.48)	1.07 (0.52, 2.21)	1.11 (0.66, 1.87)

[a]Pathogens detected in stool included astrovirus, *Campylobacter* spp., *Entamoeba histolytica*, *Giardia* spp., GII norovirus, and pathogenic *E. coli*. [b]Bivariate models with prevalence ratios (PR) presented. [c]Adjusted for neighborhood and hygiene status, per previous CMC survey. [d]Estimated for child with associated stool sample only.

Aim 3 (Vellore, India)

Methods: Model for mixed-effects logistic regression (logistic regression with a random intercept)

The model used for mixed-effects logistic regression analysis (for any (pooled) infection, as an example) was the following:

$$\text{Infection} = X_1\beta_1 + X_{\text{SW monsoon}}\beta_{\text{SW monsoon}} + X_{\text{NE monsoon}}\beta_{\text{NE monsoon}} + X_{\text{type of stool}}\beta_{\text{type of stool}} + Z_{\text{stool-child}}\gamma + \varepsilon$$

Where:

- Infection is 1/0, based on all specimens collected from children in SaniPath households.
- $X_n\beta_n$ refer to estimate-variable combinations for the variables in the model (in the example given, the first term would be some sanitation variable tested while the subsequent 3 terms are added in each model to control for season and type of stool collected (routine vs. diarrheal).
- Z is a matrix where each column is one child and each row is one stool specimen. If a stool specimen belongs to that child, that row/column combination will get a "1", otherwise a "0".
- γ is a column vector of children, assumed to have a normal distribution with a mean of 0 and a variance of "G", where is only a 1 x 1 matrix in this case (because we have only a random intercept) and thus is the variance of the random intercept.
- ε is the variance-covariance matrix of residuals, which are (in this case) unstructured.

Burden of enteric infection

To describe the burden of enteric infection during the first two years of life, histograms were constructed displaying the percent of stool collection, per child, that was positive for a given group of enteropathogens (X-axis) and the count of children (Y-axis), divided by the type of stool collected (all stool—including both routine and diarrheal stool—vs. routine stool

collection only) (Figures A3.1-A3.5). Incidence of enteric infection in general (i.e. not stratified by a particular type of enteropathogen) was high (Figure A3.1). For most children, at least half of stool specimens collected were positive for at least one enteropathogen (average of 68% of stool specimens per child were positive). The lone child showing 0% positive stools dropped out before the first year completed. Bacteria were the primary cause of infection (Figure A3.2), while children showed much lower burden of parasitic and viral infections (Figures A3.3 and A3.4), though these infections contributed to diarrhea (shown by the differences in red vs. green bars). Consistencies between the distribution of any infection and that of bacteria reflect that the majority of infections were bacterial. Parasitic infections were more prevalent than viral infections, reflected in the larger right-skew to the distribution (see differences in X-axis units), though the shape of their distributions was similar. This similarity was likely because both parasitic and viral infections are shed for longer time periods, on average, than bacteria in stool. Thus, though number of unique infections may have been much less than that of bacteria, single infections with parasites or viruses may have been detected longer in stool from a single infection (especially given the frequency of monthly stool collection and the significant contribution of viral infections to diarrheal stool collection), causing select children to have high overall detections. Burden of coinfections was relatively high as well, with less than 25 children out of the 230 having no evidence of coinfection in stool specimens during follow-up (Figure A3.5).

Figure A3.1: Enteric infection burden during the first two years of life. The percent of stool positive for at least one enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when all stool specimens (both routine and diarrheal) are presented, while green bars indicate the distribution when only routine stool is presented.

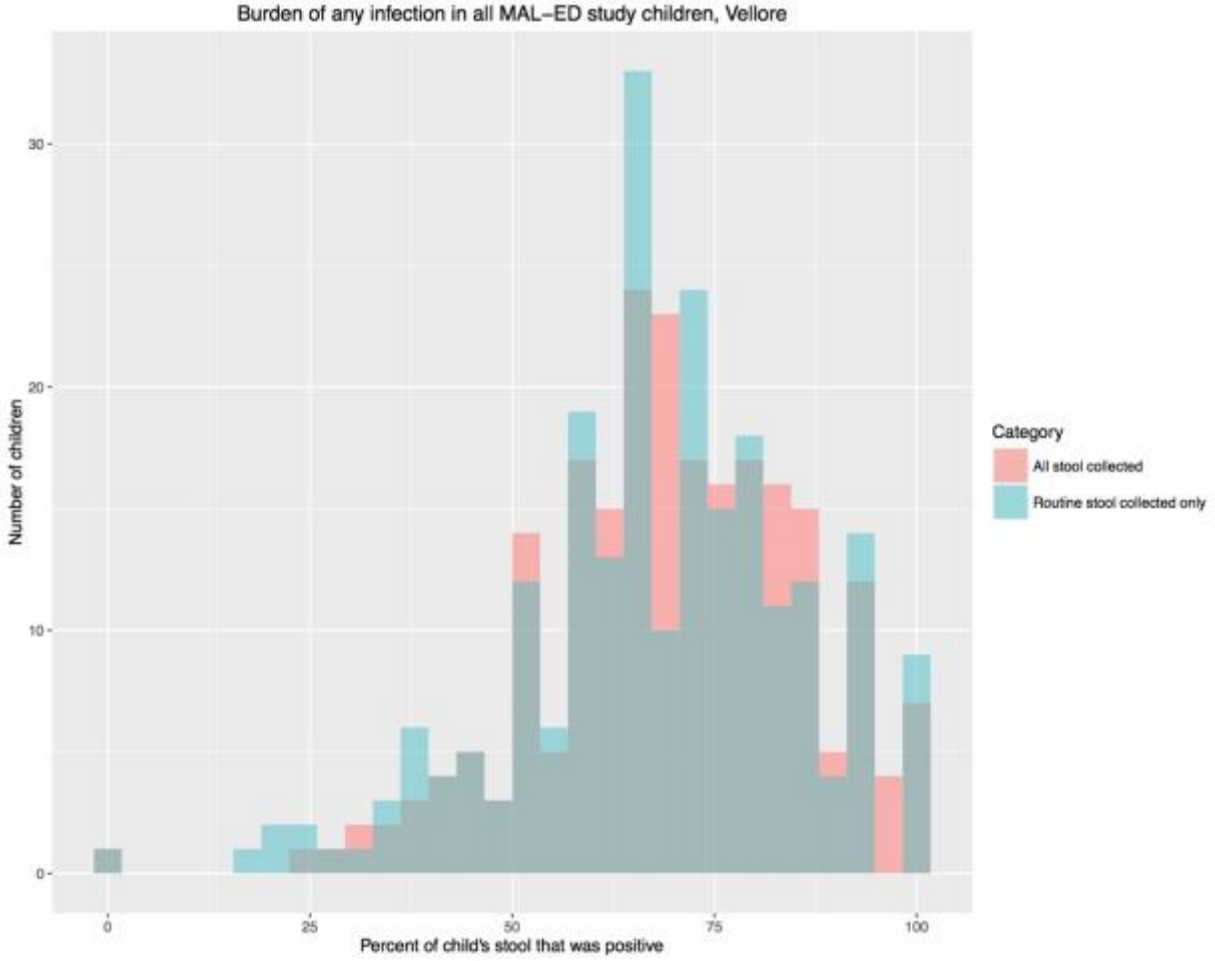


Figure A3.2: Bacterial enteric infection burden during the first two years of life. The percent of stool positive for at least one bacterial enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when all stool specimens (both routine and diarrheal) are presented, while green bars indicate the distribution when only routine stool is presented.

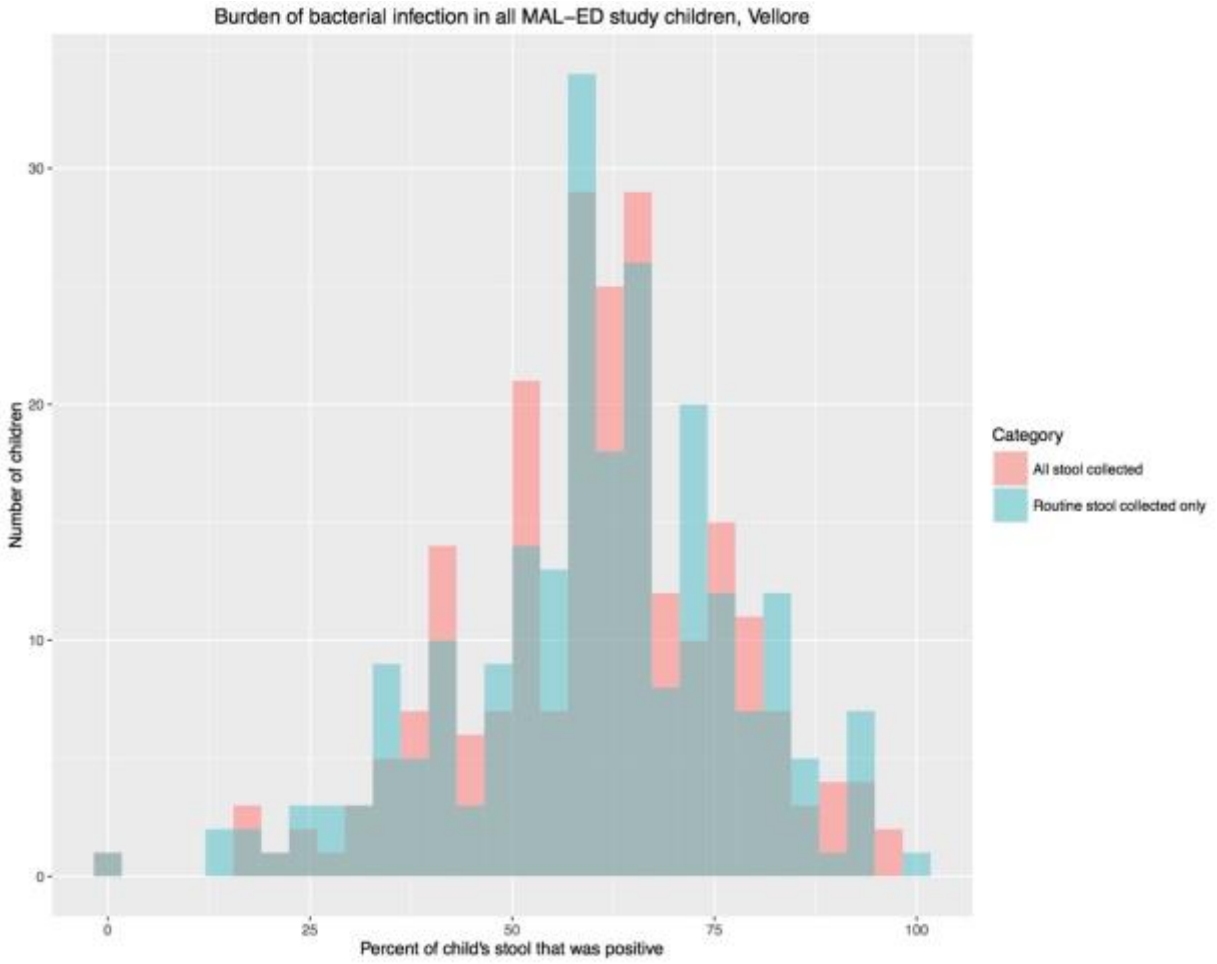


Figure A3.3: Parasitic enteric infection burden during the first two years of life. The percent of stool positive for at least one parasitic enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when all stool specimens (both routine and diarrheal) are presented, while green bars indicate the distribution when only routine stool is presented.

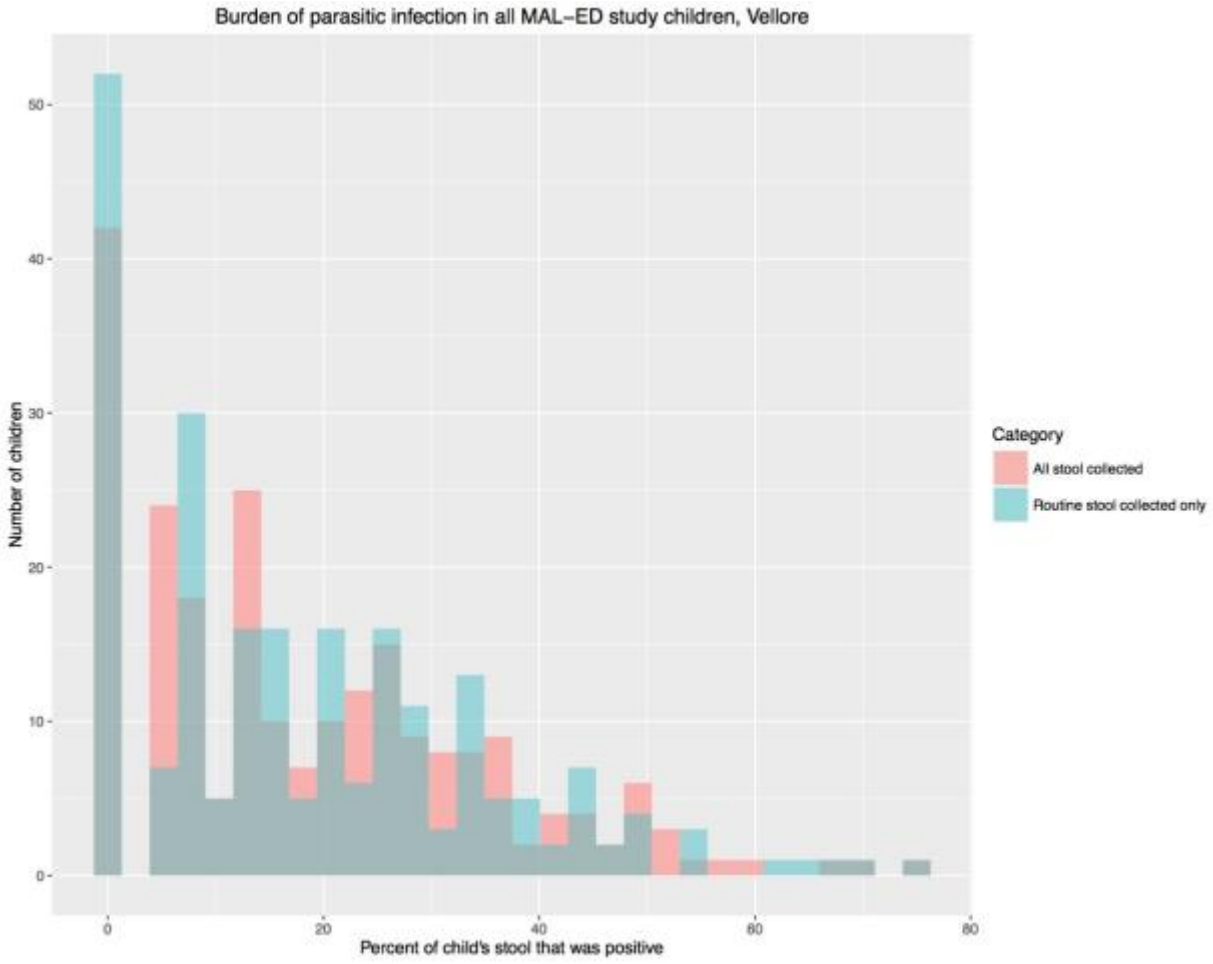


Figure A3.4: Viral infection burden during the first two years of life. The percent of stool positive for at least one viral enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when all stool specimens (both routine and diarrheal) are presented, while green bars indicate the distribution when only routine stool is presented.

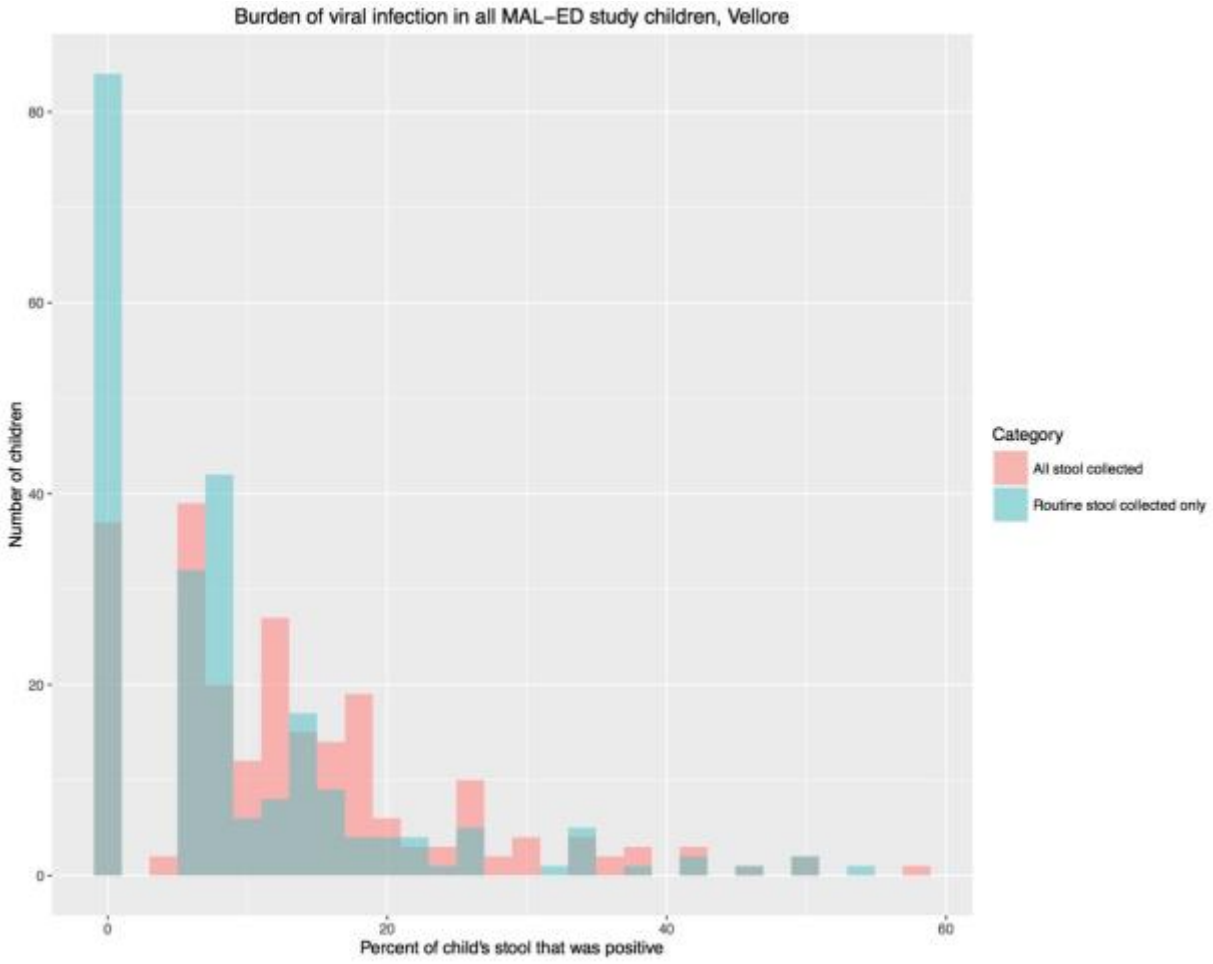
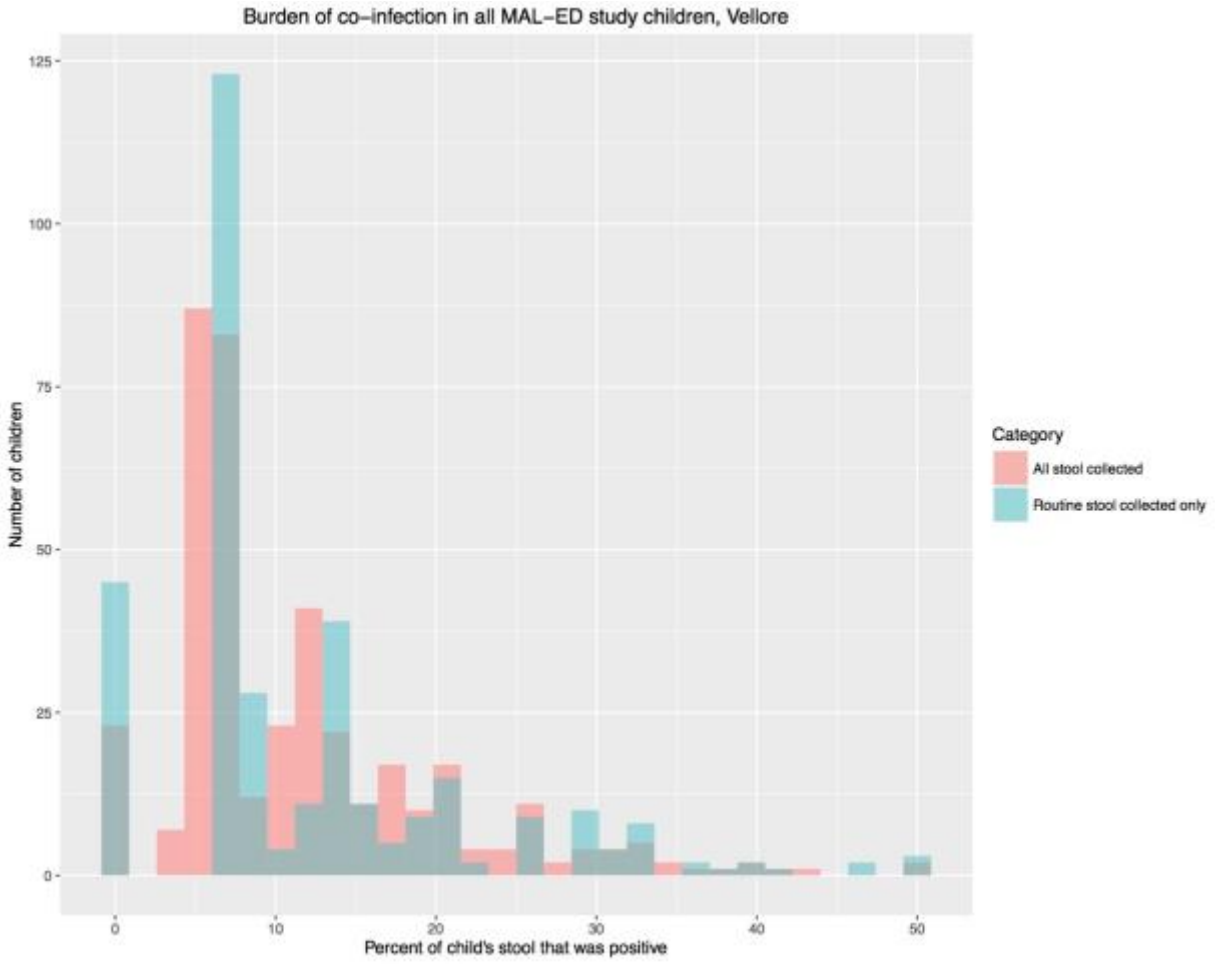


Figure A3.5: Burden of coinfection during the first two years of life. The percent of stool positive for two or more enteric pathogens is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when all stool specimens (both routine and diarrheal) are presented, while green bars indicate the distribution when only routine stool is presented.



Differences in burden of enteric infection by follow-up period

To assess differences in burden of enteric infection by follow-up period, histograms were created to show burden by the total follow-up period (two years, Figures A3.6-A3.10, red bars) and by the first year of follow-up (Figures A3.6-A3.10, green bars). Likely because most (12) specimens that were routinely collected were collected monthly during the first year and only a

few (about 4) specimens were collected the second year, there were few discrepancies between the burdens of infection during the first year and the overall burden. Though burdens for each type of infection shifted right (higher burden) when looking at the overall burden, compared with that during the first year, parasitic infection burden (Figure A3.8) appeared to shift the most, suggesting more infections during the second year of follow-up than during the first.

Figure A3.6: Burden of enteric infection by follow-up period. The percent of stool positive for at least one enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when stool specimens from both years of follow-up are presented, while green bars indicate the distribution when only the first year of follow-up (when monthly sample collection took place) is presented.

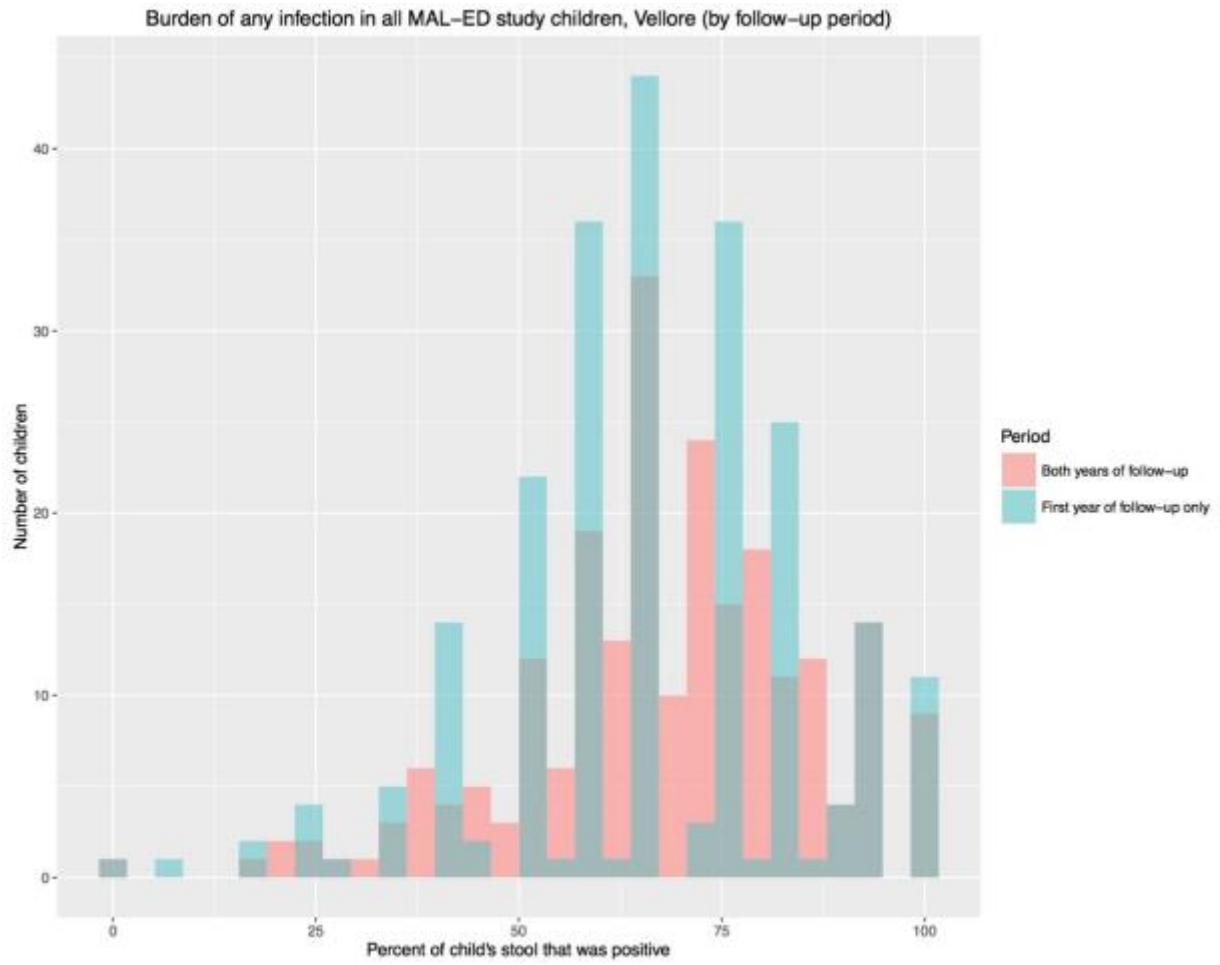


Figure A3.7: Burden of bacterial enteric infection by follow-up period. The percent of stool positive for at least one bacterial enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when stool specimens from both years of follow-up are presented, while green bars indicate the distribution when only the first year of follow-up (when monthly sample collection took place) is presented.

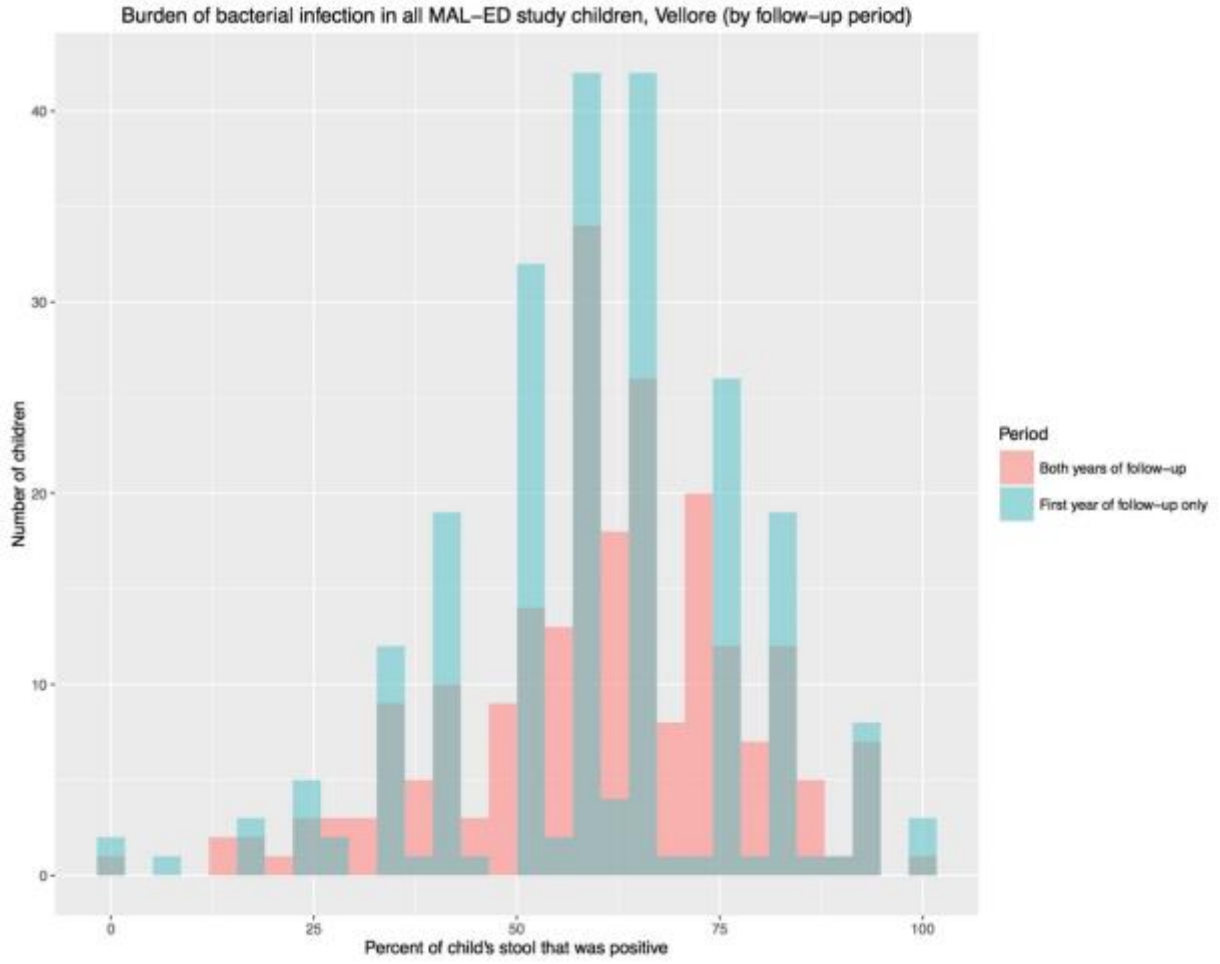


Figure A3.8: Burden of parasitic enteric infection by follow-up period. The percent of stool positive for at least one parasitic enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when stool specimens from both years of follow-up are presented, while green bars indicate the distribution when only the first year of follow-up (when monthly sample collection took place) is presented.

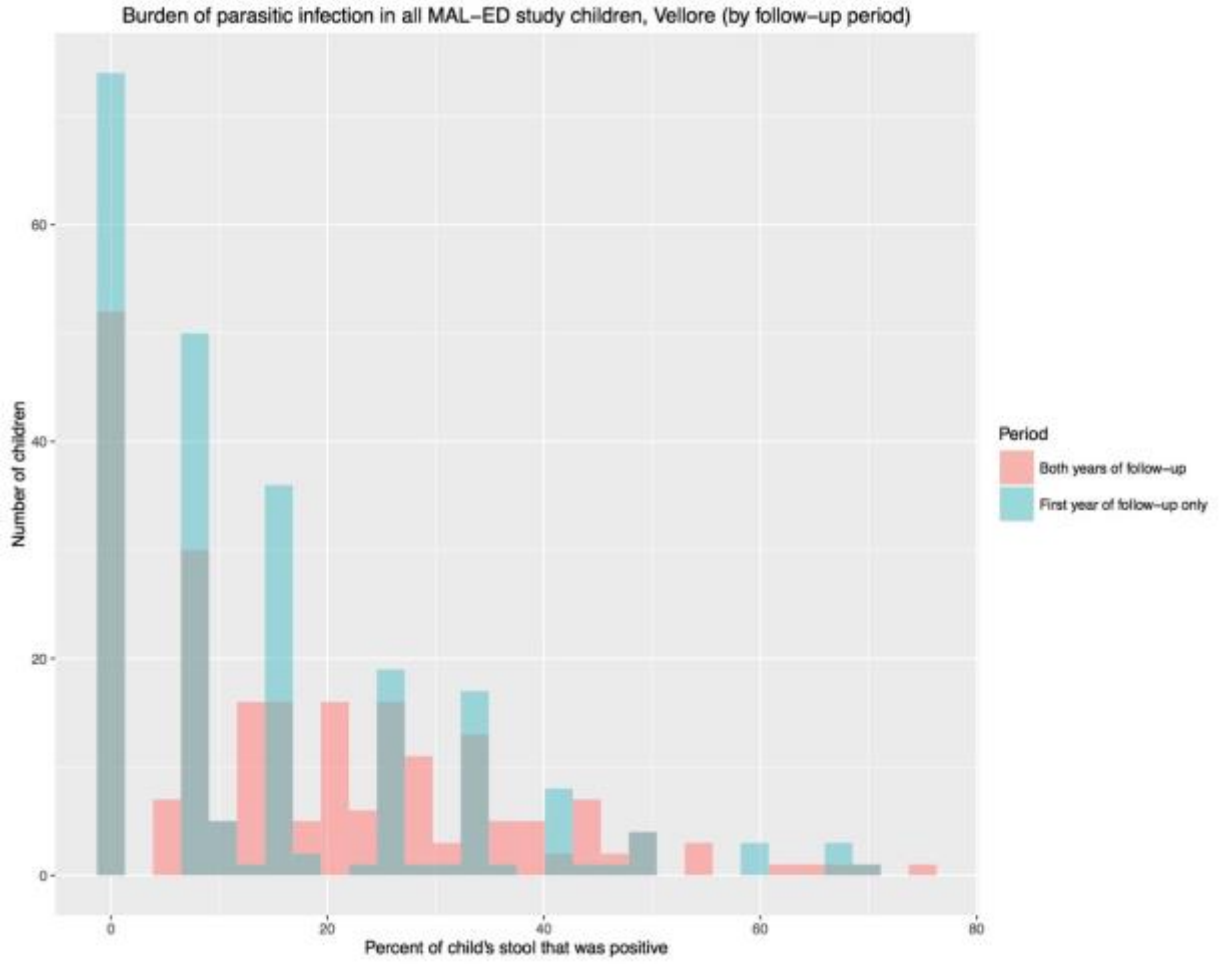


Figure A3.9: Burden of viral enteric infection by follow-up period. The percent of stool positive for at least one viral enteric pathogen is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when stool specimens from both years of follow-up are presented, while green bars indicate the distribution when only the first year of follow-up (when monthly sample collection took place) is presented.

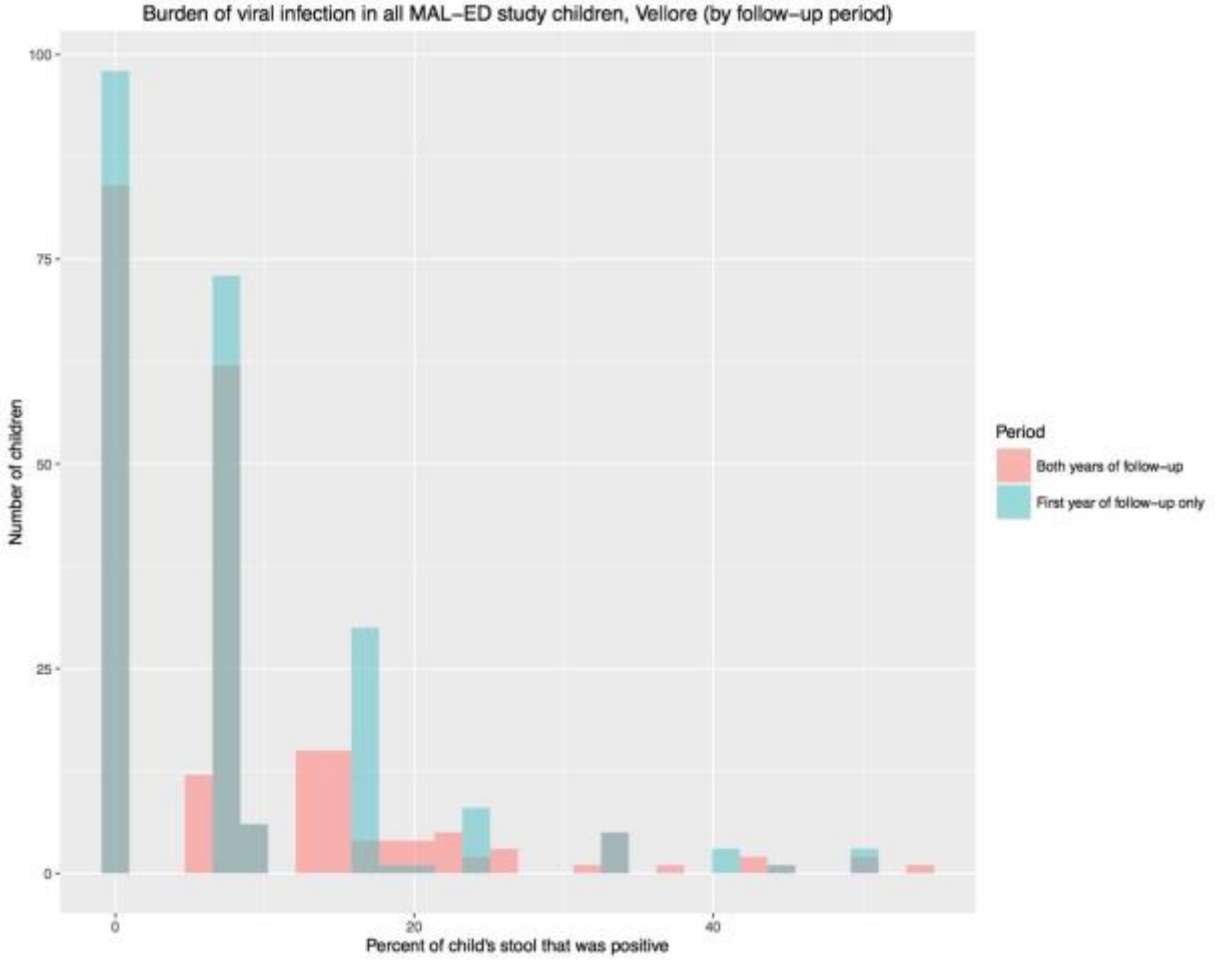
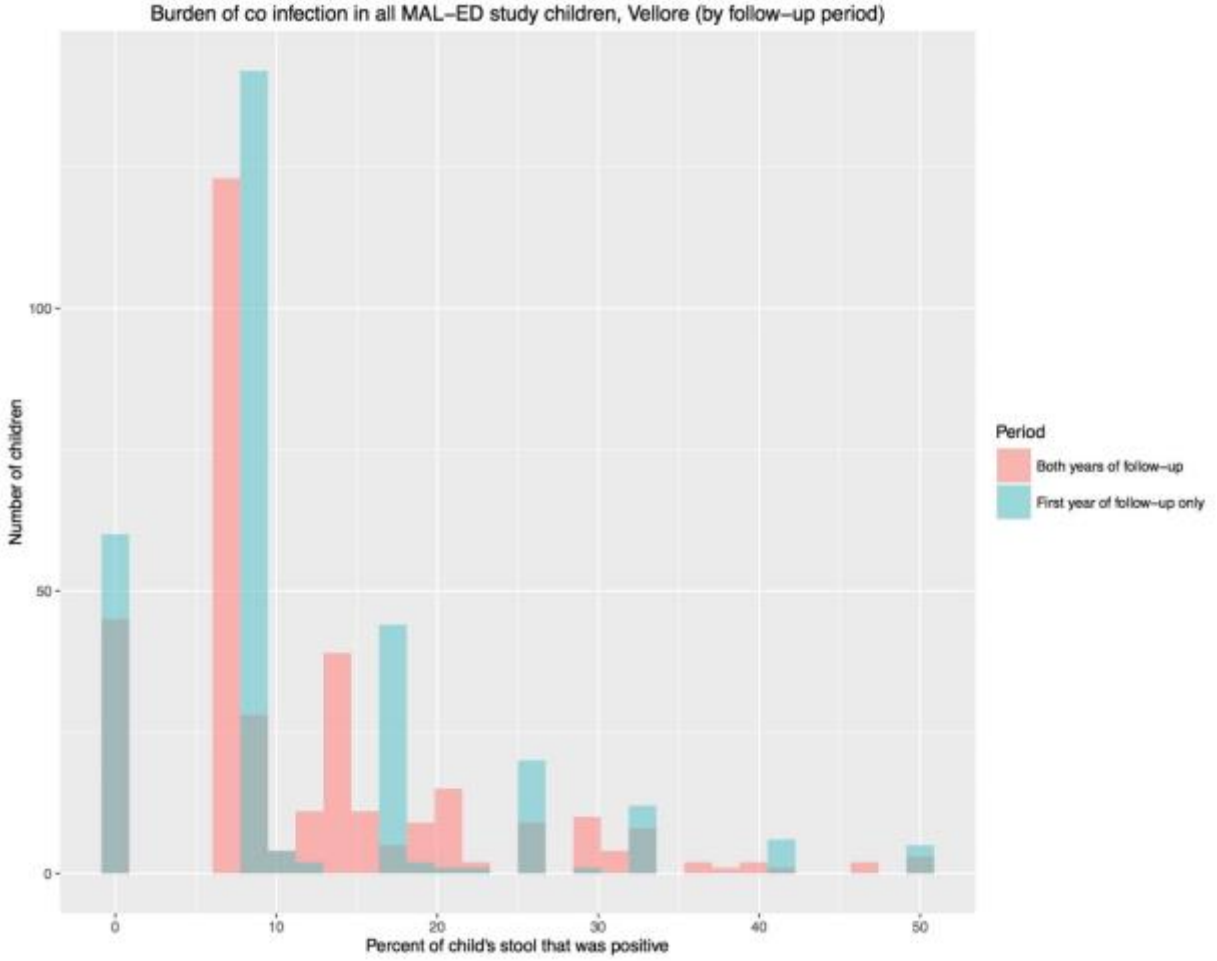


Figure A3.10: Burden of coinfection by follow-up period. The percent of stool positive for two or more enteric pathogens is shown, for each child, along the X-axis. Counts (number of children with each percentage) are on the Y-axis. Red bars indicate the distribution when stool specimens from both years of follow-up are presented, while green bars indicate the distribution when only the first year of follow-up (when monthly sample collection took place) is presented.



Incidence of enteric infection during the first two years of life

To describe seasonality of enteric infection during the 2010-2014 study period, percentages of positive stool (including both routinely-collected stool and diarrheal stool) were plotted, by month, in histograms (Figures A3.11-A3.15). Of note, enrollment in the MAL-ED study was on a rolling basis, and thus was highest in 2011-2013, thus the data for months in 2010 and 2014 are sparse. Overall, general (pooled) enteric infection incidence and bacterial infection

incidence did not show any discernable pattern or seasonality (Figures A3.11 and A3.12, respectively). Though low overall, percentages of stool specimens positive for parasites appeared to increase over time, showing no seasonal patterns (Figure A3.13). Percentages of stool specimens positive for viruses appear to cycle every 2-3 months, with high values around the end and middle of the year and lower values in-between (Figure A3.14). Counts of diarrheal events appeared to peak around these times as well, though showed less defined cyclical patterns (Figure A3.15). Across organisms, percent of positive stool specimens per month tended to increase with time, especially over the first 1-2 years of follow-up, which likely reflects increases in the number of enrolled children being followed (increases in both the denominator and numerator due to new infections). Further, individual children likely experienced increased infections due to increased fecal exposures as they aged, especially increases in parasitic infection (Figure A3.13), since they stopped breastfeeding and were more mobile outside of their household environment.

Figure A3.11: Seasonality of enteric infection (percent of positive stool by month). Follow-up time is shown on the X-axis. Percentages of stool specimens positive for at least one enteric pathogen are shown on the Y-axis. Each bar represents an individual month of follow-up (stool specimens were aggregated by month).

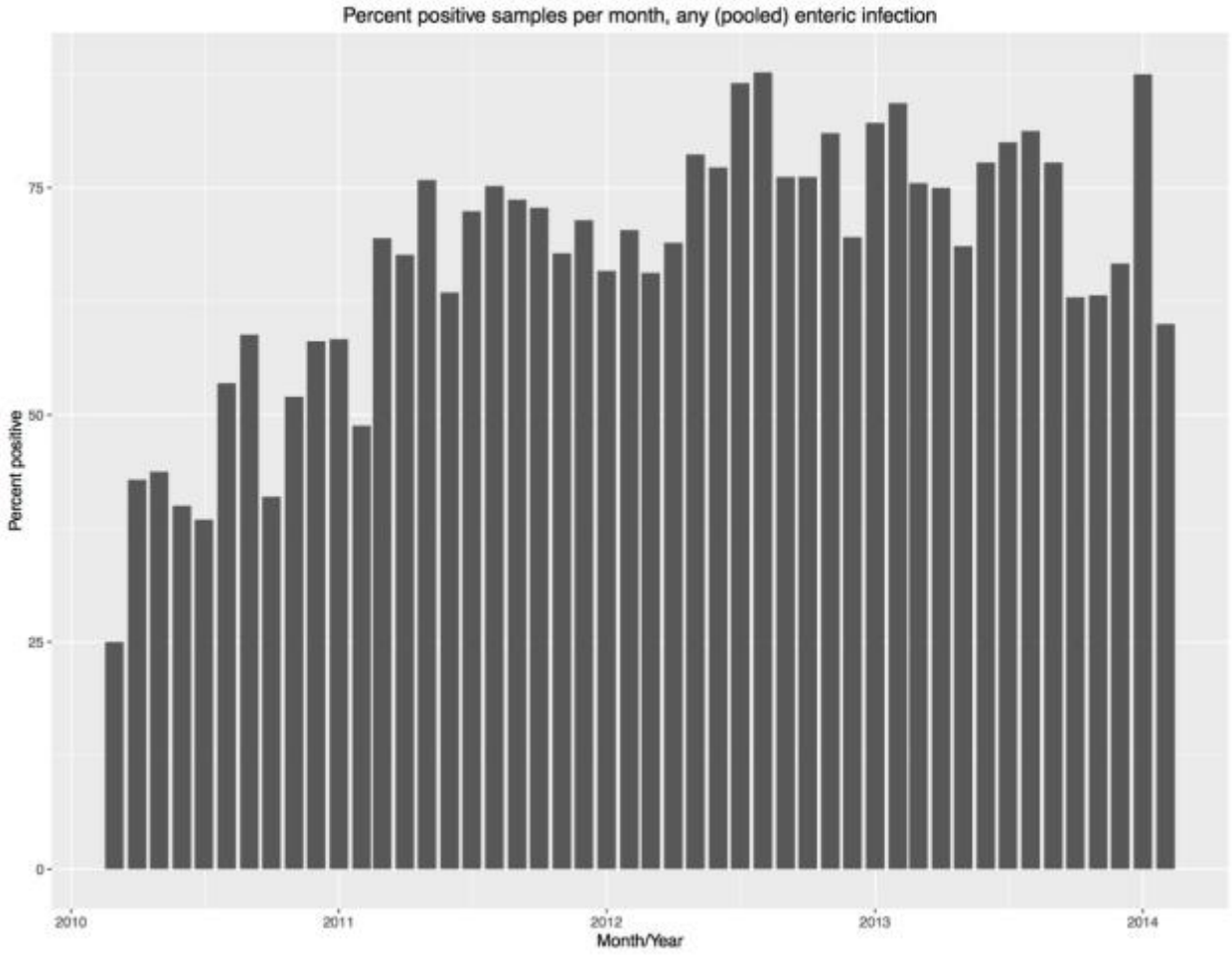


Figure A3.12: Seasonality of bacterial enteric infection. Follow-up time is shown on the X-axis. Percentages of stool specimens positive for at least one bacterial enteric pathogen are shown on the Y-axis. Each bar represents an individual month of follow-up (stool specimens were aggregated by month).

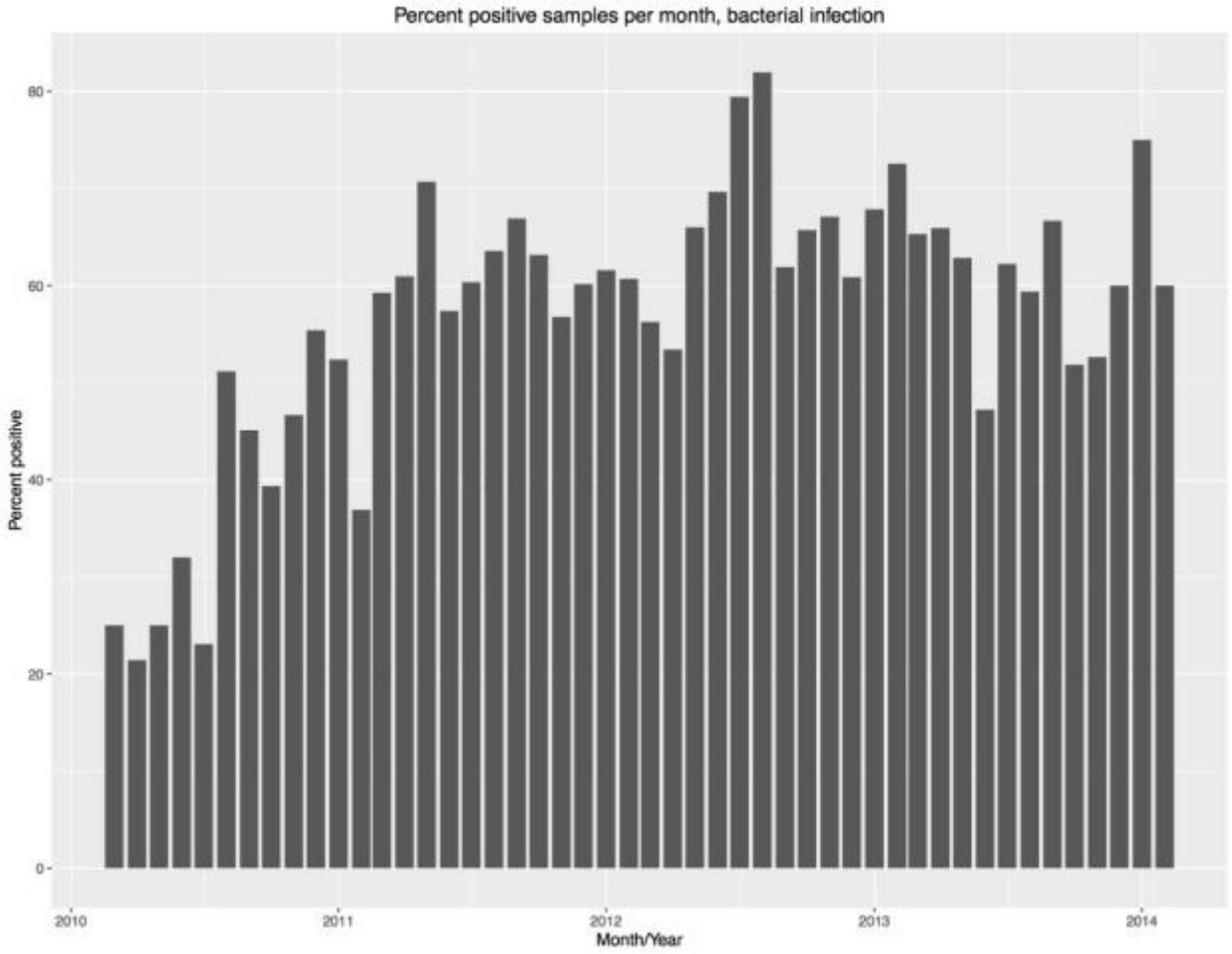


Figure A3.13: Seasonality of parasitic enteric infection. Follow-up time is shown on the X-axis. Percentages of stool specimens positive for at least one parasitic enteric pathogen are shown on the Y-axis. Each bar represents an individual month of follow-up (stool specimens were aggregated by month).

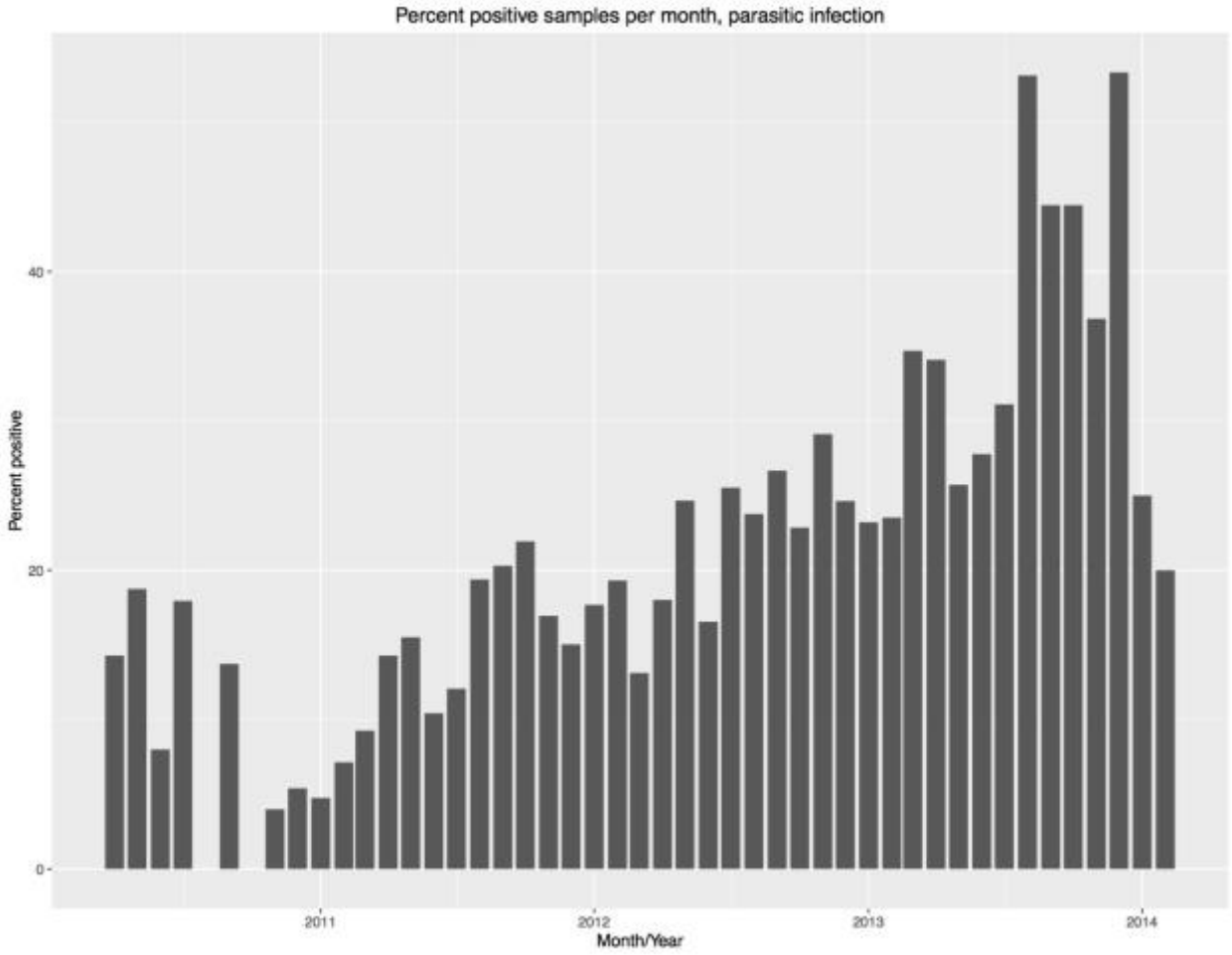


Figure A3.14: Seasonality of viral enteric infection. Follow-up time is shown on the X-axis. Percentages of stool specimens positive for at least one viral enteric pathogen are shown on the Y-axis. Each bar represents an individual month of follow-up (stool specimens were aggregated by month).

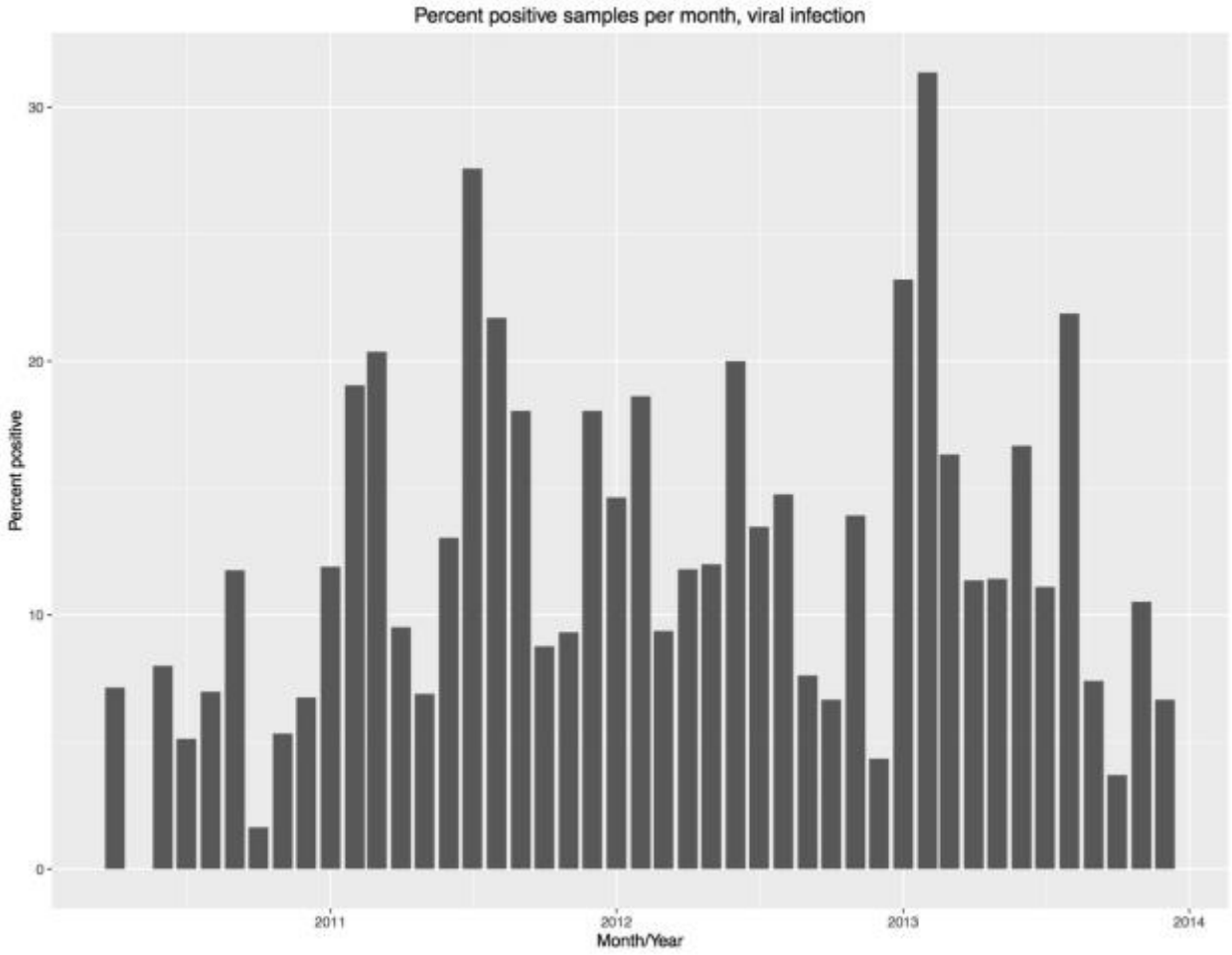
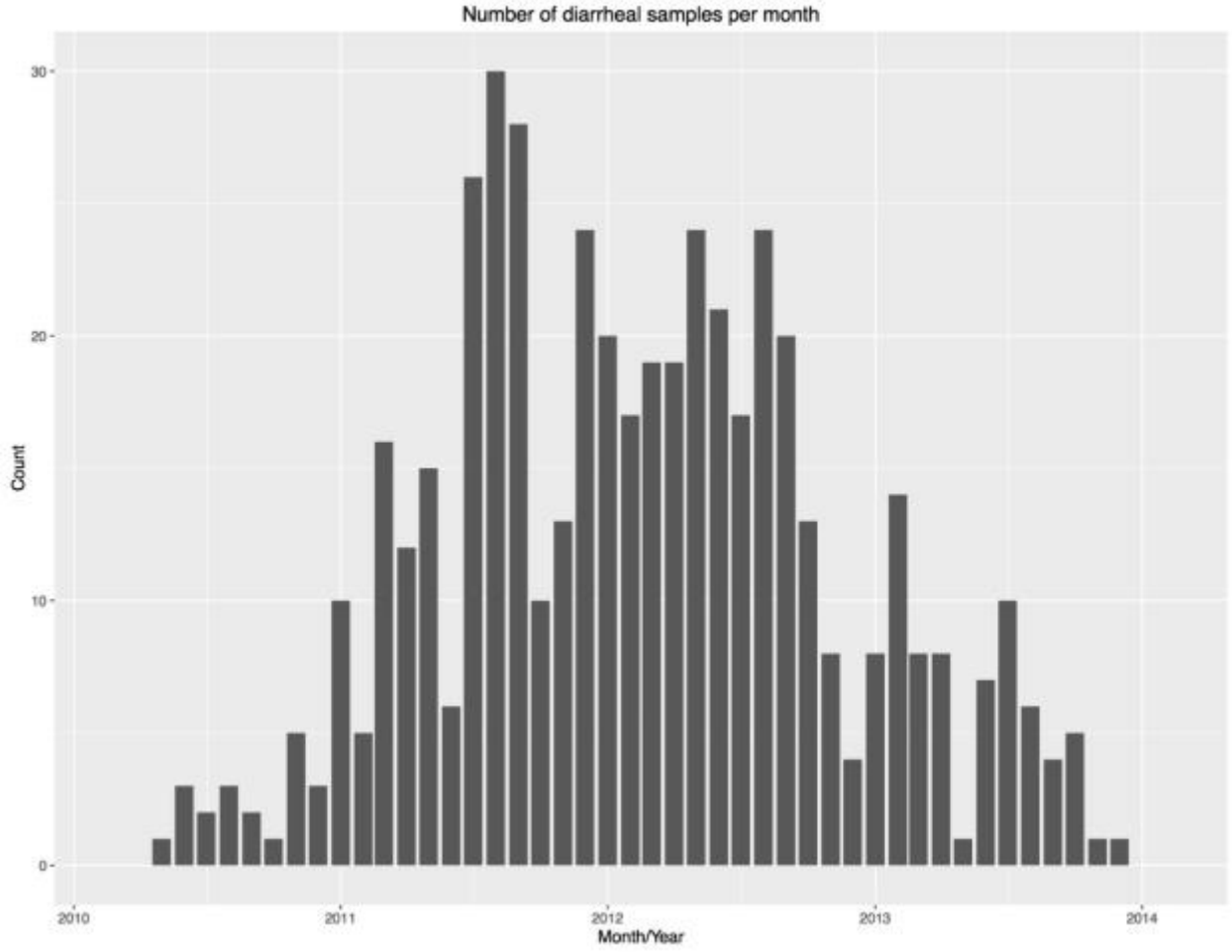


Figure A3.15: Seasonality of diarrhea (count of events per month). Follow-up time is shown on the X-axis. Number of diarrheal samples collected each month is shown on the Y-axis. Each bar represents an individual month of follow-up (analysis was aggregated by month).



Tests of global autocorrelation, local autocorrelation, and spatial clustering for burden of enteric infection

To determine whether burden (number) of enteric infections was autocorrelated or clustered in space, Global Moran's I, Anselin's Local Moran's I, and local Getis GI* tests were conducted on the percentage of stool specimens positive for enteric infection (as well as tests of pathogen group-specific outcomes) during the first year of follow-up and on counts of diarrhea. No significant positive or negative global autocorrelation was detected by Global Moran's I test. Local Moran's I tests detected a small grouping (about 5 households) of high burden of bacterial

infection during the first year (high-high grouping) near the open defecation field, though no other significant multi-household groups were detected for bacterial infection, as well as other types of infections tested, including general (pooled) enteric infection. Local Getis GI* did not detect any significant clusters of high or low infection burden as well.

Tests of global autocorrelation for residuals of final models in Aim 3

Residuals of sanitation exposure-enteric infection outcome relationships were tested by Global Moran's I in ArcMap 10.1. Using data from all time points, none of the models (bacterial, parasitic, or viral) showed evidence of significant positive spatial autocorrelation. Parasitic infection models, using this approach, showed significant negative autocorrelation ($Z = -3.31$, $p < 0.001$). Model residuals were also compared using the average of each residual at each location. Using this method, none of the models showed evidence of significant spatial autocorrelation (positive or negative).

Spatial network-based clustering analysis of enteric infection

Methods

Network-based clustering analysis of enteric infection, based on MAL-ED stool data from all MAL-ED households in the SaniPath study area and maps of the drainage network provided by CMC, was completed using GeoDaNet (Arizona State University, Tempe, AZ, USA). Global and local K functions, as well as local indicators of network cluster (LINCs), were assessed. As global and local K-functions are measures of incident cases in space, these K-functions were computed for all households on the network (as a test of "baseline clustering" since households chosen for follow-up in the MAL-ED study represented a subset of all

households). Global K-functions also compared households with greater than 50% prevalence of enteric infection during the first year of follow-up (i.e. households with children who had at least 6 out of 12 monthly stool specimens positive for at least one enteric pathogen) to those with less than 50%. An α of 0.05 was used for all tests of significance.

LINCs included local Moran's I and Getis G* tests based on the prevalence of enteric infections during the first year (i.e. percent of months positive for at least one enteric pathogen during the first year), as well as prevalence of specific infections (bacterial infections, parasitic infections, and viral infections) during the first year. Local Moran's I tests for local autocorrelation (grouping of similar values near one-another) along the network, while Getis G* tests for clustering of high or low values. Prior to computing LINCs on the drainage network, prevalence of enteric infection during the first year (number of months with a positive stool divided by 12) was computed. The drainage network was divided into segments by vertices (intersections). Each household was matched to the single drain segment nearest to it (if within 50m, otherwise the household was excluded from the analysis). Each drain segment was assigned the average prevalence of all households matched to it.

Results:

Global K-functions comparing cases (children with greater than 50% prevalence of infection during the first year) to controls (children with less than 50% prevalence of infection during the first year) showed no significant clustering at distances up to the median distance on the network (800m). Local K-functions showed significant clustering in the drains around the open defecation field at distances up to the median distance between households (300m); however, this clustering was no different from clustering of household locations along the network at baseline (data not shown).

Tests of local Moran's I showed some significant localized autocorrelation in prevalence of enteric infection along the drain network, showing areas where households high prevalence of enteric infection were near to others with similarly high prevalence (Figures A3.16—A3.18). Only high-high autocorrelation is presented, as the distribution of the MAL-ED sample of households relative to the drainage network resulted in large sections with 0% prevalence due to no households present. Testing for local autocorrelation of prevalence of enteric infection showed significant autocorrelation for households along a large stretch of drain segments south and east of the open defecation field (Figure A3.16). The same locations had significant high-high autocorrelation of bacterial infection prevalence during the first year (Figure A3.17). Testing of prevalence of parasitic infection during the first year showed a small significant section of the internal drain network in the center of the neighborhood (Figure A3.18). Prevalence of viral infection during the first year did not show significant network-based autocorrelation (data not shown).

Tests of Getis G* showed little significant local clustering of enteric infection prevalence during the first year of life (Figures A3.19—A3.21). While prevalence of bacterial infection during the first year showed significant local clustering along one of the long stretches of the drain network previously associated with significant autocorrelation (FIGURE A3.20), clusters of prevalence of any enteric infection and that of viral enteric infection during the first year were restricted to small segments of the drain network, representing few households (Figures A3.19 and A3.21, respectively). Prevalence of parasitic infection during the first year of life was not significantly clustered along the drainage network (data not shown).

Figure A3.16: Network-based clustering (local autocorrelation) of enteric infection (with any pathogen) in Old Town using local Moran's I statistic. Areas of significant ($p < 0.05$) local autocorrelation of prevalence (during the first year of life) of enteric infection are shown with red ellipses.

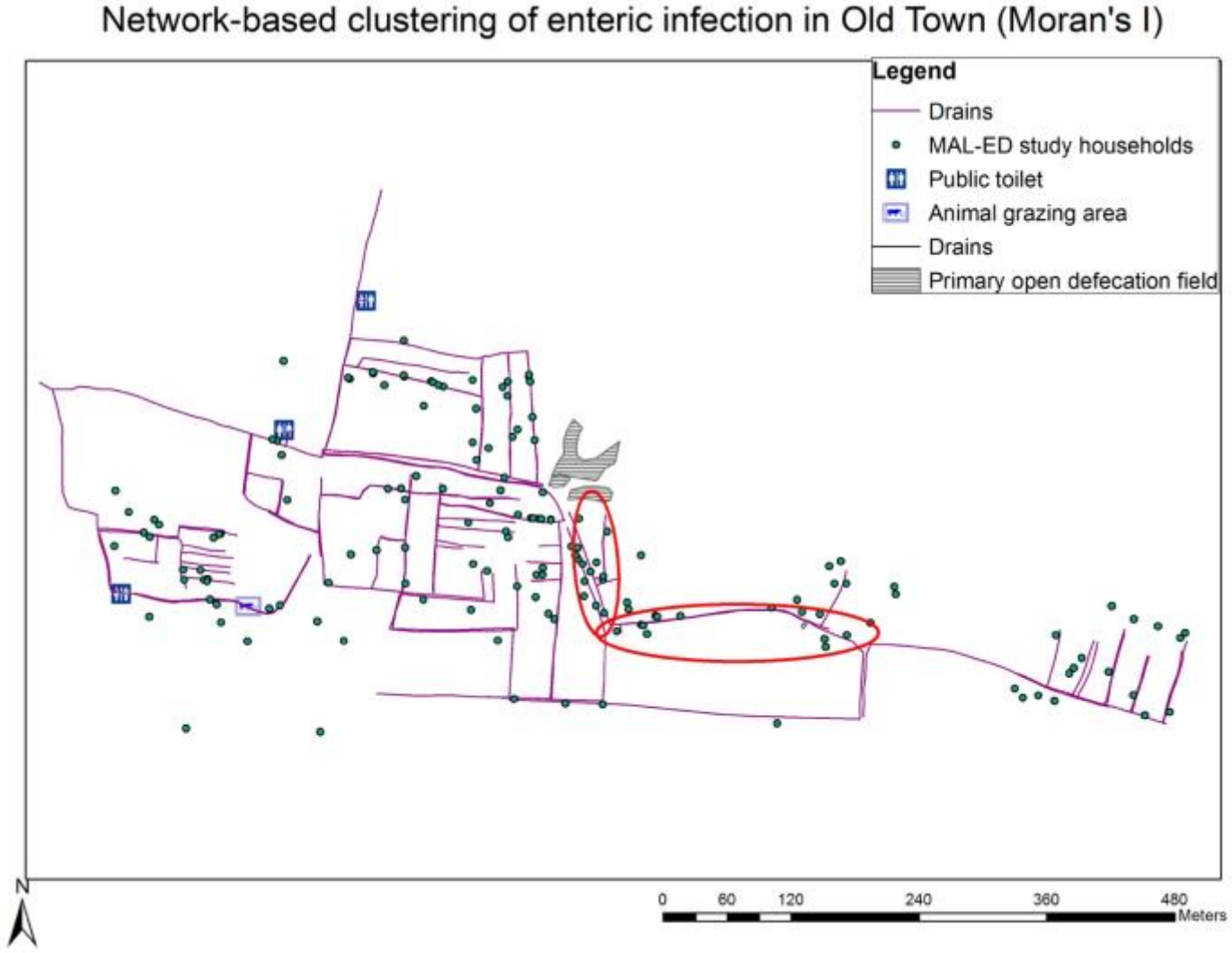


Figure A3.17: Network-based clustering (local autocorrelation) of bacterial enteric infection in Old Town using local Moran's I statistic. Areas of significant ($p < 0.05$) local autocorrelation of prevalence (during the first year of life) of bacterial infection are shown with red ellipses.

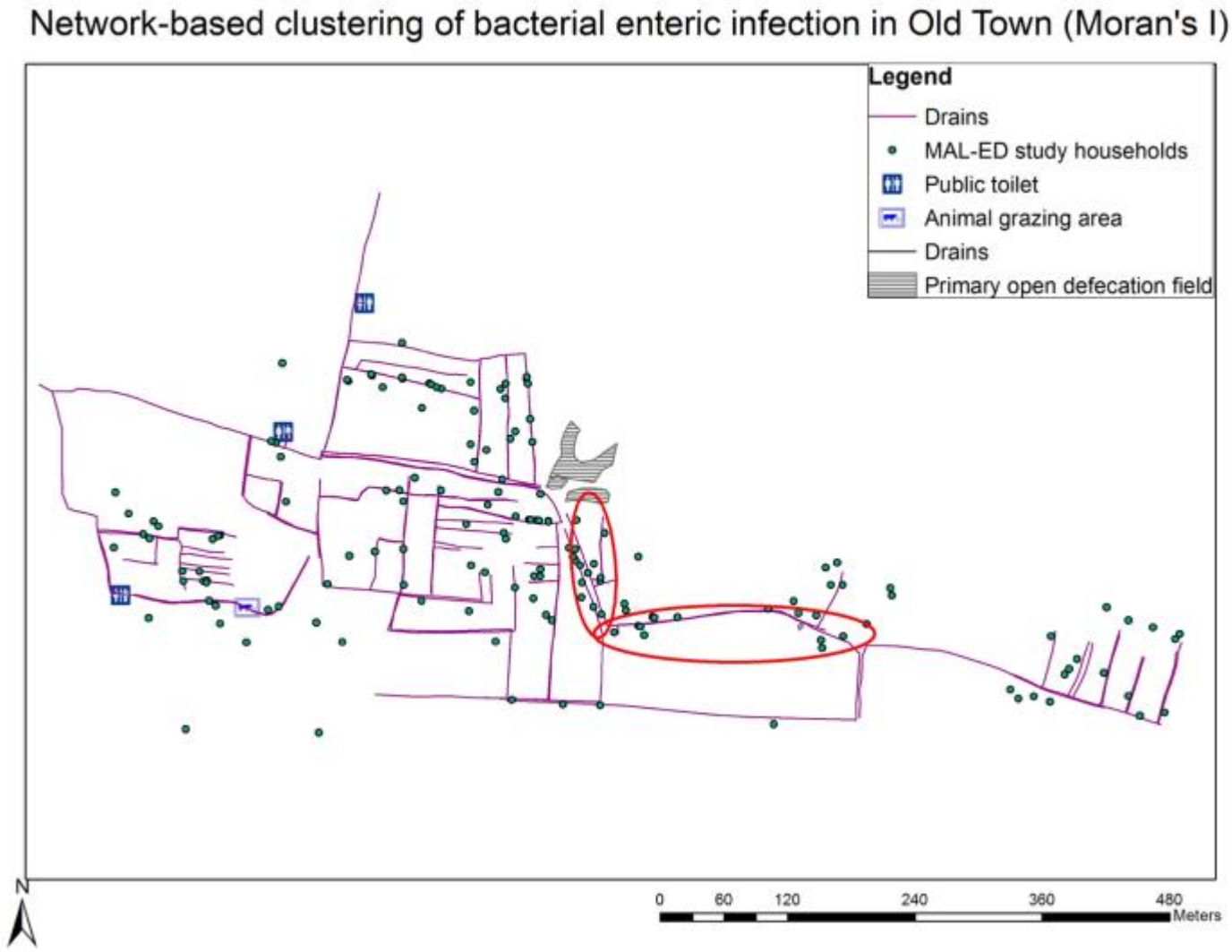


Figure A3.18: Network-based clustering (local autocorrelation) of parasitic enteric infection in Old Town using local Moran's I test. Areas of significant ($p < 0.05$) local autocorrelation of prevalence (during the first year of life) of parasitic infection are shown with red ellipses.

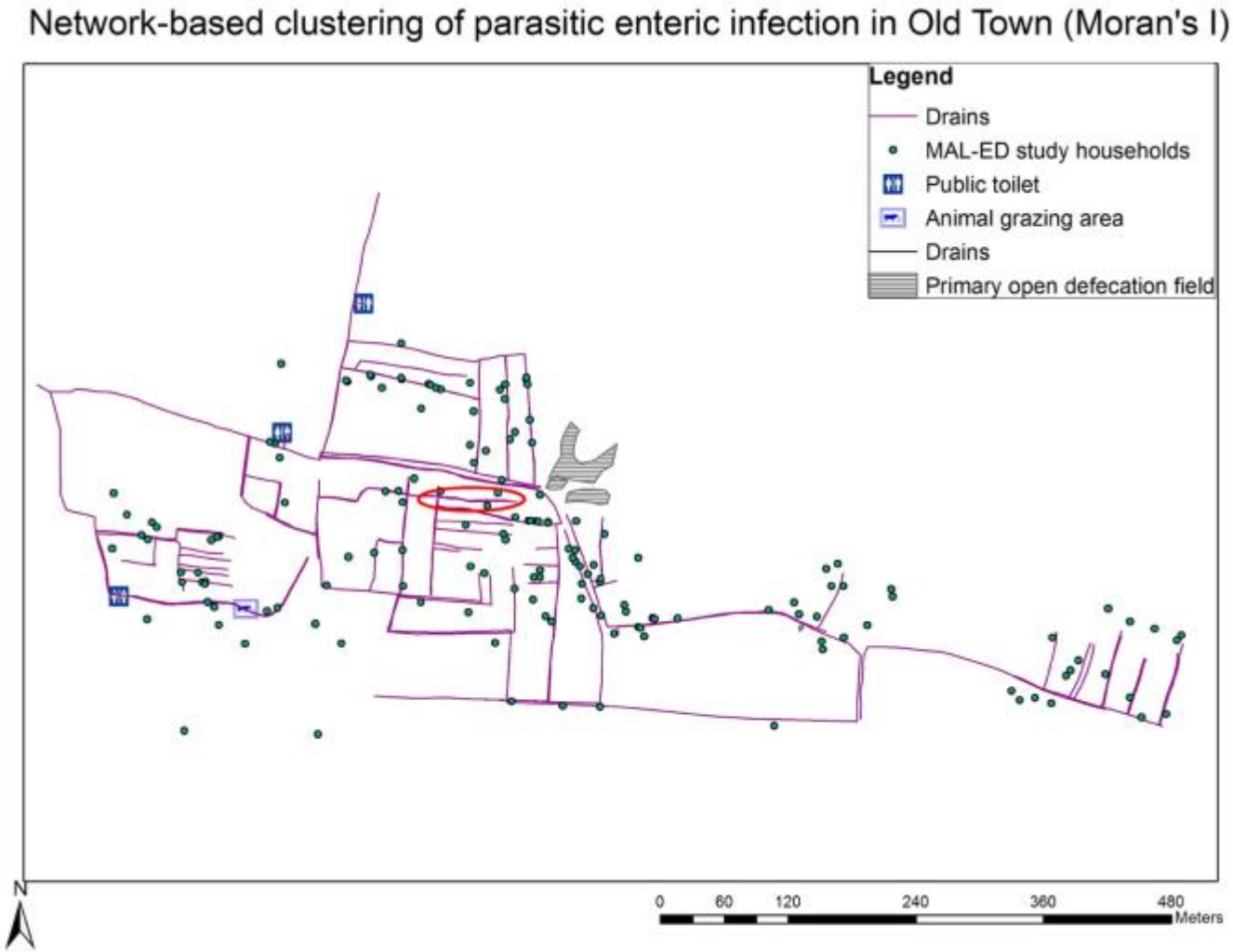


Figure A3.19: Network-based clustering of enteric infection (with any pathogen) in Old Town using Getis G* statistic. Areas of significant ($p < 0.05$) local clustering of prevalence (during the first year of life) of enteric infection are shown with red ellipses.

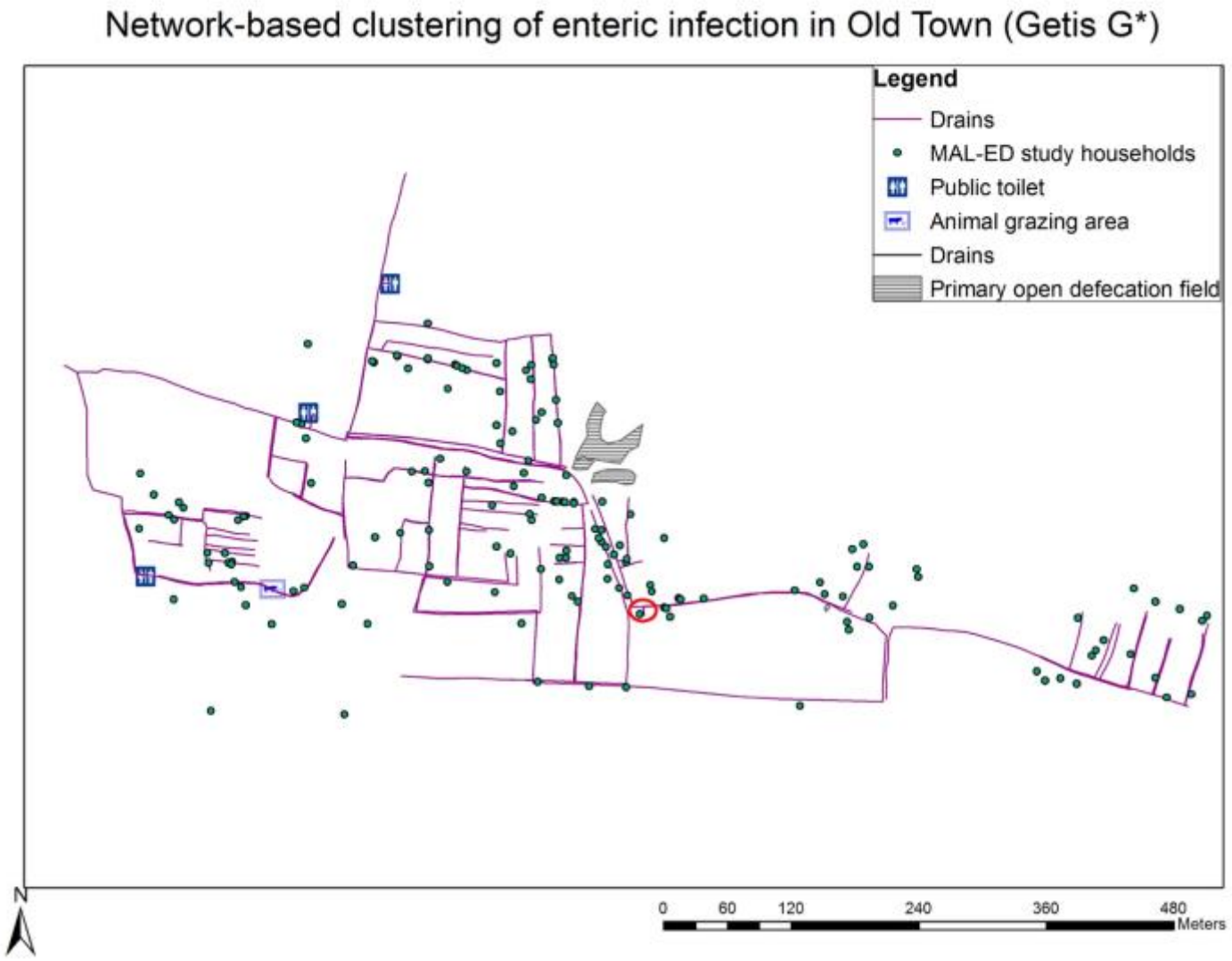


Figure A3.20: Network-based clustering of bacterial enteric infection in Old Town using Getis G* statistic. Areas of significant ($p < 0.05$) local clustering of prevalence (during the first year of life) of bacterial infection are shown with red ellipses.

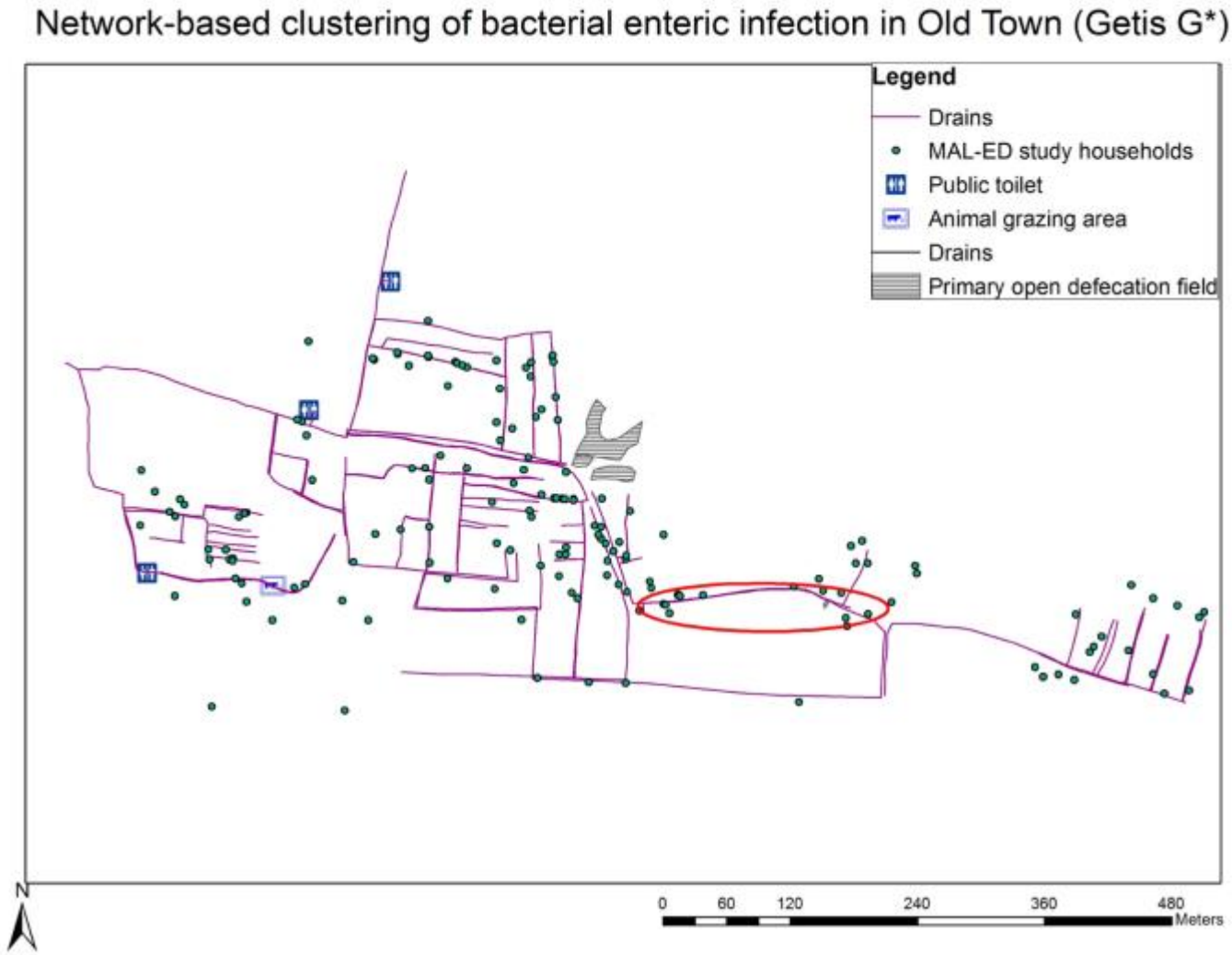
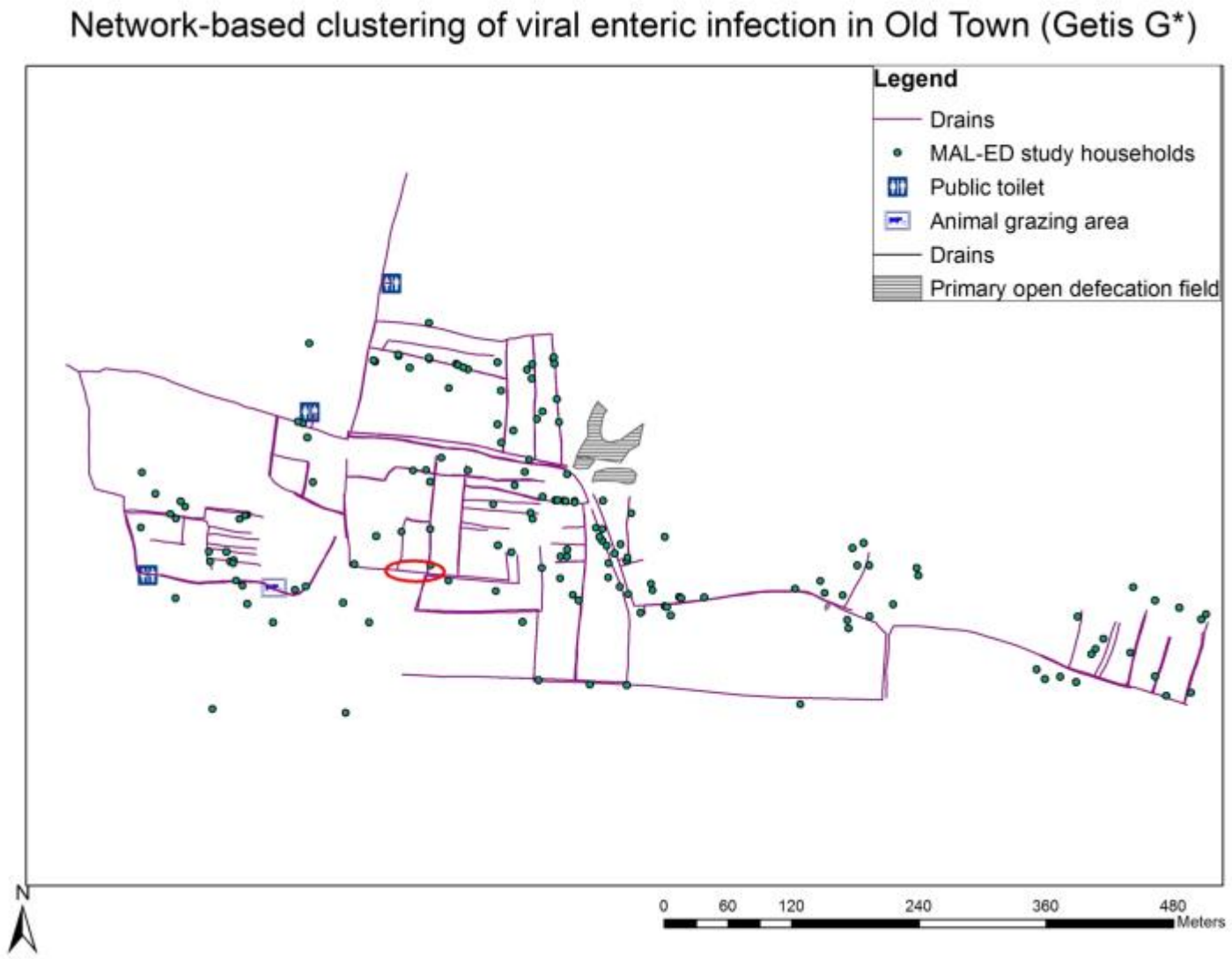


Figure A3.21: Network-based clustering of viral enteric infection in Old Town using Getis G* statistic. Areas of significant ($p < 0.05$) local clustering of prevalence (during the first year of life) of viral infection are shown with red ellipses.



Tests of spatial, temporal, and space-time clustering of coinfections

The same tests were conducted on coinfection data (with the addition of a SaTScan normal test of counts of coinfections during the first year) to assess autocorrelation and clustering. Significant positive or negative global autocorrelation of coinfections during the first year was not observed. Local Moran's I testing reveal no consistent high-high areas of more than one household, and local Getis GI* statistics were not significant. SaTScan testing revealed no significant high or low incidence clusters in the neighborhood as well.

To test for temporal variation in coinfections, a mixed-effects logistic regression model was constructed for coinfection (yes/no) as the outcome, testing the southwest and northeast monsoon seasons as individual predictors, controlling for stool type. There was no significant variation in coinfection risk by season; however, diarrheal stool was significantly associated with coinfection, when compared with routine stool collected (OR: 3.18, 95% CI: 2.33-4.35).

To test for space-time clustering of coinfections, SaTScan tests on coinfection incidence by year (2011 or 2012) were conducted. In 2011, a significant most-likely cluster of low coinfection incidence (5%) was observed in 45 households from January to June, while a significant most-likely cluster of high coinfection incidence (79%) was observed among 8 households from May to October (Figure A3.22). In 2012, a significant most-likely cluster of high coinfection incidence (86%) was observed from January to June (Figure A3.23).

Figure A3.22: Space-Time Clustering of Coinfections, Old Town, 2011. Areas of significant space-time clustering (using Kulldorff's space-time scan, $p < 0.05$) are designated with ellipses. Type of cluster (high or low) and months when incidence of coinfection was clustered are shown in the figure legend. All months are for the 2011 calendar year.

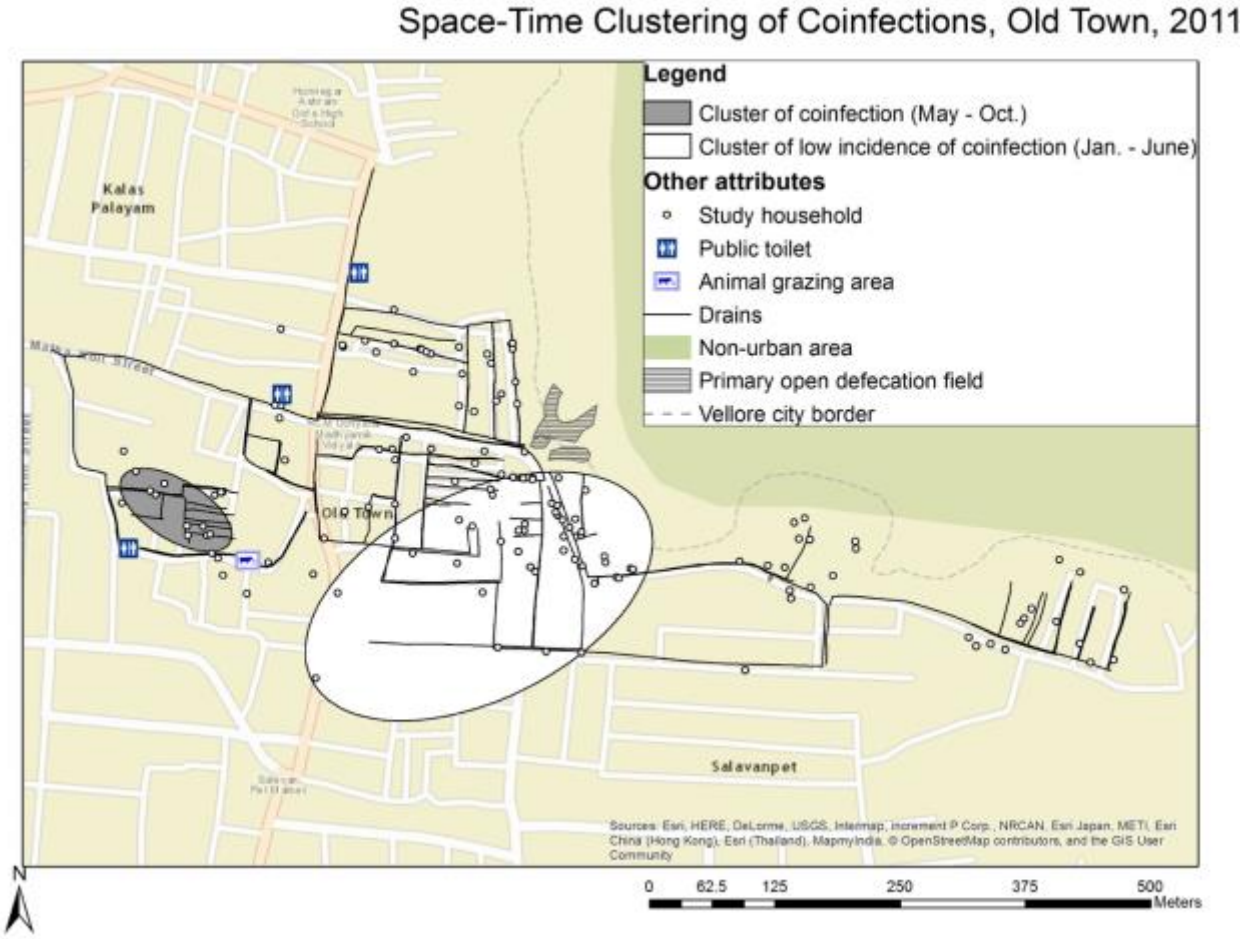
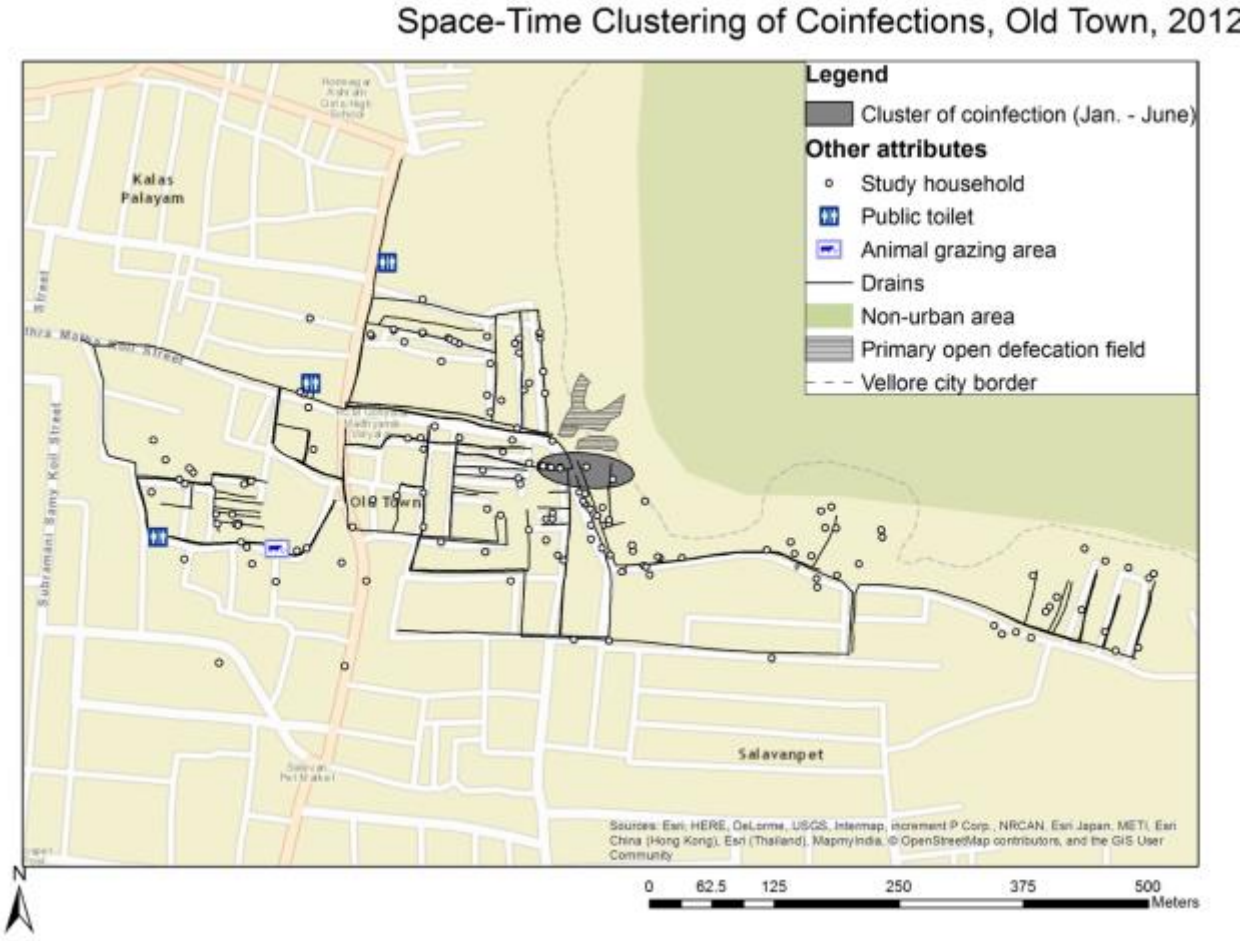


Figure A3.23: Space-Time Clustering of Coinfections, Old Town, 2012. Areas of significant space-time clustering (using Kulldorff's space-time scan, $p < 0.05$) are designated with ellipses. Type of cluster (high or low) and months when incidence of coinfection was clustered are shown in the figure legend. All months are for the 2012 calendar year.



Tests of spatial autocorrelation among mixed-effects logistic regression models

To identify whether spatial processes needed to be accounted for in mixed-effects logistic regression models in Aim 3, tests of global autocorrelation (Global Moran's I) were conducted on residuals from all time points in the final model, as well as on the average of the residuals per child (location). Tests on both sets of residuals showed no evidence of significant positive spatial autocorrelation among residuals.

Modeling of symptomatic stool

Methods:

Mixed effects Poisson regression models were used to determine whether variables associated with risk of "any enteric infection" in Aim 3 (Tables 3-5) were associated with risk of diarrheal (symptomatic) stool specifically. To do this, diarrheal stool was separated from routinely-collected stool and analyzed on its own (i.e. Poisson model predicting number of diarrheal events).

Results:

Presence of an older sibling, presence of a household toilet, residence in the drain flooding cluster (including during the northeast monsoon), any reported flood water contact, open defecation, and water treatment were tested in separate models, controlling for season. None of the variables tested were significantly associated with diarrheal stool (data not shown).

Discussion:

Given that these variables were significant risk factors for pooled enteric infection incidence, but not diarrheal events alone, this finding has significant implications for the interpretation of results of past research. For example, the observation that children in households with toilets had significantly lower risk of enteric infection, but not diarrhea alone, suggests that the failure to detect a health impact in recent randomized-controlled trials of sanitation in India may have been due to issues with the choice of health outcome (self-reported diarrhea).[246,247] Considering the significant associations between viral infection (which was not associated with sanitation) and diarrhea in this study, this finding suggests that previous discussion of the ineffectiveness of sanitation in reducing enteric infections may be of limited utility since it fails to reflect the biological and environmental mechanisms observed. Further, it

suggests that future studies must understand the etiologic causes of diarrhea and enteric infection within the study population before selecting a health outcome (or better, an intervention) in order to more appropriate measure and understand the true health impact of sanitation.

Hazard analysis of enteric infection

Methods:

To determine associations between household characteristics, neighborhood characteristics, household exposures behaviors, and the instantaneous risk (hazard) of enteric infection, Cox Proportional-Hazard models (survival curves) were fit to the data for pooled enteric infection. This model differs from the risk calculated in mixed effects logistic regression models as the latter is cumulative risk over the two year time period while the former is the instantaneous risk, and is based on (in this case) the time to first event (first enteric infection of any kind).[248]

Results:

Models for household and neighborhood characteristics, controlling for season, are shown in Table A3.1. Only children living in households with a toilet (and specifically a toilet leading to the drain) had significantly lower risk of enteric infection. Specifically, children in households with a toilet had 35% lower chance of infection (hazard) when compared with children in households without a toilet. Children in households with a toilet leading to the drain had 37% lower chance of infection (hazard) when compared with those without toilets at all.

Relationships between spatial clustering of drain flooding, seasonality, and hazard are shown in Table A3.2. Spatial clusters of drain flooding and house flooding did not show

significant associations with hazard (pooled enteric infection), both when controlling for season and when evaluating interactions with season.

Relationships between household exposure behaviors and hazard are shown in Table A3.3, controlling for season. Children in households where the adult caregiver open defecation had an 84% greater chance of enteric infection at any given time point when compared to children in households where the caregiver did not report open defecating. Children in households where the older sibling open defecated had similarly elevated chance of infection, though this was not significant.

Discussion:

Overall, fewer variables were significantly associated with the hazard (chance at any given time point) of enteric infection in general than were significantly associated with risk (cumulative, over the 2-year follow-up period) of enteric infection. Notably, household sanitation and sanitation practices were the strongly associated with hazard in these models and were also associated with enteric infection risk in earlier models, underscoring their importance in the home. Practically, this means that sanitation variables not only were associated with risk of infection, but also that children in households with a toilet did not encounter their first enteric infection until later in life. However, as noted in Aim 3, enteric infection incidence was high and all children had been infected at least once by 1 year of age.

Table A3.1: Bivariate relationships between household and neighborhood conditions and infection hazard using stool specimen data collected from children in SaniPath households, 2010-2014, Vellore, India[a]

a) Household conditions	Any pathogen Hazard Ratio (95% CI)
Demographics	
Num. people per household	0.99 (0.93, 1.07)
5-12 year old (YO) present	1.09 (0.77, 1.54)
Sanitation	
Household toilet	0.65 (0.45-0.96)†
Toilet with poor FSM[b,c]	0.62 (0.41, 0.94)†
Toilet with good FSM[c,d]	1.50 (0.53, 4.27)
Toilet with poor FSM[b,e]	0.63 (0.41, 0.95)†
Toilet with good FSM[d,e]	1.41 (0.49, 3.99)
b) Neighborhood conditions	
Drain present in front of HH	0.47 (0.20, 1.11)
Flooding	
Drain flooding	1.24 (0.87, 1.75)
House flooding	0.96 (0.65, 1.43)

[a]All models adjusted for monsoon seasons (relative to dry season); [b]Toilet discharges to an open drain ; [c]Compared to households with other toilet/FSM combinations and households without toilets; [d]Toilet excreta is contained onsite; [e]Compared to households without toilets; †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$

Table A3.2: Bivariate relationships between spatial clustering of flooding in neighborhood, seasonality, and infection hazard using stool specimen data collected from children in SaniPath households, 2010-2014, Vellore, India

	Any pathogen Hazard Ratio (95% CI)
Drain flooding cluster	
Year-round[a]	0.88 (0.58, 1.34)
Dry season (Jan. – May)	Ref.
Southwest monsoon (June – Sept.)	0.98 (0.34, 2.82)
Northeast monsoon (Oct. – Dec.)	0.92 (0.31, 2.69)
House flooding cluster	
Year-round[b]	0.79 (0.39, 1.62)
Dry season (Jan. – May)	Ref.
Southwest monsoon (June – Sept.)	0.18 (0.02, 1.74)
Northeast monsoon (Oct. – Dec.)	0.62 (0.12, 3.11)

[a]Pooled across seasons; $^{\dagger}p < 0.05$; $^{\dagger\dagger}p < 0.01$; $^{\dagger\dagger\dagger}p < 0.001$

Table A3.3: Bivariate relationships between household exposure behaviors and infection hazard using stool specimen data collected from children in SaniPath households, 2010-2014, Vellore, India[a]

	Any pathogen Hazard Ratio (95% CI)
Drain contact	
Any	1.20 (0.66, 2.16)
>10 times per month	1.45 (0.92, 2.30)
Flood water contact	
Any	1.08 (0.65, 1.77)
>10 times per month	0.83 (0.51, 1.32)
Open defecation	
<5 YO	1.40 (0.91, 2.15)
5-12 YO	1.90 (0.98, 3.69)
Adult	1.84 (1.24, 2.73)††
Public toilet use	
Any	1.00 (0.70, 1.42)
>10 times per month	1.22 (0.77, 1.95)
Municipal water consumption	
Any	0.70 (0.44, 1.10)
Daily	0.73 (0.51, 1.04)
Drinking water treatment	0.72 (0.49, 1.05)
Produce consumption	
Any	0.72 (0.50, 1.03)
Daily	1.24 (0.81, 1.89)

[a]All models adjusted for monsoon seasons; †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$

Retesting of household and neighborhood risk factors, controlling for breastfeeding status
Methods:

Breastfeeding status of the study child was recorded daily by the MAL-ED study staff, and was divided into exclusive breastfeeding, predominant breastfeeding, and partial breastfeeding during the follow-up period. Because exclusive breastfeeding has been shown to be protective against diarrhea and may reduce fecal exposures, it was included as a binary variable (exclusive breastfeeding or not exclusive breastfeeding in last 24 hours at time of specimen collection) in mixed-effects logistic regression models, with stool type and season.[249]
Results:

To understand whether breastfeeding was a potential confounder of relationships between the risk factors examined in Aim 3 and pediatric enteric infection outcomes, breastfeeding was included in mixed-effects logistic regression models for enteric infection by itself and with previously tested risk factors (Tables A3.4 – A3.6). Controlling for season of specimen collection and type of stool collected, breastfeeding at the time of specimen collection was strongly associated with reduced risk of enteric infection (OR: 0.23, 95% CI: 0.17, 0.31). Among pathogens tested, risk was similarly reduced for bacterial infections (OR: 0.26, 95% CI: 0.19, 0.36), but was not significant for viral infection risk (OR: 0.63, 95% CI: 0.37, 1.09). The model for parasitic infection risk did not converge and thus is not presented.

Associations between household and neighborhood conditions and enteric infection risk were retested, controlling for exclusive breastfeeding status (Table A3.4). Overall, breastfeeding status was not a meaningful confounder (did not change main effect estimates by more than 10% when included in the model) for these models. Of note, household toilets were associated with

borderline significantly reduced risk of parasitic infection in Aim 3; however, when controlling for exclusive breastfeeding, this relationship was statistically significant.

Associations between spatial clustering of flooding in neighborhoods and enteric infection risk were also retested, controlling for exclusive breastfeeding status (Table A3.5). Breastfeeding status was not a meaningful confounder of these relationships.

Finally, associations between household exposure behaviors and enteric infection risk were retested, controlling for breastfeeding status (Table A3.6). As with the earlier associations, breastfeeding status was not a meaningful confounder of these relationships. Controlling for breastfeeding in the models with open defecation did increase the size and statistical significance of the effects observed, however.

Discussion:

Overall, breastfeeding status was not a meaningful confounder of the existing associations between household and neighborhood risk factors and enteric infection risk, though on its own, it was a significant protective factor. Exclusive breastfeeding provides the child with essential immunologic protection and may act as a barrier to drinking water-associated exposure pathways, as children being exclusively breastfed do not receive any water or water-based products.[249,250] Given these reasons, it is somewhat surprising that breastfeeding status did not confound estimates of the association between drinking water consumption (or reported water treatment) and enteric infection risk. One would expect that risks associated with drinking water consumption would be biased towards the null in the absence of adjustment for breastfeeding, given that children who were breastfed at the time of specimen collection would have been previously mixed with children who were actually consuming the municipal water. Thus, the addition of breastfeeding to would be expected to shift effects away from the null. These shifts

are not observed, however, likely because municipal water consumption was protective for some outcomes (bacterial infection) and a risk factor for others (viral infection), though none of the associations were significant.

Table A3.4: Bivariate relationships (retested, controlling for breastfeeding) between household and neighborhood conditions and pathogen detection in stool collected from children in SaniPath households, 2010-2014, Vellore, India[a]

a) Household conditions	Bacteria OR (95% CI)	Parasite OR (95% CI)	Virus OR (95% CI)	Any pathogen OR (95% CI)
Demographics				
Num. people per household	*	*	1.08 (1.00, 1.17)	1.02 (0.95, 1.10)
5-12 year old (YO) present	1.33 (1.02, 1.74)†	1.91 (1.15, 3.18)†	1.00 (0.69, 1.45)	1.53 (1.13, 2.08)††
Sanitation				
Household toilet	0.67 (0.51, 0.88)††	0.58 (0.33, 0.99)†	1.09 (0.75, 1.58)	0.68 (0.50, 0.94)†
Toilet with poor FSM[b,c]	0.69 (0.52, 0.93)†	*	0.82 (0.55, 1.24)	0.70 (0.50, 0.99)†
b) Neighborhood conditions				
Drain present in front of HH	1.17 (0.59, 2.32)	3.98 (0.83, 19.1)	0.96 (0.37, 2.50)	1.41 (0.65, 3.06)
Flooding				
Drain flooding	1.23 (0.94, 1.61)	*	0.79 (0.55, 1.13)	1.27 (0.93, 1.73)
House flooding	1.08 (0.79, 1.49)	*	0.84 (0.54, 1.30)	1.02 (0.71, 1.48)

[a]All models adjusted for monsoon seasons (relative to dry season), exclusive breastfeeding at the time of specimen collection, and stool type (routine vs. diarrheal stool); [b]Toilet discharges to an open drain ; [c]Compared to households with other toilet/FSM combinations and households without toilets; †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$; *Model did not converge

Table A3.5: Bivariate relationships (retested, controlling for breastfeeding) between spatial clustering of flooding in neighborhood, seasonality, and pathogen detection in stool collected from children in SaniPath households, 2010-2014, Vellore, India[a]

	Bacteria OR (95% CI)	Virus OR (95% CI)	Any pathogen OR (95% CI)
Drain flooding cluster			
Year-round[b]	0.93 (0.68, 1.27)	0.79 (0.51, 1.24)	0.91 (0.63, 1.31)
Dry season (Jan. – May)	Ref.	Ref.	*
Southwest monsoon (June – Sept.)	0.96 (0.54, 1.69)	0.21 (0.08, 0.55)††	
Northeast monsoon (Oct. – Dec.)	2.19 (1.25, 4.17)†	0.72 (0.23, 2.18)	
House flooding cluster			
Year-round[b]	1.22 (0.70, 2.11)	1.27 (0.61, 2.63)	1.08 (0.58, 2.02)
Dry season (Jan. – May)	Ref.	Ref.	*
Southwest monsoon (June – Sept.)	0.63 (0.22, 1.73)	0.17 (0.02, 1.55)	
Northeast monsoon (Oct. – Dec.)	1.06 (0.34, 3.30)	1.99 (0.44, 9.12)	

[a]All models adjusted stool type (routine vs. diarrheal stool) and exclusive breastfeeding at the time of specimen collection. Parasite results not shown due to small number of events in cluster, preventing model convergence; [b]Pooled across seasons; †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$; *Model did not converge

Table A3.6: Bivariate relationships (retested, controlling for breastfeeding) between household exposure behaviors and pathogen detection in stool collected from children in SaniPath households, 2010-2014, Vellore, India[a]

	Bacteria OR (95% CI)	Parasite OR (95% CI)	Virus OR (95% CI)	Any pathogen OR (95% CI)
Drain contact				
Any	1.46 (0.95, 2.24)	2.46 (0.99, 6.08)	1.29 (0.70, 2.36)	1.55 (0.96, 2.50)
>10 times per month	1.29 (0.88, 1.90)	1.72 (0.88, 3.38)	1.16 (0.71, 1.88)	1.37 (0.88, 2.14)
Flood water contact				
Any	1.44 (1.01, 2.05)†	1.28 (0.64, 2.59)	1.41 (0.87, 2.30)	1.42 (0.96, 2.11)
>10 times per month	0.91 (0.62, 1.34)	1.64 (0.81, 3.30)	0.82 (0.48, 1.40)	0.88 (0.57, 1.36)
Open defecation				
<5 YO	1.43 (1.04, 1.98)†	*	1.06 (0.69, 1.65)	*
5-12 YO	1.73 (1.18, 2.54)††	1.23 (0.58, 2.59)	0.95 (0.58, 1.57)	1.63 (1.06, 2.54)†
Adult	1.63 (1.25, 2.14)†††	*	1.23 (0.84, 1.81)	1.62 (1.18, 2.21)††
Public toilet use				
Any	0.84 (0.64, 1.10)	*	1.04 (0.72, 1.51)	0.91 (0.66, 1.25)
>10 times per month	0.93 (0.63, 1.38)	*	1.48 (0.90, 2.43)	1.18 (0.74, 1.88)
Municipal water consumption				
Any	0.95 (0.66, 1.38)	*	1.38 (0.82, 2.33)	0.99 (0.65, 1.51)
Daily	0.95 (0.71, 1.28)	*	1.09 (0.74, 1.61)	0.91 (0.65, 1.28)
Drinking water treatment	0.75 (0.56, 0.99)†	0.57 (0.33, 0.99)†	0.75 (0.50, 1.11)	0.67 (0.49, 0.92) †
Produce consumption				
Any	0.88 (0.66, 1.18)	*	1.06 (0.71, 1.56)	0.81 (0.58, 1.13)
Daily	1.10 (0.77, 1.57)	*	1.21 (0.76, 1.91)	1.20 (0.79, 1.81)

[a]All models adjusted for monsoon seasons, exclusive breastfeeding at time of specimen collection, and stool type (routine vs. diarrheal stool); †$p < 0.05$; ††$p < 0.01$; †††$p < 0.001$; *Models did not converge

Further discussion of results from Aim 3:

Discussion of microbiological data in the context of Aim 3 findings:

The finding of significantly reduced risk of enteric infection in children in households where drinking water treatment was reported corroborates previous evidence of poor drinking water quality in Vellore.[193,214] [Kirby et al., unpublished data] Recent studies have suggested Vellore's drinking water, though reported to be treated, is not sufficiently treated and that household storage practices in many of the residences actually increase fecal contamination in the water.[193,214] Further, environmental samples collected as part of the recent SaniPath deployment in Vellore (documented in [Kirby et al., unpublished data]) showed that drinking water was a high-risk pathway because of high levels of *E. coli* detected in it. Thus, drinking water treatment, if practiced as reported by household members, may have mitigated a major environmental pathway for young children to get exposed to fecal contamination.[27]

Discussion of pathogen group-specific findings from Aim 3:

Use of toilets was observed to be significantly associated with lower bacterial, but not viral (and, to a limited extent, parasitic) infections, and corresponding open defecation was a risk factor, for a number of biological reasons. As noted in Aim 3, parasites and viruses have lower infectious doses, on average, when compared with bacteria, meaning that removal of fecal contamination from the environment via the use of toilets may not have been large enough to observe changes in incidence of those infections. Further, viruses and most parasites cannot replicate (or have limited ability to develop) outside the human host, forcing them (in addition to having lower infectious doses) to have greater environmental persistence and the ability to persist on numerous environmental surfaces.[9,12] Open defecation fields may not be the primary

exposure point for infection from these organisms. Rather, individuals' contact with each other, or with common household or public surfaces (even toilets), are more likely exposure points, as demonstrated by the association of frequent use of public toilets with increased GII norovirus risk. Thus, on the whole, the reduction in fecal contamination levels associated with use of household sanitation 1) may not have been sufficiently large to impact organisms, like viruses and parasites, with low infectious doses and 2) the use of sanitation may not have impacted the most important environmental exposure/transmission pathways for those organisms.

References

1. WHO. WHO: Diarrhoeal disease. WHO Fact Sheets. http://www.who.int/mediacentre/factsheets/fs330/en/. Published 2013.
2. Prüss-Ustün A, Bos R, Gore F, Bartram J. *Safer Water, Better Health*.; 2008.
3. Bartram J, Cairncross S. Hygiene, sanitation, and water: Forgotten foundations of health. *PLoS Med*. 2010;7(11):1-9. doi:10.1371/journal.pmed.1000367.
4. Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through improved water, sanitation and hygiene: an opportunity for stunting reduction in developing countries. *Matern Child Nutr*. 2015:n/a - n/a. doi:10.1111/mcn.12220.
5. Ngure FM, Reid BM, Humphrey JH, Mbuya MN, Pelto G, Stoltzfus RJ. Water, sanitation, and hygiene (WASH), environmental enteropathy, nutrition, and early child development: Making the links. *Ann N Y Acad Sci*. 2014;1308(1):118-128. doi:10.1111/nyas.12330.
6. Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet*. 2009;374(9694):1032-1035. doi:10.1016/S0140-6736(09)60950-8.
7. Guerrant RL, Deboer MD, Moore SR, Scharf RJ, Aldo a M. The impoverished gut--a triple burden of diarrhoea, stunting, and chronic disease. *Nat Rev Gastroenterol Hepatol*. 2014;10(4):220-229. doi:10.1038/nrgastro.2012.239.The.
8. Guerrant RL, Kosek M, Lima AAM, Lorntz B, Guyatt HL. Updating the DALYS for diarrheal disease. *Trends in Parasitol*. 2002;18(5):191-193.
9. Leclerc H, Schwartzbrod L, Dei-Cas E. Microbial agents associated with

waterborne diseases. *Crit Rev Microbiol*. 2002;28(4):371-409. doi:10.1080/1040-840291046768.

10. Bethony et. al. Soil-transmitted Helminth Infections: Ascariasis, Trichuriasis and Hookworm. *Lancet*. 2006;367:1521-1532.
11. Ziegelbauer K, Speich B, Mäusezahl D, Bos R, Keiser J, Utzinger J. Effect of Sanitation on Soil-Transmitted Helminth Infection: Systematic Review and Meta-Analysis. *PLoS Med*. 2012;9(1):e1001162. doi:10.1371/journal.pmed.1001162.
12. Kotwal G, Cannon JL. Environmental persistence and transfer of enteric viruses. *Curr Opin Virol*. 2014;4(2004):37-43. doi:10.1016/j.coviro.2013.12.003.
13. Glass RI, Parashar UD, Estes MK. Norovirus Gastroenteritis. *N Engl J Med*. 2009;361(18). doi:10.1056/NEJMra0804575.Norovirus.
14. Ayukekbong JA, Mesumbe HN, Oyero OG, Lindh M, Bergström T. Role of noroviruses as aetiological agents of diarrhoea in developing countries. *J Gen Virol*. 2015;96(8):1983-1999. doi:10.1099/vir.0.000194.
15. Lopman B, Gastañaduy P, Park GW, Hall AJ, Parashar UD, Vinjé J. Environmental transmission of norovirus gastroenteritis. *Curr Opin Virol*. 2012;2(1):96-102. doi:10.1016/j.coviro.2011.11.005.
16. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, Wu Y, Sow SO, Sur D, Breiman RF, Faruque AS, Zaidi AK, Saha D, Alonso PL, Tamboura B, Sanogo D, Onwuchekwa U, Manna B, Ramamurthy T, Kanungo S, Ochieng JB, Omore R, Oundo JO, Hossain A, Das SK, Ahmed S, Qureshi S, Quadri F, Adegbola RA, Antonio M, Hossain MJ, Akinsola A, Mandomando I, Nhampossa T, Acácio S, Biswas K, O'Reilly CE, Mintz ED, Berkeley LY,

Muhsen K, Sommerfelt H, Robins-Browne RM, Levine MM. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet*. 2013;382(9888):209-222. doi:10.1016/S0140-6736(13)60844-2.

17. Seitz SR, Leon JS, Schwab KJ, Lyon GM, Dowd M, McDaniels M, Abdulhafid G, Fernandez ML, Lindesmith LC, Baric RS, Moe CL. Norovirus infectivity in humans and persistence in water. *Appl Environ Microbiol*. 2011;77(19):6884-6888. doi:10.1128/AEM.05806-11.

18. Teunis PFM, Moe CL, Liu P, Miller SE, Lindesmith L, Baric RS, Le Pendu J, Calderon RL. Norwalk Virus: How Infectious is It? *J Med Virol*. 2008;80:1468-1476. doi:10.1002/jmv.

19. Gerhardts A, Hammer TR, Balluff C, Mucha H, Hoefer D. A model of the transmission of micro-organisms in a public setting and its correlation to pathogen infection risks. *J Appl Microbiol*. 2012;112(3):614-621. doi:10.1111/j.1365-2672.2012.05234.x.

20. Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleesschauwer B, Dopfer D, Fazil A, Fischer-Walker CL, Hald T, Hall AJ, Keddy KH, Lake RJ, Lanata CF, Torgerson PR, Havelaar AH, Angulo FJ. World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Med*. 2015;12(12):1-21. doi:10.1371/journal.pmed.1001921.

21. Jiang SC. Human adenoviruses in water: Occurrence and health implications: A

critical review. *Environ Sci Technol*. 2006;40(23):7132-7140. doi:10.1021/es060892o.

22. Hewitt J, Greening GE, Leonard M, Lewis GD. Evaluation of human adenovirus and human polyomavirus as indicators of human sewage contamination in the aquatic environment. *Water Res*. 2013;47(17):6750-6761. doi:10.1016/j.watres.2013.09.001.

23. Okhuysen PC, Dupont HL. Enteroaggregative Escherichia coli (EAEC): a cause of acute and persistent diarrhea of worldwide importance. *J Infect Dis*. 2010;202(4):503-505. doi:10.1086/654895.

24. Weintraub a. Enteroaggregative Escherichia coli: epidemiology, virulence and detection. *J Med Microbiol*. 2007;56(1):4-8. doi:10.1099/jmm.0.46930-0.

25. Platts-Mills J a, Babji S, Bodhidatta L, Gratz J, Haque R, Havt A, McCormick BJ, McGrath M, Olortegui MP, Samie A, Shakoor S, Mondal D, Lima IF, Hariraju D, Rayamajhi BB, Qureshi S, Kabir F, Yori PP, Mufamadi B, Amour C, Carreon JD, Richard S a, Lang D, Bessong P, Mduma E, Ahmed T, Lima AA, Mason CJ, Zaidi AK, Bhutta Z a, Kosek M, Guerrant RL, Gottlieb M, Miller M, Kang G, Houpt ER. Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Heal*. 2015:564-575. doi:10.1016/S2214-109X(15)00151-5.

26. Jensen BH, Olsen KEP, Struve C, Krogfelt KA, Petersen AM. Epidemiology and clinical manifestations of enteroaggregative escherichia coli. *Clin Microbiol Rev*. 2014;27(3):614-630. doi:10.1128/CMR.00112-13.

27. Wagner EG, Lanoix JN. *Excreta Disposal for Rural Areas and Small*

Communities. Geneva: WHO; 1958.

28. Montgomery MA, Elimelech M. Water and Sanitation in Developing Countries: Including Health in the Equation. *Environ Sci Technol*. 2007;41(1):17-24.
29. WHO and UNICEF. 2015 Update and MDG Assessment. 2015. http://www.wssinfo.org/.
30. Clasen T, Bostoen K, Schmidt W, Boisson S, Fung I, Jenkins M, Scott B, Sugden S, Cairncross S. *Interventions to Improve Disposal of Human Excreta for Preventing Diarrhoea (Review)*.; 2010.
31. Peal A, Evans B, Blackett I, Hawkins P, Heymans C. Fecal Sludge Management : a comparative analysis of 12 cities. *J Water, Sanit Hyg Dev*. 2014;4(4):563-575.
32. Peal A, Evans B, Blackett I, Hawkins P, Heymans C. Fecal sludge management (FSM): analytical tools for assessing FSM in cities. *J Water, Sanit Hyg Dev*. 2014;4(3):371. doi:10.2166/washdev.2014.139.
33. EAWAG. Sanitation 21. 2014;(September).
34. Briscoe J. Intervention studies and the definition of dominant transmission routes. *Am J Epidemiol*. 1984;120(3):449-455.
35. VanDerslice J, Briscoe J. Environmental interventions in developing countries: interactions and their implications. *Am J Epidemiol*. 1995;141(2):135-144. http://www.ncbi.nlm.nih.gov/pubmed/7817969.
36. Bateman OM, Smith S. *A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala*. Washington D.C.; 1991.
37. Root GPM. Sanitation, Community Environments, and Childhood Diarrhea in Rural Zimbabwe. *J Heal Popul Nutr*. 2001;19(2):73-82.

38. Shuval HI. Tilden RL. Perry BH. Grosse RN. Effect of investments in water supply and sanitation. *Bull World Health Organ*. 1981;59(2):243-248.

39. Cairncross S, Blumenthal U, Kolsky P, Moraes L, Tayeh a. The public and domestic domains in the transmission of disease. *Trop Med Int Health*. 1996;1(1):27-34. doi:10.1046/j.1365-3156.1996.d01-9.x.

40. Pullan RL, Freeman MC, Gething PW, Brooker SJ. Geographical Inequalities in Use of Improved Drinking Water Supply and Sanitation across Sub-Saharan Africa: Mapping and Spatial Analysis of Cross-sectional Survey Data. *PLoS Med*. 2014;11(4). doi:10.1371/journal.pmed.1001626.

41. Ingallinella a M, Sanguinetti G, Koottatep T, Montanger A, Strauss M. The challenge of faecal sludge management in urban areas--strategies, regulations and treatment options. *Water Sci Technol*. 2002;46(10):285-294. http://www.ncbi.nlm.nih.gov/pubmed/12479483.

42. Diallo MBC, Anceno a J, Tawatsupa B, Tripathi NK, Wangsuphachart V, Shipin O V. GIS-based analysis of the fate of waste-related pathogens Cryptosporidium parvum, Giardia lamblia and Escherichia coli in a tropical canal network. *J Water Health*. 2009;7(1):133-144. doi:10.2166/wh.2009.010.

43. Snyder RE, Jaimes G, Riley LW, Faerstein E, Corburn J. A Comparison of Social and Spatial Determinants of Health Between Formal and Informal Settlements in a Large Metropolitan Setting in Brazil. *J Urban Health*. December 2013. doi:10.1007/s11524-013-9848-1.

44. Akullian A, Ng'eno E, Matheson AI, Cosmas L, Macharia D, Fields B, Bigogo G, Mugoh M, John-Stewart G, Walson JL, Wakefield J, Montgomery JM.

Environmental Transmission of Typhoid Fever in an Urban Slum. *PLoS Negl Trop Dis*. 2015;9(12):1-14. doi:10.1371/journal.pntd.0004212.

45. Weeks JR, Getis A, Stow DA, Hill AG, Rain D, Stoler J, Lippitt C, Jankowska M, Lopez- AC, Coulter L, Ofiesh C, Engstrom R, Lopez-carr AC. Connecting the Dots Between Health , Poverty , and Place in Accra , Ghana. *Ann Assoc Am Geogr*. 2012;105(2):932-941.

46. Eisenberg JNS, Trostle J, Sorensen RJD, Shields KF. Toward a Systems Approach to Enteric Pathogen Transmission: From Individual Independence to Community Interdependence. *Annu Rev Public Heal*. 2012;April(33):239-257. doi:10.1016/j.micinf.2011.07.011.Innate.

47. Phung D, Huang C, Rutherford S, Chu C, Wang X, Nguyen M, Nguyen NH, Do CM, Nguyen TH. Temporal and spatial patterns of diarrhoea in the Mekong Delta area, Vietnam. *Epidemiol Infect*. 2015:1-10. doi:10.1017/S0950268815000709.

48. Santos CB, Araújo KC, Jardim-Botelho A, Santos MB, Rodrigues A, Dolabella SS, Gurgel RQ. Diarrhea incidence and intestinal infections among rotavirus vaccinated infants from a poor area in Brazil: a spatial analysis. *BMC Public Health*. 2014;14(1):399. doi:10.1186/1471-2458-14-399.

49. Azage M, Kumie A, Worku A, Bagtzoglou AC. Childhood diarrhea in high and low hotspot districts of Amhara Region, northwest Ethiopia: a multilevel modeling. *J Heal Popul Nutr*. 2016;35(1):13. doi:10.1186/s41043-016-0052-2.

50. Azage M, Kumie A, Worku A, Bagtzoglou AC. Childhood Diarrhea Exhibits Spatiotemporal Variation in Northwest Ethiopia: A SaTScan Spatial Statistical Analysis. *PLoS One*. 2015;10(12):1-18. doi:10.1371/journal.pone.0144690.

51. Chaikaew N, Tripathi NK, Souris M. Exploring spatial patterns and hotspots of diarrhea in Chiang Mai, Thailand. *Int J Health Geogr*. 2009;8:36. doi:10.1186/1476-072X-8-36.

52. Kulldorff M. A spatial scan statistic. *Commun Stat Meth*. 1997;26(6):1481-1496.

53. Haile D, Azage M, Mola T, Rainey R. Exploring spatial variations and factors associated with childhood stunting in Ethiopia: spatial and multilevel analysis. *BMC Pediatr*. 2016;16(1):49. doi:10.1186/s12887-016-0587-9.

54. Gao L, Zhang Y, Ding G, Liu Q, Jiang B. Identifying Flood-Related Infectious Diseases in Anhui Province, China: A Spatial and Temporal Analysis. *Am J Trop Med Hyg*. 2016;94(4):741-749. doi:10.4269/ajtmh.15-0338.

55. Anselin L. Under the hood issues in the specification and interpretation of spatial regression models. *Agric Econ*. 2002;27(3):247-267. doi:10.1016/S0169-5150(02)00077-4.

56. Fuller JA, Villamor E, Cevallos W, Trostle J, Eisenberg JNS. I get height with a little help from my friends: Herd protection from sanitation on child growth in rural Ecuador. *Int J Epidemiol*. 2016;0(0):1-10. doi:10.1093/ije/dyv368.

57. Chao DL, Park JK, Marks F, Ochiai RL, Longini IM, Halloran ME. The contribution of neighbours to an individual's risk of typhoid outcome. *Epidemiol Infect*. 2015:1-8. doi:10.1017/S0950268815000692.

58. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, Gluud C, Martin RM, Wood AJG, Sterne JAC. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ*. 2008;336(7644):601-605.

doi:10.1136/bmj.39465.451748.AD.

59. Liu J, Gratz J, Amour C, Kibiki G, Becker S, Janaki L, Verweij JJ, Taniuchi M, Sobuz SU, Haque R, Haverstick DM, Houpt ER. A Laboratory-Developed TaqMan Array Card for Simultaneous Detection of 19 Enteropathogens. *J Clin Microbiol*. 2013;51(2):472-480. doi:10.1128/JCM.02658-12.

60. Claas EC, Burnham C-A, Mazzulli T, Templeton K, Topin F. Performance of the xTAG Gastrointestinal Pathogen Panel, a Multiplex Molecular Assay for Simultaneous Detection of Bacteria, Viral, and Parasitic Causes of Infectious Gastroenteritis. *J Microbiol Biotechnol*. 2013;23(7):1041-1045.

61. Beckmann C, Heininger U, Marti H, Hirsch HH. Gastrointestinal pathogens detected by multiplex nucleic acid amplification testing in stools of pediatric patients and patients returning from the tropics. *Infection*. 2014. doi:10.1007/s15010-014-0656-7.

62. Patel A, Navidad J, Bhattacharyya S. Site-specific clinical evaluation of the Luminex xTAG gastrointestinal pathogen panel for detection of infectious gastroenteritis in fecal specimens. *J Clin Microbiol*. 2014;52(8):3068-3071. doi:10.1128/JCM.01393-14.

63. Mengelle C, Mansuy JM, Prere MF, Grouteau E, Claudet I, Kamar N, Huynh A, Plat G, Benard M, Marty N, Valentin A, Berry A, Izopet J. Simultaneous detection of gastrointestinal pathogens with a multiplex Luminex-based molecular assay in stool samples from diarrhoeic patients. *Clin Microbiol Infect*. 2013;19(10). doi:10.1111/1469-0691.12255.

64. Panchalingam S, Antonio M, Hossain A, Mandomando I, Ochieng B, Oundo J,

Ramamurthy T, Tamboura B, Zaidi AKM, Petri W, Houpt E, Murray P, Prado V, Vidal R, Steele D, Strockbine N, Sansonetti P, Glass RI, Robins-Browne RM, Tauschek M, Svennerholm AM, Kotloff K, Levine MM, Nataro JP. Diagnostic microbiologic methods in the GEMS-1 case/control study. *Clin Infect Dis*. 2012;55(SUPPL. 4):294-302. doi:10.1093/cid/cis754.

65. Houpt E, Gratz J, Kosek M, Zaidi AKM, Qureshi S, Kang G, Babji S, Mason C, Bodhidatta L, Samie A, Bessong P, Barrett L, Lima A, Havt A, Haque R, Mondal D, Taniuchi M, Stroup S, McGrath M, Lang D. Microbiologic Methods Utilized in the MAL-ED Cohort Study. *Clin Infect Dis*. 2014;59(suppl 4):S225-S232. doi:10.1093/cid/ciu413.

66. Platts-Mills JA, McCormick BJ, Kosek M, Pan WK, Checkley W, Houpt ER, Investigators M-EN. Methods of analysis of enteropathogen infection in the MAL-ED Cohort Study. *Clin Infect Dis*. 2014;59 Suppl 4(Suppl 4):S233-S238. doi:10.1093/cid/ciu408.

67. Brown J, Cumming O, Bartram J, Cairncross S, Ensink J, Holcomb D, Knee J, Kolsky P, Liang K, Liang S, Nala R, Norman G, Rheingans R, Stewart J, Zavale O, Zuin V, Schmidt W-P. A controlled, before-and-after trial of an urban sanitation intervention to reduce enteric infections in children: research protocol for the Maputo Sanitation (MapSan) study, Mozambique. *BMJ Open*. 2015;5(6):e008215. doi:10.1136/bmjopen-2015-008215.

68. Machdar E, van der Steen NP, Raschid-Sally L, Lens PNL. Application of Quantitative Microbial Risk Assessment to analyze the public health risk from poor drinking water quality in a low income area in Accra, Ghana. *Sci Total*

Environ. 2013;449:134-142.

69. Labite H, Lunani I, van der Steen P, Vairavamoorthy K, Drechsel P, Lens P. Quantitative Microbial Risk Analysis to evaluate health effects of interventions in the urban water system of Accra, Ghana. *J Water Health*. 2010;8(3):417-430. doi:10.2166/wh.2010.021.

70. Katukiza AY, Ronteltap M, van der Steen P, Foppen JW a, Lens PNL. Quantification of microbial risks to human health caused by waterborne viruses and bacteria in an urban slum. *J Appl Microbiol*. October 2013:1-17. doi:10.1111/jam.12368.

71. ten Veldhuis J a E, Clemens FHLR, Sterk G, Berends BR. Microbial risks associated with exposure to pathogens in contaminated urban flood water. *Water Res*. 2010;44(9):2910-2918. doi:10.1016/j.watres.2010.02.009.

72. de Man H, van den Berg HHJL, Leenen EJTM, Schijven JF, Schets FM, van der Vliet JC, van Knapen F, de Roda Husman a M. Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. *Water Res*. 2014;48:90-99. doi:10.1016/j.watres.2013.09.022.

73. Yajima a, Koottatep T. Assessment of E. coli and Salmonella spp. infection risks associated with different fecal sludge disposal practices in Thailand. *J Water Health*. 2010;8(2):355-364. doi:10.2166/wh.2009.310.

74. Tallon PAM, Magajna B, Lofranco C, Leung KAMTIN. Microbial Indicators of Fecal Contamination in Water: A Current Perspective. *Water, Air Soil Pollut*. 2005;166:139-166.

75. Olstadt J, Schauer JJ, Standridge J, Kluender S. A comparison of ten USEPA

approved total coliform/E. coli tests. *J Water Heal*. 2007;5(2):267-282.

76. Smith CJ, Osborn AM. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol*. 2009;67(1):6-20. doi:10.1111/j.1574-6941.2008.00629.x.

77. Torondel B, Gyekye-Aboagye Y, Routray P, Boisson S, Schimdt W, Clasen T. Laboratory development and field testing of sentinel toys to assess environmental faecal exposure of young children in rural India. *Trans R Soc Trop Med Hyg*. 2015;(March):386-392. doi:10.1093/trstmh/trv023.

78. Heijnen M, Routray P, Torondel B, Clasen T. Shared Sanitation Versus Individual Household Latrines in Urban Slums: A Cross-Sectional Study in Orissa, India. *Am J Trop Med Hyg*. 2015;93(2):263-268. doi:10.4269/ajtmh.14-0812.

79. Pickering AJ, Julian TR, Mamuya S, Boehm AB, Davis J. Bacterial hand contamination among Tanzanian mothers varies temporally and following household activities. *Trop Med Int Heal*. 2011;16(2):233-239. doi:10.1111/j.1365-3156.2010.02677.x.

80. Pickering AJ, Julian TR, Marks SJ, Mattioli MC, Boehm AB, Schwab KJ, Davis J. Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian households with and without improved sanitation. *Environ Sci Technol*. 2012;46(11):5736-5743. doi:10.1021/es300022c.

81. Pickering AJ, Davis J, Walters SP, Horak HM, Keymer DP, Mushi D, Strickfaden R, Chynoweth JS, Liu J, Blum A, Rogers K, Boehm AB. Hands , Water , and Health : Fecal Contamination in Tanzanian Communities with Improved , Non-Networked Water Supplies. *Environ Sci Technol*. 2010;44:3267-3272.

82. Mattioli MC, Boehm AB, Davis J, Harris AR, Mrisho M, Pickering AJ. Enteric pathogens in stored drinking water and on caregiver's hands in Tanzanian households with and without reported cases of child diarrhea. *PLoS One*. 2014;9(1):e84939. doi:10.1371/journal.pone.0084939.

83. Greene LE, Freeman MC, Akoko D, Saboori S, Moe C, Rheingans R. Impact of a school-based hygiene promotion and sanitation intervention on pupil hand contamination in Western Kenya: A cluster randomized trial. *Am J Trop Med Hyg*. 2012;87(3):385-393. doi:10.4269/ajtmh.2012.11-0633.

84. Gil AI, Lanata CF, Hartinger SM, Mäusezahl D, Padilla B, Ochoa TJ, Lozada M, Pineda I, Verastegui H. Fecal contamination of food, water, hands, and kitchen utensils at the household level in rural areas of Peru. *J Environ Health*. 2013;76(6):102-106. http://www.ncbi.nlm.nih.gov/pubmed/24645420.

85. Katukiza AY, Temanu H, Chung JW, Foppen JW a, Lens PNL. Genomic copy concentrations of selected waterborne viruses in a slum environment in Kampala, Uganda. *J Water Health*. 2013;11(2):358-370. doi:10.2166/wh.2013.184.

86. Gretsch SR, Ampofo JA, Baker KK, Clennon J, Null CA, Peprah D, Reese H, Robb K, Teunis P, Wellington N, Yakubu H, Moe CL. Quantification of exposure to fecal contamination in open drains in four neighborhoods in Accra, Ghana. *J Water Health*. 2015:1-13. doi:10.2166/wh.2015.138.

87. Chow MF, Yusop Z, Toriman ME. Level and transport pattern of faecal coliform bacteria from tropical urban catchments. *Water Sci Technol*. 2013;67(8):1822-1831. doi:10.2166/wst.2013.048.

88. Ekklesia E, Shanahan P, Chua LHC, Eikaas HS. Temporal variation of faecal

indicator bacteria in tropical urban storm drains. *Water Res*. 2015;68:171-181. doi:10.1016/j.watres.2014.09.049.

89. Rodríguez-Díaz J, Querales L, Caraballo L, Vizzi E, Liprandi F, Takiff H, Betancourt WQ. Detection and characterization of waterborne gastroenteritis viruses in urban sewage and sewage-polluted river waters in Caracas, Venezuela. *Appl Environ Microbiol*. 2009;75(2):387-394. doi:10.1128/AEM.02045-08.

90. Silverman AI, Nelson KL, Akrong MO, Amoah P, Drechsel P. Quantification of human norovirus GII, human adenovirus, and fecal indicator organisms in wastewater used for irrigation in Accra, Ghana. *J Water Health*. 2013;11(3):473-488. doi:10.2166/wh.2013.025.

91. Korkes F, Kumagai FU, Belfort RN, Szejnfeld D, Abud TG, Kleinman A, Florez GM, Szejnfeld T, Chieffi PP. Relationship between Intestinal Parasitic Infection in Children and Soil Contamination in an Urban Slum. *J Trop Pediatr*. 2007;55(1):42-45. doi:10.1093/tropej/fmn038.

92. Wong MS, Bundy DA. Quantitative lumbricoides assessment of contamination and Trichuris trichiura of soil by the eggs of Ascaris. *Trans R Soc Trop Med Hyg*. 1990;84:567-570.

93. Motazedian H, Pakniat A, Tavalali M. Prevalence of helminth ova in soil samples from public places in Shiraz. 2006;12:562-565.

94. Maraghi S, Jafari KM, Sadjjadi SM, Latifi SM, Zibaei M. ENVIRONMENTAL HEALTH Study on the contamination of Abadan public parks soil with Toxocara spp . eggs. 2014:2-4.

95. Ngure FM, Humphrey JH, Mbuya MNN, Majo F, Mutasa K, Govha M, Mazarura

E, Chasekwa B, Prendergast AJ, Curtis V, Boor KJ, Stoltzfus RJ. Formative research on hygiene behaviors and geophagy among infants and young children and implications of exposure to fecal bacteria. *Am J Trop Med Hyg*. 2013;89(4):709-716. doi:10.4269/ajtmh.12-0568.

96. Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PLoS One*. 2014;9(9):e107429. doi:10.1371/journal.pone.0107429.

97. Rompre A, Servais P, Baudart J, De-Roubin M-R, Patrick Laurent P. Detection and enumeration of coliforms in drinking water: current methods and emergin approaches. *J Microbiol Methods*. 2002;49(2002):31-54.

98. Feachem RG, Bradley DJ, Garelick H, Mara DD. *Sanitation and Disease: Health Aspects of Excreta and Wastewater Management*. New York: John Wiley & Sons, Inc.; 1983.

99. Moore MD, Goulter RM, Jaykus L. Human Norovirus as a Foodborne Pathogen: Challenges and Developments. 2015. doi:10.1146/annurev-food-022814-015643.

100. Liu P, Hsaio H-M, Jaykus L-A, Moe C. Quantification of Norwalk Virus Inocula: Comparison of Endpoint Titration and Real-Time Reverse Transcription-PCR Methods. *J Med Virol*. 2010;82(2010):1612-1616. doi:10.1002/jmv.

101. Wilson IG. Inhibition and Facilitation of Nucleic Acid Amplification Inhibition and Facilitation of Nucleic Acid Amplification. *Appl Env Microbiol*. 1997;63(10):3741-3751.

102. UNICEF and World Health Organization. *Progress on Drinking Water and Sanitation: 2012 Update*. (WHO/UNICEF Joint Monitoring Programme for Water

Supply and Sanitation, ed.). WHO/UNICEF; 2012.

103. Alirol E, Getaz L, Stoll B, Chappuis F, Loutan L. Urbanisation and infectious diseases in a globalised world. *Lancet Infect Dis*. 2011;11(2):131-141. doi:10.1016/S1473-3099(10)70223-1.

104. Isunju JB, Schwartz K, Schouten M a, Johnson WP, van Dijk MP. Socio-economic aspects of improved sanitation in slums: a review. *Public Health*. 2011;125(6):368-376. doi:10.1016/j.puhe.2011.03.008.

105. Victora CG, Smith PG, Vaughan JP, Nobre LC, Lombardi C, Teixeira a M, Fuchs SC, Moreira LB, Gigante LP, Barros FC. Water supply, sanitation and housing in relation to the risk of infant mortality from diarrhoea. *Int J Epidemiol*. 1988;17(3):651-654. http://www.ncbi.nlm.nih.gov/pubmed/3209344.

106. D'Souza RMD. Housing and Environmental Factors and Their Effects on the Health of Children in the Slums of Karachi, Pakiston. *J Biosoc Sci*. 1997;29(3):271-281.

107. Katukiza AY, Ronteltap M, Niwagaba CB, Foppen JW a, Kansiime F, Lens PNL. Sustainable sanitation technology options for urban slums. *Biotechnol Adv*. 2012;30(5):964-978. doi:10.1016/j.biotechadv.2012.02.007.

108. Szántó GL, Letema SC, Tukahirwa JT, Mgana S, Oosterveer PJM, van Buuren JCL. Analyzing sanitation characteristics in the urban slums of East Africa. *Water Policy*. 2012;14(4):613. doi:10.2166/wp.2012.093.

109. Mazeau A, Tumwebaze IK, Lüthi C, Sansom K. Inclusion of shared sanitation in urban sanitation coverage? Evidence from Ghana and Uganda. *Waterlines*. 2013;32(4):334-348. doi:10.3362/1756-3488.2013.034.

110. Nelson KL, Murray A. Sanitation for Unserved Populations: Technologies, Implementation Challenges, and Opportunities. *Annu Rev Environ Resour*. 2008;33(1):119-151. doi:10.1146/annurev.environ.33.022007.145142.

111. Schouten M, Mathenge R. Communal sanitation alternatives for slums. A case study of Kibera, Kenya. *Phys Chem Earth, Parts A/B/C*. 2010;35(13):815-822.

112. Fuller J a, Clasen T, Heijnen M, Eisenberg JNS. Shared Sanitation and the Prevalence of Diarrhea in Young Children: Evidence from 51 Countries, 2001-2011. *Am J Trop Med Hyg*. May 2014. doi:10.4269/ajtmh.13-0503.

113. Peprah D, Baker KK, Moe C, Robb K, Wellington N, Yakubu H, Null C. Public toilets and their customers in low-income Accra, Ghana. *Environ Urban*. 2015;27(5):1-16. doi:10.1177/0956247815595918.

114. UNICEF and World Health Organization. *Post-2015 Proposal for Sanitation Targets and Indicators*. Geneva; 2012.

115. Heijnen M, Cumming O, Peletz R, Chan GK-S, Brown J, Baker K, Clasen T. Shared Sanitation versus Individual Household Latrines: A Systematic Review of Health Outcomes. *PLoS One*. 2014;9(4):e93300. doi:10.1371/journal.pone.0093300.

116. Exley JLR, Liseka B, Cumming O, Ensink JHJ. The Sanitation Ladder, What Constitutes an Improved Form of Sanitation? *Environ Sci Technol*. 2015;49(2):1086-1094. doi:10.1021/es503945x.

117. Norman G, Pedley S, Takkouche B. Effects of sewerage on diarrhoea and enteric infections: a systematic review and meta-analysis. *Lancet Infect Dis*. 2010;10(8):536-544. doi:10.1016/S1473-3099(10)70123-7.

118. Baum R, Luh J, Bartram J. Sanitation: a global estimate of sewerage connections without treatment and the resulting impact on MDG progress. *Environ Sci Technol*. 2013;47(4):1994-2000. doi:10.1021/es304284f.

119. Ferrer A, Nguyen-Viet H, Zinsstag J. Quantification of diarrhea risk related to wastewater contact in Thailand. *Ecohealth*. 2012;9(1):49-59. doi:10.1007/s10393-012-0746-x.

120. Moraes LR, Cancio JA, Cairncross S, Huttly S. Impact of drainage and sewerage on diarrhoea in poor urban areas in Salvador, Brazil. *Trans R Soc Trop Med Hyg*. 2003;97(2):153-158. http://www.ncbi.nlm.nih.gov/pubmed/14584367.

121. Moraes LR., Cancio JA, Cairncross S. Impact of drainage and sewerage on intestinal nematode infections in poor urban areas in Salvador, Brazil. *Trans R Soc Trop Med Hyg*. 2004;98(4):197-204. doi:10.1016/S0035-9203(03)00043-9.

122. Unger A, Riley LW. Slum health: From understanding to action. *PLoS Med*. 2007;4(10):1561-1566. doi:10.1371/journal.pmed.0040295.

123. Turley R, Saith R, Bhan N, Rehfuess E, Carter B, Turley R, Saith R, Bhan N, Rehfuess E, Carter B. Slum upgrading strategies involving physical environment and infrastructure interventions and their effects on health and socio-economic outcomes (Review) Slum upgrading strategies involving physical environment and infrastructure interventions and thei. 2013;(1). doi:10.1002/14651858.CD010067.pub2.Copyright.

124. Teunis PFM, Reese HE, Null C, Yakubu H, Moe CL. Quantifying Contact with the Environment: Behaviors of Young Children in Accra, Ghana. *Am J Trop Med Hyg*. 2016. doi:10.4269/ajtmh.15-0417.

125. Barker SF, Amoah P, Drechsel P. A probabilistic model of gastroenteritis risks associated with consumption of street food salads in Kumasi, Ghana: Evaluation of methods to estimate pathogen dose from water, produce or food quality. *Sci Total Environ*. 2014;487(1):130-142. doi:10.1016/j.scitotenv.2014.03.108.

126. Amoah P, Drechsel P, Abaidoo RC, Henseler M. Irrigated urban vegetable production in Ghana: Microbiological contamination in farms and markets and associated consumer risk groups. *J Water Health*. 2007;5(3):455-466. doi:10.2166/wh.2007.041.

127. Antwi-Agyei P, Cairncross S, Peasey A, Price V, Bruce J, Baker K, Moe C, Ampofo J, Armah G, Ensink J. A farm to fork risk assessment for the use of wastewater in agriculture in Accra, Ghana. *PLoS One*. 2015;10(11):1-19. doi:10.1371/journal.pone.0142346.

128. Torgerson PR, Devleesschauwer B, Praet N, Speybroeck N, Willingham AL, Kasuga F, Rokni MB, Zhou XN, Fevre EM, Sripa B, Gargouri N, Furst T, Budke CM, Carabin H, Kirk MD, Angulo FJ, Havelaar A, de Silva N. World Health Organization Estimates of the Global and Regional Disease Burden of 11 Foodborne Parasitic Diseases, 2010: A Data Synthesis. *PLoS Med*. 2015;12(12):1-22. doi:10.1371/journal.pmed.1001920.

129. Perry S, De La Luz Sanchez M, Hurst PK, Parsonnet J. Household transmission of gastroenteritis. *Emerg Infect Dis*. 2005;11(7):1093-1096. doi:10.3201/eid1107.040889.

130. Ferrer SR, Strina A, Jesus SR, Ribeiro HC, Cairncross S, Rodrigues LC, Barreto ML. A hierarchical model for studying risk factors for childhood diarrhoea: A

case-control study in a middle-income country. *Int J Epidemiol*. 2008;37(4):805-815. doi:10.1093/ije/dyn093.

131. Haas, CN., Rose, JB., Gerba C. *Quantitative Microbial Risk Assessment*. New York: John Wiley & Sons, Inc.; 1999.

132. Simpson JM, Santo Domingo JW, Reasoner DJ. Microbial source tracking: State of the science. *Environ Sci Technol*. 2002;36(24):5279-5288. doi:10.1021/es026000b.

133. Jenkins MW, Tiwari S, Lorente M, Gichaba CM, Wuertz S. Identifying human and livestock sources of fecal contamination in Kenya with host-specific Bacteroidales assays. *Water Res*. 2009;43(19):4956-4966. doi:10.1016/j.watres.2009.07.028.

134. Ghana Statistical Service. 2010 Population & Housing Census: Accra Metropolitan District Analytical Report. 2014:78.

135. Directorate of Census Operations, Operations TN. District Census Handbook Vellore. 2011;(28). http://www.censusindia.gov.in/2011census/dchb/0966_PART_B_DCHB_0966_VARANASI.pdf.

136. John SM, Thomas RJ, Kaki S, Sharma SL, Ramanujam K, Raghava M V, Koshy B, Bose A, Rose A, Rose W, Ramachandran A, Joseph a J, Babji S, Kang G. Establishment of the MAL-ED Birth Cohort Study Site in Vellore, Southern India. *Clin Infect Dis*. 2014;59 Suppl 4(Suppl 4):S295-S299. doi:10.1093/cid/ciu390.

137. Government of India M of HA. Census of India. http://www.censusindia.gov.in/2011census/HLO/HL_PCA/Houselisting-housing-TM.html. Published 2011. Accessed July 6, 2016.

138. Mara D, Lane J, Scott B, Trouba D. Sanitation and Health. *PLoS Med*. 2010;7(11):e1000363. doi:10.1371/journal.pmed.1000363.

139. United Nations. *World Urbanization Prospects: 2014 Revision*. New York; 2015.

140. United Nations Human Settlements Program. *The Challenge of Slums*. Sterling, VA: Earthscan Publications Ltd; 2003.

141. Turley R, Saith R, Bhan N, Rehfuess E, Carter B. *Slum Upgrading Strategies Involving Physical Environment and Infrastructure Interventions and Their Effects on Health and Socio-Economic Outcomes (Review)*.; 2013.

142. Wolf J, Prüss-Ustün A, Cumming O, Bartram J, Bonjour S, Cairncross S, Clasen T, Colford JM, Curtis V, De France J, Fewtrell L, Freeman MC, Gordon B, Hunter PR, Jeandron A, Johnston RB, Mäusezahl D, Mathers C, Neira M, Higgins JPT. Systematic review: Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: Systematic review and meta-regression. *Trop Med Int Heal*. 2014;19(8):928-942. doi:10.1111/tmi.12331.

143. Baum R, Luh J, Bartram J. Sanitation: A global estimate of sewerage connections without treatment and the resulting impact on MDG progress. *Environ Sci Technol*. 2013;47(4):1994-2000. doi:10.1021/es304284f.

144. Fuller JA, Clasen T, Heijnen M, Eisenberg JNS. Shared sanitation and the prevalence of diarrhea in young children: evidence from 51 countries, 2001-2011. *Am J Trop Med Hyg*. 2014;91(1):173-180. doi:10.4269/ajtmh.13-0503.

145. Heijnen M, Rosa G, Fuller J, Eisenberg JNS, Clasen T. The geographic and demographic scope of shared sanitation: an analysis of national survey data from low- and middle-income countries. *Trop Med Int Health*. 2014;19(11):1334-1345.

doi:10.1111/tmi.12375.

146. Rheinländer T, Konradsen F, Keraita B, Gyapong M. Redefining shared sanitation. *Bull World Health Organ*. 2015:1-6.

147. Armah GE, Gallimore CI, Binka FN, Asmah RH, Green J, Ugoji U, Anto F, Brown DWG, Gray JJ. Characterisation of Norovirus Strains in Rural Ghanaian Children with Acute Diarrhoea. *J Med Virol*. 2006;78:1480-1485. doi:10.1002/jmv.

148. Reither K, Ignatius R, Weitzel T, Seidu-Korkor A, Anyidoho L, Saad E, Djie-Maletz A, Ziniel P, Amoo-Sakyi F, Danikuu F, Danour S, Otchwemah RN, Schreier E, Bienzle U, Stark K, Mockenhaupt FP. Acute childhood diarrhoea in northern Ghana: epidemiological, clinical and microbiological characteristics. *BMC Infect Dis*. 2007;7:104. doi:10.1186/1471-2334-7-104.

149. USEPA. *Method 1604: Total Coliforms and Escherichia Coli in Water by Membrane Filtration Using a Simultaneous Detection Medium (MI Medium)*.; 2002.

150. Jothikumar N, Cromeans TL, Vincent R, Lu X, Sobsey MD, Erdman DD, Hill VR. Quantitative Real-Time PCR Assays for Detection of Human Adenoviruses and Identification of Serotypes 40 and 41 Quantitative Real-Time PCR Assays for Detection of Human Adenoviruses and Identification of Serotypes 40 and 41. *Appl Env Microbiol*. 2005;6(71):3131-3136. doi:10.1128/AEM.71.6.3131.

151. Kageyama T, Kojima S, Shinohara M, Uchida K, Fukushi S, Hoshino FB, Takeda N, Katayama K. Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *J Clin*

Microbiol. 2003;41(4):1548-1557. doi:10.1128/JCM.41.4.1548.

152. WHO and UNICEF. *Progress on Sanitation and Drinking Water: 2013 Update*.; 2013.

153. Geography Department SDSU. Personal communication. 2015.

154. Genser B, Strina A, dos Santos L a, Teles C a, Prado MS, Cairncross S, Barreto ML. Impact of a city-wide sanitation intervention in a large urban centre on social, environmental and behavioural determinants of childhood diarrhoea: analysis of two cohort studies. *Int J Epidemiol*. 2008;37(4):831-840. doi:10.1093/ije/dyn101.

155. Rego RF, Barreto ML. Environmental indicators of intra-urban hetererogeneity Indicadores ambientais para detectar heterogeneidade intraurbana Indicadores ambientales para detectar la heterogeneidad intra-urbana. 2013;29(6):1173-1185.

156. Milroy C a. Evaluating sanitary quality and classifying urban sectors according to environmental conditions. *Environ Urban*. 2001;13(1):235-255. doi:10.1177/095624780101300117.

157. Heinze G, Ploner M, Dunkler D, Southworth H. logistf: Firth's bias reduced logistic regression. 2013:v. 1.2.1. https://cran.r-project.org/package=logistf.

158. Firth D. Bias Reduction of Maximum Likelihood Estimates Author (s): David Firth Published by : Biometrika Trust Stable URL : http://www.jstor.org/stable/2336755. *Biometrika*. 1993;80(1):27-38. doi:10.1093/biomet/80.1.27.

159. Kleinbaum D, Klein M. *Logistic Regression: A Self-Learning Text*. Third. New York: Springer; 2010.

160. Diallo MBC, Anceno AJ, Tawatsupa B, Tripathi NK, Wangsuphachart V, Shipin

O V. GIS-based analysis of the fate of waste-related pathogens Cryptosporidium parvum , Giardia lamblia and Escherichia coli in a tropical canal network. *J Water Health*. 2009;7(1):133-143. doi:10.2166/wh.2009.010.

161. Sidhu JPS, Ahmed W, Gernjak W, Aryal R, McCarthy D, Palmer A, Kolotelo P, Toze S. Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *Sci Total Environ*. 2013;463-464:488-496. doi:10.1016/j.scitotenv.2013.06.020.

162. Nshimyimana JP, Ekklesia E, Shanahan P, Chua LHC, Thompson JR. Distribution and abundance of human-specific Bacteroides and relation to traditional indicators in an urban tropical catchment. *J Appl Microbiol*. 2014;(Pruss 1998):n/a - n/a. doi:10.1111/jam.12455.

163. Ekklesia E, Shanahan P, Chua LHC, Eikaas HS. Associations of chemical tracers and faecal indicator bacteria in a tropical urban catchment. *Water Res*. 2015;75:270-281. doi:10.1016/j.watres.2015.02.037.

164. Leclerc H, Mossel D a, Edberg SC, Struijk CB. Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annu Rev Microbiol*. 2001;55:201-234. doi:10.1146/annurev.micro.55.1.201.

165. Pina S, Puig M, Lucena F, Jofre J, Girones R. Viral pollution in the environment and in shellfish: Human adenovirus detection by PCR as an index of human viruses. *Appl Environ Microbiol*. 1998;64(9):3376-3382.

166. Hundesa A, Maluquer De Motes C, Bofill-Mas S, Albinana-Gimenez N, Girones R. Identification of human and animal adenoviruses and polyomaviruses for determination of sources of fecal contamination in the environment. *Appl Environ*

Microbiol. 2006;72(12):7886-7893. doi:10.1128/AEM.01090-06.

167. Rochelle-Newall E, Nguyen TMH, Le TPQ, Sengtaheuanghoung O, Ribolzi O. A short review of fecal indicator bacteria in tropical aquatic ecosystems: knowledge gaps and future directions. *Front Microbiol*. 2015;6(April):308. doi:10.3389/fmicb.2015.00308.

168. Kanungo S, Sur D, Ali M, You YA, Pal D, Manna B, Niyogi SK, Sarkar B, Bhattacharya SK, Clemens JD, Nair GB. Clinical, epidemiological, and spatial characteristics of Vibrio parahaemolyticus diarrhea and cholera in the urban slums of Kolkata, India. *BMC Public Health*. 2012;12(1):830. doi:10.1186/1471-2458-12-830.

169. Munjita SM. Current Status of Norovirus Infections in Children in Sub-Saharan Africa. *J Trop Med*. 2015;2015:1-5.

170. Prüss A, Kay D, Fewtrell L, Bartram J. Estimating the burden of desease from water, sanitation and hygiene at a global level. *Environ Health Perspect*. 2002;110(5):537-542. doi:10.1289/ehp.02110537.

171. Strande L, Ronteltap M, Brdjanovic D. *Faecal Sludge Management Systems Approachs*.; 2014.

172. Emory University. SaniPath. www.sanipath.org. Published 2014. Accessed November 21, 2015.

173. John J, Giri S, Karthikeyan AS, Iturriza-Gomara M, Muliyil J, Abraham A, Grassly NC, Kang G. Effect of a single inactivated poliovirus vaccine dose on intestinal immunity against poliovirus in children previously given oral vaccine: An open-label, randomised controlled trial. *Lancet*. 2014;384(9953):1505-1512.

doi:10.1016/S0140-6736(14)60934-X.

174. Gladstone BP, Muliyil JP, Jaffer S, Wheeler JG, Le Fevre A, Iturriza-Gomara M, Gray JJ, Bose A, Estes MK, Brown DW, Kang G. Infant Mortality in an Urban Slum Birth Cohort. *Arch Dis Child*. 2008;93(6):479-484. doi:10.1136/adc.2006.114546.Infant.

175. Collinet-Adler S, Babji S, Francis M, Kattula D, Premkumar PS, Sarkar R, Mohan VR, Ward H, Kang G, Balraj V, Naumova EN. Environmental Factors Associated with High Fly Densities and Diarrhea in Vellore, India. *Appl Environ Microbiol*. 2015;81(17):6053-6058. doi:10.1128/AEM.01236-15.

176. Platts-Mills JA, Liu J, Gratz J, Mduma E, Amour C, Swai N, Taniuchi M, Begum S, Yori PP, Tilley DH, Lee G, Shen Z, Whary MT, Fox JG, McGrath M, Kosek M, Haque R, Houpt ER. Detection of Campylobacter in stool and determination of significance by culture, enzyme immunoassay, and PCR in developing countries. *J Clin Microbiol*. 2014;52(4):1074-1080. doi:10.1128/JCM.02935-13.

177. Liu P, Hill VR, Hahn D, Johnson TB, Pan Y, Jothikumar N, Moe CL. Hollow-fiber ultrafiltration for simultaneous recovery of viruses, bacteria and parasites from reclaimed water. *J Microbiol Methods*. 2012;88(1):155-161. doi:10.1016/j.mimet.2011.11.007.

178. R Core Team. R: A language and environment for statistical computing. 2015. https://www.r-project.org/.

179. Bates D, Maechler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw*. 2014;67(1):1-48. http://arxiv.org/abs/1406.5823.

180. Barreto ML, Genser B, Strina A, Teixeira MG, Assis AMO, Rego RF, Teles C a,

Prado MS, Matos SM a, Santos DN, dos Santos L a, Cairncross S. Effect of city-wide sanitation programme on reduction in rate of childhood diarrhoea in northeast Brazil: assessment by two cohort studies. *Lancet*. 2007;370(9599):1622-1628. doi:10.1016/S0140-6736(07)61638-9.

181. Kolahi A-A, Rastegarpour A, Sohrabi M-R. The impact of an urban sewerage system on childhood diarrhoea in Tehran, Iran: a concurrent control field trial. *Trans R Soc Trop Med Hyg*. 2009;103(5):500-505. doi:10.1016/j.trstmh.2008.10.016.

182. Santos C a ST, Strina A, Amorim LD, Genser B, Assis a MO, Prado MS, Barreto ML. Individual and contextual determinants of the duration of diarrhoeal episodes in preschool children: a longitudinal study in an urban setting. *Epidemiol Infect*. 2012;140(4):689-696. doi:10.1017/S0950268811000690.

183. Koné D. Making urban excreta and wastewater management contribute to cities ' economic development : a paradigm shift. *Water Policy*. 2010;12:602-610. doi:10.2166/wp.20l0.122.

184. Stanton BF, Clemens JD, Aziz KM a, Rahman M. Twenty-four-hour recall, knowledge-attitude-practice questionnaires, and direct observations of sanitary practices: A comparative study. *Bull World Health Organ*. 1987;65(2):217-222.

185. Manun'Ebo M, Cousens S, Haggerty P, Kalengaie M, Ashworth a, Kirkwood B. Measuring hygiene practices: a comparison of questionnaires with direct observations in rural Zaïre. *Trop Med Int Health*. 1997;2(11):1015-1021. doi:10.1046/j.1365-3156.1997.d01-180.x.

186. Vindigni SM, Riley PL, Jhung M. Systematic review: handwashing behaviour in

low- to middle-income countries: outcome measures and behaviour maintenance. *Trop Med Int Health*. 2011;16(4):466-477. doi:10.1111/j.1365-3156.2010.02720.x.

187. Hacker KP, Seto KC, Costa F, Corburn J, Reis MG, Ko AI, Diuk-Wasser M a. Urban slum structure: integrating socioeconomic and land cover data to model slum evolution in Salvador, Brazil. *Int J Health Geogr*. 2013;12(1):45. doi:10.1186/1476-072X-12-45.

188. Jenkins MW, Cumming O, Scott B, Cairncross S. Beyond "improved" towards "safe and sustainable" urban sanitation: assessing the design, management and functionality of sanitation in poor communities of Dar es Salaam, Tanzania. *J Water, Sanit Hyg Dev*. 2014;4(1):131. doi:10.2166/washdev.2013.180.

189. The MAL-ED Network Investigators. The MAL-ED study: a multinational and multidisciplinary approach to understand the relationship between enteric pathogens, malnutrition, gut physiology, physical growth, cognitive development, and immune responses in infants and children up to 2 years of. *Clin Infect Dis*. 2014;59 Suppl 4(Suppl 4):S193-S206. doi:10.1093/cid/ciu653.

190. Savovic J, Jones HE, Altman DG, Harris RJ, Pildal J, Als-nielsen B, Balk EM, Gluud C, Gluud LL, Ioannidis JP a, Schulz KF, Beynon R, Welton NJ, Wood L, Moher D, Deeks JJ, Sterne J a C. Research and Reporting Methods Influence of Reported Study Design Characteristics on Intervention. *Ann Intern Med*. 2012;157(6):429-438.

191. Freeman MC, Stocks ME, Cumming O, Jeandron A, Higgins JPT, Wolf J, Pruss-Ustin A, Bonjour S, Hunter PR, Fewtrell L, Curtis V. Hygiene and health: Systematic review of handwashing practices worldwide and update of health

effects. *Trop Med Int Heal*. 2014;19(8):906-916. doi:10.1111/tmi.12339.

192. Burra S. Community-designed, built and managed toilet blocks in Indian cities. *Environ Urban*. 2003;15(2):11-32. doi:10.1177/095624780301500202.

193. Brick T, Primrose B, Chandrasekhar R, Roy S, Muliyil J, Kang G. Water contamination in urban south India: household storage practices and their implications for water safety and enteric infections. *Int J Hyg Environ Health*. 2004;207(5):473-480. doi:10.1078/1438-4639-00318.

194. Kennedy-Walker R, Holderness T, Alderson D, Evans B, Barr S. Network modelling for road-based faecal sludge management. *Proc Inst Civ Eng Eng*. 2015;167(ME3):157-165.

195. Leitão JP, Almeida MDC, Simões NE, Martins A. Methodology for qualitative urban flooding risk assessment. *Water Sci Technol*. 2013;68(4):829-838. doi:10.2166/wst.2013.310.

196. Kampf G, Kramer A. Epidemiologic Background of Hand Hygiene and Evaluation of the Most Important Agents for Scrubs and Rubs. *Clin Microbiol Rev*. 2004;17(4):863-893. doi:10.1128/CMR.17.4.863.

197. Strunz EC, Addiss DG, Stocks ME, Ogden S, Utzinger J, Freeman MC. Water, Sanitation, Hygiene, and Soil-Transmitted Helminth Infection: A Systematic Review and Meta-Analysis. *PLoS Med*. 2014;11(3). doi:10.1371/journal.pmed.1001620.

198. Dillingham RA, Lima AA, Guerrant RL. Cryptosporidiosis: Epidemiology and impact. *Microbes Infect*. 2002;4(10):1059-1066. doi:10.1016/S1286-4579(02)01630-1.

199. Jones EL, Kramer A, Gaither M, Gerba CP. Role of fomite contamination during an outbreak of norovirus on houseboats. *Int J Environ Health Res*. 2007;17(2):123-131. doi:10.1080/09603120701219394.

200. Richard SA, Barrett LJ, Guerrant RL, Checkley W, Miller MA. Disease Surveillance Methods Used in the 8-Site MAL-ED Cohort Study. *Clin Infect Dis*. 2014;59(Suppl 4):S220-S224. doi:10.1093/cid/ciu435.

201. Akaike H. Likelihood of a model and information criteria. *J Econom*. 1981;16(1):3-14.

202. Kattula D, Francis MR, Kulinkina A, Sarkar R, Mohan VR, Babji S, Ward HD, Kang G, Balraj V, Naumova EN. Environmental predictors of diarrhoeal infection for rural and urban communities in south India in children and adults. *Epidemiol Infect*. February 2015:1-12. doi:10.1017/S0950268814003562.

203. Checkley W, Gilman RH, Black RE, Epstein LD, Cabrera L, Sterling CR, Moulton LH. Effect of water and sanitation on childhood health in a poor Peruvian peri-urban community. *Lancet*. 2004;363(9403):112-118. doi:10.1016/S0140-6736(03)15261-0.

204. Sobel J, Gomes T a T, Ramos RTS, Hoekstra M, Rodrigue D, Rassi V, Griffin PM. Pathogen-specific risk factors and protective factors for acute diarrheal illness in children aged 12-59 months in São Paulo, Brazil. *Clin Infect Dis*. 2004;38(11):1545-1551. doi:10.1086/420822.

205. Mascarini-Serra LM, Telles CA, Prado MS, Mattos SA, Strina A, Alcantara-Neves NM, Barreto ML. Reductions in the prevalence and incidence of geohelminth infections following a city-wide sanitation program in a Brazilian Urban Centre.

PLoS Negl Trop Dis. 2010;4(2). doi:10.1371/journal.pntd.0000588.

206. Stenström TA, Seidu R, Nelson E, Christian Z. *Microbial Exposure and Health Assessments in Sanitation Technologies and Systems*.; 2011. www.ecosanres.org.

207. Kightlinger LK, Seed JR, Kightlinger MB. Ascaris lumbricoides Intensity in Relation to Environmental , Socioeconomic , and Behavioral Determinants of Exposure to Infection in Children from Southeast Madagascar. 2013;84(3):480-484.

208. Werber D, Mason BW, Evans MR, Salmon RL. Preventing Household Transmission of Shiga Toxin–Producing Escherichia coli O157 Infection: Promptly Separating Siblings Might Be the Key. *Clin Infect Dis*. 2008;46(8):1189-1196. doi:10.1086/587670.

209. Luby SP, Halder AK, Tronchet C, Akhter S, Bhuiya A, Johnston RB. Household characteristics associated with handwashing with soap in rural Bangladesh. *Am J Trop Med Hyg*. 2009;81(5):882-887. doi:10.4269/ajtmh.2009.09-0031.

210. Nallari A. "All we want are toilets inside our homes!": The critical role of sanitation in the lives of urban poor adolescent girls in Bengaluru, India. *Environ Urban*. 2015;27(1):1-16. doi:10.1177/0956247814563514.

211. Berendes D, Beno D, Clennon J, Ghale B, Gunasekaran A, Kang G, Kartikeyan A, Kirby A, Kumar JS, Mohan V, Raj S, Roy S, Yakubu H, Moe C. Neighborhood Structure, Household Behaviors, and Their Effects on Spatial Clustering of Risk of Exposure to Fecal Contamination in Urban Flood Areas. In: *Water and Health Conference*. ; 2015:100-101.

212. Luby SP, Halder AK, Huda TM, Unicomb L, Islam MS, Arnold BF, Johnston RB.

Microbiological contamination of drinking water associated with subsequent child diarrhea. *Am J Trop Med Hyg*. 2015;93(5):904-911. doi:10.4269/ajtmh.15-0274.

213. Rosa G, Clasen T. Estimating the scope of household water treatment in low- and medium-income countries. *Am J Trop Med Hyg*. 2010;82(2):289-300. doi:10.4269/ajtmh.2010.09-0382.

214. Gopal S, Sarkar R, Banda K, Govindarajan J, Harijan BB, Jeyakumar MB, Mitta P, Sadanala ME, Selwyn T, Suresh CR, Thomas V a, Devadason P, Kumar R, Selvapandian D, Kang G, Balraj V. Study of water supply & sanitation practices in India using geographic information systems: some design & other considerations in a village setting. *Indian J Med Res*. 2009;129(3):233-241. http://www.ncbi.nlm.nih.gov/pubmed/19491414.

215. Kirby AE, Shi J, Montes J, Lichtenstein M, Moe CL. Disease Course and Viral Shedding in Experimental Norwalk Virus and Snow Mountain Virus Infection. *J Med Virol*. 2014;86:2055-2064. doi:10.1002/jmv.

216. Liu P, Yuen Y, Hsiao HM, Jaykus LA, Moe C. Effectiveness of liquid soap and hand sanitizer against Norwalk virus on contaminated hands. *Appl Environ Microbiol*. 2010;76(2):394-399. doi:10.1128/AEM.01729-09.

217. Curtis VA, Danquah LO, Aunger R V. Planned, motivated and habitual hygiene behaviour: An eleven country review. *Health Educ Res*. 2009;24(4):655-673. doi:10.1093/her/cyp002.

218. Contzen N, Meili IH, Mosler HJ. Changing handwashing behaviour in southern ethiopia: A longitudinal study on infrastructural and commitment interventions. *Soc Sci Med*. 2015;124:103-114. doi:10.1016/j.socscimed.2014.11.006.

219. Heijnen M, Greenland K. *Level of Behaviour Change Achievable by Handwashing with Soap Interventions: A Rapid Review*.; 2015. doi:10.12774/eod_hd.april2015.heijnenmetal.

220. Zambrano LD, Levy K, Menezes NP, Freeman MC. Human diarrhea infections associated with domestic animal husbandry: A systematic review and meta-analysis. *Trans R Soc Trop Med Hyg*. 2014;108(6):313-325. doi:10.1093/trstmh/tru056.

221. Bassan M, Brdjanovic D, Dangol B, Dodane P-H, Hooijmans CM, Manuel Lopez-Vazquez C, Mbeguere M, Mikhael G, Moya Diaz-Aguado B, Niwagaba CB, Magloire Kenge I, Ramsay JE, Reymond P, Robbins DM, Ronteltap M, Strande L, Tilley E. *Faecal Sludge Management*. (Strande L, Ronteltap M, Brdjanovic D, eds.). IWA; 2014.

222. Biran A, Rabie T, Schmidt W, Juvekar S, Hirve S, Curtis V. Comparing the performance of indicators of hand-washing practices in rural Indian households. *Trop Med Int Heal*. 2008;13(2):278-285. doi:10.1111/j.1365-3156.2007.02001.x.

223. Carlton EJ, Woster AP, DeWitt P, Goldstein RS, Levy K. A systematic review and meta-analysis of ambient temperature and diarrhoeal diseases. *Int J Epidemiol*. 2015;(November 2015):1-14. doi:10.1093/ije/dyv296.

224. Mellor JE, Levy K, Zimmerman J, Elliott M, Bartram J, Carlton E, Clasen T, Dillingham R, Eisenberg J, Guerrant R, Lantagne D, Mihelcic J, Nelson K. Planning for climate change: The need for mechanistic systems-based approaches to study climate change impacts on diarrheal diseases. *Sci Total Environ*. 2016;548-549(January):82-90. doi:10.1016/j.scitotenv.2015.12.087.

225. Carlton EJ, Eisenberg JNS, Goldstick J, Cevallos W, Trostle J, Levy K. Heavy rainfall events and diarrhea incidence: The role of social and environmental factors. *Am J Epidemiol*. 2014;179(3):344-352. doi:10.1093/aje/kwt279.

226. Philipsborn R, Ahmed SM, Brosi BJ, Levy K. Climatic drivers of diarrheagenic Escherichia coli: A systematic review and meta-analysis. *J Infect Dis*. 2016:1-29.

227. Kostyla C, Bain R, Cronk R, Bartram J. Seasonal variation of fecal contamination in drinking water sources in developing countries: A systematic review. *Sci Total Environ*. 2015;514:333-343. doi:10.1016/j.scitotenv.2015.01.018.

228. Widmer K, Van Ha NT, Vinitnantharat S, Sthiannopkao S, Wangsaatmaja S, Prasetiati MAN, Thanh NC, Thepnoo K, Sutadian AD, Thao HTT, Fapyane D, San V, Vital P, Hur H-G. Prevalence of Escherichia coli in surface waters of Southeast Asian cities. *World J Microbiol Biotechnol*. 2013;29:2115-2124. doi:10.1007/s11274-013-1376-3.

229. Ruiz-Moreno D, Pascual M, Emch M, Yunus M. Spatial clustering in the spatio-temporal dynamics of endemic cholera. *BMC Infect Dis*. 2010;10:51. doi:10.1186/1471-2334-10-51.

230. Fobil J, May J, Kraemer A. Assessing the relationship between socioeconomic conditions and urban environmental quality in Accra, Ghana. *Int J Environ Res Public Health*. 2010;7(1):125-145. doi:10.3390/ijerph7010125.

231. Ramani S, Banerjee I, Gladstone BP, Sarkar R, Selvapandian D, Le Fevre AM, Jaffar S, Iturriza-Gomara M, Gray JJ, Estes MK, Brown DW, Kang G. Geographic information systems and genotyping in identification of rotavirus G12 infections in residents of an urban slum with subsequent detection in hospitalized children:

emergence of G12 genotype in South India. *J Clin Microbiol*. 2007;45(2):432-437. doi:10.1128/JCM.01710-06.

232. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. Third. Philadelphia: Wolters Kluwer Health; 2008.

233. Beale L, Abellan JJ, Hodgson S, Jarup L. Methodologic issues and approaches to spatial epidemiology. *Environ Health Perspect*. 2008;116(8):1105-1110. doi:10.1289/ehp.10816.

234. Rajaraman D, Varadharajan KS, Greenland K, Curtis V, Kumar R, Schmidt W-P, Aunger R, Biran A. Implementing effective hygiene promotion: lessons from the process evaluation of an intervention to promote handwashing with soap in rural India. *BMC Public Health*. 2014;14:1179. doi:10.1186/1471-2458-14-1179.

235. Ejemot-Nwadiaro RI, Ehiri JE, Meremikwu MM, Critchley J a. Hand washing for preventing diarrhoea (Review). *Cochrane Database Syst Rev*. 2012;(2):Art. No.: CD004265. doi:http://dx.doi.org/10.1002/14651858.CD004265.pub2.

236. Dreibelbis R, Kroeger A, Hossain K, Venkatesh M, Ram PK. Behavior Change without Behavior Change Communication: Nudging Handwashing among Primary School Students in Bangladesh. *Int J Env Res Public Heal*. 2016;13(1):7-9. doi:10.3390/ijerph13010129.

237. Exum NG, Olórtegui MP, Yori PP, Davis MF, Heaney CD, Kosek M, Schwab KJ. Floors and toilets: Association of floors and sanitation practices with fecal contamination in Peruvian Amazon peri-urban households. *Environ Sci Technol*. 2016. doi:10.1021/acs.est.6b01283.

238. Glass RI, Parashar U, Patel M, Gentsch J, Jiang B. Rotavirus vaccines: Successes

and challenges. *J Infect*. 2014;68(SUPPL1):S9-S18. doi:10.1016/j.jinf.2013.09.010.

239. Breiman RF, Zaman K, Armah G, Sow SO, Anh DD, Victor JC, Hille D, Ciarlet M, Neuzil KM. Analyses of health outcomes from the 5 sites participating in the Africa and Asia clinical efficacy trials of the oral pentavalent rotavirus vaccine. *Vaccine*. 2012;30(SUPPL. 1):24-29. doi:10.1016/j.vaccine.2011.08.124.

240. Armah GE, Sow SO, Breiman RF, Dallas MJ, Tapia MD, Feikin DR, Binka FN, Steele AD, Laserson KF, Ansah NA, Levine MM, Lewis K, Coia ML, Attah-Poku M, Ojwando J, Rivers SB, Victor JC, Nyambane G, Hodgson A, Schodel F, Ciarlet M, Neuzil KM. Efficacy of pentavalent rotavirus vaccine against severe rotavirus gastroenteritis in infants in developing countries in sub-Saharan Africa: A randomised, double-blind, placebo-controlled trial. *Lancet*. 2010;376(9741):606-614. doi:10.1016/S0140-6736(10)60889-6.

241. Lopman BA, Pitzer VE, Sarkar R, Gladstone B, Patel M, Glasser J, Gambhir M, Atchison C, Grenfell BT, Edmunds WJ, Kang G, Parashar UD. Understanding reduced rotavirus vaccine efficacy in low socio-economic settings. *PLoS One*. 2012;7(8). doi:10.1371/journal.pone.0041720.

242. Petri WA, Miller M, Binder HJ, Levine MM, Dillingham R, Guerrant RL. Enteric infections, diarrhea, and their impact on function and development. *J Clin Invest*. 2008;118(4):1277-1290. doi:10.1172/JCI34005.

243. Chen D, Getis A. Point Pattern Analysis. 1998.

244. Getis A, Ord JK. The Analysis of Spatial Association by Use of Distance Statistics. *Geogr Anal*. 1992;24(3):189-206. doi:10.1111/j.1538-

4632.1992.tb00261.x.

245. Diggle PJ. A Point Process Modelling Approach to Raised Incidence of a Rare Phenomenon in the Vicinity of a Prespecified Point. *JR Stat Soc*. 1990;153(3):349-362.

246. Clasen T, Boisson S, Routray P, Torondel B, Bell M, Cumming O, Ensink J, Freeman M, Jenkins M, Odagiri M, Ray S, Sinha A, Suar M, Schmidt W-P. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *Lancet Glob Heal*. 2014;2(11):e645-e653. doi:10.1016/S2214-109X(14)70307-9.

247. Patil SR, Arnold BF, Salvatore AL, Briceno B, Ganguly S, Colford JM, Gertler PJ. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: A cluster randomized controlled trial. *PLoS Med*. 2015;11(8). doi:10.1371/journal.pmed.1001709.

248. Kleinbaum D, Klein M. *Survival Analysis*. Second edi. New York: Springer; 2005.

249. Lamberti LM, Fischer Walker CL, Noiman A, Victora C, Black RE. Breastfeeding and the risk for diarrhea morbidity and mortality. *BMC Public Health*. 2011;11 Suppl 3(Suppl 3):S15. doi:10.1186/1471-2458-11-S3-S15.

250. World Health Organization. Indicators for assessing breastfeeding practices. In: 91st ed. ; 1991.