Spring 2015

Novel scale development to assess the role of sanitation access and use on household fecal contamination in Accra, Ghana

Rebecca Lyn Ritter
University of Iowa

NOVEL SCALE DEVELOPMENT TO ASSESS THE ROLE OF SANITATION ACCESS AND USE ON HOUSEHOLD FECAL CONTAMINATION IN ACCRA, GHANA

by

Rebecca Lyn Ritter

A thesis submitted in partial fulfillment
of the requirements for the Master of Science
degree in Epidemiology in the
Graduate College of
The University of Iowa

May 2015

Thesis Supervisor: Assistant Professor Kelly Baker
Associate Professor Christine Petersen

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

MASTER'S THESIS

_____

This is to certify that the Master's thesis of

Rebecca Lyn Ritter

has been approved by the Examining Committee for
the thesis requirement for the Master of Science degree
in Epidemiology at the May 2015 graduation.

Thesis Committee: _____
Kelly Baker, Thesis Supervisor

_____
Christine Petersen, Thesis Supervisor

_____
James Torner

ACKNOWLEDGEMENTS

Thanks to our collaborators, Christine Moe, Clair Null, Habib Yakubu, Dorothy Peprah, and Kate Robb at the Emory University, George Armah at the University of Ghana, Joseph Ampofo at the Water Research Institute, Accra, Ghana, Nii Wellington at the Trend Research Group, Accra, Ghana, and Kelly K. Baker of Emory University and the University of Iowa College of Public Health.

Special thanks to the behavioral and environmental staff, neighborhood liaisons, and study participants in Ghana who made this work possible.

This work was supported by a grant from the Bill and Melinda Gates Foundation.

ABSTRACT

Diarrheal disease is one of the leading causes of mortality of children under 5 years of age. Despite this, diarrheal disease is easily preventable through adequate water, sanitation and hygiene. Sanitation access is currently classified as “improved” or “unimproved” based on level of latrine access. This does not account for differences in human behaviors, or differences in exposure risk. A sanitation score was built using behavioral and access data in order to better classify the sanitation environment of a household. Due to low levels of sanitation access and practice of open defecation in Ghana, households in four neighborhoods in Accra, Ghana were selected to participate in the data collection. Data was collected through a survey, environmental sanitary inspections and collection of hand rinse and environmental swab samples. These samples were then tested for fecal indicators, by measuring presence and concentration of *E. coli* and human Adenovirus. A novel sanitation score based on latrine access and use for each household was created. Hierarchical linear and logistic regression was used to compare the sanitation score to the environmental contamination as indicated by the *E. coli* and Adenovirus. Higher sanitation scores were significantly associated with increases in Adenovirus concentration (PR=1.6, 95%CI=1.1, 2.2). The sanitation score was not significantly associated with *E. coli* or presence of Adenovirus. Further development of a sanitation score variable could help to better understand sanitation environments.

PUBLIC ABSTRACT

Diarrheal disease is one of the leading causes of mortality of children under 5 years of age, and is easily preventable through adequate water, sanitation and hygiene. Sanitation access is currently classified as “improved” or “unimproved” based on level of latrine access. This does not account for differences in human behaviors, or differences in exposure risk. A sanitation score was built using behavioral and access data in order to better classify the sanitation environment of a household. Due to low levels of sanitation access and practice of open defecation in Ghana, households in four neighborhoods in Accra, Ghana were selected to participate in the data collection. The sanitation score was based on latrine access and defecation and waste disposal for children in each household based on surveys and observations from a caretaker in the household. This score was validated against environmental samples of hand rinses and swabs collected from each household. Human Adenovirus strains and *E. coli* in samples were used as indicators of fecal contamination in the samples. Higher sanitation scores were significantly associated with increases in Adenovirus concentration (PR=1.6, 95%CI=1.1, 2.2). The sanitation score was not significantly associated with *E. coli* or presence of Adenovirus. This suggests that sanitation access and behavior are significantly associated to exposure risk for diarrheal pathogens.

TABLE OF CONTENTS

List of Tables vi

List of Figures vii

Chapter

1. Introduction 1
2. Materials and Methods 5
3. Results 11
4. Discussion 14

Works Cited 27

Appendix

A. Assessing Sanitation-related Exposure: Fecal Indicators 29
B. Principal Components Analysis- Wealth Index 32
C. GLIMMIX Procedure 33
D. Data Collection- Survey Questions 35

LIST OF TABLES

Table

1. Sanitation score system....17
2. Socio-demographic characteristics for 71 study households in Accra, Ghana....18
3. Household indicators of child and animal fecal disposal by level of sanitation access for 71 households in Accra, Ghana....19
4. Distribution of *E. coli* and Adenovirus in swab and hand rinse samples....19
5. Hierarchical logistic regression models of sanitation score and presence or absence of *E. coli* or Adenovirus as a measure of fecal contamination, adjusted for other household covariates....22
6. Hierarchical linear regression models of sanitation score and log concentration of *E. coli* or Adenovirus as a measure of fecal contamination, adjusted for other household covariates....23

D1. Eigenvalues of PCA....34

LIST OF FIGURES

Figure

1. Distribution of Sanitation Score by Household 20
2. Comparison by latrine access level of sanitation score accounting only for child defecation practices and observed feces 21
3. Residual fit of *E. coli* logarithmic concentration and sanitation score in hand rinse samples 24
4. Residual fit of *E. coli* logarithmic concentration and sanitation score in swab samples 25
5. Residual fit of Adenovirus logarithmic concentration and sanitation score in swab samples 26

D1. Scree plot of PCA Wealth 38

D2. Scatterplot of logarithmically transformed *E. coli* concentration and logarithmically transformed Adenovirus concentration of all samples 39

Chapter 1

Introduction

Worldwide, an estimated 21 percent of deaths of all children under five years of age occur due to diarrheal diseases. Annually this accounts for an estimated 2.5 million deaths worldwide, 800,000 of which occur in Sub-Saharan Africa alone[1,2]. The greatest risk factors for diarrheal disease are inadequate sanitation, water, and hygiene[3]; which are highly prevalent in low-income countries, where poverty, unhygienic conditions, and disease are ubiquitous[1]. An estimated 37% of the world's population, or 2.6 billion people, lack access to improved facilities for disposal of human excrement, as defined by the World Health Organization and the UNICEF Joint Monitoring Program (JMP) for Water Supply and Sanitation[3-6]. Access to a sanitary latrine can ensure protection of users from exposure to feces of others; and builds barriers between human excreta and the environment.

The Millennial Development Goals (MDG) has placed a greater emphasis on tracking and improving sanitation and hygiene conditions in the developing world. Access to better sanitation has been targeted because it has been shown to be one of the most cost-effective interventions for preventing the high disease burden that occurs in developing countries[7]. The JMP and WHO defines improved and unimproved sanitation environments on a four-step ladder, households are categorized based on access to sanitation facilities. Currently all shared sanitation facilities are considered unimproved, while private facilities are considered improved. Using these definitions, in Sub-Saharan Africa, only an estimated 31% of persons have access to improved sanitation[6]. Urban dwellers that rely on shared sanitation facilities and groups of households that co-own

facilities are considered unimproved. Recent studies have suggested that current understanding of classification of improved and unimproved sanitation was too simplistic, and does not account for differences in shared latrine characteristics, nor in household defecation behavior as a part of overall sanitation safety[3-5,8-12]. Shared latrines have been defined as unimproved based on the premise that they are more likely to be poorly maintained, so users may be more likely to be exposed to the feces of other users. However, several studies have found contradictory levels of contamination from fecal indicators in household and shared facilities[4,13,14]. There is still limited evidence to understand whether shared facilities are less safe, and if so under what conditions. Furthermore, even less is known about whether shared and private facilities perform comparably at capturing and containing human waste and preventing it from entering the environment.

Emerging evidence from several studies of sanitation interventions has revealed that providing people access to a latrine structure does not necessarily prevent open defecation, and that latrine use – even for household latrines – is highly inconsistent[15]. Avoidance of unsanitary or unsafe latrines when available, lack of privacy, cost, or desire to use feces as free fertilizer can all motivate continued open defecation. Latrines shared with many other persons may be particularly poor at eliciting consistent use among all members of a community, and particularly among marginalized groups[16]. Women and children tend to be the most likely to practice open defecation because of greater limitations in resources and fear of safety, and are more often obligated to practice open defecation at night[16-18]. Children in particular tend to have higher rates of open defecation due to parental beliefs that feces of children are not able to cause disease[18,19], and

households with a higher ratio of pre-school children to adults were less likely to use facilities[19]. The definition of improved sanitation should account for a facility's structural and operational capacity to prevent exposure of users within the facility, but also its effectiveness at stimulating adequate use by a population to prevent fecal contamination of the overall environment.

The goal of this study was to test the hypothesis that having private latrine access and use of available facilities among members of a household would decrease fecal contamination within the household. The city of Accra, Ghana was selected for this study because of the high prevalence of diarrheal disease, particularly in poor households, and low levels of sanitation coverage. In Ghana alone, diarrhea causes 84,000 deaths annually, 25% of which are children under five years[1]. This may be due to extremely low levels of sanitation access in Ghana, which has the 4th lowest rate of sanitation coverage worldwide[20]. Approximately 4.8 million Ghanaians, equivalent to almost 20% of the country's population, currently do not have access to latrines and practice open defecation, while another 16 million (almost 60%) use unimproved or shared latrines[21]. Current sanitation coverage in Accra, Ghana varies by neighborhood, but is generally low. Data for this study was collected across four different neighborhoods in Accra on variables related to latrine use and access in each household as well as environmental samples. Neighborhoods were chosen that are thought to be representative of sanitation coverage, wealth, religion, and environment of greater Accra. We developed a sanitation scale which classified households upon level of access and reported and observed indicators of use. The environmental samples from each household were then used to test for an association between the sanitation score and qualitative and quantitative indicators

of household cleanliness. This sanitation score described herein could be an improvement on current classification of household sanitation environments based on both behavior and access.

Chapter 2

Materials and Methods

Data collection. The data for the Sanipath study, through which this study was supported, was collected in the city of Accra, Ghana due to the low levels of sanitation coverage and the high proportion of the urban population that relies on public latrines. The neighborhoods of Alajo, Bukom, Old Fadama, and Shiabu were chosen to be representative of the diversity of conditions, culture, and geographical and infrastructural development. This project started in March of 2012, and ran through November of 2012, with household visits occurring in each neighborhood on a cyclical basis by week. Households were randomly chosen from a list generated by a neighborhood liaison, who also facilitated the introductions between the data collectors and households in each neighborhood. Data collectors were trained by Sanipath personnel for all sample collection and survey procedures. Household participants were informed of study objectives and provided written consent to participate in the study. Consent and interviewing was performed in English, or in the participant's native language. There were three parts to the data collection, a structured observation period, a survey, and environmental sanitary inspections and sample collection. The use of multiple methods of data collection allowed this study to compare the sanitation behaviors and access in the environment with environmental markers of contaminations.

Structured observations and household surveys were collected by the same team at each location. Structured observations were used to document and tally observed sanitation-related behaviors in children. One to three children under the age of five per household were observed, and any risk or protective behaviors were recorded using the

structured observation form. Basic demographic information on the child and caregiver was collected, including age, sex, relationship to caregiver, and latrine use. Observations were done on the presence or absence of a compound latrine, the condition of the latrine, feces disposal methods, provision of sanitary materials, trash, and storage of drinking water. The household survey was given to one member, usually the primary caregiver or head of household, of each household that was observed. This survey recorded information designed to understand the economic and social behaviors of the household. Information was collected on background demographic information, such as number of people living in the household, age ranges, water access, latrine access, number of households sharing the latrine, and use practices for children less than five years of age and between 5 to 12 years of age.

Sample collection and processing. Sample collection of swabs and hand rinses were performed by a second team for each location. Teams collected up to 3 swabs and 3 hand rinses per household, pending availability of subjects and objects contacted by children. Hand rinses were collected by asking the study participant to submerge each hand up to the wrist in 500 ml PBS in a WhirlPak bag. The surfaces of the hand were gently massaged from the outside of the bag for 30 seconds, and then repeated for the second hand. Pre-moistened macrofoam swabs were used to swab a 100 by 100 cm area that the data collectors had observed the child had contact with during structured observation, or if an irregular shaped object, the entire object was swabbed. Swabs were washed twice in 8 mls of phosphate buffered saline, pH 7.2 with 0.04% Tween-80.

Human Adenovirus and *E. coli* were used to measure fecal contamination in the environment. *E. coli* is the most commonly used indicator for tracking fecal

contamination, and as common gut microbes of most mammalian species, represent fecal contamination of both human and animal. Human Adenoviruses 40/41 can be used to identify human-specific fecal contamination presence.

Swab and hand rinse samples were tested for *E. coli* presence and concentration by membrane filtration of three serial dilutions of swab elute (1ml, 0.1 ml, and 0.01ml) and three serial dilutions of hand rinse elute (100 ml, 10 ml, 1 ml). *E. coli* were enumerated using the EPA method 1604 by plating filters on BBL© MI agar and culturing overnight at 37° C. Samples with a minimum of a single colony of *E. coli* as indicated by the BBL MI© agar, were considered positive. Concentration was estimated based on colony count, those with no colonies were transformed by 0.5x the lower limit of detection, and those with colonies too numerous to count were transformed by 1.5x the upper limit of detection. Adenovirus DNA was extracted from 1.5 ml of swab elute using the Fast DNA SPIN kit for soil (MP Biomedicals, Solon, OH) plus 5 cycles of a freeze-thaw step to improve disruption of viral particles. Duplicate 5µl volumes of extract were first tested using the Quantifast Pathogen + Internal Control PCR kit to identify samples with poor amplification due to inhibition. Then samples that had positive amplification in one or both duplicates, or that showed signs of inhibition, were retested for Adenovirus using a quantitative PCR method described previously in Jothikumar 2005[22]. Amplification results with no Adenovirus detected at either stage of screening were transformed using 0.5x the lower limit of detection of the assay and were screened for consistency (different in duplicates ≤ 5 CT) to ensure within-sample variance did not skew estimates of viral concentrations. Concentrations for each sample were generated by

averaging the concentration of duplicates and back calculating concentrations to estimate the concentration per total elute or swabbed surface.

Sanitation Score. A sanitation score was created for each household which combines the household access to latrines with the reported use of latrines from the survey data. On this scale, households with what are expected to be the worst sanitation environments have the highest scores. The scale is broken down into 5 indicator categories in Table 1, the latrine access of the compound, the defecation location of children aged 5-12 in the compound, the defecation location and disposal location for children under the age of 5 living in the household, the disposal of feces for children under the age of 5 by other mothers in the compound, and the observation of feces in the compound. Each household was given a number for each category. Households with private facilities were assigned lower scores based upon the hypothesis that better latrine access would improve use and thus fecal contamination within the household. Households were penalized in the score for poor reported use of the latrine by the children aged 5-12 in the household and the reported use of other children 5-12 living in each compound. This was multiplied by the density of children 5-12 in the household in order to account for differences in household density and contribution to the sanitation environment. Households were given a higher score if children under the age of 5 in the household were allowed to defecate in places that were more likely to allow for human contact, and if feces from these children were disposed of in places that allowed for more human contact. This was also multiplied by the number of children under the age of 5 in the household in order to account for differences in household density. Household score was increased if other mother's in the compound disposed of feces of children under the

age of five in places that allow for more human contact. Households were also penalized for observed feces on the ground of the compound. The score for each category was standardized based on the potential number of points the household could have scored in each indicator category. These were then added together to create the final score. For example, if a household reported shared latrine access (score=2/3), that 2 children between 5-12 years from the household used the latrine, and the other children in the compound used the latrine (score= 0/2*3), that there is one child under 5 in the household that defecated on the ground and the feces are disposed of in the latrine, rubbish, or drain (score= 2/3 * 1), that the other mother's in the compound dispose of child feces in the drain or rubbish (score= ½), and feces were seen on the ground (score=1/1), then the total household score would be 2.83.

Statistical analysis. Descriptive statistics of study households based on neighborhoods and latrine access levels were performed, using χ^2 or Fischer's Exact test for categorical variables, and ANOVA tests for continuous variables. Hierarchical logistic and linear regression models were built separately for the hand rinse and swab samples, and used to estimate the association between the sanitation score and the presence and magnitude of environmental contamination as measured by *E. coli* and Adenovirus concentration. To correct for the skewness of the *E. coli* and Adenovirus data, concentrations were logarithmically transformed.

Data was analyzed using SAS 9.3 procedure PROC GLIMMIX. This allowed for the adjustment of random effects by household to account for potential correlation between multiple surface swab and individual hand rinse samples being taken from each household. Models were not adjusted for neighborhood level effects, as neighborhood

was correlated with latrine access levels. Covariates considered for inclusion in the model were selected *a priori,* based upon the plausible association with both sanitation access and fecal contamination exposure risks. Covariates were included in the initial model if they were independently significant at $p \leq 0.2$ in univariate analysis. Covariates in the model were kept if they significantly improved the model fit relative to other models, as measured by the AIC, or altered the effect significance of other variables in the model. Interaction terms were tested independently in each model, and were retained if $p < 0.05$. Logistic regression models were summarized using odds ratio (OR) estimates and 95% confidence intervals (95%CI). Linear regression models were summarized using prevalence ratios (PR) and 95% confidence intervals (95% CI).

Ethical considerations. Written informed consent was obtained from adult household heads enrolled in the Sanipath study. Consent for household studies involved describing the research objectives and methods to participants and requesting written consent indicating agreement to participate. Study protocols were approved by Research Review Committee and Ethical Review Committee of the University of Ghana, Legon (Ghana) and Emory University (USA) Institutional Review Boards.

Chapter 3

Results

Characteristics of households by neighborhood in Accra, Ghana. Household characteristics in the Sanipath study were varied by neighborhood location (Table 2). Neighborhood households were significantly different for level of education, religion, business presence in the home, latrine access, wealth index, and the total number of people in each household, and the number of men and children 13-17 in each household. However, they were similar in respect to drinking water source, tenancy status, and number of women and children ages 5-12 and less than 5 years of age living in the household, as well as the number of households in each compound (Table 2).

Relationship between sanitation access and usage indicators. There was a significant difference in the feces disposal location for children under the age of five based on latrine access level ($p<0.0001$) (Table 3). There were no differences in other sanitation-related factor between households based on their level of access to latrines. There was not a significant difference between the defecation locations of children 5-12 ($p=0.3$), or children under 5 ($p=0.5$). There were not significant differences for observed feces presence ($p=0.79$) or animal presence ($p=0.42$) in the compound. Latrine access level alone was not significantly associated with either environmental *E. coli* or Adenovirus (*E. coli* concentration $p=0.9$, *E. coli* presence $p=0.5$, Adenovirus concentration $p=0.4$, adenovirus presence $p=0.3$).

Prevalence of E. coli and Adenovirus. Hand rinse and surface swabs were used as indicators of the prevalence and concentration of general and human-specific fecal contamination in urban households in Accra. *E. coli* presence was found in 89% of 80

hand rinse samples collected from 47 households. The mean logarithmic *E. coli* concentration was 2.5 in hand rinse samples (Table 4). *E. coli* presence was found in 86% of 94 swab samples collected from 54 households. The mean logarithmic *E. coli* concentration was 2.0 in swab samples (Table 4). Adenovirus was found in 24% of all swab samples, and had a mean concentration after logarithmic transformation of 3.1 in swab samples (Table 4). The average Adenovirus logarithmic concentration among positive samples only was 3.6. There was poor correlation between household concentrations of *E. coli* bacteria and Adenovirus for all samples ($r=-0.1$, $p=0.2$) (Figure S2).

Sanitation Score. The standardized sanitation score was somewhat skewed to the left, with most households having a sanitation score of less than 5 (Figure 1). Sanitation behaviors of children under the age of 5 contributed most to most household sanitation scores. Analysis of the sanitation score without latrine access levels showed that there were significant differences in sanitation score based on latrine reference groups ($F=9.28$, $p<0.0001$) (Figure 2).

Association between E. coli or Adenovirus presence and sanitation score. After model adjustment, presence of fecal indicators *E. coli* and Adenovirus was not significantly associated with the household sanitation score (Table 5). Households with a higher sanitation scores were more likely to have *E. coli* presence in hand rinse samples (OR= 1.1, 95%CI= 0.8, 1.6). Households with higher sanitation scores were more likely to have *E. coli* presence in swab samples (OR=1.1, 95%CI=0.8, 1.4), but this was not significant. Households with animals present were more likely to have *E. coli* present

(OR=2.7, 95%CI= 0.6, 11.6), though this was not significant. There was no association between the sanitation score and Adenovirus presence (OR=1.0, 95%CI=0.8, 1.3).

Association between E. coli or Adenovirus concentration and sanitation score. Households with a higher sanitation score were not associated with changes in concentration of *E. coli* in hand rinse samples after adjustment (PR=1.0, 95%CI= 0.9, 1.2) (Table 6). Christian households were significantly associated with increases in concentration of *E. coli* in hand rinse samples (PR=2.2, 95%CI= 1.2, 4.0). High sanitation scores were not associated with changes in concentrations of *E. coli* in swab samples (adjPR=1.0, 95%CI= 1.0, 1.1). Animal presence was significantly associated with decreases in *E. coli* concentrations in swab samples (adjPR=0.5, 95%CI= 0.3, 0.7). Increases in sanitation score were associated with significant increases in Adenovirus concentration (adjPR=1.6, 95%CI=1.1, 2.2). Use of hierarchical linear modeling did not capture a significant amount of the variability in the model (Covariance Parameter Residual= 0.8, SE=0.2).

Chapter 4

Discussion

Classification of study households into latrine access levels, based on current WHO and JMP categories, was not associated with environmental and hand contamination as measured by *E. coli* and Adenovirus. This contradicts the current assumption that access to a private latrine reduces exposure to unsafe sanitary environments, and subsequently reduces risk of disease. If access to a private latrine truly reduced risk of exposure to organisms in human waste materials, then there should be a measurable difference in environmental and hand exposure to these organisms. The lack of significant association between latrine access levels and *E. coli* and Adenovirus in samples shows that the pathways of exposure are more complex than simply latrine access. Despite the lack of direct association, the significant difference in behavior across latrine access groups suggests that latrine access levels are important to consider in modification of human behavior. Human behavior surrounding the latrine is also important in reducing risk of exposure, because access to a latrine is not always indicative of use. Recent studies have shown that certain groups such as women and children, or those living in urban slums with often poorly maintained facilities, are at greater risk for poor sanitation practices[1,19]. Within this study population, current guidelines lack the ability to capture important differences in human behavior and environment that contribute to poor sanitary environments. The households in this study were chosen because of their high risk for open defecation and other high risk sanitation behaviors. The sanitation score, which was built with the intention of capturing variability in sanitation access and behavior in those at highest risk for poor sanitation,

children, for each household, was significantly associated with increases in Adenovirus in environmental swab samples. This suggests that the sanitation score is significantly associated with human fecal contamination in the household environment. The qualitative models of *E. coli* concentration and hand rinse and swab samples, and Adenovirus and swab samples showed no relationship between presence and sanitation score. There was also not an association between the concentrations of *E. coli* in hand rinse or swab samples and the sanitation score in the quantitative models. This suggests that there is not an association between latrine use and environmental and hand *E. coli* presence or concentration. Households that were Christian were associated with higher concentrations of *E. coli* in hand rinse samples. This may be due to the religious importance of hygiene in the Muslim population in Accra. Muslims may be more likely to wash hands after defecation in comparison to Christians. Households with animals present were associated with lower concentrations of environmental *E. coli,* while this is unexpected, it may indicate that households with animals are more likely to clean or sanitize household surfaces than those without animals, which may be more important in household *E. coli* contamination than latrine access and use. The *E. coli* models showed generally good fit, but the Adenovirus linear model did not (Figure 3-5). While 23.9% of swab samples tested positive for Adenovirus, only 5 samples (5.3%) had concentrations above the lower limit of detection of the testing methods, and so small positives were not differentiated from negative samples. The current modeling strategy may not be appropriate for this data due to lack of high concentrations, and limits our abilities to draw conclusions from the data. A large amount of variability is also still unexplained through these models,

indicating that important neighborhood or household level effects are not accounted for by differences in access and use.

This study was limited by sample size, particularly due to the lack of households with private latrines in the study. A larger sample size would have allowed the exploration of a greater number of variables in the model, and the addition of neighborhood level effects. There are significant differences in neighborhoods that may explain some variation in household contamination, such as wealth, education, and latrine access, which may not be fully accounted for in the current model. Some variables that were originally expected to be significant had to be excluded due to missing data, such as hand washing stations and business presence in the household. Open defecation may have been too prevalent in neighborhoods, which would add to the sanitation environment of the household independently of household member latrine use. This may have limited the ability to detect differences due to latrines. Use of more sensitive testing methods for Adenovirus would allow for a better understanding of the association between Adenovirus and the sanitation score. In the future, further development of a sanitation score variable could help to better understand sanitation environments. This sanitation score could be used as a starting point to create evidence based categories of "improved" and "unimproved" sanitation by redefining categories to reflect what truly contributes to exposure risk. A larger study would allow researchers to look at the contribution of more indicators of sanitation. Further, the sanitation score could be expanded to include indicators of sanitation that were not available to this study, such as hand washing stations and adult latrine use.

Table 1. Sanitation score system.

Indicator	Levels of Indicator	Penalty Points	Possible score	Adjustment	Adjusted score
Latrine Access	Private latrine in compound	0	Lat= score/ max possible score= [0, 0.5, 1]		adjLat = Lat
	Shared latrine in compound	1			
	Public latrine access/ no access	2			
Defecation location of children in compound ages 5-12	All children in compound defecate in latrine	0	Cdef= score/ max possible score= [0, 0.5, 1]	N_1= number of children 5-12 in the household	adjCdef= (Cdef)N_1
	Household children defecate in latrine, but other children in compound do not	1			
	Children don't use latrine	2			
Defecation location and feces disposal of children under 5 years	Children defecate in potty or diaper and feces are disposed of in latrine	0	Idef= score/ max possible score= [0, 0.33, 0.67, 1]	N_2 = number of children under 5 years in the household	adjIdef= (Idef)N_2
	Children defecate in potty or diaper and feces are disposed of in drain or rubbish	1			
	Children defecate on ground and feces are disposed of in latrine, rubbish, or drain	2			
	Children defecate on ground and feces are left on ground	3			
Other mother's feces disposal for children under 5 years	Feces are disposed of in any latrine	0	Oth= score/ max possible score= [0, 0.5, 1]		adjOth= Oth
	Feces are disposed of in drain or rubbish	1			
	Feces are left on ground	2			
Feces observed on ground in compound	No	0	Obs= score/ max possible score= [0,1]		adjObs= Obs
	Yes	1			
			0-5		x

Adjusted final score(x)= adjLat + adjCdef + adjIdef + adjOth + adjObs

Table 2. Socio-demographic characteristics for 71 study households in Accra, Ghana.

Neighborhood		Alajo n=18	Bukom n=17	Old Fadama n=19	Shiabu n=17	P
Education of caregiver	No Formal Education	2 (11.1)	8 (47.1)	16 (84.2)	1 (5.9)	<.0001
	Some primary	8 (44.4)	7 (41.2)	3 (15.8)	5 (29.4)	
	Completed Primary	7 (38.9)	1 (5.9)	0 (0.0)	9 (52.9)	
	Some secondary	0 (0.0)	1 (5.9)	0 (0.0)	2 (11.7)	
	Completed Secondary	1 (5.6)	0 (0.0)	0 (0.0)	0 (0.0)	
Tenancy Status (Own)		8 (44.4)	13 (76.5)	12 (63.2)	10 (58.8)	0.29
Religion	Christian	16 (88.89)	15 (88.2)	5 (26.3)	17 (100.0)	<.0001
	Moslem	2 (11.11)	1 (5.9)	14 (73.7)	0 (0.0)	
	Other	0 (0.0)	1 (5.9)	0 (0.0)	0 (0.0)	
Business Presence		10 (52.6)	8 (47.1)	11 (64.7)	3 (16.7)	0.0294
Number of Households in Compounds		11.5 (1-70)	5.3 (0-12)	4.4 (1-11)	8.2 (1-24)	0.252
Number of People in Household	Total	36.0 (6-73)	23.1 (8-50)	9.3 (2-19)	26.5 (4-64)	0.031
	Men	2.0 (0-10)	1.5 (0-5)	0.5 (0-1)	1.3 (0-4)	0.043
	Women	2.3 (1-15)	2.11 (1-4)	1.7 (1-3)	1.3 (0-3)	0.384
	Children 13-17	0.20 (0-1)	0.9 (0-3)	0.2 (0-1)	0.2 (0-1)	0.002
	Children 5-12	1.5 (0-13)	1.4 (0-6)	0.7 (0-4)	0.9 (0-3)	0.593
	Children under 5	1.0 (0-3)	1.0 (0-2)	1.1 (0-4)	1.3 (0-3)	0.785
Wealth Index		-0.35 (-1.57, 1.07)	0.19 (-0.90, 1.88)	0.30 (-1.3, 1.67)	-0.48 (-1.82, 1.67)	0.031
Latrine Access	Private latrine	1 (5.6)	0 (0.0)	1 (5.3)	3 (17.7)	0.023
	Shared compound Latrine	8 (44.4)	3 (17.7)	1 (5.3)	5 (29.4)	
	Public Latrine	9 (50.0)	14 (82.4)	17 (89.5)	9 (52.9)	
Water Source	Sachet	11 (61.1)	12 (70.6)	16 (84.2)	12 (70.6)	0.47
	Tap from pipe	7 (38.9)	5 (29.4)	3 (15.8)	5 (29.4)	

Table 3. Household indicators of child and animal fecal disposal by level of sanitation access for 71 households in Accra, Ghana.

		Private latrine in compound (n=5)	Shared latrine in compound (n=17)	No latrine/ public latrine access (n=49)
Where children 5-12 defecate	In latrine	1 (20.0)	10 (58.8)	17 (34.7)
	Other	0 (0.0)	1 (5.9)	10 (20.4)
	No child	4 (80.0)	6 (42.9)	22 (47.8)
Defecation location of children under 5	Potty/ diaper/ latrine	3 (60.0)	12 (70.6)	39 (79.6)
	Ground or drain	0.0 (0)	0.0 (0.0)	2 (4.1)
	No child	2 (40.0)	5 (29.4)	8 (16.3)
Feces disposal of children under 5 years	Latrine	1 (20.0)	11 (64.7)	6 (12.2)
	Rubbish or drain	1 (20.0)	1 (5.9)	35 (71.4)
	Ground	0 (0.0)	0 (0.0)	1 (2.0)
	No child	3 (60.0)	5 (29.4)	7 (14.3)
Feces Observed		0 (0.0)	1 (5.9)	7 (14.3)
Animal Presence		2 (40.0)	11 (64.7)	23 (46.9)

Table 4. Distribution of *E. coli* and Adenovirus in swab and hand rinse samples.

	n	*E. coli* + (%)	*E. coli* Log Mean (Median)	Adenovirus + (%)	Adenovirus Log Mean (Median)
Swabs	94	81 (86.1)	2.0 (2.1)	22 (23.9)	3.13 (3.0)
Hand Rinse	80	71 (88.8)	2.5 (2.6)		

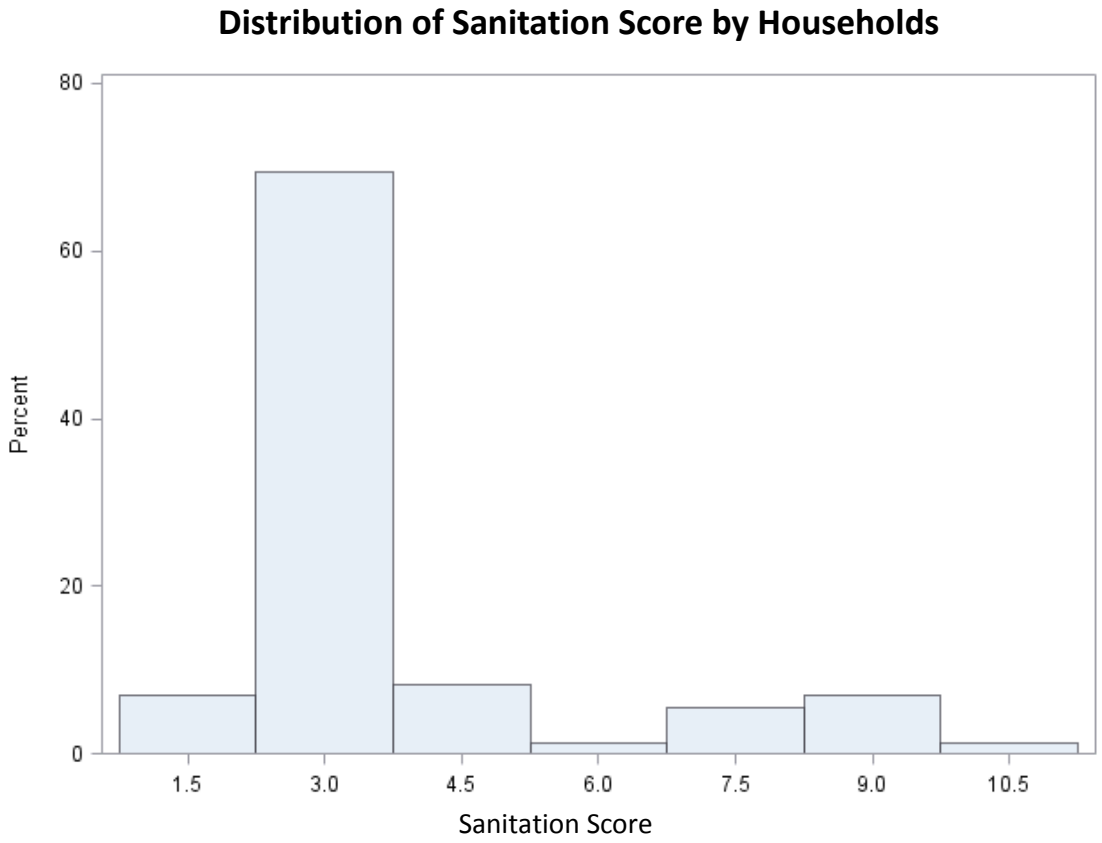

Figure 1. Distribution of the Sanitation Score based upon access level, use of facilities for feces disposal, and observed sanitary conditions in the household.

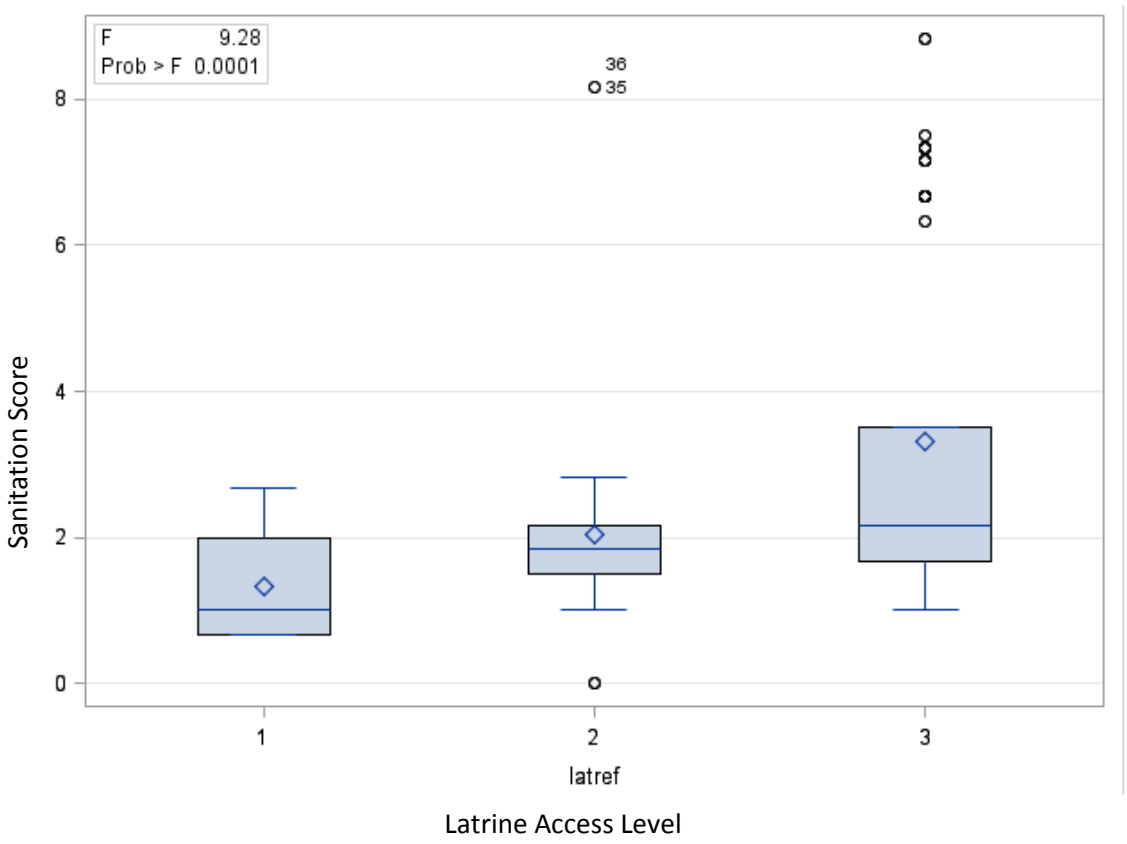

Figure 2. Comparison by latrine access level of sanitation score accounting only for child defecation practices and observed feces.

Table 5. Hierarchical logistic regression models of sanitation score and presence or absence of *E. coli* or Adenovirus as a measure of fecal contamination, adjusted for other household covariates.

	Unadjusted Effects		Adjusted Effects	
Variable	OR	CI	OR	CI
Presence of *E. coli* in Hand Rinse Samples[1]				
Score	1.1	0.8, 1.6	1.1	0.8, 1.6
Tenancy Status	2.6	0.5, 14.4		
Education Level	0.6	0.2, 1.7		
Christian vs Other Religion	1.4	0.2, 9.4		
Wealth Index	0.8	0.3, 2.2		
Animal presence	0.9	0.2, 5.5		
Sachet water vs piped water	0.8	0.1, 4.9		
Presence of *E. coli* in Swab Samples[2]				
Score	1.1	0.9, 1.5	1.1	0.8, 1.4
Tenancy Status	2.1	0.5, 7.9		
Education Level	1.4	0.7, 2.9		
Christian vs Other Religion	2.1	0.3, 14.7		
Wealth Index	1.1	0.6, 2.3		
Animal presence	3.1	0.8, 11.6	2.7	0.6, 11.6
Sachet water vs piped water	1.5	0.3, 7.9		
Presence of Adenovirus in Swab Samples[3]				
Score	1.1	0.8, 1.3	1.0	0.8, 1.3
Tenancy Status	0.9	0.3, 2.7		
Education Level	0.9	0.5, 1.6		
Christian vs Other Religion	0.8	0.2, 3.1		
Wealth Index	0.8	0.5, 1.3		
Animal presence	1.2	0.4, 3.7		
Sachet water vs piped water	0.4	0.1, 1.7	0.4	0.8, 1.3

[1]n= 80 samples, 47 households. [2, 3] n=94 samples, 54 households. OR= odds ratio; CI= 95% confidence interval. Italics refer to variables significant at $p<0.2$. Bold refers to variables significant at $p<0.05$ after adjustment.

Table 6. Hierarchical linear regression models of sanitation score and log concentration of *E. coli* or Adenovirus as a measure of fecal contamination, adjusted for other household covariates.

	Unadjusted Effects		Adjusted Effects	
Variable	PR	CI	PR	CI
Concentration of *E. coli* in Hand Rinses[1]				
Score	1.0	0.9, 1.2	1.0	0.9, 1.2
Tenancy Status	0.9	0.5, 1.7		
Education Level	1.2	0.9, 1.6		
Christian vs Other Religion	2.2	1.2, 4.0	**2.2**	**1.2, 4.0**
Wealth Index	1.1	0.8, 1.5		
Animal presence	0.6	0.4, 1.2		
Sachet water vs piped water	0.7	0.4, 1.3		
Concentration of *E. coli* in Swab Samples[2]				
Score	1.0	0.9, 1.1	1.04	1.0, 1.1
Tenancy Status	1.1	0.7, 1.7		
Education Level	0.9	0.7, 1.1		
Christian vs Other Religion	0.9	0.5, 1.5		
Wealth Index	1.0	0.8, 1.2		
Animal presence	0.5	0.3, 0.8	**0.5**	**0.3, 0.7**
Sachet water vs piped water	0.8	0.5, 1.3		
Concentration of Adenovirus in Swab Samples[3]				
Score	1.2	1.1, 1.4	**1.6**	**1.1, 2.2**
Tenancy Status	1.5	0.8, 3.0	1.5	0.9, 2.9
Education Level	1.0	0.7, 1.4		
Christian vs Other Religion	1.2	0.5, 2.7		
Wealth Index	1.0	1.0, 1.1		
Animal presence	1.5	0.8, 3.0		
Sachet water vs piped water	1.3	0.6, 2.8		

[1] n=80 samples, 76 households. [2,3] n=94 samples, 54 households.PR= prevalence ratio; CI= 95% confidence interval. Italics refer to variables significant at $p<0.2$. Bold refers to variables significant at $p<0.05$ after adjustment.

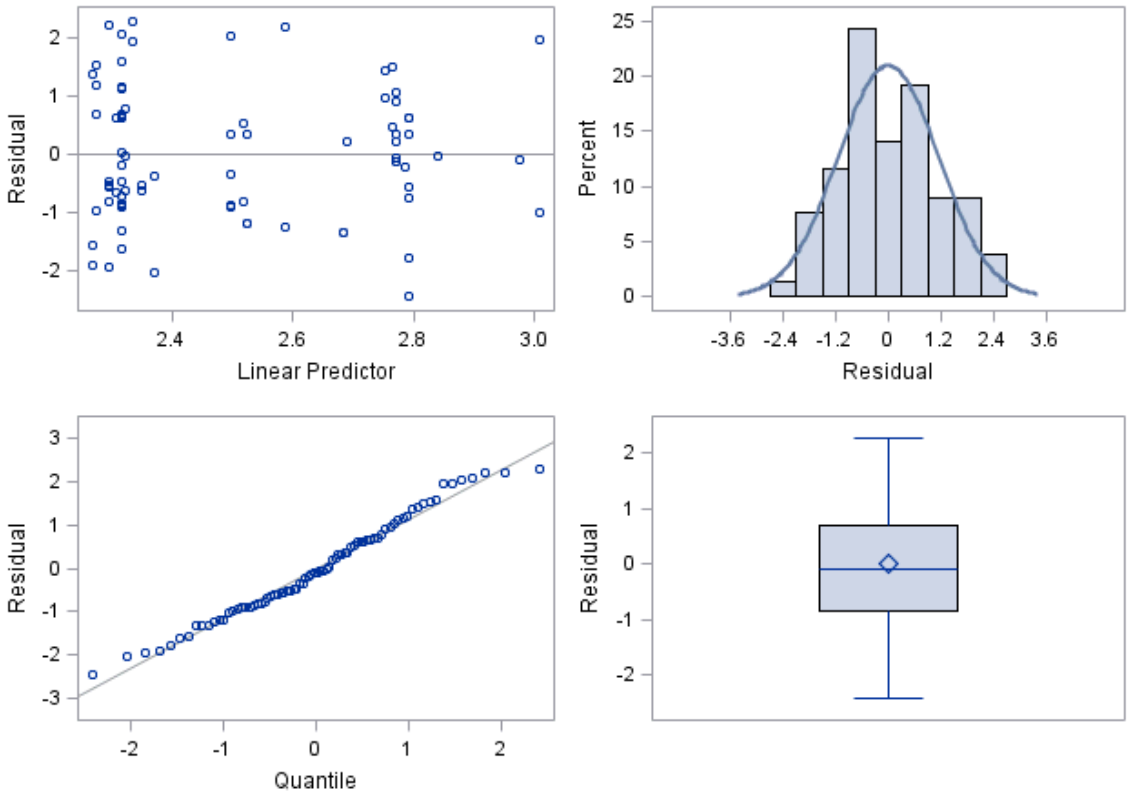

Figure 3. Residual fit of *E. coli* logarithmic concentration and sanitation score in hand rinse samples.

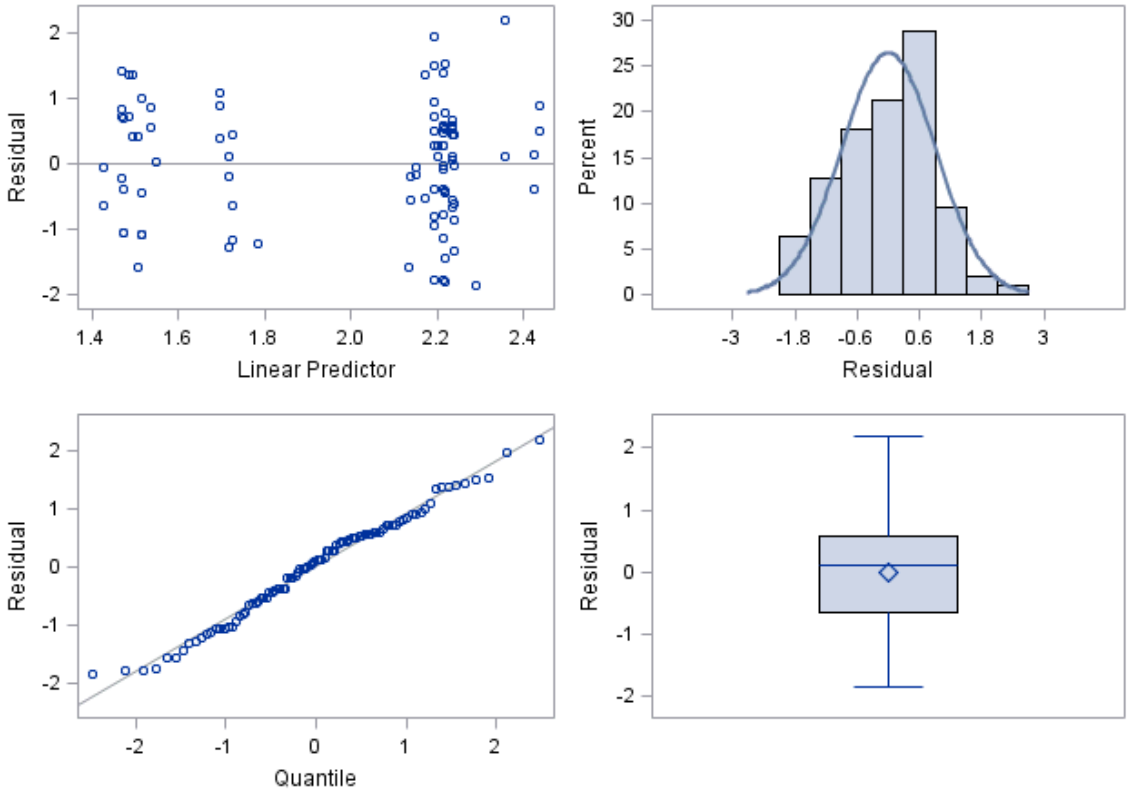

Figure 4. Residual fit of *E. coli* logarithmic concentration and sanitation score in swab samples.

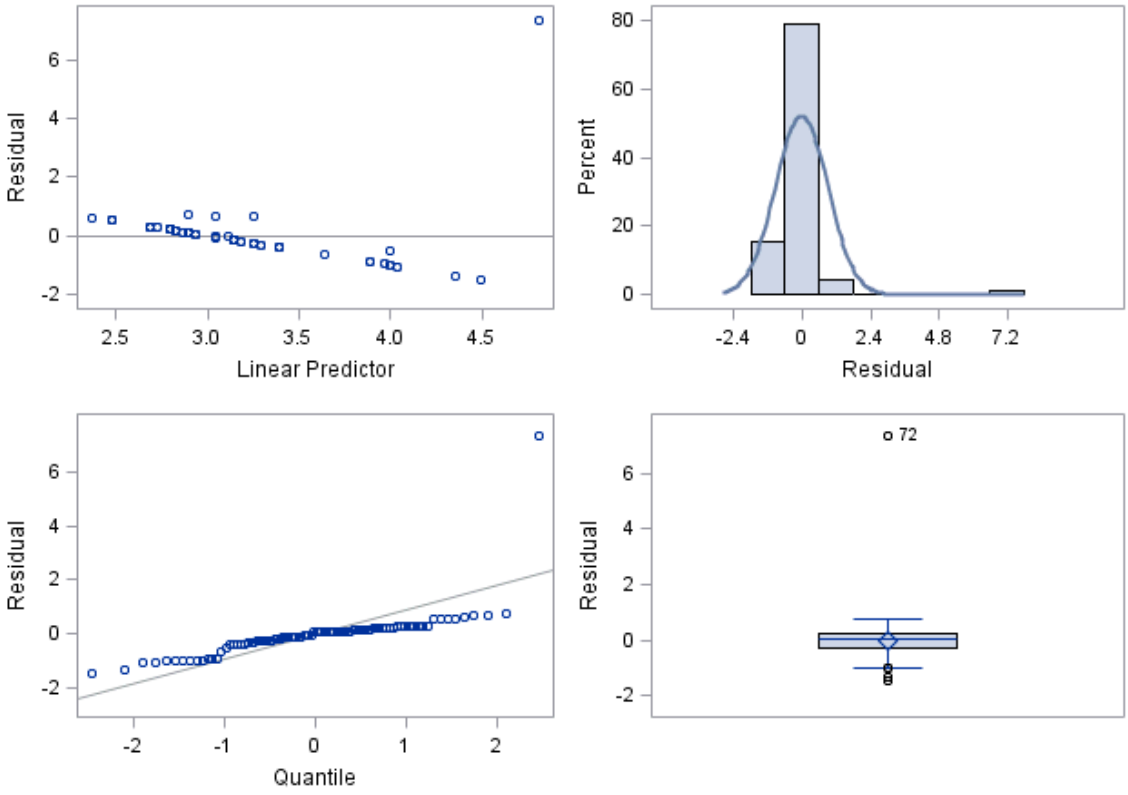

Figure 5. Residual fit of Adenovirus logarithmic concentration and sanitation score in swab samples.

Works Cited

1 Boadi, K. O. & Kuitunen, M. Environmental and health impacts of household solid waste handling and disposal practices in Third World cities: The case of the Accra Metropolitan Area, Ghana. *Journal of environmental health* **68**, 32-36 (2005).

2 Walker, C. L. *et al.* Global burden of childhood pneumonia and diarrhoea. *Lancet* **381**, 1405-1416, doi:10.1016/S0140-6736(13)60222-6 (2013).

3 Fuller, J. A., Clasen, T., Heijnen, M. & Eisenberg, J. N. Shared sanitation and the prevalence of diarrhea in young children: evidence from 51 countries, 2001-2011. *The American journal of tropical medicine and hygiene* **91**, 173-180, doi:10.4269/ajtmh.13-0503 (2014).

4 Exley, J. L., Liseka, B., Cumming, O. & Ensink, J. H. The Sanitation Ladder, What Constitutes an Improved Form of Sanitation? *Environmental science & technology*, doi:10.1021/es503945x (2015).

5 Organization, W. H. *Excreta Disposal Options* (ed Water Sanitation Health) (World Health Organization, World Health Organization).

6 Clasen, T. F. *et al.* Interventions to improve disposal of human excreta for preventing diarrhoea. *The Cochrane database of systematic reviews*, CD007180, doi:10.1002/14651858.CD007180.pub2 (2010).

7 Barreto, M. L. *et al.* Effect of city-wide sanitation programme on reduction in rate of childhood diarrhoea in northeast Brazil: assessment by two cohort studies. *Lancet* **370**, 1622-1628, doi:10.1016/S0140-6736(07)61638-9 (2007).

8 Golding, J., Greenwood, R., Mccawbinns, A. & Thomas, P. Associations between Social and Environmental-Factors and Perinatal-Mortality in Jamaica. *Paediatr Perinat Ep* **8**, 17-39, doi:DOI 10.1111/j.1365-3016.1994.tb00489.x (1994).

9 Montgomery, M. A., Desai, M. M. & Elimelech, M. Comparing the effectiveness of shared versus private latrines in preventing trachoma in rural Tanzania. *The American journal of tropical medicine and hygiene* **82**, 693-695, doi:10.4269/ajtmh.2010.09-0540 (2010).

10 Vujcic, J. *et al.* Toys and toilets: cross-sectional study using children's toys to evaluate environmental faecal contamination in rural Bangladeshi households with different sanitation facilities and practices. *Trop Med Int Health* **19**, 528-536, doi:10.1111/tmi.12292 (2014).

11 Shultz, A. *et al.* Cholera Outbreak in Kenyan Refugee Camp: Risk Factors for Illness and Importance of Sanitation. *American Journal of Tropical Medicine and Hygiene* **80**, 640-645 (2009).

12 Godana, W. & Mengistie, B. Determinants of acute diarrhoea among children under five years of age in Derashe District, Southern Ethiopia. *Rural and remote health* **13** (2013).

13 Levy, K., Nelson, K. L., Hubbard, A. & Eisenberg, J. N. Rethinking indicators of microbial drinking water quality for health studies in tropical developing countries: case study in northern coastal Ecuador. *The American journal of tropical medicine and hygiene* **86**, 499-507, doi:10.4269/ajtmh.2012.11-0263 (2012).

14 Pickering, A. J. *et al.* Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian households with and without improved sanitation. *Environmental science & technology* **46**, 5736-5743, doi:10.1021/es300022c (2012).

15 Barnard, E. E., Frost, E. B., Calvert, M. R. & Dobek, G. O. *A photographic atlas of selected regions of the Milky Way*. First paperback edition. edn, (Cambridge University Press, 2014).

16 Biran, A., Tabyshalieva, A. & Salmorbekova, Z. Formative research for hygiene promotion in Kyrgyzstan. *Health policy and planning* **20**, 213-221, doi:10.1093/heapol/czi024 (2005).
17 Team, T. W.-I. W. *Final Report on the Women's Health Study of Accra, Wave II*. (Institute for Statistical, Social and Economic Research, University of Ghana, 2011).
18 in *Water Aid* (ed Water Aid Ghana Office) (Water Aid, Accra, Ghana).
19 Biran, A., Jenkins, M. W., Dabrase, P. & Bhagwat, I. Patterns and determinants of communal latrine usage in urban poverty pockets in Bhopal, India. *Trop Med Int Health* **16**, 854-862, doi:10.1111/j.1365-3156.2011.02764.x (2011).
20 Lau, J. *Designing sanitation projects in rural Ghana*, Massachusetts Institute of Technology, (2011).
21 (eds UNICEF & World Health Organization) (World Health Organization, Geneva, 2012).
22 Jothikumar, N. *et al.* Quantitative real-time PCR assays for detection of human adenoviruses and identification of serotypes 40 and 41. *Applied and environmental microbiology* **71**, 3131-3136, doi:10.1128/AEM.71.6.3131-3136.2005 (2005).
23 Staley, C., Gordon, K. V., Schoen, M. E. & Harwood, V. J. Performance of two quantitative PCR methods for microbial source tracking of human sewage and implications for microbial risk assessment in recreational waters. *Applied and environmental microbiology* **78**, 7317-7326, doi:10.1128/AEM.01430-12 (2012).
24 Mattioli, M. C., Pickering, A. J., Gilsdorf, R. J., Davis, J. & Boehm, A. B. Hands and water as vectors of diarrheal pathogens in Bagamoyo, Tanzania. *Environmental science & technology* **47**, 355-363, doi:10.1021/es303878d (2013).
25 Weitzel, T. *et al.* Field evaluation of a rota- and adenovirus immunochromatographic assay using stool samples from children with acute diarrhea in Ghana. *Journal of clinical microbiology* **45**, 2695-2697, doi:10.1128/JCM.00562-07 (2007).
26 Williamson, W. M. *et al.* Enteric viruses in New Zealand drinking-water sources. *Water science and technology : a journal of the International Association on Water Pollution Research* **63**, 1744-1751 (2011).
27 Liu, P. *et al.* Assessment of human adenovirus removal by qPCR in an advanced water reclamation plant in Georgia, USA. *Journal of applied microbiology* **115**, 310-318, doi:10.1111/jam.12237 (2013).
28 Wu, J., Gronewold, A. D., Rodriguez, R. A., Stewart, J. R. & Sobsey, M. D. Integrating quantitative PCR and Bayesian statistics in quantifying human adenoviruses in small volumes of source water. *The Science of the total environment* **470-471**, 255-262, doi:10.1016/j.scitotenv.2013.09.026 (2014).

Appendix A
Assessing Sanitation-related Exposure: Fecal Indicators

Multiple types of bacteria and viruses can survive and replicate in the human digestive tract. While many of these organisms are harmless to human health, some cause diarrheal diseases. These organisms can persist in the environment outside of the host and go on to infect new people. Some pathogens are specific to humans, while others cause disease in a broad range of mammals. Preventing spread of diarrheal disease often relies on an understanding of the spread of these organisms in the environment. *E. coli* is the most commonly used pathogen to track fecal contamination of the environment. *E. coli* is used as a general bacterial measure of contamination, representing groups of fecal bacteria found in both humans and animals[4,14,23,24]. *E. coli* are common gut microbes of most mammalian species, and presence of *E. coli* in the environment could indicate exposure risk for diarrheal diseases from either human or animal feces, especially in sanitation-poor environments[14,24]. Certain strains of enterovirus and adenovirus can be used to identify human-specific fecal contamination presence[10,24,25]. Human adenovirus 40/41 can be used as an indicator of contamination from human feces, as this strain has typically been isolated from just humans. The presence of human adenovirus in environmental samples suggests that human waste has been introduced. Detection of these organisms together in environmental and hand samples can be used to distinguish exposure to human versus general mammalian excrement[24,25].

Past studies have detected combinations of bacterial and viral organisms to measure the risk of exposure to fecal pathogens in an environment. A study done in New Zealand used PCR to test for norovirus, enterovirus, adenovirus, rotavirus, and hepatitis E virus in order to assess ground source drinking water safety and found that rotavirus,

adenovirus, and norovirus were the most common in water samples[26]. While this study examined many viruses to assess water quality, other studies have used only adenovirus presence to represent all viral content of in the water[27,28]. One study measured the adenovirus DNA found in wastewater after treatment, as a proxy for viral content in treated water sources and found that the virus was not entirely removed during treatment, although infectivity was not tested[27]. One study tested water sources for *E. coli,* enterococci, and somatic colophage, and found that only *E. coli* presence was associated with diarrhea in a village in Ecuador, although this could have been due to testing limitations for enterococci and somatic colophage[13]. An environmental study used *E. coli* and enterovirus in soil samples taken from both latrine floors and household floors in order to monitor sanitation improvement in households with a pit latrine. This study found that *E. coli* and enterovirus concentration was higher in the household and near food preparation areas than within the actual latrine. This suggests that a simple pit latrine is not the major source of environmental pathogens in these households[14]. A study done in Tanzania evaluated latrine safety between shared and private latrines using soil samples to detect *E. coli*, helminthes, enteroviruses, and *Cryptosporidium* to estimate exposure to disease causing organisms. Latrines were categorized as "improved", "shared", and "unimproved". *E. coli* concentration was found to be highest overall in latrines that were classified as "unimproved", but in shared latrines, the concentration of *E. coli* decreased as the number of users increased. The study also found that higher concentrations of *E. coli* bacteria could be detected during the hotter season, which is historically associated with higher diarrhea[4]. Past studies have demonstrated both the

uses and limitations of using different fecal indicators in understanding environmental exposure to disease causing pathogens.

Appendix B
Principal Components Analysis- Wealth Index Variable

Principal components analysis of 8 assets was used to create a wealth index variable for each household in the dataset, which was intended to summarize the household possessions that are indicative of level of poverty. The assets used in the analysis were electricity in the household, possession of a radio, television, refrigerator, bicycle, motorcycle, car, or employment of an unrelated domestic worker. After running the initial analysis, only factors that explained at least 10% of variance were included (Table D1). Four factors met this criterion and were retained. The scree plot of the data also suggests that the first four factors are contributing the most to explain the variation, as the plot begins to plateau between factors 4 and 5 (Figure D1). These four factors are able to explain 74% of the total variance in household wealth. This analysis is problematic, however, as it distributes wealth across neighborhoods much differently than previous studies. The neighborhood of Old Fadama under this analysis is shown to have a higher level of wealth in comparison to other neighborhoods, which greatly differs from what has been found in previous studies of these neighborhoods (Baker). It is possible that the assets that were available for inclusion in the principal components analysis of wealth were not sufficient to accurately differentiate the wealth of households in this area. While this study is limited in assessment of the association between environmental and hand sanitation and wealth, there is also collinearity between wealth and latrine access, so some of the variation would be accounted for through the inclusion of the scale variable.

Appendix C
GLIMMIX Procedure

Example 1: Logistic Regression Model of Adenovirus Presence and Household Characteristics:

```
proc glimmix data=scalehhswab ;
    class  c201;
    model ad_pos1 = stanscale c201 / dist=binary solution oddsratio;
    random _residual_ / subject= locid type=cs;
    run;
```

The procedure PROC GLIMMIX was used to analyze this data in SAS 9.3. Categorical variables were named in the CLASS statement. The MODEL statement specifies dist=binary. This specifies the probability distribution of the data to a binary distribution. The solution statement is used to get the fixed-effects parameter estimates in the output, and the odds ratio statement causes the odds ratios estimates to be given. The RANDOM statement specifies the random effects in the model. The _residual_ keyword in the model stipulates that the variable after the backslash will define the random effects matrix; in this case, it specifies that the location id of the samples make up the random effects, which accounts for the clustering of samples. The TYPE = CS specifies that the compound symmetry structure be used.

Example 2: Linear Regression Model of *E. coli* Concentration and Household Characteristics:

```
proc sort data=scalehhswab;
    by locid;
    run;

proc glimmix data=scalehhswab;
    class  ;
    model ec_log= stanscale c507 / dist=normal solution ;
    random _residual_/ subject= locid type=cs;
    nloptions tech=nrridg;
    run;
```

For linear regression, the CLASS, MODEL and RANDOM statements are used in the same way as previously described. Because *E. coli* concentration is a continuous variable, the distribution is specified as normal. NLOPTIONS is used to optimize the data when there are nonlinear parameters. TECH=NRRIDG specifies that the technique for optimization used is the Newton-Raphson Ridge Optimization.

Appendix D
Data Collection- Survey Questions

Questions in this survey were asked of a caregiver in the household:

1.	Tenancy Status	1) Rent 2) Own
2.	What is your highest level of education?	1) No formal education 2) Some primary 3) Completed primary 4) Completed secondary 5) Higher than secondary 6) No response
3.	What is your religion?	1) Christian 2) Moslem 3) Traditionalist/Spiritualist 4) No religion 5) Other 6) No response
4.	Is a business run from this compound/ household?	1) Yes 2) No 3) No response
5.	How many households are in this compound? A household is people sharing a cooking pot.	
6.	How many people live in this compound?	
7.	How many people are in your household?	
8.	How many male adults (18 and older) live in your household?	
9.	How many female adults (18 and older live in your household?	
10.	How many young people ages 13-17 live in your household?	
11.	How many children ages 5-12 live in your household?	
12.	How many children under 5 live in your household?	
13.	What is your primary source of drinking water?	1) Sachet/ water bottle 2) Tap from pipe network 3) Tap from polytank 4) Well 5) Water truck 6) Harvested rainwater 7) Other specify: 8) Don't know/ no response
14.	How many latrines are on this compound?	

15.	How many households do you share a latrine with?	
16.	Where do children (ages 5-12) in your household typically defecate?	1) Compound latrine 2) Public latrine 3) In bag/ flying toilet 4) Chamber pot 5) Outside 6) Beach 7) Outside 8) Other 9) Don't know/ no response
17.	(If single household then skip) Is this the same for other children in this compound?	1) Yes 2) No 3) Don't know/ no response
18.	(If respondent has a child under 5) The last time your youngest child defecated, where did they defecate?	1) Compound latrine 2) Public latrine 3) In a potty 4) In a diaper/ nappy 5) On ground/ outside compound 6) On ground/ inside compound 7) In drain/ gutter 8) Don't know/ no response
19.	The last time your youngest child defecated, how did you dispose of feces?	1) Combined with rubbish 2) Dumped into drain/ gutter 3) Dumped into public latrine 4) Dumped into compound latrine 5) Washed diaper/ nappy 6) Left on ground 7) Don't know/ no response
20.	Do other mother's in the compound ever use potties for their children?	1) Yes 2) No 3) Don't know/ no response
21.	What kind of latrine do you have in this compound?	1) No facility/ bush/ field 2) Traditional pit latrine 3) KVIP (double) 4) VIP (single) 5) Bucket/ pan 6) Pour flush 7) Flush toilet 8) Other 9) No response
22.	(observe, don't ask) Are feces observed around compound grounds?	1) Yes 2) No

23.	(observe, don't ask) Are animals observed roaming around compound grounds?	1) Yes 2) No

Table D1. Eigenvalues of PCA

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.587	1.237	0.3233	0.323
2	1.349	0.239	0.169	0.492
3	1.110	0.219	0.139	0.631
4	0.891	0.199	0.111	0.742
5	0.693	0.091	0.087	0.829
6	0.602	0.130	0.075	0.904
7	0.472	0.176	0.059	0.963
8	0.296		0.037	1.000

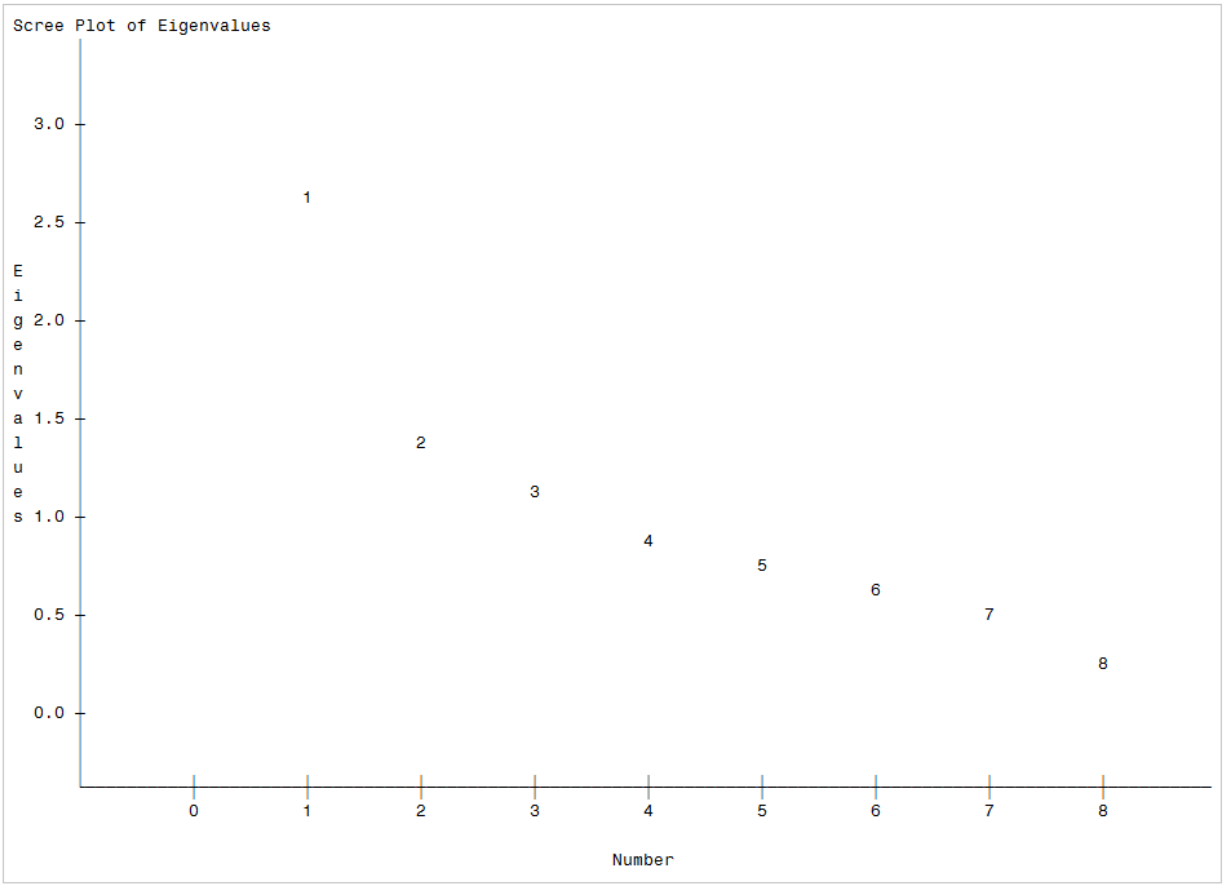

Figure D1. Scree plot of PCA Wealth

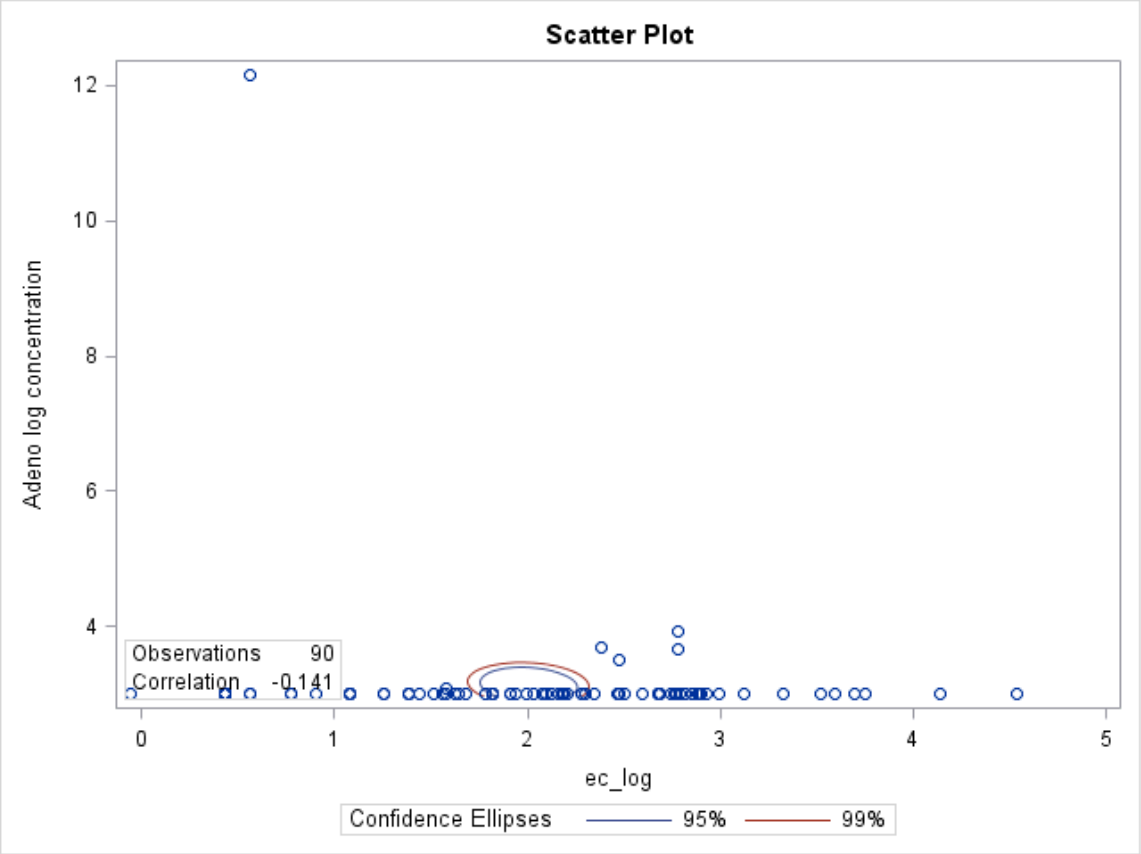

Figure D2. Scatterplot of logarithmically transformed *E. coli* concentration and logarithmically transformed Adenovirus concentration of all samples.