

## Soil Ingestion Is Associated with Child Diarrhea in an Urban Slum of Nairobi, Kenya

Valerie Bauza,<sup>1\*</sup> R. M. Ocharo,<sup>2</sup> Thanh H. Nguyen,<sup>1</sup> and Jeremy S. Guest<sup>1\*</sup>

<sup>1</sup>Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois;

<sup>2</sup>Department of Sociology and Social Work, University of Nairobi, Nairobi, Kenya

**Abstract.** Diarrhea is a leading cause of mortality in children under 5 years of age. We conducted a cross-sectional study of 54 children aged 3 months to 5 years old in Kibera, an urban slum in Nairobi, Kenya, to assess the relationship between caregiver-reported soil ingestion and child diarrhea. Diarrhea was significantly associated with soil ingestion (adjusted odds ratio = 9.9, 95% confidence interval = 2.1–47.5). Soil samples from locations near each household were also collected and analyzed for *Escherichia coli* and a human-associated *Bacteroides* fecal marker (HF183). *Escherichia coli* was detected in 100% of soil samples (mean 5.5 log colony forming units *E. coli* per gram of dry soil) and the *Bacteroides* fecal marker HF183 was detected in 93% of soil samples. These findings suggest that soil ingestion may be an important transmission pathway for diarrheal disease in urban slum settings.

### INTRODUCTION

Diarrheal disease caused by fecal pathogens is the second leading cause of mortality in children under 5 years of age, responsible for approximately 760,000 deaths annually.<sup>1</sup> Fecal pathogens are primarily transmitted through the fecal-oral pathway with commonly reported exposure points of water, fields/floor, hands, food, and flies.<sup>2</sup> However, soil ingestion due to exploratory mouthing behavior of young children is a potentially important exposure pathway that has not been well studied or characterized.

Geophagy is the intentional ingestion of soil, and typically involves specific types of soil that are selected, prepared, and then eaten.<sup>3</sup> This behavior is practiced by both children and adults and has been found to be associated with micronutrient deficiency, anemia, enteric distress, soil-transmitted helminth infections, and hunger.<sup>3–8</sup> Nonetheless, the direction of the causal relationship for many of these associations and whether geophagy serves as a risky or protective behavior is still debated.<sup>5,7</sup> Beyond the learned behaviors of geophagy, soil ingestion is also an involuntary behavior as part of mouthing and exploration in young children. There are few studies related to exploratory soil ingestion by young children in low-income countries, but recent evidence<sup>9,10</sup> suggests that it may be an important exposure point for fecal contamination leading to child illness.

A recent study in rural Zimbabwe found that the majority of household soil samples were contaminated with the fecal indicator bacteria *Escherichia coli* and 13% of infants included in the study were observed to actively ingest this contaminated soil during a 6-hour structured observation period.<sup>9</sup> Another study in peri-urban Tanzania found *E. coli* pathotype genes, a human-specific *Bacteroidales* gene, and enteric viruses present in a subset of soil samples taken from household locations, demonstrating that household soils may be contaminated with fecal pathogens.<sup>11</sup> Soil ingestion by young children has also been linked to

negative health outcomes and was found to be associated with increased risk of diarrhea in rural Kenya,<sup>12</sup> and with environmental enteropathy and stunting in rural Bangladesh.<sup>10</sup> However, these studies linking soil ingestion with negative health outcomes for young children were conducted in rural areas and do not provide a comprehensive understanding of the consequences of soil ingestion. Little is known about the prevalence and effects of exploratory soil ingestion in young children in urban areas. Studying low-income urban areas is important because the higher population density leads to increased foot traffic on soil close to households, the types and density of sanitation facilities vary relative to rural areas, and the drainage infrastructure is often poor in urban slums. These factors may impact the quantity and frequency of fecal pathogens released to the environment, affecting soil contamination. For example, in the previous study in rural Kenya that demonstrated a link between soil ingestion and diarrhea, 40% of households included in the study did not have access to any latrine,<sup>12</sup> whereas open defecation by adults is rare in urban areas (practiced by 3% of urban households in Kenya in 2015).<sup>13</sup>

The objective of this study was to assess the relationship between caregiver-reported soil ingestion events and diarrhea episodes in children under 5 years of age residing in Kibera, an urban slum of Nairobi, Kenya. To better understand factors influencing this relationship, fecal contamination levels of soil samples collected near each household were also measured by enumerating *E. coli* via membrane filtration and quantifying a human-associated *Bacteroides* fecal marker using quantitative polymerase chain reaction (qPCR). To our knowledge, this is the first time that the link between soil ingestion and diarrhea has been studied in urban areas and in households with primarily non-earth floors, as well as the first time a human-associated fecal marker has been measured in soil in this setting to determine whether human fecal contamination may be a significant contamination source of soil in urban slum areas.

### METHODS

**Study site and sampling frame.** This study was conducted in June 2015 in the Makina, Sarangombe, and Lindi wards of Kibera, the largest urban slum in Nairobi, Kenya. Purposeful sampling was used to select slum compounds (clusters of households with shared common areas and

\*Address correspondence to Valerie Bauza, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 4146 Newmark Civil Engineering Laboratory, 205 North Mathews Avenue, Urbana, IL 61801-2352, E-mail: vbauza@illinois.edu or Jeremy S. Guest, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 3221 Newmark Civil Engineering Laboratory, 205 North Mathews Avenue, Urbana, IL 61801-2352, E-mail: jsguest@illinois.edu.

often shared toilets and water sources) that were in different wards and different areas within a ward to increase the variation in sanitation, drainage, and solid waste infrastructure for compounds included in this study. Within each compound, households with children under 5 years of age were then randomly selected for study inclusion, approached for informed consent, and household interviews were conducted with the primary caregiver to obtain information about household demographics, behaviors, and child health information. A total of 54 children (aged 3 months to < 5 years) from 40 households were included in this study from 16 different compounds. The results presented here are part of a larger study of infrastructure and practices related to water, sanitation, and hygiene (WASH) in Kibera.

**Ethical approval.** Informed consent was obtained from the household's primary caregiver before enrollment. The study was approved by the Institutional Review Board of the University of Illinois at Urbana-Champaign and the National Commission for Science, Technology and Innovation in Kenya.

**Household interviews.** A questionnaire on the household's sociodemographic characteristics, diarrhea illness occurring in the past week for all children in the household, and infrastructure and behaviors related to WASH was administered to each household's primary caregiver by a Kenyan interviewer in Kiswahili or Dholuo (depending on the respondent's primary language). Diarrhea was defined as having three or more bowel movements within 24 hours. For each child, caregivers were also asked if they had observed that child putting soil, mud, clay, or sand into his or her mouth in the past 7 days. This question was asked because caregiver reported soil ingestion/geophagy events have been previously found to show good agreement with soil ingestion events observed during structured observation in Bangladesh.<sup>10</sup>

**Soil sampling.** Soil samples were collected from outdoor locations near each household entrance, for a total of 28 samples. Since a number of households had shared outdoor common space that the household entrance opened into, this provided soil samples from locations near 34 households. Soil was sampled from outside each household because only one household included in the study had an earth floor and the remaining households had a hard floor (consisting of either vinyl, concrete, wood, or tile). To be consistent among households, soil was not collected for households surrounded by concrete walkways ( $N = 6$ ), because there was no soil near the household entrance with which a child would be likely to come in frequent contact with and consume. Soil samples were collected using sterile polystyrene sampling spoons (Nasco, Fort Atkinson, WI) such that approximately 10 g of soil were collected from the surface of soil (from an area approximately 10 × 10 cm) and placed in a sterile Whirl-Pak bag (Nasco) for transport to the laboratory in a cooler with ice packs. At the time of sampling, ambient temperature ranged from 21°C to 29°C and relative humidity ranged from 41% to 76%. All samples were taken from areas that were either fully or partially shaded from sunlight.

***Escherichia coli* enumeration.** Samples were processed within 8 hours of sample collection. Each soil sample was homogenized inside of a sterile bag by hand. Two grams of soil were then handshaken for 2 minutes in 20 mL of phosphate-buffered saline, and allowed to settle for

30 seconds, following the methods recommended by Boehm and others<sup>14</sup> to recover *E. coli*. The supernatant was then poured into a sterile container, diluted, and filtered through 0.45- $\mu$ m pore size (47 mm diameter) mixed cellulose esters filters (Pall Corporation, Port Washington, NY). *Escherichia coli* was enumerated using m-ColiBlue24 Broth media (Hach, Loveland, CO), following the manufacturer's protocol approved by the U.S. Environmental Protection Agency. Plates were incubated at 35°C for 24 hours. Filtration blanks were processed daily. In addition, 5 g of soil was dried in an oven at 105°C for 24 hours to determine moisture content.

**Molecular analysis for human-specific *Bacteroides* detection.** DNA was extracted from approximately 0.25 g of soil using the PowerSoil DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, CA) following the manufacturer's guidelines. The DNA extract was stored at -20°C for up to 3 weeks, transported back to the University of Illinois in a cooler with ice packs, and stored at -80°C until further analysis.

Nucleic acid extracts were analyzed via qPCR for the human-associated *Bacteroides* fecal marker HF183, which has been previously validated in Kenya.<sup>15</sup> Each qPCR reaction for the HF183 assay was performed with a 15  $\mu$ L reaction mixture, containing 1 $\times$  final concentration of SYBR Green I dye master mix (Applied Biosystems, Waltham, MA) and 250 nM of forward and reverse primers. Previously published primers that have been verified in Kenya were used.<sup>16</sup> Samples were amplified in a 384-well plate on an Applied Biosystems 7900HT Fast Real-Time PCR System with the following thermocycle conditions: 2 minutes at 50°C, 10 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C, 45 seconds at 53°C, and 1 minute at 60°C. A dissociation curve analysis was performed with conditions set to 15 seconds at 95°C, 20 seconds at 60°C, and 15 seconds at 95°C to determine the melting temperature of amplified sequences in each sample. Seven 10-fold serial dilutions ( $3 \times 10^0$  to  $3 \times 10^6$  genes copies/ $\mu$ L) of the synthesized target DNA sequence were used to calculate amplification efficiency and gene copies per microliter DNA extract from quantification cycle threshold ( $C_q$ ) values. Standard curve dilutions were run in triplicate, environmental samples were run in duplicate, and three no-template controls were run on each plate. The target HF183 sequence was considered detected if amplification was observed in the sample well and the melting temperature was between 73°C and 76°C. The lower limit of quantification (LLOQ) was defined as the lowest concentration at which more than 50% of standard curve replicates were amplified.<sup>17</sup> Positive samples were considered within the range of quantification (ROQ) if the  $C_q$  was above the LLOQ (3 gene copies/ $\mu$ L,  $C_q < 32$ ). The remaining samples that detected the target HF183 sequence were classified as detected, but not quantified (DNQ). The average LLOQ was 1,463 gene copies per gram of dry soil. Assuming a theoretical minimum detection of three copies per PCR reaction,<sup>18</sup> the average limit of detection (LOD) is 732 gene copies per gram of dry soil. The LOD was assigned to samples classified as DNQ and half the LOD was assigned to samples with the target not detected (ND) for quantitative analysis.

PCR inhibition was assessed by examining the linearity of response in the  $C_q$  values across two 5-fold dilutions of

DNA extract, following the method in Cao and others.<sup>19</sup> For samples classified as quantifiable ( $C_q < 32$ ), 5-fold dilutions were completed using each sample's DNA extract. For samples that were classified as DNQ or ND, a subset of these samples was evaluated for inhibition by spiking  $3 \times 10^3$  copies/ $\mu\text{L}$  into the DNA extract before completing 5-fold dilutions to measure the linearity of response. Samples were considered inhibited if the  $\Delta C_q$  between 5-fold dilutions was less than 1.3, which is one cycle less than expected amplification assuming 100% efficiency, as recommended by Cao and others.<sup>19</sup>

**Statistical analysis.** Stata version 13.1 (StataCorp LP, College Station, TX) was used for all data analysis with the primary goal of determining if child soil ingestion was associated with child diarrhea in the past 7 days.  $\chi^2$ , Fischer's exact tests, and penalized maximum likelihood logistic regression (firthlogit command in Stata) were used to analyze associations between child diarrhea and child soil ingestion. Penalized maximum likelihood logistic regression was also used to control for potential confounding variables and obtain adjusted odds ratios (ORs). Penalized maximum likelihood logistic regression was used because this method addresses the problem of small sample size/sparse-data bias that can overestimate ORs for small sample sizes when using ordinary multivariable logistic regression.<sup>20,21</sup>

The following potential confounding variables were assessed for inclusion in the analysis: child's age, asset ownership index (to represent household income), number of household members, mother's educational attainment, drinking water contamination at time of visit, and whether the feces of all children under 5 years of age in the household are disposed of in a toilet or latrine. The drinking water contamination variable is a binary variable for whether any *E. coli* contamination was detected in a 100 mL sample of drinking water collected at the time of the household visit. The asset ownership index variable is an integer variable indicating the number of the following assets the household owns (or uses, in the case of electricity): bicycle, motorcycle, radio, television, computer, bed, wardrobe, electricity, and livestock. Variables for the multiple variable model were selected if their association with child diarrhea had a *P* value less than 0.2 in the binary or multiple variable model. The assets index and number of household member variables met these criteria and were included in adjusted models.

Two-sample *t* tests were used to test for statistical associations using the *E. coli* and HF183 measurements. Soil *E. coli* colony forming units (CFU) data and HF183 gene copy were log transformed before analysis to normalize their log normal distribution. For samples with *E. coli* counts that were too numerous to count ( $> 500$  CFU per filter), 550 CFU per filter volume was used.

In addition, to enable a comparison with a previous study in rural Kenya,<sup>12</sup> Pearson's correlation coefficient was calculated to test the association between caregiver-reported child diarrhea and caregiver-reported child soil ingestion.

## RESULTS

**Household characteristics.** The average household size was 5.25 (standard deviation [SD] = 1.9, range = 3–10) people, with an average of 1.4 (SD = 0.5, range = 1–3) children under 5 years of age. The majority of female heads of the household (55%,  $N = 22$ ) had completed primary education or fewer years of schooling. The average reported monthly household income of respondents was \$105 USD (SD = 65, range = 20–300). Seventeen households (42.5%) used a facility with a pour flush toilet to sewer, 20 households (50%) used a pour flush pit latrine or a pit latrine with slab, and three households (7.5%) used a pit latrine without slab. Ninety-five percent of these toilet facilities were shared with other households.

**Child soil ingestion and diarrhea.** Caregivers reported observing their child put soil into their mouth in the past 7 days for 44.4% of children ( $N = 24$ ) included in the study. Soil ingestion was reported most frequently for children in the 6- to 24-month age group, and the prevalence of reported soil ingestion decreased with increased age. It was reported least frequently for children aged 3–5 years, but was still reported for one-third of children in this age group. Diarrhea was reported for 24.1% of all children ( $N = 13$ ), with the highest levels among children aged 6–24 months and prevalence decreasing with age (Table 1).

Children who had been observed to ingest soil in past 7 days were significantly more likely to have had diarrhea in the past week compared with children who were not observed to ingest soil ( $\chi^2 = 11.2$ ,  $P = 0.001$ ). Soil ingestion and diarrhea were also significantly associated with each other (Pearson's  $r = 0.46$ ,  $P = 0.0005$ ). Penalized maximum likelihood logistic regression results (Table 2) revealed the odds of diarrhea were more than nine times higher for children who were observed to ingest soil in the past week compared with those who were not (adjusted OR = 9.9, 95% confidence interval [CI] = 2.1–47.5). This association was higher among children aged 6 months to  $< 36$  months, who were more likely to consume soil than other children (adjusted OR = 12.9, 95% CI = 1.9–88.5).

**Soil contamination.** *Escherichia coli* was detected in every soil sample, and samples had a mean of 5.5 log CFU *E. coli* (SD = 0.35) per gram of dry soil. The human-associated *Bacteroides* fecal marker HF183 was detected in 93% of samples ( $N = 26$ ). Overall, 36% of samples were quantifiable ( $N = 10$ ), 57% of samples were DNQ ( $N = 16$ ), and 7%

TABLE 1  
Child outcome characteristics for soil ingestion and diarrhea in the past 7 days

	Total	Reported soil ingestion (N [% of age group])	Reported diarrhea (N [% of age group])		
			All	Soil ingestion	No soil ingestion
All children included in study	54	24 (44.4)	13 (24.1)	11 (20.4)	2 (3.7)
Children 3–6 months	4	0 (0)	0 (0)	–	–
Children 6–24 months	24	14 (58.3)	9 (37.5)	8 (33.3)	1 (4.2)
Children 24–36 months	14	6 (42.9)	3 (21.4)	3 (21.4)	0 (0)
Children 3–5 years	12	4 (33.3)	1 (8.3)	0 (0)	1 (8.3)

TABLE 2  
Association between child soil ingestion and diarrhea from penalized maximum likelihood logistic regression models

Child diarrhea explanatory variable	Unadjusted			Adjusted		
	OR	95% CI	P value	OR	95% CI	P value
Children aged 3 months to < 5 years (N = 54)						
Soil ingestion	9.7	2.1–44.1	0.003	9.9	2.1–47.5	0.004
Asset index				0.7	0.4–1.2	0.17
Number of household members				1.4	0.9–2.1	0.13
Children aged 6 months to < 36 months (N = 38)						
Soil ingestion	14.1	2.2–92.2	0.006	12.9	1.9–88.5	0.009
Asset index				0.7	0.4–1.2	0.22
Number of household members				1.4	0.84–2.3	0.19

CI = confidence interval; OR = odds ratio.

of samples were ND ( $N = 2$ ) for the HF183 target. Among samples within the ROQ, the *HF183* gene had a mean of 4.2 log copies (SD = 0.5) per gram of dry soil. There was no correlation between *E. coli* and HF183 levels found in soil samples (Pearson's  $r = 0.06$ ,  $P = 0.79$ ).

Households with a child who was observed to consume soil were more likely to have lower levels of HF183 copies in soil samples (Table 3, mean of 3.6 log copies for households without a child observed to put soil into their mouth versus 3.1 log copies for households with a child observed to put soil into their mouth,  $t = 2.3$ ,  $P = 0.03$ ). However, there was no statistically significant difference in *E. coli* counts (mean of 5.6 log CFU *E. coli* for households without a child observed to put soil into their mouth versus 5.5 log CFU *E. coli* for households with a child observed to put soil into their mouth,  $t = 0.71$ ,  $P = 0.49$ ).

Among households in which at least one child was reported to ingest soil, there was no statistically significant difference in the average *E. coli* count in soil samples from households with at least one child reported to have had diarrhea in the past week (mean of 5.6 log CFU *E. coli* versus 5.4 log CFU *E. coli* for households with children without diarrhea,  $t = -1.04$ ,  $P = 0.35$ ) or in HF183 copies in soil samples from these two groups of households (mean of 3.2 log copies versus 3.0 log for households with children without diarrhea,  $t = -0.56$ ,  $P = 0.59$ ).

Similarly, among all households, there was no statistically significant difference in the average *E. coli* count in soil samples from households with children reported to have had diarrhea in the past week (Table 3, mean of 5.7 log CFU *E. coli* versus 5.5 log CFU *E. coli* for households with children without diarrhea,  $t = -1.69$ ,  $P = 0.13$ ) or in the average HF183 gene copies in soil samples (mean of 3.1 log copies for households with children reported to have had diarrhea in the past week versus 3.4 log copies for households with children without diarrhea,  $t = 1.07$ ,  $P = 0.30$ ).

**Quality assurance and control.** All filtration blanks, extraction blanks, and no template controls were negative.

Linearity ( $R^2 = 0.99$ ) and efficiency ( $e = 103\%$ ) of the HF183 qPCR assay were within acceptable ranges. Minor PCR inhibition was detected between undiluted and 1:5 dilutions of DNA extract, but, with the exception of two samples, no inhibition was detected between 1:5 and 1:25 dilutions. As a result,  $C_q$  results from the 1:5 dilutions were used to quantify the number of gene copies in each sample.

## DISCUSSION

This study demonstrated an association between soil ingestion and child diarrhea in an urban slum environment. This adds to the growing body of knowledge that existing WASH interventions that target improving water access, toilet/latrine access, and handwashing may not eliminate children's exposure to fecal contamination in the domestic environment.<sup>9,10,22,23</sup> Reported soil ingestion was the highest among children aged 6–24 months, which is consistent with previous studies evaluating child geophagy<sup>10</sup> and child mouthing behavior.<sup>24</sup> In addition, no soil ingestion was reported for children younger than 6 months, which is likely due to children being relatively inactive at this age.

Our results showed a positive association with soil ingestion and child diarrhea in a densely populated, urban slum area of Kenya, which is consistent with a previous study in rural Kenya<sup>12</sup> that also assessed this relationship. Although our study found a higher association of soil ingestion with diarrhea (Pearson's correlation of  $r = 0.46$  between soil ingestion and diarrhea) than the previous study in rural Kenya (significant Pearson's correlation of  $r = 0.306$ ),<sup>12</sup> this may indicate that soil ingestion is a higher risk activity in urban slum areas, but may also be a result of the relatively small sample sizes in both studies. Notably, however, our study demonstrated a link between soil ingestion and diarrhea in households without earth floors (only one household included in our study had an earth floor in the home). As there is soil in the immediate environment surrounding many of the houses in this slum and children are not

TABLE 3  
Mean and standard deviation for *Escherichia coli* CFU and HF183 gene copies per dry gram of soil for the households and  $P$  values from two-sample  $t$  tests

	Child ingesting soil in household			Child with diarrhea in household		
	Yes (N = 18)	No (N = 16)	P value	Yes (N = 9)	No (N = 25)	P value
<i>E. coli</i> (log CFU per gram of dry soil)	5.5 (0.4)	5.6 (0.3)	0.49	5.7 (0.3)	5.5 (0.4)	0.13
HF183 gene (log copies per gram of dry soil)	3.1 (0.6)	3.6 (0.8)	0.03	3.1 (0.6)	3.4 (0.8)	0.30

CFU = colony forming units; SD = standard deviation. Values reported as mean (SD).

constantly restricted to staying inside the houses, child exposure to soil is likely even in households without earth floors. To our knowledge, this is the first study demonstrating this link in a study population with primarily nonearth household flooring. This finding indicates that soil ingestion may be an important exposure route for children regardless of household floor material, and upgrading earth floors to concrete or other nonearth materials may not eliminate soil as an exposure route for fecal contamination.

Soil fecal contamination measured in this study was also higher than past studies. Soil sampled in this study had a mean of 5.5 log CFU *E. coli* per gram of dry soil, which is higher than the reported means of 4.4 log CFU *E. coli* per gram of wet soil in an urban slum in Uganda,<sup>23</sup> 3.85 log CFU *E. coli* per gram of soil in rural Bangladesh,<sup>10</sup> 2.1 log CFU *E. coli* per gram of dry soil in peri-urban Tanzania,<sup>11</sup> and 1.84 log CFU *E. coli* per gram of soil from a yard laundry area in rural Zimbabwe during dry season.<sup>9</sup> However, the study in an urban slum in Uganda was conducted during dry season and used an augur to collect soil from the top 15 cm of soil, which is different from this study in which soil was collected only from the surface where children would come into contact with it and soil may have higher levels of contamination. Samples in our study were also taken during June, during the wet season when there may be higher fecal contamination exposure than in the dry season. High soil moisture and flooding have been linked with greater survival of *E. coli* in soils,<sup>25</sup> which may result in an increase of fecal contamination of soil during the wet season compared with the dry season. Although additional studies would be beneficial for providing more conclusive comparisons between the relative amount of soil contamination in these areas, our results indicate that fecal contamination of soil may be higher in urban slums than rural and peri-urban areas. This finding demonstrates the importance of using fecal contamination exposure estimates specific to urban slums when conducting risk assessments in these settings, as using values measured in other areas will likely underestimate the level of fecal contamination exposure. Furthermore, it is critical to include soil as an exposure point in fecal pathogen exposure risk assessment, and the high variability among fecal contamination levels measured in soil from different study locations warrants site-specific soil measures be conducted as part of exposure assessments.

The human-associated *Bacteroides* fecal marker HF183 was detected in 93% of soil samples ( $N = 26$ ), indicating that human fecal contamination of the soil sampled was likely. The specific human-associated *Bacteroides* marker chosen for this study (HF183 SYBR) has been previously validated in Kenya with a sensitivity of 65% and specificity of 100%.<sup>15</sup> Within the study area, there are several potential sources that could cause soil to become contaminated with human feces. Leaking pit latrines, unhygienic disposal of feces from young children (including large garbage piles that contained used diapers), and open drainage ditches are all potential sources of human fecal contamination that we observed within the study area. Flooding in households and shared outdoor spaces where soil was sampled was common, with 60% of households reporting flooding within the past month ( $N = 24$ ). Rainfall and flooding may cause feces from these potential sources to disperse and spread human feces around the environment, contaminating nearby

soil. More research is needed to determine the relative contribution of each potential human contamination source (leaking pit latrines, unhygienic child feces disposal, open drainage ditches) to the contamination of soil near households. However, these infrastructure and behavior characteristics may also explain why we measured higher levels of fecal contamination in this urban slum setting compared with fecal contamination levels previously measured in rural and peri-urban areas.

Although we did not use direct observation to measure the quantity of soil consumed by children as part of this study, it is reasonable to leverage published estimates of soil quantity consumed by young children to estimate fecal contamination exposure by soil ingestion in our study population. In rural Zimbabwe, it was estimated that children intentionally consumed 1.25 g of soil per episode, with a mean of 11.3 soil-mouth episodes during a 6-hour observation period.<sup>9</sup> This would be equivalent to children ingesting roughly 395,000 *E. coli* per episode in our study site. Although only certain strains of *E. coli* are capable of producing illness in humans, the detection of human-specific *Bacteroides* genetic markers has been shown to be predictive of the presence of pathogenic *E. coli* in surface waters in Japan.<sup>26</sup> As the human-specific *Bacteroides* HF183 genetic marker was detected in 93% of soil samples collected in this study, it is likely that pathogenic *E. coli* are present in the soil, but it is also possible that the relationship seen in surface waters between human-specific *Bacteroides* and pathogenic *E. coli* may not be representative of the relationship in soil samples collected from our study site. The presence of pathogenic *E. coli* can be estimated in the samples collected by assuming that 8% of *E. coli* detected were pathogenic, which has been previously recommended as a low-cost method of quantifying risk from water supplies in resource-limited settings.<sup>27</sup> This would correspond to children ingesting roughly 31,600 pathogenic *E. coli* per soil mouthing episode in our study site. Although the specific estimate for the quantity of pathogenic *E. coli* consumed should be interpreted with caution as many assumptions were made in the calculation and pathogenic genes were not measured directly in this study, this rough assessment illustrates that soil is likely a significant route for fecal contamination exposure in children.

In this study, we also found that households with lower levels of human fecal contamination in soil samples (quantified by copies of the *Bacteroides* HF183 gene) were more likely to have a child who was observed to consume soil. This may indicate that caregivers living in unclean environments with greater human fecal contamination are more likely to stop their children from putting soil into their mouth. Similarly, it may indicate that caregivers who are aware that their children put soil into their mouth are more likely to keep the soil near the household cleaner and dispose of waste somewhere else. However, in-depth interviews or focus groups would be necessary to confirm this finding.

Although this study supports the general hypothesis that soil ingestion may be an important transmission pathway for diarrheal disease in urban slum settings, it also has notable limitations. Since this is a cross-sectional study, a causal relationship between soil ingestion and diarrhea could not be determined. This study also had a small sample size, which limited the statistical power. Some potential

confounding variables were included in the analysis by using penalized maximum likelihood logistic regression, but there may be other confounding factors that were not included in the model that could influence the results and increasing the sample size in future studies could allow more confounding variables to be controlled for. In addition, although soil was found to contain high levels of fecal contamination, which may link soil ingestion with diarrhea, children who put soil into their mouth may also be more likely to put other contaminated objects into their mouth. It is possible that this mouthing/exploratory behavior may lead children to mouth or ingest other contaminated objects that cause diarrhea. In that case, reported soil ingestion could potentially be a proxy of high-risk mouthing behavior. Furthermore, this study relied on information reported by caregivers, which could lead to reporting bias. Although a previous study showed good agreement between caregiver reported and directly observed soil ingestion by children,<sup>10</sup> using structured observation in addition to reported data could improve accuracy. If structured observation is not feasible, future studies should include questions related to the frequency, location, and quantity of soil observed by caregivers to be consumed by children for exposure estimates. Future studies also could improve on this work by increasing the sample size to improve the generalizability of the results and provide greater statistical power to identify associations between soil contamination and soil properties, household behaviors, and infrastructure characteristics. In addition, future studies could inform educational intervention design by including questions related to whether the caregiver perceives soil ingestion by children as a risky behavior or not (and why the caregiver feels this way), as well as if the caregiver stops the child from ingesting soil when it is observed.

Despite recent evidence that soil ingestion may be an important exposure point for fecal pathogens, interventions aimed at reducing exposure from exploratory soil ingestion are limited. A randomized controlled trial in rural Zimbabwe, referred to as the Sanitation Hygiene Infant Nutrition Efficacy trial, is currently being conducted to evaluate the use of hygienic play spaces that are mobile and placed over soil to reduce contact with fecal contamination in the environment.<sup>28</sup> The evaluation of additional interventions is needed, including interventions to create a barrier between fecal pathogens in the environment and children in low-income urban areas, as well as educational interventions to reduce risky child mouthing behavior with soil. A recent study in Bangladesh identified that caregivers rarely intervene when children are ingesting soil, stopping children from soil ingestion only 14% of the time during direct observation.<sup>10</sup> Young children require close supervision, and the presence or lack of supervision may be informed by a caregiver's child-rearing attitude, including whether limits set for children are permissive or restrictive.<sup>29</sup> However, it is not well understood if the low frequency at which parents have been observed to intervene with child soil consumption is due to the caregiver's child-rearing attitudes, a lack of knowledge about risks associated with a child ingesting soil, or other factors. A better understanding of the factors that determine whether a caregiver intervenes during child soil ingestion would be useful for designing better interventions. In addition, improvements in sanitation

infrastructure, drainage infrastructure, child feces disposal practices, and solid waste management could also reduce soil contamination from human fecal sources.

This study provides further evidence that soil is an important pathway for fecal pathogen exposure for children in low-income environments, as well as that soil is still an important exposure pathway for children living in households with non-earth flooring. The results also suggest that soil ingestion may be a more important exposure pathway for children in urban slums than in rural areas as the urban environment soil had higher levels of contamination and soil ingestion had a stronger association with diarrhea. However, as there is limited data available for each of these settings, it is also possible that other site-specific characteristics may influence soil contamination more than the general classification of rural or urban setting and site-specific soil measures are recommended for future exposure assessments. Further research is also needed to evaluate interventions that limit soil ingestion behavior in children in both rural and urban areas. Reducing soil ingestion by children should be included in holistic approaches to improve WASH that move beyond conventional water and sanitation access.

Received July 1, 2016. Accepted for publication November 27, 2016.

Published online January 16, 2017.

**Acknowledgments:** We thank Phantus Wambiya, Rodah Obondi, and Philip Omandid for their help in conducting household interviews and fieldwork, Gabrielle Levato for her help with field sampling and lab work, Vincent Madadi at the University of Nairobi for providing laboratory space, and Juliet Iwelunmor at the University of Illinois for helpful discussions. We also thank the participating households who made this study possible.

**Financial support:** Valerie Bauza was supported by the National Science Foundation Graduate Research Fellowship Program under grant no. DGE-1144245 while conducting this work, and travel and supplies were supported by the Safe Global Water Institute (SGWI) and the Department of Civil and Environmental Engineering at the University of Illinois at Urbana-Champaign.

**Authors' addresses:** Valerie Bauza, Thanh H. Nguyen, and Jeremy S. Guest, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, E-mails: vbauza@illinois.edu, thn@illinois.edu, and jsguest@illinois.edu. R. M. Ocharo, Department of Sociology and Social Work, University of Nairobi, Nairobi, Kenya, E-mail: rocharo@uonbi.ac.ke.

## REFERENCES

1. WHO, 2013. *Diarrhoeal Disease Fact Sheet*. Geneva, Switzerland: World Health Organization. Available at: <http://www.who.int/mediacentre/factsheets/fs330/en/>. Accessed December 15, 2015.
2. Curtis V, Cairncross S, Yonli R, 2000. Review: domestic hygiene and diarrhoea—pinpointing the problem. *Trop Med Int Health* 5: 26–30.
3. Kutalek R, Wewalka G, Gundacker C, Auer H, Wilson J, Haluza D, Huhulescu S, Hillier S, Sager M, Prinz A, 2010. Geophagy and potential health implications: geohelminths, microbes and heavy metals. *Trans R Soc Trop Med Hyg* 104: 787–795.
4. Young SL, Khalfan SS, Farag TH, Kavle JA, Ali SM, Hajji H, Rasmussen KM, Peltó GH, Tielsch JM, Stoltzfus RJ, 2010. Association of pica with anemia and gastrointestinal distress among pregnant women in Zanzibar, Tanzania. *Am J Trop Med Hyg* 83: 144–151.
5. Luoba AI, Wenzel Geissler P, Estambale B, Ouma JH, Alusala D, Ayah R, Mwaniki D, Magnussen P, Friis H, 2005. Earth-eating and reinfection with intestinal helminths among

- pregnant and lactating women in western Kenya. *Trop Med Int Health* 10: 220–227.
6. Young SL, Sherman PW, Lucks JB, Pelto GH, 2011. Why on earth?: evaluating hypotheses about the physiological functions of human geophagy. *Q Rev Biol* 86: 97–120.
  7. Miao D, Young SL, Golden CD, 2015. A meta-analysis of pica and micronutrient status: pica and micronutrient meta-analysis. *Am J Hum Biol* 27: 84–93.
  8. Geissler PW, Mwaniki DL, Thiong'o F, Michaelsen KF, Friis H, 1998. Geophagy, iron status and anaemia among primary school children in western Kenya. *Trop Med Int Health* 3: 529–534.
  9. Ngure FM, Humphrey JH, Mbuya MNN, Majo F, Mutasa K, Govha M, Mazarura E, Chasekwa B, Prendergast AJ, Curtis V, Boor KJ, Stoltzfus RJ, 2013. Formative research on hygiene behaviors and geophagy among infants and young children and implications of exposure to fecal bacteria. *Am J Trop Med Hyg* 89: 709–716.
  10. George CM, Oldja L, Biswas S, Perin J, Lee GO, Kosek M, Sack RB, Ahmed S, Haque R, Parvin T, Azmi IJ, Bhuyian SI, Talukder KA, Mohammad S, Faruque AG, 2015. Geophagy is associated with environmental enteropathy and stunting in children in rural Bangladesh. *Am J Trop Med Hyg* 92: 1117–1124.
  11. Pickering AJ, Julian TR, Marks SJ, Mattioli MC, Boehm AB, Schwab KJ, Davis J, 2012. Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian households with and without improved sanitation. *Environ Sci Technol* 46: 5736–5743.
  12. Shivoga WA, Moturi WN, 2009. Geophagia as a risk factor for diarrhoea. *J Infect Dev Ctries* 3: 94–98.
  13. WHO/UNICEF, 2015. *Progress on Sanitation and Drinking Water—2015 Update and MDG Assessment*. Geneva, Switzerland and New York, NY: World Health Organization and UNICEF.
  14. Boehm AB, Griffith J, McGee C, Edge TA, Solo-Gabriele HM, Whitman R, Cao Y, Getrich M, Jay JA, Ferguson D, Goodwin KD, Lee CM, Madison M, Weisberg SB, 2009. Faecal indicator bacteria enumeration in beach sand: a comparison study of extraction methods in medium to coarse sands. *J Appl Microbiol* 107: 1740–1750.
  15. Jenkins MW, Tiwari S, Lorente M, Gichaba CM, Wuertz S, 2009. Identifying human and livestock sources of fecal contamination in Kenya with host-specific Bacteroidales assays. *Water Res* 43: 4956–4966.
  16. Seurinck S, Defoirdt T, Verstraete W, Siciliano SD, 2005. Detection and quantification of the human-specific HF183 Bacteroides 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environ Microbiol* 7: 249–259.
  17. Layton BA, Cao Y, Ebentier DL, Hanley K, Ballesté E, Brandão J, Byappanahalli M, Converse R, Farnleitner AH, Gentry-Shields J, Gidley ML, Gourmelon M, Lee CS, Lee J, Lozach S, Madi T, Meijer WG, Noble R, Peed L, Reischer GH, Rodrigues R, Rose JB, Schriewer A, Sinigalliano C, Srinivasan S, Stewart J, Van De Werfhorst LC, Wang D, Whitman R, Wuertz S, Jay J, Holden PA, Boehm AB, Shanks O, Griffith JF, 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Res* 47: 6897–6908.
  18. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT, 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55: 611–622.
  19. Cao Y, Griffith JF, Dorevitch S, Weisberg SB, 2012. Effectiveness of qPCR permutations, internal controls and dilution as means for minimizing the impact of inhibition while measuring Enterococcus in environmental waters. *J Appl Microbiol* 113: 66–75.
  20. Heinze G, Puh R, 2010. Bias-reduced and separation-proof conditional logistic regression with small or sparse data sets. *Stat Med* 29: 770–777.
  21. Heinze G, Schemper M, 2002. A solution to the problem of separation in logistic regression. *Stat Med* 21: 2409–2419.
  22. Zambrano LD, Levy K, Menezes NP, Freeman MC, 2014. Human diarrhea infections associated with domestic animal husbandry: a systematic review and meta-analysis. *Trans R Soc Trop Med Hyg* 108: 313–325.
  23. Katukiza AY, Ronteltap M, van der Steen P, Foppen JWA, Lens PNL, 2014. Quantification of microbial risks to human health caused by waterborne viruses and bacteria in an urban slum. *J Appl Microbiol* 116: 447–463.
  24. Juberg DR, Alfano K, Coughlin RJ, Thompson KM, 2001. An observational study of object mouthing behavior by young children. *Pediatrics* 107: 135–142.
  25. Jamieson RC, Gordon RJ, Sharples KE, Stratton GW, Madani A, 2002. Movement and persistence of fecal bacteria in agricultural soils and subsurface drainage water: a review. *Can Biosyst Eng* 44: 1–9.
  26. Savichtcheva O, Okayama N, Okabe S, 2007. Relationships between Bacteroides 16S rRNA genetic markers and presence of bacterial enteric pathogens and conventional fecal indicators. *Water Res* 41: 3615–3628.
  27. Howard G, Pedley S, Tibatemwa S, 2006. Quantitative microbial risk assessment to estimate health risks attributable to water supply: can the technique be applied in developing countries with limited data? *J Water Health* 4: 49–65.
  28. Humphrey JH, Jones AD, Manges A, Mangwadu G, Maluccio JA, Mbuya MN, Moulton LH, Ntozini R, Prendergast AJ, Stoltzfus RJ, Tielsch JM; SHINE Trial Team, 2015. The Sanitation Hygiene Infant Nutrition Efficacy (SHINE) trial: rationale, design, and methods. *Clin Infect Dis* 61: S685–S702.
  29. Grusec JE, Danyliuk T, 2014. *Parents' Attitudes and Beliefs: Their Impact on Children's Development*. *Encyclopedia on early childhood development*. Montreal, Quebec: Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development.