Distribution Agreement

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

Signature:

________________________________ ________________

Alejandra Alvarez Date

Impact of Water Quality Variability on Associations Between Water Quality and Water Sample Characteristics, Diarrhea Outcome, and Helminth Infection, in Northern Coastal Ecuador

By

Alejandra Alvarez

Master of Public Health

Epidemiology

Karen Levy, PhD, MPH

Committee Chair

Impact of Water Quality Variability on Associations Between Water Quality and Water Sample Characteristics, Diarrhea Outcome, and Helminth Infection, in Northern Coastal Ecuador

By

Alejandra Alvarez

B.A.

Haverford College

2016

Faculty Thesis Advisor: Karen Levy, PhD, MPH

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2020

Abstract

Impact of Water Quality Variability on Associations Between Water Quality and Water Sample Characteristics, Diarrhea Outcome, and Helminth Infection, in Northern Coastal Ecuador

Background: Drinking water quality is vital for good health and affects health outcomes such as diarrhea and soil transmitted helminth infection. Spatiotemporal factors are known to affect water quality variability. Not as much research has been done on the variability effect of using categorical vs continuous water quality data in assessing health outcomes.

Methods: From the full dataset, randomly dropped, randomly kept and average datasets were created. Furthermore, a categorical variable for water quality was created in each dataset. Mixed effects linear and logistic regression were used to estimate associations to account for clustering by household for the households where there was more than one sample collected. Simple linear and logistic regression were used for the datasets that only had one sample per household.

Results: A total of 162 households were visited for this study Diarrhea occurred in 23% of households. Among urban households, the prevalence of diarrhea was 20.3% compared to 17.9% in rural areas. There was an association between continuous EC and diarrhea outcome in the average dataset. (OR = 1.07, 95% CI = 1.003 - 1.413). Alternatively, there was an association between categorical EC and diarrhea in the randomly kept dataset (OR = 4.659, 95% CI = 1.067 – 20.337). The prevalence of helminth infection in the entire dataset was 15.4% . Among rural households, the prevalence was 7.41% and among urban households, the prevalence was 24.4%. There was an association between continuous water quality and helminth infection in the randomly kept (OR = 1.512, 95% CI = 1.130 – 2.023) and average (OR = 1.869, 95% CI = 1.216 – 2.873) datasets. There was an association between categorical water quality and helminth in the average (OR = 5.921, 95% CI =1.144 – 30.652) and randomly kept (OR = 7.424, 95% CI = 1.859 – 29.652) datasets.

Conclusions: Datasets with one measurement for every household tend to result in statistically significant associations between water quality and health outcome. Furthermore, using categorical water quality variable may result in concluding that there is a stronger association than there actually is.

Impact of Water Quality Variability on Associations Between Water Quality and Water Sample Characteristics, Diarrhea Outcome, and Helminth Infection, in Northern Coastal Ecuador

By

Alejandra Alvarez

B.A.

Haverford College

2016

Faculty Thesis Advisor: Karen Levy, PhD, MPH

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2020

Acknowledgements

Foremost, I would like to thank my thesis advisor, Dr. Karen Levy, for her guidance throughout the thesis process. Secondly, I would like to thank Dr. Gwen Lee for providing data analysis suggestions. I am also grateful to all my friends at Rollins who have always encouraged and motivated me. Thank you all for always providing feedback and suggestions, as well as an open ear. This project could not have been completed without the hard work of Rollins Alumna, Dr. Aya Fanny, and her involvement in the ECUAVIDA study. Thank you, Dr. Fanny, and all those involved in the ECUAVIDA study. Finally, I would like to thank my friends and family back home for believing in me and supporting me from across the country.

Table of Contents

Background ………………………………………………………..……...............1
  Burden of disease ………………………………………………...……...1
  Association between water quality and diarrhea ……………………………...1
  Association between water quality and helminth infection …………………...2
  Water quality variability ……………………………………………………..3
Methods ………………………..…………………………………………………7
  Introduction ……………………………………………………………...7
  Population and sample ……………………………………………………7
  Research design ……………………………………………………………8
  Procedures …………………………………………………………………9
  Data analysis ……………………………………………….……………...10
  Ethical consideration ……………………………………………………..11
Results ………………………………………...…………………………...12
  Descriptive statistics ………………………………………………………15
  Association between water quality and sample characteristics ……………..16
  Association between water quality and health outcomes …..………………..20
Discussion ……………………………………..………………………..22
  Limitations …..…………………………………………………………..25
  Future directions ……………………………………………………….…26
Tables and Figures ………………………………………………………………..29
References ……………………………………………………………………42

Background

Burden of disease

Drinking water quality is vital for good health and affects health outcomes such as diarrhea and soil transmitted helminth (STH) infection. One systemic analysis aimed at quantifying the global and regional diarrhea disease burden estimated that diarrhea accounted for 1,655,944 deaths in 2016, making it the 8th leading cause of death globally for all age groups and 5th leading cause of death for children under 5 (228,057 deaths for children under 5)(1). Additionally, unsafe water was responsible for 72.1% (95% uncertainty interval: (UI) 68.2-85.0%) of diarrhea deaths for children under 5 (1). Additionally, water may be a “vehicle of [helminth] transmission” when it is contaminated with STH ova (2). Around the world, over 464 million people were infected with *Trichuris trichiura* and 819 million people with *Ascaris lumbricoides* in 2010. It is estimated that roughly 1.74 million years lived with disability are attributable to *T. trichiura* and *A. lumbricoides* (3).

Association between water quality and diarrhea

There are mixed conclusions for results finding associations between water quality and diarrhea outcome. Several systematic reviews and meta-analysis have found inconsistent results concerning the relationship between water quality and diarrhea outcome. One meta-analysis comparing categorical coliform exposure in household drinking water samples to diarrhea outcome found that roughly half of the selected studies reported that fecal contamination exposure was associated with an increased risk of diarrhea while the remaining studies concluded no risk or a reduced diarrhea risk with fecal contamination exposure (4). Additionally, another study focused on the relationship between water quality and health outcomes did not find a relationship between water quality and diarrhea outcome (pooled odds ratio (OR) = 1.12, 95% confidence interval (CI) = 0.85 – 1.48) (5). Therefore, due to the inconsistent associations between water quality and health outcomes, it is important to further investigate the relationship between *E. coli* contamination and diarrhea outcome.

One important factor to consider for water quality exposure is water treatment. While there are mixed conclusions regarding water quality and diarrhea, it is a common trend that water treated with boiling water has a protective effect on diarrhea outcome. One meta-analysis focusing on the effects of boiling drinking water and health outcomes extracted data from 27 articles of studies around the world. This meta-analysis found a protective effect of boiling water on nonspecific diarrheal outcomes (pooled OR = 0.83, 95% CI = 0.7-0.96, N=4) (6). Alternatively, one randomized controlled trial in Western Kenya that focused on assessing the effect of ceramic water filters on diarrhea prevention found that households in the intervention arm reported less diarrhea (OR = 0.86, 95% CI = 0.64 – 1.16) but did report significantly fewer health facility visits for diarrhea (OR=0.5, 95% CI = 0.30 – 0.83) (7). While different water treatments appear to produce varying levels of protection against diarrhea, perhaps the association between *E. coli* contamination and diarrhea outcome should consider also consider water treatment.

Association between water quality and helminth infection.

In addition to water quality affecting diarrhea outcomes, water quality has also been seen to affect presence of helminth infection in populations around the world. A cross-sectional study in Guatemala found that *E. coli* contaminated water was associated with a five times increased odds of STH infection (OR = 5.14, CI= 1.08 – 24.27, p-value = 0.04) (8). Additionally, the authors suggest that the effect of water treatment on helminth infection should be further investigated. Similarly, a cluster randomized control trial conducted in rural Kenya investigated the effect of different water treatment and handwashing interventions on the presence of helminth infection. This study found that the water treatment arm, which included encouraging chlorine treatment or using manual dispensers, had a lower prevalence of *A. lumbricoides* (Prevalence ratio (PR) = 0.82, 95% CI = 0.67 – 1.00) compared to the control arm (9).

Similar to the association between water quality and diarrhea outcome, not all studies find a significant association between water quality and helminth infection. The previously mentioned meta-analysis found that among the articles reporting on the association between Ascaris and boiling water, one study which took place in Sri Lanka showed a protective effect (OR = 0.33, 95% CI = 0.11-0.93), while the other study which took place in Cuba, did not (OR = 4.35, 95% CI = 1/40, 13.46) which led the authors to conclude that according to the pooled estimate (OR = 1.18, 95% CI = 0.09, 14.94, N=2) there wasn't sufficient evidence to demonstrate an effect of boiling water on helminth infection outcome (6). Therefore, strengths and confidence surrounding associations between water quality, especially after treatment, and helminth infection vary by region in the world.

Water quality variability

Factors such as landscape and time of year have been found to be associated with variability in water quality. One case study of the Mun River Basin in Thailand found that agriculture activity directly affects water pollution. Additionally, the authors found a significant difference (Pearson correlation = 0.738, $p<0.01$) in the spatial pattern of the soil nutrients when comparing the dry and rainy seasons, which affected water quality as well in that the dry season had better water quality than the rainy season (10). It is understood that different geographic locations and the activities done there may affect the immediate environment including water quality. Seasonality was also seen as a factor in water quality variability in a northern coastal Ecuador study where *E. coli* counts in the wet season were higher than the *E. coli* counts in the dry season (difference = 0.42 log) (11).

Water quality does not only vary seasonally but can also vary in short amounts of time. One study found that coastal water samples also vary significantly over a short time scale and this author also emphasizes the importance of taking multiple samples in order to better understand the microbial contamination of the water collection source and to better determine the associated health risk (12). These spatiotemporal factors were also observed in a study conducted in northern coastal Ecuador where more water quality variability was observed on an hourly basis compared to on a daily or weekly basis (11).

More recently, a meta-analysis compiling individual participant data identified one source of potential exposure assessment error as number of water quality samples. Their simulation showed that there was a difference in water quality contamination scores among the wet and dry seasons when using a median of multiple samples against using a single sample (13). Therefore, it is important to take into consideration the variation of multiple water samples as well as being mindful of how many water samples are taken when later analyzing water quality data.

In terms of drinking water quality, water source may be a factor in water contamination. One meta-analysis assessing fecal contamination of drinking water in low- and middle-income countries found that it was less likely for improved sources of drinking water to be contaminated (pooled OR = 0.15, 95% CI = 0.10- 0.21). However, the authors note that there was high heterogeneity (I_2=80.3% [95% CI 72.9–85.6]) and therefore setting of the research study should be taken into consideration. Additionally, while the pooled estimate is significant, the authors remind the readers that due to some ORs being greater than 1, improved water sources are not always effective in some settings (14).

This project recognizes the impacts different drinking water characteristics have on water quality, the effects of water quality has on health outcomes as well as the factors that are associated with water quality variability. While many meta-analysis compare associations between water quality and health outcomes, these analysis do not compare how using a continuous or categorical exposure may affect the measure of association. Furthermore, this project aims to evaluate how using different sets of observations affect these associations to more clearly understand the relationships between water quality and water characteristics and health outcomes.

Methods

Introduction

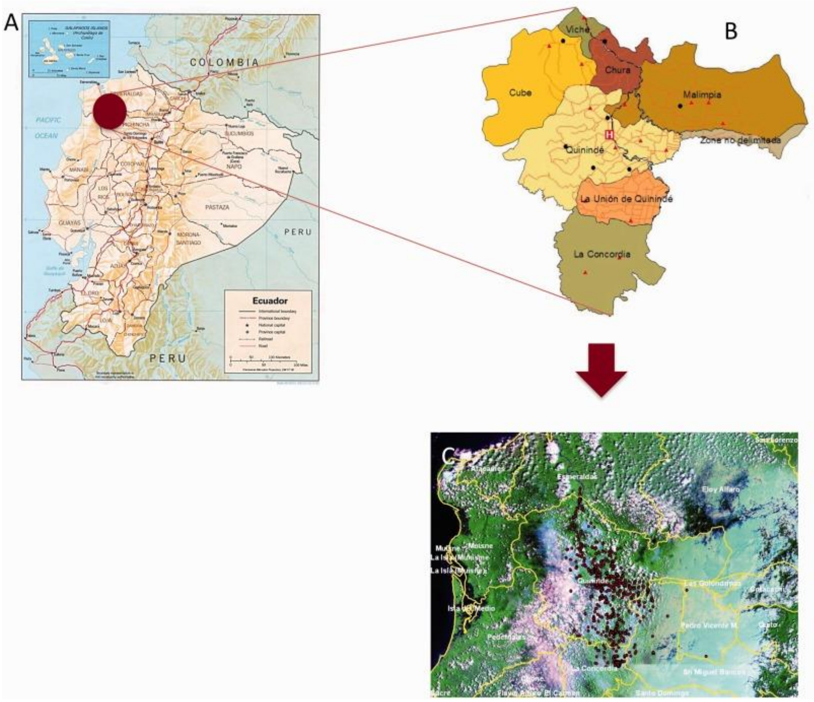
The goal of this study is to evaluate if and how measuring bacterial counts of *E. coli* in different ways, for example, taking fewer water quality samples or measuring water quality as a categorical variable as opposed to a continuous variable, affects association estimates between water sample characteristics and water quality as well as water quality and health outcomes.

There are three main methods that water quality is measured: membrane filtration (MF), most probable number (MPN), and presence-absence methods. With MF, colonies on a petri dish with the water sample are formed and counted. These colonies are represented as numbers of colony forming units per 100 mL of the sample as a continuous value. The MPN results represent the most probable number of coliform bacteria present in the water sample, not an actual count of the bacteria. After incubation with the appropriate medium, the pattern of positive results is used to estimate the concentration of coliform bacteria in the original sample by referencing statistical tables. While the statistical approximation leads to less precise results using the MPN method, it is more sensitive and applicable to turbid waters, unlike the MF method(15). The Colilert Quanti-Trays used in this study to measure water quality deliver quantification of coliforms using the MPN model (16). Finally, the presence-absence tests indicate the presence or absence of the indicator in the water and therefore this method is used when positive results are predicted to be rare and is not recommended for use in countries where contamination is common(15).

Population and sample

The water sample and household data for this research project were collected in the district (cantón) of Quinindé in Esmeraldas Province, Ecuador. According to the Development Plan for Esmeraldas, in 2010, the total population of the cantón of Quinindé was 122,570. The urban population of Quinindé was 28,928 (23%) and the rural population was 93,642 (76%) (18). It was estimated that the 2015 total population of Quinindé was 136,925. The coverage of water by a public source is estimated to be 31.4% in Quinindé (19).

The figure below (from Cooper et al., 2015), shows where in Ecuador the households are located (17).



Research Design

These data were collected in conjunction with household visits following the ECUAVIDA cohort. After receiving a study introduction and goals summary, willing participants gave verbal and written consent to participate in the study. The study activities included answering survey questions and providing water samples of household drinking water. The water samples were later analyzed for microbial contamination in the laboratory. Data on STH infection was determined by stool sample by the ECUAVIDA study. See Cooper et al. (2011, 2015), for a full description of the STH infection stool sample procedures.

Procedures

As mentioned previously, the data were collected during the ECUAVIDA cohort study household visits. In most cases, the household visits occurred at the same time as the 5- and 8-year-old follow-up visit for the ECUAVIDA visit. In other cases, additional households were visited if they were located near the households that were being visited for the routine follow-up visit.

The survey section for this study included asking the caretaker of the household a series of questions concerning access to drinking water, use of potable water, water purification methods, water storage, and occurrence of diarrhea within the past week. Diarrhea was defined as experiencing three or more loose stools per day. All storage containers stored in the house which contained water during the visit were recorded. The surveys were administered in Spanish and questions were read to the participants. The responses were recorded on the Android application Open Data Kit (www.opendatakit.org) using an Android phone (Samsung).

The water collection section of this study included asking the participants to provide drinking water samples of all drinking water sources. This included all storage containers as well as direct sources such as taps and wells. Drinking water sources were not sampled only in the rare event that potable water service was interrupted during the visit or if a single container contained water from more than one source, to reduce to classification bias. The study participant was asked to collect water how they usually would for consumption. Next, the water was poured from their drinking vessel into Whirl-Pak bags with sodium thiosulfate (NASCO Corp., Fort Atkinson, WI) for residual chlorine neutralization. Finally, the Whirl-Pak bags were placed on ice to then be transported to the laboratory.

The samples were processed within 8 hours of collection. The processed samples were tested for total coliform and *E. Coli* with Colilert Quantitray 2000 (IDEXX Laboratories Inc, Westbrook, ME). In order to read the total coliform and *E. Coli* results, the samples were first incubated at 35$_{o}$C between 24 and 48 hours. Every day, one negative control of sterile distilled water was processed and incubated along with the household samples. In the one instance when the negative control was positive, all the sample results from that day were discarded.

Data Analysis

All data analysis was performed using RStudio (Version 1.2.1335). The water sample characteristics include setting (rural vs urban), water source (rain, bottled, tap, well), water storage (yes, no), water supply (continuous, intermittent), and water treatment (none, boiling, chlorine, filter). The health outcome variables include diarrhea outcome and helminth infection and both are dichotomous (yes, no).

Bacterial counts of *E. Coli* were $\log_{10}$ transformed. For analysis purposes, bacterial counts below the limit of detection (<1 MPN/100 mL) were substituted for 0.5 MPN/ 100 mL and bacterial counts above the limit of detection (2,419.6 MPN/ 100 mL) were substituted with 2,420 MPN/ 100 mL before $\log_{10}$ transformation.

A categorical variable for bacterial count was created using the continuous bacterial count. The categorical variable was created to identify if using categorical water quality data as opposed to continuous water quality data results in different interpretations of associations between water quality and health outcomes. The categorization levels are described below:

Table 1. Description categorical variable coding according to the continuous bacterial count

Level	Continuous bacterial count
Level 1	< 1 MPN/ 100 ml
Level 2	1 – 9 MPN/ 100 mL
Level 3	10 – 99 MPN/ 100 mL
Level 4	100 – 999 MPN/ 100 mL
Level 5	> 999 MPN/ 100 mL

Along with the full dataset, a “randomly dropped” dataset and a “randomly kept” dataset were created. These datasets were created to identify changes in association estimates between the full dataset and data sets with fewer samples per household. The randomly dropped dataset demonstrated what the association would have been if one fewer samples were taken per household (for households that had more than one sample) while the randomly kept dataset demonstrated what the associations would have been if only one sample was taken per household. The randomly dropped dataset was made using Microsoft Excel. A random number generator was used to determine which of the samples per household should be dropped. For example, for households where three samples were collected, a random number generator from 1 through 3 was used to determine the observation to be dropped from the full dataset. For households where with only one observation, the observation was kept in order to include it in the analysis. This dataset was then read into RStudio for analysis. The randomly kept dataset was made using RStudio where a random observation from each household was saved into a new dataset. This left all the households with only one observation. The randomly kept and randomly dropped datasets were created by removing samples from the full dataset and therefore the categorical variable for bacterial count corresponds to that of the full dataset.

Finally, an "average" dataset was created in order to capture variability by household by summarizing the water quality for every household. The average dataset was made using RStudio where the average *E. Coli* $\log_{10}$ bacterial count was calculated and kept for each household. The categorical variable in this dataset was determined using the same criteria previously mentioned, but for the average bacterial count. Additionally, since the average bacterial count was calculated per household regardless of water sample characteristics such as source, storage, supply, and treatment, the only preserved household characteristics include the setting and whether or not diarrhea or helminth infection occurred.

Due to the non-normal distribution of the bacterial count data, the non-parametric test, Kruskal-Wallis rank sum test, was performed to determine the statistical significance in the difference between the datasets. Since multiple water samples were taken for most households, linear and logistic mixed models were used to control for clustering by household. For the datasets where there is only one entry per household (the randomly kept and average datasets), simple linear and logistic regression were used.

Ethical Considerations

The study protocol was conducted in July 2015 through September 2015. The study protocol was reviewed by the Institutional Review Board (IRB) at Emory University as well as the Universidad Central de Ecuador in Quito, Ecuador. Since this study was performed within an existing cohort population as requested by the host organization (FEPIS), the Emory University IRB waived approval for the study.

**Note: the methods sections pertaining to the sample and how information was collected was based on Dr. S Aya Fanny's thesis (20)*

Results

Descriptive statistics

From the 162 households that were visited between July 2015 and September 2015, a total of 438 drinking-water samples were collected. Of these households, 107 were located in an urban setting while 54 were located in rural settings. The most common water source among the samples was tap water (238 samples), followed by well water (109 samples), then bottled water (56 samples), river (26 samples), and finally rain (9 samples). Water treatment was uncommon with 70% of the water samples (310 samples) not treated. Boiling water was slightly more common than using chlorine as treatment (14.84% and 13.47%, respectively), and only 4 samples used filter treatment (0.91%). Two-thirds of the water samples were stored at the time of the visit. Table 1 summarizes the descriptive statistics of the water samples. The resulting characteristics of the water samples in the randomly dropped and the randomly kept data sets are also summarized in Table 1.

The distribution of *E. coli* is right skewed as displayed in the histogram in Figure 1a. The log-transformed distribution is also right skewed as displayed in the histogram in Figure 1b. The distribution of log EC is similar as seen through the means and standard deviations as well as the median and IQRs displayed in Table 1. Additionally, Figure 2 illustrate the distribution of log EC count in the form of boxplots. The full and randomly dropped datasets appear more similar in their distributions compared to the randomly kept and the average datasets. That is, the datasets with more similar sample sizes have similar distributions. After conducting a Kruskal Wallis test comparing the four datasets, resulting in a chi-squared statistic of 7.52 and p- value of 0.057, we do not have sufficient evidence to conclude that there is statistical difference between the datasets.

Almost one quarter (n=38) of the households reported diarrhea while roughly 15% (n=25) of houses reported helminth infection.

Assessing the Association Between Water Sample Characteristics and Continuous Water Quality Data

The following boxplots and tables display the distributions and associations between water sample characteristics such as setting, source, storage, supply, and treatment and water quality. Boxplots are shown to visualize the variation in the difference in distributions and central tendencies between datasets. Kruskal-Wallis rank sum test statistics (W) are reported to verify statistical significance among the groups. Finally, the linear regression model results for the different datasets are shown to demonstrate varying resulting associations between water sample characteristic and continuous water quality data.

As mentioned previously, the average dataset only preserved the setting for each household. Therefore, the average dataset only appears in the analysis of association between the water sample setting and water quality.

Association between Setting and Water Quality

The *E. Coli* counts (EC), expressed as $\log_{10}$ MPN/ 100 mL, were compared among urban and rural areas to assess if geographic location affected water quality. For all datasets, the median *E. Coli* counts in rural areas is higher than the log *E. Coli* of the water samples in the urban areas (Figure 4). On average, the rural areas had higher levels of contamination compared to urban areas (W=18,482, p-value = 0.01213).

In order to assess whether geographic location could predict water quality while controlling for clustering by household, mixed linear regression models were performed for the four datasets. No matter the dataset, there was never a significant association between setting and water quality.

Association between Water Source and Water Quality

The EC counts were compared among different water sources to assess if water source affected water quality. The distribution of EC for the different water sources is summarized in Figure 5. River water had the highest contamination of EC (median EC =2.272), followed by well water (median EC = 1.509), then tap water (median EC = 0.716) , followed by rain water (median EC = 0), and finally, bottled water median (EC = -0.125) for the full dataset. The Kruskal-Wallis test provides evidence to support that there is a statistically significant-difference in the averages among the water sources (W=51.8, $p<0.05$). A similar pattern in EC distributions by source is seen in the randomly dropped and randomly kept datasets.

When comparing the linear regression outcomes among the different datasets, there is always a positive association between river water and EC as well as well water and EC when using rain water as the reference. For river water, on average, EC is almost 1.6 units higher than rain water in the full dataset ($\beta=1.599$, $p<0.001$), almost 2 units higher in the randomly dropped dataset ($\beta=1.999$, $p<0.01$), and about 1.5 units higher than rain water in the randomly kept dataset (β=1.501, $p<0.05$). For well water, on average, EC is roughly 1.2 units higher than rain water in the full dataset (β=1.238, $p<0.01$), 1.4 units higher than rain water in the randomly dropped dataset (β=1.415, $p<0.05$), and 1.2 units higher in the randomly kept dataset (β=1.176, $p<0.05$). In summary, while the positive association between both river and well water against water quality exists no matter the dataset used, the randomly dropped dataset provides the largest mean difference in water quality between river and well water compared to rain water (Table 5).

Association between Storage and Water Quality

The boxplot in Figure 6 shows the difference in EC levels among the water samples that were stored on the day of collection and the EC levels among the water samples that were not stored. On average, the water samples that were stored (median EC = 1.086) were statistically significantly more contaminated than the water samples that were not stored (median EC= 0.491) for the water samples that were not stored in the full dataset (W = 3.9764, p-value = 0.0461). The pattern is consistent in the randomly dropped and randomly kept datasets.

When comparing the linear regression outcomes among all the datasets, no matter the dataset, there is no statistically significant association between water storage and water quality (Table 6).

Association between Supply and Water Quality

The boxplot in Figure 7 shows the distributions of EC for water samples from a continuous water source and an intermittent water source. Water samples from intermittent supply (median = 0.716) were higher compared to those form continuous supply sources (median = - 0.301). On average, the difference between these groups is significant (W = 7.723, p-value = 0.0055)

Based on the linear regression outcomes, only the randomly dropped dataset produced a statistically significant association between water supply and water quality (Table 7). In the randomly dropped dataset, on average, the households with intermittent water supply showed a 0.469 $\log_{10}$EC unit increase in water quality compared to those with continuous water supply.

Association between Treatment and Water Quality

The association between water treatment method and water quality was assessed. The boxplot shown in Figure 8 shows that the water samples which were boiled generally have lower EC levels compared to the other treatment methods. The Kruskal-Wallis test confirms that there is a statistically significant difference between these groups (W=21.09, p-value = 0.0001).

Boiling water was statistically significantly associated with water quality in the full ($\beta=-0.529$, $p<0.001$) and randomly dropped ($\beta=-0.465$, $p<0.05$) datasets, where it showed a protective effect, but did not show a statistically significant association in the randomly kept dataset (β=-0.473, $p>0.05$). Chlorine treatment was statistically significantly associated with water quality only in the randomly kept dataset and compared to no treatment, appeared to, on average, have higher EC counts (β=0.699, $p<0.05$) (Table 8).

Assessing the association between water quality and health outcome

The following boxplots and tables display the distributions and associations between water sample quality in terms of EC and health outcomes such as diarrhea outcome or helminth infection. The tables show the odds ratios (OR) and confidence intervals (CI) which were calculated from the given beta values. The point and whisker plots are shown to visualize the variation in ORs and CIs among the different datasets. Only water quality and health outcomes are included in these models. Therefore, in addition to comparing the full, randomly dropped, and randomly kept datasets, the average dataset can be used since water quality and health outcomes are present in this dataset.

Assessing the association between water quality and diarrhea outcome using a continuous exposure variable

As mentioned above (Table 1), diarrhea occurred in roughly 23% of households. Among urban households, the prevalence of diarrhea was 20.3% compared to 17.9% in rural areas. After performing a Wilcoxon rank sum test with a p-value of 0.5477, we do not have sufficient evidence to observe a statistically significant difference the proportion of diarrhea in the two groups. Across all datasets, the median EC value was higher where there was an occurrence of diarrhea.

Table 9 below shows that only the average dataset provides a statistically significant association between EC and diarrhea outcome (OR = 1.07, 95% CI = 1.003 - 1.413). Furthermore, estimating the associations between diarrhea outcome and water quality by setting, the statistically significant association between water quality and diarrhea outcome remains only among the rural setting. Alternatively, while the full dataset did not produce a statistically significant result as a whole, there is a statistically significant association between diarrhea outcome and water quality in the urban households in the full dataset (OR = 1.185, 95% CI (1.180 – 1.190)).

Assessing the association between water quality and diarrhea outcome using a categorical exposure variable

When using the categorical EC exposure variable (defined in the methods section) and setting the lowest exposure category as the reference group, the association between water quality and diarrhea outcome results in very wide confidence intervals (Table 10). Using a categorical predictor variable only produces a statistically significant association between the highest level of EC contamination and diarrhea in the randomly kept dataset (OR = 4.659 (1.067 – 20.337)).

Assessing the association between water quality and helminth infection using a continuous exposure variable

The prevalence of helminth infection in the entire dataset was 15.4% (n=25). Among rural households, the prevalence was 7.41% (n=4) and among urban households, the prevalence was 24.4% (n=21). In other words, almost all of the helminth infection cases occurred in urban households. The median EC was higher among households where helminth infection was present as opposed to where it was not present across all datasets (Figure 10).

When estimating the association between water quality and helminth infection, only the randomly kept (OR = 1.512, 95% CI = 1.130 – 2.023) and Average (OR = 1.869, 95% CI = 1.216 – 2.873) datasets resulted in statistically significant increased odds of water quality and helminth infection. The statistically significant association remained only in the urban setting in the Average dataset (OR = 1.915, 95% CI = 1.179 – 3.111) (Table 11).

Assessing the association between water quality and helminth infection using a categorical exposure variable

When using the categorical EC exposure variable and setting the lowest exposure category as the reference group, the association between water quality and helminth infection results in very high p-values and wide confidence intervals. Using a categorical predictor variable only results in statistically significant associations when comparing

Level 4 exposure to Level 1 exposure in the average (OR = 5.921, 95% CI =1.144 – 30.652) and randomly kept (OR = 7.424, 95% CI = 1.859 – 29.652) datasets (Table 12).

Discussion

This study not only estimated the associations between water sample characteristics and water quality, water quality and health outcomes, but also compared how these associations varied when using the water quality data in different ways. While the associations between water characteristics and water quality have been researched, as have the associations between water quality and health outcome, they often conclude varying results. Furthermore, there is no standard way of looking at *E. coli* (sometimes dichotomous, other times continuous). This study aimed at comparing how the associations vary depending on continuous and categorical variables for water quality. Additionally, the water quality data varied in terms of number of water samples representing water quality per household or if a central measure of tendency was used to represent the water quality for that household.

Water sample characteristics and E. coli

As visualized through the side by side boxplots summarizing EC by water characteristic, the datasets had similar medians and distributions for their respective water sample characteristic, Furthermore, the pattern of the distributions remained the same throughout the datasets. For example, boiling water produced the lowest median EC while filtered water produced the highest EC no matter the dataset.

Estimating the association between water source and water quality showed that compared to rain water, river and well water consistently had an increased association to higher EC. The increased EC in river water could be due to human activities near the river as explained by the case study of the Mun River Basin in Thailand (Zhao, 2018). When estimating the association between water treatment and water quality, boiling water, on average, was associated with lower EC in both the Full and Randomly Dropped datasets but the statistically significant association no longer exists in the Randomly Kept dataset.

Differences when looking at E. coli as a continuous variable

The logistic regression models show that there were statistically significant associations between water quality and health outcomes only for the datasets where there was one entry per household: diarrhea outcome, Average dataset; helminth infection, Randomly Kept and Average datasets. While clustering for household was accounted for by using mixed models, the datasets with multiple samples for each household did not result in statistically significant associations between EC and health outcomes.

Furthermore, the datasets that showed a general statistically significant associations between water quality and health outcome were then further investigated to see if there was a difference in this association between urban and rural settings. For diarrhea outcome, among the Average dataset only, there was an increased odds of diarrhea outcome for higher EC in rural (OR = 1.985, 95% CI: 1.000 – 3.940) and urban (OR = 1.278 , 95% CI: 0.825 – 1.979) settings (Table X). Similarly, rural settings had higher odds of helminth infection from occurring in rural areas (Random Keep OR = 2.55, 95% CI: 0.753 – 8.645, Average OR = 2.930, 95% CI: 0.862 – 9.962) compared to urban areas (Random Keep OR = 1.427, 95% CI: 0.955 – 2.132; Average OR = 1.915, 95% CI :

1.179 – 3.111) (Tables X-X). Similar to previous studies, perhaps human activities common in these areas affect the water quality and therefore the health outcome.

Differences when looking at E. coli as an ordinal variable

The associations between EC and health outcomes when using a categorical EC variable resulted in very unstable estimates as presented by the extremely high p-values and wide confidence intervals (Tables 10, 11). Despite the high p- values, the general pattern was higher levels of EC category had higher ORs, which could be predicted.

Differences between E. coli as a continuous and categorical variable

There is a lack of statistically significant associations between EC and diarrhea when using both the continuous and ordinal EC variable. Additionally, when using both the continuous and ordinal EC variables for helminth outcome, only the Average and Randomly Kept datasets resulted in statistically significant associations between EC and helminth infection (Table 12). However, it is important to note that the estimated association appears to be stronger when using the categorical EC variable. Also, interestingly, while the randomly kept dataset did not produce a statistically significant association between continuous EC exposure and diarrhea (1.254, 95% CI = 0.923 – 1.702), there was a statistically significant association between in the ordinal EC exposure when comparing level 5 exposure to level 1 exposure (OR= 4.659, 95% CI =

1.067 – 20.337). Therefore, an association may appear to be stronger than it actually is if using categorical exposure variables.

Limitations

There are several limitation in this study. First, since this had a cross-sectional design, risk of health outcomes could not be estimated. Furthermore, taking samples at different time periods could further inform on water quality variability by time. Secondly, the water quality samples which were taken at home may not always correspond to the health outcomes since household members are exposed to other sources of water other than the household water. For example, many children who attend school may have exposure to diarrhea and helminth risk factors in school as opposed to at home. Additionally, the models used for the analysis did not take other covariates such as nail trimming (Novianty, 2018) or handwashing habits (Pasaribu, 2019) into consideration which could affect the association estimate. Also, diarrhea within the last week, one of the health outcome variables, was prone to recall bias. Finally, the logistic regression used to estimate the categorical associations was limited because not all models were able to converge given the categorization. In the future, other types of analysis should be used to assess the association between ordinal water quality exposures and health outcomes.

Future Directions

The analyses conducted for this study lead to the conclusions that:

- There does not appear to be any loss of information when using categorical data as predictor EC variable as opposed to continuous data,
- Using categorical EC variable may result in concluding that there is a stronger association than there actually is, and
- Datasets with one measurement for every household tend to result in statistically significant associations between water quality and health outcome.

Future studies where the goal is to assess water quality should continue to take multiple samples per household since taking just one sample per household could lead to a statistically significant result when in fact, the variability in multiple samples does not allow for such certain claims. Additionally, future studies that are deciding between continuous or categorical water collection methods should feel certain in the ability of categorical data to be used in the data analysis steps.

Tables and Figures

Table 2. Summary of household characteristics for all data collected

Characteristic	N	%
Geographic location		
Urban	107	66.46%
Rural	54	33.54%
Diarrhea		
Present	38	23.60%
Not Present	123	76.40%
Helminth infection		
Present	25	15.43%
Not Present	137	84.57%

Table 3. Summary of distribution of water sample characteristics by dataset

Dataset	**Full**		**Randomly Dropped**		**Randomly Kept**		**Average**	
Characteristic	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Source								
Rain	9	2.05%	5	1.68%	4	2.47%	-	-
Bottled	56	12.79%	36	12.12%	24	14.81%	-	-
River	26	5.94%	17	5.72%	10	6.17%	-	-
Tap	238	54.34%	171	57.58%	78	48.15%	-	-
Well	109	24.89%	68	22.90%	46	28.40%	-	-
Treatment								
None	310	70.78%	201	67.68%	119	73.46%	-	-
Boiling	65	14.84%	52	17.51%	23	14.20%	-	-
Chlorine	59	13.47%	42	14.14%	18	11.11%	-	-
Filter	4	0.91%	2	0.67%	2	1.23%	-	-
Storage								
Yes	290	66.21%	203	68.35%	69	45.59%	-	-
No	148	33.79%	94	31.65%	93	57.41%	-	-
Supply								
Continuous	49	11.19%	32	10.77%	26	16.15%	-	-
Intermittent	243	44.48%	174	58.59%	77	47.83%	-	-
NA	146	33.33%	91	20.64%	58	26.02%	-	-
Mean EC log10 (SD)	0.992 (1.172)		0.995 (1.170)		0.921 (1.202)		0.989 (1.008)	
Median EC log 10 (IQR)	0.924 (-0.301, 1.934)		0.989 (-0.301,1.884)		0.716 (-0.301,1.911)		0.754 (0.163, 1.793)	

Figure 1. Distribution of EC for all samples

Figure 1a. Distribution of *E. Coli*

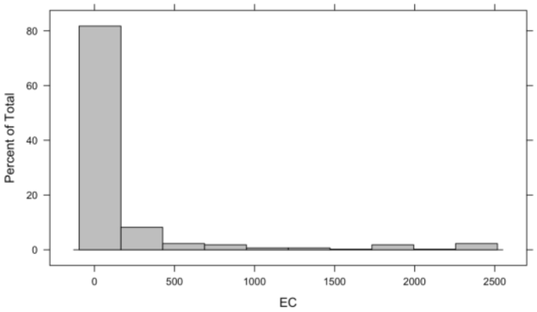


Figure 1b. Distribution of $\log_{10}$ *E. Coli*

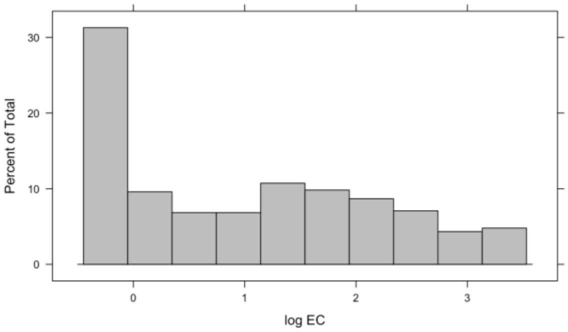


Figure2. Distribution of EC by dataset, boxplot

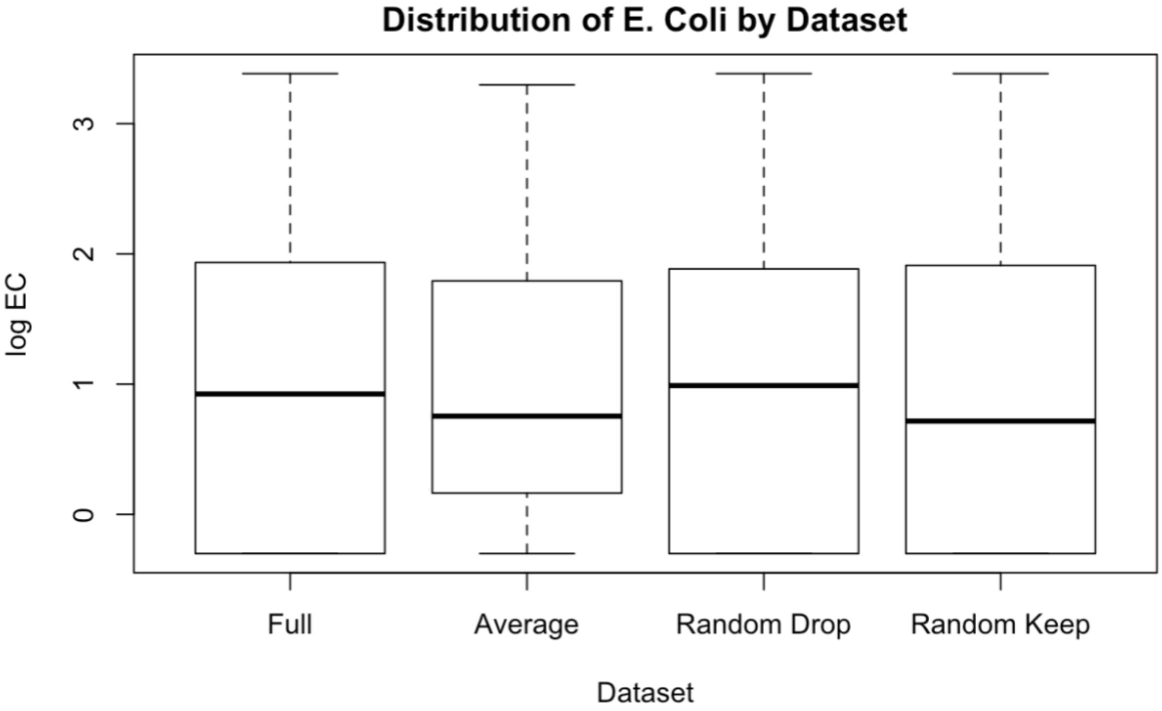
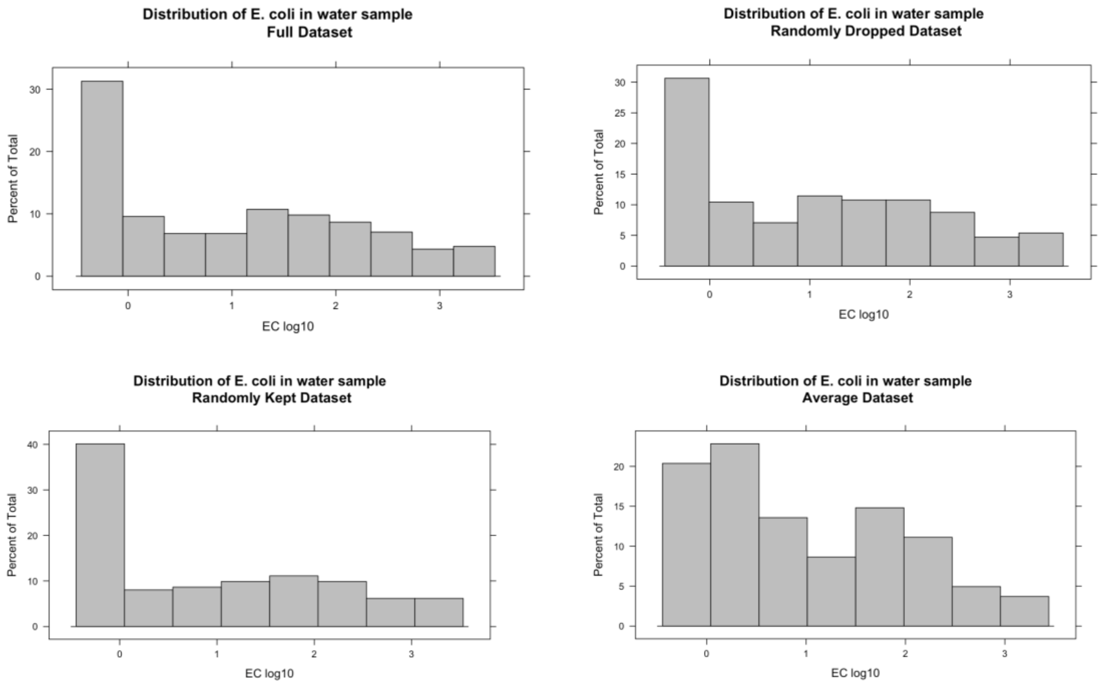


Figure 3. Distribution of EC by dataset, histograms



Distribution of E. coli in water sample
Full Dataset

Distribution of E. coli in water sample
Randomly Dropped Dataset

Distribution of E. coli in water sample
Randomly Kept Dataset

Distribution of E. coli in water sample
Average Dataset

Figure 4. Distribution of $\log_{10}$ *E. coli* counts for all datasets by setting

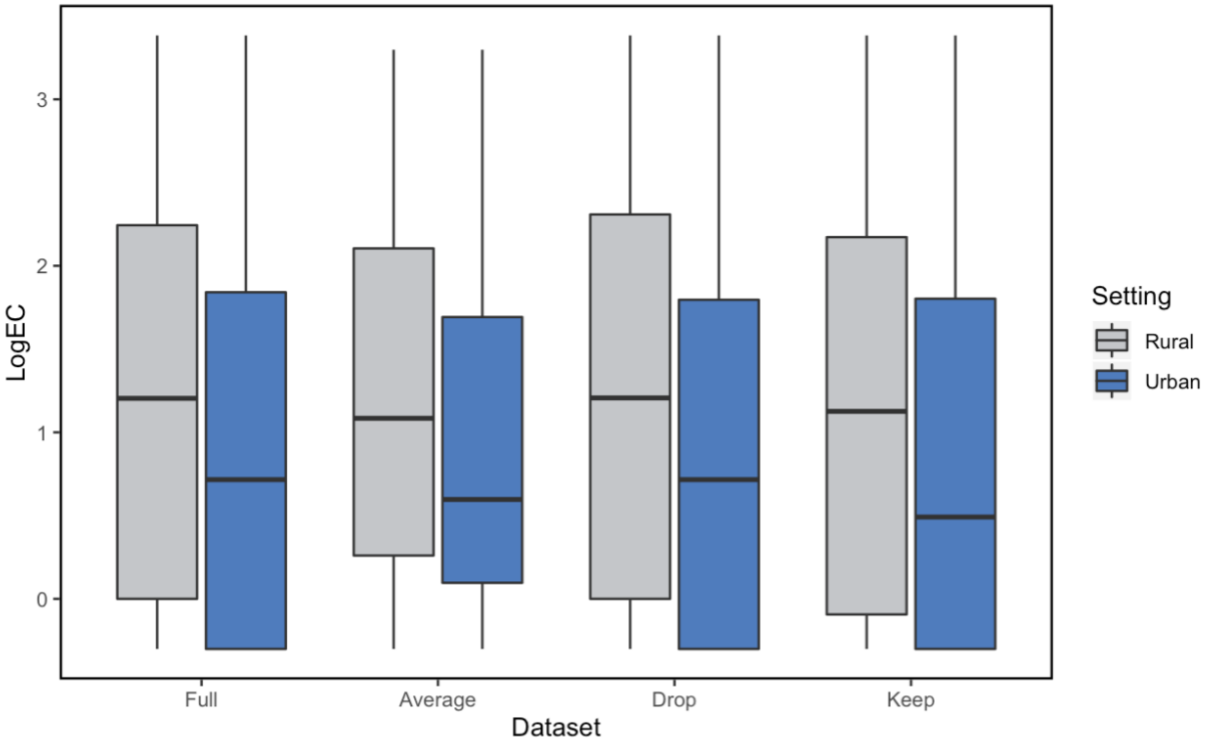


Table 4 Association between Setting and Water Quality

(Reference = Urban)

Dataset	Beta (SE)
Full	0.258 (0.163)
Random Drop	0.229 (0.180)
Random Keep	0.331 (0.199)
Average	0.276 (0.168)

Figure 5. Distribution of $\log_{10}$ *E. Coli* for full, random drop, and random keep datasets by source

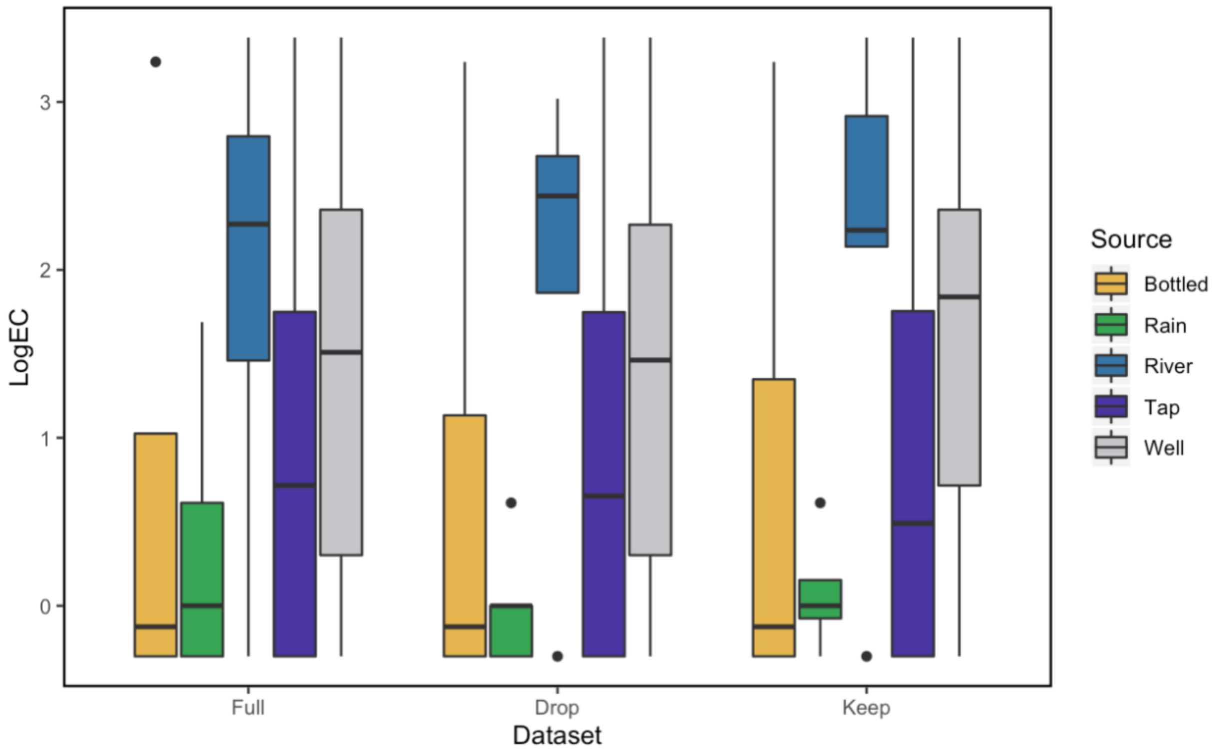


Table 5. Association between Water Source and Water Quality

(Reference = Rain)

Dataset	Beta (SE)	
Full	Bottled	0.281 (0.449)
	River	1.599 *** (0.475)
	Tap	0.680 (0.433)
	Well	1.238 ** (0.442)
Random Drop	Bottled	0.485 (0.564)
	River	1.999 ** (0.617)
	Tap	0.871 (0.543)
	Well	1.415 * (0.554)
Random Keep	Bottled	-0.114 (0.602)
	River	1.501 * (0.659)
	Tap	0.498 (0.571)
	Well	1.176 * (0.581)
Note: *p<0.05; **p<0.01; ***p<0.001 No * : p>0.05		

Figure 6. Distribution of $\log_{10}$ *E. Coli* for full, random drop, and random keep datasets by storage

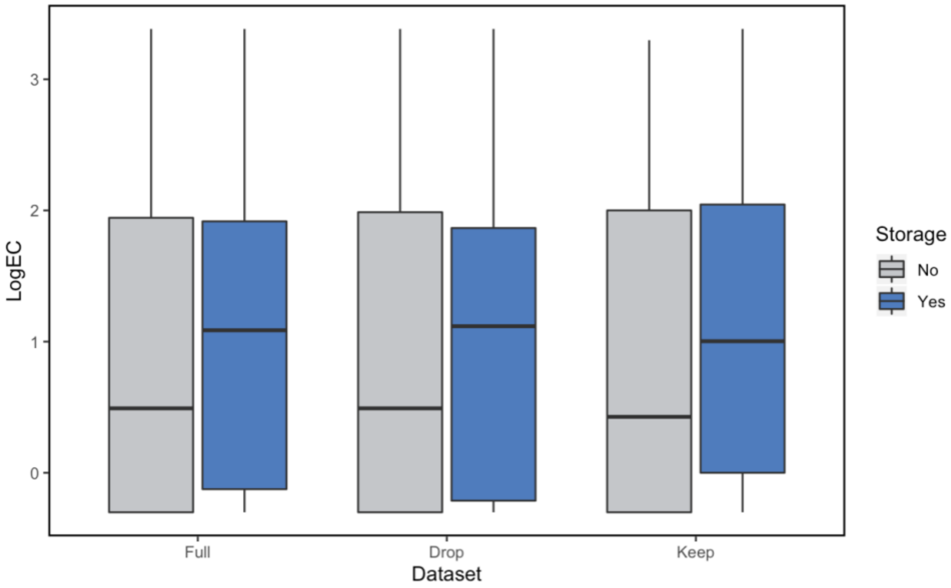


Table 6. Association between Storage and Water Quality

(Reference = No)

Dataset	Beta (SE)
Full	0.143 (0.103)
Random Drop	0.102 (0.138)
Random Keep	0.266 (0.190)
Note: *p<0.05; **p<0.01; ***p<0.001 No * : p>0.05	

Figure 7. Distribution of $\log_{10}$ *E. Coli* for full, random drop, and random keep datasets by supply

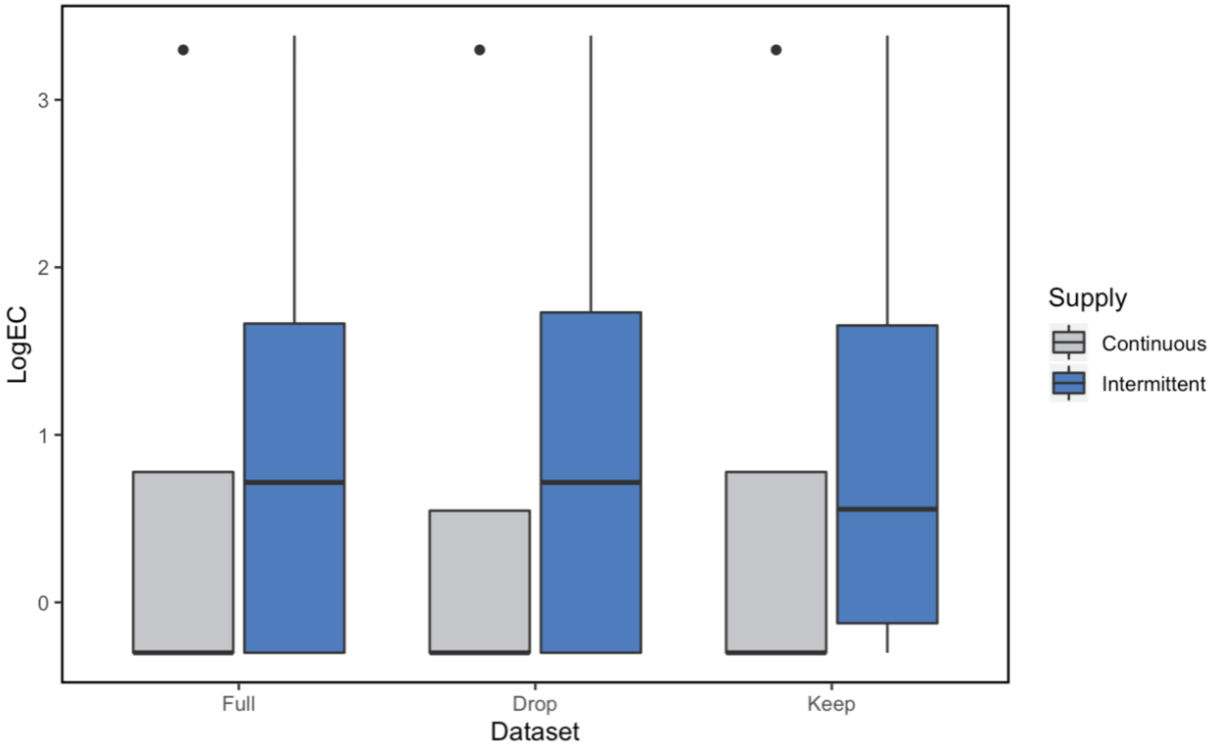


Table 7. Association between Supply and Water Quality

(Reference = Continuous)

Dataset	Beta (SE)
Full	0.331 (0.222)
Random Drop	0.469 * (0.217)
Random Keep	0.184 (0.257)
Note: *p<0.05; **p<0.01; ***p<0.001 No * : p>0.05	

Figure 8. Distribution of $\log_{10}$ *E. Coli* for full, random drop, and random keep datasets by treatment

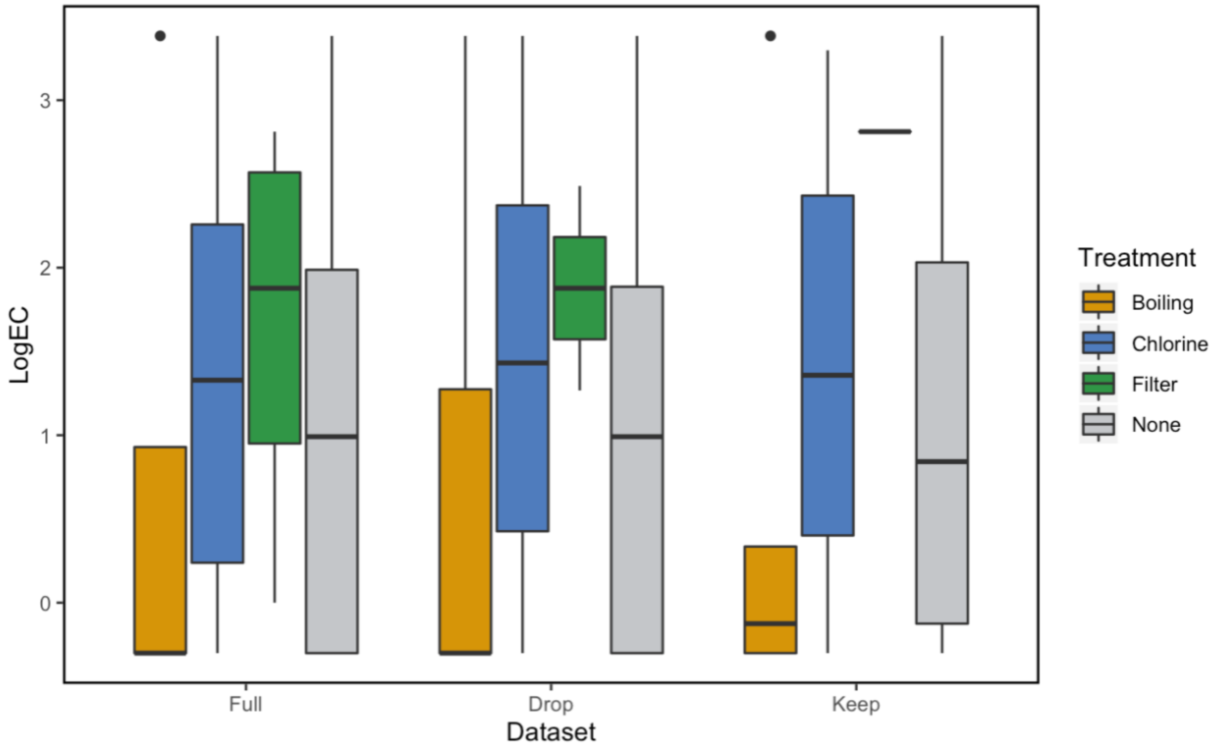


Table 8. Association between Treatment and Water Quality

(Reference = None)

Dataset	Beta (SE)	
Full	Boiling	-0.529*** (0.156)
	Chlorine	0.096 (0.218)
	Filter	0.391 (0.548)
Random Drop	Boiling	-0.465 * (0.187)
	Chlorine	0.363 (0.267)
	Filter	0.546 (0.719)
Random Keep	Boiling	-0.473 (0.267)
	Chlorine	0.699* (0.296)
	Filter	0.979 (0.835)
Note: *p<0.05; **p<0.01; ***p<0.001 No * : p>0.05		

Assessing the association between water quality and health outcome

Figure 9. Distribution of $\log_{10}$ *E. Coli* for all datasets by diarrhea outcome

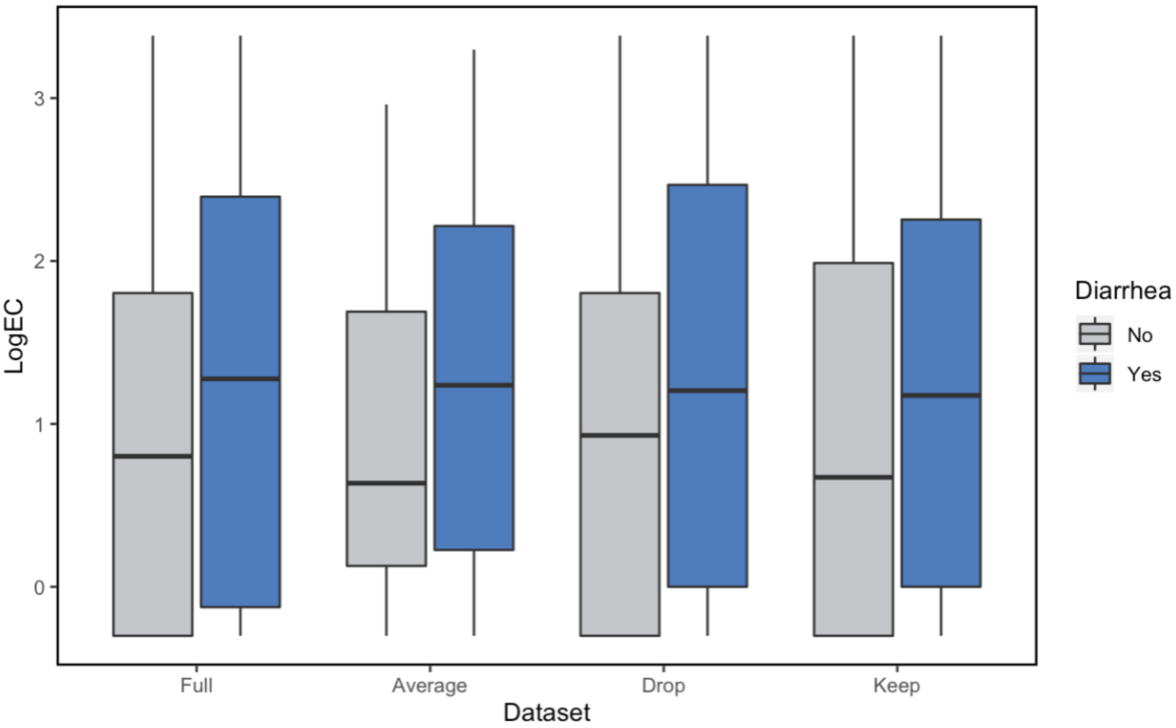


Table 9. Association between log EC and diarrhea outcome

Dataset	OR (CI)	p-value
Full	1.259 (0.293 – 5.417)	0.757
Random Drop	1.559 (0.356 – 6.830)	0.556
Random Keep	1.254 (0.923 – 1.702)	0.147
Average	1.442 (1.007 – 2.064)	0.046
Among Rural Setting		
Full	1.483 (0.107 – 20.599)	0.769
Random Drop	1.546 (0.112 – 21.263)	0.744
Random Keep	1.460 (0.839 – 2.538)	0.180
Average	1.985 (1.000 – 3.940)	0.050
Among Urban Setting		
Full	1.185 (1.180 – 1.190)	< 0.05
Random Drop	1.558 (0.260 – 9.315)	0.627
Random Keep	1.240 (0.853 – 1.802)	0.271
Average	1.278 (0.825 – 1.979)	0.271

Table 10. Association between categorical EC and diarrhea outcome

(Reference = Level 1)

Dataset	EC Level	OR (CI)	p-value
Full			
	Level 2	1.255 (0.0113– 139.358)	0.925
	Level 3	0.838 (0.00699 – 100.346)	0.942
	Level 4	1.726 (0.0150–198.531)	0.822
	Level 5	2.959 (0.00718–121.963)	0.724
Random Drop			
** The model could not converge under with categorical exposure for this dataset			
Random Keep			
	Level 2	0.966 (0.333 – 2.803)	0.950
	Level 3	0.932 (0.332 – 2.696)	0.896
	Level 4	1.242 (0.439 – 3.518)	0.683
	Level 5	4.659 (1.067 – 20.337)	0.041
Average			
	Level 2	1.122 (0.356 – 3.529)	0.844
	Level 3	1.173 (0.338 – 4.077)	0.801
	Level 4	1.760 (0.494 – 6.273)	0.383
	Level 5	--	--

-- In the average dataset, the Level 5 OR , CI and p- value were not able to be generated.

Figure 10. Distribution of $\log_{10}$ *E. Coli* for all datasets by helminth infection

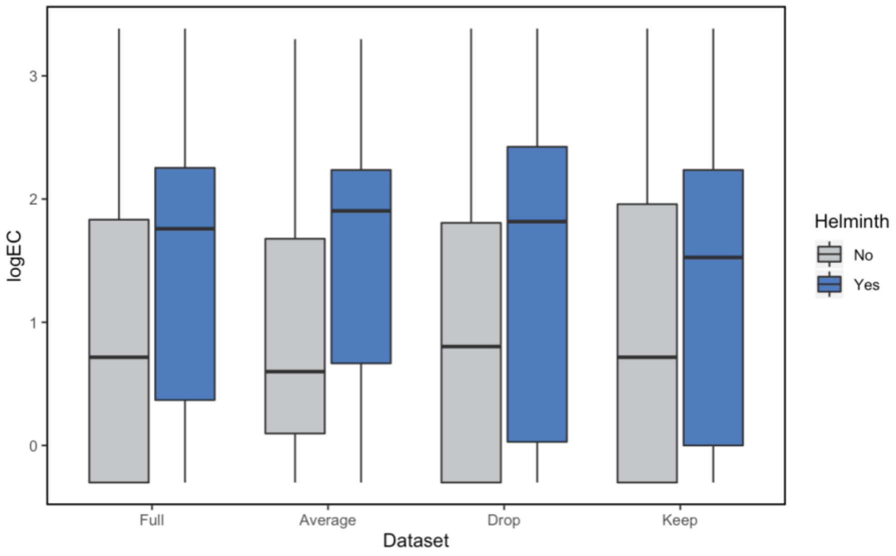


Table 11. Association between log EC and helminth outcome

Dataset	OR (CI)	p-value
Full	1.447 (0.244 – 8.563)	0.684
Random Drop	1.488 (0.237 – 9.33)	0.671
Random Keep	1.512 (1.130 – 2.023)	0.005
Average	1.869 (1.216 – 2.873)	0.044
Among Rural Setting		
Full	1.930 (0.010 – 369.200)	0.806
Random Drop	2.160 (0.0074 – 6.25 $x10_2$)	0.790
Random Keep	2.55 (0.753 – 8.645)	0.132
Average	2.930 (0.862 – 9.962)	0.085
Among Urban Setting		
Full	1.478 (0.215 – 10.180)	0.691
Random Drop	1.522 (0.204 – 11.314)	0.681
Random Keep	1.427 (0.955 – 2.132)	0.083
Average	1.915 (1.179 – 3.111)	0.009

Table 12. Association between categorical EC and Helminth Infection

(Reference = Level 1)

Dataset	EC Level	OR (CI)	p-value
Full			
	Level 2	0.629 (0.0002– 2.41x10^{3})	0.912
	Level 3	1.629 (0.0064 – 4.13 x10^{3})	0.863
	Level 4	2.983 (0.0011–1.02 x10^{3})	0.714
	Level 5	2.874 (0.00718–7.31 x10^{3})	0.792
Random Drop			
	Level 2	0.482 (0.0048 – 4.82 x10^{3})	0.877
	Level 3	1.092 (0.0228) – 4.19 x10^{3})	0.977
	Level 4	2.951 (0.0084 – 1.04 x10^{3})	0.718
	Level 5	2.571 (0.0013 – 5.87 x10^{3})	0.811
Random Keep			
	Level 2	1.021(0.162 – 6.452)	0.983
	Level 3	4.083 (0.977 -17.063)	0.054
	Level 4	7.424 (1.859 – 29.652)	0.005
	Level 5	8. 167 (1.335 – 49.95)	0.023
Average			
	Level 2	1.271 (0.240 – 6.734)	0.778
	Level 3	2.822 (0.538 –14.807)	0.220
	Level 4	5.921 (1.144 – 30.652)	0.034
	Level 5	4.167 (0.285 – 60.929)	0.297

Table 13. Summary of Associations Between Water Quality (Continuous and Categorical) Against Health Outcome (Diarrhea and Helminth Infection)

	Continuous	Level 2	Level 3	Level 4	Level 5
Full					
Diarrhea	1.259	1.255	0.838	1.726	2.959
Helminth	1.447	0.629	1.629	2.983	2.874
Random Drop					
Diarrhea	1.559	-	-	-	-
Helminth	1.488	0.482	1.092	2.951	2.571
Random Keep					
Diarrhea	1.254	0.966	0.932	1.242	**4.659**
Helminth	**1.512**	1.021	4.083	**7.424**	**8.167**
Average					
Diarrhea	**1.442**	1.122	1.173	1.760	-
Helminth	**1.869**	1.271	2.822	**5.921**	4.167

Bold values represent statistically significant ORs.

Purple cells represent an OR less than 1.

Green cells represent an OR greater than 1.

“-“ represent cells that did not produce an OR given the categorical exposure

References:

1. GBD 2016 Diarrhoeal Disease Collaborators (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet. Infectious diseases*, *18*(11), 1211–1228. https://doi.org/10.1016/S1473-3099(18)30362-1
2. Izurieta, R., Reina-Ortiz, M. and Ochoa-Capello, T. 2018. Trichuris trichiura. In: J.B. Rose and B. JiménezCisneros, (eds) Global Water Pathogen Project. http://www.waterpathogens.org (Robertson, L (eds) Part 4 Helminths) http://www.waterpathogens.org/book/trichuris-trichiura Michigan State University, E. Lansing, MI, UNESCO. https://doi.org/10.14321/waterpathogens.43 Acknowledgements: K.R.L. Young, Project Design editor; Website Design (http://www.agroknow.com)
3. Pullan, R. L., Smith, J. L., Jasrasaria, R., & Brooker, S. J. (2014). Global numbers of infection and disease burden of soil transmitted helminth infections in 2010. *Parasites & vectors*, *7*, 37. https://doi.org/10.1186/1756-3305-7-37
4. Gruber, J. S., Ercumen, A., & Colford, J. M., Jr. (2014). Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PLoS One*, *9*(9), e107429. doi:10.1371/journal.pone.0107429
5. Gundry, S., Wright, J., & Conroy, R. (2004). A systematic review of the health outcomes related to household water quality in developing countries. *J Water Health*, *2*(1), 1-13. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15384725
6. Cohen, A., & Colford, J. M. (2017). Effects of Boiling Drinking Water on Diarrhea and Pathogen-Specific Infections in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis. *The American journal of tropical medicine and hygiene*, *97*(5), 1362–1377. https://doi.org/10.4269/ajtmh.17-0190
7. Morris, J. F., Murphy, J., Fagerli, K., Schneeberger, C., Jaron, P., Moke, F., Juma, J., Ochieng, J. B., Omore, R., Roellig, D., Xiao, L., Priest, J. W., Narayanan, J., Montgomery, J. M., Hill, V., Mintz, E., Ayers, T. L., & O'Reilly, C. E. (2018). A Randomized Controlled Trial to Assess the Impact of Ceramic Water Filters on Prevention of Diarrhea and Cryptosporidiosis in Infants and Young Children-Western

Kenya, 2013. *The American journal of tropical medicine and hygiene*, *98*(5), 1260–1268. https://doi.org/10.4269/ajtmh.17-0731

8. Matanock, A., Lu, X., Derado, G., Cuéllar, V. M., Juliao, P., Alvarez, M., … Roy, S. L. (2018). Association of water quality with soil-transmitted helminthiasis and diarrhea in Nueva Santa Rosa, Guatemala, 2010. *Journal of Water and Health*, *16*(5), 724–736. doi: 10.2166/wh.2018.207
9. Pickering, A. J., Njenga, S. M., Steinbaum, L., Swarthout, J., Lin, A., Arnold, B. F., Stewart, C. P., Dentz, H. N., Mureithi, M., Chieng, B., Wolfe, M., Mahoney, R., Kihara, J., Byrd, K., Rao, G., Meerkerk, T., Cheruiyot, P., Papaiakovou, M., Pilotte, N., Williams, S. A., … Null, C. (2019). Effects of single and integrated water, sanitation, handwashing, and nutrition interventions on child soil-transmitted helminth and Giardia infections: A cluster-randomized controlled trial in rural Kenya. *PLoS medicine*, *16*(6), e1002841. https://doi.org/10.1371/journal.pmed.1002841
10. Zhao, Z., Liu, G., Liu, Q., Huang, C., & Li, H. (2018). Studies on the Spatiotemporal Variability of River Water Quality and Its Relationships with Soil and Precipitation: A Case Study of the Mun River Basin in Thailand. *International journal of environmental research and public health*, *15*(11), 2466. https://doi.org/10.3390/ijerph15112466
11. Levy, K., Hubbard, A. E., Nelson, K. L., & Eisenberg, J. N. (2009). Drivers of water quality variability in northern coastal Ecuador. Environmental Science & Technology, 43(6), 1788-1797. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19368173
12. Boehm, A. B. (2007). Enterococci Concentrations in Diverse Coastal Environments Exhibit Extreme Variability. *Environmental Science & Technology*, *41*(24), 8227–8232. doi: 10.1021/es071807v
13. Goddard, F. G. B., Chang, H. H., Clasen, T. F., & Sarnat, J. A. (2020). Exposure measurement error and the characterization of child exposure to fecal contamination in drinking water. *Npj Clean Water*, *3*(1). doi: 10.1038/s41545-020-0063-9
14. Bain, R., Cronk, R., Wright, J., Yang, H., Slaymaker, T., & Bartram, J. (2014). Fecal contamination of drinking-water in low- and middle-income countries: a systematic

review and meta-analysis. *PLoS medicine*, *11*(5), e1001644. https://doi.org/10.1371/journal.pmed.1001644

15. WHO. (2011). Guidelines for Drinking Water Quality.Volume 3
16. IDEXX. (n.d.). Retrieved from https://www.idexx.com/en/water/water-products-services/quanti-tray-system/
17. Cooper, P. J., Chico, M. E., Platts-Mills, T. A., Rodrigues, L. C., Strachan, D. P., & Barreto, M. L. (2015). Cohort Profile: The Ecuador Life (ECUAVIDA) study in Esmeraldas Province, Ecuador. International Journal of Epidemiology, 44(5), 1517-1527. doi:10.1093/ije/dyu128
18. Gobierno Autonomico Descentralizado Municipal del Cantón Quinindé. (2015). Plan de Ordenamiento Territorial del GAD Municipal de Quinindé. Retrieved from http://app.sni.gob.ec/sni-link/sni/PORTAL_SNI/data_sigad_plus/sigadplusdocumentofinal/0860000160001_PDOT%20ACTUALIZACI%C3%93N%202015_18-08-2015_12-32-18.pdf
19. SENPLADES. (2014). Agua Potable y Alcantarillado para Eradicar la Pobreza en Ecuador. In SENPLADES (Ed.). Quito, Ecuador. Retrieved from https://www.planificacion.gob.ec/wp-content/uploads/downloads/2014/09/FOLLETO-Agua-SENPLADES.pdf
20. Fanny, S.A. (2016). Water, Worms and Weird Diseases:
    Water Quality Variability and Pediatric Health Outcomes in Northern Coastal Ecuador. Emory University