Distribution Agreement

Signature:

________________________________ _______________

Andrew Thornton Date

The Effect of Household Water, Sanitation, and Hygiene Characteristics on the Quality of Water at the Point of Consumption in Nueva Santa Rosa, Guatemala

By

Andrew Thornton

Master of Public Health

Global Epidemiology

____

Juan Leon, PhD, MPH

Committee Chair

____

Sharon Roy, MD, MPH

Committee Member

The Effect of Household Water, Sanitation, and Hygiene Characteristics on the Quality of Water at the Point of Consumption in Nueva Santa Rosa, Guatemala

By

Andrew Thornton

B.S., University of Wisconsin – Madison, 2008

Faculty Thesis Advisor: Juan Leon, 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 Global Epidemiology

2011

Abstract

The Effect of Household Water, Sanitation, and Hygiene Characteristics on the Quality of Water at the Point of Consumption in Nueva Santa Rosa, Guatemala

By Andrew Thornton

Background: Water contamination at the point of consumption has important implications to human health, particularly as it relates to diarrheal illness. Past studies have found water contamination to increase from the water source to the point of consumption, and there is a need to better understand this relationship.

Goal: The goal of our study was to evaluate potential differences in the contamination of water at both the source and point of consumption, and to determine risk factors that may lead to the contamination of consumed drinking water in the "*municipio*" of Nueva Santa Rosa, Guatemala.

Methods: We conducted a randomized field study in which we interviewed 210 households regarding water, sanitation, and hygiene characteristics and practices. We collected water samples at both the source and point of consumption from 139 households and tested them for free chlorine residual levels, total coliforms, and generic *E. coli*.

Results: Chlorination was found in only 2 water samples at both the source and point of consumption. Contamination with total coliforms was found in more than 97% of samples at both the source and point of consumption. Contamination with *E. coli* was found in more than 72% of samples at both the source and point of consumption. No significant difference was found in contamination levels between the source and point of consumption for total coliforms and *E. coli*. Multivariate linear regression found that of all water, sanitation, and hygiene characteristics, only the amount of *E. coli* in source water was significantly (positively) associated with the amount of *E. coli* in water at the point of consumption.

Conclusion: This investigation highlights the importance of water treatment to protect against fecal contamination, and the need for point-of-use treatment to protect drinking water at the point of consumption in the absence of a chlorinated supply.

The Effect of Household Water, Sanitation, and Hygiene Characteristics on the Quality of Water at the Point of Consumption in Nueva Santa Rosa, Guatemala

By

Andrew Thornton

B.S., University of Wisconsin – Madison, 2008

Faculty Thesis Advisor: Juan Leon, 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 Global Epidemiology

2011

Acknowledgements:

First and foremost I would like thank my family for their endless love and encouragement. Mom, Dad, Amy, and Matt, I couldn’t have done it without you. The last year has been a tough one for all us, but we are stronger for it. Thank you for always being there.

To my faculty advisor, Juan Leon, I can’t thank you enough for your mentorship over the past two years. Thank you for imparting your organization and writing skills, and for your patience, guidance, and the many, many hours spent in meetings and with revisions. I sincerely hope to work with you again in the future.

To my field advisor, Sharon Roy, you have played an invaluable role in my education and professional development while at Emory. Thank you for giving me the tremendous opportunity to be a part of our study, and for your dedicated support and mentorship,

Thank you to all of my friends at Emory for two great years. In particular, I would like to thank Sydney Hubbard and Toshiko Dignam. Your friendship has meant a lot to me, and I can’t thank you enough for your support and lasting impact on the person that I am.

I would also like to thank Jaymin Patel and Victoria Cuéllar. This thesis would not have been possible without you, and Guatemala would not have been nearly as fun. I look forward to more *chapeetings* in the future!

Finally, I would like to thank the many people involved with our study in Guatemala, including staff from the CDC, the Universidad del Valle de Guatemala, our wonderful field team, and all of the study participants in Nueva Santa Rosa. This would not have been possible without you.

Dedication:

I dedicate this thesis, in its entirety, to my mom, Nancy Blyler. Without your unconditional love and support, I would not be where I am today. Your courage and grace are my inspiration, and my strength. I love you, and I think about you every day.

We must accept finite disappointment, but we must never lose infinite hope.
- Martin Luther King Jr.

Table of Contents

LITERATURE REVIEW

Water is essential to life; however, water contamination is a persistent global problem. The World Health Organization (WHO) estimates that 13% of the world's population (850 million people) does not have access to an improved water source [1]. This results in millions of people consuming water that may be of lower biological quality. The consumption of this water, in concert with lack of access to sanitation and inadequate hygiene, places people at risk for diarrheal diseases [2]. WHO reports that most of the world's diarrheal disease (88%) is attributable to unsafe water, sanitation, and hygiene [2]. In 2004, diarrheal disease was the world's 5th leading cause of death, which resulted in 2.2 million deaths [3], and 99% of these deaths occurred in the developing world [4].

Maintaining water free from fecal contamination is central to improved health. A meta-analysis focusing on water, sanitation, and hygiene (WASH) interventions and diarrheal illness found that water quality interventions reduced diarrhea frequency by 31% [5]. Current evidence suggests that improving water quality at the point-of-consumption (POC) is more important to reducing diarrheal illness than the quality at the source, as a clean source is of limited value if contamination occurs after collection [6]. A number of studies have investigated this relationship between water quality at the source and POC. In general, in the absence of interventions aimed at improving water quality, water contamination has been shown to increase from the source to POC. One meta-analysis found an increase in contamination in more than half of the studies included, and did not find any significant decrease in contamination. [7]. Our study will further explore the relationship between water quality at the source and POC, and will also investigate the factors that affect the quality of water at its POC.

Measurement of fecal contamination of water: fecal indicator bacteria versus pathogen detection

Fecal contamination of drinking water can cause diarrheal illness, and therefore there is a need to be able to accurately test for the presence of fecal contamination. It is possible to test water directly for diarrheal pathogens. However, the large variety and low prevalence of pathogens coupled with testing costs makes it prohibitive to test for them all. Therefore, fecal contamination of water is generally measured by testing for indicators of fecal contamination, such as certain viruses and bacteria that can be found in the gastrointestinal tract of animals and/or humans. Viral indicators may include, among others, *Bacteroides fragilis*, phages, and somatic and male-specific coliphages. Bacterial indicators include total coliforms (TC), thermotolerant (fecal) coliforms, *Escherichia coli* (EC), and fecal streptococci (reviewed in [8]). Our investigation measured both TC and EC, which will be the focus of this section.

There are advantages and disadvantages to using indicators to test for fecal contamination. One of the advantages to using indicator bacteria is that testing for these bacteria is quick and relatively inexpensive (reviewed in [8]). In addition, by testing for indicator bacteria, it is not necessary to test for the myriad of other bacteria that may be present in a water sample to confirm fecal contamination. According to WHO and United States Environmental Protection Agency (EPA) guidelines, neither TC nor EC should be present in treated drinking water [8, 9]). However, the presence of indicator bacteria is not always associated with the presence of fecal contamination. While both TC and EC are universally present in high numbers in human and warm-blooded animal feces, there are several types of TC bacteria that are present in natural soils and water (reviewed in [10]). As a result, a water sample might test positive for TC if there has been inadequate water treatment, bacterial re-growth, post-treatment contamination, a nutrient-rich water supply, and/or biofilm formation (reviewed in [8]). Conversely, EC is rarely found in the absence of fecal pollution, though it has been found in natural tropical water systems (reviewed in [11]). Growth of EC is also unlikely in water distribution systems (reviewed in [8]). Although EC is more specific than TC for fecal contamination, one limitation of EC measurements is that EC is not a good indicator for a number of enteric pathogens including *Cryptosporidium parvum*, *Giardia lamblia*, *Yersinia enterocolitica*, and enteric viruses (reviewed in [8, 11, 12]). Despite drawbacks, indicator bacteria, and EC in particular, are still powerful tools for assessing the fecal contamination of water.

Post-Source Water Contamination

In general, water contamination can occur in either the public domain or the domestic domain. The public domain includes public places of work, schooling, commerce and recreation, and in the fields and streets [13]. This contamination generally occurs at the water source. The domestic domain includes areas normally occupied and under control of the household, and is not limited to solely the interior of the home [13]. Water contamination in the domestic domain (post-source contamination) is the focus of this investigation. This contamination typically occurs when households use water supplies that require transport and/or storage, such as surface water and networked systems with intermittent supplies; contamination can also occur at the POC. Risk factors for the contamination of post-source water in the domestic domain are classified into several categories including: household water characteristics, sanitation and hygiene characteristics, and socioeconomic factors.

Post-Source Water Contamination Risk Factors

Household Water Characteristics

In the absence of safe water storage or POC treatment interventions, the microbiological quality of source water may play a role in water quality post-source, however findings are not consistent. Studies have shown that post-source water contamination occurs regardless of water quality at source [5, 7, 14-18]. However, a meta-analysis found that water contamination generally increased after collection and that this increase in contamination between the source and POC was proportionally greater when the source water was clean [7]. Other studies have found that high levels of contamination at the source may lead to a decrease in the amount [19] or percentage [20] of EC contamination in stored water because of bacterial die-off over time. However, the study by Momba and Notshe was done under laboratory conditions where re-contamination was unlikely. Conversely, increased post-source contamination has also been associated with the amount of time water spends in household storage [21-23]. This contamination may be the result of a lack of safe water storage practices, but may also be caused by microbial growth in the water.

In addition to water storage time, other water storage characteristics also play a role in post-source water contamination. The presence of a cover on the storage container has been shown to reduce TC and fecal coliform counts by as much as 50% [24, 25]. Smaller openings in water storage containers have also been associated with reduced fecal contamination of water at the POC [19, 22]. Studies have shown that extracting water with utensils or hands instead of using a spigot is a risk factor for contamination for stored water [15, 26]. The Centers for Disease Control and Prevention (CDC) recommends that home water storage containers have an opening with a screw-on lid. The opening should facilitate both cleaning and filling, and be 6-9 cm in width to discourage the insertion of hands or utensils. CDC also recommends a spigot or tap for water extraction [27]. The protective effect of safe water storage containers, combined with water treatment, has been shown in multiple studies [28, 29].

Household water treatment, when performed correctly, has also been shown to reduce microbial contamination in stored water. Such methods include chlorination, solar disinfection (SODIS), and boiling [30]. The positive effects of chlorination with sodium hypochlorite on microbial water quality have been shown in numerous studies [22, 28, 31, 32]. However, chlorination is not as effective against viruses and protozoa and can cause odor and a bad taste when the water contains a high concentration of organic matter (reviewed in [33]). SODIS has also been shown to be an effective method to improve water quality, though it requires relatively clear water and an area with substantial solar radiation (reviewed in [33]). Boiling can also improve the water quality of stored water by inactivating bacteria, viruses, and protozoa. However the household cost associated with boiling can be economically prohibitive and environmentally unsustainable [14, 34, 35]. Additionally, boiling provides no residual protection against fecal contamination and water can easily be re-contaminated [14, 36]. Therefore, while households have many options for treating water at POC, chlorination may be the most effective option.

Sanitation and Hygiene Factors

Proper sanitation and hygiene are intrinsically linked and play an important role in post-source water quality. WHO defines improved sanitation as facilities that ensure the hygienic separation of human excreta from human contact [1]. Without such separation, there is an increase in both the presence of fecal matter in the environment and human exposure to the fecal matter. Consequently, the risk of microbial contamination in the household increases. For example, a study in rural Botswana in which 70% of the participating households had limited sanitation (no toilet, dirty toilet, or toilet not in use) found that 31% of washed and unwashed plates, 29% of dishcloths, and 40% of baby bottles were contaminated with fecal coliforms [37]. Several studies have shown connections between such contaminated household utensils and post-source water contamination. For example, studies from Peru and Bolivia found that household drinking glasses were contaminated with EC [14, 16]. Studies have also found that fecally contaminated water was used to rinse or wash drinking cups and utensils, which increases the potential for household water contamination [14, 38]. Thus, household water quality can be affected by sanitary conditions, as shown in another study that found a lack of basic sanitation increased the likelihood of fecal indicator bacteria contaminating stored water [39].

The contamination of stored water may also result from human contact with stored water and/or water storage containers. The relationship between hand hygiene and stored water quality has been shown in several studies. A study in Tanzania found a significant correlation between mean EC levels found on the hands of household members and the mean EC levels found in household stored water [15]. Hand hygiene is of particular concern because physical hand contact with water is often unavoidable during water storage and collection [22, 38, 40, 41]. However, the magnitude of the impact of hand hygiene on water quality is not consistently demonstrated across all such studies. For example, a study in Thailand found that an intervention that included hand washing education and the provision of improved water storage containers with taps resulted in a significant improvement in the quality of aggregated stored household water, though the improvement was not significant with hand washing education alone [42].

Socioeconomic Factors

Household socioeconomic (SES) and demographic factors may also help predict post-source water quality. Parental education, for example, may be indirectly associated with water quality at the POC. A study conducted in Tanzania found that maternal hand contamination was related to the amount of fecal contamination in stored water samples, and maternal educational attainment was associated with hand contamination [15]. In Peru, access to clean water at the POC has been related to household wealth. Households with children younger than age 5 years that were living in extreme poverty were found to have less access to clean water [43]. In addition, the location of the home, whether urban or rural, may be of importance. Urban areas with high population density have poor environmental health, which can lead to post-source water contamination [7]. In summary, the hygiene and sanitation characteristics associated with education, household wealth, and location may indirectly relate to water quality at POC.

In conclusion, there are a number of factors that can affect the quality of water at its POC. These factors include water quality at source, water storage and extraction methods, household sanitation, home and personal hygiene, and socioeconomic factors. All of these factors likely play a role in the amount of contamination likely to be found in water at the POC.

Study Site: Nueva Santa Rosa, Guatemala

To investigate the relationship between water quality at the source and POC we chose to conduct a study in the *municipio* of Nueva Santa Rosa, Guatemala. The site is relatively close to CDC's Regional Office for Central America and Panama (CDC-CAP) and collaborating laboratories, which are located in Guatemala City, Guatemala. In addition, The International Emerging Infections Program (IEIP) of the CDC-CAP has been conducting population-based surveillance at the hospital in the *departamento* of Santa Rosa since March 2007, and at all six health centers and health posts in the *municipio* of Nueva Santa Rosa.

Nueva Santa Rosa is located in the *departamento* of Santa Rosa. As of 2002 the population of Nueva Santa Rosa was 29,957. Most of the population is *Ladino* (mixture of Spanish decent with Indigenous groups) with an additional 3,338 inhabitants (approximately) belonging to the *Xinca* indigenous group [44]. Spanish is the common language. Improved water and sanitation are fairly common within the *municipio*. In 1998-1999 in the southeast zone of Guatemala (which includes Nueva Santa Rosa), only an estimated 8% of households used surface water as their main water source, and 25% had no sanitation facilities [45].

Statement of Need, Project Goals and Aims, Significance

Statement of Need and Goal

There is a need to understand the quality of water, both at its source and at its POC in the *municipio* of Nueva Santa Rosa, in southeastern Guatemala. This includes understanding factors that might influence water quality. This knowledge will be used to develop evidence-based recommendations to guide future health initiatives of the Guatemalan Ministry of Public Health and Social Welfare and other international organizations in rural Guatemala. The goal of this project is to evaluate potential differences in quality of water at the source and POC, and to determine factors that may lead to the contamination of consumed drinking water among households in Nueva Santa Rosa, Guatemala, during the summer of 2010.

Project Aims

The project presented in this thesis was one component of a pilot water, sanitation, and hygiene (WASH) study that was embedded within a study to determine the community-level prevalence and incidence of disease (diarrhea, influenza-like-illness, and soil-transmitted helminthiasis) in Nueva Santa Rosa. As such, the design of the pilot

WASH study was constrained by the sample size required for the main study and was not fully powered. However, within that constraint, the aims of this project were:

1) To compare the water quality between the source and POC with respect to presence of EC and TC.

2) To model the relationship between the quality of drinking water at its POC (with respect to presence of EC) and several household and WASH characteristics. These characteristics include household water, sanitation, and hygiene characteristics, and household socioeconomic and demographic characteristics.

Significance

The results of this investigation will be important to government and to researchers in many ways. The investigation will provide a better understanding of water quality at both the source and POC in Guatemala and other marginalized areas. The study will also highlight themes that need to be addressed in order to improve POC water quality and public health. In particular, the findings will provide information regarding the possible causes of water contamination at the POC, and how to best target interventions to counter this contamination.

METHODS

Household Selection

Our study was conducted from July–August 2010 in the *municipio* of Nueva Santa Rosa. Nueva Santa Rosa is located in the *departamento* of Santa Rosa, Guatemala. In the 2002 National Census, the population of Nueva Santa Rosa was estimated to be 29,957 persons in 6,189 households (Instituto Nacional de Estadistica Guatemala [INE], 2002).

The unit of analysis for this study was the household. To create a sampling frame of households, we obtained high-resolution aerial photographs of Nueva Santa Rosa taken in 2006 by the Instituto Geográfico Nacional (National Geographic Institute) of Guatemala and overlaid a series of 200m x 200m grids and a topographic map with GIS coordinates. Potential residential structures (households, kitchens, dining rooms, bedrooms, living rooms) were identified according to roof size (between $25m^2$ - $150m^2$ in urban areas, between $16m^2$ - $150m^2$ in rural areas), and were digitized to generate GIS coordinates for each roof. Each potential residential-associated roof was represented by a dot placed on top of the roof in the aerial photos. Grids containing 52 or more potential residential-associated roofs were classified as urban. We selected a simple random sample of potential residential-associated roofs in Nueva Santa Rosa to interview. If the dotted roof corresponded to a house, kitchen, dining room, bedroom, or living room, the household associated with this structure was interviewed. If more than one household was associated with a selected structure, then all households were included to maintain an equal probability sample of households.

Institutional Review Board (IRB) ethical approval for this investigation was granted by the IRB review board at the Centers for Disease Control and Prevention and the Universidad del Valle de Guatemala, in Guatemala City. The Emory IRB determined that because the author was under CDC IRB approval the project did not require Emory University IRB review (Appendix A). The Emory IRB also determined that IRB review was not required for the secondary data analysis of this data because it did not meet the definition of "Research involving Human Subjects."

Field Methods

Data was collected over four weeks in July–August 2010. Household surveys and sample collection were conducted by 8 data collection teams. Household surveys were based on existing surveys created and used by the WHO and UNICEF [46], the Demographic Health Survey (DHS) [47] and the CDC (unpublished documents). The investigators added and removed questions as necessary to address study aims. Consent was given by all participating households. Consent was given either via the household representative's signature, or thumbprint, if illiterate. If a thumbprint was given, the signature of a witness was also obtained. Households were given a copy of the consent form. After obtaining household consent, one enumerator on the team interviewed the household representative, usually the female head of household or eldest daughter, about disease in the household (diarrhea, influenza-like illness, and soil-transmitted helminthiasis) and household water, sanitation, and hygiene practices. The other enumerator conducted an environmental assessment of household water, sanitation, and hygiene infrastructure. The assessment documented the type of water source(s) and the type of household sanitation facility according to the WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation (JMP) [1]. The enumerator also collected 100mL water samples at both the source of the drinking water (e.g., surface water, well, faucet) and the POC.

Source water was defined as water at its point of collection by the household. Source water from a body of water, such as a lake or river, was collected via the container commonly used by the household member(s) to withdraw water. Source water from taps and bottled water was collected directly in the water collection container brought by the field staff. Initially, source water samples for bottled water were incorrectly collected via a household drinking glass. At the end of the study, enumerators visited these households and re-collected the bottled “source” water sample correctly using a water collection container brought by the field staff.

POC water was defined as water served in a household drinking vessel. To collect water at the POC, the enumerator asked the interviewee for a glass of water, and water from this glass was poured into the collection container. In households that had tap water supplies, samples at the POC were taken only when tap water was stored because of intermittent supplies of piped tap water. In these households, water at the POC was collected from a glass of water provided by the interviewee that was taken from the tap water stored in the household (as described previously).

The enumerator also tested the chlorine residual of water collected at both the source and the POC using the Hach Free Chlorine Test Kit (Hach, Loveland, CO, USA). For each sample tested, the water was poured into two vials and a packet of reagents was emptied into one of the vials. Both vials were placed in a color comparator and a color disc was rotated until the color of the water in the reagent tube matched the color disc, corresponding to mg/L of chlorine.

At each interview, enumerators observed hand washing technique to determine if six basic hand washing steps, as defined by the CDC, were performed. These steps included using water, using soap or a cleanser, rubbing hands together, washing both hands, washing hands for 31+ seconds, and drying the hands hygienically (on a clean towel or air drying)[48]. Interviewees that performed at least five of the steps were categorized to have washed their hands "correctly."

Laboratory Methods

All water samples collected in the field were kept on ice in coolers until their arrival at the laboratory at the Universidad del Valle de Guatemala in Guatemala City. The next day, water samples were analyzed using the Colilert® Most Probable Number (MPN) method utilizing the Quanti-Tray® Enumeration Procedure (IDEXX, Westbrook, Maine, USA). This procedure consists of mixing the contents of a Colilert® Pack with a 100mL water sample, which is poured in a Quanti-Tray 2000, sealed, and incubated for 24 hours. The tray contains wells which turn yellow if positive for TC and fluorescent if positive for EC. The number of positive wells corresponds to the MPN. Positive wells were counted by two different laboratory staff and inconsistencies were re-checked. When applicable, the MPN was then multiplied by a dilution factor. Test results were recorded by laboratory staff into the study's Excel database. The test limits of detection for both TC and EC were < 1 MPN/100mL and > 2,419.2 MPN/100mL, which can change according to dilution factor. The lower limit of detection in our study for TC and EC was < 1 MPN/100mL. The upper limit of detection was 328,319 MPN/100mL for total coliforms and > 241,920 MPN/100mL for EC. Due to extreme contamination, several MPN estimates reached the detection limit and were recorded with a greater than (>) in the database. In order to include these in the analyses, these estimates were rounded to the next highest whole number. Many MPN estimates also reached the lower limit of detection, < 1, and were recoded to zero for the analyses.

Data Collection and Management

All survey data were collected in the field on Personal Digital Assistants (PDAs), which were backed up each night. Data collected in the laboratory were entered in Excel databases by laboratory staff. Throughout the study, data were checked for quality. When necessary, staff re-visited households to correct data inconsistencies. During data cleaning, implausible and conflicting answers were set to missing. At the discretion of the principal investigator, these answers were corrected when staff could specifically recall the household or when responses from other questions in the survey indicated a different response.

Analysis

To facilitate analysis, several variables were re-coded, including type of water source, type of household sanitation facility, and household wealth. The type of water source was classified as bottled water, piped water on the premises (in home or on patio), other improved water (public taps or standpipes, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection), and unimproved water (unprotected dug well, unprotected spring, water cart, tanker truck, and surface water). This classification scheme is adapted from the WHO/UNICEF JMP for Water Supply and Sanitation “water quality ladder.” In this ladder, the JMP has listed bottled water as an unimproved source [1]. However, the JMP considers bottled water as an improved water source if the household also uses another improved source for cooking and for hygienic purposes. We categorized bottled water separately because it was significantly less contaminated than unimproved water sources (see Results). The type of household sanitation facility was classified as improved (flush or pour-flush toilet/latrine, ventilated improved pit [VIP] latrine, pit latrine with slab, or composting toilet), shared (improved facilities shared by two or more households), unimproved (pit latrines without slabs or platforms, hanging latrines, and bucket latrines), or a lack of sanitation facilities. These classifications were adapted from the WHO/UNICEF JMP for Water Supply and Sanitation “sanitation facility ladder” [1].

Household wealth was categorized into equal quintiles through a principal components analysis of household belongings and construction materials. Monthly income was excluded from this analysis because of missing values.

Throughout the analysis, $P < 0.05$ was considered to be statistically significant. MPN estimates for TC and EC were not normally distributed, and were $\log_{10}$ transformed in an attempt to achieve normality. A 1 was added to all MPN estimates prior to transformation because an estimate of 0 MPN could not be transformed. Following transformation, TC followed a normal distribution, but EC did not. For consistency, non-parametric tests of significance were used for both indicators.

The Wilcoxon-Mann-Whitney Test was used to compare the median TC and EC contamination between the correctly and incorrectly sampled bottled specimens at the “source” to evaluate which group of samples to use in subsequent analyses. The Wilcoxon Signed Rank Sum Test was used to compare differences in contamination between the source and POC at the household level. To test whether the amount of contamination differed between water from different source types, levels of contamination among multiple groups were tested by the Kruskal-Wallis test followed by a Tukey all pairwise post hoc comparisons test.

Multivariate stepwise linear regression was used to identify household water, sanitation, and hygiene characteristics that were significant predictors of EC contamination in household water at the POC. Fecal contamination of consumed water is linked to diarrheal illness [5], and we modeled EC contamination because it is a better predictor of fecal contamination than TC (reviewed in ([8]). We modeled water contamination at the POC (rather than at the source) because water at the POC provides a better measure of the quality of water that is consumed. Variables were included in the predictive models because they were significantly associated with the outcome in unadjusted bivariate analysis, were biologically plausible, and were risk factors identified in other studies (see Literature Review).

Interaction terms were created for possible interactions between all predictors. Each interaction term was assessed for significance independently in a model containing the lower order predictors of the interaction term and adjusted for all confounders. To select final significant interaction terms, all significant interaction terms from this analysis were included in both forward and backward selection models using the *PROC REG* procedure. Multivariate stepwise linear regression was then used to identify significant predictors of EC contamination. Two predictive stepwise models were run. One model included the significant interaction terms and the other model did not. All data were analyzed using SAS version 9.2 (Cary, NC).

RESULTS

The goal of our study was to determine the quality of household water, both at its source and at its POC and identify factors that might influence water quality in the *municipio* of Nueva Santa Rosa, in southeastern Guatemala. To meet this goal our study aimed to (1) evaluate potential differences in the quality of water at both the source and POC, and (2) determine risk factors that may lead to the contamination of consumed drinking water.

Data were collected from a total of 210 households in Nueva Santa Rosa. Five households were excluded due to incomplete records, and of the 205 remaining households, 139 provided water samples at both the source and POC, making them eligible for subsequent analysis for this study. Forty-one samples of bottled source water were re-collected due to incorrect sampling. Analysis using the Wilcoxon-Mann-Whitney Test indicated that that there was no statistically-significant difference in median levels of TC ($P = 0.731$) and EC ($P = 0.343$) between the incorrectly and correctly sampled bottled water specimens within each household. Consequently, the test results for the correctly sampled specimens were substituted for the incorrectly sampled specimens in subsequent analyses despite the likelihood that the household was using a different bottled water source at the time of re-sampling.

To better understand the characteristics of the population surveyed, we performed univariate analyses on household demographics and other characteristics (Table 1). In general, households were of low socioeconomic status with low maternal education levels. Most households identified their ethnicity as *Ladino* and did not live in urban areas. Household size averaged 4.5 people, and on average, each household had 0.5

children younger than 5 years of age. Bottled water and piped water to the household were the most common water sources, and the majority of households did not treat water before drinking. At the source, two samples of household piped water had a measurable free chlorine residual level (Table 2). At the POC, two samples of bottled water had a measurable free chlorine residual level. Close to 75% of households used improved sanitation facilities, with a small portion living without sanitation facilities. In summary, households in the study were generally poor and uneducated, though most had both improved water sources and sanitation facilities.

Comparison of Water Quality Between the Source and Point of Consumption

To determine differences in contamination with total coliforms (TC) and *E. coli* (EC) between the source and POC, we quantified the amount of TC and EC in water samples and compared them using the Wilcoxon Signed Rank Sum Test. The prevalence of samples contaminated with TC (i.e., TC > 0 MPN) was uniformly high (> 91%) in both the source water and POC samples regardless of water type (i.e., bottled, piped on premises, other improved, and unimproved) (Figure 1). The prevalence of contamination with EC (i.e., EC > 0 MPN) was lower overall (> 30%) (Figure 2). For both TC and EC, there was no significant difference in median contamination levels (MPN) between the source and POC, both overall and when considered by type of water source (Figures 1 and 2). In conclusion, water sample median contamination levels for TC and EC did not vary between the source and POC.

Among source samples, median TC contamination levels at the source ranged from 2.42 Log_{10} MPN (263 MPN) for bottled water to 3.48 Log_{10} MPN (3,007 MPN) for other improved water sources (Table 3). Among POC samples, these ranged from 2.74 Log_{10} MPN (547 MPN) for bottled water to 4.03 Log_{10} MPN (10,618 MPN) for unimproved water sources. At both the source and POC, bottled water samples were significantly less contaminated than water from both unimproved and other improved water sources. At POC, samples of piped water to the dwelling were also significantly less contaminated than unimproved water source samples.

Among the source samples, median EC contamination levels ranged from 0 Log_{10} MPN (0 MPN) for bottled water to 2.26 Log_{10} MPN (181 MPN) for other improved water sources (Table 3). At POC these ranged from 0 Log_{10}MPN (0 MPN) for bottled water to 1.90 Log_{10}MPN (78 MPN) for other improved water sources. At both the source and POC, bottled water was significantly less contaminated than water from unimproved and other improved sources, and from water piped to the dwelling. In addition, at both the source and POC, water piped to the dwelling was significantly less contaminated than water from other improved sources. At the POC, water piped to the dwelling was also significantly less contaminated than water from unimproved sources.

In conclusion, there were significant differences in the amount of contamination with TC and EC among source and POC samples. In general, bottled water was of the highest quality, followed by water piped to the dwelling.

Predictors of Water Quality at the Point of Consumption

To investigate the relationship between household water, sanitation, and hygiene characteristics and the amount of EC at the POC, we performed bivariate analyses. All predictors used in the bivariate analyses were considered because they were biologically plausible for contributing to contamination at the POC and were risk factors identified in other studies (see Literature Review). From the bivariate analyses, we found significant negative associations between levels of EC and bottled water sources and water piped to the premise, when compared to unimproved water sources (Table 4). Significant negative associations with EC contamination were also identified for storage containers with openings less than 3cm in diameter, covered storage containers, improved household sanitation (compared to no sanitation facilities), sanitary disposal of baby feces, and correct hand washing. We found significant positive associations between the levels of EC at the POC and the levels of EC in the source water, storage containers located on the floor, the time it takes to collect water, and the presence of feces in the hand washing area. In conclusion, bivariate analyses identified multiple significant predictors of water contamination at the POC among household water, sanitation, and hygiene characteristics.

We also investigated the relationship between our *a priori* confounders, the household demographic characteristics, and the levels of EC contamination at the POC through bivariate analyses. From these bivariate analyses, we found significant negative associations between the levels of EC at the POC and household wealth quintile, *Indigena* compared to *Xinca* ethnicity, and maternal education beyond primary school as compared to having no formal education (Table 5). We found significant positive associations between the levels of EC at the POC and the number of people living in the household and the number of children younger than 5 years of age in the household. In conclusion, most demographic characteristics investigated were significantly associated with EC contamination at the POC using bivariate analyses.

To determine variables that significantly predict contamination with EC at the POC, we constructed two multivariate stepwise linear regression models (Table 6). One model included the assessment of interaction terms between predictor variables and the other did not. Household water, sanitation, and hygiene characteristics that were statistically significant in unadjusted bivariate analyses were included in these models (Table 4). All confounders (Table 5), independent of their significance, were also included in the multivariate predictive models. Two non-significant predictors that approached statistical significance were also included. These predictors were the presence of a hand washing station within 10 m of household sanitation facilities and the presence of animals in the house. Both multivariate models (with and without interaction) identified levels of EC at the source as a significant positive predictor. The parameter estimates for EC at the source were 0.36 in the model with interaction terms and 0.33 in the model without interaction terms. These results indicate that, in the model without interaction, an increase of approximately 1 $\log_{10}$ MPN of EC / 100mL in the source samples was associated with a 0.33 $\log_{10}$ MPN / 100mL increase at the POC, adjusted for the other variables in the model. When anti-logged, the results indicate that, in the model without interaction, an increase from 1 to 10 MPN of EC / 100 mL in the source samples was associated with a 1.13 MPN / 100 mL increase at the POC, adjusted for the other variables in the model. No confounders were significant in the model with interaction terms. However, in the model without interaction terms, three confounders were significantly associated with EC contamination at the POC: household wealth quintile and *Ladino* ethnicity (as compared to *Xinca*) were significantly negatively associated with EC contamination and the number of children younger than 5 years of age was significantly positively associated with contamination. In conclusion, both multivariate predictive models identified EC contamination levels at the source as a positive predictor for EC contamination levels at POC.

DISCUSSION

The goal of our study was to determine the quality of household water, both at its source and at its POC, and identify factors that might influence water quality in the *municipio* of Nueva Santa Rosa in southeastern Guatemala. For both TC and EC, there was no significant difference in contamination levels between the source and point of consumption, both overall, and when considered by type of water source. Multivariate linear regression found the amount of EC in source water to be positively associated with the amount of EC found in water at the POC.

Comparison of Water Quality Between the Source and Point of Consumption

The differences in the median levels of TC and EC contamination were not statistically different between the source and POC, both overall and when considered by type of water source. One mechanism to explain the lack of significant change between the source and POC is that additional household contamination may not have occurred during storage or at the POC. However, this scenario is questionable given unsafe household water handling and hygiene practices observed in a number of households, including lack of water treatment before consumption, uncovered water storage containers, and wide-mouthed water storage containers ($\geq$3 cm) (Table 1).

Much of the literature has found significant increases in the amount of contamination going from source to POC [5, 7, 14-18]; however, our findings of non-significant differences are supported in other studies. This includes a study in Pakistan that found no significant difference in the amount of EC found at the source and POC

[19]. It also includes a meta-analysis that found a non-significant difference between the source and POC for TC and fecal coliforms in 40% (6/15) of studies with a high percentage of contaminated source samples [7]. The meta-analysis also found contamination at the POC to be significantly higher when the amount of contamination at the source was low. These results parallel our findings for TC, and suggest that any additional contamination after source collection may not have been significant due to high source contamination.

Predictors of Water Quality at the Point of Consumption

Our investigation found the amount of EC in the source to be the main predictor of the amount of EC in water at the POC in two separate models, adjusted for several potential confounders, including household wealth quintile, ethnicity, mother's education level, number of persons in the household, number of children younger than 5 years of age in the household, and urbanization (Table 6).

Given the lack of differential in EC levels between the source and POC (Figure 2) and the lack of water chlorination in Nueva Santa Rosa (Table 2), this finding was foreseeable. It indicates that, in the absence of a significant differential, the amount of EC found in the water at the source was strongly related to the amount of EC found at the POC. This finding is supported by other studies that also found the quality of source water to be a significant predictor of the quality of water at the POC [18, 49]. Likewise, both studies also found water storage factors, such as covering the water container and the location of the water storage container, not to be significant predictors of POC water quality. They also did not find evidence of water chlorination at the source or POC.

The contamination of water with EC that we found at the POC is likely the result of fecally contaminated water sources and the lack of water chlorination. Water chlorination could be very effective at reducing the amount of microbial contamination in both networked water supplies and at the POC given that 93% of households in Nueva Santa Rosa (Table 1) stored drinking water. WHO recommends that municipal water supplies undergo a number of treatments to render it safe for public consumption. These treatments include coagulation, flocculation, sedimentation, filtration, and disinfection (reviewed in [8]). Chlorination is the most commonly used disinfection method and, in order to maintain water quality through the distribution network, CDC recommends a free chlorine residual greater than or equal to 0.5 mg/L [50]. Ideally, networked water supplies would provide drinking water to the entire population of Guatemala and all networked water supplies would be chlorinated, although this will be challenging and take time. Therefore, household water treatment is also an important intervention in the interim. A number of different household water treatment options exist, including point-of-use chlorination, boiling, and SODIS (solar disinfection). Point-of-use chlorination, in addition to its low cost, has the added advantage of residual protection against re-contamination [51]. Chlorination of household water has the potential to reduce the amount of bacteria and viruses in water by a $\log_{10}$ reduction 6 [8], and its protective effect on water quality has been shown in a number of studies [22, 28, 31, 32]. CDC recommends that, after 24 hours in storage, water should have a free chlorine residual between 0.2 mg/L and 2.0 mg/L to maintain water quality [50]. This recommendation is part of CDC's Safe Water System (SWS) program. The SWS also recommends the safe storage of treated water and behavior change communication focused on sanitation and hygiene [51].

Our investigation supported the WHO/UNICEF JMP "water quality ladder." The ladder names piped water on the household premise as the water supply most protected from outside contamination (e.g., fecal contamination) by the nature of its construction or through active interventions, followed by other improved water sources, then by unimproved water sources. Bottled water, as defined by the JMP is not an improved source unless the household uses another improved source for cooking and hygienic purposes [1]. Our survey did not specify the type of water used for cooking and hygiene, but we chose to keep bottled water as a separate category because it was generally found to be less contaminated with TC and EC than water from piped household connections, other improved water, and unimproved water at both the source and POC (Table 3). Water from households that had taps on the premise was also frequently found to be significantly less contaminated (particularly with respect to EC) than water from unimproved sources and other improved sources at both the source and POC (Table 3). This supports the "water quality ladder," however our findings showed that water from improved water sources may not necessarily meet WHO and EPA water quality guidelines which state that neither TC nor EC should be present in treated drinking water [8, 9]. Close to three-quarters of households in Nueva Santa Rosa regularly consume water that is contaminated with EC, despite much of this water coming from "improved" sources (Figure 2). This again emphasizes the need for treatment, both at the source and POC, to maintain safe drinking water supplies.

Strengths and Limitations

Our study was distinctive because, in addition to evaluating water quality in households with water sources located outside the home, it also evaluated water quality in households whose sources included bottled water and piped water. Most studies investigating the difference in contamination between the source and POC only have considered water that was transported from a source outside of the home. Sixty six percent of our samples were from either bottled water or from piped water to the household premise. WHO has found that using bottled water sources as a primary water source is becoming increasingly more common in the developing world. As of 2005, bottled water was used by more than 31% of urban populations and 10% of rural populations in Guatemala (reviewed in [1]). As bottled water and household tap connections become more common, understanding the differential in water quality from the source (in this case water inside the bottle or water directly from the tap) to the POC for these water types will be of greater importance. An additional strength to our study was the relative proximity of the laboratory at the Universidad del Valle de Guatemala to the study site. This facilitated sample testing and ensured the quality of the analyzed water samples. Our study was also strengthened by the redundancy of our questionnaire. Household interviews involved two enumerators who both asked questions and made observations regarding water sources, sanitation, and hygiene facilities. Therefore, data inconsistencies that would have otherwise have been missed with only one enumerator were identified and dealt with appropriately. Furthermore, our field staff consisted of several enumerators who had previously worked with CDC in Guatemala and were familiar with the PDA technology, which further assured the reliability of survey data.

The sampling methodology that we employed, using aerial photographs, was both a strength and limitation. The advantage to using this methodology is that we were able to produce a random sample of households. However, many potential residential structures were also systematically ineligible to be included because of the methodology. This included structures that were obscured by tree cover in photographs and structures built after 2006 (the year the photos were taken). Additionally, vacant structures and those demolished after the photographs were taken decreased our sample size. Furthermore, because our study was part of a larger pilot study, the investigation into water quality was not designed with a large enough sample size to reach statistical significance. Therefore, the number of observations that we were able to use in our multivariate models was low for the number of predictors that we considered. Another limitation was that “source” water samples were not initially collected correctly for bottled water and had to be re-collected. We found that there was no significant difference in the amount of EC contamination between the new and old samples. Nevertheless, the substitution of the new samples into the analyses of the differential in water quality between the source and POC may have affected our results. We also lost data on chlorination at the source, mostly because we did not evaluate the chlorine residual levels from the re-collected bottled water samples. Chlorination data from water samples of all types were missing at POC as well, possibly due to supervisory issues during collection. However, given the almost complete lack of water chlorination in the rest of the samples, it is likely that there was also no free chlorine residual in the missing samples. An additional limitation associated with water testing was the rounding required when MPNs reached the kit detection limits. We rounded to the next highest whole number for samples at the upper detection limit, a conservative estimate that may have affected the results of the analysis. We also rounded MPNs at the detection minimum (< 1) to zero, and then added 1 for log transformation, which changed the interpretation of the results slightly.

Conclusion

In summary, we that found water contamination levels did not significantly differ between the source and POC for both TC and EC. We found a lack of chlorination at both the source and POC, and found that the amount of EC in source water was positively associated with the amount of EC found in water at the POC. These findings provide a better understanding of the complex relationship between water quality at the source and the POC. In particular, they highlight the effect of deficient water treatment and emphasize the fact that households in Nueva Santa Rosa regularly consume water that is fecally contaminated. These results will be important to the Ministry of Public Health and Social Welfare of Guatemala (MSPAS) as they create policy and take actions to counter their public health challenges.

PUBLIC HEALTH IMPLICATIONS AND FUTURE DIRECTIONS

- Drinking water at both the source and POC was of highest quality (with respect to TC and EC) in bottled water, followed by piped water to the household premise. This supports the JMP recommendations encouraging access to improved water supplies.
- Water contamination at the source was the only predictor found to be significantly associated with water contamination at the POC. This emphasizes the need for safe drinking water sources and the need to protect water sources from fecal contamination.
- Nearly three-quarters of households in Nueva Santa Rosa regularly consume drinking water that does not meet WHO and EPA water quality standards, despite much of this water coming from "improved" water sources. Water samples at both the source and POC were also found to not be chlorinated. Therefore, interventions and programs focused on improving the quality of drinking water in Nueva Santa Rosa could focus on:
    - Protecting water sources from sources of fecal contamination;
    - Municipal water treatment and disinfection (chlorination) for networked water systems;
    - Point-of-use water treatment through chlorination to both disinfect contaminated water and provide a free chlorine residual to protect against re-contamination.

- Given the established link between water contamination at the POC and diarrheal illness, there may be a substantial amount of diarrheal illness in Nueva Santa Rosa on account of the high levels of fecal contamination in drinking water at the POC.
- We are currently compiling a report for the Ministry of Public Health and Social Welfare of Guatemala (MSPAS) detailing the results of this study, and highlighting the challenged water quality and lack of free chlorine residual levels in drinking water at the POC. This information will assist MSPAS to formulate policy and take actions to address these public health issues.
- Future directions should also include replicating this investigation in other locations using larger sample sizes to assess how the findings from fully powered studies compare to those of this pilot study. These future studies would ideally be done both in Guatemala and in other regions to compare the results in different settings. It will also be important to attempt to include more households that utilize bottled water and piped water in order to better understand the source-POC relationship with these water sources as their use becomes more common.

REFERENCES

1. WHO and UNICEF, *Progress on Drinking Water and Sanitation Special Focus on Sanitation*. 2008, World Health Organization and UNICEF: Geneva.
2. WHO, *The world health report 2002 - Reducing Risks, Promoting Healthy Life*. 2002, World Health Organization: Geneva.
3. WHO, *The global burden of disease: 2004 update.* 2008, World Health Organization: Geneva.
4. Prüss-Üstün, A., et al., *Safer water, better health cost benefits and sustainability of interventions to protect and promote health*. 2008, World Health Organization: Geneva.
5. Fewtrell, L., et al., *Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis.* Lancet Infect Dis, 2005. **5**(1): p. 42-52.
6. Clasen, T., et al., *Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis.* BMJ, 2007. **334**(7597): p. 782.
7. Wright, J., S. Gundry, and R. Conroy, *Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use.* Trop Med Int Health, 2004. **9**(1): p. 106-17.
8. WHO, *Guidelines for Drinking-water Quality Third Edition Incorporating the First and Second Addenda*. 2008, World Health Organization: Geneva.
9. Environmental Protection Agency, *National Primary Drinking Water Regulations*. 2009.

10. WHO, *Guidelines for drinking-water quality, 2nd edition*. 1997, World Health Organization: Geneva.

11. Tallon, P., et al., *Microbial indicators of faecal contamination in water: A current perspective.* Water Air and Soil Pollution, 2005. **166**(1-4): p. 139-166.

12. Leclerc, H., L. Schwartzbrod, and E. Dei-Cas, *Microbial agents associated with waterborne diseases.* Crit Rev Microbiol, 2002. **28**(4): p. 371-409.

13. Cairncross, S., et al., *The public and domestic domains in the transmission of disease.* Trop Med Int Health, 1996. **1**(1): p. 27-34.

14. Oswald, W.E., et al., *Fecal contamination of drinking water within peri-urban households, Lima, Peru.* Am J Trop Med Hyg, 2007. **77**(4): p. 699-704.

15. Pickering, A.J., et al., *Hands, water, and health: fecal contamination in Tanzanian communities with improved, non-networked water supplies.* Environ Sci Technol, 2010. **44**(9): p. 3267-72.

16. Rufener, S., et al., *Quality of drinking-water at source and point-of-consumption--drinking cup as a high potential recontamination risk: a field study in Bolivia.* J Health Popul Nutr, 2010. **28**(1): p. 34-41.

17. Eshcol, J., P. Mahapatra, and S. Keshapagu, *Is fecal contamination of drinking water after collection associated with household water handling and hygiene practices? A study of urban slum households in Hyderabad, India.* J Water Health, 2009. **7**(1): p. 145-54.

18. Trevett, A.F., R.C. Carter, and S.F. Tyrrel, *Water quality deterioration: A study of household drinking water quality in rural Honduras.* International Journal of Environmental Health Research, 2004. **14**(4): p. 273-283.

19. Jensen, P.K., et al., *Domestic transmission routes of pathogens: the problem of in-house contamination of drinking water during storage in developing countries.* Trop Med Int Health, 2002. **7**(7): p. 604-9.

20. Momba, M.N.B. and T.L. Notshe, *The microbiological quality of groundwater-derived drinking water after long storage in household containers in a rural community of South Africa.* Journal of Water Supply Research and Technology-Aqua, 2003. **52**(1): p. 67-77.

21. Han, A.M., et al., *Contamination of drinking water during collection and storage.* Trop Geogr Med, 1989. **41**(2): p. 138-40.

22. Roberts, L., et al., *Keeping clean water clean in a Malawi refugee camp: a randomized intervention trial.* Bull World Health Organ, 2001. **79**(4): p. 280-7.

23. Khairy, A.E., et al., *The sanitary condition of rural drinking water in a Nile Delta village. I. Parasitological assessment of 'zir' stored and direct tap water.* J Hyg (Lond), 1982. **88**(1): p. 57-61.

24. Chidavaenzi, M.T., et al., *An evaluation of water urns to maintain domestic water quality.*, in *24th WEDC Conference*, J. Pickford, Editor. 1998, WEDC: Islamabad, Pakistan. p. 249 - 253.

25. Mazengia, E., et al., *Effective and culturally acceptable water storage in Zimbabwe: maintaining the quality of water abstracted from upgraded family wells.* Journal of Environmental Health, 2002. **64**(8): p. 15-8, 25.

26. Vanderslice, J. and J. Briscoe, *All Coliforms Are Not Created Equal - a Comparison of the Effects of Water Source and in-House Water Contamination*

on Infantile Diarrheal Disease. Water Resources Research, 1993. **29**(7): p. 1983-1995.

27. CDC, *Safe Water System Handbook. Safe Water Systems for the Developing World: A Handbook for Implementing Household-Based Water Treatment and Safe Storage Projects*. 2001, US Department of Health & Human Services, Centers for Disease Control and Prevention: Atlanta, GA.
28. Quick, R.E., et al., *Narrow-mouthed water storage vessels and in situ chlorination in a Bolivian community: a simple method to improve drinking water quality.* Am J Trop Med Hyg, 1996. **54**(5): p. 511-6.
29. Quick, R.E., et al., *Diarrhea prevention through household-level water disinfection and safe storage in Zambia.* Am J Trop Med Hyg, 2002. **66**(5): p. 584-9.
30. Sobsey, M., *Managing Water in the Home: Accelerated Health Gains from Improved Water Supply*. 2002, World Health Organization: Geneva.
31. Semenza, J.C., et al., *Water distribution system and diarrheal disease transmission: a case study in Uzbekistan.* Am J Trop Med Hyg, 1998. **59**(6): p. 941-6.
32. Luby, S., et al., *A low-cost intervention for cleaner drinking water in Karachi, Pakistan.* Int J Infect Dis, 2001. **5**(3): p. 144-50.
33. Mintz, E., et al., *Not just a drop in the bucket: expanding access to point-of-use water treatment systems.* Am J Public Health, 2001. **91**(10): p. 1565-70.
34. de Koning, H.W., K.R. Smith, and J.M. Last, *Biomass fuel combustion and health.* Bull World Health Organ, 1985. **63**(1): p. 11-26.

35. Gilman, R.H. and P. Skillicorn, *Boiling of drinking-water: can a fuel-scarce community afford it?* Bull World Health Organ, 1985. **63**(1): p. 157-63.

36. Luby, S.P., et al., *Limited effectiveness of home drinking water purification efforts in Karachi, Pakistan.* Int J Infect Dis, 2000. **4**(1): p. 3-7.

37. Kaltenthaler, E.C., B.S. Drasar, and C.W. Potter, *The use of microbiology in the study of hygiene behaviour.* Microbios, 1996. **88**(354): p. 35-43.

38. Pinfold, J.V., *Faecal contamination of water and fingertip-rinses as a method for evaluating the effect of low-cost water supply and sanitation activities on faeco-oral disease transmission. I. A case study in rural north-east Thailand.* Epidemiol Infect, 1990. **105**(2): p. 363-75.

39. Ologe, J., *Household Water in rural Kwara. In: Water, Engineering and Development in Africa: Proceedings of the 15th WEDC Conference.*, in *15th WEDC Conference*. 1989: Kano, Nigeria. p. 96-99.

40. Blum, D., et al., *The Imo State (Nigeria) Drinking Water Supply and Sanitation Project, 1. Description of the project, evaluation methods, and impact on intervening variables.* Trans R Soc Trop Med Hyg, 1990. **84**(2): p. 309-15.

41. Hoque, B.A., et al., *Research methodology for developing efficient handwashing options: an example from Bangladesh.* J Trop Med Hyg, 1995. **98**(6): p. 469-75.

42. Pinfold, J.V., *Faecal contamination of water and fingertip-rinses as a method for evaluating the effect of low-cost water supply and sanitation activities on faeco-oral disease transmission. II. A hygiene intervention study in rural north-east Thailand.* Epidemiol Infect, 1990. **105**(2): p. 377-89.

43. Miranda, M., et al., *[State of the quality of drinking water in households in children under five years in Peru, 2007-2010]*. Rev Peru Med Exp Salud Publica, 2010. **27**(4): p. 506-11.

44. INE. *City Population*. 2002; Available from: http://www.citypopulation.de/Guatemala.html.

45. MeasureDHS, *Guatemala National Maternal and Child Health Survey 1998–1999*. 1999, Measured DHS.

46. WHO and UNICEF, *Core questions on drinking-water and sanitation for household surveys*. 2006, World Health Organization and UNICEF: Geneva.

47. Measure DHS, *DHS Model Questionnaire – Phase 6 (2008-2013). Questionnaires: Household, Woman's, and Man's* 2010, Measure DHS.

48. CDC. *Handwashing: Clean Hands Save Lives*. May, 2011 [cited 2011 July]; Available from: http://www.cdc.gov/handwashing/.

49. Young, B. and J. Briscoe, *A Case-Control Study of the Effect of Environmental Sanitation on Diarrhea Morbidity in Malawi*. Journal of Epidemiology and Community Health, 1988. **42**(1): p. 83-88.

50. CDC, *Chlorine Residual Testing Fact Sheet, CDC SWS Project*. 2008.

51. CDC, *Household Water Treatment Options in Developing Countries: Household Chlorination*. 2008.

TABLES

Table 1. Demographic and Household Characteristics of 139 Households Surveyed in Nueva Santa Rosa, Guatemala, 2010

Variable (N)	n (%)[§] or Mean ± SD
Number of People in Household (139)	4.45 ± (2.20)
Number of Children Younger than 5 Years (139)	0.47 ± (0.70)
Household Wealth Quintile^ (135)	2.09 ± (1.43)
Monthly Income* (129)	
Less than $124	89 (69)
$124 - $372	35 (27)
Greater than $372	5 (4)
Ethnicity (139)	
Xinca	8 (6)
Indigena	4 (3)
Ladino	123 (91)
Live in Urban Area[ψ] (139)	29 (21)
Mother's Education Level (139)	
None	27 (19)
Part/All Primary	92 (66)
Part/All Secondary	14 (10)
Trade School/University	6 (4)
Type of Water Source[Ж] (139)	
Bottled Water	46 (33)
Piped Water on Premise	47 (34)
Other Improved Water Source	30 (22)
Unimproved	16 (12)
Store Water (205)	188 (93)
Treat Water Before Drinking (139)	44 (32)
Storage Container Covered (132)	106 (80)
Storage Container Opening Less than 3cm (132)	56 (42)
Type of Household Sanitation† (139)	
Improved	99 (71)
Shared (Improved)	8 (6)
Unimproved	24 (17)
No Facilities	8 (6)
Interviewee Washes Hands After Defecating or Changing Baby (139)	85 (61)

§ Percentages may not add to 100% because of rounding

^ Household Wealth Quintile was categorized into equal quintiles, with zero being the lowest wealth quintile and four being the highest wealth quintile (see Methods)

* Value converted from August, 2010 *Quetzales* to August, 2010 US Dollars

ψ "Urban" defined as households located in 200m x 200m map grids that contained 52 or more potential residential-associated roofs (See Methods)

Ж “Piped Water on Premise” includes tap connections inside the house or on the patio. “Other Improved Sources” include public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. “Unimproved Sources” include unprotected springs, surface water, and water from water trucks.

† “Improved” includes flush or pour-flush toilet/latrine, ventilated improved pit [VIP] latrine, pit latrine with slab, or composting toilet. “Shared” includes improved facilities shared by two or more households. “Unimproved” includes pit latrines without slab or platform, hanging latrines, and bucket latrines.

Table 2. Free Chlorine Residual Levels in Water Samples Taken at the Source and Point of Consumption in Households in Nueva Santa Rosa, Guatemala, 2010

		Free Chlorine Residual Level			
	Type of Water[§] (N)	0.0 mg/L	0.1 mg/L	0.2 mg/L - 0.4 mg/L	≥ 0.5 mg/L
Source^	Bottled Water (5)	5	0	0	0
	Piped Water on Premise (47)	45	0	1	1
	Other Improved Water Source (30)	30	0	0	0
	Unimproved (10)	10	0	0	0
	Total (92)	90	0	1	1
Point of Consumption*	Bottled Water (20)	18	1	1	0
	Piped Water on Premise (24)	24	0	0	0
	Other Improved Water Source (22)	22	0	0	0
	Unimproved (14)	14	0	0	0
	Total (80)	78	1	1	0

§ "Piped Water on Premise" includes tap connections inside the house or on the patio. "Other Improved Sources" include public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. "Unimproved Sources" include unprotected springs, surface water, and water from water trucks.

^ Missing 47 observations at Source: 6 Unimproved and 41 Bottled. Bottled Water observations missing due to re-collection of bottled water samples (See Methods).

* Missing 59 observations at Point of Consumption: 26 Bottled, 23 Piped, 8 Other Improved, 2 Unimproved

Table 3. Water Sample Contamination with Total Coliforms (TC) and *E. coli* (EC) at the Source and Point of Consumption in Nueva Santa Rosa, Guatemala, 2010

		Total Coliform Contamination (MPN)					
		Source			**Point of Consumption**		
Type of Water Source§	N (%)^	Median TC*	25th and 75 Percentile	$10^{\text{Median TC}} - 1$	Median TC*	25th and 75 Percentile	$10^{\text{Median TC}} - 1$
Bottled Water	46 (33)	2.42	(1.62, 3.39)	263	2.74	(1.70, 3.39)	547
Piped Water on Premise	47 (34)	3.24	(2.04, 3.84)	1,751	3.24	(2.23, 4.16)	1,751
Other Improved	30 (22)	3.48ψ	(2.79, 3.92)	3,007	3.44ψ	(3.05, 4.20)	2,748
Unimproved	16 (12)	3.33ψ	(2.76, 4.62)	2,118	4.03ψЖ	(3.63, 4.77)	10,618
Overall	139	3.08	(2.21, 3.92)	1,209	3.24	(2.26, 4.12)	1,733

		***E. coli* Contamination (MPN)**					
		Source			**Point of Consumption**		
Type of Water Source§	N (%)^	Median EC*	25th and 75 Percentile	$10^{\text{Median EC}} - 1$	Median EC*	25th and 75 Percentile	$10^{\text{Median EC}} - 1$
Bottled Water†	46 (33)	0.00	(0.00, 0.30)	0	0.00	(0.00, 0.91)	0
Piped Water on Premise	47 (34)	1.16ψ	(0.48, 2.03)	13	1.16ψ	(0.48, 1.78)	13
Other Improved	30 (22)	2.26ψЖ	(1.49, 2.54)	181	1.90ψЖ	(1.08, 2.52)	78
Unimproved	16 (12)	1.70ψ	(1.31, 2.14)	49	1.63ψЖ	(1.21, 2.92)	41
Overall	139	0.98	(0.00, 2.11)	9	1.03	(0.00, 1.85)	10

§ "Piped Water on Premise" includes tap connections inside the house or on the patio. "Other Improved Sources" include public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. "Unimproved Sources" include unprotected springs, surface water, and water from water trucks.

^ Percentages may not add to 100% because of rounding

* Units are Log_{10}((MPN *E. coli* + 1) / 100mL) or Log10((MPN Total Coliforms + 1)

ψ Indicates $p < 0.05$ compared to Bottled Water

Ж Indicates $p < 0.05$ compared to Piped Water to Dwelling

† A value of 0 corresponds to < 1, the lower limit of detection

Table 4. Unadjusted Relationship Between Household Water, Sanitation, and Hygiene Characteristics and *E. coli* Contamination at the Point of Consumption in Nueva Santa Rosa, Guatemala, 2010

Variable§	Beta^	Standard Error	p-value
Type of Water Source*			
Unimproved	Ref	Ref	-
Bottled Water	-1.56ψ	0.27	< 0.001
Piped Water on Premise	-0.86ψ	0.27	0.002
Other Improved Water Source	-0.22	0.29	0.442
Storage Container Opening Less than 3cm	-0.4ψ	0.18	0.030
Storage Container Covered	-0.58ψ	0.23	0.011
Type of Household Sanitation			
No Facilities	Ref	Ref	-
Improved	-1.2ψ	0.38	0.002
Shared (Improved)	-0.35	0.51	0.492
Unimproved	-0.52	0.42	0.221
Sanitary Disposal of Baby's Feces	-0.58ψ	0.24	0.016
Correct Hand Washing	-0.4ψ	0.20	0.047
Amount of *E. coli* in Source Water^	0.57ψ	0.07	< 0.001
Storage Container Located on Floor, in Reach of Children	0.49ψ	0.23	0.042
Time it Takes to Collect Water and Return	0.03ψ	0.01	0.042
Feces in the Hand Washing Area	0.86ψ	0.20	< 0.001
Treat Water Before Drinking	-0.29	0.20	0.150
Day Without Source Water in the Last 2 Weeks	-0.07	0.24	0.770
Place to Wash Hands Within 10 m of Sanitation Facilities	-0.43	0.27	0.110
Interviewee Washes Hands After Defecating or Changing Baby	0.17	0.19	0.380
Animals Stay Inside the House During the Day or Night (With Exception of Dogs or Cats)	0.63	0.34	0.065
Interviewee Washes Hands Before Eating, Cooking, or Serving Kids	0.63	0.78	0.420

§ All variables coded as 0 = No, 1 = Yes, except *E. coli* in Source Water and Time to Collect Water and Return, which were continuous variables

^ Units of Beta and *E. coli* in Source Water are Log_{10}((MPN *E. coli* + 1) / 100mL)

* "Piped Water on Premise" includes tap connections inside the house or on the patio. "Other Improved Sources" include public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. "Unimproved Sources" include unprotected springs, surface water, and water from water trucks.

ψ Indicates significant predictors of *E. coli* contamination ($p < 0.05$) at point of consumption

Table 5. Unadjusted Relationship Between Confounders and *E. coli* Contamination at the Point of Consumption in Nueva Santa Rosa, Guatemala, 2010

Confounders§	Beta^	Standard Error	p-value
Household Wealth Quintile	-0.37*	0.06	< 0.001
Ethnicity			
Xinca	Ref	Ref	-
Indigena	-1.73*	0.64	0.008
Ladino	-1.31	0.38	0.876
Mother's Education Level			
None	Ref	Ref	-
Part/All Primary	0.2	0.22	0.363
Part/All Secondary	-0.95*	0.33	0.005
Trade School/University	-1.16*	0.46	0.012
Number of People in Household	0.13*	0.04	0.002
Number of Children Younger than 5 Years	0.47*	0.13	< 0.001
Live in Urban Area	0.04	0.23	0.850

§ All variables coded as 0 = No, 1 = Yes, except Number of People in Household and Number of Children Younger than 5 Years which were continuous, and Household Wealth Quintile which was categorized into equal quintiles, with zero being the lowest wealth quintile and four being the highest wealth quintile (see Methods)

^ Units of Beta are Log_{10}((MPN *E. coli* + 1) / 100mL)

* Indicates significant predictors of *E. coli* contamination (p < 0.05) at point of consumption

Table 6. Final Adjusted and Reduced Regression Models for Significant Predictors of *E. coli* Contamination at the Point of Consumption in Nueva Santa Rosa, Guatemala, 2010

	Model with Interaction Terms N = 112			**Model without Interaction Terms N = 131**		
Variable[§]	Beta^	Standard Error	p-value	Beta^	Standard Error	p-value
E. coli in Source Water^*	0.36	0.10	< 0.001	0.33	0.09	< 0.001
Type of Water Source[ψЖ]						
Unimproved	Ref	-	-			
Bottled Water	-2.03	0.59	< 0.001			
Piped Water on Premise	-2.06	0.61	0.001			
Other Improved Water Source	-1.70	0.46	< 0.001			
Sanitary Disposal of Baby's Feces[Ж]	-1.94	0.47	< 0.001			
Storage Container Covered[Ж]	1.00	0.37	0.009			
Place to wash hands within 10 m of sanitation facilities[Ж]	1.27	0.34	< 0.001			
Confounders†						
Household Wealth Quintile	-0.10	0.07	0.197	-0.18	0.07	0.006
Ethnicity						
Xinca	Ref	-	-	Ref	-	-
Indigena	-0.53	0.57	0.356	-0.40	0.56	0.483
Ladino	-0.42	0.30	0.170	-0.72	0.33	0.028
Mother's Education Level						
None	Ref	-	-	Ref	-	-
Part/All Primary	0.04	0.20	0.848	0.07	0.20	0.729
Part/All Secondary	-0.33	0.30	0.274	-0.36	0.32	0.256
Trade School/University	-0.50	0.47	0.292	-0.39	0.40	0.337
Number of People in Household	0.02	0.04	0.553	0.01	0.03	0.747
Number of Children Younger than 5 Years	0.10	0.12	0.398	0.26	0.12	0.027
Live in Urban Area	0.17	0.19	0.360	0.11	0.20	0.580

Interaction Terms[Ж]				
Sanitary Disposal of Baby's Feces x Unimproved Water	Ref	-	-	
Sanitary Disposal of Baby's Feces x Bottled Water	2.04	0.63	0.002	
Sanitary Disposal of Baby's Feces x Piped Water on Premise	2.19	0.64	0.001	
Sanitary Disposal of Baby's Feces x Other Improved Water	1.93	0.57	0.001	
Storage Container Covered x Place to Wash Hands Within 10 m of Sanitation Facilities	-1.45	0.43	0.001	
Adjusted R^2			0.548	0.407

§ All variables coded as 0 = No, 1 = Yes, except *E. coli* in Source Water, Time to Collect Water and Return, Number of People in Household, and Number of Children Younger than 5 Years which were continuous. Household Wealth Quintile was categorized into equal quintiles, with zero being the lowest wealth quintile and four being the highest wealth quintile (see Methods)

^ Units of Beta and E. coli in Source Water are Log10((MPN *E. coli* + 1) / 100mL)

* Indicates predictors not forced into model whose final p-value was less than 0.05. All predictors entered at 0.05 and removed at 0.05 from the model.

ψ "Piped Water on Premise" includes tap connections inside the house or on the patio. "Other Improved Sources" include public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. "Unimproved Sources" include unprotected springs, surface water, and water from water trucks.

Ж All interaction terms and their lower order terms forced into the model

† All confounders forced into the model

FIGURES

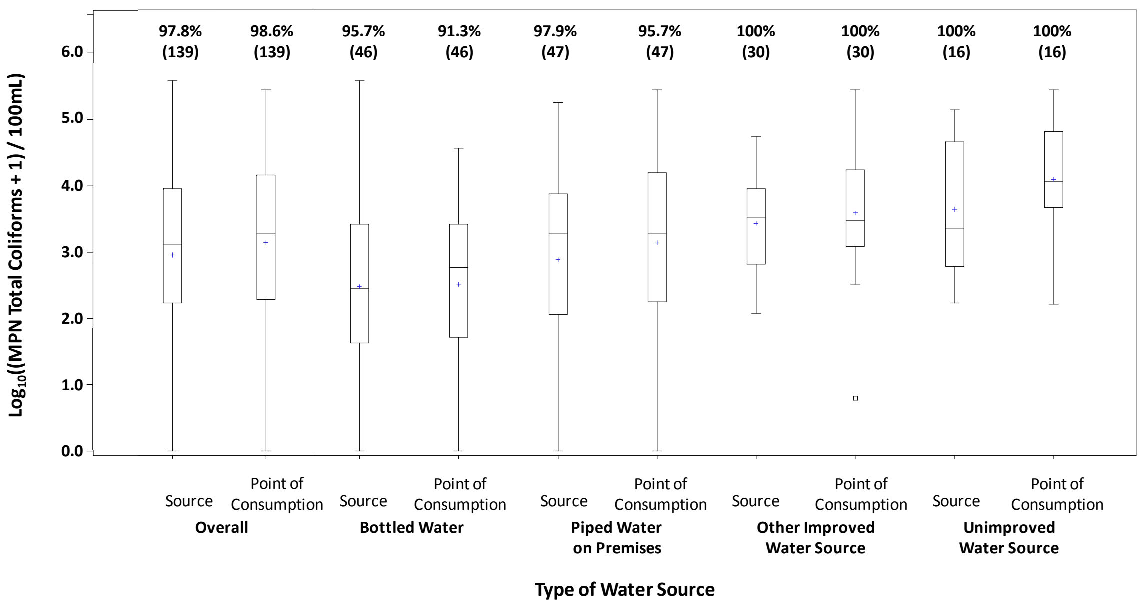

Figure 1. Total coliform contamination levels do not vary significantly between source and point of consumption. Y-axis represents the $\log_{10}$ of 1 plus the Most Probable Number (MPN) of Total Coliforms per 100mL water. X-axis represents the type of water source. “Piped Water on Premises” includes tap connections inside the house or on the patio. “Other Improved Sources” include public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. “Unimproved Sources” include unprotected springs, surface water, and water from water trucks. Boxes represent the interquartile range (IQR) from the lower quartile (25^{th} percentile) to the upper quartile (75^{th} percentile). The “+” inside the box represents the mean of the data while the line represents the median. Whiskers represent minimum and maximum values that fall within 1.5IQR of the lower and upper quartiles. Values beyond the whiskers are outliers, identified as small squares. Percentages above bars indicate the percentage of water samples testing positive for total coliforms. Numbers in parentheses represent sample sizes.

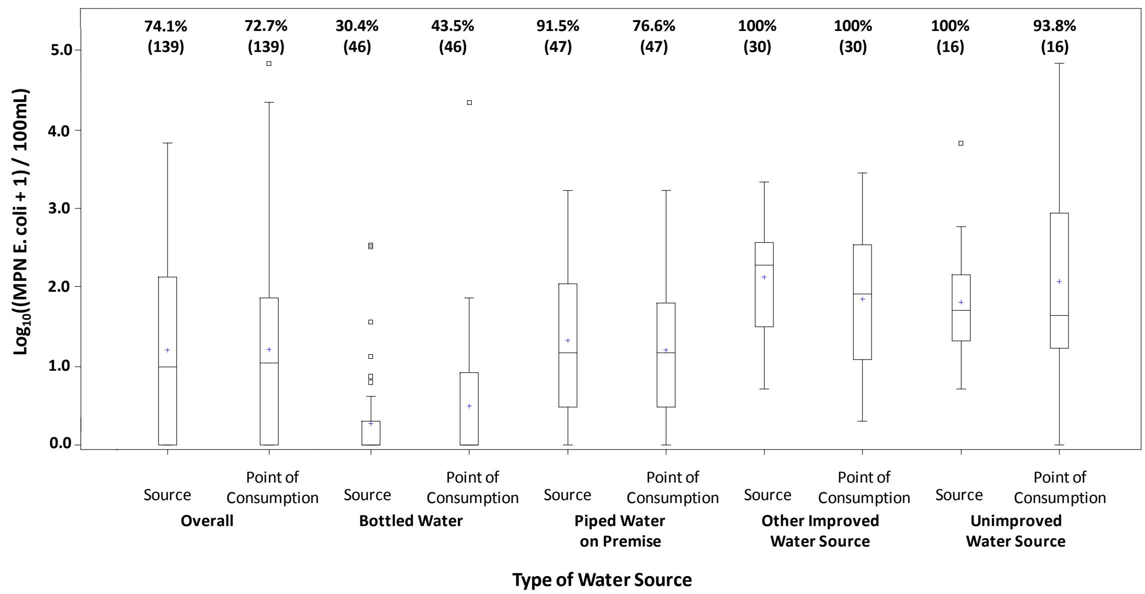

Figure 2. *Escherichia coli* contamination levels do not vary significantly between source and point of consumption. Y-axis represents the $\log_{10}$ of 1 plus the Most Probable Number (MPN) *E. coli* per 100mL water. X-axis represents the type of water source. "Piped Water on Premise" includes tap connections inside the house or on the patio. "Other Improved Sources" include public taps, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection. "Unimproved Sources" include unprotected springs, surface water, and water from water trucks. Boxes represent the interquartile range (IQR) from the lower quartile (25^{th} percentile) to the upper quartile (75^{th} percentile). The "+" inside the box represents the mean of the data while the line represents the median. Whiskers represent minimum and maximum values that fall within 1.5IQR of the lower and upper quartiles. Values beyond the whiskers are outliers, identified as small squares. Percentages above bars indicate the percentage of water samples testing positive for *E. coli*. Numbers in parentheses represent sample sizes.

APPENDIX A: IRB CLEARANCE

EMORY UNIVERSITY

Institutional Review Board

TO: Andrew Thornton
Principal Investigator

DATE: April 19, 2010

RE: **Notification of Submission Determination: No IRB Review Required**
WASH Away Neglected Tropical Diseases: Impact of Water, Sanitation, & Hygiene (WASH) Interventions on Neglected Tropical Diseases in Santa Rosa, Guatemala

The above-referenced study has been vetted by the Institutional Review Board (IRB), and it was determined that it does not require Emory University IRB review. The Principal Investigator is not conducting his own independent research as an agent of Emory University but is working under a Centers for Disease Control and Prevention (CDC) IRB approved study. Mr. Thornton's roles and responsibilities while working on the study will be covered under the CDC IRB approval of the study.

Please note that any changes to the protocol could conceivably alter the status of this research. Accordingly, any substantive changes in the protocol should be presented to the IRB for consideration prior to their implementation in the research.

Sincerely,

Carol Corkran, MPH, CIP
Senior Research Protocol Analyst
This letter has been digitally signed

Emory University
1599 Clifton Road, 5th Floor - Atlanta, Georgia 30322
Tel: 404.712.0720 - Fax: 404.727.1358 - Email: irb@emory.edu - Web: http://www.emory.edu/irb
An equal opportunity, affirmative action university

EMORY UNIVERSITY | Institutional Review Board

TO: Andrew Thornton
Principal Investigator

DATE: December 17, 2010

RE: **Notification of Submission Determination: No IRB Review Required**
Analysis of the Relationship between Water, Sanitation, and Hygiene and the Burden of Diarrheal Disease, Influenza-Like Illness, and Infection with Soil-Transmitted Helminthes in Nueva Santa Rosa, Guatemala

The above-referenced study has been vetted by the Institutional Review Board (IRB), and it was determined that it does not require IRB review because it does not meet the definition of "Research involving Human Subjects" under applicable federal regulations. Based on the information included in the research protocol, the goal of this secondary data analysis is to understand the quality of water in Municipio of Nueva Santa Rosa, Guatemala, and how it affects the health of its citizens. The PI will use existing data collected by the CDC. The PI will not have access to identifiable data or coded-links to identifiers now or in the future. Accordingly, IRB review is not required.

45 CFR Section 46.102(f)(2) defines "Research involving Human Subjects" as follows:

> Human subject means a living individual about whom an investigator (whether professional or student) conducting research obtains:
>
> (1) data through intervention or interaction with the individual, or
> (2) identifiable private information
>
> Intervention includes both physical procedures by which data are gathered (for example, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes. Interaction includes communication or interpersonal contact between investigator and subject. Private information includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record). Private information must be individually identifiable (i.e., the identity of the subject is or may be ascertained by the investigator or associated with the information) in order for obtaining the information to constitute research involving human subjects.

Please note that any changes to the protocol could conceivably alter the status of this research under the federal regulations cited above. Accordingly, any substantive changes in the protocol should be presented to the IRB for consideration prior to their implementation in the research.

Sincerely,

Carol Corkran, MPH, CIP
Senior Research Protocol Analyst
This letter has been digitally signed

Emory University
1599 Clifton Road, 5th Floor - Atlanta, Georgia 30322
Tel: 404.712.0720 - Fax: 404.727.1358 - Email: irb@emory.edu - Web: http://www.emory.edu/irb
An equal opportunity, affirmative action university