Environmental Drivers of Water Quality and Waterborne Disease in the Tropics with a Particular Focus on Northern Coastal Ecuador

by

Karen Levy

B.A. (Stanford University) 1995
M.S. (University of California, Berkeley) 2002
M.P.H. (University of California, Berkeley) 2006

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Environmental Science, Policy, and Management
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Joseph N.S. Eisenberg, Co-Chair
Professor Wayne M. Getz, Co-Chair
Professor Vincent H. Resh
Professor Kara L. Nelson

Fall 2007

The dissertation of Karen Levy is approved:

Co-Chair ____________________________________ **Date** ______________

Co-Chair ____________________________________ **Date** ______________

____________________________________ **Date** ______________

____________________________________ **Date** ______________

University of California, Berkeley
Fall 2007

Environmental Drivers of Water Quality and Waterborne Disease in the Tropics with a Particular Focus on Northern Coastal Ecuador

by Karen Levy

ABSTRACT

Environmental Determinants of Water Quality and Waterborne Disease in the Tropics with a Particular Focus on Northern Coastal Ecuador

by

Karen Levy

Doctor of Philosophy in Environmental Science, Policy, and Management

University of California, Berkeley

Professor Joseph N.S. Eisenberg, co-chair

Professor Wayne M. Getz, co-chair

Infectious diseases, especially waterborne diseases, are inextricably connected to environmental conditions. Not only is the transmission of the organisms that cause them mediated by their immediate environment, but larger scale ecological and social contexts often play a role in determining the incidence of these diseases in a community. Diarrheal illness remains one of the deadliest scourges on the planet, accounting for 2.5 million annual deaths of children worldwide. The body of research in this dissertation addresses the role of environmental drivers of water quality and waterborne disease in tropical countries, with a focus on the Borbón watershed in the northwestern Ecuadorian province of Esmeraldas. Chapter One concentrates on how best to assess microbial contamination in tropical waters, evaluating five techniques for

measuring water quality based on their reliability, utility, and practicality for use in field situations. None of the indicators tested were ideal, and none exhibited a clear and consistent association with diarrheal disease incidence in the household. However, the three *E.coli* indicators performed better overall than the other indicator organisms. Chapter Two takes up the question of recontamination of water in households, describing the results of a controlled experiment to assess contamination of drinking water between the source and point-of-use. On average a more than half-log reduction of indicator organisms was observed between the source of drinking water and its point-of-use, followed by an average 0.2-log increase during storage. Chapter Three addresses sources of variability in water quality at varying timescales over the course of one year in one study village. The results suggest that a "runoff effect," influenced by peak rainfall events, operates at a seasonal timescale, whereas a "concentration effect," influenced by local contamination events, operates at a daily timescale in determining surface source water quality. For household water samples, both seasonality and household level factors affect variability in water quality. The theme of seasonality is carried into Chapter Four, which presents a systematic review and meta-analysis on the seasonal epidemiology of rotavirus in the tropics, concluding that rotavirus responds to changes in climate in the tropics, peaking at the colder and drier times of the year.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Like a child raised in an extended family in Ecuador this dissertation is the product of a collaboration of many many people who helped along the way.

I feel extremely fortunate to have fallen in with Joe Eisenberg and “The Ecuador Project” just when I was seeking a research project at the intersection of health, development, and environmental issues. I found a project that challenged me, taught me new analytical, laboratory, and field skills, and fulfilled my passion for interdisciplinary studies. I also found a mentor who has the patience and perspective to distinguish the distal from the proximal not only in analysis but also in daily life. Joe has always treated me as nothing less than a colleague. I hope that he has not only transmitted to me the ability to generate focused research results and thoughtful analysis, but also that he has infected me with some of his decency as a person. I trust that I will continue to name him in my network of people with whom I share food and talk to about important matters.

I thank Wayne Getz for taking me into his research group when it was time for a new home in ESPM. Wayne not only survived the Five★ Star luxury hotel accommodations in Borbón, but also dared to enjoy some tasty *ratones del monte*. I am thankful for his cheerful support, for entrusting me with full autonomy to do my work, and for cultivating a friendly and challenging (in the good way) lab environment. I have also appreciated the advice and camaraderie of all the “Getzies” who have passed through the lab during my time there.

Kara Nelson’s feedback and critique vastly improved this dissertation, and I am grateful for her input at all stages of the project. She has also served as a role-model and inspiration for how to be a successful academic, contributing member of society, fashion maven, and loving mom all at the same time. I look forward to future overlaps in the Venn diagrams of our lives.

Alan Hubbard has played the role of the wizard behind the stats curtain. I’m not sure what I did to deserve him accepting my cellphone calls at all hours of the day, but it must have been good. I am forever indebted to him and promise to pay him back with plenty of good gossip.

I thank Vince Resh for introducing me to Robert Desowitz, whose writings not only made studying for my oral exams much more fun, but also helped me put the study of infectious diseases into a societal context. His presence as an active and engaged member of the department and his extra efforts to mentor students do not go unnoticed.

Steve Welters has served as a sounding board for academic and life advice since 2001, and held my hand through critical points as I made my way through my PhD program, for which I am deeply grateful. Others who have advised me along various stages of my graduate career path include Claudia Carr, Jeff Romm, and Lee Riley. Jim Trostle, Rick Danielson, Gabriel Trueba, Stefan Wuertz, and Scott Meschke all provided extensive input in the planning of the research involved in this dissertation. Khalid Kadir and Sarah Brownell taught me lab techniques for measuring water quality in Berkeley.

I could not have completed this work without the help of countless people in Ecuador, especially the amazing EcoDess field team. William Cevallos was instrumental in coordinating field logistics and has also been a great friend throughout this project. Nancy Castro was also a wonderful *chiquilina*. I hope to continue to work with them in the future. Karina Ponce helped me with microbiology techniques and also graciously and patiently accepted my periodic intrusions into her space. Melecio Quintero (a.k.a. “El Toro Negro”) shuttled me around and valiantly protected me on the rivers and roads of the Borbón region. I could do without his 6am wake up calls but would never replace his charming smile or his sweet voice belting out Christian songs above the din of the motor. Deny Tenorio, Marixa Renteria, Geovanny Hurtado, Chelo Ortiz, and Jessica Renteria all collected water samples for me, served as my cultural ambassadors, and provided many adventures. Ana Estupiñan took good care of her *pollitos*, and Elizabeth Falconi entered data for me and held down the fort in Quito. I also shared much laughter and camaraderie with Carmen Campaña, Betty Ayoví, Betty Corozo, Simon Quimí, and Francisco Robinzon. In Quito, everyone at the USFQ microbiology lab was quite helpful, especially Deisy Parrales.

I feel quite lucky that Felisa Caicedo agreed to work as my field assistant in Colon Eloy. She diligently collected and processed weekly water samples from the Estero Maria and seven households for a full year. Junior Mina Estupiñan and Patricio Bueno both processed many water samples and were great *compañeros* on the *Equipo del Agua*. I thank all of the village health promoters, especially Amelia Preciado, Maria Ayoví, and Adalín Valencia, who collected data for me each week on water treatment practices in Colon Eloy, and Mirtha Campaz, Denni Orobio, Carmen Ayoví, and Lucrecia Palacios, who assisted me in Telembí and in Playa de Oro.

Of course this study would also not have been possible if not for the graciousness of the people in the study villages who participated in the study and opened their homes to me. Many incredible women cooked meals and provided a home base in the communities: Doña Maura in Telembí, Doña Ruty in Colon Eloy, Doña Neura in Playa de Oro, and many others. Jairo Ayoví has been an incredible friend, helped me in many ways, and taught me how real "local villagers" can be. I also thank my *compadres* Juan Pablo and Emilia for honoring me with the chance to be a godmother to La Niña Llaemi.

I am indebted to many fellow gringos who provided tremendous amounts of assistance in the field. Sarah Bates showed me how to put out laboratory fires, both real and metaphorical, with a certain nonchalance. I have no idea how I would have set up my field lab, learned the techniques I used for water analysis, collected data, or survived the tarimbita ride of the Disco Caribe, the Burbuja, and the Mariscal without her. She is a gem. Others with whom I was graced to share travel and water sampling adventures include Felipe Lowenthal, Sangam Tiwari (my pinch-hitter in Colon Eloy), Laura McLaughlin, Owen Solberg, Darlene Bhavnani, Nina Townsend, Marylin Rodriguez, and Betsy Cowan. Liz, Sophia, and Thea have also been great pals in Ecuador, Berkeley, and Michigan.

I have had the good fortune to share many meals, laughter, conversations, questions, outdoor romps, and other fun and angst with many of my ESPM, SPH, and other grad school peers over the years, especially Arielle Levine, Logan Hennessey, Elizabeth Havice, Sang Lee, Amanda King, Jason Delborne, Jess Kaslow, Letty Brown, Brinda Sarathy, Molly Smith, George Wittemeyer, Shirli Bar-David, Paul Cross, Jamie Lloyd-Smith, Leo “take me to the river” Polansky, Allison Bidlack, Holly Ganz, Sadie Ryan, Manish Desai, Jim Scott, Greg Wilson, Caroline Strömberg, Serena Bodman, Abby Reyes, Jen Dunne, Katrina Schneider, and many many others. My thoughts about development, ecology, globalization, rivers, water, infectious diseases, and countless other topics have been enriched by conversations with all of them.

The Friday Night Dinner Gang, despite our comings and goings, has been a second family for me-- Heidi “Blaze the Trail” Ballard, Stephanie “Just Jump through the Hoops” Ewing, Douglas “We’ll take Three More Just Like Him”

Fischer, Georgianna and Phinneas Fischwing, Lance “Interpretive Dance” Criley, both the new and the old Heather Swartz, and Saint Adriana Sulak. Our mutual excitement, explorations, and conversations about the natural world will continue to serve as a waypoint and compass. I am especially glad for the camaraderie of Andie and Steph in the final push, and for having Heidi as our coach every step of the way.

Others who saw me through to the finish line include Cole Coffee and the sticker crew, most notably the savior of my sanity—Ar “not a statistical package” Killoran. Her encouragement has inspired me to try and make my creative life involve more than just stickers. I was just tickled to have Suzy “Cookie” McIlroy as a housemate and cheerleader in my final year, and that Dave, Misti, Becky, and Isadora Mae Groves moved into the neighborhood and back into my life.

Other important figures along the road include Neo Martinez, who always provided me with a shelter from the stresses of academia while at the same time challenging my thoughts and fostering the exchange of ideas. I have shared many inspiring conversations over the years with Nicole Heller and Tomas Matza, and look forward to many more as our academic and life paths unfold. I thank Anne Wojcicki for chats about science and society, and for always being my #1 fan.

Had I rubbed a genie's bottle or Buddha's belly and created from scratch a friend to live out these grad school years I never could have dreamed up the particularly unique combination of colorful, loyal, dedicated, fun, caring, helpful, healing, hilarious, brilliant, thoughtful, inspiring, tall, vegan, beautiful friend that I have been so graced to have in Tracey Brieger. She is an incredible soul and I thank her for the all-hours chats, parties, costumes, travels, and most of all total acceptance.

My family is my most profound blessing. They have supported and encouraged me through all of my intellectual and worldly pursuits, tolerated my whims, and borne the brunt of my sometimes overfull social schedule. Ema passed onto me a profound appreciation for biology and has been such a strong role-model that it never even gave me pause to think that I could pursue a career in science. I hope that Aba passed onto me a cornucopia of attributes including patience, humor, getting experiments right, grace under pressure, and the ability to throw "pearls to the swine". Together they provide a model that I strive to emulate in generosity of love, curiosity, passion for science, art, food, and thought. I feel lucky to be their daughter. Tali and Naomi are my cheerleaders, my shoulders to cry on, my rocks. They are an inspiration. I don't know what or who or how or why I would be without them. Lee's constant humor and gentle ribbing, Ben's steady flow of Johnny Cash tunes, and the miracles of Devon and Kiva have been especially gratifying additions to the family over the past few years.

This dissertation is dedicated to my second family in Ecuador—Jorge, Marixa, Yaritza, Karen, and Thais Ayoví. They opened their home to me and shared what few riches they have—*cacao*, *coco*, *verde*, *guava, chontaduro, raton del monte*, *armadillo*, *pollo*, and most importantly, *amistad.* They also represent the people in Ecuador and throughout the world who deserve the highest quality drinking water. This work is an effort not to forget that.

~~~

Funding for this work was provided by generous grants from the National Institutes of Allergy and Infectious Diseases (R01#AI050038), the Pacific Rim Research Program of the University of California Office of the President, The University of California Center for Occupational and Environmental Health, the U.C. Berkeley Center for Latin American Studies, and the Switzer Environmental Leadership Program.
~~~

“WHOEVER wishes to investigate medicine properly, should proceed thus: in the first place to consider the seasons of the year, and what effects each of them produces for they are not all alike, but differ much from themselves in regard to their changes. Then the winds, the hot and the cold, especially such as are common to all countries, and then such as are peculiar to each locality. We must also consider the qualities of the waters, for as they differ from one another in taste and weight, so also do they differ much in their qualities. … One ought to consider … the waters which the inhabitants use, whether they be marshy and soft, or hard, and running from elevated and rocky situations, and then if saltish and unfit for cooking; and the ground, whether it be naked and deficient in water, or wooded and well watered, and whether it lies in a hollow, confined situation, or is elevated and cold; and the mode in which the inhabitants live, and what are their pursuits, whether they are fond of drinking and eating to excess, and given to indolence, or are fond of exercise and labor, and not given to excess in eating and drinking. From these things [s]he must proceed to investigate everything else.”

~ Hippocrates, 400 B.C.E.

On Air Water & Places

“Quite apart from details of what I have to say, everyone can surely agree that a fuller comprehension of humanity’s ever-changing place in the balance of nature ought to be part of our understanding of history, and no one can doubt that the role of infectious diseases in the natural balance has been and remains of key importance.”

~ William H. McNeill, 1977

Plagues and Peoples

INTRODUCTION

Throughout history, the fate of human societies has often been determined by infectious diseases, many of which are mediated by the environmental context in which people live (Barrett et al. 1998; Crosby 1986; Dubos 1959; McMichael 2001; McNeill 1976). As the world gets more crowded and natural resources become increasingly stressed, problems of infectious diseases have, and will, become increasingly salient. In recent years the relationships between environmental change and the spread of infectious diseases have become more apparent as large-scale environmental changes become more widespread (Daszak et al. 2000; Morse 1995; Patz et al. 2004; Patz et al. 2000; Weiss & McMichael 2004; Wilson 2001).

I began my graduate school career with an interest in exploring questions of how exploitation of natural resources affects resource-dependent communities, and my path has taken me on a tour of many different fields, from the study of social movements and theories of international development to hydrology and geomorphology to epidemiology and environmental microbiology. While seemingly disparate topics at first glance, all of these fields play an important role in the web of influences, causation, and interconnectivity between environmental change, population growth, poverty, globalization, and community health. The study of infectious diseases inherently integrate these fields because these incredible biological systems highlight the interconnections between humans and their immediate environments as well as their larger socio-political contexts. Understanding infectious diseases must therefore be an interdisciplinary endeavor, because of the complexity of the biological, ecological, and human systems involved.

The Borbón watershed in the northwestern coastal Ecuadorian province of Esmeraldas provides an excellent study system for exploring infectious disease processes and community health in their larger socio-ecological context. Most of the work detailed in this dissertation was carried out in this area. In this region, approximately 125 villages (ranging in size from 20-800 inhabitants) lie along three rivers, the Río Cayapas, Río Santiago, and Río Onzole. All of these rivers flow toward the town Borbón, a town of approximately 5,000 people that serves as the main population center of the region.

The forests of northwest Ecuador are considered to be one of the world's top-ten "hotspots of biodiversity" due to the particularly large numbers of endemic species at risk of extinction (Myers et al. 2000). The region, populated primarily by Afro-Ecuadorians (Whitten 1965), is undergoing intense environmental and social change due to the construction of a new highway along the coast, which connects previously remote villages to the outside world. Construction of the road was completed from Borbón westward to the coast in 1996 and from Borbón eastward to the Andes in 2003. Secondary and tertiary dirt roads off of this twolane asphalt highway are continually being built, mostly for logging, and the area has come to be known as one of the world's top ten deforestation fronts (Myers 1993; Sierra 1999).

In the umbrella project under which this dissertation was carried out, strong trends in infection rates and diarrhea were seen in villages across a gradient of remoteness for viral, bacterial, and protozoan marker pathogens, suggesting that more remote villages have lower infection rates than villages closer to the road. This can be explained by social factors such as increased social connectedness and social capital in more remote villages and higher movement of people into and out of villages closer to the road, which provides opportunities for pathogen incursion (Eisenberg et al. 2006).

Additionally, deforestation may play a role in the observed effect, since the villages close to the road have experienced higher rates of deforestation than the more inland forests. Deforestation can affect community health in several ways. Forest clearing leads to habitat loss of forest-dwelling animals, which may reduce protein intake amongst people that once relied on these animals as a food source. Nutritional status plays a large role in susceptibility to diarrheal disease (Brown 2003). Economic pressures also follow land-clearing, since the natural resource base that villagers depended upon for selective logging has been depleted (Sierra et al. 2003). In conversations with people living in villages near the road, "*Ya no hay bosque*" ("There is no more forest") is on oft-cited source of poverty and economic problems. Vegetative landcover following deforestation also alters hydrologic regimes, which can affect source water quality; I return to this in Chapter 3 of this dissertation. Deforestation is a complex phenomenon affecting many potential causal pathways to disease in this region, and merits further attention in the future.

Causality in infectious diseases is multi-faceted and multi-scaled. Infectious disease risk for a given individual is dependent on the disease status of others in the community, as well as previous exposure to infection and subsequent acquired immunity. In order to address these complexities, epidemiologists are increasingly turning to alternative analytical approaches (e.g., Longini et al. 1988) and frameworks (e.g., Koopman & Lynch 1999; Smith et al. 2005; Susser & Susser 1996) that do not assume disease risks for an individual are independent from those of the community, a standard assumption for epidemiologic models of non-infectious processes.

Richard Levins sums up the complexities of understanding causal connections in infectious disease eloquently (Levins 1995):

> Causation must be understood in the broad sense as residing in much larger wholes than are usually considered by the microbiological or clinical paradigms. Thus, an epidemic is 'caused' by a microorganism of a particular biology, in an environment where it can survive, coming in contact with an exposed and vulnerable population, under conditions that permit successful transmission and infection, allowing enough reproduction within a host to produce disease in the individual and sufficient propagation to initiate enough new cases to affect a population. From this perspective, the current pandemic of cholera can be recognized as

> being possibly ‘caused’ by plankton blooms increasing the populations of Vibrio cholerae, international shipping transporting them in coastal ballast water, the dismantling of social services in Latin America, and the reluctance of governments to acknowledge outbreaks that might affect the tourist trade, among other factors. The plankton blooms can be related to eutrophication of coastal waters owing to erosion, agricultural fertilizers and urban sewage as well as the warming of the seas, and the dismantling of social services can be related to budgetary crises resulting from Third World debt and World Bank insistence on progress through impoverishment.

Similarly, the Borbón region allows one to consider causal factors determining health of human communities from the cellular to the global scales. Diarrheal disease in these villages can be considered to be affected by: genetic exchange of virulence factors between bacteria within the human gut; human-to-human transmission of pathogens within and between households due to poor hygiene; corruption of water at the community level due to poor sanitation; systemic regional socio-cultural issues related to racism and lack of investment in education that limit the ability of communities to address issues of poor water and sanitation infrastructure; introduction of new pathogens and changing social structure due to increased movement of people in the region as a result of road-building; deforestation and subsequent loss of protein sources, medicinal plants, cultural knowledge, economic stability, and local autonomy over natural resources; global consumption patterns that provide the impetus for logging; and systems of international monetary structures in place that allow development banks to finance roads that bleed natural resources into global markets without being held accountable for the ensuing changes in local livelihoods. To claim that community health and incidence of disease in this region is not a result of the interplay of all of these factors would be foolish and arrogant.

Yet while health of human communities is a complex problem that is clearly a result of the multiplicative, rather than additive, effects of these various factors, it is necessary to isolate these factors in order to more fully understand them. While it is important to take note of the larger context in which a biological question is situated, progress in understanding these diseases and the ways in which they might be controlled will be limited if we focus only on large-scale complexities. Scientific inquiry requires focused questions and the testing of specific hypotheses in order to advance knowledge.

The chapters of this dissertation therefore represent an effort to focus on specific questions related to determinants of water quality and waterborne disease in the tropics, but should be understood in the context of the broader question of how the environment, broadly construed, affects health and disease in human communities.

Waterborne Disease

Diarrheal diseases remain one of the deadliest, yet most denied, scourges on our planet. Despite improving trends, diarrhea still accounts for 2.5 million annual deaths of children worldwide (Kosek et al. 2003). To put this in perspective, this means that *every day* twice as many children die of diarrheal illness than were killed in the September 11th attacks on the United States. But there is virtually no mention of diarrheal disease in the popular news media, and very little discussion even in the scientific literature. *Science* magazine's 2006 cover story on global health did not once mention waterborne disease as a leading killer worldwide (Cohen 2006).

The World Health Organization estimates that diarrhea causes 4% of all deaths and 5% of health loss to disability worldwide (World Health Organization (WHO) 2003). Most of the burden falls on children, who account for 90% of the deaths and suffer most from the aggravated malnutrition, stunting, and cognitive problems associated with waterborne illness. According to the Pan-American Health Organization (2002), acute diarrheal disease accounts for 9% of annual registered deaths in Ecuador. Ambitious aims have been set to increase the proportion of people with access to safe drinking water (UN Millennium Project 2005), but to address these goals more attention will need to be paid to the study of transmission and prevention of diarrheal illnesses.

Much is already known about the control of diarrheal disease from a biological and engineering perspective, but many questions remain about how best intervene to prevent diarrheal illness. While the global community has an imperative to use the knowledge already in hand to address the pressing problems of waterborne illness, it is also important to recognize the role of continued research, and to acknowledge how scientific advances have improved our understanding of how to best address the problem of waterborne illness.

The most important advance has been the introduction of Oral Rehydration Therapy (ORT) in the early 1980s (Kosek et al. 2003; Victora et al. 2000). ORT, consisting of the oral administration of sodium, a carbohydrate and water as a mechanism for reducing diarrhea-associated fluid loss and dehydration, has been a major factor in reducing global child mortality. The estimated median number of annual deaths from diarrhea fell from 4.6 million in 1982 to 3.0 million in 1992 and to 2.5 million as estimated in 2000, despite world population growth and the inclusion of China in the most recent analysis. The increased use of ORT over the past two decades has been cited as largely responsible for this decline in mortality from diarrheal disease worldwide (Kosek et al. 2003). Other important interventions likely to have had an impact on mortality caused by diarrhea include the promotion of breastfeeding, improved supplemental feeding, female education, and immunization against measles (Victora et al. 2000). Promising technologies that may lead to future reductions in the global burden of diarrheal disease include the use of zinc therapy (Bhutta et al. 2000), and the introduction of new vaccines against rotavirus (Glass & Parashar 2006).

However, while diarrhea-associated mortality has declined, morbidity has remained high over the last four decades (Kosek et al. 2003). Understanding transmission and prevention of diarrheal disease remains a field of active research. Until the late 1980s, people assumed that poor quality drinking water was the primary source of diarrheal illness (Curtis et al. 2000). However, reviews of the literature by Esrey and others came to the conclusion that other factors, such as available water supply, sanitation, and hygiene practices were equally or more important as water quality in determining the burden of diarrhea (Esrey & Habicht 1986; Esrey et al. 1991).

More recently, researchers in the field of water, sanitation, and hygiene studies have focused on the roles of contamination within versus outside the home. Cairncross et al. (1996) characterize the "public" and "domestic domains" as separate localities with distinct transmission patterns. They suggest that both domains are important, because transmission in the public domain can allow a single case to cause a large epidemic; but transmission in the domestic domain is less dramatic and often ignored, although it may account for a substantial number of cases. The contribution of contamination in public versus private domains has often been debated in terms of the relative importance of "source" versus "point-of-use" water quality. Many studies have shown that the bacteriological quality of drinking water significantly declines after collection (Wright et al. 2004), suggesting that safer household water storage and treatment (point-of-use) should be the recommended focus of intervention efforts (Clasen & Bastable 2003; Gundry et al. 2004; Mintz et al. 1995).

Systematic reviews of the literature on the effectiveness of different interventions provide some conflicting information about the importance of improvements to sanitation in reducing diarrheal disease incidence. Some reviews (Esrey 1996; Gundry et al. 2004; VanDerslice & Briscoe 1995) find evidence for interactions between water supply, water quality, and sanitation, whereas others (Clasen et al. 2007; Fewtrell et al. 2005) find little or no evidence to support the importance of improved sanitation in reducing the burden of diarrheal disease. VanDerslice and Briscoe (2003) point out the importance of the interactions between different routes of transmission. They suggest that the positive impact of improved water quality is greatest for families living under good sanitary conditions and improving drinking water quality would have no effect in neighborhoods with very poor environmental sanitation. This result is supported by a recent modeling showing that the effect of intervening on one transmission pathway depends on the magnitude of other transmission pathways, and therefore, when community sanitation is poor, water quality improvements may have minimal impact (Eisenberg et al. 2007).

What seems clear from the literature on water, sanitation, and hygiene studies is that multiple interacting routes of infection exist, and that all three components of this famous triad are important in transmission of enteric pathogens. More studies are needed that move beyond the household level of analysis, towards a consideration of factors in the community and broader environment in which transmission of these pathogens occurs. As mentioned above, epidemiological analysis tends to focus on proximal risk factors because these can be more easily established and tested, but broader circumstances determine exposure to these proximal risk factors. The chapters of this dissertation therefore represent an investigation of the role of environmental drivers of water quality and waterborne disease in tropical countries, with a focus on the Borbón watershed in the northwestern Ecuadorian province of Esmeraldas. In it I attempt to untangle some of the complex relationships among climate, water, microbes, and humans at varying scales of influence.

In embarking upon this study, I was surprised to find that not only do many questions remain about how water quality affects the transmission of waterborne disease in the tropics, but much uncertainty still exists about how to even go about measuring the quality of water in the tropics. In Chapter One, therefore, I report on a study comparing five different techniques for measuring microbial contamination of tropical waters. These techniques were evaluated based on their reliability, utility, and practicality for use in field situations.

In Chapter Two, I use data on two of these indicators to explore the question of recontamination of water in households, describing the results of a controlled experiment to assess contamination of drinking water between the source and point-of-use. The analysis describes the relative importance of community-level source water quality versus factors determining recontamination of water within the household.

In Chapter Three I address sources of variability in water quality at varying timescales over the course of one year in one study village, focusing on the role of climatic variations in determining water quality both at the source and in the household. Seasonality is known to be an important factor in determining the incidence of infectious diseases (Altizer et al. 2006), and will only become more salient as the effects of climate change become more pronounced. Exploring the way that populations of microbes and the diseases that they cause change seasonally also provides insight into the key environmental variables affecting these organisms and diseases.

The theme of seasonality is carried into Chapter Four, where I explore how climatological variables affect rotavirus disease in particular. In this chapter I take a step even further back, and present a systematic review and meta-analysis on the seasonal epidemiology of rotavirus in the tropics as a whole.

I conclude with some general insights gleaned from the research results, and from the process of carrying out this research.

Literature Cited:

Altizer, S., A. Dobson, P. Hosseini, P. Hudson, M. Pascual, and P. Rohani. 2006. Seasonality and the dynamics of infectious diseases. Ecology Letters 9:467-4848.

Barrett, R., C. W. Kuzawa, T. McDade, and G. J. Armelagos. 1998. Emerging and Re-emerging Diseases: The Third Epidemiologic Transition. Annual Review of Anthropology 27:247-271.

Bhutta, Z. A., S. M. Bird, R. E. Black, K. H. Brown, J. M. Gardner, A. Hidayat, F. Khatun, R. Martorell, N. X. Ninh, M. E. Penny, J. L. Rosado, S. K. Roy, M. Ruel, S. Sazawal, and A. Shankar. 2000. Therapeutic effects of oral zinc in acute and persistent diarrhea in children in developing countries: pooled analysis of randomized controlled trials. The American Journal of Clinical Nutrition 72:1516-1522.

Brown, K. H. 2003. Diarrhea and Malnutrition. The Journal of Nutrition 133:328S-332S.

Cairncross, S., U. Blumenthal, P. Kolsky, L. Moraes, and A. Tayeh. 1996. The public and domestic domains in the transmission of disease. Tropical Medicine and International Health 1:17-34.

Clasen, T., W. P. Schmidt, T. Rabie, I. Roberts, and S. Cairncross. 2007. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. British Medical Journal 334:782-791.

Clasen, T. F., and A. Bastable. 2003. Faecal contamination of drinking water during collection and household storage: the need to extend protection to the point of use. Journal of Water and Health 1:109-115.

Cohen, J. 2006. The New World of Global Health. Science 311:162-167.

Crosby, A. 1986. Ecological Imperialism: The Biological Expansion of Europe, 900-1900. Cambridge University Press, New York.

Curtis, V., S. Cairncross, and R. Yonli. 2000. Domestic hygiene and diarrhoea -- pinpointing the problem. Tropical Medicine and International Health 5:22-32.

Daszak, P., A. A. Cunningham, and A. D. Hyatt. 2000. Emerging Infectious Diseases of Wildlife-- Threats to Biodiversity and Human Health. Science 287:443-449.

Dubos, R. 1959. Mirage of Health: Utopias, Progress, and Biological Change. Harper & Bros., New York.

Eisenberg, J. N. S., W. Cevallos, K. Ponce, K. Levy, S. J. Bates, J. C. Scott, A. Hubbard, N. Vieira, P. Endara, M. Espinel, G. Trueba, L. W. Riley, and J. Trostle. 2006. Environmental change and infectious disease: How new roads affect the transmission of diarrheal pathogens in rural Ecuador. Proceedings of the National Academy of Sciences 103:19460-19465.

Eisenberg, J. N. S., J. C. Scott, and T. C. Porco. 2007. Integrating public health control strategies: Balancing water sanitation, and hygiene interventions to reduce diarrheal disease burden. . American Journal of Public Health 97:846-852.

Esrey, S. A. 1996. Water, waste, and well-being: a multicountry study. American Journal of Epidemiology 146:359-361.

Esrey, S. A., and J.-P. Habicht. 1986. Epidemiologic Evidence for Health Benefits from Improved Water and Sanitation in Developing Countries. Epidemiological Reviews 8:117-128.

Esrey, S. A., J. B. Potash, L. Roberts, and C. Shiff. 1991. Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma. Bulletin of the World Health Organization 69:609-621.

Fewtrell, L., R. B. Kaufmann, D. Kay, W. Enanoria, L. Haller, and J. M. Colford, Jr. 2005. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. Lancet Infectious Diseases 5:42-52.

Glass, R. I., and U. D. Parashar. 2006. The promise of new rotavirus vaccines. New England Journal of Medicine 354:75-77.

Gundry, S., J. Wright, and R. Conroy. 2004. A systematic review of the health outcomes related to household water quality in developing countries. Journal of Water and Health 2:1-13.

Koopman, J. S., and J. W. Lynch. 1999. Individual Causal Models and Population System Models in Epidemiology. American Journal of Public Health 89:1170-1174.

Kosek, M., C. Bern, and R. L. Guerrant. 2003. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bulletin of the World Health Organization 81:197-204.

Levins, R. 1995. Toward an integrated epidemiology. Trends in Ecology & Evolution 10:304.

Longini, I. M., J. S. Koopman, M. Haber, and G. A. Cotsonis. 1988. Statistical inference for infectious diseases: Risk-specific household and community transmission parameters. American Journal of Epidemiology 128:845-859.

McMichael, T. 2001. Human Frontiers, Environments and Disease: Past patterns, uncertain futures. Cambridge University Press, Cambridge.

McNeill, W. H. 1976. Plagues and Peoples. Doubleday, New York.

Mintz, E. D., F. M. Reiff, and R. V. Tauxe. 1995. Safe Water Treatment and Storage in the Home. Journal of the American Medical Association 273:948-953.

Morse, S. S. 1995. Factors in the Emergence of Infectious Diseases. Emerging Infectious Diseases 1:7-15.

Myers, N. 1993. Tropical forests: The main deforestation fronts. Environmental Conservation 20:9-16.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. G. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853-858.

Pan American Health Organization (PAHO) 2002. Regional Core Health Data System-Country Health Profile 2002:

Ecuador.http://www.paho.org/English/DD/AIS/cp_218.htm.Accessed 9/19/2003.

Patz, J. A., P. Daszak, G. M. Tabor, A. A. Aguirre, M. Pearl, J. Epstein, N. D. Wolfe, A. M. Kilpatrick, J. Foufopoulos, D. Molyneux, D. J. Bradley, and Members of the Working Group on Land Use Change and Disease Emergence. 2004. Unhealthy Landscapes: Policy Recommendations on Land Use Change and Infectious Disease Emergence. Environmental Health Perspectives 112:1092-1098.

Patz, J. A., T. K. Graczyk, N. Geller, and A. Y. Vittor. 2000. Effects of environmental change on emerging parasitic diseases. International Journal for Parasitology 30:1395-1405.

Sierra, R. 1999. Traditional resource-use systems and tropical deforestation in a multiethnic region in North-west Ecuador. Environmental Conservation 26:136-145.

Sierra, R., M. Tirado, and W. Palacios. 2003. Forest-Cover Change from Labor- and Capital-Intensive Commercial Logging in the Southern Chocó Rainforests. The Professional Geographer 55:477-490.

Smith, K. F., A. P. Dobson, F. E. McKenzie, L. A. Real, D. L. Smith, and M. L. Wilson. 2005. Ecological theory to enhance infectious disease control and public health policy. Frontiers in Ecology 3:29-37.

Susser, M., and E. Susser. 1996. Choosing a Future for Epidemiology: II. From Black Box to Chinese Boxes and Eco-Epidemiology. American Journal of Public Health 86:674-677.

UN Millennium Project 2005. Investing in Development: A Practical Plan to Achieve the Millennium Development Goals. Earthscan, New York.

VanDerslice, J., and J. Briscoe. 1995. Environmental Interventions in Developing Countries: Interactions and their Implications. American Journal of Epidemiology 141:135-144.

Victora, C. G., J. Bryce, O. Fontaine, and R. Monasch. 2000. Reducing deaths from diarrhoea through oral rehydration therapy. Bulletin of the World Health Organization 78:1246-1255.

Weiss, R. A., and A. J. McMichael. 2004. Social and environmental risk factors in the emergence of infectious diseases. Nature Medicine 10:S70-S76.

Whitten, N. E., Jr. 1965. Class, Kinship and Power in an Ecuadorian Town: The Negroes of San Lorenzo. Stanford University Press, Stanford.

Wilson, M. L. 2001. Ecology and Infectious Disease. Pages 283-324 in J. L. Aron, and J. A. Patz, editors. Ecosystem Change and Public Health: A Global Perspective. The Johns Hopkins University Press, Baltimore.

World Health Organization (WHO) 2003. Water-related Diseases.http://www.who.int/water_sanitation_health/diseases/diarrhoea/en/.Accessed 9/19/03.

Wright, J., S. Gundry, and R. Conroy. 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Tropical Medicine and International Health 9:106-117.

“Have no fear of perfection - you'll never reach it.”

~ Salvador Dali

CHAPTER ONE

INDICATORS OF DRINKING WATER QUALITY FOR TROPICAL WATERS: CASE STUDY IN NORTHERN COASTAL ECUADOR

Introduction

Contaminated drinking water is a major contributor to the problem of diarrheal disease, which continues to plague children worldwide. Despite improving trends, diarrhea still accounts for 2.5 million annual deaths of children worldwide (Kosek et al. 2003). Ambitious aims have been set to increase the proportion of people with access to safe drinking water (UN Millennium Project 2005), but to address these goals methods are needed to assess what constitutes drinking water quality.

The use of organisms to indicate fecal pollution of waters began in the late 1800s, and the coliform group of bacteria in particular has been used as such indicators since the turn of the century, with *E. coli* as the most widely accepted indicator organism in use today (Hazen & Toraznos 1990). Traditionally, the criteria for an ideal indicator of water quality include: useful for all types of water; present whenever enteric pathogens are present; survives longer than the hardiest enteric pathogen; does not grow in natural waters; has an easy-to-perform testing method; grows with a density that has some direct relationship to the degree of fecal pollution; and member of the intestinal microflora of warm-blooded animals (Bonde 1977; Gerba 2000).

Methods for assessing the microbiological quality of water were developed in temperate areas of the world and have subsequently been widely employed in tropical areas, despite differences in physio-chemical, biological, and social and economic differences in tropical and temperate source waters (Hazen & Toraznos 1990). In tropical regions, where diarrheal disease poses the biggest health burden, little research has been carried out about appropriate use of indicators of water quality. Indeed, many of the organisms that practitioners have traditionally used to assess fecal contamination of waterways (total and fecal coliforms, *E.coli*, enterococci) have been found to naturally occur in tropical areas, bringing into question their utility as indicators of fecal contamination (Fujioka et al. 1988; Hazen & Toraznos 1990; Rivera et al. 1988; Solo-Gabriele et al. 2000; Toranzos 1991).

Studies comparing different methods have focused on sensitivity of the test in comparison to standard methods (Eckner 1998; Vail et al. 2003), recovery efficiency, sensitivity and specificity (Pagel et al. 1982; Santiago-Mercado & Hazen 1987), comparison of percentage of samples failing to meet established water quality standards (Noble et al. 2003) or ability to detect low levels of the target organism while also suppressing the growth of non-target organisms(Olstadt et al. 2007). While these are all useful metrics for their respective purposes, here we take an alternate approach, focusing our evaluation of the indicators on their appropriateness for health research studies in tropical settings. In this paper, we expand upon the traditional criteria for a useful indicator of water quality to include criteria particularly appropriate for a tropical context. Our evaluation of indicators focuses on their utility in field situations, where ideal laboratory conditions may not exist. We evaluate five techniques (measuring three different indicator organisms) for assessing water quality in northern coastal Ecuador, based on their reliability, utility, and practicality for use in field situations where much research in tropical health takes place.

Methods

Study Site

The study area lies in northern coastal Ecuador, in the province of Esmeraldas, Canton Eloy Alfaro (Figure 1.1). The 125 villages in this region are situated along the Santiago, Cayapas, and Onzole rivers, all of which drain into the major population center of Borbón. Little sanitation infrastructure exists in these communities. While some people utilize private or community latrines, in household surveys that we have carried out 60% of people dispose of human waste out in the open, by digging a hole, or directly into the rivers. These same rivers serves as the primary water source for 68% of households, and 60% of households reported drinking their water without treating it. High rates of diarrheal disease have been observed in this study area (Eisenberg et al. 2006; Endara et al. 2007; Vieira et al. 2007).

Sample Collection

Water samples were collected in five villages: two with a simple piped water system that transports surface water, two that rely on unimproved surface water from fast-flowing rivers, and one that relies on unimproved surface water from a small stream. In addition to their primary water source (tap or surface water), some villagers also use simple wells and/or collect rainwater as source waters for drinking. Samples were collected from waters identified as drinking water sources by household members and from storage containers within the house. The number of samples collected from each type of source (both directly from the source and from storage containers in the household) is shown in Table 1.1. Sample collection and processing took place between March, 2005 – March, 2006.

Village and household water samples were collected in conjunction with a case-control study of diarrhea incidence in each village. Over the course of each 15-day visit, all cases of diarrhea (defined as three loose stools in a 24-hour period) were identified through daily visits to the households in the community. For every household with a case of diarrhea, a control household without a case of diarrhea was randomly selected. Household drinking water samples and source water samples were collected for as many case and control households as possible (Table 1.2). The Institutional Review Boards of U.C. Berkeley and Universidad San Francisco de Quito approved all contact with human subjects.

Sample Processing

Samples were collected in a manner consistent with the way users collect and serve their drinking water. All samples were collected in Whirlpak bags (Nasco) and kept on ice until processed, within 24 hours. A field laboratory was set up in a house or health dispensary in the village in which the samples were collected, using a bunsen burner inside a modular field hood made from plexiglass and metal to avoid contamination. Agar plates were poured at a microbiology lab in Quito, wrapped individually in Parafilm, packaged in plastic bags, and then transferred to the field site in coolers. Five different assays of water quality were used: Petrifilms (3M) for *E.Coli*, membrane filtration for *E.Coli* using m-Coliblue media (Millipore) and mI agar (BD Difco), membrane filtration for *Enterococci* using mEI agar (BD Difco), and double-layer somatic coliphage using CN-13 *E.coli* as the host strain (ATCC#700609). Enterococci plates were incubated using an egg incubator. A generator was used where electricity was not available. All results were read after 24 hours of incubation at 30±2°C (*E.coli*) or 41±2°C (enterococci)

Several additional indicator organisms and tests of water quality were considered but not employed for a variety of reasons, including lack of specificity, excessive growth, and difficulties encountered with utilization in a field context. *Clostridium* *perfringens* has been suggested as a good indicator for tropical waters (Fujioka & Shizumura 1985). This organism, however, grows only under anaerobic conditions, which would have been difficult to create in the field. *E.coli* was favored over thermotolerant coliforms because of its increased specificity. Furthermore, problems were experienced with growth of fecal coliforms in pilot tests using mFC media (Millipore). The cost of the DelAgua water testing kit prohibited its use for testing of thermotolerant coliforms, given that several different indicators were being tested. Originally, both double and single layer agar methods were piloted for coliphage analysis, but the double agar layer was preferred because of difficulties with melting the agar in the field using the single layer method. An alternate host strain of *E.coli* (DH10) was also tested, but results were found to be inconsistent. Male-specific (F+) coliphage were not assayed because of the requirement that the host bacteria be in log-phase growth. Methods eliminated because of excessive growth during pilot testing include Petrifilm *Enterobacteriaciae* count plates, the PathoScreen Field Test Kit for Hydrogen Sulfide producing bacteria (Hach Method 10032), and Colilert presence/absence tests. Because of the high number of positive results encountered, these methods were not considered discriminatory enough for our purposes. Use of the Colilert Quantitray system (IDEXX) would have been difficult and costly in the field context due to the need for use of the Quanti-Tray Sealer. Multiple tube methods were ruled out because of the glassware involved.

It should be emphasized that testing was optimized for limitations on space during transport by bus, car, and dugout canoe. This is not a trivial concern when considering the types of conditions involved in setting up a field laboratory. Financial considerations are also crucial in a developing country context. Thus the indicator assays chosen were the best suite of indicators given the constraints of our study, which may be representative of constraints of other studies carried out in similar contexts. Below we further describe the assay techniques used.

Petrifilms:

Petrifilm *E.coli*/Total Coliform plates consist of plastic films with grids that are coated with gelling agents and Violet Red Bile nutrients, an indicator of glucuronidase activity (3M 2007). Petrifilms were inoculated with 1 mL of water spread over the gel and incubated for 24 hours. Blue colonies were counted as *E. coli*.

Membrane Filtration:

A sample of water was passed through a 47 mm diameter 0.45 μm cellulose filter (Millipore) and then rinsed with a phosphate-buffered saline solution (pH 7.4 ± 0.2) before transferring to a growth medium plate. The stainless steel membrane filtration apparatus (Millipore) was dipped in alcohol, flame-sterilized, and cooled between each sample. Growth media used included mEI agar (BD Difco; prepared according to EPA Method 1600) (U.S. EPA 2002a), mI agar (BD Difco;

prepared according to EPA Method 1604) (U.S. EPA 2002b), and m-Coliblue (mCB) ampules (Millipore) (U.S. EPA 1999).

Coliphage:

A 2 or 3 mL sample of water was mixed in a sterile tube with three drops of CN-13 *E.Coli* (ATTC# 700609), which was grown up overnight in tryptic soy broth (TSB) with 0.01% Naladixic Acid (Acros). Glass tubes containing 7 mL of 0.7% tryptic soy agar (TSA) plus 0.2% $MgSO_4$ were heated in a water bath and allowed to cool to approximately 50ºC before adding 0.07 mL of a 0.01% solution of Naladixic Acid. The water and *E.Coli* mixture was then added to the agar, agitated, and poured over 100x15mm plates of 1.5% TSA agar containing 0.2% $MgSO_4$ and 0.01% Naladixic Acid. Plates were sealed with Parafilm, inverted, and incubated at air temperature (approximately 28 - 30ºC). Results were read after 18 and 24 hours of incubation. These methods were based on the enumeration procedure described in EPA Method 1601 (U.S. EPA 2001).

<u>Volume of Water Filtered</u>

As noted above, logistical limitations of the research context limited the amount of supplies that could be transported to the field, requiring optimization of protocols for testing the water samples. Thus serial dilutions were not possible because of the multiple tubes that would have been required to transport, and also because of considerations of trade-offs between total numbers of samples versus the need for duplication of assays. Thus, in the majority of cases only one volume was processed for each sample.

In the majority of samples (62-76%, depending on the culture media), a volume of 10 mL was filtered through the membrane filtration unit. To minimize non-detectable results, if a sample was suspected of being particularly clean (rain water, treated drinking water), a volume of 50 mL was filtered; this occurred in 15-25% of samples (depending on the culture media). To minimize the number of results too numerous to count (TNTC), if a sample was suspected of being particularly contaminated (based on previous samples from the same source), a smaller volume (usually 5 mL but in a few cases 1 mL) was filtered; this occurred in 8-22% of samples. If the level of contamination was uncertain, more than one volume was filtered for the same sample; this only occurred in only 1% of samples. In these cases, if one of the volumes tested had a result of zero or TNTC, only the sample in the countable range was included in the analysis. If both results were in a countable range, or if both resulted in a count of zero, the result for both volumes were summed and divided by the total volume sampled. If both were TNTC, then the smaller volume was used in the denominator. In the case of petrifilms, if a sample was expected of being particularly clean, the test was carried out in triplicate; this occurred in 24% of samples. The results of the three tests were summed and divided by 3 mL. In the case of coliphage, a volume of 2 mL (23% of samples) or 3 mL (77% of samples) was used.

In all cases the number of CFUs or PFUs was normalized by the volume of water processed, and multiplied by 100 to get a standardized total count per 100 mL. Non-detects were included in the analysis as one-half of the lower detection limit. A value of 450/plate was assigned to the TNTC results, since the highest reported count was 400/plate. Note that these values for the lower and upper detection limits varied depending on the volume of water processed. For petrifilms, possible results therefore ranged from 16.7 CFU/100 mL (halfway between zero and the lower detection limit of 1 CFU/3 mL x 100 mL) to 45,000 CFU/100 mL (the upper limit of 450 CFU/1 mL x 100 mL), whereas for the membrane filtration techniques, possible results ranged from 1 CFU/100 mL (halfway between zero and the lower detection limit of 1 CFU/50 mL) and 9,000 (the upper limit of 450 CFU/5 mL x 100 mL). The choice of volume to process was based on knowledge of the sample, in order to limit the number of non-detect or TNTC results. Of the samples processed with volumes >10 mL, only 1-3% (depending on the culture media) reached the upper detection limit, and only 0-2% of the samples processed with volumes <10 mL reached the lower detection limit. Thus, the decision to sample a smaller or larger volume and the resulting change in the range of detectable values did not affect a large number of samples. The impact of the variability in volume of water filtered in the different assays is further discussed below.

Quality Control

Some basic quality control measures were utilized. Daily negative controls were used with each test, and if contamination was detected the results for that assay were not used for that day; this occurred only for the coliphage assays. A duplicate of one randomly selected water sample was carried out for all tests for half of the days when labwork was carried out. To test the specificity of the assay for the organism being tested, several isolates were subjected to further testing. Of 117 *E.coli* isolates, 95% (111) also tested positive for *E.Coli* using the Colilert (IDEXX) presence/absence test, which is based on activity of *β*-glucuronidase. Using Kovac's reagent, which tests for the ability to cleave indole from tryptophan, 91% (49 of 54 isolates) were confirmed as *E. coli*. Following the procedures outlined in EPA Method 1600 (U.S. EPA 2002a), 76% (28 of 37 isolates grown on mEI agar) were confirmed as enterococci.

Evaluation of Results

Indicators were evaluated for performance for the criteria below. All analysis was carried out using STATA 9.0 (StataCorp LP, College Station, TX).

I. Reliability

a) Detection Limits

To distinguish one sample from another, indicator assays should be able to detect a range of concentrations and results should be in a consistently countable range. We therefore report the number and percentage of test results falling within a countable range in the main dataset, as well as the results of an experiment to compare the detection limits of the different techniques. In this experiment, nine serial dilutions were carried out on a 50 mL volume of sterile water spiked with human feces and analyzed with the five microbiological assays. The original concentration of indicator organisms in the spiked samples was unknown; the goal was to compare the relative sensitivity and detection limits of the indicator assays to one another. In these assays, a 10 mL sample was used for all membrane filtration techniques, a 3 mL sample for coliphage, and a 1 mL sample for the petrifilms.

b) Consistency of Results

To test for reproducibility of indicator test results, repeated laboratory assays of the same sample (lab duplicates) and repeated samples of the same water source (field duplicates) were compared. Agreement of paired samples of 41 field duplicates and 49 lab duplicates were assessed using both Spearman's Rank Correlation and percentage agreement of presence-absence results. Additionally, we report cross-comparisons of results of the different tests for all samples tested.

II. Utility

a) Presence in Natural Waters

To test the ability of the indicators to distinguish human versus non-human contamination, samples were taken from waters at two sites considered to be free of human contamination: a) a fast-flowing river upstream of any human settlement; and b) surface water flowing over a vertical rock face upstream of human traffic. The distributions of 36 samples from these relatively “undisturbed” sites were compared to those of surface water samples downstream of human habitation using the Wilcoxin rank-sum (Mann Whitney) test.

b) Detection of Contamination Gradient

Samples were collected for 11 consecutive days at the same hour at four collection sites along a small stream located in one village: i) upstream of the village, ii) upstream of the major population center of the village, iii) in the center of town, and iv) at the very downstream end of the town. A gradient of contamination was expected from the point furthest upstream to that furthest downstream in the village. Results from these four sites were plotted using a regression line of best fit and also analyzed using a generalized estimating equation to account for correlation amongst individual sampling days(Liang & Zeger 1986). Robust standard errors were estimated to protect the inference against misspecification of this model.

c) Disease Prediction

An important test of an indicator of water quality is its ability to predict diarrheal disease incidence, since that is ultimately the outcome of interest for health studies. Because more than one sample was often collected for the same household and case was defined at the household level, we pooled results for all drinking water samples for a given household and used this number in the logistic regression analysis. The log of the median value was used in logistic models, with binary case status as the outcome variable. The median household value was used because it is a more robust measure of central tendency than the mean. Data were log-transformed to evaluate the effect of an order of magnitude, rather than a unit change in microbial contamination. The logistic model determines the odds ratio, which is a measure of the power of each indicator to predict diarrheal disease outcome. An odds ratio of 1 indicates that diarrheal disease is equally likely regardless of exposure to indicator organisms in the water. Any value above unity suggests that increasing exposure to indicator organisms increases a household's odds of one of its members having a case of diarrhea. Household exposure was examined with logistic regression at several levels: a) for samples of source water only; b) for samples of water from the source and from household storage containers combined; c) for only samples from household storage containers; and d) for only samples from household storage containers identified by household members as drinking water. If exposure to indicator organisms at the household level were associated with increased diarrheal disease, we would expect the logistic regression to result in an odds ratio above unity. We would also expect an increase in this association from levels a to d, as water type becomes more specific to drinking water consumed in the home.

Following Moe (1991), a stratified analysis was also carried out to determine the risk ratio of a household becoming a case depending on five categories of exposure (0, 1-10, 11-100, 101-1000, and >1000 CFUs/100 mL), as well as depending on high versus low exposure (>1000 vs. <1000 CFUs/100 mL).

III. Practicality

a) Cost

As discussed above, cost of water quality testing is a crucial consideration in the developing country context. Price per test for each indicator is provided and tests are ranked from least to most expensive.

b) Ease of Use

Ease of use in a field situation is also a consideration of paramount importance for research in remote regions or locations without a developed infrastructure. Indicators are therefore ranked according to ease of use in the field.

Results

A total of 1,405 samples were collected and processed, although not all samples were processed with each assay.

I. Reliability

a) Detection Limits

Because they were not limited by sample volume, the three membrane filtration techniques had fewer non-detectable results and therefore a higher percentage of results in a countable range (Table 1.3). Similar results were seen with the experimental serial dilutions. Of the results of ten serial dilutions from eight different samples spiked with human feces, 60% of petrifilm and 78% of coliphage assays had results below the detection limit, compared to 38-53% of the samples analyzed with membrane filtration techniques. This suggests that, as would be expected, the techniques that employed a smaller sampling volume resulted in a non-detectable result more often.

In the detection limits experiment, samples for the comparitive dilution series were spiked with the feces from three different individuals. Enterococci was not present in the samples taken from one of these individuals on sampling days two months apart, even when *E.coli* and coliphage were isolated from his fecal samples in high quantities. Enterococci was isolated from the other two individuals on the same sampling day. This suggests that enterococci is not found in the feces of this individual.

b) Consistency of Results

The duplicate test results for both field and laboratory duplicates are shown in Table 1.4. Surprisingly, across all indicator assays laboratory duplicates showed less consistency than field duplicates. Petrifilms showed the poorest consistency for laboratory duplicates. The results for mCB media must be interpreted with caution because of the small sample sizes.

In Table 1.5, the agreement of test results is shown for all techniques assessed. As might be expected, the *E.coli* tests have the highest agreement with one another, and the coliphage has the lowest agreement with any of the other tests, since it is the only non-bacterial indicator organism. These values are similar to correlations between indicators found by Moe (1991), which ranged from 0.58-0.85.

II. Utility

a) Presence in Natural Waters

The assays that best distinguished between water samples from sites upstream of human contamination versus samples from sites downstream of human contamination were the three tests for *E.coli*: membrane filtration with mCB media (p=0.0001), petrifilms (p=0.01), and mI media (p=0.05). A high amount of enterococci was cultured from the samples taken from the vertical rock face. No significant difference was seen between these samples when assayed for enterococci (using mEI media) or somatic coliphage (Table 1.6).

b) Detection of Contamination Gradient

Results of the 11 consecutive days of sampling are shown in Figure 1.2. The two assays for *E.coli* (membrane filtration with mI agar and petrifilms) performed better than the assays for enterococci and somatic coliphage, in that a significant increase in contamination level was observed from upstream to downstream sites, as expected. Assays for *E.coli* using mCB media were not carried out for this analysis.

c) Disease Prediction

The two-week period prevalence and characteristics of each of the villages visited is summarized in Table 1.2. The results of the logistic analysis suggest that none of the indicator assays are particularly good predictors of disease outcome. Almost all of the 95% confidence intervals of the odds ratios cross unity (Figure 1.3) suggesting that a one order of magnitude increase in exposure to indicator organisms does not have a significant effect on disease outcome. The only exception seen was for *E.coli* measured by mI media, for which significant results were seen for samples from household containers, and for samples from household containers and source waters for a particular house. Using this assay, a one unit increase in log household median counts of *E.coli* in household containers resulted in a 0.29 (0.02-0.65) increased odds of a household having a case of diarrhea; a one unit increase in log household median counts of *E.coli* in household containers combined with samples collected at the source resulted in a 0.18 (0.02-0.37) increased odds of a case of diarrhea in the household. For all of the estimates of association between exposure to indicator organisms in source water and diarrheal disease outcome, the point estimate of the odds ratio is below 1.0, although this effect was not significant. The stratified analyses (results not shown) did not reveal a dose-response effect with different categories, nor a threshold effect with two categories of exposure (<1000 vs >1000 CFUs/100 mL).

III. Practicality

a) Cost

Petrifilms had the lowest cost ($1.00/test), followed by membrane filtration with mI agar ($1.40/test), coliphage ($1.80/test), membrane filtration with mCB ($2.20/test), and mEI ($2.40/test).

b) Ease of Use

Petrifilms were by far the easiest test to carry out and could be handled by someone with very little microbiology experience. Because of the amount of material required and the complexity of the protocols involved (maintaining a host organism, warming up agar tubes, etc.) coliphage was by far the most difficult test to carry out in the field. Of the three membrane filtration techniques, mCB was the easiest to use because no media had to be prepared ahead of time, as with the mI and mEI agars.

Discussion

In an effort to compare the indicators, a summary evaluation is presented in Table 1.7. We assigned each indicator a score from zero to three for each evaluation category, and an overall score for performance of the indicator was calculated by dividing the total number accrued by the total possible number. While this evaluation scheme is subjective, it provides a way to compare the indicators and assess their performance. A justification of the scores is provided here.

I. Reliability

a) Detection limits: The three membrane filtration assays can be used with any volume of water so can be calibrated to the water quality of a given area, so they were each assigned a score of three, with the exception of Enterococci (mEI agar), which only received a score of two because of the false negative result for the spiked sample experiment. The petrifilms and coliphage analysis both were given a score of zero, because given their low sample volumes they would not be able to detect contamination in cleaner water sources.

b) Consistency: There was little basis for differentiating the assays by consistency of results. None of the assays were perfect, so the highest mark assigned was two, and the *E.coli* results for petrifilms had low consistency for laboratory duplicates, so it received only a score of one. Not enough samples were duplicated for the samples tested using mCB agar to evaluate this assay for consistency.

II. Utility

a) Growth in Natural Waters: Neither enterococci nor somatic coliphage could distinguish between samples taken from sites upstream versus downstream of human contamination. The three other tests distinguished between the two, but geometric mean counts of 39-300 CFUs/100 mL were still observed in sites free of human contamination, so they only received a score of one. None of the indicators can differentiate between human and animal contamination, so the observed presence in these waters upstream of human contamination was likely caused by recent contamination with animal feces. However, the three *E.coli* indicators more successfully distinguished between upstream and downstream sites.

b) Detection of Contamination Gradient: The coliphage and enterococci assays did not detect a gradient of contamination, whereas the other two indicators showed a significant increase between upstream and downstream sites within the village, although the slope coefficient was small and day-to-day variability was observed, so they were only assigned a score of two.

c) Disease Prediction: All five assays failed to detect an association between unit changes in geometric mean counts of indicator organisms and diarrheal disease in the household. Only *E.coli* as measured with mI agar detected an association between diarrheal disease in the household and order of magnitude changes in median counts of indicator organisms. The association was fairly small and inconsistent between categories of samples analyzed, so a score of only one was assigned. Petrifilms exhibited an even greater association with disease outcome (higher point estimate for the odds ratio), but this relationship was not significant; the confidence intervals for this indicator assay were greater because of the larger range of values over which petrifilms were estimated (detection limit ranged from 16.7-45,000 CFUs/100 mL). Disease prediction is further discussed below.

III. Practicality

a) Cost: Ideally, indicators of water quality would need to cost less than $1/sample to be practical in a developing country context, and efforts are underway to create a test that can be manufactured for $0.10 (Wright & Khush 2007). Petrifilms were the assay that came the closest to meeting that goal, so a score of two was assigned, *E.coli* analyzed with mI agar and the coliphage assays fell between $1-2 so a score of one was assigned, and assays of enterococci (mEI) and *E.coli* analyzed with mCB media cost over $2 per test, so they received a score of zero.

b) Ease of Use: Petrifilms were very simple to use, earning the maximum of three for this evaluation criteria. Membrane filtration is not easy to carry out in a remote field laboratory, but of the three assays, mCB media was easiest to work with because of the easy-pour ampules; there was no need to pour agar plates ahead of time and the plates also do not require a UV light to be read. Coliphage was cumbersome and difficult to carry out in the field.

Overall Evaluation

According to the evaluation criteria used here, none of the indicators were ideal, although the three *E.coli* indicators performed better overall than the other indicator organisms. While all of the evaluation criteria are important, disease prediction is the most interesting when choosing an indicator for public health research. *E.coli* as measured by mI agar was the only indicator that showed any significant association with diarrheal disease outcome; petrifilms showed an even greater association, but the point estimates were not significant due to the high uncertainty of the estimate (Figure 1.3). Thus, of the three indicator organisms tested, *E.coli* appears to predict disease outcome in the household better than the other organisms tested, although none exhibited a strong association with disease outcome.

Membrane filtration, particularly with mI agar, was the best technique for identifying an association between *E.coli* and diarrheal disease outcome. Of the membrane filtration techniques, using mCB media is the most convenient for the field, although other researchers have found problems with this culture medium (Jensen et al. 2004; Olstadt et al. 2007).

Of the three techniques used to assay *E.coli*, petrifilms were the easiest to carry out but unfortunately this technique is limited by small sample volume. In situations with high levels of contamination, such as our found in our study area, they can be useful for water quality testing, but in sites with improved water supplies their utility would be limited. A previous study comparing petrifilms to mCB, mTEC, and the IDEXX Quanti-Tray system, which all assay *E.coli*, found comparable results when 1-mL samples were consistently used (Vail et al. 2003). Petrifilms can also be useful as a pedagological tool for community educational seminars about drinking water quality. We used petrifilms to teach community health workers with no microbiology training about water contamination and the effects of water treatment. They proved to be a very effective teaching tool, because the health workers could observe bacterial growth in their water supply.

Not only did the enterococci assay have no association with disease outcome but it was less successful than *E.coli* at detecting a difference between sites upstream versus downstream of human contamination, or at detecting a contamination gradient, suggesting that it might be more naturally prevalent in surface waters. The fact that enterococci was consistently absent from one individual's feces is disconcerting, given that presence in human feces is a key requirement of indicator organisms. There is evidence that incidence of various enterococcal species in human feces varies with the geographical location and diets of human subjects (Noble 1978), and in a study testing for enterococci in the human gut enterococci was absent from 43% of neonates and 13% of adults tested (Chenoweth & Schaberg 1990).

Somatic coliphage showed reasonably consistent results, but did not perform well according to the other evaluation criteria. Carrying out coliphage assays in the field was logistically challenging and cannot be recommended. In situations where better laboratory facilities are available, coliphage (specifically FRNA coliphages) might prove useful (Luther & Fujioka 2004).

Indicator Organisms and Diarrheal Disease

There are several possible explanations for the lack of association observed between the indicators assessed and diarrheal disease. First, it is inherently difficult to establish relationships between diarrheal disease and drinking water quality due to lack of specificity in measures of both outcome and exposure. Diarrhea is a non-specific symptom, rather than a disease caused by one particular etiologic agent. In fact only a fraction of diarrheal illness can be traced to a particular infectious agent. Some diarrhea is known to be non-infectious in nature and rather caused by lactose intolerance, irritable bowel syndrome, and various other conditions. The bacterial, viral and protozoan agents of diarrheal illness can be transmitted not only by water but also by food, fomites, personal contact, and sometimes even the air. An individual's nutritional status, immunological naïveté, and genetic factors also play a large role in determining disease outcome. Thus, it is unsurprising that poor associations exist between measures of water quality and diarrheal disease.

In a recent systematic review of water quality and diarrhea, no clear relationship was found between diarrhea and microbial contamination of drinking water at the point of use, although interventions to improve water quality were found to be effective in reducing rates of diarrheal illness (Gundry et al. 2004). Jensen et al (2004) also found no association between bacterial drinking water quality and incidence of diarrhea in a one-year prospective study in rural Pakistan. Even in temperate areas, there is poor evidence relating indicators of drinking water quality to health outcomes (Colford et al. 2007; National Research Council 2005).

Another possible explanation for the weak associations observed here between indicator organisms and diarrheal disease is that the number of cases of diarrhea and number of household water samples might not have been high enough to detect an association, due to lack of statistical power. This can be observed in the differing results for the *E.coli* assays using mCB media compared to those using mI agar. The point estimates are similar, yet the confidence intervals are greater for the mCB assays, due to the lower sample sizes. The estimates for petrifilms were almost significant, despite the high uncertainty in the estimates (Figure 1.3). With even higher sample sizes, it is possible that all of the assays for *E.coli* might show a significant relationship with disease. The assay for enterococci, on the other hand, had higher sample sizes than the mI assay for

E.coli and still showed no association with disease outcome. The coliphage assay also had similar sample sizes to the mI assay, and had a similar sample volume compared to petrifilms, yet showed the least significant associations with disease of any of the indicators tested. Thus, of the indicator organisms tested here, *E.coli* appears to be the best predictor of diarrheal disease.

A similar study comparing several different indicators of fecal contamination in a tropical environment (in Cebu, The Philippines) found a significant association with disease for enterococci using mE followed by testing for esculin hydrolysis on esculin-iron agar, and for *E.coli* using mTEC agar (Moe et al. 1991). This study used a serial cross-sectional case study design, and had the advantage of a higher total number of cases of diarrhea (277) and number of water samples tested (more than 2000), which may have led to the increased significance of results. Using the same analysis techniques as used here (logistic regression of case status versus log of the indicator density per 100 mL), the authors found an odds ratio of 1.18 (1.05-1.31) and 1.14 (0.99-1.31) for case status as predicted by *E.coli* and enterococci, respectively. The results also showed an exposure threshold effect of water quality on disease rates, with exposure of >1000 CFUs/100 mL associated with an increased risk of disease. A similar categorical threshold effect was not observed here, but this might have been due to the lower number of samples in the highest exposure category.

The ability to predict diarrhea using enterococci was also not observed here, which might have had to do with the differences in water samples tested. Our results suggest that enterococci might naturally occur in surface waters in higher numbers than *E.coli*, which might explain why enterococci was not a good predictor of disease in this region. Furthermore, the sources of water tested by Moe et al. (1991) came from protected springs, wells, boreholes, and a piped water supply, as opposed to unprotected surface waters. The majority of samples tested were from boreholes, which were cleaner overall than the source waters we tested in this study. This may have increased the power to detect a differential effect of clean versus contaminated waters.

The final issue related to measuring indicators is measurement variability, which is especially relevant when assessing surface water supplies. Household water quality is influenced by original source water quality, turbidity, matrix effects, periodic contamination events, die-off, and other factors. High variability in estimates of indicator concentrations have been observed in recreational waters at several timescales (Boehm et al. 2002). Yet many epidemiological studies have been based on a low water quality sampling frequency, and sampling is not carried out over a long period of time. Significant differences have also been observed between the ten USEPA-approved tests to detect and quantify total coliforms and *E.coli* (Olstadt et al. 2007).

Given the problems with measuring indicators of fecal contamination as a proxy for disease risk, in what situations might indicators be useful? Indicators as a proxy measurement to assess the relative risk of disease under differing conditions, as used here and in many other health studies (Colford et al. 2007; Gundry et al. 2004; Schwab 2007), is one application, but the practice of using indicator organisms to detect fecal contamination was developed, mostly in temperate areas, with the goal of examining the effectiveness of water treatment practices. Indicators have subsequently been applied in other contexts, such as studies of waterborne disease in developing countries.

Indicators can be used to determine the level of microbial contamination of source water, whether existing water treatment processes are adequate, and whether the integrity of the distribution system has been breached (National Research Council 2005). *Process indicators* demonstrate the efficacy of a process such as water treatment; *Fecal indicators* can indicate the specific presence of fecal contamination; and *Model organisms* can be used to indicate or model the presence and behavior of a particular type of pathogen (Ashbolt et al. 2001).

Thus, even if clear disease-exposure relationships are difficult to establish, indicator organisms can still be useful for testing relative contamination levels of different water sources or containers, for assessing recontamination in the household, for assessing seasonal changes in source waters, and for addressing other basic questions about relative safety of different water sources and usage practices. Indicator organisms can still be quite useful when assessing the effectiveness or efficacy of in-home treatment devices or source water supply improvements, and for monitoring for intrusion of sewage in these systems. They can be useful as process indicators about whether a device works. However, when evaluating treatment processes using indicators, care should be taken not to equate water quality improvements with disease reductions. Until indicators can be made more specific, studies of interventions should maintain focus on disease incidence as the outcome measure of interest, rather than using indicators as a proxy.

Most studies of water quality focus on the individual or household level, but given the interactions between water quality and sanitation (VanDerslice & Briscoe 1995), it might make more sense to evaluate water quality at the community level rather than engage in household-level risk factor analysis. Given the non-linear dynamics of infectious diseases (Eisenberg et al. 2007) and lack of specificity of exposure levels to particular drinking water containers in many developing country contexts, further insight might be gained from exploring higher-, community-level relationships between water quality and disease. This is further explored in Chapter 3.

Ultimately to be useful in health studies, indicators of water contamination need to be more specific, and should not be limited to detecting gross contamination with large sample sizes. But research is also sorely needed to explore new ideas for indicator organisms and assays of fecal contamination. For example, using novel technologies such as microarrays we may be able to look for organisms in water that have associations with disease and develop quick, cheap, easy methods to test for these organisms in the water rather than relying on the organisms we currently know how to culture. We must begin to think outside the box and devote resources to this research if we are to have any hope of making a dent in the Millennium Development Goals.

Literature Cited

3M 2007. Petrifilm E.Coli/Coliform Count Plate Interpretation Guide.www.3m.com/microbiology.Accessed 5/1/07.

Ashbolt, N. J., W. O. K. Grabow, and M. Snozzi. 2001. Indicators of microbial water quality. Pages 289-314 in L. Fewtrell, and J. Bartram, editors. Water Quality: Guidelines, Standards and Health. World Health Organization / IWA Publishing, London.

Boehm, A. B., S. B. Grant, J. H. Kim, S. L. Mowbray, C. D. McGee, C. D. Clark, D. M. Foley, and D. E. Wellman. 2002. Decadal and Shorter Period Variability of Surf Zone Water Quality at Huntington Beach, California. Environmental Science & Technology 36:3885-3892.

Bonde, G. J. 1977. Bacterial indication of water pollution. Advances in Aquatic Microbiology 1:273-364.

Chenoweth, C., and D. Schaberg. 1990. The Epidemiology of Enterococci. European Journal of Clinical Microbiology & Infectious Diseases 9:80-89.

Colford, J. M., Jr., T. J. Wade, K. C. Schiff, C. C. Wright, J. F. Griffith, S. K. Sandhu, S. Burns, M. D. Sobsey, G. Lovelace, and S. B. Weisberg. 2007. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination. Epidemiology 18:27-35.

Eckner, K. F. 1998. Comparison of membrane filtration and multiple-tube fermentation by the Colilert and Enterolert methods for detection of waterborne coliform bacteria, *Escherichia coli*, and Enterococci used in drinking and bathing water quality monitoring in Southern Sweden. Applied and Environmental Microbiology 64:3079-3083.

Eisenberg, J. N. S., W. Cevallos, K. Ponce, K. Levy, S. J. Bates, J. C. Scott, A. Hubbard, N. Vieira, P. Endara, M. Espinel, G. Trueba, L. W. Riley, and J. Trostle. 2006. Environmental change and infectious disease: How new roads affect the transmission of diarrheal pathogens in rural Ecuador. Proceedings of the National Academies of Science 103:19460-19465.

Eisenberg, J. N. S., J. C. Scott, and T. C. Porco. 2007. Integrating public health control strategies: Balancing water sanitation, and hygiene interventions to reduce diarrheal disease burden. American Journal of Public Health 97:846-852.

Endara, P., G. Trueba, O. D. Solberg, S. J. Bates, K. Ponce, W. Cevallos, J. Matthijnssens, and J. N. S. Eisenberg. 2007. Symptomatic and Subclinical Infection with Rotavirus P[8]G9, Rural Ecuador. Emerging Infectious Diseases 13:574-580.

Fujioka, R., and L. K. Shizumura. 1985. *Clostridium perfringens*, a reliable indicator of stream water quality. J. Wat. Poll. Control

Fed. 57, 986–992. Journal of the Water Pollution Control Federation 57:986-992.

Fujioka, R. S., K. Tenno, and S. Kansako. 1988. Naturally occurring fecal coliforms and fecal streptococci in Hawaii's freshwater streams. Toxicity Assessment 3:613-630.

Gerba, C. P. 2000. Indicator Microorganisms in R. M. Maier, I. L. Pepper, and C. P. Gerba, editors. Environmental Microbiology. Academic Press, San Diego.

Gundry, S., J. Wright, and R. Conroy. 2004. A systematic review of health outcomes related to household water quality in developing countries. Journal of Water and Health 2:1-13.

Hazen, T. C., and G. A. Toraznos. 1990. Tropical Source Water. Pages 32-53 in G. A. McFeters, editor. Drinking Water Microbiology: Progress and Recent Developments. Springer-Verlag, New York.

Jensen, P. K., G. Jayasinghe, W. van der Hoek, S. Cairncross, and A. Dalsgaard. 2004. Is there an association between bacteriological drinking water quality and childhood diarrhoea in developing countries? Tropical Medicine and International Health 9:1210-1215.

Kosek, M., C. Bern, and R. L. Guerrant. 2003. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bulletin of the World Health Organization 81:197-204.

Liang, K., and S. L. Zeger. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73:13-22.

Luther, K., and R. Fujioka. 2004. Usefulness of monitoring tropical streams for male-specific RNA coliphages. Journal of Water and Health 2:171-181.

Moe, C., M. Sobsey, G. Samsa, and V. Mesolo. 1991. Bacterial indicators of risk of diarrhoeal disease from drinking-water in the Philippines. Bulletin of the World Health Organization 69:305-317.

National Research Council. 2005. Health Effects Assessment. Pages 53-108 in Committee on Indicators for Waterborne Pathogens; Board on Life Sciences; Water Science and Technology Board, editor. Indicators for Waterborne Pathogens. The National Academies Press, Washington, D.C.

Noble, C. J. 1978. Carriage of group D streptococci in the human bowel. Journal of Clinical Pathology 31:1182-1186.

Noble, R. T., D. F. Moore, M. K. Leecaster, C. D. McGee, and S. B. Weisberg. 2003. Comparison of total coliform, fecal coliform, and enterococcus bacterial indicator response for ocean recreational water quality testing. Water Research 37:1637-1643.

Olstadt, J., J. J. Schauer, J. Standridge, and S. Kluender. 2007. A comparison of ten USEPA approved total coliform/*E.coli* tests. Journal of Water and Health 5:267-282.

Pagel, J. E., A. A. Qureshi, D. M. Young, and L. T. Vlassoff. 1982. Comparison of Four Membrane Filter Methods for Fecal Coliform Enumeration. Applied and Environmental Microbiology 43:787-793.

Rivera, S., T. C. Hazen, and G. A. Toranzos. 1988. Isolation of fecal coliforms from pristine sites in a tropical rain forest. Applied Environmental Microbiology 54:513-517.

Santiago-Mercado, J., and T. C. Hazen. 1987. Comparison of Four Membrane Filter Methods for Fecal Coliform Enumeration in Tropical Waters. Applied and Environmental Microbiology 53:2922-2928.

Schwab, K. J. 2007. Are Existing Bacterial Indicators Adequate for Determining Recreational Water Illness in Waters Impacted by Nonpoint Pollution? Epidemiology 18:21-22.

Solo-Gabriele, H. M., M. A. Wolfert, T. R. Desmarais, and C. J. Palmer. 2000. Sources of *Escherichia coli* in a Coastal Subtropical Environment. Applied and Environmental Microbiology 66:230-237.

Toranzos, G. A. 1991. Current and possible alternate indicators of fecal contamination in tropical waters: a short review. Environmental Toxicology & Water Quality 6:121-130.

U.S. EPA. 1999. National Primary and Secondary Drinking Water Regulations: Analytical methods for chemical and microbiological contaminants and revisions to laboratory certification requirements; Final Rule. Pages 67449-67467 in U.S. Environmental Protection Agency, editor. Federal Register.

U.S. EPA. 2001. Method 1601: Male-specific (F+) and Somatic Coliphage in Water by Two-step Enrichment Procedure. U.S. Environmental Protection Agency Office of Water, Washington, D.C.

U.S. EPA. 2002a. Method 1600: Enterococci in Water by Membrane Filtration Using membrane-Enterococcus Indoxyl-β-D-Glucoside Agar (mEI). U.S. Environmental Protection Agency Office of Water, Washington, D.C.

U.S. EPA. 2002b. Method 1604: Total Coliforms and Escherichia coli in Water by Membrane Filtration Using a Simultaneous Detection Technique (MI

Medium). U.S. Environmental Protection Agency Office of Water, Washington, D.C.

UN Millennium Project 2005. Investing in Development: A Practical Plan to Achieve the Millennium Development Goals. Earthscan, New York.

Vail, J. H., R. Morgan, C. R. Merino, F. Gonzales, R. Miller, and J. L. Ram. 2003. Enumeration of waterborne *Escherichia coli* with Petrifilm plates: Comparison to standard methods. Journal of Environmental Quality 32:368-373.

VanDerslice, J., and J. Briscoe. 1995. Environmental interventions in developing countries: interactions and their implications. American Journal of Epidemiology 141:135-144.

Vieira, N., S. J. Bates, O. D. Solberg, K. Ponce, R. Howsmon, W. Cevallos, G. Trueba, L. Riley, and J. N. S. Eisenberg. 2007. High prevalence of Enteroinvasive Escherichia coli isolated in a remote region of northern coastal Ecuador. American Journal of Tropical Medicine and Hygiene 76:528-533.

Wright, J., and R. Khush. 2007. AQUATEST Activities. Aquatest Research Newsletter 1.

Table 1.1: Sources of water as sampled a) from containers stored in the household and b) directly from the source

	a) Household Containers	b) Source
Water Source	**# Samples**	**# Samples**
Harvested Rainwater	104	6
Unprotected Well	25	14
Piped (Untreated)	259	95
Surface (River)	122	49
Surface (Small Stream)	117	16
Total	627	180

Table 1.2: Characteristics of study villages at the time of each visit. A water treatment plant was installed in Borbón between the first and the second visits.

Village	River Basin	Water System	# House-holds	Population	Month/Year of Visit	Total # Cases	2-week Prevalence per household	# Case Households Sampled	# Control Households Sampled
Borbón	Santiago-Cayapas-Onzole	Piped (pre-water treatment plant)	202*	877*	7/05	40	0.20	8	8
Borbón	Santiago-Cayapas-Onzole	Piped (post-water treatment plant)	202*	862*	3/06	25	0.12	23	24
Colon Eloy	Estero Maria	Surface	170	709	8/05	16	0.09	12	18
Playa de Oro	Santiago	Piped surface	56	235	9/05	3	0.05	3	15
Guayabal	Santiago	Surface	27	107	9/05	3	0.11	3	7
Telembí	Cayapas	Piped surface	46	268	3/05	7	0.39	6	15
Telembí	Cayapas	Piped surface	58	290	1/06	12	0.26	11	12
Trinidad	Cayapas	Surface	15	79	3/05	1	0.07	1	12
Trinidad	Cayapas	Surface	18	89	1/06	3	0.05	3	3

***These numbers represent a sample of the population of approximately 5000 people living in approximately 1000 houses**

Table 1.3: Number and percentage of tests falling within a countable range.

Assay	Organism		Below Detection Limit		Above Detection Limit		In Countable Range	
Membrane Filtration		N	#	%	#	%	#	%
mEI media	Enterococci	1401	176	12.6%	74	5.3%	1151	82.2%
mI media	*E.coli*	1276	194	15.2%	131	10.3%	951	74.5%
mCB media	*E.coli*	585	134	22.9%	39	6.7%	412	70.4%
Other Technique								
Petrifilm	*E.coli*	1404	506	36.0%	36	2.1%	868	61.8%
Coliphage	Somatic coliphage	1117	431	38.6%	10	0.9%	676	60.5%

Table 1.4: Consistency of test results for duplicate tests. Field duplicates are duplicate samples from the same source and Lab duplicates are duplicate assays of the same sample (ρ = Spearman's rank correlation coefficients; P/A = % presence absence agreement) .

	Assay	Field Duplicates	Lab Duplicates
Membrane Filtration	**Enterococci (mEI agar)**	ρ = 0.72 P/A = 0.88 (*n*= 41)	ρ = 0.68 P/A = 0.90 (*n*= 41)
	E.coli **(ml agar)**	ρ = 0.78 P/A = 0.95 (*n*= 41)	ρ = 0.64 P/A = 0.96 (*n*= 49)
	E.coli **(mCB media)**	ρ = 0.72 P/A = 1.00 (*n*= 7)	ρ = 0.25 P/A = 1.00 (*n*= 9)
Other technique	***E.coli*** **(petrifilms)**	ρ = 0.78 P/A = 0.78 (*n*= 40)	ρ = 0.49 P/A = 0.65 (*n*= 49)
	Somatic Coliphage	ρ = 0.78 P/A = 0.83 (*n*= 36)	ρ = 0.72 P/A = 0.77 (*n*= 44)

Table 1.5: Agreement among test results using the five different assays. (ρ = Spearman's rank correlation coefficients; P/A = % presence absence agreement)

		E.coli			Enterococci	Coliphage
		mI media	**mCB media**	**petrifilms**	**mEI media**	**phage**
E.coli	**mI media**	1.00 (*n*=1276)				
	mCB media	ρ = 0.86 P/A = 0.87 (*n*=585)	1.00 (*n*=585)			
	petrifilms	ρ = 0.83 P/A = 0.76 (*n*=1272)	ρ = 0.78 P/A = 0.76 (*n*=583)	1.00 (*n*=1404)		
Entero-cocci	**mEI media**	ρ = 0.72 P/A = 0.88 (*n*=1274)	ρ = 0.74 P/A = 0.82 (*n*=585)	ρ = 0.65 P/A = 0.73 (*n*=1397)	1.00 (*n*=1401)	
Coli-phage	**phage**	ρ = 0.60 P/A = 0.69 (*n*=1113)	ρ = 0.62 P/A = 0.70 (*n*=536)	ρ = 0.57 P/A = 0.70 (*n*=1115)	ρ = 0.51 P/A = 0.66 (*n*=1112)	1.00 (*n*=1117)

Table 1.6: Ability to detect difference between samples from sites upstream vs. downstream of human contamination (*p*-value tests difference between samples using a Two Sample Wilcoxin Rank-Sum (Mann Whitney) test)

			UPSTREAM SITES			DOWNSTREAM SITES		
Organism	**Assay**	***p*-value**	**# Samples**	**Geometric Mean**	**Median (95% C.I.)**	**# Samples**	**Geometric Mean**	**Median (95% C.I.)**
E.coli	**mCB**	**0.0001**	25	39	30 (13-77)	13	415	400 (190-572)
E.coli	**petrifilms**	**0.01**	36	176	217 (119-381)	81	546	400 (300-783)
E.coli	**mI**	**0.05**	36	300	260 (123-473)	60	507	525 (259-674)
Enterococci	**mEI**	0.44	36	567	650 (391-908)	81	435	580 (372-720)
Somatic Coliphage	**Phage**	0.60	32	125	117 (33-300)	51	91	67 (33-133)

Table 1.7: Summary Scores for Assays of Water Quality

	Enterococci (mEI)	*E.coli* (ml)	*E.coli* (mCB)	*E.coli* (petrifilms)	Somatic Coliphage
I. RELIABILITY					
a. Detection Limits	2	3	3	0	0
b. Consistency	2	2	--	1	2
II. UTILITY					
a. Growth in Natural Waters	0	1	1	1	0
b. Detection of Contamination Gradient	0	2	--	2	0
c. Disease Prediction	0	1	0	0	0
III. PRACTICALITY					
a. Cost	0	1	0	2	1
b. Ease of Use	1	1	2	3	0
OVERALL SCORE	5/21 24%	11/21 52%	6/15 40%	9/21 43%	3/21 14%

Figure 1.1: Map of Study Region in Northwestern Ecuador. Dots represent villages in the region. The Santiago, Cayapas, and Onzole Rivers drain into the main town of Borbón (pop. 5000).

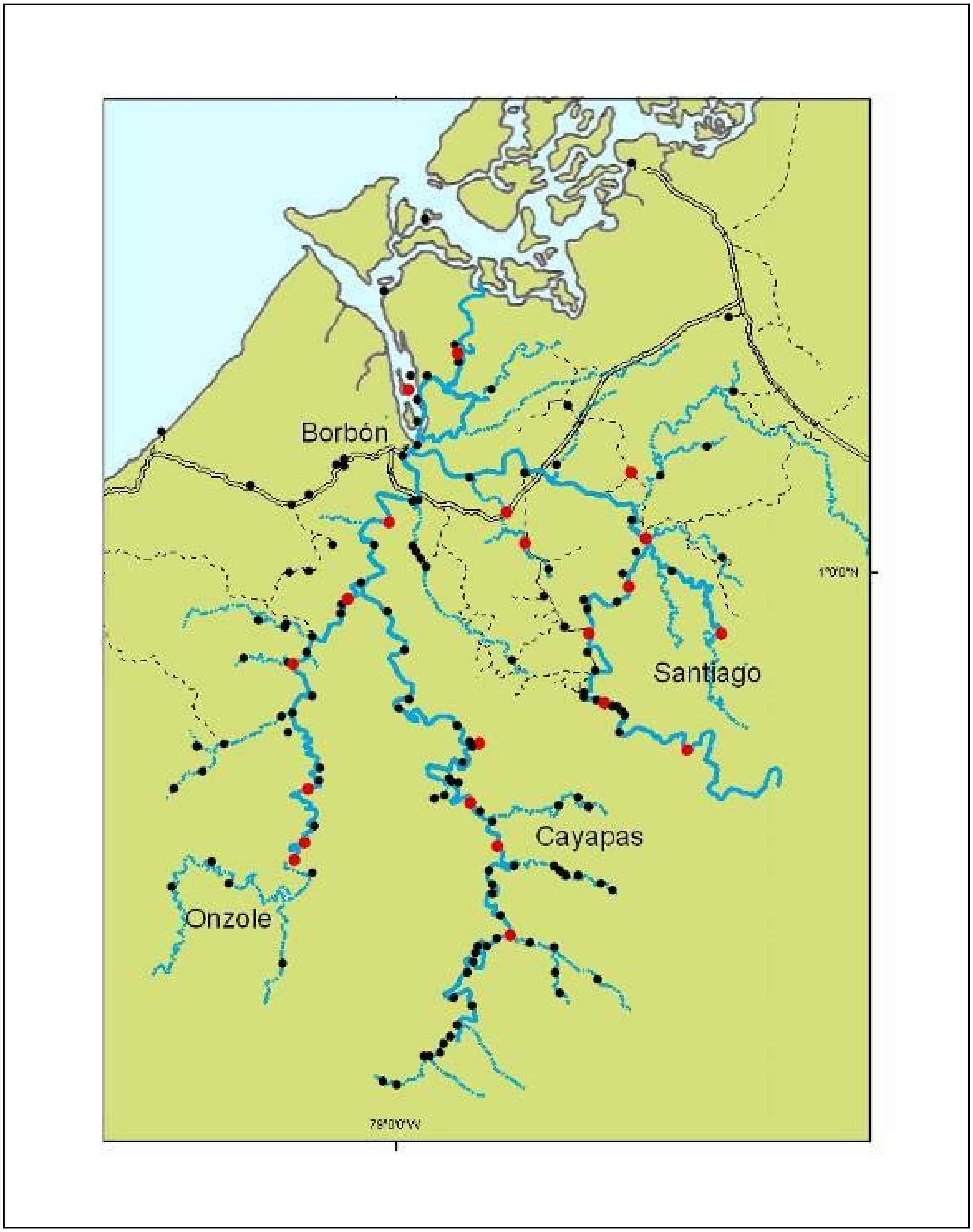


Figure 1.2: Results of Sample Analysis Along a Contamination Gradient. Location 1 is directly upstream of the village, Location 2 is upstream of the main population center of the village, Location 3 is in the center of the village, and Location 4 is at the downstream end of the village. Samples were collected at these sites for ten consecutive days. Dashed lines represent individual days of sampling, and the solid line represents the overall trend based on a simple linear regression. Beta values are reported for the results of the generalized estimating equation regression, which accounts for correlation among sampling days.

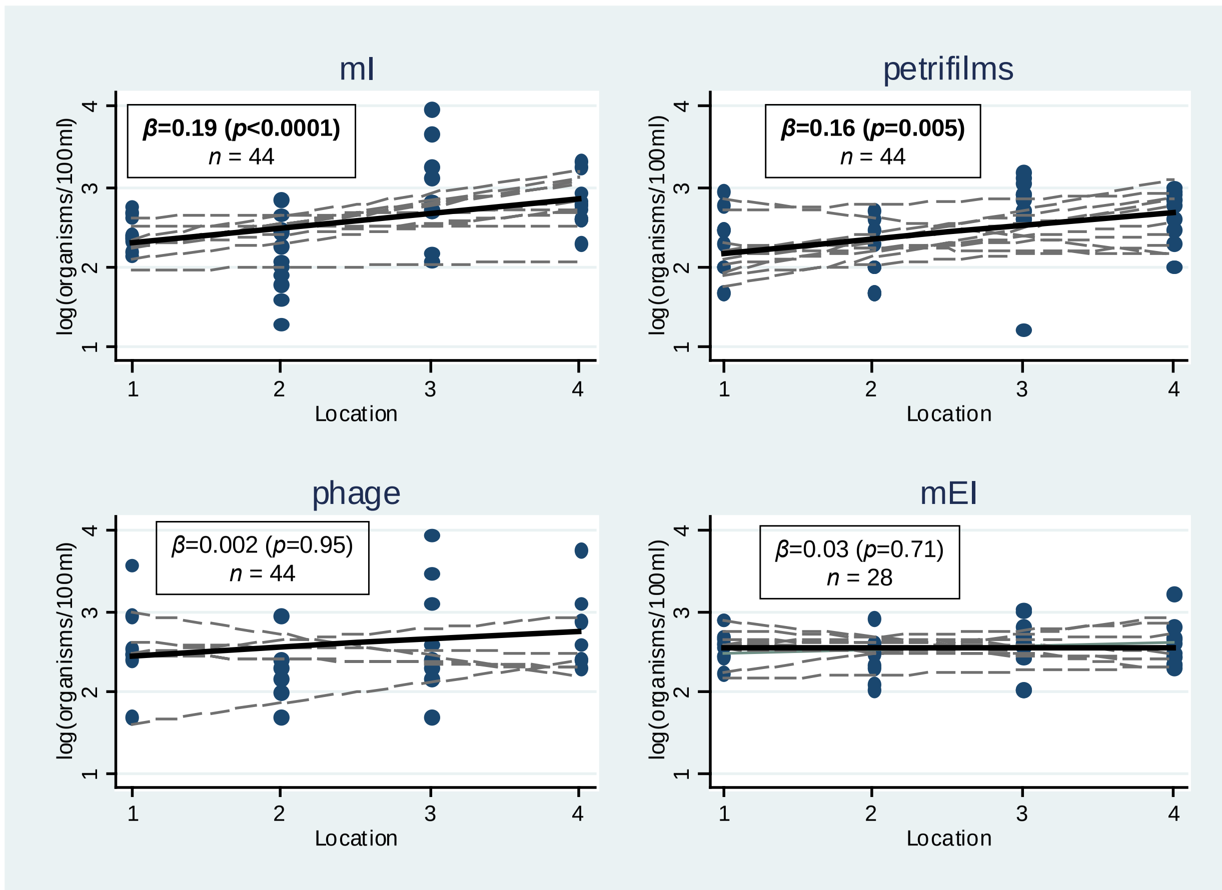
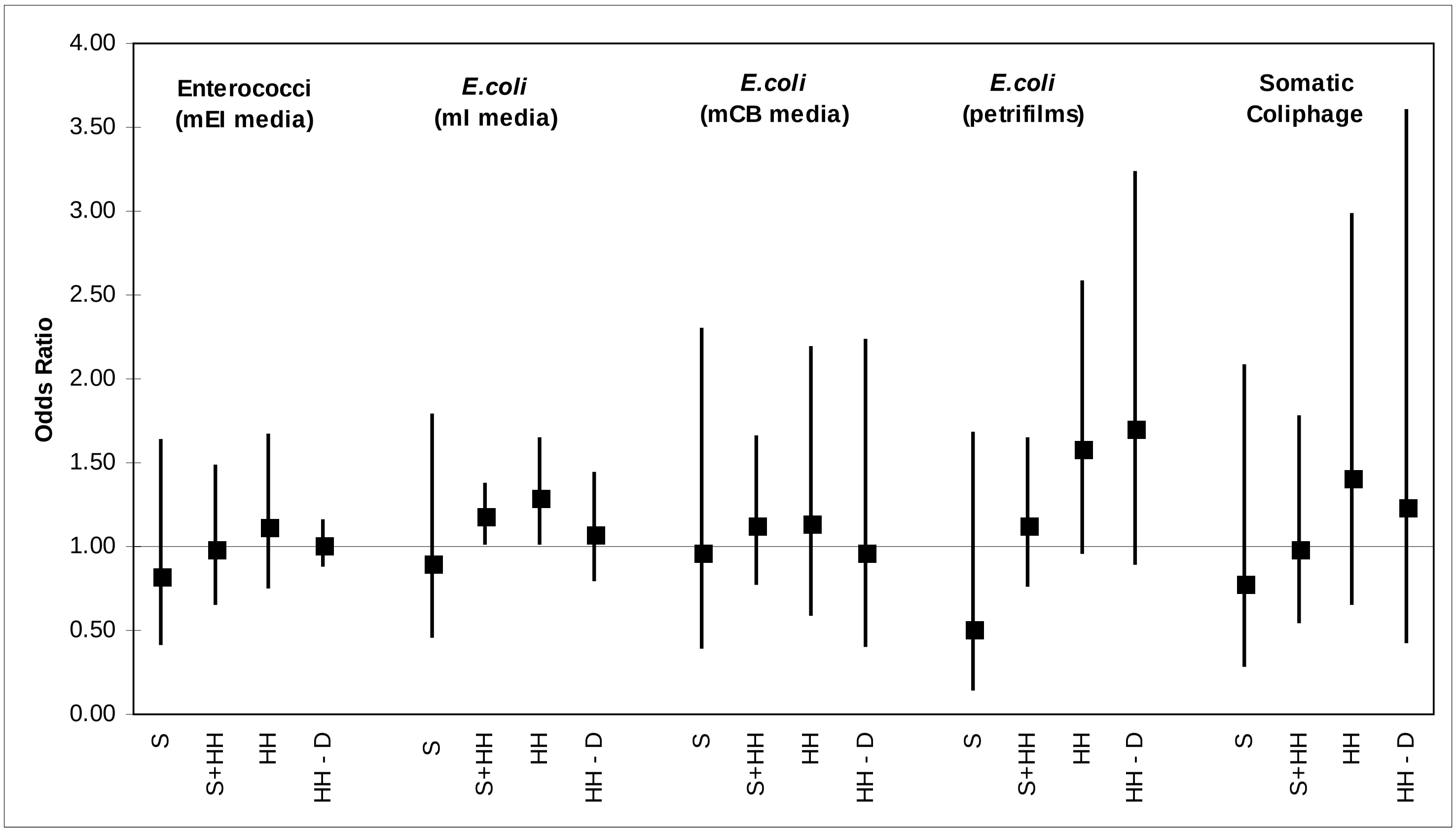


Figure 1.3: Comparison of indicators' ability to predict disease outcome. Odds ratios and 95% confidence intervals are shown, based on values calculated with logistic regression of household case status versus the log10 of the median number of indicator organisms/100ml for samples collected in a particular household. Results are shown for regressions using four different categories of water associated with a particular household: S refers to samples collected from the original water source; S + HH refers to samples collected at the source and from containers within the household; HH refers to samples collected only from containers within the house; and HH-D refers to samples collected from within the house that specifically were identified as drinking water. Sample sizes are shown below each line.



“And I wish to give an account of the other kinds of waters, namely, of such as are wholesome and such as are unwholesome, and what bad and what good effects may be derived from water; for water contributes much towards health. … The best are those which flow from elevated grounds, and hills of earth; these are sweet, clear, and can bear a little wine; they are hot in summer and cold in winter, for such necessarily must be the waters from deep wells. … Rain waters, then, are the lightest, the sweetest, the thinnest, and the clearest; for originally the sun raises and attracts the thinnest and lightest part of the water, as is obvious from the nature of salts; for the saltish part is left behind owing to its thickness and weight, and forms salts; but the sun attracts the thinnest part, owing to its lightness, and he abstracts this not only from the lakes, but also from the sea, and from all things which contain humidity, and there is humidity in everything; and from man himself the sun draws off the thinnest and lightest part of the juices.”

~ Hippocrates, 400 B.C.E.

On Air, Water and Places

CHAPTER TWO

DETERMINANTS OF HOUSEHOLD DRINKING WATER QUALITY: A CONTROLLED STUDY IN NORTHERN COASTAL ECUADOR

Introduction

Worldwide, 1.1 billion people still did not have access to safe drinking water in 2002 (United Nations 2005), and every day more than 6,500 children die from diarrheal illness (Kosek et al. 2003). If we are to move towards the Millennium Development Goals of halving the number of people without access to safe water by 2015 (United Nations 2005), a variety of different interventions may be necessary, since water quality and water usage patterns depend on environmental, social, economic, and cultural characteristics of a given area. To design the most appropriate interventions to improve water quality and supply for a variety of contexts, information is needed to assess the characteristics of water contamination under differing environmental conditions.

Many researchers have observed that storing water in the household leads to a deterioration of water quality, due to recontamination in the home. Even if families have a source of clean drinking water, water may become contaminated in the home due to poor hygiene and water-handling practices, transfer between collection and storage containers, and storage in unclean containers. Factors known to affect recontamination of water in the home include size of the storage vessel mouth (e.g., Mintz et al. 1995), transfer of water between containers from collection to storage (e.g.,Lindskog & Lindskog 1988), hand-water contact and dipping of utensils (e.g.,Hammad & Dirar 1982; Trevett et al. 2005), and bacterial regrowth within the storage container (e.g., Momba & Kaleni 2002). Studies have also shown that organisms can prosper in biofilms in containers (Jagals et al. 2003).

Wright et al. (2004) carried out a systematic meta-analysis of 57 studies measuring bacteria counts for source water and stored water in the home to assess how contamination varied between different settings. They conclude that bacteriological quality of drinking water significantly declines after collection in many settings, although they note considerable variability between settings in the extent of this post-collection contamination. However, few of these studies follow water in a household over time, and even fewer use proper controls to assess how water quality changes when stored outside of the environs of the household. In general, mean source water quality is simply compared to mean household stored water quality, or household samples are matched to specific sources, but sample collection is not matched in time. In studies where controls have been used (e.g.,Roberts et al. 2001), the control containers have not been paired with samples stored in the household environs. Another potential methodological problem with previous studies is that most rely on self-reported data on water source, which might introduce bias because people are likely to misrepresent where they get their water (Wright et al. 2004).

We describe the results of a controlled experiment to compare microbiological contamination of drinking water between the source and point-of-use in northern coastal Ecuador. In this study, we sampled water from the same source at the same time as members of the study households filled their containers. We also followed this water over time, comparing microbiological contamination of water in containers filled at the time of the visit that were stored in the household with containers filled with the same source water that were stored in controlled conditions. In addition to assessing differences in water quality between source and point-of-use samples, we also explore the influence of a series of potential covariates on determining water quality in these samples.

Methods

Study Area

This study was carried out in northern coastal Ecuador, in the province of Esmeraldas, Canton Eloy Alfaro, and in five villages situated along the Santiago, Cayapas, and Onzole rivers (Figure 1.1). Two of these villages rely on simple piped water systems, two rely on surface water from fast-flowing rivers, and one relies on surface water from a small stream. In addition to their primary water source (tap or surface water), some villagers also used simple wells or collected rainwater as source waters for drinking.

Little sanitation infrastructure exists in these communities. While some people utilize private or community latrines, according to surveys we have carried out in the region, 60% of people dispose of human waste out in the open, by digging a hole, or directly into the river. This same river serves as the primary water source for 68% of households, and 60% of households reported drinking their water without treating it. High rates of diarrheal disease have been observed in this study area (Eisenberg et al. 2006; Vieira et al. 2007).

Villages and household water samples were collected in conjunction with a case-control study of diarrhea incidence in each village. Over the course of each 15-day visit, all cases of diarrhea (defined as three loose stools in a 24-hour period) were identified through daily visits to the households in the community. For every household with a case of diarrhea, a control household without a case of diarrhea was randomly selected. Household drinking water samples and source water samples were collected for case and control households. Sample collection and processing took place between March, 2005 – March, 2006. The Institutional Review Boards of U.C. Berkeley and Universidad San Francisco de Quito approved all contact with human subjects.

Sample Processing

Samples were collected in a manner consistent with how users collect and serve their drinking water. Container openings and taps were not sterilized prior to sampling. All samples were collected in Whirlpak bags (Nasco) and kept on ice until processed, within 24 hours. Culturing was carried out in a field laboratory set up in a house or health dispensary in the villages in which samples were collected, using a modular field hood made from plexiglass and metal to avoid contamination. Plates were incubated using an egg incubator and generator where electricity was not available. Agar plates were poured at a microbiology lab in Quito, wrapped individually in Parafilm and packaged in plastic bags, then transferred to the field site in coolers.

Water quality was evaluated for microbiological contamination using membrane filtration. A sample of water was passed through a 47-mm diameter 0.45 μm cellulose filter (Millipore) and then rinsed with a phosphate-buffered saline solution (pH 7.4 ± 0.2) before transferring to a growth medium plate. The stainless steel membrane filtration apparatus (Millipore) was dipped in alcohol, flame-sterilized, and cooled between each sample. *E.coli* was detected using mI agar (BD Difco; prepared according to EPA Method 1604) (U.S. EPA 2002b) and enterococci was detected with mEI agar (BD Difco; prepared according to EPA Method 1600) (U.S. EPA 2002a). Plates were counted after 24 hours of incubation at approximately 30±2°C (*E.coli*) and 41±2°C (enterococci).

For a description of data management, see Chapter 1.

<u>Analysis</u>

Samples were collected from source waters (surface water, well, tap, rain) and point-of-use storage containers within the households. Figure 2.1 shows a schematic of the sampling schemes used to assemble the three datasets used in the analysis. All analysis was carried out using Stata 9.0 (StataCorp LP, College Station, TX).

Dataset 1: Water stored in the household. In all case and control households, between one and three water containers were sampled upon the first visit to the household, depending on the amount of stored water available at the time of visit. These data were merged with all other samples taken from households (under the sampling schemes used to create Datasets 2 and 3, see description below) to get a complete dataset of all samples taken from containers stored in the household.

Using this dataset, various covariates were analyzed using linear regression to estimate their effects on log indicator concentrations (CFU/100 mL). Covariates considered included community- and household-level variables: *community size* (number of houses per village), *community sanitation* (percent of individuals in the village who stated that they used improved sanitation, i.e., latrines or septic tanks), and *crowding* (number of people living in the household at the time of the visit). We also included several container-level variables: *water source* (rain, well, piped, river, or small stream), *treatment* (none, boiled, chlorinated, or left to settle), *container type* (small- versus large-mouthed), *covered* (whether or not the container was covered or capped), and *storage time* (number of days since the container was filled). The standard errors of the regression coefficients were adjusted for intra-group correlation amongst samples collected within the same village visit, which represents the highest level of possible correlation amongst the samples. Village visit was used rather than village to account for potential differences in visits for the villages that were visited more than once.

Dataset 2: Source water followed over time in matched household and control storage containers. In 59 households, water was collected from the same source at the same time as the household member filled their water container, and a sample of this source water was also stored in a control container similar to those most commonly used in the households (a 10 gallon plastic jerry can). The household container was marked and re-sampled daily for one to five days, until the family had finished using the water collected on the day of the visit to the source. Control containers, which were kept covered and in controlled conditions in the field laboratory, were re-sampled in parallel. This laboratory had similar environmental conditions to the households (i.e., no air conditioning, open ceilings, etc.). Control containers were sterilized with boiling water between samples. This study design allowed for a controlled assessment of die-off and recontamination events, comparing source waters to both control and household samples.

Geometric mean values were compared for samples taken directly from the source, samples taken from household containers, and samples taken from control containers. The mean of paired log differences was also calculated for samples from sources versus household containers, sources versus control containers, and household containers versus control containers. The significance of these paired differences was tested with one-sided matched paired *t*-tests. For the calculations of differences between matched pairs, the final value for the household and control containers was used, in order to create only one value pair per household. For example, if household and control containers were sampled for three consecutive days, only the value on the third day was used. Graphs of indicator concentrations in both types of container were visually inspected to ensure that in the majority of cases the trends in behavior of these two container treatments after collection were correlated over time, i.e., that the differences in the last samples from the control and household containers was representative of the differences between these two types of samples for each of the other days.

Dataset 3: Household water followed over time in household storage containers. Thirty-six of the household containers sampled upon the initial visit to the household were resampled daily for up to five days. These data were merged with the data from the household containers of the paired samples (Dataset 2), beginning with the first day of sample collection in the household, to assess recontamination in households over time. Note that in this dataset, all sampling begins in the household, so reductions between the source and household due to settling or die-off would have already occurred before the first day of sampling. Graphs and linear regressions were carried out to assess contamination of containers over time. These results were stratified by whether or not they experienced recontamination between the first and last sampling (difference greater versus less than zero), as well as other covariates.

Results

Analysis Using Dataset 1:

The overall geometric means of samples stratified by source and various container-level variables can be seen in Table 2.1. In the regression analyses (Table 2.2), the community and household factors assessed were less informative in explaining the variability seen in the water quality outcomes than the variables describing characteristics of the container. Water treatment was the most important explanatory variable: boiling and chlorine (when reported by members of the household as a form of treatment) were significantly associated with decreased counts of indicator organisms compared to no treatment, although the effect of settling was not significant. This significant effect of treatment was seen despite the small percentage of water samples that had received treatment of any kind (only 22% of samples). Additionally, container type and whether or not a container was covered at the time of sampling, both showed an effect for enterococci. Water source showed an effect for *E.coli*.

Storage time showed a significant, but non-meaningful effect (β=0.00), for both indicators.

Analysis Using Dataset 2:

Source water was found to be significantly more contaminated than water in the household. Source waters had a geometric mean of >200 CFU/100 mL for both enterococci and *E.coli*, whereas samples from containers stored in the household had a geometric mean of approximately 100 CFU/100 mL for both indicator organisms. Samples from control containers had even lower counts, on the order of 60 CFU/100 mL (Table 2.3).

The analysis of differences between the paired samples (Table 2.4) provides further insight. Difference between source and control samples can be considered to reflect reduction of indicator organisms, due to settling or die-off. A more than 0.5 log reduction on average was observed for both enterococci and *E.coli*. Differences between source and household samples reflect reductions of indicator organisms in the home. On average, a 0.4 log reduction was observed for both indicator organisms. This is also reflected in the differences between samples from household and control containers; samples from containers stored in households were on average 0.2 logs more contaminated than their matched controls, which reflects contamination within the household environment. Overall source waters had significantly higher concentrations of indicator organisms than household or containers, and samples from household containers had significantly higher concentrations than samples from control containers. It should be noted that there was significant heterogeneity in the occurrence of recontamination among households; only 65.5% (enterococci) and 51.7% (*E.coli*) of household-container pairs were more contaminated in the household than in the control container.

Analysis of Dataset 3:

Evidence of recontamination in the home was also seen in the household containers followed in time, although again, this did not occur in all cases. Increasing contamination between the first and last day of sampling was observed in 46.4% (enterococci) and 32.8% (*E.coli*) of these containers. Most of the containers resampled over time were only sampled one day apart, because the family had finished the water in the container or due to logistical difficulties of re-visiting the house. Between the first and second day of sampling, 42.8% (enterococci) and 45.5% (*E.coli*) of containers increased in level of contamination. The overall trend for all containers, as well as for just containers that experienced recontamination, can be seen in Figure 2.2. To explain some of the heterogeneity seen in recontamination of containers, regressions of log indicator counts (CFU/100 mL) against days of storage in the household were carried out. Of all the covariates assessed, the slope coefficient of the regression of large-mouthed containers most closely matches the slope coefficient of the regression for recontaminated samples only (Table 2.5), suggesting that size of container opening might be an important factor in recontamination of containers.

Discussion

To our knowledge this is the first study to evaluate contamination between source and point-of-use drinking water quality by sampling the same source water as collected by households in real time, and to follow its fate over the course of several days of storage within a household. It is also the first study to use paired controls to assess changes in contamination levels over time in the home. On average, we observed greater than 0.5 log reductions of indicator organism concentrations from the source of drinking water to its point of use, followed by a 0.2 log increase in approximately half of households sampled during storage and use. Given the trend reported in the literature showing a tendency for household water samples to be more contaminated than the source waters from which they draw (Wright et al. 2004), the recontamination we observed in the home was expected. The higher overall levels of contamination observed at the source, on the other hand, contradicts this trend.

This result is not unprecedented, however, and likely reflects the poor quality of source water in our study communities. Studies that have compared recontamination under variable initial conditions have shown that the quality of source water affects the extent of recontamination observed in the home. For example, in a study in Venda, South Africa, Verweij et al. (1991) observed a 10-

to 15-fold increase in fecal coliform counts between source and storage in water collected from boreholes; but in water samples from unprotected springs, which exhibited high initial coliform counts (approximately 300 CFU/100 mL), they observed a two-fold decline in counts over four hours of storage. Musa et al. (1999) found contrasting results for different types of communities in a study in northern Sudan. In rural villages and nomadic areas, where people depended on poor quality (>100 CFU/100 mL) source waters, fecal coliform counts were lower in storage containers than at the source, whereas in three peri-urban communities, where municipally source water was of reasonable quality (<10 CFU/100 mL), fecal coliform counts were significantly higher in home storage containers, suggesting contamination in the household. Only the result showing recontamination in the household was reported in the review by Wright et al. (2004).

In the review by Wright et al. (2004), water quality deterioration from the source to the point-of-use was found to be greater for studies of uncontaminated water sources, but most of the studies in this review had high initial water quality (i.e., low counts of indicator organisms) at the source. Recontamination was also less pronounced in homes with poorer quality source water. Within any given population, there often appears to be a subset of households in which the quality of stored water improves compared to the quality of source water (Wright, pers. comm., 3/21/07). For example, while VanDerslice and Briscoe (1993) observed a net increase in fecal coliform counts in over half of source-household sample pairs, they observed a net decrease in counts in 16% of households and no net change in 32%. Our methods of sampling water concurrently with household members and following these particular containers over time eliminates the possibility for bias in reporting of levels of contamination at the source; and this elimination of bias might partially explain why our results differ from many others reported in the literature.

Post-collection reductions in microbial contamination have also been observed in laboratory studies. Tomkins et al. (1978) observed a marked fall in coliform counts following overnight storage in earthenware containers in a study in northern Nigeria where rural villagers relied on water from both protected and unprotected wells. Mazengia et al. (2002) reported significant reductions in bacterial loads in water urns stored in a laboratory setting compared to the source wells from which they were drawn. These reductions corresponded with declines in turbidity. In one study in rural South Africa, type of container was shown to affect rates of removal of organisms during storage: indicator organisms persisted in borehole water in polyethylene containers for longer than they did in galvanized steel containers (Momba & Notshe 2003). Studies of organisms resident in water containers would further elucidate these factors.

The observed reductions in bacterial loads could be due to settling of organisms to the bottom of storage containers or die-off of these organisms caused by predation by other microorganisms, lack of nutrients, or other factors contributing to inhospitable conditions in the container. This is an important distinction, because organisms that settle out could become resuspended and consumed, thus maintaining their ability to cause infection, whereas die-off of organisms would imply loss of infectivity. In a study in Malawi, shaking of containers led to a three-fold increase in detection of indicator organisms in unimproved buckets, suggesting that bacteria that had settled to the bottom of the storage container were still viable upon resuspension (Roberts et al. 2001). Future studies should focus on distinguishing between removal and inactivation in storage containers, both in the water column and sediment, as well as resuspension. Furthermore, the behavior of indicator organisms should be compared to that of actual pathogens. A key factor influencing removal is likely to be the particle association of indicators and pathogens, and the settling velocity of those particles. Given the high turbidity in source waters in this study, settling within containers in the home likely explains at least part of the reduction in bacterial loads we observed.

Following the initial reductions between source and point-of-use, we observed an increase in contamination in some households, as can be observed in Figure 2.2 and as evidenced by differences between household and control samples (Tables 2.3- 2.4). However, increased contamination within the home environment was only observed in about half of all containers assessed. It should be noted that the use of control containers allowed for an assessment of the total change in contamination occurring within the household, factoring in the amount of organisms lost to settling and die-off as observed in the control containers. Since most studies do not take such reductions into account, they might actually underestimate the extent of recontamination in the household environment, or missing it altogether.

The results of the regression analyses (Table 2.2) suggest that water treatment by boiling and chlorination was associated with reduced contamination. Larger mouths and covered containers were associated with decreased water quality when measured with enterococci. All water sources were significantly more contaminated than rain water when assessed with *E.coli*, but only river water was significantly more contaminated when enterococci were used.

The slope of regressions of indicator concentrations for large-mouthed containers most closely matched that of containers exhibiting recontamination overall, suggesting that mouth size may be a large factor in determining whether a container becomes recontaminated in the home. These results are consistent with previous studies showing that factors related to the container, such as mouth size and providing a cover, are key factors in determining quality of stored water (Mintz et al. 1995).

With the growing recognition of the issue of household recontamination, many authors have recommended focusing interventions on improving water quality at the point of use rather than improving water supply or water quality (Clasen &

Bastable 2003; Mintz et al. 2001; Reiff et al. 1996). A wide range of interventions aimed at improving drinking water in the home have been proposed, including improving vessels (e.g., Hammad & Dirar 1982; Mintz et al. 1995; Roberts et al. 2001), decontaminating drinking water using chlorine (Mintz et al. 2001; Mintz et al. 1995; Quick et al. 1999; Reiff et al. 1996), sunlight (Conroy et al. 1999), ceramic filtration (Clasen et al. 2006), and coagulation plus chlorination (Rangel et al. 2003). Taking into account initial conditions in source water quality in a given region is important for determining the most appropriate strategy for in-home decontamination. For example, in this region coagulation followed by chlorination might be the most effective at reducing pathogen concentrations, because the high turbidity waters might reduce the efficacy of direct chlorination and solar disinfection, and rapidly clog ceramic filters. Introducing containers with smaller mouths would also decrease the potential for recontamination in the home.

It is also important to note that, while water treatment at the point-of-use has been shown to be an effective strategy in intervention trials, this effect is not universally observed (Clasen et al. 2007; Fewtrell et al. 2005). Taking into account initial conditions may explain some of the heterogeneity seen in the effectiveness of point-of-use interventions. The results of this study suggest that surface source waters are more contaminated than water in the home, and that in-home contamination may be a smaller factor compared to initial source water quality in determining the quality of drinking water in the home. While we observe recontamination in the home, we see more than twice as much reduction between the source and point-of-use as we see recontamination at the point-of-use.

In our study villages, we often observe children and adults drinking water directly from the stream. In areas such as the one we studied, with poor sanitation and poor source water quality, where villagers drink straight from the stream, improving water quality in the home may not be sufficient to break the cycle of transmission of waterborne pathogens. Given the nonlinear nature of transmission of waterborne diseases and the complex set of interdependent pathways by which enteric pathogens are transmitted (Eisenberg et al. 2007), focusing solely on household interventions without reducing the sources of contamination in the community may not be as effective as implementing integrated control strategies that include sanitation and improvement of water quality at the source. Furthermore, intervening only at the household level would ignore the health risks of bathing in contaminated waters.

While recent reviews have found little or no evidence that the efficacy of water quality interventions was related to levels of sanitation (Clasen et al. 2007; Fewtrell et al. 2005), others have suggested that the efficacy of household water quality interventions depends on the level of sanitation within the target community (Esrey & Habicht 1986; Gundry et al. 2004; VanDerslice & Briscoe 1995).

In an ideal world, interventions aimed at improving drinking water would be generalizable to all situations. However, the results of this study suggest that the optimal intervention strategy may depend on initial source water quality. In areas where initial source water quality is poor, in-home water treatment and safe water storage may need to be augmented by efforts to improve sanitation and/or source water quality.

Literature Cited:

Clasen, T., W. P. Schmidt, T. Rabie, I. Roberts, and S. Cairncross. 2007. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. British Medical Journal 334:782-791.

Clasen, T. F., and A. Bastable. 2003. Faecal contamination of drinking water during collection and household storage: the need to extend protection to the point of use. Journal of Water and Health 1:109-115.

Clasen, T. F., S. Nadakatti, and S. Menon. 2006. Microbiological performance of a water treatment unit designed for household use in developing countries. Tropical Medicine and International Health 11:1399-1405.

Conroy, R. M., M. E. Meegan, T. Joyce, K. McGuigan, and J. Barnes. 1999. Solar disinfection of water reduces diarrhoeal disease: an update. Archives of disease in childhood 81:337-338.

Eisenberg, J. N. S., W. Cevallos, K. Ponce, K. Levy, S. J. Bates, J. C. Scott, A. Hubbard, N. Vieira, P. Endara, M. Espinel, G. Trueba, L. W. Riley, and J. Trostle. 2006. Environmental change and infectious disease: How new roads affect the transmission of diarrheal pathogens in rural Ecuador. Proceedings of the National Academies of Science 103:19460-19465.

Eisenberg, J. N. S., J. C. Scott, and T. C. Porco. 2007. Integrating public health control strategies: Balancing water sanitation, and hygiene interventions to reduce diarrheal disease burden. American Journal of Public Health 97:846-852.

Esrey, S. A., and J.-P. Habicht. 1986. Epidemiologic Evidence for Health Benefits from Improved Water and Sanitation in Developing Countries. Epidemiological Reviews 8:117-128.

Fewtrell, L., R. B. Kaufmann, D. Kay, W. Enanoria, L. Haller, and J. M. Colford, Jr. 2005. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. Lancet Infectious Diseases 5:42-52.

Gundry, S., J. Wright, and R. Conroy. 2004. A systematic review of the health outcomes related to household water quality in developing countries. Journal of Water and Health 2:1-13.

Hammad, Z. H., and H. A. Dirar. 1982. Microbiological Examination of Sebeel Water. Applied and Environmental Microbiology 43:1238-1243.

Jagals, P., C. Jagals, and T. C. Bokako. 2003. The effect of container-biofilm on the microbiological quality of water used from plastic household containers. Journal of Water and Health 1:101-108.

Kosek, M., C. Bern, and R. L. Guerrant. 2003. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bulletin of the World Health Organization 81:197-204.

Lindskog, R. U. M., and P. A. Lindskog. 1988. Bacteriological contamination of water in rural areas: an intervention study from Malawi. Journal of Tropical Medicine and Hygiene 91:1-7.

Mazengia, E., M. T. Chidavaenzi, M. Bradley, M. Jere, C. Nhandara, D. Chigunduru, and E. C. Murahwa. 2002. Effective and culturally acceptable water storage in Zimbabwe: Maintaining the quality of water abstracted from upgraded family wells. Journal of Environmental Health 65:15-18.

Mintz, E. D., J. Bartram, P. Lochery, and M. Wegelin. 2001. Not Just a Drop in the Bucket: Expanding access to point-of-use water treatment systems. American Journal of Public Health 91:1565-1570.

Mintz, E. D., F. M. Reiff, and R. V. Tauxe. 1995. Safe Water Treatment and Storage in the Home. Journal of the American Medical Association 273:948-953.

Momba, M. N. B., and P. Kaleni. 2002. Regrowth and survival of indicator microorganisms on the surfaces of household containers used for the storage of drinking water in rural communities of South Africa. Water Research 36:3023-3028.

Momba, M. N. B., and T. L. Notshe. 2003. 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 52:67-77.

Musa, H. A., P. Shears, S. Kafi, and S. K. Elsabag. 1999. Water quality and public health in northern Sudan: a study of rural and per-urban communities. Journal of Applied Microbiology 87:676-682.

Quick, R. E., L. V. Venczel, E. D. Mintz, L. Soleto, J. Aparicio, M. Gironaz, L. Hutwagner, K. Greene, C. Bopp, K. Maloney, D. Chavez, M. Sobsey, and R. V. Tauxe. 1999. Diarrhoea prevention in Bolivia through point-of-use water treatment and safe storage: a promising new strategy. Epidemiology and Infection 122:83-90.

Rangel, J. M., B. Lopez, M. A. Mejia, C. Mendoza, and S. Luby. 2003. A novel technology to improve drinking water quality: a microbiological evaluation of in-home flocculation and chlorination in rural Guatemala. Journal of Water and Health 1:15-22.

Reiff, F. M., M. Roses, L. Venczel, R. Quick, and V. M. Witt. 1996. Low-cost safe water for the world: a practical interim solution. Journal of Public Health Policy 17:389-408.

Roberts, L., Y. Chartier, O. Chartier, G. Malenga, M. Toole, and H. Rodka. 2001. Keeping clean water clean in a Malawi refugee camp: a randomized intervention trial. Bulletin of the World Health Organization 79:280-287.

Tomkins, A. M., B. S. Drasar, A. K. Bradley, and W. A. Williamson. 1978. Water supply and nutritional status in rural Northern Nigeria. Transactions of the Royal Society of Tropical Medicine and Hygiene 72:239-243.

Trevett, A. F., R. C. Carter, and S. F. Tyrrel. 2005. Mechanisms leading to post-supply water quality deterioration in rural Honduran communities. International Journal of Hygiene and Environmental Health 208:153-161.

U.S. EPA. 2002a. Method 1600: Enterococci in Water by Membrane Filtration Using membrane-Enterococcus Indoxyl-β-D-Glucoside Agar (mEI). U.S. Environmental Protection Agency Office of Water, Washington, D.C.

U.S. EPA. 2002b. Method 1604: Total Coliforms and Escherichia coli in Water by Membrane Filtration Using a Simultaneous Detection Technique (MI

Medium). U.S. Environmental Protection Agency Office of Water, Washington, D.C.

United Nations. 2005. The Millennium Development Goals Report 2005. United Nations Department of Public Information, New York.

VanDerslice, J., and J. Briscoe. 1993. 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 29:1983-1995.

VanDerslice, J., and J. Briscoe. 1995. Environmental Interventions in Developing Countries: Interactions and their Implications. American Journal of Epidemiology 141:135-144.

Verweij, P. E., M. van Egmond, D. J. Bac, J. G. van der Schroeff, and R. P. Mouton. 1991. Hygiene, skin infections nd types of water supply in Venda, South Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene 85:681-684.

Vieira, N., S. J. Bates, O. D. Solberg, K. Ponce, R. Howsmon, W. Cevallos, G. Trueba, L. Riley, and J. N. S. Eisenberg. 2007. High prevalence of Enteroinvasive Escherichia coli isolated in a remote region of northern coastal Ecuador. American Journal of Tropical Medicine and Hygiene 76:528-533.

Wright, J., S. Gundry, and R. Conroy. 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Tropical Medicine and International Health 9:106-117.

Table 2.1: Geometric means of indicator organism concentrations (CFU/100 mL) in household containers, stratified by characteristics. Reported *p*-values test equality of means using t-tests (binary variables) or anovas (variables with multiple categories).

	N	Enterococci	*E.coli*
Container Type			
Small-mouthed	372	74	41
Large-mouthed	260	110	65
		***p*=0.001**	*p*=0.68
Container Covered?			
Yes	415	62	45
No	228	161	69
		***p*<0.0001**	***p*=0.009**
Water Treatment			
None	500	122	81
Boiling	48	14	11
Chlorine	42	26	19
Let settle	6	48	120
		***p*=0.05**	*p*=0.25
Water Source			
Rain	104	74	9
Well	25	64	45
Tap	259	86	51
River	122	242	272
Estero	117	58	91
		***p*=0.003**	***p*<0.0001**

Table 2.2: Effects of covariates on quality of water stored in household containers, as measured using log values of *E.coli* and enterococci concentrations (CFU/100 mL) as the outcome variable. Unadjusted values report the results of univariate analyses; Adjusted values report the results of multivariate analysis, including all covariates in the model.

Variable	Level	Description	n		Enterococci vs. Variable unadjusted	Enterococci vs. Variable adjusted	*E.col* vs. Variable unadjusted	*E.col* vs. Variable adjusted
Crowding	household	# people in HH	155	β	0.06	0.00	0.03	-0.03
				s.e.	0.01	0.01	0.02	0.01
				p-value	**0.002**	0.970	0.253	0.155
Community Size	commty	# houses in village	5	β	0.00	0.00	0.00	0.00
				s.e.	0.00	0.00	0.00	0.00
				p-value	0.58	0.58	0.97	0.28
Sanitation	commty	sanitation index	5	β	0.01	0.01	0.00	0.01
				s.e.	0.01	0.01	0.00	0.00
				p-value	0.288	**0.029**	0.827	0.097
Covered	container	1 = covered	415					
		2 = uncovered	228	β	0.41	0.29	0.16	0.08
				s.e.	0.07	0.07	0.11	0.08
				p-value	**0.000**	**0.008**	0.191	0.358
Water Source	container	1=rain	104					
		2=well	25	β	-0.06	-0.37	0.71	1.19
				s.e.	0.28	0.46	0.28	0.12
				p-value	0.831	0.448	**0.040**	**0.001**
		3=piped	259	β	0.07	0.14	0.75	0.90
				s.e.	0.12	0.17	0.16	0.12
				p-value	0.590	0.430	**0.002**	**0.001**
		4=river	122	β	0.51	0.39	1.48	0.95
				s.e.	0.19	0.16	0.30	0.07
				p-value	0.497	**0.052**	**0.002**	**<0.0001**
		5=small stream	117	β	-0.10	0.09	1.01	1.18
				s.e.	0.14	0.33	0.14	0.31
				p-value	0.497	0.789	**<0.0001**	**0.018**
Treatment	container	0=none	500					
		2=boiled	48	β	-0.94	-0.97	-0.88	-1.56
				s.e.	0.09	0.14	0.15	0.16
				p-value	**<0.0001**	**<0.0001**	**0.001**	**0.001**
		3=chlorination	42	β	-0.67	-0.92	-0.62	-1.10
				s.e.	0.26	0.04	0.17	0.18
				p-value	**0.028**	**0.000**	**0.007**	**0.004**
		4=left to settle	6	β	-0.40	-0.19	0.18	0.33
				s.e.	0.56	0.77	0.32	0.44
				p-value	0.489	0.815	0.604	0.498
Container	container	1=small mouth	372					
		2=large mouth	260	β	0.17	0.37	-0.21	0.23
				s.e.	0.13	0.11	0.12	0.12
				p-value	0.230	**0.013**	0.127	0.124
Storage Time	container	# days since filled	602	β	0.00	0.00	0.00	0.00
				s.e.	0.00	0.00	0.00	0.00
				p-value	**0.021**	**0.006**	**0.028**	**0.011**

Table 2.3: Overall levels of contamination at the source, and in household and control containers (geometric means of CFUs/100ml). *N*=59.

	Enterococci	***E.coli***
source	227.9	227.1
household	93.8	102.6
control	55.6	63.9

Table 2.4: Mean paired log differences between water samples from a) source and control containers; b) source and household containers; and c) household and control containers. *P*-values report results of one-sided matched paired t-tests comparing log values for sample pairs.

	N	**Enterococci**	***E.coli***
Control Reductions (source – control)	59	0.65 (±0.12) $p<0.0001$	0.57 (±0.10) $p<0.0001$
In-Home Reductions (source – house)	59	0.42 (±0.14) $p=0.002$	0.37 (±0.13) $p=0.004$
Total Household Contamination (house – control)	58	0.24 (±0.09) $p=0.006$	0.18 (±0.11) $p=0.06$

Table 2.5: Slope coefficients (and 95% confidence interval) for regressions of water quality (log indicator concentrations (CFU/100 mL)) versus days of storage for a) all containers; b) all containers exhibiting recontamination; c) containers with treated water; d) large-mouthed containers; and e) uncovered containers

	a) All containers (n=162)	**b) Recontaminated** (n=74)	**c) Treated** (n=37)	**d) Large Mouth** (n=42)	**e) Uncovered** (n =119)
Enterococci	0.09 -0.07 - 0.25	0.43 0.21 - 0.65	0.05 -0.4 - 0.5	0.39 0.13 - 0.64	0.12 -0.06 - 0.31
E.coli	-0.03 -0.19 - 0.13	0.32 0.07 - 0.56	-0.08 -0.54 - 0.38	0.15 -0.11 - 0.42	0.03 -0.15 - 0.21

Figure 2.1: Overview of sampling schemes and datasets used in the analysis. POU refers to point-of-use, or household samples. Labels a, b, c refer to household storage containers already filled with water at the time of the initial visit to the household; some of these were followed over time. Label x refers to containers filled at the same time as the control container, for which a source sample was also collected; these were all followed over time. $X_{i,POU}$ were stored in the household, whereas $X_{i,\,lab}$ were stored under controlled conditions. Subscripts refer to the day of sampling in the household.

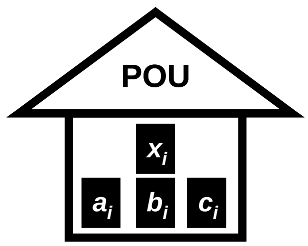
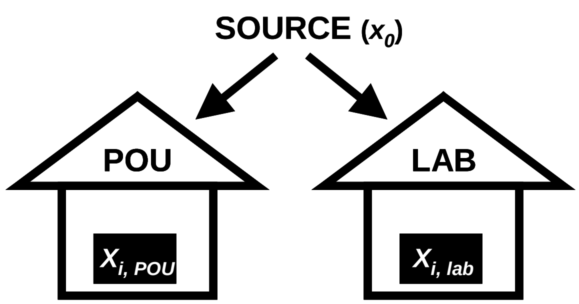
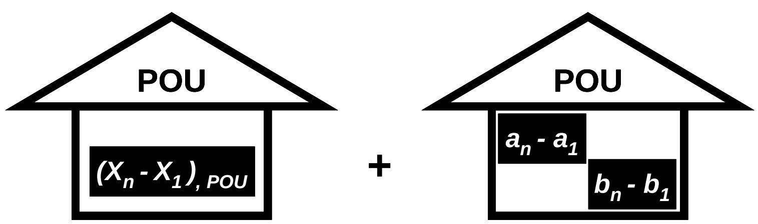
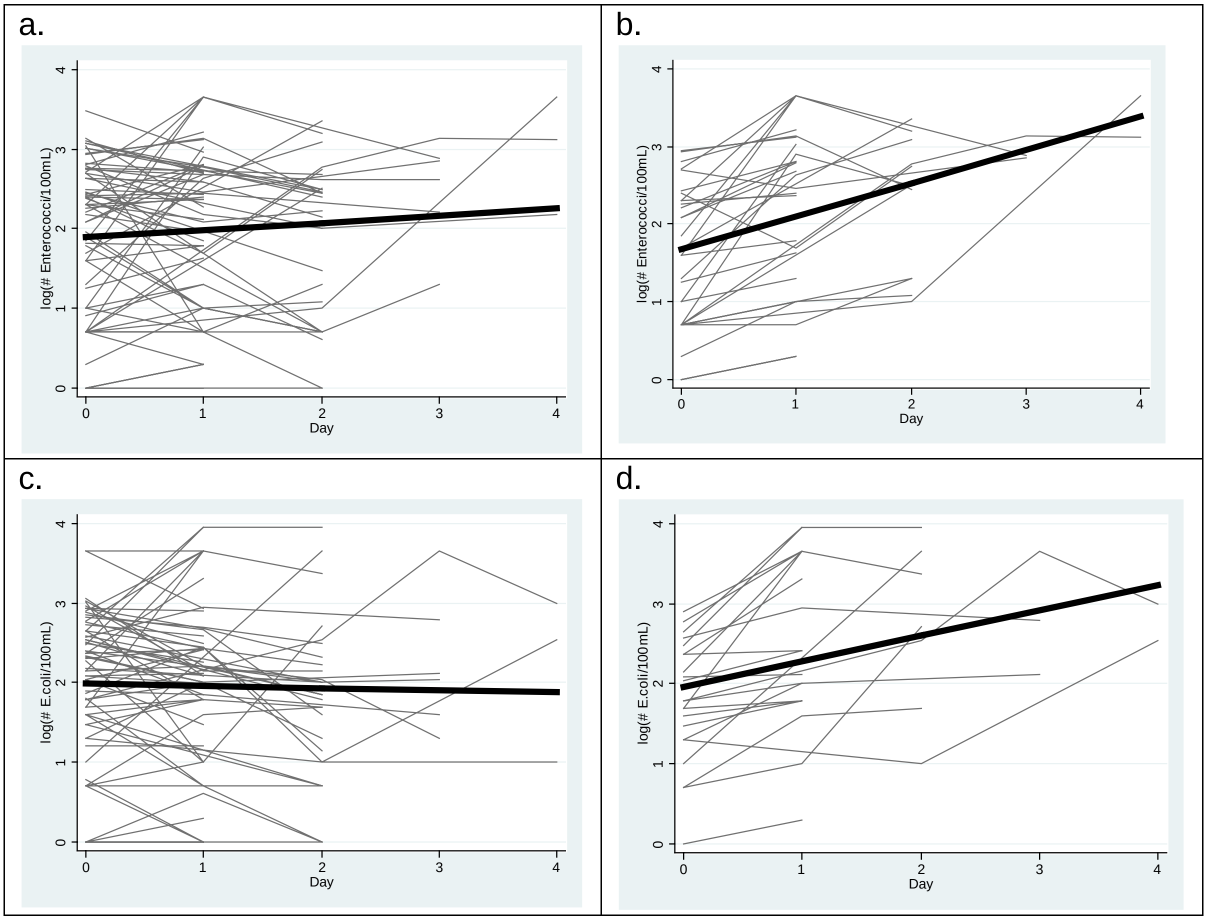
<table>
<tr><th>Data-set</th><th>Schematic of Sampling Scheme</th><th>Description of Sampling Scheme</th></tr>
<tr><td>1</td><td>
</td><td>Water stored in the household.</td></tr>
<tr><td>2</td><td>
</td><td>Source water followed over time in matched household and control storage containers.</td></tr>
<tr><td>3</td><td>
</td><td>Difference between first and last day of sampling in household storage containers.</td></tr>
</table>

Figure 2.2: Contamination over time within households, for all containers (a and c) and just containers exhibiting recontamination, i.e., difference between the first and last sampling greater versus than zero (b and d), using enterococci (a and b) and *E.coli* (c and d) as water quality indicators. Grey lines indicate unique containers and black line shows the regression fit.



"In contrast to the experimenter, the epidemiologist has to deal with biological phenomena in all their natural complexity. [S]he must try to recognize the relationships most common in certain specified ecological situations and to derive from this knowledge the methods of control having the best statistical chance of being useful in each particular situation."

~ René Dubos, 1959

Mirage of Health: Utopias, Progress, and Biological Change

CHAPTER THREE

Seasonal Drivers of Variability in Water Quality in Northern Coastal Ecuador

Introduction

The microbiological safety of recreational and drinking waters is commonly measured using fecal indicator bacteria. However, concentrations of these indicator organisms are spatiotemporally variable, and most sampling is too infrequent to transcend this granularity, making interpretation of data difficult (National Research Council 2005). In developing countries, large variations in concentrations of indicator organisms have been observed at different times of the year, especially for in-house drinking water storage containers (Jensen et al. 2004). In industrialized countries, high degrees of variability in indicator concentrations were observed in near coastal waters on timescales ranging from minutes to decades (Boehm et al. 2002). This variability presents challenges to the use of indicator organism concentrations to guide water quality interventions and regulations. Information is therefore needed on drivers of this variability in order to better understand how to interpret water quality data.

Seasonal changes are known to be a major source of variability in indicator data. Peaks of microbial contamination in waterways have been associated with rainfall in the nearshore marine environment (Boehm et al. 2002), with rainfall and soil moisture content in subtropical rivers (Solo-Gabriele et al. 2000), with rainfall and suspended solid content in temperate streams (George et al. 2004), and with peak rainfall, peak streamflow events, and peak turbidity measurements in rivers in northern latitudes (Dorner et al. 2007). Associations between rainfall and fecal contamination in drinking water sources have been demonstrated in Nicaragua (Sandiford et al. 1989), Uganda (Howard et al. 2003), and Malawi (Lindskog & Lindskog 1988). Peaks of fecal indicator counts in surface water sources were observed at the transition from dry to wet season in a study in Sierra Leone (Wright 1986), and the same effect was observed for a study of several different water sources in Nigeria (Blum et al. 1987). In a study in Gambia, 10- to 100-fold increases in indicator organism concentrations in unprotected wells were observed with the onset of the rains and high levels were sustained throughout the rainy season (Barrell & Rowland 1979).

These peaks in microbial contamination following heavy rainfall are likely due to either the direct flushing of fecal material from the surrounding environment into the stream, mobilization of bacteria resident in soils, or a combination of both of these phenomena. This “runoff effect” is exacerbated in areas with poor sanitation, where a constant input of fecal material is available to be transported in the environment. Areas with poor sanitation often also have inadequate protection of drinking water sources such as surface water and shallow wells, which increases the impact of runoff on contamination of drinking water.

A competing, not mutually exclusive, phenomenon is the “concentration effect,” wherein dry conditions cause an increase in the density of microorganisms due to lack of flushing by rains. The concentration effect has been used to explain cases of dry season epidemics of diarrheal disease (Drasar et al. 1978). Wright et al. (1986) suggest that this effect can more precisely be considered an “accruing-input effect” because the mechanism involved is not in fact concentration of microorganisms but rather the lack of dilution of a continual input of fecal bacteria. These bacteria may originate from poor sanitation, from survival and growth in riverbank soils, or from people bathing and washing in the stream. Rainfall therefore has the potential to increase microbial contamination at both extremes: high rainfall increases flow of bacteria into the stream due to the runoff effect (which can later be flushed out if river flows are high enough), and low rainfall can increase the concentration of bacteria in the stream due to lack of dilution, which for simplicity we refer to here as the “concentration effect.” The relative magnitude of these two effects can vary depending on the season.

Rainfall also affects disease incidence. Seasonality is known to be an important factor in determining the incidence of infectious diseases (Altizer et al. 2006), and diarrheal diseases are no exception. Waterborne disease outbreaks have been associated with peak rainfall events in the United States (Curriero et al. 2001), Canada (Thomas et al. 2006), and with both high and low extremes of rainfall in Fiji (Singh et al. 2001). Increased waterborne disease has been associated with major floods in Bangladesh and elsewhere (Kondo et al. 2002; Qadri et al. 2005; Schwartz et al. 2006). In Guatemala, increased diarrhea was recorded during months with highest rainfall, although this relationship was less apparent in a town with improved piped water (Shiffman et al. 1976).

Monitoring of water quality using indicator organisms is conducted with the ultimate goal of preventing the transmission of waterborne disease. However, to interpret the data provided by these indicators we must understand sources of variability in their measurement in both space and time in each particular context. Understanding this variation has implications for when and how often to collect water samples, for providing recommendations on whether water is safe to drink or bathe in, and for sorting out true variability versus uncertainty in measurement caused by limitations of testing technology.

In this manuscript we address these questions by systematically examining sources of variability in water quality measurements in source and household water samples at varying timescales over the course of one year in a rural Ecuadorian village, where high rates of diarrheal disease have been observed (Eisenberg et al. 2006; Vieira et al. 2007). Specifically, we explore environmental drivers of seasonal, day-to-day, and hour-to-hour variability in water quality measurements of surface waters used as drinking water source, analogous to Boehm et al's (2002) study in marine recreational waters; we also examine seasonal and other drivers of variability of water quality in the home.

Methods

Water samples were collected in the village of Colon Eloy, a town of approximately 700 inhabitants living in 170 houses in the northwestern Ecuadorian province of Esmeraldas. Villagers use unimproved surface water from the Estero Maria, a small stream, as their primary source of drinking water. In surveys we carried out in the village, 71% of people interviewed report using the estero as their primary water source, and more than half (54-69%, depending on the season) drink their water without treating it. Alternate sources of drinking water include harvesting of rainwater and use of private unprotected hand-dug wells (reported as the primary drinking water source by 10% and 13%, respectively, of people surveyed). Some community members also import water from the larger Santiago River nearby during the dry season, when reduced flows are experienced in the Estero Maria. Inadequate sanitation infrastructure exists in this community and only 65% of houses have private latrines.

Water Quality Testing

Between January 2005 – March 2006, weekly samples were collected from five sites along the Estero Maria, one just upstream of the village (Site 1), one upstream of the major population center (Site 2), two more in the center of town (Sites 3 and 5), and one at the downstream end of the village (Site 6). Locations of sampling sites are shown in Figure 3.1. One day per month, samples were taken three times on the same day at the three middle sites (Sites 2, 3, and 5).

Weekly samples were also collected from seven households randomly selected using a block randomization scheme to ensure a distribution of houses throughout the village. The distribution of these houses is also indicated in Figure 3.1. To assess day-to-day variability, daily samples were collected for 11 consecutive days in the dry season from four stream sites at the same time each day. Samples from Sites 1, 2, and 4 were collected between 9:00-12:00 and samples from Site 6 were collected between 14:00-16:00. To assess hour-to-hour variability, samples were collected every 90 minutes between 5:00 – 23:00 and every 180 minutes between 23:00-5:00 for four consecutive days, once in the wet season and once in the dry season, at Site 4. These different sampling schemes are summarized in Table 3.1.

All samples were collected in Whirlpak bags (Nasco), immediately placed on ice, and processed within 24 hours. Petrifilm Coliform - *E.coli* count plates (3M) were used to detect and quantify *E.coli* colonies in the samples. These plates consist of plastic films with grids that are coated with gelling agents and Violet Red Bile nutrients, an indicator of glucuronidase activity (3M). Petrifilms were inoculated with 1 mL of water and incubated at 30ºC ± 2ºC for 24 hours.

Blue colonies were counted as *E. coli*. If a sample was suspected of being particularly clean (rain water, treated drinking water), the test was carried out in triplicate and the results of the three tests were summed and divided by 3 mL; this occurred in 15% of the weekly samples and none of the daily or hour-to-hour samples. Non-detects were included in the analysis as one-half of the lower detection limit. The plates that had too many colonies to count were assigned a value of 100 CFU/plate for the weekly samples collected over the course of the year, and a value of 450 CFU/plate for the samples collected on a day-to-day and hour-to-hour basis. This discrepancy is due to two different observers collecting the data for the different parts of the study. The number of samples above and below the detection limit for each of the sampling schemes is summarized in Table 3.1. In any analysis described herein that includes samples from both sampling schemes an upper limit of 100 CFU/plate was assigned. The number of *E.coli* colonies was multiplied by 100 to get a standardized total count per 100 mL. Possible results therefore ranged from 16.7 or 50 CFU/100 mL (halfway between zero and the lower detection limit of 33 or 100 CFU/100 mL) to 10,000 or 45,000 CFU/100 mL, depending on the analysis. These values were $\log_{10}$-transformed for use in the analysis.

Data Collection on Environmental and Household Covariates

Water temperature, pH, and electroconductivity were measured at source waters at the time of collection using a handheld device (Hanna Instruments, Ann Arbor, MI). The number of people in the river was also noted. River level and water clarity were measured each week at the most upstream location in the village (Site 1). Distance from a fixed point above the river (a cement bridge) to the water level was measured to calculate river height, and a Secchi disk dropped into the water from the same point was used to determine water clarity. Values for precipitation were taken from a rain gauge (Onset Computer Corporation, Bourne, MA) in the town of Borbón, twelve kilometers away. For analysis, the total rainfall accumulated by calendar week was used.

For each water sample collected in the households, a form was used to record the type of container in which the water was stored, whether or not the container was covered at the time of sampling, the source of the water (rain, well, Estero Maria, or Santiago River), how the water had been treated (no treatment, boiled, chlorinated, left to settle), how long the water had been stored in that container, and if the water was used for drinking (only asked in half of the samples). The first two variables were observed and the latter four were self-reported by household informants. Community health workers also visited each household in the village once per week and recorded water treatment practices as reported by a key informant in the household. Not all the houses were visited on the same day each week so data were aggregated by calendar week. The Institutional Review Boards at U.C. Berkeley and the Universidad San Francisco de Quito approved protocols for interaction with human subjects.

Data Analysis

Weekly geometric mean *E.coli* counts/100 mL for source and household water quality were calculated for all stream sampling sites and all households, respectively. Seasonal geometric mean counts of *E.coli*/100 mL were also calculated, both at the household and at stream sampling sites; we considered July – December the dry season and January – June as the rainy season.

Generalized estimating equations were used to estimate the association between water quality and various explanatory variables, controlling for correlation within sampling sites and within households, and using a log-linear (Poisson regression) model (Liang & Zeger 1986). Robust standard errors were specified to protect the inference against misspecification of this model. *E.coli* concentration (CFU/100 mL) was used as the outcome measure, and samples of source waters were analyzed separately from water sampled from household containers.

Explanatory variables were first modeled individually, and then combined into a multivariate model to produce adjusted coefficients. For the source samples, environmental covariates included *total weekly rainfall* (inches of rainfall summed over the calendar week), *river height* and *river clarity* (measured weekly at Sampling Site 1), *pH*, *electroconductivity* (mS/cm), *water temperature* (°C), and *number of people in the river* at the time of sample collection (measured for each source water sample). For the household samples, covariates included observed *container type* (small vs. large mouth), reported duration of *storage time* (hours), reported water *source*, reported *treatment*, whether the container was *covered* at the time of collection, and *weekly rainfall*. Additionally, whether the stored water was used as *drinking water* in the home was tested univariately; because this variable is not a driver of water quality and because this question was only asked for half of the samples it was not included in the multivariate analysis.

Lags of 1 to 4 weeks were tested for weekly rainfall, and a lag of one week was tested for river height and river clarity. The lag that best predicted water quality was used in the multivariate model. Lags were not tested for pH, electroconductivity, water temperature or number of people in the river. Because river level and river clarity lie in the causal pathway between rainfall and water quality, these variable were omitted from the multivariate analysis. The models were stratified by season to assess any interacting effects of season and rainfall.

Most variables describing the household and source samples were measured for each sample, and therefore have repeated values for the same date (one for each of the five stream sites sampled each week and up to three from each of seven households). Only weekly rainfall, river level, and water clarity were measured weekly, rather than on a per sample basis. Time series regression techniques therefore do not apply, because each outcome measure does not have a unique date of sampling associated with it. To adjust for temporal autocorrelation we carried out regressions of the geometric mean water quality across a) all sampling sites and b) all households for each week versus weekly rainfall (the variable with the most significant effect in the multivariate model). The inference on the associations was adjusted to account for the possibility of serial dependence of the residuals (a common problem when time-series are compared) by using a Newey regression approach (Newey & West 1987).

To assess the relative importance of different sources of variance in estimating mean water quality, we also fit several linear mixed models, with *E.coli* concentration (CFU/100 mL) as the outcome measure, and hour of day when the sample was collected (in 3-hour blocks) as a fixed effect (Laird & Ware 1982). The random effects in a mixed model are not directly estimated but are summarized according to their estimated variances and covariances, and therefore provide a way to compare the relative contribution of each of the sources of variance in determining water quality. We assessed the proportion of variance in the mean outcome attributable hierarchically (see model structure below) to: *(a) sampling site*; *(b) season* (rainy vs. dry); *(c) month* (1-15); *(d) week* (1-64); *(e) day* (1/28/05 – 3/02/06); and *(f) time of day* as represented by 3-hour blocks. To examine how Y, the concentration of *E.coli* (CFU/100 mL), varies from μ, the mean of Y, we used the following hierarchical model (as represented by the ordering to the subscripts):

(1) $Y_{abcdef} = \mu + b_1 x_1 + \alpha_a + \beta_{ab} + \gamma_{abc} + \delta_{abcd} + \zeta_{abcde} + \eta_{abcdef} + \varepsilon,$

where:

b_1 = regression coefficient;

x_1 = time of day (in 3-hr blocks);

α_a = variance attributable to sampling site;

β_{ab} = variance attributable to season given fixed sampling site;

γ_{abc} = variance attributable to month given fixed season & sampling site;

δ_{abcd} = variance attributable to week given fixed month, season, & sampling site;

ζ_{abcde} = variance attributable to day given fixed week, month, season, & sampling site;

η_{abcdef} = variance attributable to 3-hr block given fixed day, week, month, season, & sampling site; and

ε = residual variance.

In the model, temporal variables are nested under sampling site so the variance of the mean for each sampling site is measured over all seasons, months, weeks, days, and 3-hour intervals. All of the other variables are nested under sampling site.

To further discriminate the sources of variance in the residuals of this model, i.e., the variability unexplained by the factors included in the model, two additional models were analyzed. One includes a variable to assess the variance observed between *(g)* duplicate samples collected from the same location, and another to assess the variance observed between *(h)* duplicate laboratory analyses of the same sample. These models are represented by the following equations:

(2) $Y_{efg} = \mu + \zeta_e + \eta_{ef} + \theta_{efg} + \varepsilon$

(3) $Y_{efh} = \mu + \zeta_e + \eta_{ef} + \lambda_{efh} + \varepsilon$

where:

ζ_e = variance attributable to day;

η_{ef} = variance attributable to 3-hr block given fixed day;

and

ε = residual variance.

Date and hour were also included as variables in these two models in an effort to relate the inference to the first model. Separate models were necessary to estimate these terms due to the low number of field and lab duplicate samples (*n*=5 pairs and 10 pairs, respectively); had these terms been estimated with Model 1, that model would have been limited to this small number of samples. The combination of Models 2 and 3 was also tested, combining all duplicates of either category. The proportion of variance stemming from each variable in the model was assessed by calculating the variance of that factor divided by the total variance of the model. All analysis was carried out using STATA 9.0 (StataCorp LP, College Station, TX).

Results

Geometric means for all samples, for source samples, and for household samples, stratified by season, are shown in Table 3.2. In all cases, water contamination was higher in the wet season than in the dry season. A stronger seasonal difference was seen in source samples than in household samples. In four out of five source sites there was a significant difference between seasons, whereas only one of seven households had a significant difference. An overall trend of increasing contamination from Sampling Site 1 (upstream of the main human settlement) to Sampling Site 6 (downstream end of the village) was also observed (nonparametric test for trend across ordered groups: $p<0.0001$).

The variability at different timescales is shown in Figures 3.2 through 3.4. Hour-to-hour measurements over the course of four days show that *E.coli* counts can vary extensively, from 200-45,000 CFU/100 mL (mean: 5,383, σ: 11,363) in the dry season and from 300-45,000 CFU/100 mL (mean: 6,614; σ:11,401) in the wet season (Figure 3.2). A daily pattern was also detected, with peak counts at 16:00 in the dry season (t-test of difference from geometric mean for those sampling days: p=0.04) (Figure 3.3a), and two less distinctive peaks seen in the wet season, at 10:00, p=0.14; and 14:00, p=0.14) (Figure 3.3b). Values significantly lower than the geometric mean were observed during the dry season in the morning (at 05:00, p=0.06; 06:00, p=0.03; 07:00, p=0.004; 09:00, p=0.002) and late at night (at 23:00, p=0.009); in the wet season significantly lower values were observed during some of the morning hours (at 06:00, p=0.013; 08:00, p=0.02) and in the evening (19:00, p=0.008).

The stream overtopped its banks during the wet season. During the flood, *E.coli* counts were high, despite the fact that it occurred during nighttime hours. The flood also affected the peak hours for *E.coli* counts; before the flood highest levels of microbial contamination were observed between 10:00-11:00, and after the flood at approximately 12:00, although these peaks were not significantly different from the geometric mean. Water quality also showed extensive variability over the course of 11 days of sampling at four sites along the Estero Maria, from 17-2,800 CFU/100 mL (mean: 477; σ:502) (Figure 3.4).

The impact of various explanatory variables on water quality as estimated by generalized estimating equations for source samples is shown in Table 3.3. For source samples, in the unadjusted analysis of both seasons combined, a lag of three weeks had the strongest association with water quality outcome, even when accounting for multiple testing for different lags (Bonferonni-corrected level of significance: $p = 0.01$); this lag was therefore used in the subsequent multivariate and stratified analyses. The visual display confirms that the three-week lag on weekly rainfall is appropriate, because many of the peaks in water quality and rainfall match in time when the graph is shifted by three weeks (Figure 3.5). The results of the Newey West regressions also suggest that the 3-week lag is the best fit (Table 3.4). As mentioned above, river level and river clarity are not included in the multivariate model because they are in the causal pathway between rainfall and water quality. The correspondence of these three variables can be seen in Figure 3.6. Controlling for the *time of day* when the sample was collected did not affect the outcome of the model, and was therefore not included.

When both seasons are combined, the adjusted coefficient of 0.07 for weekly rainfall with a three week lag suggests that a one inch increase in rainfall was associated with a 7% increase in *E.coli* counts ($p<0.0001$). Post-estimation linear combination estimates suggest that a 10-inch increase in three-week-lagged rainfall was associated with a doubling of *E.coli* counts. This effect was not seen in the dry season ($\beta=0.00$, $p=0.99$), but was strongly evident in the wet season (β=0.08, p<0.0001). Electroconductivity also showed a significant effect in the wet season but not in the dry season, nor in both seasons combined. No other factors were significant when adjusting for the effects of the other variables. However, when *E.coli* concentrations for just the four days of hour-to-hour sampling were analyzed in separate Poisson regressions, ignoring the weekly and daily samples, number of people in the river at the time of sample collection was highly significant (β=0.18, p<0.0001). This suggests that each additional person in the river at the time of collection was associated with an 18% increase in *E.coli* counts.

In the household samples, a lag in weekly rainfall of two weeks was used, because this lag showed the strongest association with water quality (Table 3.5). In the households, the opposite effect was seen from that of the source samples. The adjusted coefficient suggests that a one-inch *decrease* in rainfall was associated with an 10% increase in *E.coli* counts (p<0.0001). This effect increased in the wet season. For household samples, the Newey West regressions did not show a significant result for any of the lags on weekly rainfall (Table 3.4). Other factors affecting household water quality included source of water (rain and well water were protective in both the dry and wet season, whereas only rain water was protective in the wet season); container (large-mouthed containers were associated with more contamination during the dry season only); and whether or not the container was kept covered. A significantly higher percentage of households reported treating their water during the dry season (mean: 46%, range: 24-64%) than during the wet season (mean: 31%, range: 0-71%); (one-sided t-test: p=0.0003).

The proportion of variance represented by each factor is shown in Table 3.6. Only 8% of variance was represented by sampling site. Season given fixed sampling site represents 14% of variability. Given fixed season, month-to-month variability explained little more of the variance (1%), but week-to-week fluctuations explained much more (30%). Given fixed season, month, and week, day-to-day variability explained little more of the variance (4%). Given everything else fixed, time of day explained 7%. A large amount of variance (35%) remained unexplained by the model.

Some of this unexplained variance was due to uncertainty stemming from the limitations of the measurement technology. Models 2 and 3 address this uncertainty by estimating the amount of the residual variance in Model 1 that can be explained by field and lab duplication. The coefficients representing date and time in these models can be considered to represent the variance explained by the variables in Model 1, and the remaining variance can be considered to be the variance attributable to measurement error and unexplained residual variance. The proportion of these two variances due to duplication of samples can roughly be considered the amount of the residual error from Model 1 attributable to uncertainty in measurement. More than half of the 35% residual variance from Model 1 could thus be explained by field duplication (0.09/0.18, or 50%), lab duplication (0.40/0.55, or 73%), or a combination of the two (0.17/0.20, or 85%). These percentages do not add up to 100% because they are from distinct models. The estimates exhibit a high degree of variability but they all suggest that a large proportion of the unexplained residual in the first model can be attributable to field and lab reproducibility, or uncertainty in measurement.

Discussion

In this paper we have explored sources of variability in water quality as measured by *E.coli* concentrations in source and household water samples at varying timescales over the course of one year in a rural Ecuadorian village.

Explaining Variability in Surface Source Water Quality

We observed large variations in concentrations of *E.coli* at different times of the day, week, and year. We observed just as much variability in water quality as measured by *E.coli* on an hour-to-hour basis (Figure 3.2) as on a day-to-day (Figure 3.4) or season-to-season (Figure 3.5) basis. In fact, hour-to-hour variability was greater than day-to-day variability at a fixed hour (Figures 3.2 and 3.4, Table 3.6). The greatest amount of variation in mean *E.coli* counts in the surface source waters was explained by week-to-week variability (30%), followed by seasonal variability (14%).

The high proportion of variance explained by week-to-week variability was likely driven by spikes in rainfall from one week to the next (Figure 3.5), although part of the weekly variability may be due to differences in measurement at different times of the day, which can affect peak indicator levels, as shown in Figure 3.3. We controlled for this by including the time of day at which the weekly samples were collected as a fixed effect in Model 1. However, to the extent to which the variability at the hour scale is random, confounding due to time of day may not be controllable with this method.

The seasonal effects observed also appear to be mostly driven by changes in rainfall (Table 3.3), although rainfall was more correlated with water quality in the wet season than in the dry season. Several mechanisms could explain the impact of rainfall on microbial contamination in the waterways. Increased rainfall may raise the counts of *E.coli* in the stream due to a runoff effect, whereby fecal material in the village environment is flushed into the river by the rains. Alternatively, scouring of riverbank soils due to heavy rains and increased river flows may release *E.coli* that are attached to sediments or that have grown in the soil. In small streams, the fraction of fecal coliforms attached to particles has been shown to increase with suspended solid content in the water (George et al. 2004).

In the dry season (during which it still rains periodically), peaks in microbial contamination appeared to occur mostly independently of rainfall pulses (Figure 3.5, Table 3.3). One explanation for this would be that both a “runoff effect” and a “concentration effect” are at play in this village. That is, during the rainy season, the river is subject to periodic pulses of fecal contamination that are driven by rainfall events, whereas during the dry season, observed peaks in contamination are driven by local contamination events that do not get flushed out of the river as quickly by higher wet weather rains or river flows. Overall levels of contamination were higher in the wet season than the dry season, suggesting that the runoff effect may be stronger than the concentration effect.

We would have expected to see a more immediate impact of rains on water quality in the stream if the “runoff effect” were occurring in the village, rather than a three-week delay. However, in reconsidering the data collection strategy, it appears that such a trend might not be detectable given the weekly sampling scheme. If fecal material present in the village were being flushed into the stream with rains, this effect might only be seen on a day-to-day basis, and the weekly sampling might have been too coarse to capture this effect. Indeed, the rains might initially flush fecal material into the stream but eventually dilute it. During the four-day intensive (hour-to-hour) sampling, increased levels of microbial contamination were observed immediately following the flood event during the hours of the day that normally experience lower contamination. This post-flood increase in contamination was followed by an overall decrease immediately after the flood event, suggesting that existing contamination might have been flushed out of the system in the short term.

We observed levels of microbial contamination to have an association with rainfall lagged by three weeks. This effect might be explained by the regrowth of *E.coli* in riverbank soils, due to cycles of wetting and drying, nutrient deposition during high flow conditions, and predation or competition with the indigenous soil microflora community. *E.coli* and other indicators of fecal contamination have been shown to survive and regrow in tropical (Byappanahalli & Fujioka 2004; Desmarais et al. 2002; Fujioka & Hardina 1995; Solo-Gabriele et al. 2000) and even temperate soils (Ishii et al. 2006). Several authors have found that the ability of *E.coli* and other indicator organisms to multiply in the soil is dependent on soil moisture content. In locations where indicator organisms have been cultured in the soil, the highest concentrations of indicator organisms have been found closest to the edge of riverbanks (Desmarais et al. 2002; Solo-Gabriele et al. 2000). At the same time, Solo-Gabriele et al. (2000) found that *E.coli* was capable of multiplying by several orders of magnitude upon a decrease in moisture content. They found that *E.coli* concentrations increased two days after storms ceased, explaining that the storm waters might have flushed *E.coli* from the soil banks and then it took roughly two days to increase to levels that affected water column to a noticeable degree. The authors posit that soil moisture appeared to be associated with predator survival, and *E.coli* can survive at lower soil moisture than its predators. Byappanahalli and Fujioka (2004) also suggest that interactions with major indigenous microflora played a large role in determining populations of *E.coli* in the soil.

Soil characteristics and nutrient availability are also factors. High organic content and a greater fraction of fine sediments, from deposition of suspended organic material, have been found to be more conducive for regrowth of *E.coli* (Desmarais et al. 2002). Soil provides suboptimal growth conditions for *E.coli* (Fujioka & Hardina 1995); nonetheless, it appears that the appropriate combination of factors (low competition and predation, availability of nutrients, and optimal soil moisture content) can enable significant growth of *E.coli* in the environment. In our study system, a combination of wetter conditions delaying optimal soil desiccation and the particular competitive cycles among *E.coli* and the indigenous microflora of the soil might be responsible for the three week lag in response time. Further research on the microbiology of these soils would be necessary to explore these ideas further.

The wide variation we observed on an hour-to-hour basis is also likely caused by the interplay of several factors. We posit that localized fecal contamination contributed by people bathing and washing in the river plays a large role. The number of people in the river at the time of collection was not a significant factor when assessed on a weekly basis over the course of a full year, but it was an important factor when just the data from the four-day intensive (hour-to-hour) sampling were analyzed. Each additional person in the river was associated with increases in *E.coli* levels of around 20%. Thus, in addition to poor sanitation in the village, bathing and washing clothes in the river are likely primary sources of fecal contamination in the surface waters. Peak *E.coli* counts occurred during daytime hours, despite the fact that solar radiation is known to reduce indicator counts (Boehm et al. 2002). In the dry season, the peak of contamination occurred at around 16:00 (t-test for difference from geometric mean: p=0.04), with decreased counts observed in the morning hours (05:00 – 09:00 and late at night (23:00). In the wet season two less distinct peaks were seen, one around 10:00 and another around 14:00 (not statistically significant), with decreased counts also observed in the morning hours (06:00-08:00) and in the evening (19:00). While we did not collect data on the total number of people in the river at all hours in the dry and wet seasons, we posit that the cumulative effect of people in the river all day might play a larger role in the dry season than in the wet season. This lends support for the role of the concentration effect in the dry season.

Explaining Variability in Stored Household Water Quality

For household water samples, both seasonality and household-level factors affected variability in water quality. Water source, container type, and whether the container was kept covered were consistently significant factors affecting household water quality in both seasons, whereas rainfall was significant in the wet season, but not in the dry season. Other studies have found these same factors to be important in determining household water quality (e.g., Mintz et al. 1995).

Treatment of water by boiling, chlorination, or settling did not show a significantly protective effect against microbial contamination in the samples taken from the seven households surveyed (and in some cases were actually a risk factor, although this effect was not significant) (Table 3.4). This may be due to misclassifaction due to respondent error (treatment practices were recorded as reported by key informants in the household), poor post-treatment storage practices leading to recontamination of water after treatment, or, in the case of chlorine, rapid dissipation of the chlorine residual due to reactions with other compounds (e.g. organic matter) (LeChevallier et al. 1980). In another study in Colon Eloy, no significant reductions were seen in counts of indicator organisms of households chlorinating their drinking water as compared to households that did not treat their water (McLaughlin et al. In Review). Boiling has also been shown to be an ineffective form of treatment due to recontamination during water handling (Oswald et al. In Press).

We found household water quality overall to be significantly less contaminated than source water quality, both in the wet and dry seasons (Table 3.2 and Figure 3.7). Geometric mean *E.coli* counts in the home were 260 CFU/100 mL, compared to 1,182 CFU/100 mL in the surface source water. In most regions, water stored in the home is significantly more contaminated than source water (Wright et al. 2004). However, most studies of contamination between source and point-of-use have been carried out in areas with improved water supplies. In locations with highly contaminated source waters, this result is less apparent

(Musa et al. 1999; VanDerslice & Briscoe 1995; Verweij et al. 1991). The result observed here is also similar to what we have observed throughout this region, where we see on average a more than half-log reduction of indicator organisms between the source of drinking water to its point-of-use. This may be due to settling or die-off of bacteria from turbid waters. (See Chapter 2 for further discussion.) In a future study, it would be useful to characterize daily and hourly variability of household drinking water from a variety of different sources in a manner similar to the way in which we characterized this temporal variability for source waters.

Interestingly, in contrast to what is observed in the source samples, rainfall is protective against contamination in the home. We posit that this may be due to a higher turnover rate of water stored in the home in the wet season, which would limit the opportunity for recontamination, or increased water available for hygiene. Unfortunately we did not collect data on the frequency with which households refill storage containers in the wet and dry seasons, so we cannot comment on this definitively. Another possibility is that households rely more on rain- and well-water in the wet season, both of which were significantly protective against microbial contamination (Table 3.4) in the wet season. Interestingly, data on water treatment trends across the village show that a higher percentage of households treat their water during the dry season than during the wet season. However, as mentioned above, treatment is not always effective, which might explain why treatment did not improve water conditions in periods of low rain.

Uncertainty in Measurement

While much of the variability in contamination levels in the home and at the source can be explained by temporal variability, environmental factors, and water storage and usage practices, a large amount of unexplained variability still exists. The analysis of duplicate samples in Models 2 and 3 suggests that uncertainty in measurement is responsible for more than half of the residual unexplained variance in Model 1. Unfortunately, the small number of duplicate sample pairs tested limits this analysis. In a future study, duplicate samples could be tested for a larger number of the samples, at all timescales, so that all of these sources of variability could be included in the same model analysis.

The large amount of measurement error may be due to the methods used to culture *E.coli*. Each petrifilm plate only tests a 1 mL water sample, which limits the precision of the analysis, especially in assessing the lower ranges of contamination. This was especially important in assessing lower contamination levels in the household samples (Table 3.1). A similar analysis using membrane filtration or most probable number techniques would offer more precision and might lead to lower estimates of uncertainty, although results with petrifilms were comparable to standard methods when 1 mL samples were consistently used (Vail et al. 2003).

Applications

The results shown here have practical applications. First, in this and other villages with inadequate sanitation that depend on untreated surface water, even in the absence of growth of indicator organisms in a given sample, water may be unsafe to drink. The main source of this contamination is likely to be feces, due to poor sanitation and use of the river for bathing and washing. Also, while the results of this study show that water can potentially be contaminated at any time of day, there are certain hours of the day and conditions under which water is more likely to be contaminated. Community health workers could redouble their efforts aimed at encouraging people to treat their drinking water during extreme conditions: after heavy rainfalls and during drought. In the absence of more comprehensive interventions involving in-home water treatment, safe storage devices, and sanitation improvements, community health workers could also encourage people to collect their water in the morning, before the onset of extensive human activity in the river, and urge villagers to keep their water storage containers covered.

Second, in measuring water quality, these results suggest that an analysis of more than one grab sample is needed to provide an accurate assessment of water quality at any given time. As suggested by Jensen et al. (2004), the relatively low sampling frequency of water quality in epidemiological studies may often fail to provide accurate estimates of the varying levels of fecal contamination that people are exposed to through drinking water. This may explain in part why indicators of water quality have failed to show a correlation with diarrheal disease, even when interventions to improve water quality are seen to be effective in preventing diarrheal disease (Gundry et al. 2004). While there are always trade-offs in designing a sampling scheme, in order to increase confidence in indicators of water quality, replication of results will be necessary. Boehm et al. (2002) suggest that the geometric mean standard, in which the geometric mean of multiple samples collected over a specified window of time, would be a better way to evaluate water quality.

This study also shows the importance of climatic drivers on water quality at both the household and village scales. Heavy rainfall events, flooding, and increased temperature associated with global climate change are all expected to affect the incidence of waterborne diseases (Hunter 2003), and climate change is expected to increase the variability of rainfall. Factors such as water supply and sanitation appear to modify the effect of climate on diarrheal disease, and should be considered when predicting the impact of climate change on diarrheal disease in areas with poor sanitation infrastructure. The effect of rainfall in other regions has been shown to vary depending on water source (Sandiford et al. 1989; Shiffman et al. 1976). Sanitation may therefore be interacting with rainfall in causing increased microbial contamination in unprotected wells and source waters following heavy rains, and should be considered in efforts to reduce waterborne illness.

Other anthropogenic changes in the landscape might also be affecting water quality in this village. Loss of forest cover may indirectly increase variability of water flow patterns and water quality in the stream. In 1996, a new highway was completed along the northern Ecuadorian coast to facilitate logging in the region, which is now considered one of the top 10 deforestation fronts in the world (Myers 1993). The village of Colon Eloy is situated along a secondary road that branches off of this main road, and has suffered high amounts of clearing of its surrounding forests. Higher rates of diarrheal disease, especially those with bacterial etiologies, have been observed in this region in villages closer to the new road (Eisenberg et al. 2006). Deforestation is known to affect local climatic and hydrological patterns (Gentry & Lopez-Parodi 1980; Neill et al. 2006), and decreased vegetative cover may lead to “flashier” conditions, with higher incidence of both flooding and drought in the stream.

In conclusion, it appears that several interacting variables, at many different spatial and temporal scales, determine water quality in this village. By analyzing sources of the observed variability in water quality measurements, we have distinguished some of the drivers of variability in water quality versus uncertainty in the measurements themselves. Rainfall and river level play a central role in determining water quality, but the effects of water levels act out differently at different timescales. Use of the river is also an important factor in determining hour-to-hour variability. Water storage and use patterns appear to play a more central role in determining microbial contamination in household containers. In household samples, rainfall is actually protective against microbial contamination, perhaps because water is replaced more often in the wet season, limiting the opportunity for recontamination in the home. People might also collect water from different locations along the river during the wet season. A large amount of uncertainty is associated with these measurements, suggesting that geometric means of indicator counts from multiple samples taken at different times would be a more appropriate measure of water quality than single grab samples. The close relationship of water quality and environmental factors suggests that climatic variables interact with sanitation and should be taken into account in considerations of how anthropogenic changes affect the transmission potential for diarrheal diseases.

Literature Cited

3M Petrifilm *E.Coli*/Coliform Count Plate Interpretation Guide. www.3m.com/microbiology. Accessed 5/1/07.

Altizer, S., A. Dobson, P. Hosseini, P. Hudson, M. Pascual, and P. Rohani. 2006. Seasonality and the dynamics of infectious diseases. Ecology Letters 9:467-4848.

Barrell, R. A. E., and M. G. M. Rowland. 1979. The relationship between rainfall and well water pollution in a West African (Gambian) village. Journal of Hygiene 83:143-150.

Blum, D., S. R. A. Huttly, J. I. Okoro, C. Akujobi, B. R. Kirkwood, and R. G. Feachem. 1987. The bacteriological quality of traditional water sources in north-eastern Imo State, Nigeria. Epidemiology and Infection 99:429-437.

Boehm, A. B., S. B. Grant, J. H. Kim, S. L. Mowbray, C. D. McGee, C. D. Clark, D. M. Foley, and D. E. Wellman. 2002. Decadal and Shorter Period Variability of Surf Zone Water Quality at Huntington Beach, California. Environmental Science & Technology 36:3885-3892.

Byappanahalli, M., and R. Fujioka. 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. Water Science and Technology 50:27-32.

Curriero, F. C., J. A. Patz, J. B. Rose, and S. Lele. 2001. The Association Between Extreme Precipitation and Waterborne Disease Outbreaks in the

United States, 1948-1994. American Journal of Public Health 91:1194-1199.

Desmarais, T. R., H. M. Solo-Gabriele, and C. J. Palmer. 2002. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Applied and Environmental Microbiology 68:1165-1172.

Dorner, S. M., W. B. Anderson, T. Gaulin, H. L. Candon, R. M. Slawson, P. Payment, and P. M. Huck. 2007. Pathogen and indicator variability in a heavily impacted watershed. Journal of Water and Health 5:241-257.

Drasar, B. S., A. M. Tomkins, and R. G. Feachem. 1978. Seasonal aspects of diarrhoeal diseases. Seasonal Dimensions to Rural Poverty. Institute for Development Studies, Sussex.

Eisenberg, J. N. S., W. Cevallos, K. Ponce, K. Levy, S. J. Bates, J. C. Scott, A. Hubbard, N. Vieira, P. Endara, M. Espinel, G. Trueba, L. W. Riley, and J. Trostle. 2006. Environmental change and infectious disease: How new roads affect the transmission of diarrheal pathogens in rural Ecuador. Proceedings of the National Academies of Science 103:19460-19465.

Fujioka, R., and C. Hardina. 1995. Soil: the environmental source of *Escherichia coli* and enterococci in Hawaii's streams. Environmental Toxicology and Water Quality 6:185-195.

Gentry, A. H., and J. Lopez-Parodi. 1980. Deforestation and Increased Flooding of the Upper Amazon. Science 210:1354-1356.

George, I., A. Anzil, and P. Servais. 2004. Quantification of fecal coliform inputs to aquatic systems through soil leaching. Water Research 37:611-618.

Gundry, S., J. Wright, and R. Conroy. 2004. A systematic review of health outcomes related to household water quality in developing countries. Journal of Water and Health 2:1-13.

Howard, G., S. Pedley, M. Barrett, M. Nalubega, and K. Johal. 2003. Risk factors contributing to microbiological contamination of shallow groundwater in Kampala, Uganda. Water Research 37:3421-3429.

Hunter, P. R. 2003. Climate change and waterborne and vector-borne disease. Journal of Applied Microbiology 94:37S-46S.

Ishii, S., W. B. Ksoll, R. E. Hicks, and M. J. Sadowsky. 2006. Presence and Growth of Naturalized *Escherichia coli* in Temperate Soils from Lake Superior Watersheds. Applied and Environmental Microbiology 72:612-621.

Jensen, P. K., G. Jayasinghe, W. van der Hoek, S. Cairncross, and A. Dalsgaard. 2004. Is there an association between bacteriological drinking water quality and childhood diarrhoea in developing countries? Tropical Medicine and International Health 9:1210-1215.

Kondo, H., N. Seo, T. Yasuda, M. Hasizume, Y. Koido, N. Ninomiya, and Y. Yamamoto. 2002. Post-flood--infectious diseases in Mozambique. Prehospital Disaster Medicine 17:126-133.

Laird, N. M., and J. H. Ware. 1982. Random-effects models for longitudinal data. Biometrics 38:963-974.

LeChevallier, M. W., T. M. Evans, and R. J. Seidler. 1980. LeChevallier, M. W., Evans, T. M., Seidler, R. J. 1980 Effect of Turbidity on Chlorination

Efficiency and Bacterial Persistence in Drinking Water. Applied and Environmental Microbiology 42 (1), 159-167. Applied and Environmental Microbiology 42:159-167.

Liang, K., and S. L. Zeger. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73:13-22.

Lindskog, R. U. M., and P. A. Lindskog. 1988. Bacteriological contamination of water in rural areas: an intervention study from Malawi. Journal of Tropical Medicine and Hygiene 91:1-7.

McLaughlin, L. A., K. Levy, N. K. Beck, J. S. Meschke, and J. N. Eisenberg. In Review. Observations on the Effectiveness of Point-Of-Use Chlorination in Rural Ecuador.

Mintz, E. D., F. M. Reiff, and R. V. Tauxe. 1995. Safe Water Treatment and Storage in the Home. Journal of the American Medical Association 273:948-953.

Musa, H. A., P. Shears, S. Kafi, and S. K. Elsabag. 1999. Water quality and public health in northern Sudan: a study of rural and per-urban communities. Journal of Applied Microbiology 87:676-682.

Myers, N. 1993. Tropical forests: The main deforestation fronts. Environmental Conservation 20:9-16.

National Research Council. 2005. Introduction and Historical Background. Pages 17-52 in Committee on Indicators for Waterborne Pathogens; Board on Life Sciences; Water Science and Technology Board, editor. Indicators for Waterborne Pathogens. The National Academies Press, Washington, D.C.

Neill, C., L. A. Deegan, S. M. Thomas, C. L. Haupert, A. V. Krusche, V. M. Ballester, and R. L. Victoria. 2006. Deforestation alters the hydraulic and biogeochemical characteristics of small lowland Amazonian streams. Hydrological Processes 20:2563-2580.

Newey, W. K., and K. D. West. 1987. A Simple, Positive, Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55:703-708.

Oswald, W., A. Lescano, C. Bern, M. Calderon, L. Cabrera, and R. Gilman. In Press. Fecal Contamination of Drinking Water Within Periurban Households, Lima, Peru. American Journal of Tropical Medicine and Hygiene.

Qadri, F., A. I. Khan, A. S. Faruque, Y. A. Begum, F. Chowdhury, G. B. Nair, S. M. A, D. A. Sack, and A. M. Svennerholm. 2005. Enterotoxigenic *Escherichia coli* and *Vibrio cholerae* diarrhea, Bangladesh, 2004. Emerging Infectious Diseases 11:1104-1107.

Sandiford, P., A. C. Gorter, G. Davey Smith, and J. P. C. Pauw. 1989. Determinants of drinking water quality in rural Nicaragua. Epidemiology and Infection 102:429-438.

Schwartz, B. S., J. B. Harris, A. I. Khan, R. C. Larocque, D. A. Sack, M. A. Malek, A. S. Faruque, F. Qadri, S. B. Calderwood, S. P. Luby, and E. T. Ryan. 2006. Diarrheal epidemics in Dhaka, Bangladesh, during three consecutive floods: 1988, 1998, and 2004. American Journal of Tropical Medicine and Hygiene 74:1067-1073.

Shiffman, M. A., R. Schneider, A. G. Turner, and R. W. Helms. 1976. Seasonality in Water Related Intestinal Disease in Guatemala. International Journal of Biometerology 20:223-229.

Singh, R. B. K., S. Hales, N. de Wet, R. Raj, M. Hearnden, and P. Weinstein. 2001. The Influence of Climate Variation and Change on Diarrheal Disease in the Pacific Islands. Environmental Health Perspectives 109:155-159.

Solo-Gabriele, H. M., M. A. Wolfert, T. R. Desmarais, and C. J. Palmer. 2000. Sources of *Escherichia coli* in a Coastal Subtropical Environment. Applied and Environmental Microbiology 66:230-237.

Thomas, M. K., D. F. Charron, D. Waltner-Toews, C. Schuster, A. R. Maarouf, and J. D. Holt. 2006. A role of high impact weather events in waterborne disease outbreaks in Canada, 1975-2001. International Journal of Environmental Health Research 16:167-180.

Vail, J. H., R. Morgan, C. R. Merino, F. Gonzales, R. Miller, and J. L. Ram. 2003. Enumeration of waterborne *Escherichia coli* with Petrifilm plates: Comparison to standard methods. Journal of Environmental Quality 32:368-373.

VanDerslice, J., and J. Briscoe. 1995. Environmental Interventions in Developing Countries: Interactions and their Implications. American Journal of Epidemiology 141:135-144.

Verweij, P. E., M. van Egmond, D. J. Bac, J. G. van der Schroeff, and R. P. Mouton. 1991. Hygiene, skin infections nd types of water supply in Venda, South Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene 85:681-684.

Vieira, N., S. J. Bates, O. D. Solberg, K. Ponce, R. Howsmon, W. Cevallos, G. Trueba, L. Riley, and J. N. S. Eisenberg. 2007. High prevalence of Enteroinvasive *Escherichia coli* isolated in a remote region of northern coastal Ecuador. American Journal of Tropical Medicine and Hygiene 76:528-533.

Wright, J., S. Gundry, and R. Conroy. 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Tropical Medicine and International Health 9:106-117.

Wright, R. C. 1986. The seasonality of bacterial quality of water in a tropical developing country (Sierra Leone). Journal of Hygiene 96:75-82.

Table 3.1: Summary of Datasets Used in the Analysis

Dataset	Sampling Location(s)	Sampling Dates	Sampling Frequency	# Samples Analyzed	Lower Detection Limit	# Results Below Detection Limit	Upper Detection Limit	# Results Above Detection Limit
Hour-to-Hour: Dry Season	Site 4	9/1/05 - 9/4/05	Every 90 min*	82	50	0	45,000	6 (7.3%)
Hour-to-Hour: Wet Season	Site 4	1/27/06 - 1/30/06	Every 90 min*	76	50	0	45,000	5 (6.6%)
Day-to-Day	Sites 1, 2, 4, 6	8/12/05 - 8/22/05	Daily	51	50	1 (2.0%)	45,000	0
Week-to-Week: Source	Sites 1, 2, 3, 5, 6	1/28/05 - 3/2/06	Weekly	306	16.7	20 (6.5%)	10,000	50 (16.3%)
Week-to-Week: Household	7 Households	1/28/05 - 3/2/06	Weekly	980	16.7	349 (35.6%)	10,000	53 (5.4%)

** Between 23:00 - 5:00 samples were collected only every 180 minutes*

Table 3.2: Geometric Mean Values by Season for weekly samples collected over 25 weeks in the dry season and 40 weeks in two wet seasons. Reported *p*-values are for the Mann-Whitney non-parametric test of equality of distribution between seasons. Note that Sample Site 4 was not sampled on a weekly basis and is therefore not included here.

	OVERALL	**DRY SEASON**	**WET SEASON**	***p-value***
All Weekly Samples	**375 (*n*=1,251)**	**293 (*n*=541)**	**451 (*n*=710)**	**0.0025**
Weekly Samples: Sources Only	**1,182 (*n*=238)**	**680 (*n*=106)**	**1,775 (*n*=132)**	**<0.0001**
Sampling Site 1	556 (*n*=53)	336 (*n*=25)	871 (*n*=28)	**0.0225**
Sampling Site 2	596 (*n*=48)	249 (*n*=21)	1,081 (*n*=27)	**0.0004**
Sampling Site 3	1,944 (*n*=50)	1,672 (*n*=21)	2,154 (*n*=29)	0.5755
Sampling Site 5	2,413 (*n*=46)	1,528 (*n*=20)	3,339 (*n*=26)	**0.0036**
Sampling Site 6	1,380 (*n*=41)	613 (*n*=19)	2,661 (*n*=22)	**0.0074**
Weekly Samples: Households Only	**260 (*n*=1,009)**	**224 (*n*=430)**	**291 (*n*=579)**	0.0770
Household 1	335 (*n*=139)	184 (*n*=61)	544 (*n*=78)	**0.0020**
Household 2	325 (*n*=141)	315 (*n*=60)	332 (*n*=81)	0.8939
Household 3	229 (*n*=144)	186 (*n*=63)	270 (*n*=81)	0.1890
Household 4	197 (*n*=140)	163 (*n*=62)	227 (*n*=78)	0.3316
Household 5	216 (*n*=145)	194 (*n*=58)	253 (*n*=87)	0.3832
Household 6	583 (*n*=147)	474 (*n*=60)	671 (*n*=87)	0.3423
Household 7	136 (*n*=153)	131 (*n*=66)	142 (*n*=87)	0.8115

Table 3.3: Results of generalized estimating equation regressions testing the effect of various covariates on counts of *E. coli* /100 mL for source samples, controlling for correlation within sampling sites. Unadjusted coefficients report results of univariate analyses, whereas Adjusted Coefficients report the results of multivariate analyses. Results are shown for the entire year and also stratified by season.

	BOTH SEASONS								DRY SEASON								WET SEASON							
Variable	Unadj. Coeff.	p-value	95% C.I.	N	Adj. Coeff.	p-value	95% C.I.	N	Unadj. Coeff.	p-value	95% C.I.	N	Adj. Coeff.	p-value	95% C.I.	N	Unadj. Coeff.	p-value	95% C.I.	N	Adj. Coeff.	p-value	95% C.I.	N
Dry Season (baseline)																								
Rainy Season	0.47	**0.03**	0.05 - 0.89	306				291								119								172
pH	-0.10	0.58	-0.47 - 0.26	300	0.06	0.71	-0.23 - 0.34		0.16	0.67	-0.57 - 0.89	123	0.10	0.55	-0.23 - 0.44		0.26	**0.03**	0.02 - 0.51	177	-0.12	0.70	-0.74 - 0.49	
Electroconductivity (mS/cm)	-1.44	0.61	-6.97 - 4.09	296	0.05	0.99	-7.37 - 7.47		0.72	0.91	-12.42 - 13.87	121	0.32	0.96	-12.42 - 13.06		7.89	**0.00**	3.25 - 12.53	175	11.93	**0.01**	3.17 - 20.69	
Water Temp (deg. C)	-0.05	0.40	-0.17 - 0.07	300	0.00	0.97	-0.07 - 0.06		0.03	0.72	-0.14 - 0.20	123	0.04	0.66	-0.14 - 0.22		0.06	**0.02**	0.01 - 0.10	177	0.05	0.29	-0.05 - 0.15	
# People in River	0.03	**0.00**	0.02 - 0.05	295	0.03	0.11	-0.01 - 0.06		0.04	0.56	-0.08 - 0.15	121	0.05	0.35	-0.05 - 0.14		0.02	0.23	-0.02 - 0.06	174	0.03	0.16	-0.01 - 0.07	
River Level (m)	-0.15	**0.05**	-0.30 - 0.00	301					0.95	0.68	-3.58 - 5.47	123					-0.38	**0.00**	-0.47 - -0.30	178				
River Level - 1 wk. Lag	-0.11	0.23	-0.30 - 0.07	296					-0.41	0.77	-3.12 - 2.30	123					-0.33	**0.00**	-0.51 - -0.15	173				
River Clarity (m)	-0.42	0.18	-1.03 - 0.19	301					-0.52	0.77	-4.03 - 2.98	123					-0.99	0.06	-2.03 - 0.04	178				
River Clarity - 1 wk. Lag	0.12	0.25	-0.09 - 0.32	296					-2.85	0.16	-6.83 - 1.12	123					-0.01	0.93	-0.18 - 0.17	173				
Weekly Rainfall (in)	-0.01	0.52	-0.05 - 0.03	306					0.11	0.80	-0.72 - 0.93	128					-0.07	**0.00**	-0.12 - -0.03	178				
Weekly Rainfall - 1 wk. Lag	-0.04	0.47	-0.14 - 0.06	306					0.21	0.69	-0.83 - 1.25	128					-0.15	**0.00**	-0.23 - -0.07	178				
Weekly Rainfall - 2 wk. Lag	0.00	0.98	-0.09 - 0.09	306					-0.06	0.87	-0.73 - 0.62	128					-0.09	**0.02**	-0.17 - -0.01	178				
Weekly Rainfall - 3 wk. Lag	0.08	**0.00**	0.03 - 0.13	306	0.07	**0.00**	0.06 - 0.09		0.03	0.91	-0.54 - 0.60	128	0.00	0.99	-0.75 - 0.74		0.05	**0.02**	0.01 - 0.10	178	0.08	**0.00**	0.05 - 0.11	
Weekly Rainfall - 4 wk. Lag	0.02	0.26	-0.02 - 0.06	301					-0.60	**0.02**	-1.10 - -0.11	128					0.00	0.89	-0.04 - 0.04	173				

Table 3.4: Results of Regressions of Geometric Mean Counts of E.coli across a) Sources and b) Households for each week against Weekly Rainfall with Various Lags, adjusting for serial autocorrelation with Newey-West standard errors.

	Lag	Coefficient	s.e.	p-value	lower CI	upper CI	N
a) SOURCE SAMPLES	0	52.2	76.5	0.50	-101.9	206.2	48
	1	-22.8	82.5	0.78	-188.9	143.3	48
	2	60.1	103.9	0.57	-149.0	269.1	48
	3	192.3	73.2	**0.01**	44.9	339.6	48
	4	28.8	74.1	0.70	-120.4	178.0	47
b) HOUSEHOLD SAMPLES	0	-43.2	24.5	0.08	-92.5	6.0	48
	1	-37.6	27.2	0.17	-92.4	17.2	48
	2	-52.5	28.8	0.08	-110.4	5.5	48
	3	-22.0	20.4	0.29	-63.0	19.0	48
	4	8.2	37.1	0.83	-66.5	82.9	47

Table 3.5: Results of generalized estimating equation regressions testing the effect of various covariates on counts of E.coli/100 mL in household samples, controlling for correlation within households. Unadjusted coefficients report results of univariate analyses, whereas Adjusted Coefficients report the results of multivariate analyses. Results are shown for the entire year and also stratified by season.

	BOTH SEASONS								DRY SEASON								WET SEASON							
Variable	**Unadj. Coeff.**	**p-value**	**95% C.I.**	**N**	**Adj. Coeff.**	**p-value**	**95% C.I.**	**N**	**Unadj. Coeff.**	**p-value**	**95% C.I.**	**N**	**Adj. Coeff.**	**p-value**	**95% C.I.**	**N**	**Unadj. Coeff.**	**p-value**	**95% C.I.**	**N**	**Adj. Coeff.**	**p-value**	**95% C.I.**	**N**
Dry Season (baseline)								874								382								492
Rainy Season	0.13	0.46	-0.22 - 0.48	980																				
Container: Small-mouthed (baseline)																								
Container: Large-mouthed	0.17	0.28	-0.14 - 0.47	937	-0.03	0.86	-0.37 - 0.31		0.53	**0.05**	-0.01 - 1.07	395	0.31	0.33	-0.31 - 0.93		-0.06	0.72	-0.38 - 0.26	542	-0.35	0.08	-0.74 - 0.04	
Storage time (hours)	-0.01	**0.00**	-0.01 - 0.00	906	0.00	0.18	-0.01 - 0.00		-0.01	**0.05**	-0.02 - 0.00	391	0.00	0.96	-0.01 - 0.01		-0.01	**0.01**	-0.02 - 0.00	515	-0.01	0.12	-0.02 - 0.00	
Source: Estero Maria (baseline)																								
Source: Rainwater	-1.10	**0.00**	-1.48 - -0.72	960	-0.58	**0.00**	-0.95 - -0.21		-1.53	**0.00**	-2.08 - -0.98	405	-1.04	**0.00**	-1.41 - -0.67		-1.17	**0.00**	-1.49 - -0.85	555	-0.81	**0.00**	-1.09 - -0.52	
Source: Well water	0.13	0.74	-0.66 - 0.93		-0.94	**0.00**	-1.46 - -0.42		0.01	0.99	-1.50 - 1.52		-2.53	**0.00**	-3.91 - -1.15		0.22	0.32	-0.21 - 0.66		-0.14	0.65	-0.74 - 0.46	
Source: Santiago River	-0.72	**0.01**	-1.26 - -0.18		-0.31	0.32	-0.91 - 0.29		-0.65	**0.04**	-1.28 - -0.02		0.07	0.83	-0.59 - 0.74			--	-- --	--				
Treatment: None (baseline)																								
Treatment: Boiled	0.01	0.97	-0.35 - 0.36	963	0.15	0.29	-0.13 - 0.42		-0.28	0.37	-0.89 - 0.33	405	0.16	0.59	-0.42 - 0.74		0.26	**0.01**	0.07 - 0.46	558	0.19	0.26	-0.14 - 0.53	
Treatment: Chlorinated	0.11	0.81	-0.82 - 1.05		0.17	0.60	-0.46 - 0.79		0.73	**0.00**	0.54 - 0.92		0.58	**0.02**	0.10 - 1.06		0.09	0.88	-1.02 - 1.20		-0.07	0.89	-1.06 - 0.92	
Treatment: Left to Settle	-0.28	0.59	-1.29 - 0.74		-0.08	0.89	-1.28 - 1.11		--	--	-- --	--					-0.25	0.64	-1.27 - 0.78		-0.70	0.37	-2.23 - 0.83	
Uncovered (baseline)																								
Covered	0.47	**0.04**	0.03 - 0.91	956	0.47	**0.00**	0.18 - 0.77		0.96	**0.00**	0.58 - 1.34	403	0.91	**0.00**	0.43 - 1.39		0.18	0.49	-0.34 - 0.70	553	0.27	0.09	-0.04 - 0.58	
Drinking Water (baseline)																								
Not drinking water	0.65	**0.00**	0.23 - 1.07	409					0.62	**0.01**	0.17 - 1.07	247					0.62	0.31	-0.59 - 1.83	162				
Weekly Rainfall (in)	-0.09	**0.00**	-0.14 - -0.04	980					0.07	0.83	-0.54 - 0.67	408					-0.14	**0.01**	-0.25 - -0.04	572				
Weekly Rainfall - 1 wk. Lag	-0.04	0.16	-0.09 - 0.01	980					0.37	**0.01**	0.09 - 0.65	408					-0.09	0.10	-0.20 - 0.02	572				
Weekly Rainfall - 2 wk. Lag	-0.12	**0.00**	-0.15 - -0.08	980	-0.10	**0.00**	-0.16 - -0.04		-0.21	0.57	-0.93 - 0.51	408	-0.03	0.94	-0.83 - 0.76		-0.21	**0.00**	-0.31 - -0.10	572	-0.21	**0.00**	-0.31 - -0.11	
Weekly Rainfall - 3 wk. Lag	-0.03	0.42	-0.12 - 0.05	980					-0.57	0.13	-1.30 - 0.16	408					-0.06	0.20	-0.14 - 0.03	572				
Weekly Rainfall - 4 wk. Lag	0.02	0.27	-0.01 - 0.05	959					-0.23	0.47	-0.84 - 0.39	408					0.01	0.65	-0.04 - 0.07	551				

Table 3.6: Results of the Linear Mixed Models. Numbers reported are the fraction of the variance attributable to each variable in the model. Letters refer to the subscripts in Model Equations (1), (2), and (3).

	a	b	c	d	e	f	g	h	I	
	Sampling Site	**Season**	**Month**	**Week**	**Date**	**Time of Day**	**Field Duplicate**	**Lab Duplicate**	**Residual**	***N***
Model 1	0.07	0.16	0.004	0.28	0.05	0.07			0.37	517
Model 2					0.61	0.21	0.09		0.09	10
Model 3					0.04	0.40		0.40	0.15	20
Models 2 & 3					0.37	0.43	0.17		0.03	

Figure 3.1: Map of Colon Eloy indicating Water Collection Sites. Source sampling sites are numbered and indicated in black boxes; Household sites are indicated with a large circle. Note that location of household sites is slightly offset in order to protect the identity of human subjects. Weekly samples were collected at Sites 1, 2, 3, 5, and 6. Daily samples were collected for 11 consecutive days at Sites 1, 2, 4, and 6. Hourly samples were collected for 4 consecutive days in the wet and dry seasons at Site 4. Arrows indicate direction of river flow.

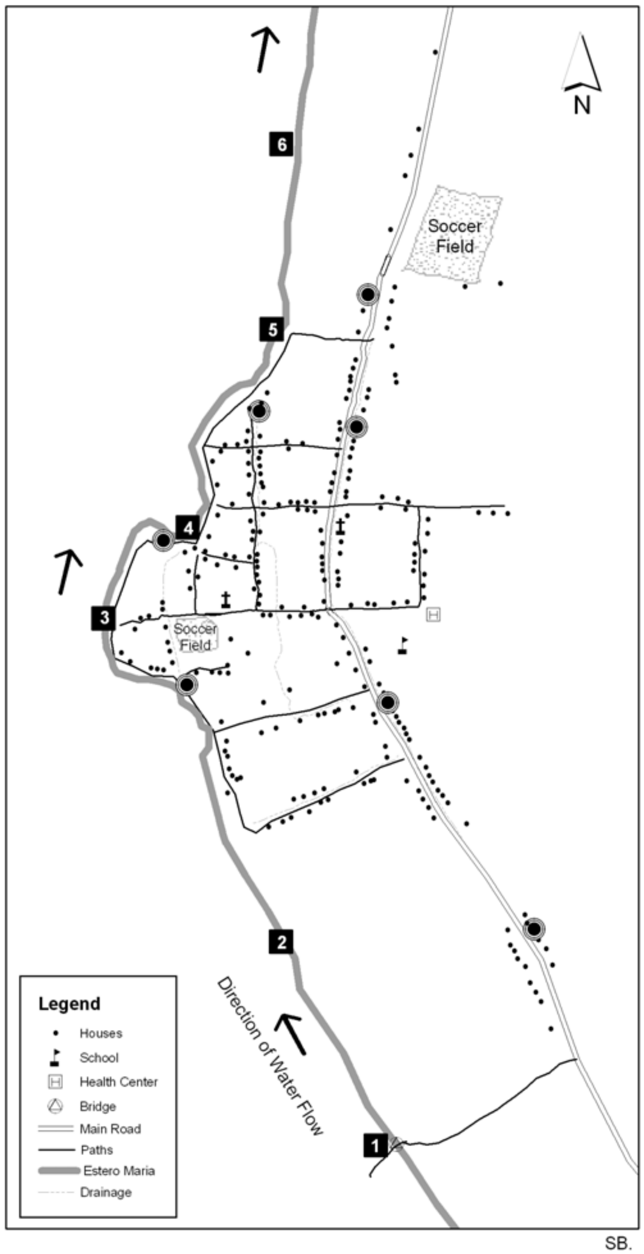
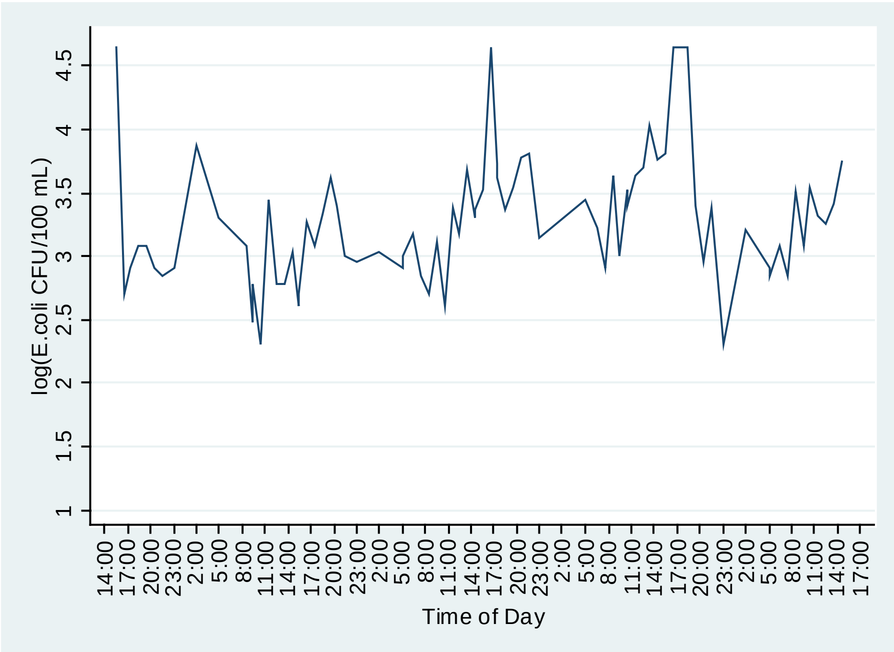


Figure 3.2: Hour-to-Hour Variability in Water Quality over the Course of Four Consecutive Days in Estero Maria, sampled at Sampling Site 4 in: a) the Dry Season (September 1-4, 2005); b) the Wet Season (January 27-30, 2006). Note that during the wet season the stream flooded; the time for which it flowed at greater than bankful flow is noted. Upper limit of detection is 4.65 (45,000 CFU/100 mL); Lower limit of detection is 1.70 (50 CFU/100 mL).

a)



b)

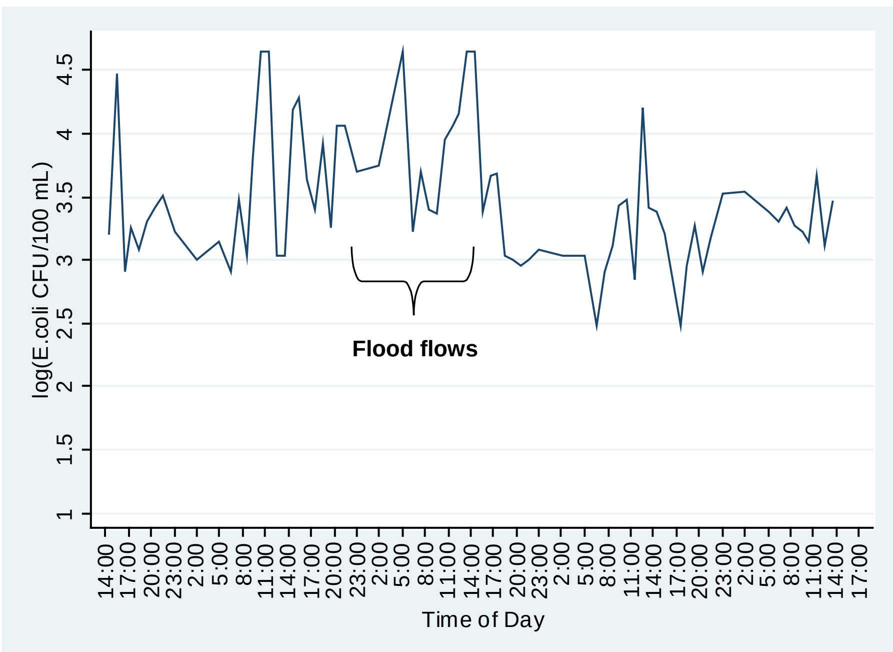


Figure 3.3: Log_{10} of Geometric Mean Counts of *E.coli* by Hour of the Day: a) Dry Season; and b) Wet Season. Upper limit of detection is 4.65 (45,000 CFU/100 mL); Lower limit of detection is 1.70 (50 CFU/100 mL). Geometric mean value for the season is shown with a dotted line.

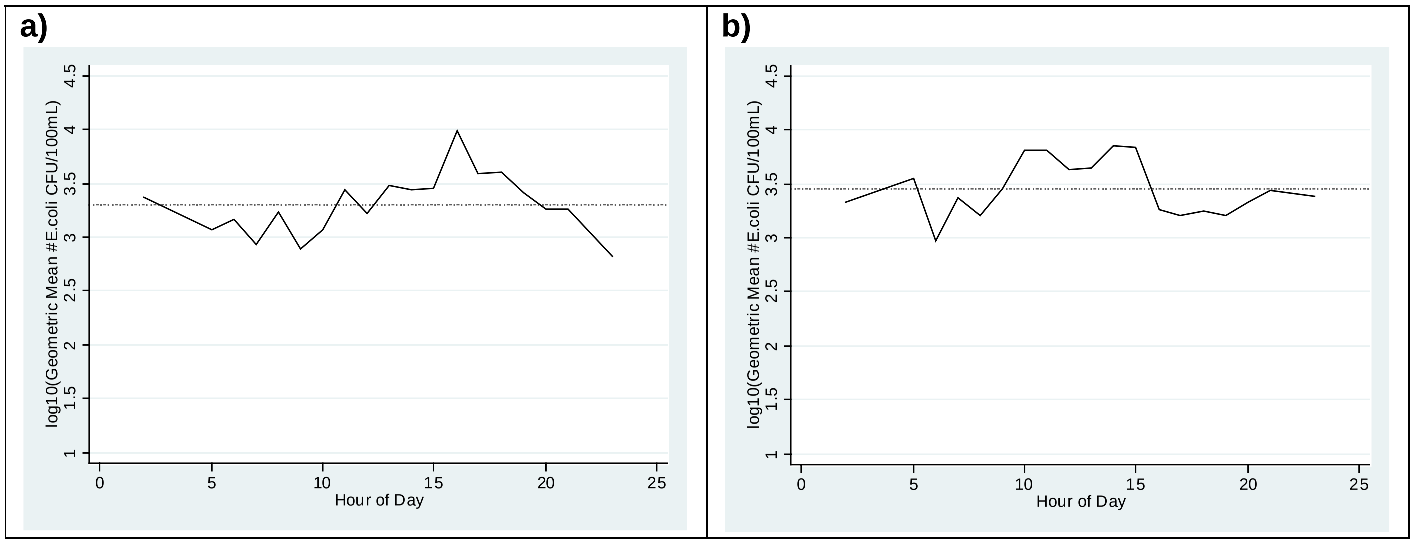


Figure 3.4: Day-to-Day Variability in Water Quality at Four Stream Sampling Sites on Estero Maria over the Course of 12 Days (8/11/05-8/23/05) during the dry season: a) Sampling Site 1; b) Sampling Site 2; c) Sampling Site 4; d) Sampling Site 6. Sites 1, 2, and 4 were sampled between 09:00-12:00 each day and Site 6 was sampled between 14:00-16:00. Upper limit of detection is 4.65 (45,000 CFU/100 mL); Lower limit of detection is 1.70 (50 CFU/100 mL).

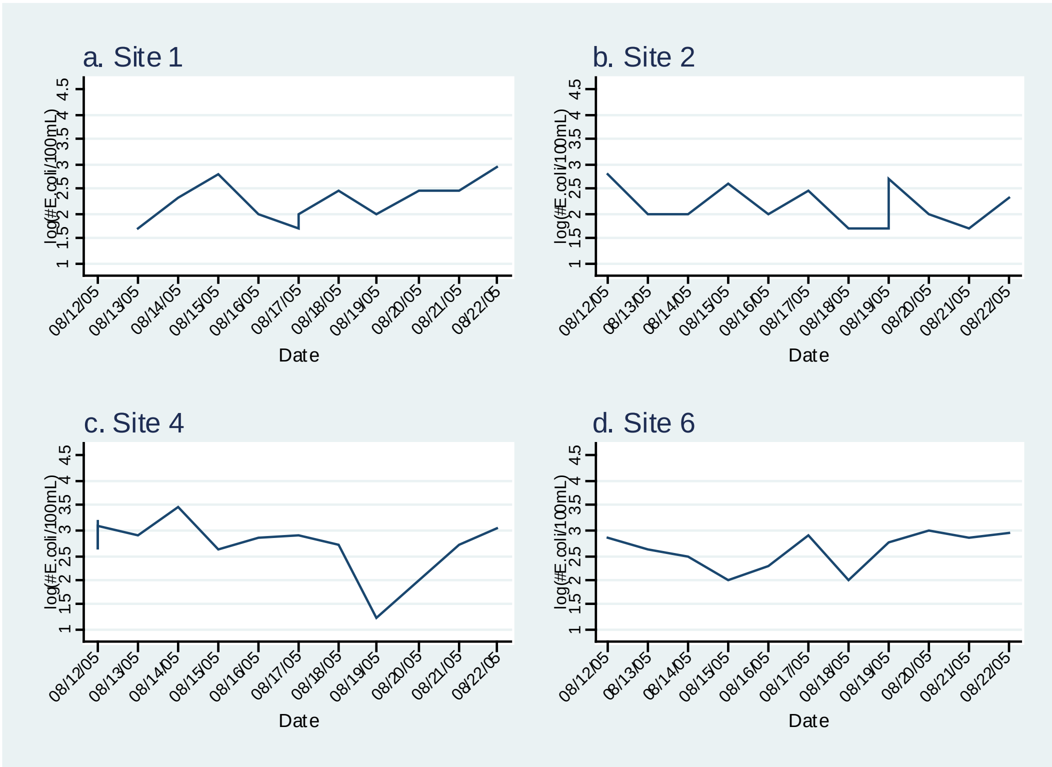


Figure 3.5: Variation in Geometric Mean Counts of *E.coli* in Source Waters over the course of the year (grey lines) and Variations in Rainfall (black lines): a) Weekly rainfall; b) Weekly rainfall lagged by three weeks. Data and lowess running-mean smooth (bandwidtch 0.8) lines are shown. For geometric mean counts of *E.coli,* upper limit of detection is 10,000 CFU/100 mL; Lower limit of detection is 17 CFU/100 mL.

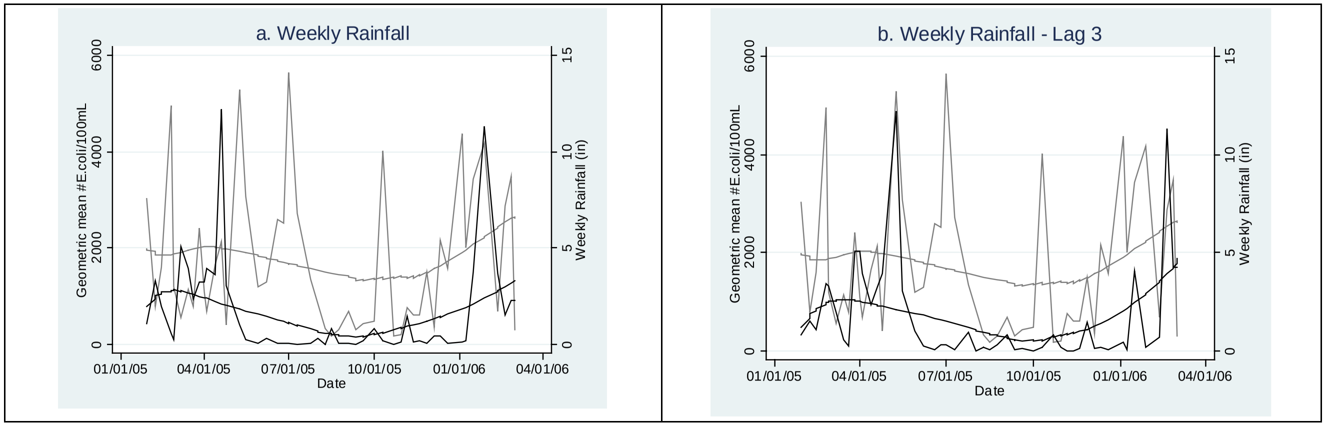


Figure 3.6: Correspondence of Variations in Weekly Measurements of Rainfall (inches; solid black line), River level (meters; dotted grey line), and River clarity (meters; dashed grey line) over the period of study.

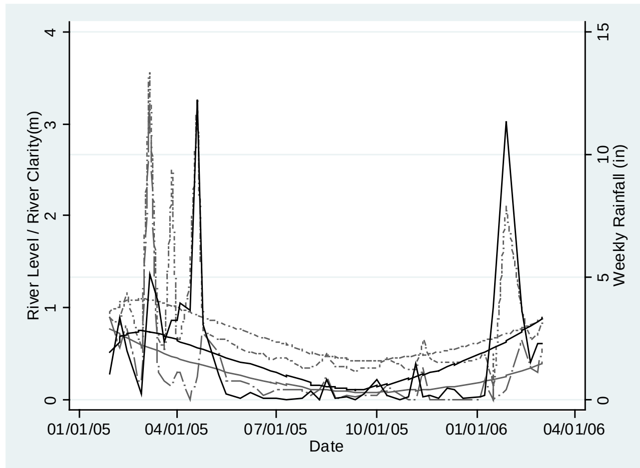
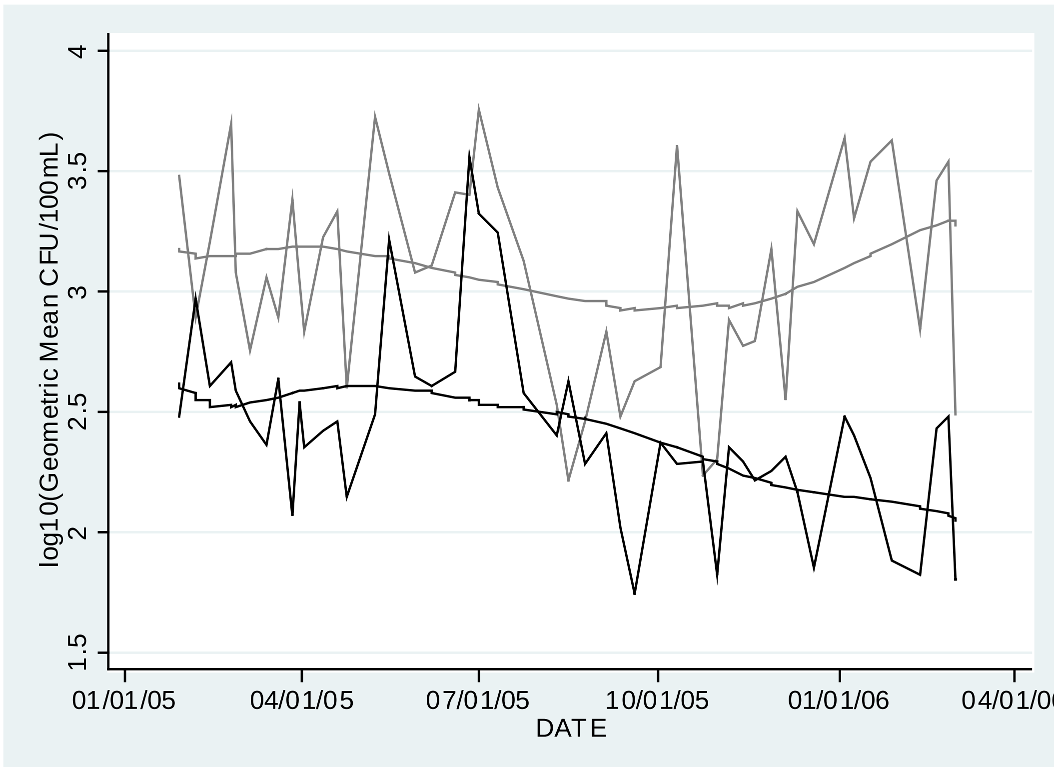


Figure 3.7: Variation in Geometric Mean Counts of *E.coli* in Water from Surface Source Waters (grey line) and Household Containers (black line) over the course of the year. Data and lowess running-mean smooth (bandwidtch 0.8) lines are shown. For geometric mean counts of *E.coli,* upper limit of detection is 10,000 CFU/100 mL; Lower limit of detection is 17 CFU/100 mL.



“Live each season as it passes; breathe the air, drink the drink, taste the fruit, and resign yourself to the influences of each.”

~ Henry David Thoreau

CHAPTER FOUR

SEASONALITY OF ROTAVIRUS DISEASE IN THE TROPICS: A SYSTEMATIC REVIEW AND META-ANALYSIS

Introduction

Throughout the world, rotavirus is the single most important viral agent of acute diarrhea in children. Rotavirus accounts for approximately 39% of all cases of severe diarrhea and over 600,000 deaths worldwide each year (Parashar et al. 2006). Unlike many bacteriological agents of diarrheal illness, rotavirus occurs in both temperate and tropical areas. Rotavirus infections are universal diseases of children, regardless of the level of hygiene that prevails or the quality of food and water. Children in developed countries are infected at the same ages and as frequently as those in less developed countries (Cook et al. 1990). However, an estimated 82% of the 1,205 children who die each day from rotavirus disease live in the poorest countries of the world, possibly because of inadequate access to hydration therapy and a greater prevalence of malnutrition (Parashar et al. 2003). A review of 43 studies from 15 countries on the African continent found rotavirus to be the single most common cause of diarrhea, responsible for about one-fourth of all diarrhea cases identified in both hospital patients and outpatients (Cunliffe et al. 1998).

Rotavirus is believed to be spread predominantly through fecal-oral transmission, through water, person-to-person contact, or contaminated environmental surfaces, although respiratory secretions have also been speculated as a source of infection (Parashar et al. 1998). Rotaviruses are highly infectious. Replication within the intestinal tract can result in shedding of ≥ 10^{10} infectious particles (PFU) per ml of feces and the infectious dose for the human small intestine has been calculated as only approximately 10 PFU/ml (Bishop 1996). Rotaviruses are very durable in the environment and can survive for weeks on surfaces and in potable and recreational water (Ansari et al. 1991).

Rotavirus exhibits distinct seasonality, and has been known as “winter diarrhea” in some parts of the world. “Winter gastroenteritis” and “winter vomiting disease” were recognized illnesses of early childhood before rotavirus was identified and found to be their cause (Cook et al. 1990). However, the most recent review of the global seasonality of rotavirus infections concluded that the winter seasonality of rotavirus infections is too simple a generalization. The authors of this study systematically reviewed 34 epidemiological studies of childhood diarrhea from a wide range of countries, and concluded that globally in temperate zones, rotavirus is certainly more common in the cooler months, but the seasonal peaks of the infections can vary broadly and occur from autumn to spring. Strict winter seasonality was common only in the Americas and was the exception in other parts of the world. In the tropics (between latitudes 23º27' north and south of the equator), this seasonality, defined by calendar year, was even less distinct.

Of ten surveys conducted within 10º north or south of the equator, eight exhibited no distinct seasonal trend. Farther from the equator, at latitudes 10º to 23º27' north and south, five of six studies showed a distinct seasonal peak—two in winter, two in the autumn and winter, and one in the autumn (Cook et al. 1990).

Despite these inconclusive results for the tropical belt, many other authors have referred to clustering during the cool, dry season in the tropics (Haffejee 1995). In a systematic review of rotavirus in Africa, rotavirus was detected year-round in nearly every country and generally exhibited distinct peaks during the dry months. Peaks were more common during dry periods than wet periods, but this pattern was not consistent for every country (Cunliffe et al. 1998). Because of the high burden of rotavirus disease in developing countries and because many of these countries lie in the tropical belt, further analysis of the seasonal patterns of rotavirus in tropical countries can shed light on the epidemiology of this important disease. Two live oral vaccines, prepared by GlaxoSmithKline (GSK) and Merck, have completed promising large scale clinical trials and are currently being introduced on a global scale (Glass & Parashar 2006; Ruiz-Palacios et al. 2006; Vesikari et al. 2006). The next generation of rotavirus vaccines will have greatest impact in developing countries where the disease burden is greatest (Glass et al. 2005; Glass et al. 2006; Parashar et al. 2003). This information on the role of climatic drivers of rotavirus transmission in the tropics is important in increasing our understanding of the epidemiology and transmission of rotavirus disease.

In their analysis, Cook et al. (1990) defined winter as December-March in the northern hemisphere and June-September in the southern hemisphere. While this classification of season by month is appropriate for temperate climates, in tropical zones distinct seasonality might occur that does not conform to these monthly classifications. In the tropics, seasonality may be driven by pressure belts and local air circulation patterns rather than changes in the amount of sunlight during different parts of the year. Therefore local climatological factors might provide more insight into seasonal patterns of rotavirus infection in this region of the world.

At the time that Cook et al. (1990) carried out their review, attempts to relate rotavirus disease incidence to climatological factors such as rainfall, humidity, and temperature had not provided any conclusive results. In the 16 years since the publication of their global review, many more studies have been carried out on rotavirus incidence, in anticipation of a vaccine for this disease. In an effort to increase our understanding of the transmission and epidemiology of rotavirus disease, this paper updates the knowledge of seasonality of rotavirus infection in tropical countries through a systematic review and meta-analysis of the relationship between monthly rotavirus prevalence and climatological variables (temperature, rainfall, and relative humidity) for those same months.

Methods

Search Strategy & Selection Criteria

Studies published between 1974-1988 were identified by Cook et al. (1990), and those carried out from 1988 through December 2005 were identified through a PubMed search for “rotavirus and season.” Additionally, reference lists from several key reviews of rotavirus epidemiology were scanned for identification of relevant papers (Ansari et al. 1991; Cunliffe et al. 1998; Parashar et al. 2003).

The criteria for study selection used by Cook et al. (1990) included: 1) conducted continuously for one year or more; 2) more than 50 confirmed cases of rotavirus diarrhea reported; and 3) monthly data on the proportion of all patients with diarrhea caused by rotavirus described. These same criteria were employed, although total number of monthly rotavirus cases was used, rather than proportion of diarrhea patients testing positive for rotavirus, to avoid specious results driven by seasonal changes in other diarrheal pathogens. In particular, bacterial pathogens in particular are known to peak in summer months. Assuming that each study carried out consistent case identification for all months of the year, rotavirus case count data would be reliable within each particular study. In addition, five further criteria were required for inclusion: 4) study site within 24º north or south of the equator; 5) monthly data on at least one of the following climatological variables was reported or available from another source: temperature, rainfall, or relative humidity; 6) study location was confined to a geographic area to which a single weather data set would apply (i.e., no country-wide studies); 7) study was written in either English or Spanish; and 8) study had a surveillance (rather than case-control or cohort) design to minimize variation in approaches to case identification; papers describing the results of case-control studies were only included if the authors reported rotavirus incidence for cases identified through hospital diarrhea surveillance separately from controls. Cohort studies were excluded. A total of 26 articles fit these criteria and were included in the analysis presented here.

All data reported in the published papers was taken directly from tables or extracted from graphs using DigitizeIt software (I. Bormann; www.digitizeit.de). Monthly climatological data for each region was taken from the published paper or from online databases of historical climatological data of the U.S. National Climatic Data Center (2005a; 2005b; 2005c). Latitude, altitude, average yearly temperature and average yearly rainfall (or number of days of rain per year for the study location) were taken from Weatherbase (2004).

Analysis

To estimate the association of climatological variables and the frequency of rotavirus, univariate linear regression was carried out for each study location, using monthly data. The inference on the associations was adjusted to account for the possibility of serial dependence of the residuals (a common problem when time-series are compared) by using a Newey regression approach (Newey & West 1987). A 5-month and 12-month lag in residual auto-correlations was assumed for studies that were carried out for less than or equal to and greater than 15 months, respectively; the resulting statistical inference was not sensitive to assumptions on the correlation lag.

Studies were also combined to get an overall (average) association between monthly rotavirus incidence (count) and mean temperature, mean rainfall, and relative humidity using a generalized estimating equation (GEE) approach, controlling for residual (within study) correlation and using a log-linear (Poisson regression) link (Liang & Zeger 1986). An auto-regressive (AR1) working correlation model was used, with robust standard errors to protect the inference against misspecification of this model. This approach accounts for correlation stemming from between-study variation as well as serial correlation in monthly data points within studies. Because studies differ in their mean numbers of rotavirus (due both to different incidences as well as different and sometimes unknown denominators) we also included “study” as a categorical (fixed effect) variable in the model. This can be thought of as stratifying by study and thus assumes studies have different underlying (baseline) mean counts of rotavirus (due in part to having different population sizes), but assumes the association (relative rate for a unit increase) of the climate variables is the same for all studies. If this is not true, the resulting estimate of the single association can be thought of as a weighted average relative rate. Because the climatological variables are highly related to one another, separate analyses were carried out for each.

A random effects model was used for assessing within- and between-study heterogeneity in rotavirus incidence and overall heterogeneity between studies in the association of rotavirus incidence and climatological variables was tested with a meta-analysis function. Covariates including altitude, latitude, average monthly rainfall, average temperature, range in temperature, range in rainfall, duration of study, and total number of rotavirus cases reported were explored for patterns that explained residual variability. Specifically, we included multiplicative interaction terms of categorical study*climatological variable and tested the overall significance of this interaction. We examined whether adding other study-specific variables changed the significance of the overall test of study*climatological variable interaction. All statistical analyses were carried out using STATA 8.0 (StataCorp, College Station, TX).

Results

A total of 26 studies from 15 countries between latitudes 24ºS and 24ºN were included in the analysis (Table 4.1). These studies took place between 1975 and 2003, with study duration ranging from one to ten years. In some studies, enrollment was limited to children, and in others all ages were included. However, most of the patients were children because of the predominance of rotavirus disease in infants and young children.

In the univariate analyses, an inverse relationship between monthly rotavirus incidence and climatological variables was consistently found in the data; 75%, 73%, and 100% of studies showed a negative association with temperature, rainfall, and relative humidity, respectively (of these, 65%, 55% and 60% showed a statistically significant negative association). For the same respective variables, only 10%, 18%, and 0% had significant positive correlations.

The pooled results are shown in Table 4.2. These models suggest that on average, across all studies analyzed, for every 7ºC increase in mean temperature (the median temperature range amongst the studies), the incidence of rotavirus decreases by 51% (IRR=0.49, p<0.0001); for every 31 cm increase in mean monthly rainfall (the median rainfall range), the incidence of rotavirus decreases by 21% (IRR=0.79; p=0.001); and for every 22% increase in relative humidity (the median range), the incidence of rotavirus decreases by 44% (IRR=0.56; p=0.018).

The meta-analysis revealed significant heterogeneity (p<0.0001) in studies for all climatological variables assessed. This heterogeneity between studies can be assessed visually in Figure 4.1, which plots the results of the univariate regression analyses, showing Newey West standard errors. Examining the residual variance of rotavirus incidence in a simple random effects models (one for each of the climatological variables of interest) within the random effects (GLS) model, 12-39% of unexplained variation in the model was seen within studies, and 61-88% was seen between studies (Table 4.2). The most relevant heterogeneity was that for the *associations* of incidence of rotavirus and climatological variables and this was significant for all the climate variables. When the interaction of these climatological variables and study was assessed (as described in the methods section), none of the covariates explored (altitude, latitude, average monthly rainfall, average temperature, range in temperature, range in rainfall, duration of study, and total number of rotavirus cases reported) could explain the between-study heterogeneity seen in the variability of association of climatological variables on the incidence of rotavirus.

Discussion

The results of this review suggest that rotavirus infections tend to peak under cool, dry conditions in the tropics. Negative associations between monthly rotavirus prevalence and climotological variables analyzed predominate in both the univariate and pooled analyses. A total of 23 studies (89%) showed an inverse correlation with at least one climatological variable (18 studies, or 69% statistically significant), compared with 10 (39%) showing a positive correlation (six studies, or 23% statistically significant) with at least one climatological variable (Table 4.1). According to the pooled GEE analysis, low values of all climatological variables predicted increased monthly incidence of rotavirus disease in patients with gastroenteritis, with temperature as the strongest predictor.

Using data on the total number of monthly rotavirus cases, rather than proportion of diarrhea patients testing positive for rotavirus, avoided the potential for reporting on patterns driven by seasonal changes in other diarrheal pathogens. However, we should note that similar results were found with both approaches (total case count and proportion of diarrhea cases testing positive for rotavirus).

The effect of seasonal changes on rotavirus incidence seen here is not as extreme in the tropics as it is in temperate areas of the world. Rotavirus is found year-round in the tropics with peaks and valleys, whereas incidence often goes to zero in some months in temperate areas. One explanation for this phenomenon is that less climatic variability exists in tropical climates and zones, so variations in climatological variables are not large enough to cause the observed effect. Still, the fact that rotavirus persists year-round in tropical areas of the world, and that rotavirus responds to climatic changes in many different climatic zones throughout the world, suggests that it is not an absolute temperature or humidity level that rotavirus prefers, but rather a relative change in climatic conditions.

We see a large amount of heterogeneity both within and between studies in the pooled analysis. The significant heterogeneity suggests that we would expect to see a stronger effect, and therefore have greater predictive power, if we could reduce some of the sources of variation between the different studies reviewed, such as socioeconomic status of patients, sampling scheme, diagnostic methods used, lengths of studies, numbers of participants sampled, populations of study regions, and differing climatic conditions at each study location. While all studies lie within the latitudes defined as the tropics, various climatic regimes (e.g., rainforest versus semi-arid) prevail in the different settings and at different altitudes, potentially confounding the results. We were unable to account for these differences in our analysis of potential covariates.

Understanding Rotavirus Transmission

The heterogeneity in effect observed in the pooled analysis is not surprising given that this analysis did not take into account additional factors potentially affecting rotavirus transmission, such as sanitation and hygiene practices or flood peaks. Several authors of the articles reviewed noted multiple peaks in rotavirus incidence as affected by the monsoon rains (Table 4.1). Flooding in conjunction with poor sanitation could augment the waterborne component of rotavirus transmission, obfuscating the seasonal patterns, which might be driven more by other routes of transmission, such as the air or fomites.

Strong evidence suggests that rotavirus is a waterborne pathogen. The virus can retain its infectivity for several days in aqueous environments, and waterborne spread has been implicated in a number of rotavirus outbreaks (Ansari et al. 1991). However, the high rates of infection in the first three years of life regardless of sanitary conditions, the failure to document fecal-oral transmission in several outbreaks of rotavirus diarrhea, and the dramatic spread of rotavirus over large geographic areas in the winter in temperate zones suggests that water alone may not be responsible for all rotavirus transmission (Parashar et al. 1998). No direct evidence shows that fomites and environmental surfaces play a role in the spread of rotavirus gastroenteritis, but indirect evidence shows that these possess a strong potential for spreading rotavirus gastroenteritis. Rotaviruses can remain viable on inanimate surfaces for several days when dried from a fecal suspension (Ansari et al. 1991).

Many authors have also suggested that rotavirus spreads through the air. In nosocomial outbreaks of rotavirus gastroenteritis, many patients show symptoms of upper respiratory tract infection before the onset of diarrhea, although the rotavirus has not been isolated from the respiratory tract (Ansari et al. 1991). Other evidence also points to a respiratory mode of spread. For example, large outbreaks have been recorded in both mid-Pacific island groups and Native American reservations where affected population groups were widely separated geographically. In both of these cases, each population had its own water supply and the rotavirus epidemic spread rapidly with little or no contact between the various groups (Haffejee 1995). In the United States, seasonal rotavirus activity occurs in a sequential manner, beginning first in the Southwest from October through December and ending in the Northeast in April or May (Torok et al. 1997). This pattern of seasonal spread is similar to other respiratory viruses such as influenza and measles (Haffejee 1995). Bishop (1996) proposes that the explanation for these patterns might lie in the airborne spread of aerosolized particles that are ingested, rather than through respiratory tract infections.

We speculate that the airborne component of rotavirus transmission might be responsible for the seasonal pattern of rotavirus disease. A relative drop in humidity and rainfall combined with drying of soils in higher temperatures might increase the aerial transport of dried, contaminated fecal material (in the form of droplet nuclei), and might also lead to increased formation of dust, which could provide a substrate for the virus particles. Increased burning of organic materials during the dry season may also increase the amount of particulates in the air. Airborne particles could settle out and infect water supplies or environmental surfaces, or could be ingested. Particles carried on fomites such as clothing could also play a role in large-scale dispersal of rotavirus organisms.

Some mechanical force would likely be required for aerosolization to occur, and wind might play this role, as well as help disperse the particles once formed. In a four-year study of rotavirus in Pune, India, a tight correlation was seen between number of days with easterly wind and the number of rotavirus diarrhea cases as functions of time (Purohit et al. 1998). Further research on the relationship between wind patterns and rotavirus incidence might shed further light on this potential transmission mechanism.

Conclusions

The widely cited conclusion of Cook et al. (1990) that seasonality of rotavirus infections in the tropics is less distinct than in temperate zones was based on a definition of winter as December-March in the northern hemisphere and June-September in the southern hemisphere. The results of this review suggest that local climatological variables provide a better predictor of seasonality in the tropics, where weather patterns differ from those in the temperate zones. This review reveals a trend for rotavirus to occur in the cool, dry seasons in tropical countries, as observed in temperate zones. These results suggest that paying close attention to *local* climatic conditions will improve our understanding of the transmission and epidemiology of rotavirus disease.

Literature Cited

Ansari, S. A., V. S. Springthorpe, and S. A. Sattar. 1991. Survival and vehicular spread of human rotaviruses: possible relation to seasonality of outbreaks. Reviews of Infectious Diseases 13:448-461.

Bingnan, F., L. E. Unicomb, Z. Rahim, N. N. Banu, G. Podder, J. D. Clemens, F. P. L. Van Loon, M. R. Rao, A. Malek, and S. R. Tzipori. 1991. Rotavirus-Associated Diarrhea in Rural Bangladesh: Two-Year Study of Incidence and Serotype Distribution. Journal of Clinical Microbiology 29:1359-1363.

Bishop, R. F. 1996. Natural history of human rotavirus infection. Archives of Virology, Supplementum 12:119-128.

Biswas, R., D. J. Lyon, E. A. Nelson, D. Lau, and P. J. Lewindon. 1996. Aetiology of acute diarrhoea in hospitalized children in Hong Kong. Tropical Medicine and International Health 1:679-683.

Brown, D. W., M. M. Mathan, M. Mathew, R. Martin, G. M. Beards, and V. I. Mathan. 1988. Rotavirus epidemiology in Vellore, south India: group, subgroup, serotype, and electrophoretype. Journal of Clinical Microbiology 26:2410-2414.

Chan, P. K. S., J. S. Tam, E. A. S. Nelson, K. S. C. Fung, F. A. B. Adeyemi-Doro, T. F. Fok, and A. F. Cheng. 1998. Rotavirus infection in Hong Kong: epidemiology and estimates of disease burden. Epidemiology and Infection 120:321-325.

Cook, S. M., R. I. Glass, C. W. LeBaron, and M. S. Ho. 1990. Global seasonality of rotavirus infections. Bulletin of the World Health Organization 68:171-177.

Cunliffe, N. A., P. E. Kilgore, J. S. Bresee, A. D. Steele, N. Luo, C. A. Hart, and R. I. Glass. 1998. Epidemiology of rotavirus diarrhoea in Africa: a review to assess the need for rotavirus immunization. Bulletin of the World Health Organization 76:525-537.

El Assouli, S. M., K. A. Mohammed, and Z. M. Banjar. 1995. Human rotavirus genomic RNA electropherotypes in Jeddah, Saudi Arabia from 1988 to 1992. Annals of Tropical Paediatrics 15:45-53.

Espejo, R. T., E. Calderón, N. González, A. Salomon, A. Martuscelli, and P. Romero. 1979. Presence of Two Distinct Types of Rotavirus in Infants and Young Children Hospitalized with Acute Gastroenteritis in Mexico City, 1977. The Journal of Infectious Diseases 139:474-477.

Glass, R. I., J. S. Bresee, R. Turcios, T. K. Fischer, U. D. Parashar, and A. D. Steele. 2005. Rotavirus vaccines: targeting the developing world. The Journal of Infectious Diseases 192:S160-S166.

Glass, R. I., and U. D. Parashar. 2006. The promise of new rotavirus vaccines. New England Journal of Medicine 354:75-77.

Glass, R. I., U. D. Parashar, J. S. Bresee, R. Turcios, T. K. Fischer, M. A. Widdowson, B. Jiang, and J. R. Gentsch. 2006. Rotavirus Vaccines: Current Prospects and Future Challenges. The Lancet 368:323-332.

Haffejee, I. E. 1995. The epidemiology of rotavirus infections: a global perspective. Journal of Pediatric Gastroenterology and Nutrition 20:275-286.

Hieber, J. P., S. Shelton, J. D. Nelson, J. Leon, and E. Mohs. 1978. Comparison of Human Rotavirus Disease in Tropical and Temperate Settings. American Journal of Diseases in Children 132:853-858.

Liang, K., and S. L. Zeger. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73:13-22.

Marjoribanks, H. C., M. C. Croxson, N. Potoi, and A. R. Bellamy. 1988. Infantile gastroenteritis in Western Samoa. The New Zealand Medical Journal 101:195-197.

Moe, K., and J. A. Shirley. 1982. The Effects of Relative Humidity and Temperature on the Survival of Human Rotavirus in Faeces. Archives of Virology 72:179-186.

Mpabalwani, M., H. Oshitani, F. Kasolo, K. Mizuta, N. Luo, N. Matsubayashi, G. Bhat, H. Suzuki, and Y. Numazaki. 1995. Rotavirus gastro-enteritis in hospitalized children with acute diarrhoea in Zambia. Annals of Tropical Paediatrics 15:39-43.

Muhe, L., B. Fredrikzon, and D. Habte. 1986. Clinical profile of rotavirus enteritis in Ethiopian children. Ethiopian Medical Journal 24:1-6.

Mutanda, L. N. 1980. Epidemiology of Acute Gastroenteritis in Early Childhood in Kenya III> Distribution of the Aetiological Agents. The East African Medical Journal 57:317-326.

Mutanda, L. N., S. N. Kinoti, W. Gemert, and E. O. Lichenga. 1984. Age Distribution and Seasonal Pattern of Rotavirus Infection in Children in Kenya. Journal of Diarrheal Disease Research 2:147-150.

Nelson, E. A. S., J. Tam, S., J. S. Bresee, K.-H. Poon, C.-H. Ng, K.-S. Ip, T. C. Mast, P. K. S. Chan, U. D. Parashar, T.-F. Fok, and R. I. Glass. 2005. Estimates of Rotavirus Disease Burden in Hong Kong: Hospital-Based Surveillance. The Journal of Infectious Diseases 195:S71-79.

Newey, W. K., and K. D. West. 1987. A Simple, Positive, Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55:703-708.

Parashar, U. D., J. S. Bresee, J. R. Gentsch, and R. I. Glass. 1998. Rotavirus. Emerging Infectious Diseases 4:561-570.

Parashar, U. D., C. J. Gibson, J. S. Bresee, and R. I. Glass. 2006. Rotavirus and severe childhood diarrhea. Emerging Infectious Diseases 12:304-306.

Parashar, U. D., E. G. Hummelman, J. S. Bresee, M. A. Miller, and R. I. Glass. 2003. Global Illness and Deaths Caused by Rotavirus Disease in Children. Emerging Infectious Diseases 9:565-572.

Pereira, H. G., A. C. Linhares, J. A. Candeias, and R. I. Glass. 1993. National laboratory surveillance of viral agents of gastroenteritis in Brazil. Bulletin of the Pan-American Health Organization 27:224-233.

Purohit, S. G., S. D. Kelkar, and V. Simha. 1998. Time series analysis of patients with rotavirus diarrhoea in Pune, India. Journal of Diarrheal Disease Research 16:74-83.

Rosa e Silva, M. L., F. G. Naveca, and I. Pires de Carvalho. 2001. Epidemiological aspects of rotavirus infections in Minas Gerais, Brazil. The Brazilian Journal of Infectious Diseases 5:215-222.

Ruiz-Palacios, G. M., I. Perez-Schael, F. R. Velazquez, H. Abate, T. Breuer, S. C. Clemens, B. Cheuvart, F. Espinoza, P. Gillard, B. L. Innis, Y. Cervantes, A. C. Linhares, P. Lopez, M. Macias-Parra, E. Ortega-Barria, V. Richardson, D. M. Rivera-Medina, L. Rivera, B. Salinas, N. Pavia-Ruz, J. Salmeron, R. Ruttimann, J. C. Tinoco, P. Rubio, E. Nunez, M. L. Guerrero, J. P. Yarzabal, S. Damaso, N. Tornieporth, X. Saez-Llorens, R. F. Vergara, T. Vesikari, A. Bouckenooghe, R. Clemens, B. De Vos, and M. O'Ryan. 2006. Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis. New England Journal of Medicine 354:11-22.

Soenarto, Y., T. Sebodo, R. Ridho, H. Alrasjid, J. E. Rohde, H. C. Bugg, G. L. Barnes, and R. F. Bishop. 1981. Acute diarrhea and rotavirus infection in newborn babies and children in Yogyakarta, Indonesia, from June 1978 to June 1979. Journal of Clinical Microbiology 14:123-129.

Stoll, B. J., R. I. Glass, M. I. Huq, M. U. Khan, J. E. Holt, and H. Banu. 1982. Surveillance of patients attending a diarrhoeal disease hospital in Bangladesh. British Medical Journal 285:1185-1188.

Suzuki, H., T. Sato, S. Kitaoka, F. Tazawa, T. Konno, Y. Amano, A. A. Alprecht, E. Gutierrez Vera, J. L. Villalta, Y. Numazaki, and N. Ishida. 1986.

Epidemiology of Rotavirus in Guayaquil, Ecuador. American Journal of Tropical Medicine and Hygiene 35:372-375.

Tam, J. S., W. W. S. Kum, B. Lam, C. Y. Yeung, and M. H. Ng. 1986. Molecular epidemiology of human rotavirus infection in children in Hong Kong. Journal of Clinical Microbiology 23:660-664.

Torok, T. J., P. E. Kilgore, M. J. Clarke, R. C. Holman, J. S. Bresee, and R. I. Glass. 1997. Visualizing geographic and temporal trends in rotavirus activity in the United States, 1991 to 1996. National Respiratory and Enteric Virus Surveillance System Collaborating Laboratories. The Pediatric Infectious Disease Journal 16:941-946.

U.S. National Climatic Data Center 2005a. Global Climate Observing System.http://www.ncdc.noaa.gov/gsn/gsn.Accessed November, 2005.

U.S. National Climatic Data Center 2005b. The Global Historical Climatology Network.http://lwf.ncdc.noaa.gov/cgi-bin/res40.pl.Accessed November, 2005.

U.S. National Climatic Data Center 2005c. Monthly Global Surface Data.http://cdo.ncdc.noaa.gov/pls/plclimprod/cdomain.DS3500?datasetabbv=DS3500.Accessed November, 2005.

Unicomb, L. E., F. Bingnan, Z. Rahim, N. N. Banu, J. G. Gomes, G. Podder, M. H. Munshi, and S. R. Tzipori. 1993. A one-year survey of rotavirus strains from three locations in Bangladesh. Archives of Virology 132:201-208.

Unicomb, L. E., P. Kilgore, A. S. G. Faruque, J. D. Hamadani, G. J. Fuchs, M. J. Albert, and R. Glass. 1997. Anticipating rotavirus vaccines: hospital-based surveillance for rotavirus diarrhea and estimates of disease burden in Bangladesh. The Pediatric Infectious Disease Journal 16:947-951.

Urquidi, V. 1989. Molecular epidemiology of human rotavirus infection in Coro, Venezuela. Acta Científica Venezolana 40:33-39.

Velazquez, F. R., G. Luna, R. Cedillo, J. Torres, and O. Munoz. 2004. Natural rotavirus infection is not associated to intussusception in Mexican children. Pediatr Infect Dis J 23:S173-178.

Vesikari, T., D. O. Matson, P. Dennehy, P. Van Damme, M. Santosham, Z. Rodriguez, M. J. Dallas, J. F. Heyse, M. G. Goveia, S. B. Black, H. R. Shinefield, C. D. Christie, S. Ylitalo, R. F. Itzler, M. L. Coia, M. T. Onorato, B. A. Adeyi, G. S. Marshall, L. Gothefors, D. Campens, A. Karvonen, J. P. Watt, K. L. O'Brien, M. J. DiNubile, H. F. Clark, J. W. Boslego, P. A. Offit, and P. M. Heaton. 2006. Safety and efficacy of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine. N Engl J Med 354:23-33.

Viera de Torres, B., R. Mazzali de Ilja, and J. Esparza. 1978. Epidemiological Aspects of Rotavirus Infection in Hospitalized Venezuelan Children with Gastroenteritis. American Journal of Tropical Medicine and Hygiene 27:567-572.

Weatherbase 2004.http://www.weatherbase.com/.Accessed May 2004.

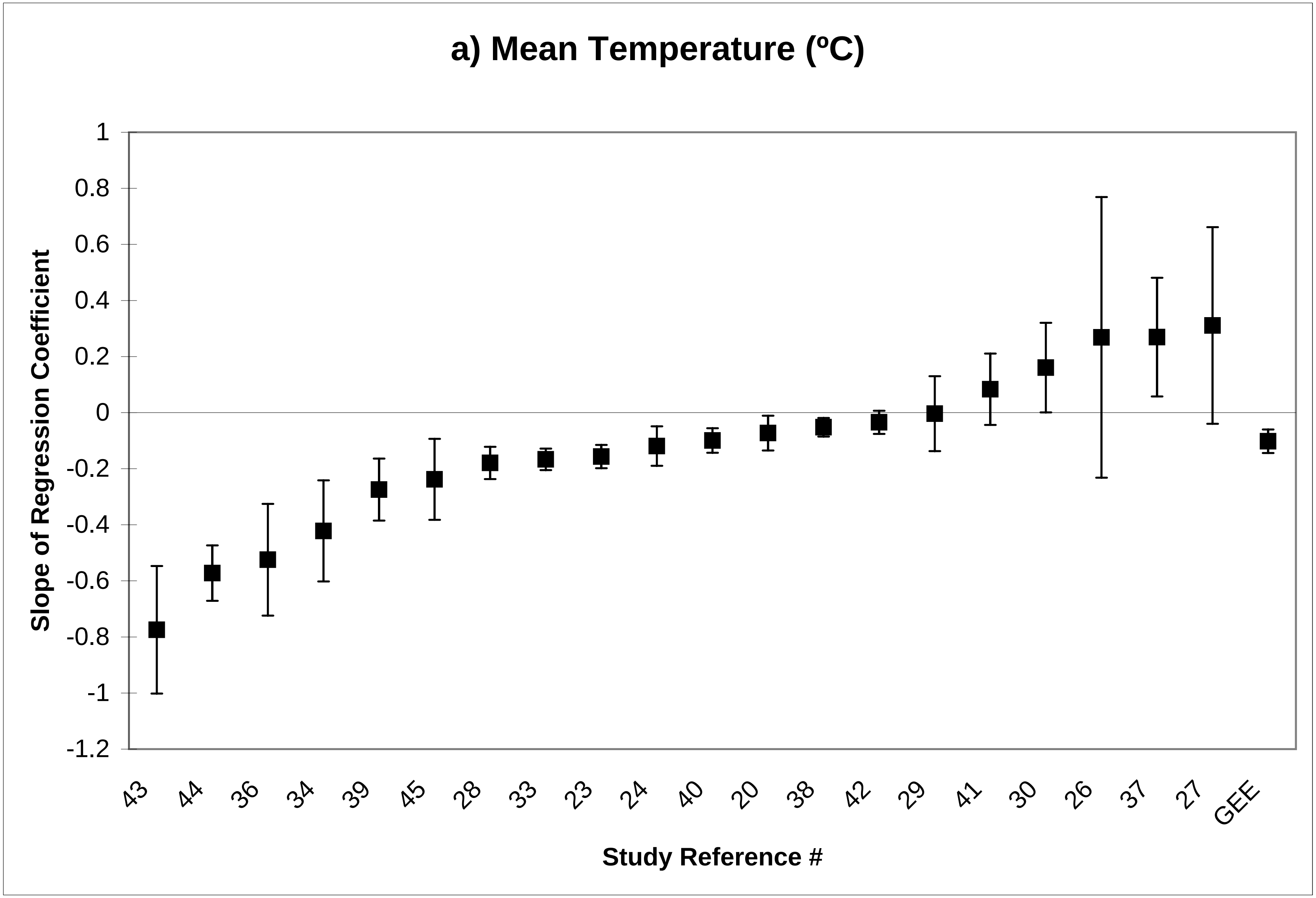
Table 4.1: Summary of Studies included in Review. Sign of coefficient of slope of regression between monthly rotavirus incidicence and monthly temperature (Temp), monthly rainfall (Rain) and relative humidity (RH) given in all cases where data was available. Asterisks indicate a significant (p<0.05) slope.

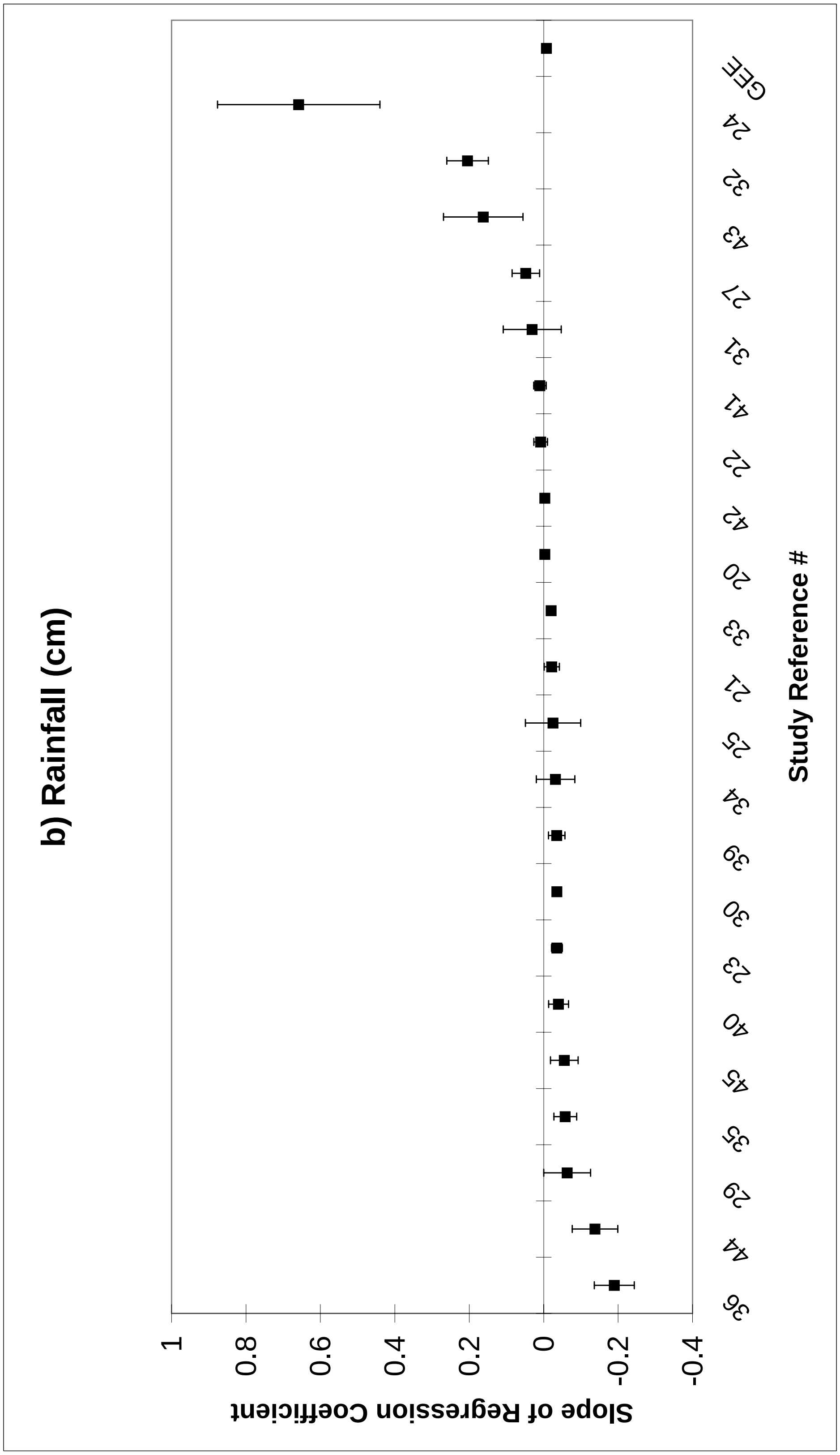
Ref.	City	Country	Latitude	Altitude	Avg. rain (cm)	Avg. temp (C)	Period of Study	Total # RV Seen	Ages included	Setting	Temp	Rain	RH	Peak Rotavirus Season as Noted by Authors
Bingnan et al. 1991	Matlab	Bangladesh	23N	10	235	25	1/87 - 5/89	898	all ages	2 diarrhea treatment centers + 1 hospital	- *	- *		Peaks in the dry winter months and coinciding with a major flood
Biswas et al. 1996	Hong Kong	China	22N	24	218	23	8/94 - 7/95	130	<5	hospital		- *		Peak in the cooler, drier months
Brown et al. 1988	Vellore	India	13N	204	127	28	8/83 - 7/85	163	<3	hospital		+		Peak in winter and SW monsoon season
Chan et al. 1998	Hong Kong	China	22N	24	218	23	1/87 - 12/96	2281	all ages	hospital	- *	- *		Peak in winter
El Assouli et al. 1995	Jeddah	Saudi Arabia	22N	11	4.5	28	3/88 - 12/92	523	infants & young children	hospital	- *	+ *		Increase in incidence in the cooler months
Espejo et al. 1977	Mexico City	Mexico	19N	2233	63	17	12/76 - 1/78	62	<5	2 hospitals		-		More frequent in autumn than winter
Hieber et al. 1978	San Jose	Costa Rica	10N	920	179	20	4/76 - 3/77	49	<8	hospital + outpatient clinic	+		-	Higher in dry season
Marjoribanks et al. 1988	Apia	Western Samoa	14S	2	285	26	6/83 - 6/85	67	<2	hospital	+	+ *		More common in the dry season
Moe et al. 1982	Yangon	Myanmar	17N	28	(147 days)	27	1/02 - 12/03	923	<5	tertiary care hospital	- *		- *	Cool, dry season
Mpabalwani et al. 1995	Lusaka	Zambia	15S	1153	(60 days)	21	1/92 - 12/92	256	<5	hospital	-	- *		Higher rate in the dry season
Muhe et al. 1986	Addis Ababa	Ethiopia	9N	2354	121	16	12/83 - 12/84	109	<3	hospital	+ *	- *		Year-round
Mutanda et al. 1980	Nairobi	Kenya	1S	1623	75	19	5/75 - 4/76	66	all ages	hospital		+		Lowest during the long rains
Mutanda et al. 1984	Nairobi	Kenya	1S	1623	75	19	12/81 - 6/83	133	<2	hospital		+ *	-*	Peak in times of dry, hot weather & low relative humidity
Nelson et al. 2005	Hong Kong	China	22N	24	218	23	4/01 - 3/03	1764	<5	4 hospitals	- *	- *	- *	Clear winter seasonal peak; infrequently identified in the summer season
Pereira et al. 1993	Rio de Janeiro	Brazil	23S	6	109	25	1/81 - 12/84	265	<5	hospital	- *	-		Peak in "winter" or dry season
Purohit et al. 1998	Pune	India	18N	559	70	24	7/92 - 6/96	266	<5	hospital		- *		More common in cooler months
Rosa e Silva et al. 2001	Minas Gerais	Brazil	22S	939	150	18	1/98 - 12/98	80	<5	hospital	- *	- *		Peak in coldest driest month
Soenarto et al. 1981	Yogyakarta	Indonesia	8S	106	218	25	6/78 - 6/79	126	<12	hospital	+ *		-	Decrease when moved from dry to wet conditions
Stoll et al. 1982	Dhaka	Bangladesh	24N	9	197	26	12/79 - 11/80	635	all ages	hospital	- *			Detection most common in winter
Suzuki et al. 1986	Guayaquil	Ecuador	2S	9	108	26	4/78 - 12/81	376	infants & young children	hospital	- *	- *		Detection rate higher in the dry season than the wet season
Tam et al. 1986	Hong Kong	China	22N	24	218	23	6/83 - 5/84	256	<5	hospital	- *	- *		Peaked during the winter months
Unicomb et al. 1993	Dhaka	Bangladesh	24N	9	197	26	7/89 - 6/90	756	all ages	hospital	+	+		Peak from July - January
Unicomb et al. 1997	Dhaka	Bangladesh	24N	9	197	26	1/90 - 12/93	1561	<5	hospital	-	-		Distinct peak in cool dry months
Urquidi et al. 1989	Coro	Venezuela	11N	17	41	28	12/84 - 12/87	115	<5	hospital	- *	+ *		Peaks in cooler months
Velazquez et al. 2004	Mexico City	Mexico	19N	2233	63	17	12/99 - 2/01	76	<2	5 hospitals	- *	- *		Marked increase in fall-winter
Vierra de Torres et al. 1978	Caracas	Venezuela	10N	834	(105 days)	23	11/75 - 12/76	121	<5	hospital	- *	- *		More frequent November - February

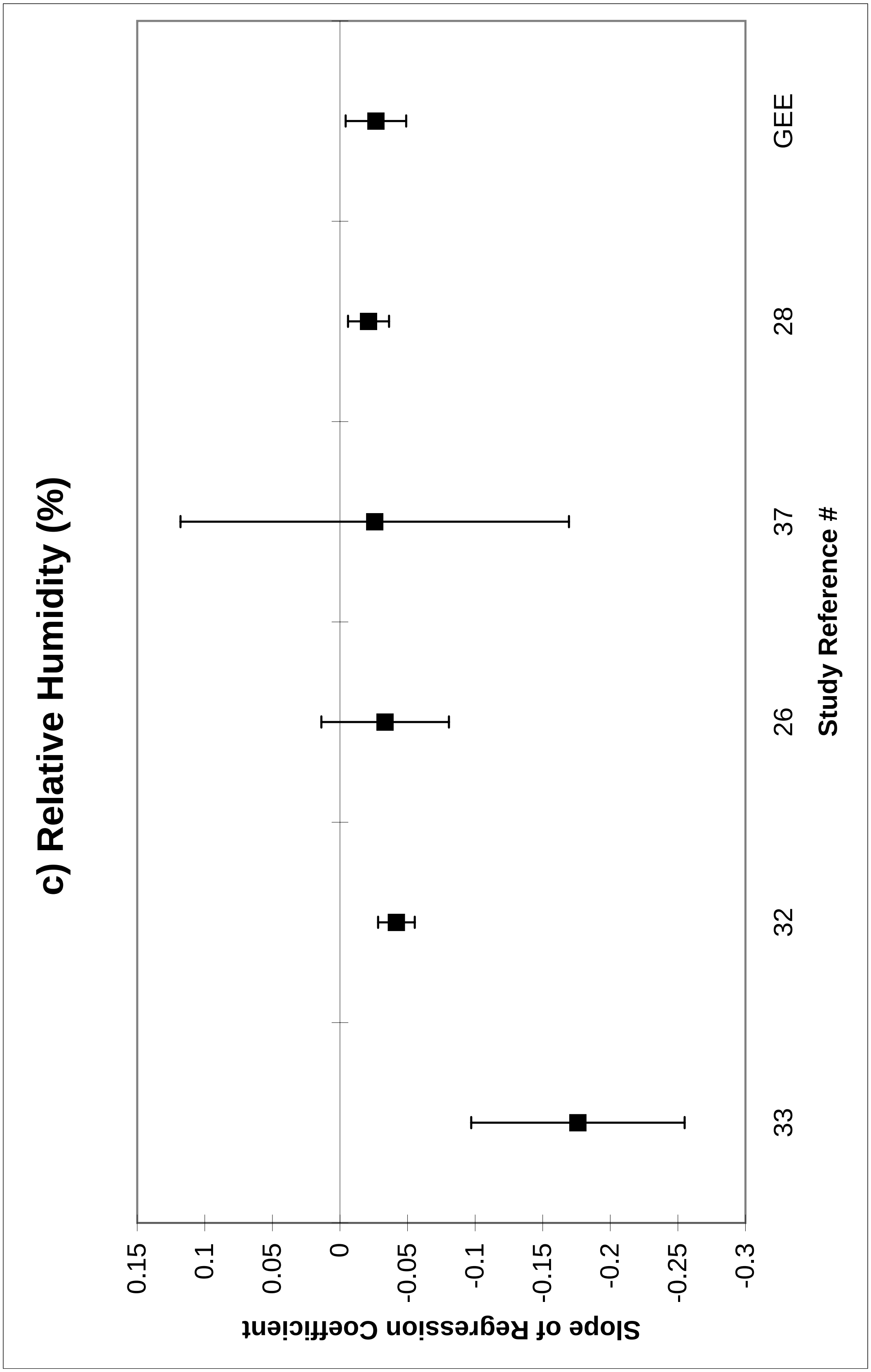
Table 4.2: Results of Pooled Analyses. Under GEE are the results of the Generalized Estimating Equation, and under GLS are the results of the Random Effects Model used to assess within- and between-study variability.

		GEE			GLS	
Climatological Variable	**# Studies Analyzed**	**Coefficient**	**Standard Error**	***p*-value**	**σ_u (between)**	**σ_e (within)**
Mean Temperature	25	**-0.102**	**0.021**	**<0.0001**	87.8%	12.2%
Monthly Rainfall	26	**-0.008**	**0.002**	**0.001**	83.8%	16.2%
Relative Humidity	5	**-0.027**	**0.011**	**0.018**	61.0%	39.0%

Figure 4.1: Heterogeneity plots of effect estimates, showing 95% confidence intervals using Newey West standard errors, for regressions of rotavirus incidence versus: (a) monthly mean temperature (ºC); (b) monthly rainfall (cm); and (c) relative humidity (%). The pooled effect from the generalized estimating equation (GEE) analysis is also shown.


a) Mean Temperature (ºC)
Slope of Regression Coefficient
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
43
44
36
34
39
45
28
33
23
24
40
20
38
42
29
41
30
26
37
27
GEE
Study Reference #


b) Rainfall (cm)
Slope of Regression Coefficient
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
Study Reference #
36
44
29
35
45
40
23
30
39
34
25
21
33
20
42
22
41
31
27
43
32
24
GEE


c) Relative Humidity (%)
Slope of Regression Coefficient
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
-0.25
-0.3
33
32
26
37
28
GEE
Study Reference #

“By the end of the nineteenth century all major cities of the western world had done something to come up to the new level of sanitation and water management that had been pioneered in Great Britain, 1848-1854. Urban life became far safer from disease than ever before as a result. Not merely cholera and typhoid but a host of less serious waterborne infections were reduced sharply. One of the major causes of infant mortality thereby trailed off toward statistical insignificance.

~ William McNeil, 1976

Plagues & Peoples

—

“Much work remains to be done before safe water is made as widely available as tobacco, alcohol, or carbonated soft drinks”

~ Eric Mintz et al., 2001

American Journal of Public Health

CONCLUSIONS

The reviews of the literature, field research, analysis, writing, and thought that have gone into the production of this dissertation have led me to two competing conclusions.

First, the association between water quality and diarrheal disease in areas with endemic diarrhea is highly complex and difficult to untangle given the current status of measurement technology. This is seen in the substance of the chapters of this dissertation. Chapter One details some of the inherent difficulties of measuring fecal contamination of source and household drinking water quality. Chapter Two shows that initial source water quality is an important factor in determining drinking water quality in the home, potentially more so than recontamination. These results and a careful review of the literature suggest that conclusions drawn from studies where people have access to improved water sources may not apply under circumstances where source water is highly contaminated. Chapter Three describes sources of variability and uncertainty in determining water quality in surface source waters and water stored in the home, as well as some of the complexities of interactions between water quality, climate, and sanitation. Chapter Four extends this analysis of climatic drivers of diarrheal illness, exploring the roles that rainfall, temperature, and relative humidity play in determining the incidence of rotavirus disease in the tropics.

I hope that the results presented here provide some insight about the role of environmental and seasonal factors in determining water quality and waterborne disease incidence in the tropics. Yet the results also show how difficult and complex a task it is to untangle the determinants of water quality and waterborne disease in the tropics. The study of diarrheal disease suffers from the problem of a lack of a "smoking gun." Diarrhea is a non-specific symptom, rather than a disease caused by one particular etiologic agent. It can be caused by infectious and non-infectious factors. The pathogens responsible for infectious diarrhea can be transmitted via food, water, personal contact, fomites, and sometimes the air. An individual's nutritional status, immunological naïveté, and genetic factors also play a large role in determining disease outcome. Water quality in a home and in a community plays a central role in determining the incidence of diarrheal disease, but a host of interacting biological, chemical, environmental and social factors are also at play.

This research is a contribution to our understanding, but so much more research is needed to fully understand the relative roles of sanitation, hygiene, and improvements to water sources in determining water quality, and also the extent to which water quality impacts diarrheal disease incidence. This topic is woefully understudied given that diarrheal disease still claims the lives of 2.5 million each year (Kosek et al. 2003). Reviews on interventions to improve water quality, water supply, sanitation, and other factors known to impact diarrheal disease have been limited to small numbers (<50 total) of studies on the topic (Clasen et al. 2007; Fewtrell et al. 2005; Gundry et al. 2004). The WHO spent only $7-10 million on controlling diarrheal diseases in 2004-05, compared to $216.7 million on HIV and $123.2 million on tuberculosis (Gerlin 2006). Yet diarrheal disease competes with these other illnesses in its contributions to the burden of childhood illness. In developed countries, resources continue to be devoted to research and analysis of water quality and waterborne illness, despite the fact that waterborne illness no longer poses a major public health problem in these countries. Between 1989-2000, a total of 278 outbreaks of waterborne disease occurred in the United States, including those caused by both recreational and drinking waters (National Research Council 2005). The number of outbreaks of waterborne disease in any given developing country in a single year would be several orders of magnitude higher than the number experienced in the U.S. over this twelve year period.

This brings me to my second conclusion, which is that the problem of waterborne disease is quite simple to understand, if not to solve. Diarrheal diseases are first and foremost diseases of poverty. That diarrhea still plagues us, given how much we know about how it is spread, and about how to prevent its transmission, remains a global scandal.

Debates ensue on the relative efficacy and effectiveness of various approaches to intervention: sanitation improvements versus increasing water supply versus water quality interventions in the home versus hygiene education. Yet despite these deliberations, everyone agrees on one conclusion: intervention works. All types of efforts to improve water quality, hygiene, and sanitation have independently been shown to be effective in preventing diarrhea (Fewtrell et al. 2005). We can argue about relative effect sizes but ultimately we must recognize that the burden of diarrheal disease depends on water, sanitation, *and* hygiene, and efforts to improve all three nodes of this famous triad are needed. The burden of diarrheal disease decreased in developed countries only after all three were improved.

One of the most promising recent advances in our understanding of drinking water quality is the recognition of the importance of recontamination of water in the household during storage, after collection (Wright et al. 2004). Many different types of “point-of-use” devices have been proposed to tackle the problem of in-home recontamination, and interventions to improve water quality in the home appear to be effective in most cases (Clasen et al. 2007; Fewtrell et al. 2005; Sobsey 2002). These developments and new devices are encouraging, and have added new hope to those involved in the field of drinking water in developing countries.

Sobsey (2002) sums up the current opinion in the field, that “Waiting for the provision of piped, microbiologically safe community water systems to the many people lacking such services is an inappropriate response to the basic need for safer drinking water that can be met on at least a provisional basis by available technologies. Effective measures are needed immediately to provide at-risk populations with safer water at the household level until the long-term goal of providing safe, piped, community water supplies can be achieved." In a comprehensive review of water quality interventions in settings with endemic diarrhea, Clasen and colleagues also write, "The cost and effort of combining the water quality intervention with improved hygiene, water storage, water supply, or sanitation may not be justified on the basis of a synergistic effect on diarrheal disease" (Clasen et al. 2006). I agree about the need to focus on practical, realistic, achievable goals and the promise of in-home water treatment to achieve improvements in water quality in the short-term. Yet it is important to recognize the implications of these kinds of statements.

As Paul Farmer writes, there is a distinction between declaring diseases "too expensive to treat" versus declaring people "too poor to treat" (Farmer 2001). "Appropriate technology" may only be appropriate because of the exceedingly limited funds available for "sustainable" projects.

The new focus on in-home devices to improve water quality hold much promise, but these developments should not make us complacent about the need to improve the other links in the chain of diarrheal disease transmission, which are tied to such local, regional and global phenomena such as sanitation, deforestation, climate change. It is important, while treating the immediate problems, to not lose sight of the distal factors driving the situation with which we are faced.

Literature Cited

Clasen, T., W. P. Schmidt, T. Rabie, I. Roberts, and S. Cairncross. 2007. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. British Medical Journal **334**:782-791.

Clasen, T. F., I. Roberts, T. Rabie, W.-P. Schmidt, and S. Cairncross. 2006. Interventions to improve water quality for preventing diarrhoea. Cochrane Database of Systematic Reviews.

Farmer, P. 2001. Infections and Inequalities: The Modern Plagues. University of California Press, Berkeley.

Fewtrell, L., R. B. Kaufmann, D. Kay, W. Enanoria, L. Haller, and J. M. Colford, Jr. 2005. Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. Lancet Infectious Diseases **5**:42-52.

Gerlin, A. 2006. A Simple Solution. TIME Magazine:42-47.

Gundry, S., J. Wright, and R. Conroy. 2004. A systematic review of the health outcomes related to household water quality in developing countries. Journal of Water and Health **2**:1-13.

Kosek, M., C. Bern, and R. L. Guerrant. 2003. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bulletin of the World Health Organization **81**:197-204.

National Research Council. 2005. Introduction and Historical Background. Pages 17-52 in Committee on Indicators for Waterborne Pathogens; Board on

Life Sciences; Water Science and Technology Board, editor. Indicators for Waterborne Pathogens. The National Academies Press, Washington, D.C.

Sobsey, M. 2002. Managing water in the home: accelerated health gains from improved water supply. World Health Organization, Geneva.

Wright, J., S. Gundry, and R. Conroy. 2004. Household drinking water in developing countries: a systematic review of microbiological contamination between source and point-of-use. Tropical Medicine and International Health **9**:106-117.