Environmental Science & Technology

Characterization of Natural and Affected Environments

Fecal Fingerprints of Enteric Pathogen Contamination in Public Environments of Kisumu, Kenya associated with Human Sanitation Conditions and Domestic Animals

Kelly K. Baker, Reid Senesac, Daniel Sewell, Ananya Sen Gupta, Oliver Cumming, and Jane Mumma

Environ. Sci. Technol., **Just Accepted Manuscript** • DOI: 10.1021/acs.est.8b01528 • Publication Date (Web): 14 Aug 2018

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Fecal Fingerprints of Enteric Pathogen Contamination in Public Environments of Kisumu, Kenya associated with Human Sanitation Conditions and Domestic Animals

Kelly K. Baker, [1*] Reid Senesac, [1] Daniel Sewell, [2] Ananya Sen Gupta, [3] Oliver Cumming,[4] Jane Mumma [5]

[1] Department of Occupational and Environmental Health, University of Iowa, Iowa City, Iowa, United States

[2] Department of Biostatistics, University of Iowa, Iowa City, Iowa, United States

[3] Department of Electrical Engineering, University of Iowa, Iowa City, Iowa, United States

[4] Department of Disease Control, London School of Hygiene and Tropical Medicine, London, United Kingdom

[5] Department of Community Nutrition, Great Lakes University of Kisumu, Kisumu, Kenya

* Corresponding author: Kelly K. Baker, University of Iowa College of Public Health, 145 S. Riverside Dr., S316, Iowa City, IA 52246; (001) 319-384-4008; kelly-k-baker@uiowa.edu.

Abstract

Young children are infected by a diverse range of enteric pathogens in high disease burden settings, suggesting pathogen contamination of the environment is equally diverse. This study aimed to characterize across and within neighborhood diversity in enteric pathogen contamination of public domains in urban informal settlements of Kisumu, Kenya, and to assess the relationship between pathogen detection patterns and human and domestic animal sanitation conditions. Microbial contamination of soil and surface water from 166 public sites in three Kisumu neighborhoods was measured by enterococcus assays and qRT-PCR for nineteen enteric pathogens. Regression was used to assess the association between observed sanitary indicators of contamination with enterococcus and pathogen presence and concentration, and pathogen diversity. Seventeen types of pathogens were detected in Kisumu public domains. Enteric pathogens were co-detected in 33% of soil and 65% of surface water samples. Greater pathogen diversity was associated with the presence of domestic animal feces, but not with human open defecation, deteriorating latrines, flies, or disposal of human feces. Sanitary conditions were not associated with enterococcus bacteria, specific pathogen concentrations, or “any pathogen”. Young children played at 40% of observed sites. Managing domestic animal feces may be required to reduce enteric pathogen environmental contamination in high-burden settings.

INTRODUCTION

Recent multi-country epidemiological studies of diarrheal disease etiology in high-burden settings have shown a wide diversity of enteric pathogen infections within and between individual children, and between populations and population sub-groups (i.e. by age) [1, 2]. The diversity in pathogens causing childhood enteric infections shows that children under the age of five years (<5yr) in these settings are chronically exposed to a range of environmentally transmitted enteric pathogens. Relatively little is known about the extent of environmental enteric pathogen contamination in these settings, and the patterns with which enteric pathogen contamination occurs over space and time. A handful of studies have reported high frequencies of detection of different types of enteric pathogens in Indian and Tanzanian household environments, confirming household exposure pathways pose a risk to children [3-5]. However, recent multi-million dollar randomized, controlled trials of household water, sanitation, and hygiene (WASH) interventions found inconsistent impact of household WASH on pediatric diarrhea [6-8]. This suggests that children may be infected by non-household exposure pathways, such as through play in public areas near their household. In low-income settings where sanitation coverage is low, neighborhood public areas are often used for open defecation and disposal of untreated human waste by many households, as well as for domestic animal husbandry [9]. Thus, levels of pathogen contamination in public areas may far exceed contamination levels in privately owned domestic areas, and could pose a disproportionally high risk of infection by enteric pathogens for exposed children.

The overarching goal of the Social Microbes Study is to examine enteric pathogen transmission patterns in high disease burden settings, from sources of fecal contamination through the environment to children, using pathogen distribution and diversity as indicators of transmission pathways. This is the first Social Microbes report to be released, which focuses on testing the hypothesis that public areas within low-income neighborhoods with high disease burden are contaminated by a diverse set of enteric

pathogens, and to test whether observed indicators of human and animal feces contamination would be associated with increased presence and total diversity in pathogens. We also measured how frequently <5yr children are observed in public areas, and whether we saw unsafe exposure behaviors, to assess whether public domain exposures are viable pathways for <5yr enteric infection.

Environmental fecal contamination is usually measured by general fecal indicator bacteria, fecal source tracking markers, or focused detection of specific pathogens of interest as indicators of exposure risk. Fecal *E. coli* and enterococci bacteria are popular, low-cost indicators, but typically correlate poorly with other infectious pathogens in sewage, ecological soils, surfaces, and waters [10-13]. They may be particularly unsuitable indicators for measuring risks from feces in ecological systems where gut bacteria become a naturalized part of soil, water, and surface microbial communities [14]. Host-specific fecal source tracking markers have improved capacity to distinguish human versus animal fecal contamination of the environment, but the reliability of fecal source tracking markers for predicting infectious pathogens is currently unclear [10, 15]. Selection of one or a few specific types of enteric pathogens is often used to understand transmission patterns of specific types of pathogens [16, 17]. However, dozens of enteric viruses, bacteria, protozoans, and helminthic species circulate at different times of the year in endemic settings, so reliance on any one to predict risks from all pathogens transmitted by fecal-oral exposure pathways is risky [1, 2]. In this study, environmental samples were tested for a wide range of common enteric viruses, bacteria, protozoans, and helminths involved in fecal-oral disease transmission in high-burden settings to reduce the likelihood of exposure misclassification (classifying samples as uncontaminated based upon one indicator) or inaccurate concentration estimates for environmental exposure pathways.

Pathogen diversity was an important systems-level evaluation metric in this study that was adopted based upon the theory that diversity in enteric pathogen contamination would occur not just across high disease burden neighborhoods, but also at fine spatial scales. If environmental microbial

contamination by human and animal feces is pervasive in high disease burden settings, with multiple pathogens circulating at any given time, then there is a strong chance that multiple enteric pathogens co-occur in some environments (e.g. open defecation sites) at the same time. Exposure of children to these uniquely hazardous environments could increase their risk of infection by multiple pathogen types. The risk of exposure to any given pathogen from soil or surface water is influenced by the recent presence of a fecal source, but also by the pathogen's capacity to persist in specific environments, like soil or surface water, and overall infection rate in a host population. The evolutionary traits that govern persistence of microbes may be unique to members of a phylogenetic group (e.g. non-obligate bacterial replication in the environment), as well as reflect species-specific ecological adaptations. Examining both the pathogen-specific and higher order taxonomic differences in pathogens in the environment could provide insight as to the types of pathogens potentially transmitted via different child behaviors, e.g. ingesting soil vs water. From a programming standpoint, capacity to identify high-risk areas or exposure pathways, in a milieu of existing elevated background contamination, could improve how well investments reduce <5yr enteric disease in high-burden settings. Pathogen diversity could improve the identification of high- versus low-risk areas or exposure pathways, and the fecal sources contributing to environmental contamination.

This manuscript describes enteric pathogen detection frequencies and diversity patterns in soil and surface water from public areas of three low-income, urban neighborhoods of Kisumu, Kenya with low sanitation coverage. Then, we examine the relationship between human and domestic animal fecal sources and pathogen detection and diversity patterns.

MATERIALS AND METHODS

Study Population. Kisumu is the third largest city in Kenya, and has a population of approximately 409,928 inhabitants [18]. Up to 60% of the city's population resides in peri-urban informal settlements , which have emerged due to economic migration and a lack of affordable housing [19, 20]. Kisumu County has a high prevalence of diarrhea (18% two-week period prevalence) with most cases of diarrhea occurring in children less than three years of age [21]. The child mortality rate is 105 deaths per 1,000 live births and the prevalence of severe childhood stunting (> -2 standard deviations below the reference norm) is approximately 25% [21]. This study took place in three established informal settlements, Nyalenda A (population density (pd) = 8,953 persons per square kilometer (km^2)) Nyalenda B (pd = 6,886/km^2), and Obunga (pd=1,913 km^2). Co-habitation with domestic animals is common, and most residents rely upon open defecation or sanitation facilities shared by eight or more households [22, 23].

Pilot of Environmental Sampling and Microbiology Methods. Prior to conducting the primary study in Kenya, a pilot sampling project was conducted in Iowa in June 2015 (similar climate conditions) to understand how often fecal indicators and enteric pathogens would be detected in a watershed impacted by agriculture, free-range animal management (cows), and concentrated animal feeding operations (pigs). This information is presented in Supplemental documents as a comparative baseline for understanding the importance of fecal indicator and pathogen contamination in Kisumu, Kenya.

Sample site selection. A cross-sectional observational survey with environmental sampling was performed in Kenya in July 2015. The study was designed to ensure sampling sites were randomly distributed across special areas of interest to optimize measurement of variability in pathogen diversity and co-occurrence patterns and to prevent introduction of observer bias in selection of environmental sampling locations. Neighborhood boundary parameters were visualized utilizing Batchgeo (Google), and these spaces were defined geographically by the northernmost, westernmost, easternmost, and southernmost latitude or longitude values in rectangular form. For each neighborhood, sixty latitude and longitude pairs that fell within these specified ranges were randomly generated utilizing the

website, geomidpoint.com, prior to field based data collection (Figure S1). GPS coordinates were entered into a Waytracker mobile phone app, and daily routes were identified to navigate between ~ 15 coordinates per day. Observers navigated to the coordinates (+/- 3 meters), which was considered the center of a "site", defined as all area falling within a 25 meter radius around the central coordinates. If the coordinates fell within a private household yard or business, the nearest set of coordinates outside that private space but within 25 meters of the random coordinates were identified. Sites were visited at various times during the day (morning and afternoon), with site visits lasting approximately 20 minutes.

Public site observation. Study teams systematically documented sanitary conditions at each site using an observational survey implemented via mobile phone app (Fieldlogs). Observers documented landscape features such as surface waters, grasslands and altitude. Development of infrastructure was noted, including roads, drains, dams, industry, housing, public water sources, and presence, physical condition, and hygiene of public or communal latrines. Indicators of human open defecation or unsafe disposal of excreta included observed "flying toilets" (plastic bags containing excreta), used diapers, piles of human feces, emptying of latrines into drains or land around the latrine, septage emptied from a latrine next to the latrine, and visual confirmation of an adult or child actively defecating in the open. Presence and type of domestic animals and their feces were recorded. Observers recorded whether during that ~ 10-15 minute spot observation, children approximately < 5 yrs were observed in the public area, the number present, and any behaviors that would result in hand or mouth contact with environmental fomites (touching soil, surface water, animals, or objects on the ground, swimming, eating food, eating dirt, mouthing hands). Additionally, enumerators documented whether any children were defecating in the open at the time of observation. Ecological conditions that could influence pathogen presence and persistence, specifically daily temperature and relative humidity were extracted from NOAA data collected at the Kisumu Airport. Observers did not include area within private housing or businesses adjoining or overlapping with the site area to avoid potential bias in conditions observed

from privately held property. However, private conditions that impacted public space, such as drains or sewage leaching from the household or animals roaming between public and private areas, were recorded.

Sample collection. Standard Operating Protocols were implemented to ensure standardized hygienic sampling and processing of environmental materials. Approximately five grams of soil was collected at every site by inserting an alcohol-sterilized scoop into the ground at a 45° angle to a depth of 5cm (half the length of the scoop) and transferring the soil into a sterile WhirlPak® bag (Sigma-Aldrich Corp., St. Louis, MO, USA). Surface water was collected, if present, by skimming water into a WhirlPak bag. Collection bags were stored on ice packs in a cooler and transported to the laboratory within six hours of collection. In Kenya, four samples were collected at seven randomized sites in each neighborhood to account for anticipated variance in pathogen distributions at public sites caused by the final spatial scale of sites per neighborhood.

Indicator analysis. One gram of soil was measured into 10 mls of Phosphate Buffered Saline (PBS), vortexed for 30 seconds, and then mixed on a rotator for 20 minutes. Solid matter was allowed to settle for 5 minutes and elute was removed for Enterococci assays. Enterococci are recommended by the US Environmental Protection Agency for identifying fecal material in fresh and marine recreational waters [24], although both enterococci and *E. coli* are considered inadequate indicators in tropical settings like Hawaii [15, 25]. There is no global recommendation on appropriate indicators for tropical waters, so enterococci were chosen based upon the EPA policy. Enterococcus were enumerated by vacuum filtration of three serial dilutions of surface water (10 millilter (mL), 1 mL, 0.1 mL) or soil (1 mL, 0.1 mL, 0.01 mL) rinse through a white gridded 0.45 µm mixed cellulose esters filter (Product No. GSWP04700, Millipore, Billerica, MA), and culturing filters for 18-24 hours at 37°C on mEI agar (EPA method 1600). Colony forming units (cfu) of enterococci were counted according to manufacturer's recommendations.

DNA and RNA extraction. DNA and RNA was extracted from 0.5 grams of soil using the FastDNA and FastRNA SPIN kits for Soil (MP Biomedicals, Solon, OH), including a bead beating step. A 10 mL volume of surface water was processed by adjusting water to 2.5 mM MgCl2 and pH 10.0, and vacuum filtering through a 0.45 µm mixed cellulose esters filter. This volume was chosen to ensure we could systematically sample a wide range of surface waters (flood water puddles, drains, rivers) representing potential exposure hazards in Kisumu. This method was chosen because it more efficiently removes inhibitors than ultrafiltration and is practical for laboratories with limited capacity and for isolation of many different types of pathogens [26, 27]. Filters were frozen at -20°C for 1-4 weeks and transported to the University of Iowa where they were stored at -80°C until extraction (1 week). Filters were cut in half and processed using the FastDNA and FastRNA for Soil kits (equivalent to 5 mls water).

Quantitative PCR on Environmental Samples. The concentration and quality of DNA and RNA was measured using a Nanodrop UV-Vis spectrophotometer (Thermo Scientific, DE, USA). Duplicate 6 microliter (µl) and 0.6 µl volume of nucleic acid extract were tested for inhibition using the QuantiFast Pathogen Internal Control kit (Qiagen, Germantown, MD). These methods effectively removed most amplification inhibitors (Table S1). Samples that were not inhibited were analyzed by combining 20 µl of DNA and RNA each (total 40 µl) with AgPath polymerase and then performing quantitative reverse transcription Polymerase Chain Reaction (qRT-PCR) using a nineteen pathogen microfluidic TaqMan Array Card (TAC) on a QuantiStudio 12K Flex Real-Time PCR System with Array Card block (ThermoFisher, Chicago, IL) [28]. Inhibited samples were diluted 1:10 with diH20 before TAC analysis. The card format included five types of viruses (adenovirus 40-41, astrovirus, sapovirus, norovirus GII, rotavirus), fourteen types of bacteria (enteroaggregative *E. coli* (EAEC), enteropathogenic *E. coli* (EPEC), enteropathogenic *E. coli* (ETEC), shiga toxin expressing *E. coli* (STEC), *C. jejuni*, Shigella, *V. Cholerae*, *S. enterica*, *C. difficile*), three protozoans (*Cryptosporidium spp.*, *G. lamblia*, *E. histolytica*), and two helminths (Ascaris, Trichuris). Exponential curves and multicomponent plots were visually examined to

validate positive amplification. Gene targets with real amplification in one well were re-analyzed and considered positive if amplification was detected again. Gene targets with cycle threshold (CT) values over 35 were cross-validated by performing a 14 cycle pre-amplification reaction with pathogen-specific primers (0.2 μM each) to increase the starting concentration of pathogen DNA/cDNA in a sample [29], and reanalyzed by qRT-PCR. Samples were classified as positive for a specific gene target if amplification was verified at a lower cycle threshold, and negative if unverified. This resulted in presence and concentration data for 24 pathogen gene targets representing 19 types of pathogen taxa per sample (detection of either gene target for ETEC, EPEC, EAEC, and STEC was positive at a taxa level).

The concentration of each pathogen per reaction volume was estimated by comparing the CT for a pathogen gene target in a sample to a standard curve generated by qRT-PCR of a six to seven-fold serial dilution of a positive control of known concentration for each of the 19 qRT-PCR targets (Table S2). For samples that were repeated using the pre-amplification step, the initial concentration was estimated by comparison to a standard processed by pre-amplification. If samples determined to be positive still reflected signs of significant inhibition, concentrations were excluded from analysis to avoid biasing statistical estimates of mean and standard deviation. Otherwise, final concentrations of gene copy per gram of soil or per mL of surface water were generated by multiplying the concentration of each gene target by the dilution factors introduced by processing. Concentrations for samples with no detectable amplification were not transformed to avoid inflating statistics caused by relatively high methodological lower limit of detection (LLOD) for some pathogens.

Statistical analysis. Data were analyzed using SAS Version 9.4 (SAS Institute, Cary NC). Descriptive statistics of variables were reported as proportions or mean and standard deviation.

Microbial contamination outcomes of interest included the (1) presence of enterococcus fecal indicator bacteria, (2) $\log_{10}$-number of enterococcus cfu, (3) presence of any pathogen gene targets, (4) "Pathogen Diversity" defined by the sum count of all unique types of enteric pathogens detected in an

individual soil or water sample (either target for pathogenic bacteria defined as one group), and (5) $\log_{10}$ gene copy of pathogen gene targets.

Exposure variables representing potential sources of fecal contamination were: any human or animal feces observed versus none observed; indicators of open defecation vs. none observed; the presence of any type of domestic animal present versus none observed; flies present versus none observed; and a categorical variable for latrines present in good condition, latrines present in bad condition, and no latrines present. Frequently latrines in both good and bad condition were present at a site. In this case the site was classified as "latrines in bad condition" based upon the default assumption that just one deteriorating latrine could introduce contamination into soil and water.

Potential confounders included: type of sample (soil versus water), altitude in meters, relative humidity, temperature in Celsius, and landscape use (housing area versus undeveloped).

Generalized linear mixed models (glmm) with binary log link, robust standard errors, and random intercept for site ID and neighborhood with exchangeable correlation structure to account for spatial clustering at each level were used to test whether sites with any of the exposure variables were more likely to have at least one type of enteric pathogen detected. Identical glmm models with Poisson log link were used to test whether sites with any of the same potential fecal sources were more likely to have increased sum counts in Pathogen Diversity. Censored regression models were used to assess whether exposure variables were associated with increased concentration of enteric pathogen gene copy number to account for left-censoring of data. Each modeling process involved analyzing the effect of each exposure variable independently on the pathogen outcome, while adjusting for landscape type, daily temperature, relative humidity, altitude, and type of sample (soil vs. water). Full models included all variables simultaneously to measure individual effects, accounting for co-occurrence of exposure variables at a site. Based upon full model results, a sub-analysis of association between specific types of

domestic animals and Pathogen Diversity was performed by replacing the binary Domestic Animals variable with presence/absence variables for specific categories of animals.

A one-sided Mantel test using the geographic distances between sites was used to test for spatial clustering of sites with compositional similarity between pairs of sites in terms of the number of pathogens detected at both sites.

RESULTS

Characteristics of study sites. A total of 166 public sites were inspected for this study. Ecological conditions and the proportion of sites developed with housing were relatively similar in the three surveyed neighborhoods, with the exception of higher mean elevation in Obunga, more sites with indicators of open defecation or open feces disposal in Nyalenda B, fewer sites with shared latrines in Obunga, and more latrines in good structural condition in Nyalenda B (Table 1).

Children less than five years of age were observed at 40% (66 of 166) sites (median 2 children per site, range 1-10), including infants (defined as unable to stand and walk) at 9% (15 of 166) of sites. At least one <5 child was observed crawling on the ground or sitting on the ground playing in dirt or mud at 24 sites, with a third of these observations (n=8) recording infants crawling or sitting in the dirt. Children were observed playing with objects in the dirt at 4 sites. One infant and one child ~12-24 months were observed playing in a water puddle or drain water at 2 sites. Children were seen mouthing their hands at 3 sites (multiple ages), and one infant was eating soil at 1 site. Open defecation was not observed.

Microbial Detection, Concentration, and Neighborhood-Level Diversity in Kenya. A total of 185 public site soil and 51 water samples were collected and tested for enterococcus and enteric pathogens. Enterococcus colonies were isolated from 100% of surface water samples and 74.6% of soil samples (Figure 1). At least one type of enteric pathogen was detected in 92% of water samples and 82% of soil

samples. This included twenty gene markers representing sixteen types of enteric pathogens in water and twenty gene markers representing seventeen types of pathogens in soil. The most common pathogens in soil were Cryptosporidium (67.0%), Giardia (17.8%), ETEC (13.5%), EAEC (10.8%), and EPEC (8.6%). The most common pathogens in water were Cryptosporidium (70.0%), EAEC (58.0%), ETEC (56.0%), EPEC (46%), and human adenovirus (36.0%). The relative abundance of different biological phyla and types of pathogens differed between pooled soil and water samples. Protozoan organisms accounted for 61.2% of microbial organisms in soil, followed by 30.6% pathogenic bacteria, 5.8% helminths, and 2.3% viruses (Figure 2). In contrast, pathogenic bacteria were more abundant in water (57.5%), followed by 26.9% protozoans, 5.8% viruses, and 2.3% helminths.

Enterococcus concentrations were about three logs higher in concentration in water than in soil (Table S3). Mean pathogen concentrations in positive samples were two to five logs higher than the limit of detection, although wide standard deviations for all pathogens reflected variability in concentration ranges (Table S3). Concentrations of pathogens in one gram of soil were consistently one to two logs higher in concentration than in 1 mL of surface water. In soil, the adenovirus 40/41 hexon, ETEC *eltB*, Salmonella *invA*, and STEC *stx2* genes were all detected at > 7.0 $\log_{10}$ concentration per gram, whereas *C. difficile tcdB* was the lowest concentration at 4.2 $\log_{10}$. In surface water, adenovirus 40/41 hexon was the highest at 6.6 $\log_{10}$ per mL, followed by EAEC *aaiC*, ETEC *eltB*, and STEC *stx1* and *stx2* genes at 5.6 $\log_{10}$ per mL. Astrovirus capsid, Norovirus GI/GII ORF1-2, and Campylobacter *cadF* at < 3 $\log_{10}$ per mL.

Within-site Enteric Pathogen Diversity. Two or more enteric pathogens were co-detected in 35% (64/185) of soil samples (median 1; range 0 - 7), and 69% (35/51) of surface water samples (median 4; range 0 - 10) (Figure 3). Patterns of pathogen co-detection were heterogeneous, with only 2 patterns (*Cryptosporidium spp*., EPEC, ETEC, EAEC) and (*Cryptosporidium spp*., *Giardia spp*., ETEC, EAEC) reoccurring twice at a total of two sites each among the 35 water samples with >=2 pathogens detected. Of the 64 soil samples with >=2 pathogens co-detected, *Cryptosporidium spp*., ETEC, and EAEC occurred

three times, *Cryptosporidium spp*., *Giardia spp*., and ETEC twice, and the remainder were various combinations of *Cryptosporidium spp*. with other types of pathogens. The *Cryptosporidium spp*. assays might have detected species that do not typically infect humans[30]. If *Cryptosporidium spp*. results are excluded from the analysis, then 44% (82/185) of soils and 80% (41/51) contained at least one pathogen, and two or more enteric pathogens were co-detected in 18% (33/185) of soil samples (median 0; range 0 - 6), and 65% (33/51) of surface water samples (median 3; range 0 - 9).

Association between Sanitary Conditions and Enterococcus contamination. Observation of feces, indicators of human open defecation, domestic animals, presence and poor condition of latrines, and flies were not associated with the presence or increased concentration of enterococcus indicator bacterial of fecal contamination of water and soil, after adjusting for development of the terrain, relative humidity, altitude, ambient temperature, and type of sample (Table 2).

Association between Sanitation Conditions and Pathogen Diversity. Observation of feces, indicators of human open defecation, domestic animals, presence and poor condition of latrines, and flies were not associated with the detection of any type of enteric pathogen, after adjusting for development of the terrain, relative humidity, altitude, ambient temperature, and type of sample (Table 3). However, the presence of domestic animals was significantly associated with increased diversity in enteric pathogens in fully adjusted models. Observation of feces, indicators of human open defecation, presence and poor condition of latrines, and flies were not associated with increased diversity. Sub-analysis of animal types in the fully adjusted model found that chickens (adjOR=1.64; 1.22, 2.22), cattle (adjOR=1.36; 1.08, 1.72), and goats and sheep (adjOR=1.39; 1.00, 1.94) were associated with increased pathogen diversity, while pigs (adjOR=0.80; 0.36, 1.77) and dogs (adjOR=1.20; 0.87, 1.65) were not significantly associated. No association was observed between different fecal sources and increased concentration of individual types of pathogens (not shown).

Spatial Clustering of sites with High Pathogen Diversity. High pathogen diversity was not statistically associated with lower distance between pairs of sites (p-value = 0.22).

DISCUSSION

This study confirmed that neighborhood landscapes in Kisumu, Kenya are contaminated by many enteric virus, bacteria, protozoan, and helminth enteric pathogen species, at both neighborhood and localized (within 25 meter radius) levels of spatial scale. At least one type of pathogen was detected at 80% of public sites, with one third of soil and three-quarters of surfaces water co-contaminated by multiple taxa of enteric pathogens. Even if Cryptosporidium detections are excluded, in the event that the species detected were not human-infective, multiple pathogens were still detected in one-fifth of soils and two-thirds of waters. This evidence addresses major knowledge gaps about enteric pathogen co-occurrence patterns in the environment, especially for public areas that are “ground zero” for contamination from open defecation and animal feces. The most abundant types of pathogen taxa varied for soil and surface water, with protozoans being most abundant in soil and pathogenic bacteria being most abundant in surface water. Domestic animals (specifically chickens, cows, and goats/sheep) – rather than human sources of fecal waste – were associated with increased pathogen diversity, whereas enterococcus bacteria, the presence of any pathogen type, and concentrations of individual enteric pathogen taxa were not significantly associated with human and animal sources of fecal contamination. Finally, this study validated that <5yr old children play in these contaminated public settings, confirming that child exposure to and infection by pathogens in public areas is plausible. Therefore, interventions that prevent neighborhood-level animal fecal contamination may be necessary for reducing enteric disease burden in < 5yr children in Kisumu. These interventions could include

creation of protected child play spaces with finished, cleanable floors, promoting the safe disposal of animal waste, or animal penning designs that safely collect animal waste away from children.

These findings confirmed our hypothesis that pathogen diversity in Kisumu neighborhoods would mirror the leading etiologic causes of symptomatic and asymptomatic childhood infections in low-income, high-burden endemic countries in Asia and Africa [1, 2, 28]. Many of the most common causes of moderate to severe diarrhea in children in Kenya *(Cryptosporidium spp.*, ETEC ST, adenovirus 40/41, *C. jejuni*)[1, 31] were the most frequently detected in the environment in Kisumu, although Rotavirus and Shigella/EIEC were rare. This suggests that widespread contamination by those pathogens in the environment is responsible for causing enteric infection in children. Our findings also confirmed our hypothesis that multi-pathogen contamination of soil and water would be detected at locations where children play. This further suggests that children could ingest multiple pathogens during play in just one area outside the household. Pathogen diversity is a novel approach for characterizing environmental exposure risks in LICs, and for identifying potential fecal sources associated with environmental contamination. However, microbial community characterization has been used for decades to understand ecological health and to monitor the impact of natural or man-made actions on system function [32-34]. In public health, microbiome community characterization has become a platform for understanding human susceptibility to enteric infection, symptomology and severity of disease, and for understanding vaccine failure [35, 36]. A few studies have used the microbial community approach to understand the relationships between environmental microbiota, fecal indicator bacteria and pathogenic bacteria in surface water and soil, although did not examine infrastructure conditions driving this contamination [37, 38]. Rather than sequencing, we used customized multi-pathogen qPCR tools for disease-targeted detection of the viral, bacterial, protozoan, and helminthic causes of pediatric diarrhea [28]. Additionally, we accounted for potential interrelated conditions underlying pathogen co-occurrence in our analysis, rather than treating pathogen occurrence as an isolated, independent event.

The inclusion of pathogen diversity as indicator of interrelated contamination conditions was important for understanding the relative intensity and determinants of contamination of public areas across Kisumu, as well as between Kenya and a reference site in Iowa. Consistent with prior studies, enterococcus were omnipresent in Kenya *and* Iowa soil and surface water, even though evidence of human fecal contamination was absent in Iowa, and animal feces was less common [15]. Evidence on fecal microbes in soil is limited, but our detection rates were similar to at least one other study [39]. The lack of association between human and animal fecal source risk factors with enterococci in Kenya reinforces that enterococcus indicators are not optimal tools for predicting or quantifying exposure to enteric pathogens or for fecal source tracking. An "any pathogen" indicator was also not statistically associated with fecal sources. However, the stark differences in overall pathogen detection frequencies between Kenya and an Iowa watershed impacted by farm and agriculture helped calibrate our expectations as to what contamination patterns should look like in a setting where open defecation is absent, and highlighted how alarming the pathogen detection levels in Kenya are by comparison. Sampling at different times, for example different seasons, or sampling from agriculturally-impacted areas in a tropical state (e.g. Florida), may have produced different results.

Neighborhood-level pathogen contamination in Kenya corresponded with the frequent observation of human and animal feces in public areas. These neighborhoods have low levels of household sanitation coverage and domestic animal ownership is common in Kisumu [22, 23]. We expected open defecation, dilapidated latrines, and domestic animals to be associated with pathogen diversity, given the evidence of their role in diarrheal in children [40, 41]. In spite of widespread human feces in Kisumu, only domestic animals were associated with increased diversity in pathogens in the environment. Domestic animal reservoirs may contribute to pediatric disease burden by facilitating a rapid turnover in the types of pathogen species that children are exposed to, a situation where exposure-based immunity provides little benefit. We are unaware of evidence of this hypothesis, but

386 theoretically domestic animals may be more important disease vectors than humans. First animals
387 (dogs, chickens, pigs, ducks) are more likely to be coprophagous (eat feces than humans, and ergo have
388 higher risks of infection [42]. Second, animals are less likely to be treated for diarrheal symptoms of enteric
389 infection than children [43], and ergo shed pathogens for longer due to persistent infection.Third, dozens
390 of animals may be kept by a household versus one or a few children [22], ergo density of potential animal
391 disease vectors is higher than child vectors. The lack of association between human sanitation and
392 pathogen presence or diversity might reflect either lower levels of pathogen infection rates in humans,
393 relative to animals. We think that the lack of association with all sanitary conditions and pathogen
394 concentration is due to ecological conditions (humidity, UV, properties of soil) playing a strong
395 mediating role on pathogen persistence and fate in the environment.

396 This study has several limitations. It is cross-sectional and cannot establish the direction of
397 causality. Pathogen contamination could be a proxy for sites with more abundant animal food sources
398 (feces, grassland), which thus attract domestic animals or flies to those sites. Even if that is the case, this
399 is likely a circular relationship that involves animals defecating at the feeding site. Our observational
400 criteria may not have distinguished well between unsafe, that is deteriorated and/or sometimes unused,
401 latrines and safe latrines, that is latrines which effectively contain excreta and prevent release into the
402 environment. Without an in-depth inspection of the underground integrity of the latrine pits, the degree
403 to which excreta is effectively contained cannot be assessed. Our sample size may have not provided
404 sufficient power to detect important associations between fecal sources and microbial outcomes. There
405 was no prior source of information to predict frequencies of enteric pathogen co-occurrence, especially
406 for nineteen enteric pathogens.

407 Inherent heterogeneity in pathogen distributions in the environment and insufficient spacing
408 scale between sites (sites too far apart) may explain why we did not find many repeat patterns of
409 pathogen co-occurrence or spatial “hot zones”. Such patterns would provide even more precise metrics

for linking widespread contamination to specific fecal sources, compared to the relatively simple approach of simply summing the number of pathogen types detected. In the handful of repeat patterns noted, the pathogens detected were the most common overall, suggesting presence is more a function of probability and sample size than impacts from types of fecal sources. In particular, the most common pathogen gene detected was *Cryptosporidium spp*. 18S, which could indicate contamination by species rarely detected in humans (*C. meleagridis, C. canis, C. felis, C. muris, C. suis*), rather than the *C. hominus* or *C. parvum* species commonly linked with human infection[30]. This may overestimate the abundance of clinically-relevant pathogens in public areas of Kisumu, although even after excluding Cryptosporidium-positive counts, pathogen detection and co-detection was still very high. Research involving larger numbers of environmental samples is needed to understand whether co-occurrence patterns of specific groups of pathogens is common. Either way, the substantial amount of heterogeneity in pathogen detection patterns highlights the risk for misclassification of environmental exposures if the presence of an individual pathogen type is used as a proxy for the presence of any fecal pathogens, or fecal contamination overall. While some associations with trends towards significance may surpass significance cutoff thresholds with larger sample sizes, we are confident that it wouldn't change the overall conclusion that domestic animals are important contributors to multi-pathogen contamination in public domains of Kisumu.

Last, detection of pathogen DNA or RNA in soil and water by PCR does not confirm viability or infectivity of pathogens. PCR methods may detect extracellular DNA or intact, but non-infectious cells. By filtering sample elute through a 0.45 µm pore size filter, our sampling methods may have removed some extracellular material, although non-viable cellular DNA or RNA could still be trapped on the filter. PCR was chosen as the most systematic way to detect viral, bacterial, protozoan and helminth classes of fecal pathogens, many of which have no alternative detection methods. Culture-based assays could have been used for some bacteria, but may have been equally inaccurate due to the presence of viable

but non-recoverable bacteria (e.g. *Vibrio cholerae, Shigella dysenteriae*, *Camylobacter jejuni*, and Enterotoxigenic *E. coli*) [17, 44-46]. Since we did not adopt a pre-enrichment step, it is possible that our enterococcus concentration data underestimates actual live concentration of enterococcus in soil or water. This also would introduce inter-method variability in detection methods across viral, bacterial, protozoan, and helminth taxa. Adjusting concentrations for pathogen decay under these ecological conditions may improve the accuracy of estimated PCR concentrations. However, there is no such information for many enteric pathogens included on our assay, so such an approach could not be systematic. Also problematic is that feces contamination by animals and humans in Kisumu is ongoing. Failing to counter-adjust pathogen concentration estimates for periodic reintroduction of pathogens in the environment would underestimate final concentration estimates. While some detected pathogen DNA detected in Kisumu may be non-viable, it is more likely to be relatively recent contamination because free and cellular DNA/RNA is typically degraded by native soil or water microbiota within the matter of days or weeks [47-52].

In conclusion, this study addresses two neglected realms in the WASH sector. First, we demonstrate that public domains are highly contaminated and may pose a substantial risk of enteric pathogen exposure for children. Young children will continue to play in soil and surface water in the public domain because they live in crowded conditions and at a certain age are drawn to engage in social play with their peers. In light of this, WASH interventions should invest in improving public sanitary conditions to prevent child exposure to enteric pathogens, and possibly to reduce the childhood enteric disease burden that persists in many low and middle income countries. Second, even if high levels of household WASH coverage are achieved, and human open defecation is eliminated, a high baseline level of enteric disease caused by contact with animal feces will persist. Domestic animal management must be included in the WASH agenda to reduce pathogen contamination at the neighborhood level, and potentially within households as well. Forthcoming research by this group is

examining the generalizability of these findings in other geographical settings using improved assessment methods. Yet, the conditions in Kisumu – low-income urban settlements with high population density and limited public health infrastructure – are common in low-income settlements around the world, suggesting pervasive pathogen contamination of the environment in such settings is a universal problem.

Supplementary Documents Table of Contents

A. Figure S1. Location of public domain sampling sites in three neighborhoods of Kisumu, Kenya.
B. Assessment of RT-PCR inhibition
C. Table S1. Number of soil and surface water samples from Kenya with evidence of inhibition of the QuantiFast Internal Control.
D. RT-PCR Standard curves
E. Table S2. Source of reference DNA or RNA use for standard curves in this study.
F. Back calculations of final concentrations.
G. Microbial Detection and Diversity in Iowa watershed.
H. Figure S2. Detection of *E. coli* bacteria, bacterial 16S DNA, and enteric virus, bacteria, protozoan, or helminth pathogens in soil (A) and water (B) from an agricultural Iowa watershed.
I. Table S3. Concentration of enterococcus indicator bacteria and enteric pathogen DNA and RNA gene copy in soil and surface water from Kisumu neighborhood public domains.

AKNOWLEDGMENTS:

This study was supported by start-up funding from the University of Iowa to PI Baker, a career development award from the University of Iowa Environmental Health Sciences Research Center (NIH

P30 ES005605), and a pilot grant from the University of Iowa Water Sustainability Initiative. We appreciate the support of Lucy DesJardin and Nancy Hall at the Iowa State Hygiene laboratory who allowed our group access to a ViiA7 thermocyler, and Julius Ochieno, Lily Lukorito, Horace Phoxydee, Mandy Larson, Lena Swander, Gocale Nicoue, and Kevin Tsai assisted with the collection of data, development of assays, or data management. Craig Ellermeier at the University of Iowa generously donated several strains of *Clostridium difficile* for development of standard controls.

All authors declare that they have no competing financial interests in the publication of this study.

CONTRIBUTIONS

Conceived of this study: KKB, ASG. Acquired funding: KKB, ASG. Designed study or study tools: KKB, RS, OC, JM. Collected data: KKB, RS, JM. Statistical analysis: DS, RS, KKB. Drafted manuscript: KKB. Edited manuscript for intellectual content: all authors. Provided final approval of the version to be published: all co-authors.

REFERENCES

1. Kotloff, K. L.; Nataro, J. P.; Blackwelder, W. C.; Nasrin, D.; Farag, T. H.; Panchalingam, S.; Wu, Y.; Sow, S. O.; Sur, D.; Breiman, R. F.; Faruque, A. S.; Zaidi, A. K.; Saha, D.; Alonso, P. L.; Tamboura, B.; Sanogo, D.; Onwuchekwa, U.; Manna, B.; Ramamurthy, T.; Kanungo, S.; Ochieng, J. B.; Omore, R.; Oundo, J. O.; Hossain, A.; Das, S. K.; Ahmed, S.; Qureshi, S.; Quadri, F.; Adegbola, R. A.; Antonio, M.; Hossain, M. J.; Akinsola, A.; Mandomando, I.; Nhampossa, T.; Acacio, S.; Biswas, K.; O'Reilly, C. E.; Mintz, E. D.; Berkeley, L. Y.; Muhsen, K.; Sommerfelt, H.; Robins-Browne, R. M.; Levine, M. M., 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), 209-22.
2. Platts-Mills, J. A.; Babji, S.; Bodhidatta, L.; Gratz, J.; Haque, R.; Havt, A.; McCormick, B. J.; McGrath, M.; Olortegui, M. P.; Samie, A.; Shakoor, S.; Mondal, D.; Lima, I. F.; Hariraju, D.; Rayamajhi, B. B.; Qureshi, S.; Kabir, F.; Yori, P. P.; Mufamadi, B.; Amour, C.; Carreon, J. D.; Richard, S. A.; Lang, D.;

Bessong, P.; Mduma, E.; Ahmed, T.; Lima, A. A.; Mason, C. J.; Zaidi, A. K.; Bhutta, Z. A.; Kosek, M.; Guerrant, R. L.; Gottlieb, M.; Miller, M.; Kang, G.; Houpt, E. R.; Investigators, M.-E. N., Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Health* **2015,** *3*, (9), e564-75.

3. Pickering, A. J.; Julian, T. R.; Marks, S. J.; Mattioli, M. C.; Boehm, A. B.; Schwab, K. J.; 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), 5736-43.

4. Mattioli, M. C.; Pickering, A. J.; Gilsdorf, R. J.; Davis, J.; Boehm, A. B., Hands and water as vectors of diarrheal pathogens in Bagamoyo, Tanzania. *Environmental science & technology* **2013,** *47*, (1), 355-63.

5. Odagiri, M.; Schriewer, A.; Daniels, M. E.; Wuertz, S.; Smith, W. A.; Clasen, T.; Schmidt, W. P.; Jin, Y.; Torondel, B.; Misra, P. R.; Panigrahi, P.; Jenkins, M. W., Human fecal and pathogen exposure pathways in rural Indian villages and the effect of increased latrine coverage. *Water Res* **2016,** *100*, 232-44.

6. Humphrey, J. H.; Prendergast, A. J.; Ntozini, R.; Gladstone, M.; Colford, J. In *The Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial*, American Society for Tropical Medicine and Hygiene Annual Meeting, Baltimore, MD, November 7, 2017, 2017; American Society for Tropical Medicine and Hygiene: Baltimore, MD, 2017.

7. Null, C.; Stewart, C. P.; Pickering, A. J.; Dentz, H. N.; Arnold, B. F.; Arnold, C. D.; Benjamin-Chung, J.; Clasen, T.; Dewey, K. G.; Fernald, L. C. H.; Hubbard, A. E.; Kariger, P.; Lin, A.; Luby, S. P.; Mertens, A.; Njenga, S. M.; Nyambane, G.; Ram, P. K.; Colford, J. M., Jr., Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *Lancet Glob Health* **2018,** *6*, (3), e316-e329.

8. Luby, S. P.; Rahman, M.; Arnold, B. F.; Unicomb, L.; Ashraf, S.; Winch, P. J.; Stewart, C. P.; Begum, F.; Hussain, F.; Benjamin-Chung, J.; Leontsini, E.; Naser, A. M.; Parvez, S. M.; Hubbard, A. E.; Lin, A.; Nizame, F. A.; Jannat, K.; Ercumen, A.; Ram, P. K.; Das, K. K.; Abedin, J.; Clasen, T. F.; Dewey, K. G.; Fernald, L. C.; Null, C.; Ahmed, T.; Colford, J. M., Jr., Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *Lancet Glob Health* **2018,** *6*, (3), e302-e315.

9. Medgyesi, D. N.; Brogan, J. M.; Sewell, D. K.; Creve-Coeur, J. P.; Kwong, L. H.; Baker, K. K., Where Children Play: Young child exposure to environmental hazards during play in public areas in a transitioning internally displaced persons community in Haiti. *International Journal of Environmental Research and Public Health* **2018,** *15*, (8), 1682.

10. Harwood, V. J.; Levine, A. D.; Scott, T. M.; Chivukula, V.; Lukasik, J.; Farrah, S. R.; Rose, J. B., Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Appl Environ Microbiol* **2005,** *71*, (6), 3163-70.

11. Ferguson, A. S.; Layton, A. C.; Mailloux, B. J.; Culligan, P. J.; Williams, D. E.; Smartt, A. E.; Sayler, G. S.; Feighery, J.; McKay, L. D.; Knappett, P. S.; Alexandrova, E.; Arbit, T.; Emch, M.; Escamilla, V.; Ahmed, K. M.; Alam, M. J.; Streatfield, P. K.; Yunus, M.; van Geen, A., Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater. *Sci Total Environ* **2012,** *431*, 314-22.

12. Haramoto, E.; Katayama, H.; Oguma, K.; Yamashita, H.; Tajima, A.; Nakajima, H.; Ohgaki, S., Seasonal profiles of human noroviruses and indicator bacteria in a wastewater treatment plant in Tokyo, Japan. *Water Sci Technol* **2006,** *54*, (11-12), 301-8.

13. Verbyla, M. E.; Iriarte, M. M.; Mercado Guzman, A.; Coronado, O.; Almanza, M.; Mihelcic, J. R., Pathogens and fecal indicators in waste stabilization pond systems with direct reuse for irrigation: Fate and transport in water, soil and crops. *Sci Total Environ* **2016,** *551-552*, 429-37.

14. Desmarais, T. R.; Solo-Gabriele, H. M.; Palmer, C. J., Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. *Appl Environ Microbiol* **2002,** *68*, (3), 1165-72.

15. Byappanahalli, M. N.; Nevers, M. B.; Korajkic, A.; Staley, Z. R.; Harwood, V. J., Enterococci in the environment. *Microbiol Mol Biol Rev* **2012,** *76*, (4), 685-706.
16. Islam, M. S.; Talukder, K. A.; Khan, N. H.; Mahmud, Z. H.; Rahman, M. Z.; Nair, G. B.; Siddique, A. K.; Yunus, M.; Sack, D. A.; Sack, R. B.; Huq, A.; Colwell, R. R., Variation of toxigenic Vibrio cholerae O1 in the aquatic environment of Bangladesh and its correlation with the clinical strains. *Microbiol Immunol* **2004,** *48*, (10), 773-7.
17. Alam, M.; Sultana, M.; Nair, G. B.; Siddique, A. K.; Hasan, N. A.; Sack, R. B.; Sack, D. A.; Ahmed, K. U.; Sadique, A.; Watanabe, H.; Grim, C. J.; Huq, A.; Colwell, R. R., Viable but nonculturable Vibrio cholerae O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proc Natl Acad Sci U S A* **2007,** *104*, (45), 17801-6.
18. KNBS *The 2009 Kenya Population and Housing Census*; Kenyan National Bureau of Statistics: Kenya, 2010.
19. Habitat, U., Situation analysis of informal settlements in Kisumu. *Cities Without Slums Sub-Regional Programme for Eastern and Southern Africa. Kenya Slum Upgrading Programme* **2005**.
20. Kenya Slum Update Programme; UN Human Settlements Programme *Situational Analysis of Informal Settlements in Kisumu*; Nairobi, Kenya, 2005.
21. Statistics, K. N. B. o. *Nyanza Province Multiple Indicator Cluster Survey 2011*; Kenyan National Bureau of Statistics: Nairobi, Kenya, 2013.
22. Barnes, A. N.; Mumma, J.; Cumming, O., Role, ownership and presence of domestic animals in peri-urban households of Kisumu, Kenya. *Zoonoses Public Health* **2018,** *65*, (1), 202-214.
23. Simiyu, S.; Swilling, M.; Cairncross, S.; Rheingans, R., Determinants of quality of shared sanitation facilities in informal settlements: case study of Kisumu, Kenya. *BMC Public Health* **2017,** *17*, (1), 68.
24. EPA)., U. S. E. P. A. U. S. Recreational water quality criteria. http://water.epa.gov/scitech/swguidance/standards/criteria/health/recreation/index.cfm (1/26/2018),
25. Griffith, J. F.; Cao, Y.; McGee, C. D.; Weisberg, S. B., Evaluation of rapid methods and novel indicators for assessing microbiological beach water quality. *Water Res* **2009,** *43*, (19), 4900-7.
26. Ahmed, W.; Harwood, V. J.; Gyawali, P.; Sidhu, J. P.; Toze, S., Comparison of concentration methods for quantitative detection of sewage-associated viral markers in environmental waters. *Appl Environ Microbiol* **2015,** *81*, (6), 2042-9.
27. Haramoto, E.; Katayama, H.; Asami, M.; Akiba, M., Development of a novel method for simultaneous concentration of viruses and protozoa from a single water sample. *J Virol Methods* **2012,** *182*, (1-2), 62-9.
28. Liu, J.; Gratz, J.; Amour, C.; Kibiki, G.; Becker, S.; Janaki, L.; Verweij, J. J.; Taniuchi, M.; Sobuz, S. U.; Haque, R.; Haverstick, D. M.; Houpt, E. R., A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. *J Clin Microbiol* **2013,** *51*, (2), 472-80.
29. Ishii, S.; Kitamura, G.; Segawa, T.; Kobayashi, A.; Miura, T.; Sano, D.; Okabe, S., Microfluidic quantitative PCR for simultaneous quantification of multiple viruses in environmental water samples. *Appl Environ Microbiol* **2014,** *80*, (24), 7505-11.
30. Stroup, S. E.; Roy, S.; McHele, J.; Maro, V.; Ntabaguzi, S.; Siddique, A.; Kang, G.; Guerrant, R. L.; Kirkpatrick, B. D.; Fayer, R.; Herbein, J.; Ward, H.; Haque, R.; Houpt, E. R., Real-time PCR detection and speciation of Cryptosporidium infection using Scorpion probes. *Journal of medical microbiology* **2006,** *55*, (Pt 9), 1217-22.
31. Liu, J.; Platts-Mills, J. A.; Juma, J.; Kabir, F.; Nkeze, J.; Okoi, C.; Operario, D. J.; Uddin, J.; Ahmed, S.; Alonso, P. L.; Antonio, M.; Becker, S. M.; Blackwelder, W. C.; Breiman, R. F.; Faruque, A. S.; Fields, B.; Gratz, J.; Haque, R.; Hossain, A.; Hossain, M. J.; Jarju, S.; Qamar, F.; Iqbal, N. T.; Kwambana, B.; Mandomando, I.; McMurry, T. L.; Ochieng, C.; Ochieng, J. B.; Ochieng, M.; Onyango, C.; Panchalingam, S.; Kalam, A.; Aziz, F.; Qureshi, S.; Ramamurthy, T.; Roberts, J. H.; Saha, D.; Sow, S. O.; Stroup, S. E.; Sur,

604 D.; Tamboura, B.; Taniuchi, M.; Tennant, S. M.; Toema, D.; Wu, Y.; Zaidi, A.; Nataro, J. P.; Kotloff, K. L.;
605 Levine, M. M.; Houpt, E. R., Use of quantitative molecular diagnostic methods to identify causes of
606 diarrhoea in children: a reanalysis of the GEMS case-control study. *Lancet* **2016,** *388*, (10051), 1291-301.
607 32. Balint, M.; Bahram, M.; Eren, A. M.; Faust, K.; Fuhrman, J. A.; Lindahl, B.; O'Hara, R. B.; Opik, M.;
608 Sogin, M. L.; Unterseher, M.; Tedersoo, L., Millions of reads, thousands of taxa: microbial community
609 structure and associations analyzed via marker genes. *FEMS Microbiol Rev* **2016,** *40*, (5), 686-700.
610 33. Xiao, X. Y.; Wang, M. W.; Zhu, H. W.; Guo, Z. H.; Han, X. Q.; Zeng, P., Response of soil microbial
611 activities and microbial community structure to vanadium stress. *Ecotoxicol Environ Saf* **2017,** *142*, 200-
612 206.
613 34. Brandt, K. K.; Amezquita, A.; Backhaus, T.; Boxall, A.; Coors, A.; Heberer, T.; Lawrence, J. R.;
614 Lazorchak, J.; Schonfeld, J.; Snape, J. R.; Zhu, Y. G.; Topp, E., Ecotoxicological assessment of antibiotics: A
615 call for improved consideration of microorganisms. *Environ Int* **2015,** *85*, 189-205.
616 35. Harris, V. C.; Armah, G.; Fuentes, S.; Korpela, K. E.; Parashar, U.; Victor, J. C.; Tate, J.; de Weerth,
617 C.; Giaquinto, C.; Wiersinga, W. J.; Lewis, K. D.; de Vos, W. M., Significant Correlation Between the Infant
618 Gut Microbiome and Rotavirus Vaccine Response in Rural Ghana. *J Infect Dis* **2017,** *215*, (1), 34-41.
619 36. David, L. A.; Weil, A.; Ryan, E. T.; Calderwood, S. B.; Harris, J. B.; Chowdhury, F.; Begum, Y.;
620 Qadri, F.; LaRocque, R. C.; Turnbaugh, P. J., Gut microbial succession follows acute secretory diarrhea in
621 humans. *MBio* **2015,** *6*, (3), e00381-15.
622 37. Mwaikono, K. S.; Maina, S.; Sebastian, A.; Schilling, M.; Kapur, V.; Gwakisa, P., High-throughput
623 sequencing of 16S rRNA Gene Reveals Substantial Bacterial Diversity on the Municipal Dumpsite. *BMC*
624 *Microbiol* **2016,** *16*, (1), 145.
625 38. Sun, H.; He, X.; Ye, L.; Zhang, X. X.; Wu, B.; Ren, H., Diversity, abundance, and possible sources of
626 fecal bacteria in the Yangtze River. *Appl Microbiol Biotechnol* **2017,** *101*, (5), 2143-2152.
627 39. Ben Said, L.; Klibi, N.; Dziri, R.; Borgo, F.; Boudabous, A.; Ben Slama, K.; Torres, C., Prevalence,
628 antimicrobial resistance and genetic lineages of Enterococcus spp. from vegetable food, soil and
629 irrigation water in farm environments in Tunisia. *J Sci Food Agric* **2016,** *96*, (5), 1627-33.
630 40. Penakalapati, G.; Swarthout, J.; Delahoy, M. J.; McAliley, L.; Wodnik, B.; Levy, K.; Freeman, M. C.,
631 Exposure to Animal Feces and Human Health: A Systematic Review and Proposed Research Priorities.
632 *Environ Sci Technol* **2017,** *51*, (20), 11537-11552.
633 41. Freeman, M. C.; Garn, J. V.; Sclar, G. D.; Boisson, S.; Medlicott, K.; Alexander, K. T.; Penakalapati,
634 G.; Anderson, D.; Mahtani, A. G.; Grimes, J. E. T.; Rehfuess, E. A.; Clasen, T. F., The impact of sanitation
635 on infectious disease and nutritional status: A systematic review and meta-analysis. *Int J Hyg Environ*
636 *Health* **2017,** *220*, (6), 928-949.
637 42. Dock, G.; Bass, C. C., *Hookworm disease; etiology, pathology, diagnosis, prognosis, prophylaxis,*
638 *and treatment*. C.V. Mosby Co.: St. Louis, 1910.
639 43. Kagira, J. M.; Kanyari, P. W., Questionnaire survey on urban and peri-urban livestock farming
640 practices and disease control in Kisumu municipality, Kenya. *J S Afr Vet Assoc* **2010,** *81*, (2), 82-6.
641 44. Rollins, D. M.; Colwell, R. R., Viable but nonculturable stage of Campylobacter jejuni and its role
642 in survival in the natural aquatic environment. *Appl Environ Microbiol* **1986,** *52*, (3), 531-8.
643 45. Islam, M. S.; Hossain, M. A.; Khan, S. I.; Khan, M. N.; Sack, R. B.; Albert, M. J.; Huq, A.; Colwell, R.
644 R., Survival of Shigella dysenteriae type 1 on fomites. *J Health Popul Nutr* **2001,** *19*, (3), 177-82.
645 46. Lothigius, A.; Sjoling, A.; Svennerholm, A. M.; Bolin, I., Survival and gene expression of
646 enterotoxigenic Escherichia coli during long-term incubation in sea water and freshwater. *J Appl*
647 *Microbiol* **2010,** *108*, (4), 1441-9.
648 47. Romanowski, G.; Lorenz, M. G.; Sayler, G.; Wackernagel, W., Persistence of free plasmid DNA in
649 soil monitored by various methods, including a transformation assay. *Appl Environ Microbiol* **1992,** *58*,
650 (9), 3012-9.

651 48. Dejean, T.; Valentini, A.; Duparc, A.; Pellier-Cuit, S.; Pompanon, F.; Taberlet, P.; Miaud, C.,
652 Persistence of environmental DNA in freshwater ecosystems. *PLoS One* **2011,** *6*, (8), e23398.
653 49. Bae, S.; Wuertz, S., Decay of host-associated Bacteroidales cells and DNA in continuous-flow
654 freshwater and seawater microcosms of identical experimental design and temperature as measured by
655 PMA-qPCR and qPCR. *Water Res* **2015,** *70*, 205-13.
656 50. Eichmiller, J. J.; Borchert, A. J.; Sadowsky, M. J.; Hicks, R. E., Decay of genetic markers for fecal
657 bacterial indicators and pathogens in sand from Lake Superior. *Water Res* **2014,** *59*, 99-111.
658 51. Kim, M.; Wuertz, S., Survival and persistence of host-associated Bacteroidales cells and DNA in
659 comparison with Escherichia coli and Enterococcus in freshwater sediments as quantified by PMA-qPCR
660 and qPCR. *Water Res* **2015,** *87*, 182-92.
661 52. Rogers, S. W.; Donnelly, M.; Peed, L.; Kelty, C. A.; Mondal, S.; Zhong, Z.; Shanks, O. C., Decay of
662 bacterial pathogens, fecal indicators, and real-time quantitative PCR genetic markers in manure-
663 amended soils. *Appl Environ Microbiol* **2011,** *77*, (14), 4839-48.

664

665 Table 1. Ecological and Sanitation characteristics of public domains of three Kisumu neighborhoods.

Observed Sanitary Indicators	Nya-A, N=55	Nya-B, N=53	Obu., N=58	Total N=166
Ecological Factors				
Landscape developed with housing, n (%) Landscape with open lots/fields, n (%)	40 (72.7%) 15 (27.3%)	39 (73.6%) 14 (26.4%)	45 (77.6%) 13 (22.4%)	124 (74.7%) 42 (75.3%)
Relative humidity, percent	61.6 (3.1)	60.1 (0.4)	62.8 (1.1)	61.6 (2.2)
Altitude, meters	1,147.3 (3.9)	1,145.8 (4.7)	1,162.1 (8.8)	1,152 (9.7)
Temperature, Celsius	24.4 (1.7)	23.4 (1.1)	22.1 (0.99)	23.3 (1.4)
Surface Water present, n (%)	15 (27.3%)	11 (20.8%)	14 (26.4%)	40 (24.1%)
Potential Fecal Sources				
Human/animal feces on ground, n (%)	37 (67.3%)	36 (67.9%)	44 (75.9%)	117 (70.5%)
Human open defecation indicators, n (%)	13 (23.6%)	19 (35.8%)	15 (25.9%)	47 (28.3%)
Domestic Animals present, n (%) - Pig - Chicken - Cattle - Dog - Sheep/Goat	41 (75.6%) 2 (3.6%) 22 (39.3%) 8 (14.3%) 9 (16.1%) 5 (8.9%)	42 (79.3%) 2 (3.8%) 23 (43.4%) 11 (20.8%) 13 (24.5%) 12 (22.6%)	45 (77.6%) 2 (3.5%) 34 (58.6%) 12 (20.7%) 6 (10.3%) 13 (22.4%)	128 (77.1%) 6 (3.6%) 79 (47.3%) 31 (18.6%) 28 (16.8%) 30 (17.9%)
Flies Present, n (%)	51 (96.2%)*	50 (94.3%)	56 (96.6%)	157 (95.7%)*
Shared or Public Latrine On-site, n (%) - Functional Condition	45 (81.8%) 0 (0%)	44 (83%) 9 (17.0%)	39 (67.2%) 0 (0%)	130 (78.3%) 9 (5.4%)

- Dilapidated Condition	44 (100%)	36 (67.9%)	39 (100%)	121 (72.9%)
Children under 5 present, n (%)	24 (43.6%)	17 (32.1%)	31 (53.5%)	72 (43.4%)

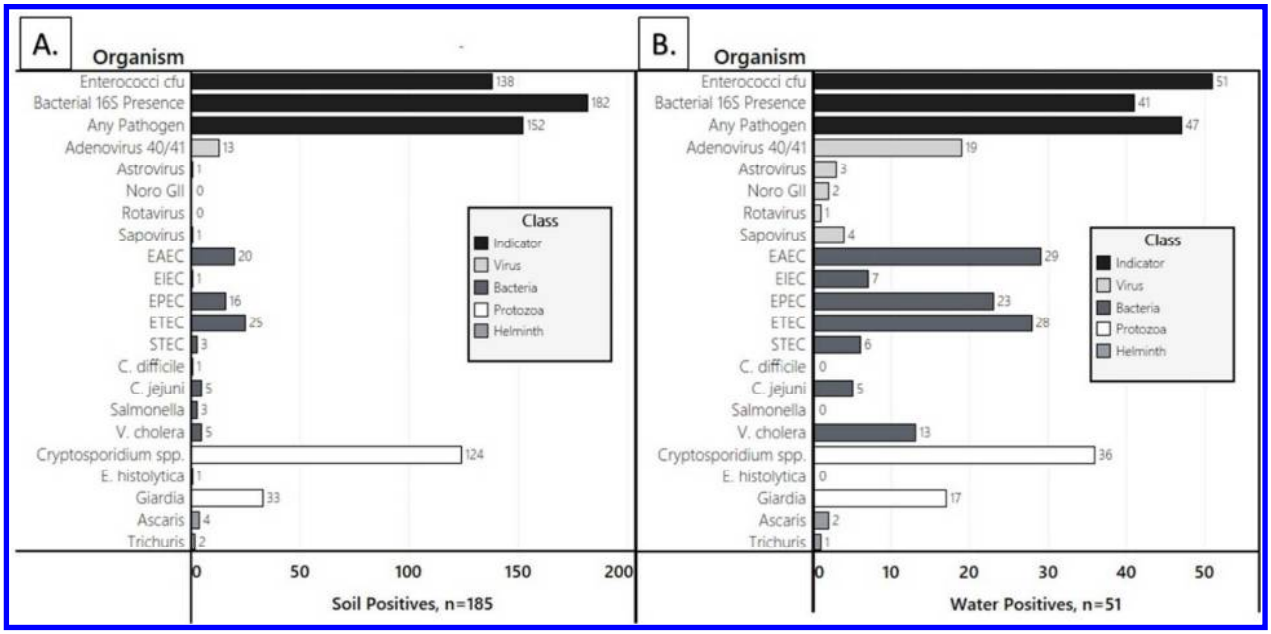
666 [a] Total of 145 single sample sites and 21 multi-sample (more than one soil sample) sites. Nyalenda A

667 (Nya-A); Nyalenda B (Nya-B); Obunga (Obu.).

668 * Denominator = 163.

669 Figure 1. Detection frequency for enterococci bacteria, bacterial 16S DNA, and enteric virus, bacteria,

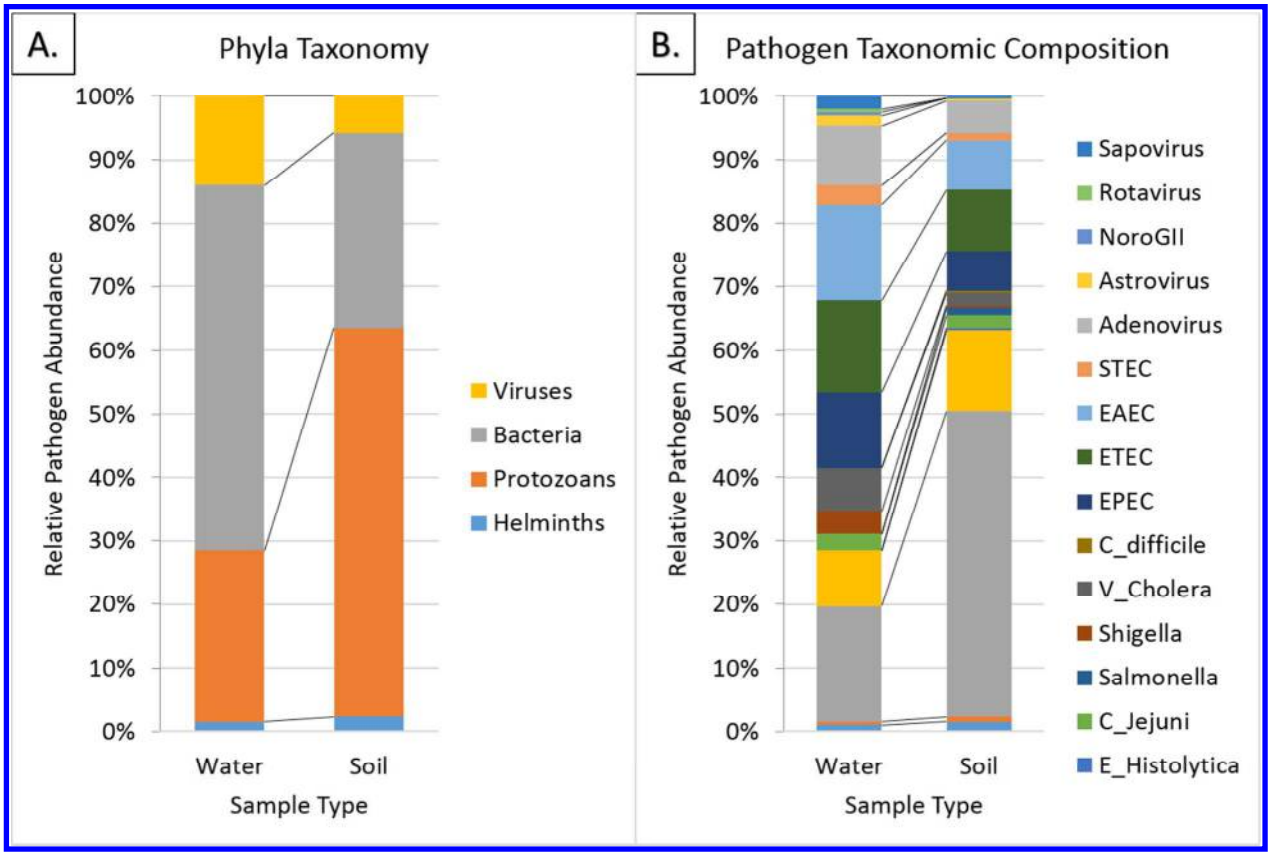
670 protozoan, or helminth pathogens in soil (A) and water (B) from Kisumu neighborhoods.



671

672 Figure 2. Relative abundance of pathogen gene copies for phylogenetic groups (A) and categories of

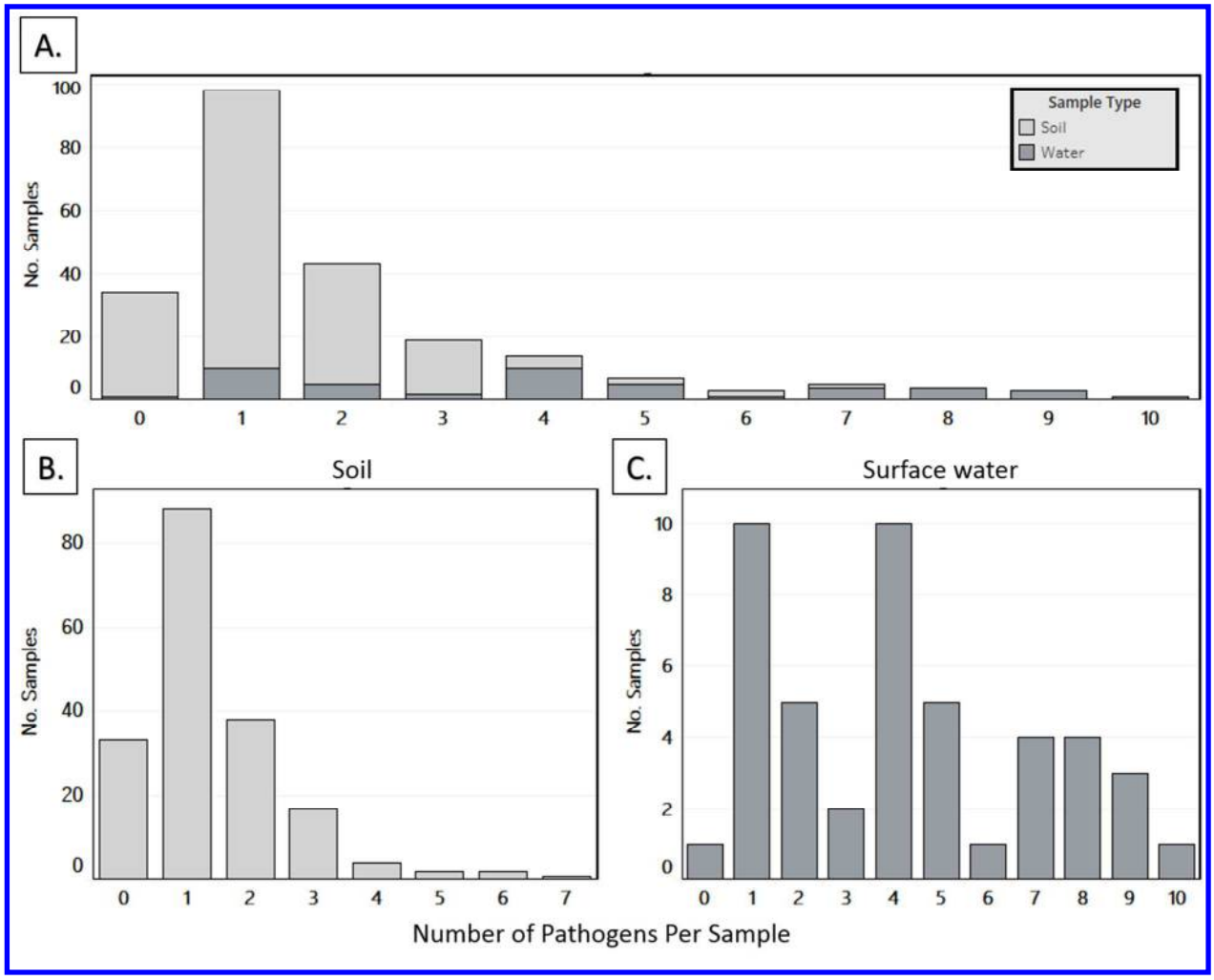
673 enteric pathogens (B) in soil and water from Kisumu neighborhoods.



674

675 Figure 3. Histogram of diversity in enteric pathogens detected in soil or surface water combined (A), soil

676 (B), and surface water (C) from Kisumu neighborhoods.



677

678 Table 2. Association between human sanitation and animal feces sources and the presence and

679 concentration of enterococcus in soil from public domains from Kisumu neighborhood public domains.

Presence of Enterococcus			
Exposure	% (*n* positive/N)	Unadj. OR (95% C.I.)	Adj. OR (95% C.I.)
Observed feces on ground	80.7% (125/155)	2.67 (0.41, 17.44)	2.42 (0.62, 9.45)
None	76.1% (54/71)	Ref.	Ref.
Observed open defecation indicators	71.0% (44/62)	1.42 (0.31, 6.55) [b]	0.86 (0.22, 3.28)
None	82.3% (135/164)	Ref.	Ref.
Domestic Animals	83.0% (146/176)	2.88 (0.35, 24.03)	3.42 (0.69, 17.03)
None	66.0% (33/50)	Ref.	Ref.
Flies	79.9% (171/214)	1.73 (0.11, 27.31)	0.72 (0.05, 11.39)
None	57.1% (4/7)	Ref.	Ref.
Latrine condition - No latrines	71.1% (32/45)	Ref.	Ref.
Latrine condition - Good condition	66.7% (8/12)	0.78 (0.10, 5.85)	0.59 (0.05, 7.40)
Latrine condition - Bad condition	82.5% (139/169)	1.07 (0.19, 6.03)	1.03 (0.19, 5.56)
Log_{10} Concentration of Enterococcus			
Exposure	Mean (SD)	Unadj. RR (95% C.I.)	Adj. RR (95% C.I.)
Observed feces on ground	2.5 (2.0)	1.65 (0.92, 2.96)	1.25 (0.66, 2.37)
None	2.4 (2.1)	Ref.	Ref.
Observed open defecation indicators	2.1 (2.0)	1.63 (0.87, 3.05) [b]	1.18 (0.60, 2.33)
None	2.6 (2.0)	Ref.	Ref.
Domestic Animals	2.6 (2.0)	1.82 (0.86, 3.85)	1.50 (0.63, 3.56)
None	2.0 (2.1)	Ref.	Ref.

Flies	2.5 (2.0)	**3.02 (1.39, 6.56)**	1.88 (0.74, 4.76)
None	0.6 (0.9)	Ref.	Ref.
Latrine condition - No latrines	1.9 (2.0)	Ref.	Ref.
Latrine condition - Good condition	1.2 (1.6)	0.67 (0.23, 1.94)	0.67 (0.22, 2.09)
Latrine condition - Bad condition	2.7 (2.0)	1.50 (0.70, 3.22)	1.31 (0.57, 3.01)

680 [a] All models are adjusted for landscape features, relative humidity, altitude, and ambient temperature.
681 Adjusted models also include all five exposure variables at the same time. Standard Deviation (SD); Odds
682 Ratio (OR); Risk Ratio (RR); Confidence Interval (CI); Reference (Ref.). Surface water not included
683 because all samples were positive. Statistical significance (P<0.05) indicated using bold typeface.

684 [b] The unadjusted OR and RR for this model reflect significant confounding by an ecological variable.
685 Bivariate association for open defecation indicators and presence of enterococcus is unadj. OR=0.50
686 (95% C.I. 0.19, 1.34), and the bivariate model for open defecation indicators and log10 concentration of
687 enterococcus is unadj. RR=0.92 (95% C.I. 0.48, 1.76).

688

689 Table 3. Association between human sanitation and animal feces sources with the presence and

690 diversity in enteric pathogens from Kisumu neighborhood public domains.

Presence of Any Pathogen			
Exposure	% (*n* positive/N)	Unadj. OR (95% C.I.)	Adj. OR (95% C.I.)
Observed feces on ground	83.9% (130/155)	1.41 (0.41, 4.87)	1.01 (0.25, 4.04)
None	83.1% (59/71)	Ref.	1.02 Ref.
Observed open defecation indicators	82.3% (51/62)	1.46 (0.41, 5.19)	0.98 (0.22, 4.52)
None	84.2% (138/164)	Ref.	Ref.
Domestic Animals	85.3% (151/177)	1.60 (0.43, 5.91)	1.53 (0.28, 8.54)
None	77.6% (38/49)	Ref.	Ref.
Flies	84.7% (182/215)	17.1 (0.70, 418.0)	5.15 (0.20, 130.1)
None	42.9% (3/7)	Ref.	Ref.
Latrine condition - No latrines	84.4% (38/45)	Ref.	Ref.
Latrine condition - Good condition	66.7% (8/12)	0.16 (0.01, 2.72)	0.70 (0.10, 4.99)
Latrine condition - Bad condition	84.6% (143/169)	0.78 (0.17, 3.54)	0.98 (0.21, 4.52)
Pathogen Diversity			
Exposure	Mean (SD)	Unadj. RR (95% C.I.)	Adj. RR (95% C.I.)
Observed feces on ground	1.9 (1.9)	1.04 (0.80, 1.34)	1.0 (0.75, 1.33)
None	1.8 (2.0)	Ref.	2.0 Ref.
Observed open defecation indicators	1.4 (1.5)	0.95 (0.70, 1.29)	0.88 (0.63, 1.22)
None	2.0 (2.1)	Ref.	Ref.
Domestic Animals	2.0 (2.1)	**1.39 (1.03, 1.89)**	**1.44 (1.01, 2.06)**
None	1.3 (1.2)	Ref.	Ref.

Flies	1.9 (2.0)	2.35 (1.00, 5.56)	2.00 (0.84, 4.79)
None	0.6 (0.8)	Ref.	Ref.
Latrine condition - No latrines	1.5 (1.7)	Ref.	Ref.
Latrine condition - Good condition	0.8 (0.8)	0.77 (0.39, 1.51)	0.78 (0.40, 1.53)
Latrine condition - Bad condition	2.0 (2.0)	1.12 (0.79, 1.58)	1.08 (0.74, 1.57)

691 [a] Adjusted for landscape features, relative humidity, altitude, and ambient temperature. Unadjusted
692 (unadj.); Adjusted (Adj.); Standard Deviation (SD); Odds Ratio (OR); Risk Ratio (RR); Confidence Interval
693 (CI); Reference (Ref.). Statistical significance (P<0.05) indicated using bold typeface.

694 J.

Example of a public site in Kisumu impacted by multiple potential sources of enteric pathogens: a communal latrine, a domestic goat, and an informal drain used for water management and and waste disposal.

285x508mm (96 x 96 DPI)