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

Just Accepted

1   Fecal Fingerprints of Enteric Pathogen Contamination in Public Environments of Kisumu, Kenya

2   associated with Human Sanitation Conditions and Domestic Animals

3

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

5

6   [1] Department of Occupational and Environmental Health, University of Iowa, Iowa City, Iowa, United

7   States

8   [2] Department of Biostatistics, University of Iowa, Iowa City, Iowa, United States

9   [3] Department of Electrical Engineering, University of Iowa, Iowa City, Iowa, United States

10   [4] Department of Disease Control, London School of Hygiene and Tropical Medicine, London, United

11   Kingdom

12   [5] Department of Community Nutrition, Great Lakes University of Kisumu, Kisumu, Kenya

13

14   * Corresponding author: Kelly K. Baker, University of Iowa College of Public Health, 145 S. Riverside Dr.,

15   S316, Iowa City, IA 52246; (001) 319-384-4008; kelly-k-baker@uiowa.edu.

16

17

18 **Abstract**

19   Young children are infected by a diverse range of enteric pathogens in high disease burden

20   settings, suggesting pathogen contamination of the environment is equally diverse. This study aimed to

21   characterize across and within neighborhood diversity in enteric pathogen contamination of public

22   domains in urban informal settlements of Kisumu, Kenya, and to assess the relationship between

23   pathogen detection patterns and human and domestic animal sanitation conditions. Microbial

24   contamination of soil and surface water from 166 public sites in three Kisumu neighborhoods was

25   measured by enterococcus assays and qRT-PCR for nineteen enteric pathogens. Regression was used to

26   assess the association between observed sanitary indicators of contamination with enterococcus and

27   pathogen presence and concentration, and pathogen diversity. Seventeen types of pathogens were

28   detected in Kisumu public domains. Enteric pathogens were co-detected in 33% of soil and 65% of

29   surface water samples. Greater pathogen diversity was associated with the presence of domestic animal

30   feces, but not with human open defecation, deteriorating latrines, flies, or disposal of human feces.

31   Sanitary conditions were not associated with enterococcus bacteria, specific pathogen concentrations,

32   or "any pathogen". Young children played at 40% of observed sites. Managing domestic animal feces

33   may be required to reduce enteric pathogen environmental contamination in high-burden settings.

34   **INTRODUCTION**

35   Recent multi-country epidemiological studies of diarrheal disease etiology in high-burden

36   settings have shown a wide diversity of enteric pathogen infections within and between individual

37   children, and between populations and population sub-groups (i.e. by age) [1, 2]. The diversity in

38   pathogens causing childhood enteric infections shows that children under the age of five years (<5yr) in

39   these settings are chronically exposed to a range of environmentally transmitted enteric pathogens.

40   Relatively little is known about the extent of environmental enteric pathogen contamination in these

41   settings, and the patterns with which enteric pathogen contamination occurs over space and time. A

42   handful of studies have reported high frequencies of detection of different types of enteric pathogens in

43   Indian and Tanzanian household environments, confirming household exposure pathways pose a risk to

44   children [3-5]. However, recent multi-million dollar randomized, controlled trials of household water,

45   sanitation, and hygiene (WASH) interventions found inconsistent impact of household WASH on

46   pediatric diarrhea [6-8]. This suggests that children may be infected by non-household exposure pathways,

47   such as through play in public areas near their household. In low-income settings where sanitation

48   coverage is low, neighborhood public areas are often used for open defecation and disposal of

49   untreated human waste by many households, as well as for domestic animal husbandry [9]. Thus, levels of

50   pathogen contamination in public areas may far exceed contamination levels in privately owned

51   domestic areas, and could pose a disproportionally high risk of infection by enteric pathogens for

52   exposed children.

53   The overarching goal of the Social Microbes Study is to examine enteric pathogen transmission

54   patterns in high disease burden settings, from sources of fecal contamination through the environment

55   to children, using pathogen distribution and diversity as indicators of transmission pathways. This is the

56   first Social Microbes report to be released, which focuses on testing the hypothesis that public areas

57   within low-income neighborhoods with high disease burden are contaminated by a diverse set of enteric

58   pathogens, and to test whether observed indicators of human and animal feces contamination would be

59   associated with increased presence and total diversity in pathogens. We also measured how frequently

60   <5yr children are observed in public areas, and whether we saw unsafe exposure behaviors, to assess

61   whether public domain exposures are viable pathways for <5yr enteric infection.

62   Environmental fecal contamination is usually measured by general fecal indicator bacteria, fecal

63   source tracking markers, or focused detection of specific pathogens of interest as indicators of exposure

64   risk. Fecal *E. coli* and enterococci bacteria are popular, low-cost indicators, but typically correlate poorly

65   with other infectious pathogens in sewage, ecological soils, surfaces, and waters [10-13]. They may be

66   particularly unsuitable indicators for measuring risks from feces in ecological systems where gut bacteria

67   become a naturalized part of soil, water, and surface microbial communities [14]. Host-specific fecal

68   source tracking markers have improved capacity to distinguish human versus animal fecal contamination

69   of the environment, but the reliability of fecal source tracking markers for predicting infectious

70   pathogens is currently unclear [10, 15]. Selection of one or a few specific types of enteric pathogens is often

71   used to understand transmission patterns of specific types of pathogens [16, 17]. However, dozens of

72   enteric viruses, bacteria, protozoans, and helminthic species circulate at different times of the year in

73   endemic settings, so reliance on any one to predict risks from all pathogens transmitted by fecal-oral

74   exposure pathways is risky [1, 2]. In this study, environmental samples were tested for a wide range of

75   common enteric viruses, bacteria, protozoans, and helminths involved in fecal-oral disease transmission

76   in high-burden settings to reduce the likelihood of exposure misclassification (classifying samples as

77   uncontaminated based upon one indicator) or inaccurate concentration estimates for environmental

78   exposure pathways.

79   Pathogen diversity was an important systems-level evaluation metric in this study that was

80   adopted based upon the theory that diversity in enteric pathogen contamination would occur not just

81   across high disease burden neighborhoods, but also at fine spatial scales. If environmental microbial

4

82   contamination by human and animal feces is pervasive in high disease burden settings, with multiple

83   pathogens circulating at any given time, then there is a strong chance that multiple enteric pathogens

84   co-occur in some environments (e.g. open defecation sites) at the same time. Exposure of children to

85   these uniquely hazardous environments could increase their risk of infection by multiple pathogen

86   types. The risk of exposure to any given pathogen from soil or surface water is influenced by the recent

87   presence of a fecal source, but also by the pathogen's capacity to persist in specific environments, like

88   soil or surface water, and overall infection rate in a host population. The evolutionary traits that govern

89   persistence of microbes may be unique to members of a phylogenetic group (e.g. non-obligate bacterial

90   replication in the environment), as well as reflect species-specific ecological adaptations. Examining both

91   the pathogen-specific and higher order taxonomic differences in pathogens in the environment could

92   provide insight as to the types of pathogens potentially transmitted via different child behaviors, e.g.

93   ingesting soil vs water. From a programming standpoint, capacity to identify high-risk areas or exposure

94   pathways, in a milieu of existing elevated background contamination, could improve how well

95   investments reduce <5yr enteric disease in high-burden settings. Pathogen diversity could improve the

96   identification of high- versus low-risk areas or exposure pathways, and the fecal sources contributing to

97   environmental contamination.

98   This manuscript describes enteric pathogen detection frequencies and diversity patterns in soil

99   and surface water from public areas of three low-income, urban neighborhoods of Kisumu, Kenya with

100   low sanitation coverage. Then, we examine the relationship between human and domestic animal fecal

101   sources and pathogen detection and diversity patterns.

102

103   **MATERIALS AND METHODS**

104   **Study Population.** Kisumu is the third largest city in Kenya, and has a population of approximately

105   409,928 inhabitants [18]. Up to 60% of the city's population resides in peri-urban informal settlements ,

106   which have emerged due to economic migration and a lack of affordable housing [19, 20]. Kisumu County

107   has a high prevalence of diarrhea (18% two-week period prevalence) with most cases of diarrhea

108   occurring in children less than three years of age [21]. The child mortality rate is 105 deaths per 1,000 live

109   births and the prevalence of severe childhood stunting (> -2 standard deviations below the reference

110   norm) is approximately 25% [21]. This study took place in three established informal settlements,

111   Nyalenda A (population density (pd) = 8,953 persons per square kilometer (km^2)) Nyalenda B (pd =

112   6,886/km^2), and Obunga (pd=1,913 km^2). Co-habitation with domestic animals is common, and most

113   residents rely upon open defecation or sanitation facilities shared by eight or more households [22, 23].

114   **Pilot of Environmental Sampling and Microbiology Methods.** Prior to conducting the primary study in

115   Kenya, a pilot sampling project was conducted in Iowa in June 2015 (similar climate conditions) to

116   understand how often fecal indicators and enteric pathogens would be detected in a watershed

117   impacted by agriculture, free-range animal management (cows), and concentrated animal feeding

118   operations (pigs). This information is presented in Supplemental documents as a comparative baseline

119   for understanding the importance of fecal indicator and pathogen contamination in Kisumu, Kenya.

120   **Sample site selection.** A cross-sectional observational survey with environmental sampling was

121   performed in Kenya in July 2015. The study was designed to ensure sampling sites were randomly

122   distributed across special areas of interest to optimize measurement of variability in pathogen diversity

123   and co-occurrence patterns and to prevent introduction of observer bias in selection of environmental

124   sampling locations. Neighborhood boundary parameters were visualized utilizing Batchgeo (Google),

125   and these spaces were defined geographically by the northernmost, westernmost, easternmost, and

126   southernmost latitude or longitude values in rectangular form. For each neighborhood, sixty latitude

127   and longitude pairs that fell within these specified ranges were randomly generated utilizing the

128   website, geomidpoint.com, prior to field based data collection (Figure S1). GPS coordinates were

129   entered into a Waytracker mobile phone app, and daily routes were identified to navigate between ~ 15

130   coordinates per day. Observers navigated to the coordinates (+/- 3 meters), which was considered the

131   center of a "site", defined as all area falling within a 25 meter radius around the central coordinates. If

132   the coordinates fell within a private household yard or business, the nearest set of coordinates outside

133   that private space but within 25 meters of the random coordinates were identified. Sites were visited at

134   various times during the day (morning and afternoon), with site visits lasting approximately 20 minutes.

135   **Public site observation.** Study teams systematically documented sanitary conditions at each site using

136   an observational survey implemented via mobile phone app (Fieldlogs). Observers documented

137   landscape features such as surface waters, grasslands and altitude. Development of infrastructure was

138   noted, including roads, drains, dams, industry, housing, public water sources, and presence, physical

139   condition, and hygiene of public or communal latrines. Indicators of human open defecation or unsafe

140   disposal of excreta included observed "flying toilets" (plastic bags containing excreta), used diapers,

141   piles of human feces, emptying of latrines into drains or land around the latrine, septage emptied from a

142   latrine next to the latrine, and visual confirmation of an adult or child actively defecating in the open.

143   Presence and type of domestic animals and their feces were recorded. Observers recorded whether

144   during that ~ 10-15 minute spot observation, children approximately < 5 yrs were observed in the public

145   area, the number present, and any behaviors that would result in hand or mouth contact with

146   environmental fomites (touching soil, surface water, animals, or objects on the ground, swimming,

147   eating food, eating dirt, mouthing hands). Additionally, enumerators documented whether any children

148   were defecating in the open at the time of observation. Ecological conditions that could influence

149   pathogen presence and persistence, specifically daily temperature and relative humidity were extracted

150   from NOAA data collected at the Kisumu Airport. Observers did not include area within private housing

151   or businesses adjoining or overlapping with the site area to avoid potential bias in conditions observed

152   from privately held property. However, private conditions that impacted public space, such as drains or

153   sewage leaching from the household or animals roaming between public and private areas, were

154   recorded.

155   **Sample collection.** Standard Operating Protocols were implemented to ensure standardized hygienic

156   sampling and processing of environmental materials. Approximately five grams of soil was collected at

157   every site by inserting an alcohol-sterilized scoop into the ground at a 45° angle to a depth of 5cm (half

158   the length of the scoop) and transferring the soil into a sterile WhirlPak® bag (Sigma-Aldrich Corp., St.

159   Louis, MO, USA). Surface water was collected, if present, by skimming water into a WhirlPak bag.

160   Collection bags were stored on ice packs in a cooler and transported to the laboratory within six hours of

161   collection. In Kenya, four samples were collected at seven randomized sites in each neighborhood to

162   account for anticipated variance in pathogen distributions at public sites caused by the final spatial scale

163   of sites per neighborhood.

164   **Indicator analysis.** One gram of soil was measured into 10 mls of Phosphate Buffered Saline (PBS),

165   vortexed for 30 seconds, and then mixed on a rotator for 20 minutes. Solid matter was allowed to settle

166   for 5 minutes and elute was removed for Enterococci assays. Enterococci are recommended by the US

167   Environmental Protection Agency for identifying fecal material in fresh and marine recreational waters

168   [24], although both enterococci and *E. coli* are considered inadequate indicators in tropical settings like

169   Hawaii [15, 25]. There is no global recommendation on appropriate indicators for tropical waters, so

170   enterococci were chosen based upon the EPA policy. Enterococcus were enumerated by vacuum

171   filtration of three serial dilutions of surface water (10 millilter (mL), 1 mL, 0.1 mL) or soil (1 mL, 0.1 mL,

172   0.01 mL) rinse through a white gridded 0.45 μm mixed cellulose esters filter (Product No. GSWP04700,

173   Millipore, Billerica, MA), and culturing filters for 18-24 hours at 37°C on mEI agar (EPA method 1600).

174   Colony forming units (cfu) of enterococci were counted according to manufacturer's recommendations.

175   **DNA and RNA extraction.** DNA and RNA was extracted from 0.5 grams of soil using the FastDNA and

176   FastRNA SPIN kits for Soil (MP Biomedicals, Solon, OH), including a bead beating step. A 10 mL volume of

177   surface water was processed by adjusting water to 2.5 mM MgCl2 and pH 10.0, and vacuum filtering

178   through a 0.45 µm mixed cellulose esters filter. This volume was chosen to ensure we could

179   systematically sample a wide range of surface waters (flood water puddles, drains, rivers) representing

180   potential exposure hazards in Kisumu. This method was chosen because it more efficiently removes

181   inhibitors than ultrafiltration and is practical for laboratories with limited capacity and for isolation of

182   many different types of pathogens [26, 27]. Filters were frozen at -20°C for 1-4 weeks and transported to

183   the University of Iowa where they were stored at -80°C until extraction (1 week). Filters were cut in half

184   and processed using the FastDNA and FastRNA for Soil kits (equivalent to 5 mls water).

185   **Quantitative PCR on Environmental Samples.** The concentration and quality of DNA and RNA was

186   measured using a Nanodrop UV-Vis spectrophotometer (Thermo Scientific, DE, USA). Duplicate 6

187   microliter (µl) and 0.6 µl volume of nucleic acid extract were tested for inhibition using the QuantiFast

188   Pathogen Internal Control kit (Qiagen, Germantown, MD). These methods effectively removed most

189   amplification inhibitors (Table S1). Samples that were not inhibited were analyzed by combining 20 µl of

190   DNA and RNA each (total 40 µl) with AgPath polymerase and then performing quantitative reverse

191   transcription Polymerase Chain Reaction (qRT-PCR) using a nineteen pathogen microfluidic TaqMan

192   Array Card (TAC) on a QuantiStudio 12K Flex Real-Time PCR System with Array Card block

193   (ThermoFisher, Chicago, IL) [28]. Inhibited samples were diluted 1:10 with diH20 before TAC analysis. The

194   card format included five types of viruses (adenovirus 40-41, astrovirus, sapovirus, norovirus GII,

195   rotavirus), fourteen types of bacteria (enteroaggregative *E. coli* (EAEC), enteropathogenic *E. coli* (EPEC),

196   enteropathogenic *E. coli* (ETEC), shiga toxin expressing *E. coli* (STEC), *C. jejuni*, Shigella, *V. Cholerae*, *S.*

197   *enterica*, *C. difficile*), three protozoans (*Cryptosporidium spp.*, *G. lamblia*, *E. histolytica*), and two

198   helminths (Ascaris, Trichuris). Exponential curves and multicomponent plots were visually examined to

199   validate positive amplification. Gene targets with real amplification in one well were re-analyzed and

200   considered positive if amplification was detected again. Gene targets with cycle threshold (CT) values

201   over 35 were cross-validated by performing a 14 cycle pre-amplification reaction with pathogen-specific

202   primers (0.2 μM each) to increase the starting concentration of pathogen DNA/cDNA in a sample [29], and

203   reanalyzed by qRT-PCR. Samples were classified as positive for a specific gene target if amplification was

204   verified at a lower cycle threshold, and negative if unverified. This resulted in presence and

205   concentration data for 24 pathogen gene targets representing 19 types of pathogen taxa per sample

206   (detection of either gene target for ETEC, EPEC, EAEC, and STEC was positive at a taxa level).

207   The concentration of each pathogen per reaction volume was estimated by comparing the CT

208   for a pathogen gene target in a sample to a standard curve generated by qRT-PCR of a six to seven-fold

209   serial dilution of a positive control of known concentration for each of the 19 qRT-PCR targets (Table S2).

210   For samples that were repeated using the pre-amplification step, the initial concentration was estimated

211   by comparison to a standard processed by pre-amplification. If samples determined to be positive still

212   reflected signs of significant inhibition, concentrations were excluded from analysis to avoid biasing

213   statistical estimates of mean and standard deviation. Otherwise, final concentrations of gene copy per

214   gram of soil or per mL of surface water were generated by multiplying the concentration of each gene

215   target by the dilution factors introduced by processing. Concentrations for samples with no detectable

216   amplification were not transformed to avoid inflating statistics caused by relatively high methodological

217   lower limit of detection (LLOD) for some pathogens.

218   **Statistical analysis.** Data were analyzed using SAS Version 9.4 (SAS Institute, Cary NC). Descriptive

219   statistics of variables were reported as proportions or mean and standard deviation.

220   Microbial contamination outcomes of interest included the (1) presence of enterococcus fecal

221   indicator bacteria, (2) $\log_{10}$-number of enterococcus cfu, (3) presence of any pathogen gene targets, (4)

222   "Pathogen Diversity" defined by the sum count of all unique types of enteric pathogens detected in an

223   individual soil or water sample (either target for pathogenic bacteria defined as one group), and (5) $\log_{10}$

224   gene copy of pathogen gene targets.

225   Exposure variables representing potential sources of fecal contamination were: any human or

226   animal feces observed versus none observed; indicators of open defecation vs. none observed; the

227   presence of any type of domestic animal present versus none observed; flies present versus none

228   observed; and a categorical variable for latrines present in good condition, latrines present in bad

229   condition, and no latrines present. Frequently latrines in both good and bad condition were present at a

230   site. In this case the site was classified as "latrines in bad condition" based upon the default assumption

231   that just one deteriorating latrine could introduce contamination into soil and water.

232   Potential confounders included: type of sample (soil versus water), altitude in meters, relative

233   humidity, temperature in Celsius, and landscape use (housing area versus undeveloped).

234   Generalized linear mixed models (glmm) with binary log link, robust standard errors, and

235   random intercept for site ID and neighborhood with exchangeable correlation structure to account for

236   spatial clustering at each level were used to test whether sites with any of the exposure variables were

237   more likely to have at least one type of enteric pathogen detected. Identical glmm models with Poisson

238   log link were used to test whether sites with any of the same potential fecal sources were more likely to

239   have increased sum counts in Pathogen Diversity. Censored regression models were used to assess

240   whether exposure variables were associated with increased concentration of enteric pathogen gene

241   copy number to account for left-censoring of data. Each modeling process involved analyzing the effect

242   of each exposure variable independently on the pathogen outcome, while adjusting for landscape type,

243   daily temperature, relative humidity, altitude, and type of sample (soil vs. water). Full models included

244   all variables simultaneously to measure individual effects, accounting for co-occurrence of exposure

245   variables at a site. Based upon full model results, a sub-analysis of association between specific types of

11

246   domestic animals and Pathogen Diversity was performed by replacing the binary Domestic Animals

247   variable with presence/absence variables for specific categories of animals.

248   A one-sided Mantel test using the geographic distances between sites was used to test for

249   spatial clustering of sites with compositional similarity between pairs of sites in terms of the number of

250   pathogens detected at both sites.

251

252   **RESULTS**

253   **Characteristics of study sites.**   A total of 166 public sites were inspected for this study. Ecological

254   conditions and the proportion of sites developed with housing were relatively similar in the three

255   surveyed neighborhoods, with the exception of higher mean elevation in Obunga, more sites with

256   indicators of open defecation or open feces disposal in Nyalenda B, fewer sites with shared latrines in

257   Obunga, and more latrines in good structural condition in Nyalenda B (Table 1).

258   Children less than five years of age were observed at 40% (66 of 166) sites (median 2 children

259   per site, range 1-10), including infants (defined as unable to stand and walk) at 9% (15 of 166) of sites. At

260   least one <5 child was observed crawling on the ground or sitting on the ground playing in dirt or mud at

261   24 sites, with a third of these observations (n=8) recording infants crawling or sitting in the dirt. Children

262   were observed playing with objects in the dirt at 4 sites. One infant and one child ~12-24 months were

263   observed playing in a water puddle or drain water at 2 sites. Children were seen mouthing their hands at

264   3 sites (multiple ages), and one infant was eating soil at 1 site. Open defecation was not observed.

265   **Microbial Detection, Concentration, and Neighborhood-Level Diversity in Kenya.** A total of 185 public

266   site soil and 51 water samples were collected and tested for enterococcus and enteric pathogens.

267   Enterococcus colonies were isolated from 100% of surface water samples and 74.6% of soil samples

268   (Figure 1). At least one type of enteric pathogen was detected in 92% of water samples and 82% of soil

269   samples. This included twenty gene markers representing sixteen types of enteric pathogens in water

270   and twenty gene markers representing seventeen types of pathogens in soil. The most common

271   pathogens in soil were Cryptosporidium (67.0%), Giardia (17.8%), ETEC (13.5%), EAEC (10.8%), and EPEC

272   (8.6%). The most common pathogens in water were Cryptosporidium (70.0%), EAEC (58.0%), ETEC

273   (56.0%), EPEC (46%), and human adenovirus (36.0%). The relative abundance of different biological

274   phyla and types of pathogens differed between pooled soil and water samples. Protozoan organisms

275   accounted for 61.2% of microbial organisms in soil, followed by 30.6% pathogenic bacteria, 5.8%

276   helminths, and 2.3% viruses (Figure 2). In contrast, pathogenic bacteria were more abundant in water

277   (57.5%), followed by 26.9% protozoans, 5.8% viruses, and 2.3% helminths.

278   Enterococcus concentrations were about three logs higher in concentration in water than in soil

279   (Table S3). Mean pathogen concentrations in positive samples were two to five logs higher than the limit

280   of detection, although wide standard deviations for all pathogens reflected variability in concentration

281   ranges (Table S3). Concentrations of pathogens in one gram of soil were consistently one to two logs

282   higher in concentration than in 1 mL of surface water. In soil, the adenovirus 40/41 hexon, ETEC *eltB*,

283   Salmonella *invA*, and STEC *stx2* genes were all detected at > 7.0 $\log_{10}$ concentration per gram, whereas

284   *C. difficile tcdB* was the lowest concentration at 4.2 $\log_{10}$. In surface water, adenovirus 40/41 hexon was

285   the highest at 6.6 $\log_{10}$ per mL, followed by EAEC *aaiC*, ETEC *eltB*, and STEC *stx1* and *stx2* genes at 5.6

286    $\log_{10}$ per mL. Astrovirus capsid, Norovirus GI/GII ORF1-2, and Campylobacter *cadF* at < 3 $\log_{10}$ per mL.

287   **Within-site Enteric Pathogen Diversity.** Two or more enteric pathogens were co-detected in 35%

288   (64/185) of soil samples (median 1; range 0 - 7), and 69% (35/51) of surface water samples (median 4;

289   range 0 - 10) (Figure 3). Patterns of pathogen co-detection were heterogeneous, with only 2 patterns

290   (*Cryptosporidium spp.*, EPEC, ETEC, EAEC) and (*Cryptosporidium spp.*, *Giardia spp.*, ETEC, EAEC)

291   reoccurring twice at a total of two sites each among the 35 water samples with >=2 pathogens detected.

292   Of the 64 soil samples with >=2 pathogens co-detected, *Cryptosporidium spp.*, ETEC, and EAEC occurred

13

293   three times, *Cryptosporidium spp.*, *Giardia spp.*, and ETEC twice, and the remainder were various

294   combinations of *Cryptosporidium spp.* with other types of pathogens. The *Cryptosporidium spp.* assays

295   might have detected species that do not typically infect humans[30]. If *Cryptosporidium spp.* results are

296   excluded from the analysis, then 44% (82/185) of soils and 80% (41/51) contained at least one

297   pathogen, and two or more enteric pathogens were co-detected in 18% (33/185) of soil samples

298   (median 0; range 0 - 6), and 65% (33/51) of surface water samples (median 3; range 0 - 9).

299   **Association between Sanitary Conditions and Enterococcus contamination.** Observation of feces,

300   indicators of human open defecation, domestic animals, presence and poor condition of latrines, and

301   flies were not associated with the presence or increased concentration of enterococcus indicator

302   bacterial of fecal contamination of water and soil, after adjusting for development of the terrain,

303   relative humidity, altitude, ambient temperature, and type of sample (Table 2).

304   **Association between Sanitation Conditions and Pathogen Diversity.** Observation of feces, indicators of

305   human open defecation, domestic animals, presence and poor condition of latrines, and flies were not

306   associated with the detection of any type of enteric pathogen, after adjusting for development of the

307   terrain, relative humidity, altitude, ambient temperature, and type of sample (Table 3). However, the

308   presence of domestic animals was significantly associated with increased diversity in enteric pathogens

309   in fully adjusted models. Observation of feces, indicators of human open defecation, presence and poor

310   condition of latrines, and flies were not associated with increased diversity. Sub-analysis of animal types

311   in the fully adjusted model found that chickens (adjOR=1.64; 1.22, 2.22), cattle (adjOR=1.36; 1.08, 1.72),

312   and goats and sheep (adjOR=1.39; 1.00, 1.94) were associated with increased pathogen diversity, while

313   pigs (adjOR=0.80; 0.36, 1.77) and dogs (adjOR=1.20; 0.87, 1.65) were not significantly associated. No

314   association was observed between different fecal sources and increased concentration of individual

315   types of pathogens (not shown).

316   **Spatial Clustering of sites with High Pathogen Diversity.** High pathogen diversity was not statistically

317   associated with lower distance between pairs of sites (p-value = 0.22).

318

319   **DISCUSSION**

320   This study confirmed that neighborhood landscapes in Kisumu, Kenya are contaminated by

321   many enteric virus, bacteria, protozoan, and helminth enteric pathogen species, at both neighborhood

322   and localized (within 25 meter radius) levels of spatial scale. At least one type of pathogen was detected

323   at 80% of public sites, with one third of soil and three-quarters of surfaces water co-contaminated by

324   multiple taxa of enteric pathogens. Even if Cryptosporidium detections are excluded, in the event that

325   the species detected were not human-infective, multiple pathogens were still detected in one-fifth of

326   soils and two-thirds of waters. This evidence addresses major knowledge gaps about enteric pathogen

327   co-occurrence patterns in the environment, especially for public areas that are "ground zero" for

328   contamination from open defecation and animal feces. The most abundant types of pathogen taxa

329   varied for soil and surface water, with protozoans being most abundant in soil and pathogenic bacteria

330   being most abundant in surface water. Domestic animals (specifically chickens, cows, and goats/sheep)

331   – rather than human sources of fecal waste – were associated with increased pathogen diversity,

332   whereas enterococcus bacteria, the presence of any pathogen type, and concentrations of individual

333   enteric pathogen taxa were not significantly associated with human and animal sources of fecal

334   contamination. Finally, this study validated that <5yr old children play in these contaminated public

335   settings, confirming that child exposure to and infection by pathogens in public areas is plausible.

336   Therefore, interventions that prevent neighborhood-level animal fecal contamination may be necessary

337   for reducing enteric disease burden in < 5yr children in Kisumu. These interventions could include

15

338   creation of protected child play spaces with finished, cleanable floors, promoting the safe disposal of

339   animal waste, or animal penning designs that safely collect animal waste away from children.

340   These findings confirmed our hypothesis that pathogen diversity in Kisumu neighborhoods

341   would mirror the leading etiologic causes of symptomatic and asymptomatic childhood infections in

342   low-income, high-burden endemic countries in Asia and Africa [1, 2, 28]. Many of the most common causes

343   of moderate to severe diarrhea in children in Kenya *(Cryptosporidium spp.*, ETEC ST, adenovirus 40/41,

344   *C. jejuni*)[1, 31] were the most frequently detected in the environment in Kisumu, although Rotavirus and

345   Shigella/EIEC were rare. This suggests that widespread contamination by those pathogens in the

346   environment is responsible for causing enteric infection in children. Our findings also confirmed our

347   hypothesis that multi-pathogen contamination of soil and water would be detected at locations where

348   children play. This further suggests that children could ingest multiple pathogens during play in just one

349   area outside the household. Pathogen diversity is a novel approach for characterizing environmental

350   exposure risks in LICs, and for identifying potential fecal sources associated with environmental

351   contamination. However, microbial community characterization has been used for decades to

352   understand ecological health and to monitor the impact of natural or man-made actions on system

353   function [32-34]. In public health, microbiome community characterization has become a platform for

354   understanding human susceptibility to enteric infection, symptomology and severity of disease, and for

355   understanding vaccine failure [35, 36]. A few studies have used the microbial community approach to

356   understand the relationships between environmental microbiota, fecal indicator bacteria and

357   pathogenic bacteria in surface water and soil, although did not examine infrastructure conditions driving

358   this contamination [37, 38]. Rather than sequencing, we used customized multi-pathogen qPCR tools for

359   disease-targeted detection of the viral, bacterial, protozoan, and helminthic causes of pediatric diarrhea

360   [28]. Additionally, we accounted for potential interrelated conditions underlying pathogen co-occurrence

361   in our analysis, rather than treating pathogen occurrence as an isolated, independent event.

16

362   The inclusion of pathogen diversity as indicator of interrelated contamination conditions was

363   important for understanding the relative intensity and determinants of contamination of public areas

364   across Kisumu, as well as between Kenya and a reference site in Iowa. Consistent with prior studies,

365   enterococcus were omnipresent in Kenya *and* Iowa soil and surface water, even though evidence of

366   human fecal contamination was absent in Iowa, and animal feces was less common [15]. Evidence on fecal

367   microbes in soil is limited, but our detection rates were similar to at least one other study [39]. The lack of

368   association between human and animal fecal source risk factors with enterococci in Kenya reinforces

369   that enterococcus indicators are not optimal tools for predicting or quantifying exposure to enteric

370   pathogens or for fecal source tracking. An "any pathogen" indicator was also not statistically associated

371   with fecal sources. However, the stark differences in overall pathogen detection frequencies between

372   Kenya and an Iowa watershed impacted by farm and agriculture helped calibrate our expectations as to

373   what contamination patterns should look like in a setting where open defecation is absent, and

374   highlighted how alarming the pathogen detection levels in Kenya are by comparison. Sampling at

375   different times, for example different seasons, or sampling from agriculturally-impacted areas in a

376   tropical state (e.g. Florida), may have produced different results.

377   Neighborhood-level pathogen contamination in Kenya corresponded with the frequent

378   observation of human and animal feces in public areas. These neighborhoods have low levels of

379   household sanitation coverage and domestic animal ownership is common in Kisumu [22, 23]. We expected

380   open defecation, dilapidated latrines, and domestic animals to be associated with pathogen diversity,

381   given the evidence of their role in diarrheal in children [40, 41]. In spite of widespread human feces in

382   Kisumu, only domestic animals were associated with increased diversity in pathogens in the

383   environment. Domestic animal reservoirs may contribute to pediatric disease burden by facilitating a

384   rapid turnover in the types of pathogen species that children are exposed to, a situation where

385   exposure-based immunity provides little benefit. We are unaware of evidence of this hypothesis, but

17

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

18

410   for linking widespread contamination to specific fecal sources, compared to the relatively simple

411   approach of simply summing the number of pathogen types detected. In the handful of repeat patterns

412   noted, the pathogens detected were the most common overall, suggesting presence is more a function

413   of probability and sample size than impacts from types of fecal sources. In particular, the most common

414   pathogen gene detected was *Cryptosporidium spp*. 18S, which could indicate contamination by species

415   rarely detected in humans (*C. meleagridis, C. canis, C. felis, C. muris, C. suis*), rather than the *C. hominus*

416   or *C. parvum* species commonly linked with human infection[30]. This may overestimate the abundance of

417   clinically-relevant pathogens in public areas of Kisumu, although even after excluding Cryptosporidium-

418   positive counts, pathogen detection and co-detection was still very high. Research involving larger

419   numbers of environmental samples is needed to understand whether co-occurrence patterns of specific

420   groups of pathogens is common. Either way, the substantial amount of heterogeneity in pathogen

421   detection patterns highlights the risk for misclassification of environmental exposures if the presence of

422   an individual pathogen type is used as a proxy for the presence of any fecal pathogens, or fecal

423   contamination overall. While some associations with trends towards significance may surpass

424   significance cutoff thresholds with larger sample sizes, we are confident that it wouldn't change the

425   overall conclusion that domestic animals are important contributors to multi-pathogen contamination in

426   public domains of Kisumu.

427   Last, detection of pathogen DNA or RNA in soil and water by PCR does not confirm viability or

428   infectivity of pathogens. PCR methods may detect extracellular DNA or intact, but non-infectious cells.

429   By filtering sample elute through a 0.45 μm pore size filter, our sampling methods may have removed

430   some extracellular material, although non-viable cellular DNA or RNA could still be trapped on the filter.

431   PCR was chosen as the most systematic way to detect viral, bacterial, protozoan and helminth classes of

432   fecal pathogens, many of which have no alternative detection methods. Culture-based assays could

433   have been used for some bacteria, but may have been equally inaccurate due to the presence of viable

434   but non-recoverable bacteria (e.g. *Vibrio cholerae, Shigella dysenteriae*, *Camylobacter jejuni*, and

435   Enterotoxigenic *E. coli*) [17, 44-46]. Since we did not adopt a pre-enrichment step, it is possible that our

436   enterococcus concentration data underestimates actual live concentration of enterococcus in soil or

437   water. This also would introduce inter-method variability in detection methods across viral, bacterial,

438   protozoan, and helminth taxa. Adjusting concentrations for pathogen decay under these ecological

439   conditions may improve the accuracy of estimated PCR concentrations. However, there is no such

440   information for many enteric pathogens included on our assay, so such an approach could not be

441   systematic. Also problematic is that feces contamination by animals and humans in Kisumu is ongoing.

442   Failing to counter-adjust pathogen concentration estimates for periodic reintroduction of pathogens in

443   the environment would underestimate final concentration estimates. While some detected pathogen

444   DNA detected in Kisumu may be non-viable, it is more likely to be relatively recent contamination

445   because free and cellular DNA/RNA is typically degraded by native soil or water microbiota within the

446   matter of days or weeks [47-52].

447   In conclusion, this study addresses two neglected realms in the WASH sector. First, we

448   demonstrate that public domains are highly contaminated and may pose a substantial risk of enteric

449   pathogen exposure for children. Young children will continue to play in soil and surface water in the

450   public domain because they live in crowded conditions and at a certain age are drawn to engage in

451   social play with their peers. In light of this, WASH interventions should invest in improving public

452   sanitary conditions to prevent child exposure to enteric pathogens, and possibly to reduce the childhood

453   enteric disease burden that persists in many low and middle income countries. Second, even if high

454   levels of household WASH coverage are achieved, and human open defecation is eliminated, a high

455   baseline level of enteric disease caused by contact with animal feces will persist. Domestic animal

456   management must be included in the WASH agenda to reduce pathogen contamination at the

457   neighborhood level, and potentially within households as well. Forthcoming research by this group is

458   examining the generalizability of these findings in other geographical settings using improved

459   assessment methods. Yet, the conditions in Kisumu – low-income urban settlements with high

460   population density and limited public health infrastructure – are common in low-income settlements

461   around the world, suggesting pervasive pathogen contamination of the environment in such settings is a

462   universal problem.

463

464   **Supplementary Documents Table of Contents**

477

489

CONTRIBUTIONS

491   Conceived of this study: KKB, ASG. Acquired funding: KKB, ASG. Designed study or study tools: KKB, RS,

492   OC, JM. Collected data: KKB, RS, JM. Statistical analysis: DS, RS, KKB. Drafted manuscript: KKB. Edited

493   manuscript for intellectual content: all authors. Provided final approval of the version to be published:

494   all co-authors.

495

REFERENCES

497   1.   Kotloff, K. L.; Nataro, J. P.; Blackwelder, W. C.; Nasrin, D.; Farag, T. H.; Panchalingam, S.; Wu, Y.;
498   Sow, S. O.; Sur, D.; Breiman, R. F.; Faruque, A. S.; Zaidi, A. K.; Saha, D.; Alonso, P. L.; Tamboura, B.;
499   Sanogo, D.; Onwuchekwa, U.; Manna, B.; Ramamurthy, T.; Kanungo, S.; Ochieng, J. B.; Omore, R.;
500   Oundo, J. O.; Hossain, A.; Das, S. K.; Ahmed, S.; Qureshi, S.; Quadri, F.; Adegbola, R. A.; Antonio, M.;
501   Hossain, M. J.; Akinsola, A.; Mandomando, I.; Nhampossa, T.; Acacio, S.; Biswas, K.; O'Reilly, C. E.; Mintz,
502   E. D.; Berkeley, L. Y.; Muhsen, K.; Sommerfelt, H.; Robins-Browne, R. M.; Levine, M. M., Burden and
503   aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric
504   Multicenter Study, GEMS): a prospective, case-control study. *Lancet* **2013,** *382*, (9888), 209-22.
505   2.   Platts-Mills, J. A.; Babji, S.; Bodhidatta, L.; Gratz, J.; Haque, R.; Havt, A.; McCormick, B. J.;
506   McGrath, M.; Olortegui, M. P.; Samie, A.; Shakoor, S.; Mondal, D.; Lima, I. F.; Hariraju, D.; Rayamajhi, B.
507   B.; Qureshi, S.; Kabir, F.; Yori, P. P.; Mufamadi, B.; Amour, C.; Carreon, J. D.; Richard, S. A.; Lang, D.;

508 Bessong, P.; Mduma, E.; Ahmed, T.; Lima, A. A.; Mason, C. J.; Zaidi, A. K.; Bhutta, Z. A.; Kosek, M.;
509 Guerrant, R. L.; Gottlieb, M.; Miller, M.; Kang, G.; Houpt, E. R.; Investigators, M.-E. N., Pathogen-specific
510 burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED).
511 *Lancet Glob Health* **2015,** *3*, (9), e564-75.
512 3.   Pickering, A. J.; Julian, T. R.; Marks, S. J.; Mattioli, M. C.; Boehm, A. B.; Schwab, K. J.; Davis, J.,
513 Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian households with
514 and without improved sanitation. *Environ Sci Technol* **2012,** *46*, (11), 5736-43.
515 4.   Mattioli, M. C.; Pickering, A. J.; Gilsdorf, R. J.; Davis, J.; Boehm, A. B., Hands and water as vectors
516 of diarrheal pathogens in Bagamoyo, Tanzania. *Environmental science & technology* **2013,** *47*, (1), 355-
517 63.
518 5.   Odagiri, M.; Schriewer, A.; Daniels, M. E.; Wuertz, S.; Smith, W. A.; Clasen, T.; Schmidt, W. P.; Jin,
519 Y.; Torondel, B.; Misra, P. R.; Panigrahi, P.; Jenkins, M. W., Human fecal and pathogen exposure
520 pathways in rural Indian villages and the effect of increased latrine coverage. *Water Res* **2016,** *100*, 232-
521 44.
522 6.   Humphrey, J. H.; Prendergast, A. J.; Ntozini, R.; Gladstone, M.; Colford, J. In *The Sanitation
523 Hygiene Infant Nutrition Efficacy (SHINE) Trial*, American Society for Tropical Medicine and Hygiene
524 Annual Meeting, Baltimore, MD, November 7, 2017, 2017; American Society for Tropical Medicine and
525 Hygiene: Baltimore, MD, 2017.
526 7.   Null, C.; Stewart, C. P.; Pickering, A. J.; Dentz, H. N.; Arnold, B. F.; Arnold, C. D.; Benjamin-Chung,
527 J.; Clasen, T.; Dewey, K. G.; Fernald, L. C. H.; Hubbard, A. E.; Kariger, P.; Lin, A.; Luby, S. P.; Mertens, A.;
528 Njenga, S. M.; Nyambane, G.; Ram, P. K.; Colford, J. M., Jr., Effects of water quality, sanitation,
529 handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-
530 randomised controlled trial. *Lancet Glob Health* **2018,** *6*, (3), e316-e329.
531 8.   Luby, S. P.; Rahman, M.; Arnold, B. F.; Unicomb, L.; Ashraf, S.; Winch, P. J.; Stewart, C. P.; Begum,
532 F.; Hussain, F.; Benjamin-Chung, J.; Leontsini, E.; Naser, A. M.; Parvez, S. M.; Hubbard, A. E.; Lin, A.;
533 Nizame, F. A.; Jannat, K.; Ercumen, A.; Ram, P. K.; Das, K. K.; Abedin, J.; Clasen, T. F.; Dewey, K. G.;
534 Fernald, L. C.; Null, C.; Ahmed, T.; Colford, J. M., Jr., Effects of water quality, sanitation, handwashing,
535 and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised
536 controlled trial. *Lancet Glob Health* **2018,** *6*, (3), e302-e315.
537 9.   Medgyesi, D. N.; Brogan, J. M.; Sewell, D. K.; Creve-Coeur, J. P.; Kwong, L. H.; Baker, K. K., Where
538 Children Play: Young child exposure to environmental hazards during play in public areas in a
539 transitioning internally displaced persons community in Haiti. *International Journal of Environmental
540 Research and Public Health* **2018,** *15*, (8), 1682.
541 10.   Harwood, V. J.; Levine, A. D.; Scott, T. M.; Chivukula, V.; Lukasik, J.; Farrah, S. R.; Rose, J. B.,
542 Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health
543 protection. *Appl Environ Microbiol* **2005,** *71*, (6), 3163-70.
544 11.   Ferguson, A. S.; Layton, A. C.; Mailloux, B. J.; Culligan, P. J.; Williams, D. E.; Smartt, A. E.; Sayler,
545 G. S.; Feighery, J.; McKay, L. D.; Knappett, P. S.; Alexandrova, E.; Arbit, T.; Emch, M.; Escamilla, V.;
546 Ahmed, K. M.; Alam, M. J.; Streatfield, P. K.; Yunus, M.; van Geen, A., Comparison of fecal indicators with
547 pathogenic bacteria and rotavirus in groundwater. *Sci Total Environ* **2012,** *431*, 314-22.
548 12.   Haramoto, E.; Katayama, H.; Oguma, K.; Yamashita, H.; Tajima, A.; Nakajima, H.; Ohgaki, S.,
549 Seasonal profiles of human noroviruses and indicator bacteria in a wastewater treatment plant in Tokyo,
550 Japan. *Water Sci Technol* **2006,** *54*, (11-12), 301-8.
551 13.   Verbyla, M. E.; Iriarte, M. M.; Mercado Guzman, A.; Coronado, O.; Almanza, M.; Mihelcic, J. R.,
552 Pathogens and fecal indicators in waste stabilization pond systems with direct reuse for irrigation: Fate
553 and transport in water, soil and crops. *Sci Total Environ* **2016,** *551-552*, 429-37.
554 14.   Desmarais, T. R.; Solo-Gabriele, H. M.; Palmer, C. J., Influence of soil on fecal indicator organisms
555 in a tidally influenced subtropical environment. *Appl Environ Microbiol* **2002,** *68*, (3), 1165-72.

556   15.   Byappanahalli, M. N.; Nevers, M. B.; Korajkic, A.; Staley, Z. R.; Harwood, V. J., Enterococci in the
557   environment. *Microbiol Mol Biol Rev* **2012,** *76*, (4), 685-706.
558   16.   Islam, M. S.; Talukder, K. A.; Khan, N. H.; Mahmud, Z. H.; Rahman, M. Z.; Nair, G. B.; Siddique, A.
559   K.; Yunus, M.; Sack, D. A.; Sack, R. B.; Huq, A.; Colwell, R. R., Variation of toxigenic Vibrio cholerae O1 in
560   the aquatic environment of Bangladesh and its correlation with the clinical strains. *Microbiol Immunol*
561   **2004,** *48*, (10), 773-7.
562   17.   Alam, M.; Sultana, M.; Nair, G. B.; Siddique, A. K.; Hasan, N. A.; Sack, R. B.; Sack, D. A.; Ahmed, K.
563   U.; Sadique, A.; Watanabe, H.; Grim, C. J.; Huq, A.; Colwell, R. R., Viable but nonculturable Vibrio
564   cholerae O1 in biofilms in the aquatic environment and their role in cholera transmission. *Proc Natl Acad*
565   *Sci U S A* **2007,** *104*, (45), 17801-6.
566   18.   KNBS *The 2009 Kenya Population and Housing Census*; Kenyan National Bureau of Statistics:
567   Kenya, 2010.
568   19.   Habitat, U., Situation analysis of informal settlements in Kisumu. *Cities Without Slums Sub-*
569   *Regional Programme for Eastern and Southern Africa. Kenya Slum Upgrading Programme* **2005**.
570   20.   Kenya Slum Update Programme; UN Human Settlements Programme *Situational Analysis of*
571   *Informal Settlements in Kisumu*; Nairobi, Kenya, 2005.
572   21.   Statistics, K. N. B. o. *Nyanza Province Multiple Indicator Cluster Survey 2011*; Kenyan National
573   Bureau of Statistics: Nairobi, Kenya, 2013.
574   22.   Barnes, A. N.; Mumma, J.; Cumming, O., Role, ownership and presence of domestic animals in
575   peri-urban households of Kisumu, Kenya. *Zoonoses Public Health* **2018,** *65*, (1), 202-214.
576   23.   Simiyu, S.; Swilling, M.; Cairncross, S.; Rheingans, R., Determinants of quality of shared
577   sanitation facilities in informal settlements: case study of Kisumu, Kenya. *BMC Public Health* **2017,** *17*,
578   (1), 68.
579   24.   EPA)., U. S. E. P. A. U. S. Recreational water quality criteria.
580   http://water.epa.gov/scitech/swguidance/standards/criteria/health/recreation/index.cfm (1/26/2018),
581   25.   Griffith, J. F.; Cao, Y.; McGee, C. D.; Weisberg, S. B., Evaluation of rapid methods and novel
582   indicators for assessing microbiological beach water quality. *Water Res* **2009,** *43*, (19), 4900-7.
583   26.   Ahmed, W.; Harwood, V. J.; Gyawali, P.; Sidhu, J. P.; Toze, S., Comparison of concentration
584   methods for quantitative detection of sewage-associated viral markers in environmental waters. *Appl*
585   *Environ Microbiol* **2015,** *81*, (6), 2042-9.
586   27.   Haramoto, E.; Katayama, H.; Asami, M.; Akiba, M., Development of a novel method for
587   simultaneous concentration of viruses and protozoa from a single water sample. *J Virol Methods* **2012,**
588   *182*, (1-2), 62-9.
589   28.   Liu, J.; Gratz, J.; Amour, C.; Kibiki, G.; Becker, S.; Janaki, L.; Verweij, J. J.; Taniuchi, M.; Sobuz, S.
590   U.; Haque, R.; Haverstick, D. M.; Houpt, E. R., A laboratory-developed TaqMan Array Card for
591   simultaneous detection of 19 enteropathogens. *J Clin Microbiol* **2013,** *51*, (2), 472-80.
592   29.   Ishii, S.; Kitamura, G.; Segawa, T.; Kobayashi, A.; Miura, T.; Sano, D.; Okabe, S., Microfluidic
593   quantitative PCR for simultaneous quantification of multiple viruses in environmental water samples.
594   *Appl Environ Microbiol* **2014,** *80*, (24), 7505-11.
595   30.   Stroup, S. E.; Roy, S.; McHele, J.; Maro, V.; Ntabaguzi, S.; Siddique, A.; Kang, G.; Guerrant, R. L.;
596   Kirkpatrick, B. D.; Fayer, R.; Herbein, J.; Ward, H.; Haque, R.; Houpt, E. R., Real-time PCR detection and
597   speciation of Cryptosporidium infection using Scorpion probes. *Journal of medical microbiology* **2006,**
598   *55*, (Pt 9), 1217-22.
599   31.   Liu, J.; Platts-Mills, J. A.; Juma, J.; Kabir, F.; Nkeze, J.; Okoi, C.; Operario, D. J.; Uddin, J.; Ahmed,
600   S.; Alonso, P. L.; Antonio, M.; Becker, S. M.; Blackwelder, W. C.; Breiman, R. F.; Faruque, A. S.; Fields, B.;
601   Gratz, J.; Haque, R.; Hossain, A.; Hossain, M. J.; Jarju, S.; Qamar, F.; Iqbal, N. T.; Kwambana, B.;
602   Mandomando, I.; McMurry, T. L.; Ochieng, C.; Ochieng, J. B.; Ochieng, M.; Onyango, C.; Panchalingam,
603   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 (%)	40 (72.7%)	39 (73.6%)	45 (77.6%)	124 (74.7%)
Landscape with open lots/fields, n (%)	15 (27.3%)	14 (26.4%)	13 (22.4%)	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 (%)	41 (75.6%)	42 (79.3%)	45 (77.6%)	128 (77.1%)
- Pig	2 (3.6%)	2 (3.8%)	2 (3.5%)	6 (3.6%)
- Chicken	22 (39.3%)	23 (43.4%)	34 (58.6%)	79 (47.3%)
- Cattle	8 (14.3%)	11 (20.8%)	12 (20.7%)	31 (18.6%)
- Dog	9 (16.1%)	13 (24.5%)	6 (10.3%)	28 (16.8%)
- Sheep/Goat	5 (8.9%)	12 (22.6%)	13 (22.4%)	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 (%)	45 (81.8%)	44 (83%)	39 (67.2%)	130 (78.3%)
- Functional Condition	0 (0%)	9 (17.0%)	0 (0%)	9 (5.4%)

27

- 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.
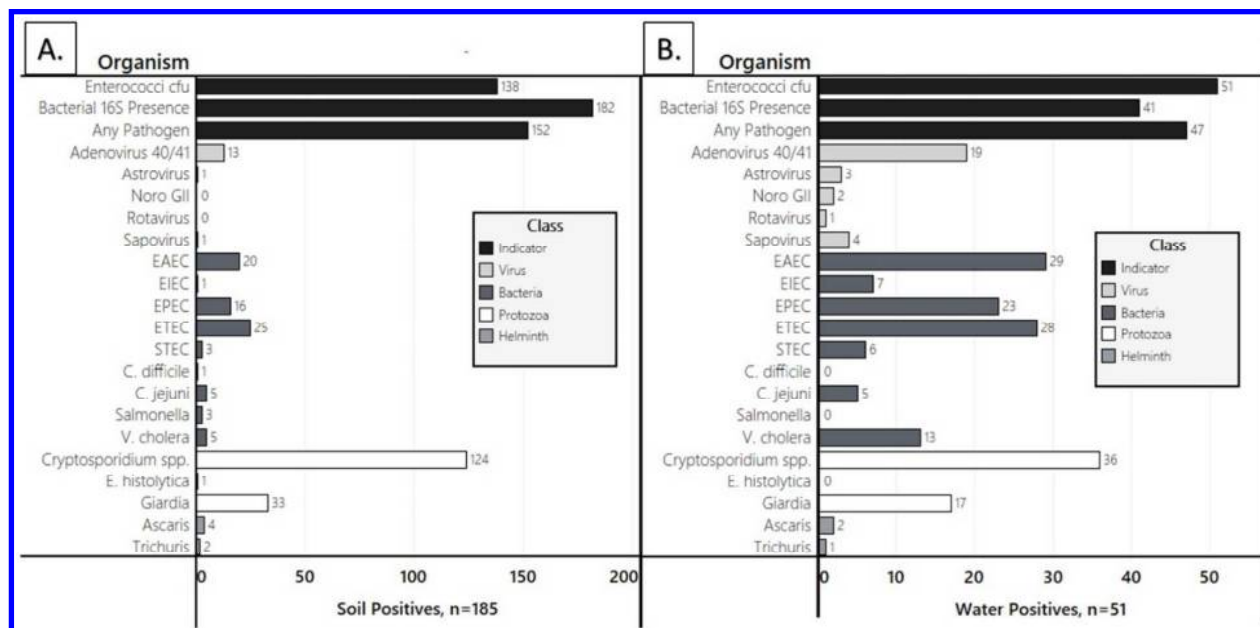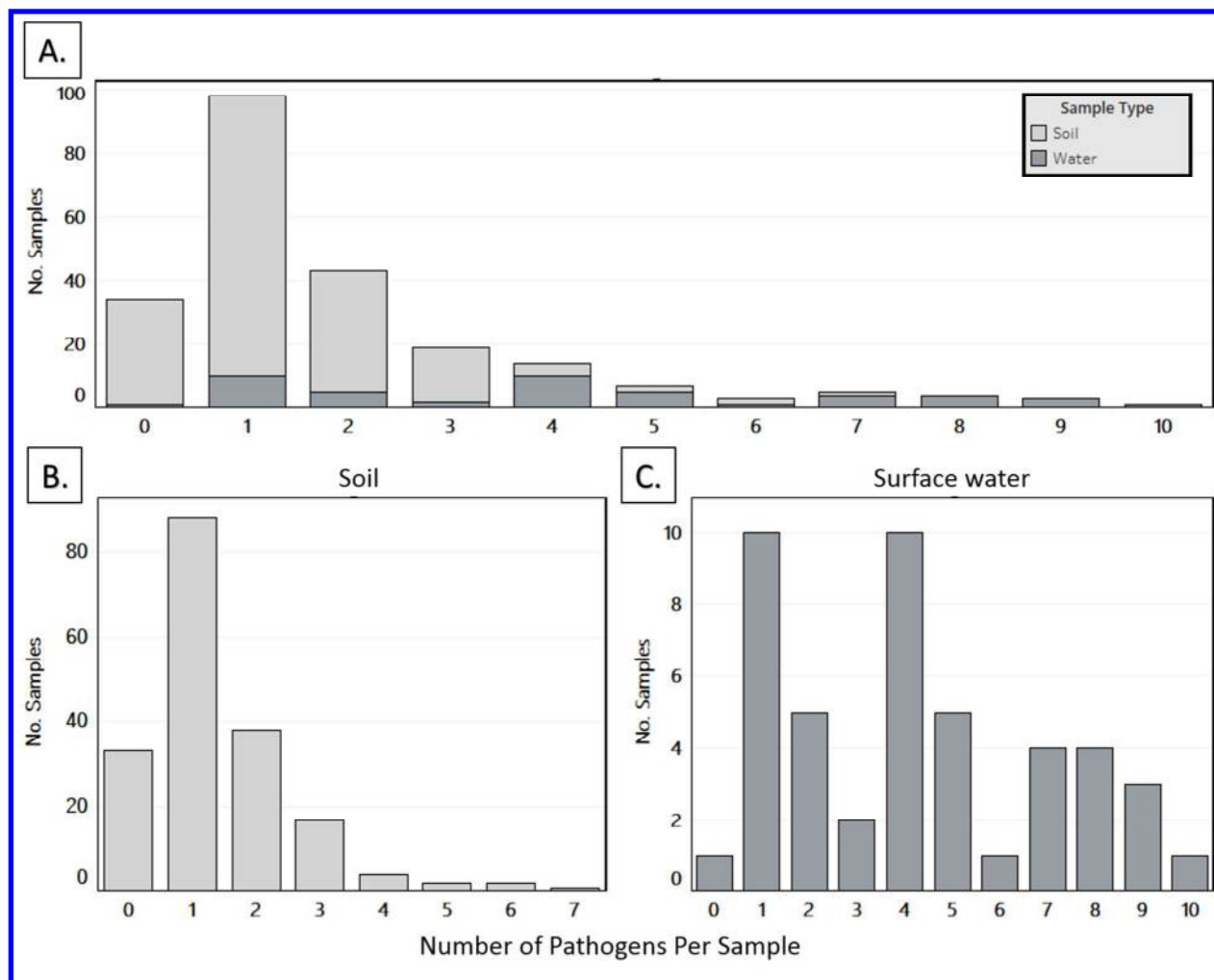

671

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.
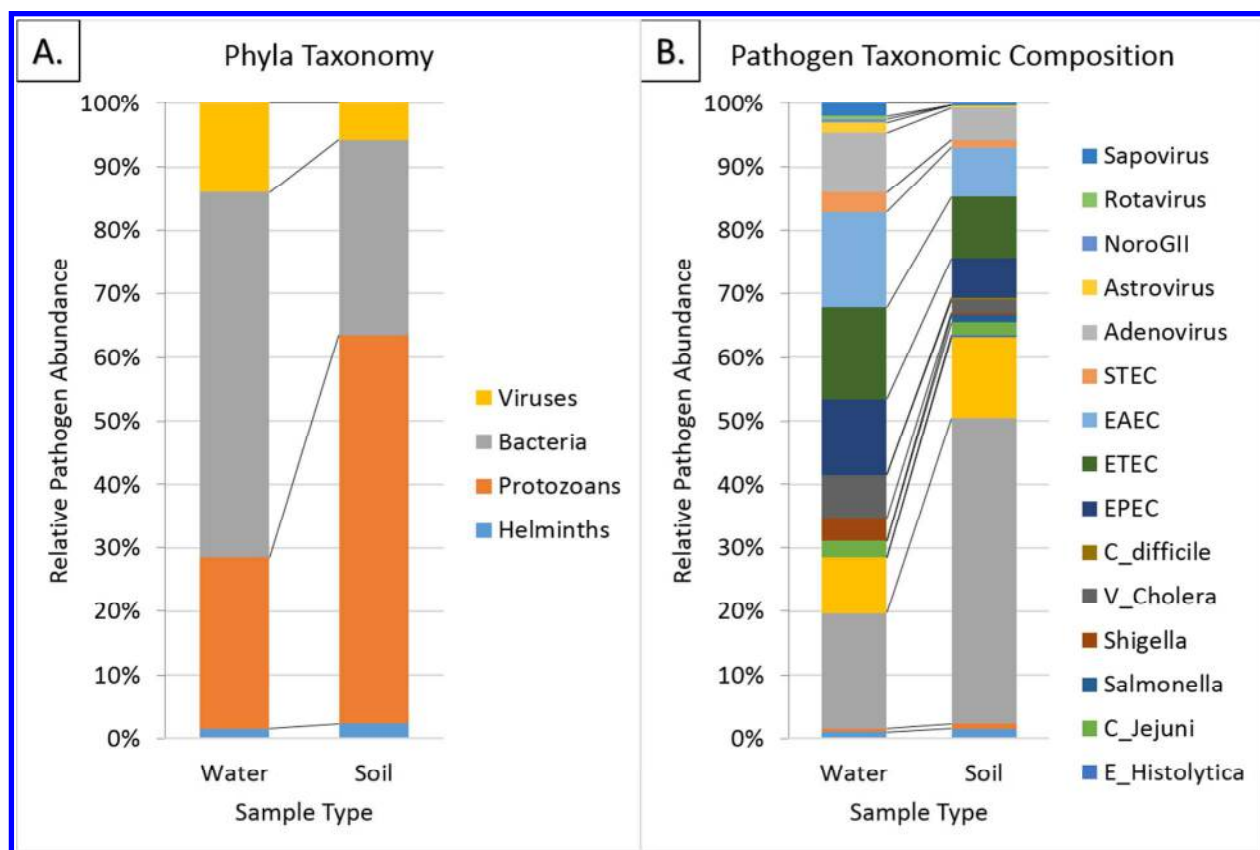
674

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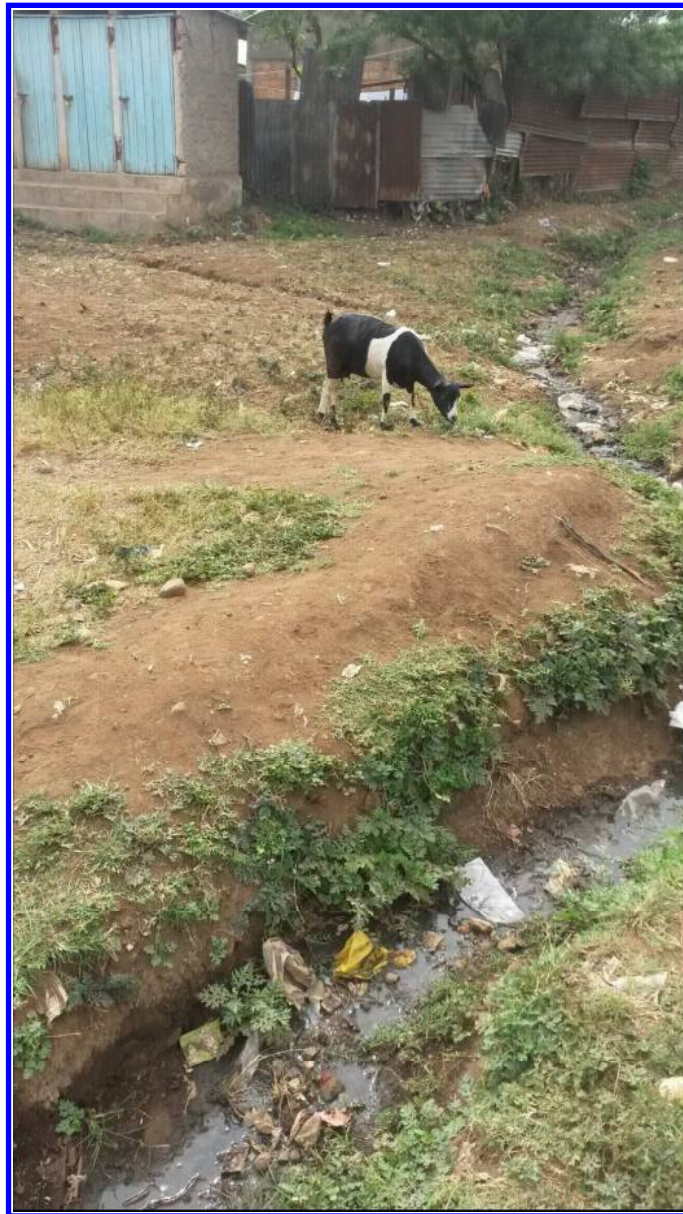
Example of a public site in Kisumu impacted by multiple potential sources of enteric pathogens: a communal latrine, a domestic goat, and an informal drain used for water management and and waste disposal.

285x508mm (96 x 96 DPI)