In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by the professor under whose direction it was written when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. In the absence of the professor, the dean of the Graduate School may grant permission. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

_____

Sarah Ceaser Tinker

Drinking Water and
Gastrointestinal Illness in Atlanta,
1993 – 2004

By

Sarah Ceaser Tinker
Doctor of Philosophy
Department of Epidemiology

_____
Paige E. Tolbert, Ph.D.
Advisor

_____
Christine L. Moe, Ph.D.
Committee Member

_____
Mitchel Klein, Ph.D.
Committee Member

_____
W. Dana Flanders, M.D., D.Sc.
Committee Member

_____
Appiah Amirtharajah, Ph.D., M.S.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____
Date

Drinking Water and Gastrointestinal Illness in Atlanta, 1993 – 2004

By

Sarah Ceaser Tinker
B.A., Northwestern University, 2000
M.P.H., Emory University, 2002

Advisor: Paige E. Tolbert, Ph.D.

An Abstract of
A dissertation submitted to the Faculty of the Graduate School of
Emory University in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Department of Epidemiology

2007

Previous research has suggested municipal drinking water may contribute to endemic gastrointestinal (GI) illness in the U.S., but the results were inconsistent, and the burden of GI illness attributable to drinking water contamination remains unclear. Three studies were conducted to examine the population impact of multiple surrogates of drinking water quality in Atlanta, Georgia. These analyses made use of an extensive emergency department (ED) database containing information on more than 10 million visits made to 41 hospitals between 1993 and 2004. The first of these studies considered the association of GI illness with an estimate of the time taken by drinking water to travel from the treatment plant to the end user (water residence time). The second study examined the role of the drinking water treatment plant itself as a risk factor for GI illness, as source water quality and treatment methods differ by plant. The final study examined the association between turbidity, the primary indicator of drinking water quality used by utilities, and ED visits for GI illness using time-series methods. The results support roles for both the raw water source and the distribution system as sites of drinking water contamination. Filtered water turbidity, a primary water quality measure used by the utilities, did not appear to predict risk. Overall, these studies suggest that a low level of GI illness in Atlanta may be attributable to drinking water exposure, particularly among young children and the elderly.

Drinking Water and Gastrointestinal Illness in Atlanta, 1993 – 2004

By

Sarah Ceaser Tinker
B.A., Northwestern University, 2000
M.P.H., Emory University, 2002

Advisor: Paige E. Tolbert, Ph.D.

A dissertation submitted to the Faculty of the Graduate School of
Emory University in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Department of Epidemiology

2007

ACKNOWLEDGMENTS

I am indebted to many people for their contributions, collaboration, and support on this dissertation and the work leading to its completion. I was fortunate to have Paige Tolbert as my dissertation committee chair, advisor, and mentor. She provided me with invaluable guidance and afforded me so many opportunities to grow as a researcher. I would particularly like to thank her for serving as a wonderful example of how to succeed as a mother working in academia. My understanding of drinking water and its importance in so many aspects of daily life is due to wonderful instruction and mentoring by Christine Moe, whose dedication to bringing clean water to the world is truly an inspiration. Working with Dana Flanders was an immensely educational, enjoyable, and humbling experience. His enthusiastic study of epidemiology and ethical philosophy regarding its applications will continue to serve as an example of how to practice epidemiology throughout my career. I am appreciative of Mitch Klein for his tireless dedication to his students and for the many hours he spent instructing, collaborating, and sparring with me on topics of epidemiology and life. My confidence as researcher was enriched by his encouragement.

The Atlanta water quality and health study would never have been possible without the collaboration of many people. Amit Amirtharajah contributed a deep understanding of Atlanta's drinking water infrastructure and his recruitment efforts with the drinking water utilities were vital to the success of the project. It was a privilege to work with Phillip Singer, who served as a wonderful bridge between the worlds of engineering and epidemiology. I would also like to thank Jim Uber for his work on developing innovative metrics for quantifying disease risk from drinking water and for

his enthusiasm about the potential partnership of drinking water engineering and epidemiology.

Working with the SOPHIA study team has been an immensely enjoyable and educational experience. I would like to thank Kristi Metzger and Jennifer Peel for serving as wonderful mentors and for laying the foundation for a tremendous project. Stefanie Sarnat has served as an advisor and friend and working with her has been a true pleasure. I am also grateful to the many research assistants who made compilation of the immense water quality dataset possible, particularly Aimee Cunningham and Catherine Kroll.

I would like to thank my fellow PhD students for their support and friendship. I am especially grateful for my friendship with Allison Curry, my epi partner-in-crime. She taught me how to learn epidemiology and helped me maintain my sanity. I am a better epidemiologist and friend for having met her.

I am so appreciative of my amazing family for their continued support and love through my many, many years of education. My parents, Jim and Nancy Ceaser, have never ceased to celebrate even my smallest academic achievement and never let me doubt my ability to succeed. My sister, Elizabeth, was always there to put a smile on my face when things got tough. I could not have made it through this experience without the love, counsel, and encouragement of my wonderful husband, Stuart. He is an equal partner in my accomplishments. Finally, I am thankful to my daughter, Lilly, who reminds me why public health is important and who serves as my guiding light.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Historically, drinking water has served as the source for some of the deadliest outbreaks of infectious disease. Typhoid fever and cholera were spread through contaminated water, and until the latter half of the 19th century, there remained a constant threat of acquiring these illnesses through the simple and necessary act of consuming water. In the 19th century, the link between infectious disease and water consumption was elucidated, and over time, communities began protecting their source water and treating their drinking water so as to minimize contamination. The U.S. Centers for Disease Control and Prevention (CDC) listed the provision of clean water as one of the most important advances in public health in the 20th century (U.S. Centers for Disease Control and Prevention 1999). The United States now supplies its residents with water of high quality, and diseases like cholera and typhoid fever are rare in this country.

Despite the enormous resources dedicated to keeping the U.S. water supply safe, waterborne disease outbreaks serve as a reminder of the potential for infectious disease to spread through the drinking water supply. The most dramatic example of this potential is the 1993 outbreak of cryptosporidiosis in Milwaukee, in which an estimated 403,000 people became ill, and 50 people died (Hoxie et al. 1997; MacKenzie et al. 1994). The source of the epidemic was traced to inadequate treatment of the city's drinking water (Eisenberg et al. 2005).

As drinking water quality has improved with respect to infectious disease risk, a concomitant increase in risk for adverse pregnancy outcomes and cancer has been

postulated from exposure to the chemicals used in disinfecting the water (Mills et al. 1998; Tardiff, Carson, and Ginevan 2006). The overall safety of drinking water is not simply a matter of performing as much treatment as possible, but has rather become a balancing of the risks from microorganisms against those from chemical products used in treatment. It is important to quantify the risk of infectious disease from drinking water in the U.S. in order to assess the adequacy of current treatment guidelines and requirements and to aid in the development of new policies, all with the aim of minimizing total health risk.

The importance of this task is highlighted by a major initiative of the U.S. Environmental Protection Agency (EPA) to develop an estimate of the endemic waterborne disease burden in the U.S., using results from published studies. Their findings have recently been published (Calderon, Craun, and Levy 2006). This task was made particularly challenging due to the conflicting results of many of the studies. The first major randomized-controlled trial examining this question, in which families were followed for development of gastrointestinal (GI) illness after being randomized to receive conventionally treated drinking water or a home-unit providing additional treatment, found that up to 40% of GI illness could be attributed to drinking water exposure (Payment, Richardson et al. 1991; Payment et al. 1997). Other randomized-controlled trials failed to find an association between drinking water and endemic GI illness (Hellard et al. 2001; Colford Jr. et al. 2005).

Evidence from observational studies also contributes to the collective knowledge about the burden of waterborne GI illness. A method which has been employed to examine this association is the time-series analysis, in which variations in water quality

over time are examined in relation to variations in disease occurrence through quantified healthcare utilization. There were two studies of this design conducted in Milwaukee in response to the 1993 *Cryptosporidium* outbreak (Morris et al. 1996; Morris, Naumova, and Griffiths 1998). Both found an association between turbidity, a measure of the cloudiness of the water and a rough proxy for microbial contamination, and hospital utilization for GI illness, even before the outbreak period. In 1997 Schwartz, et al. published results of an investigation of drinking water turbidity in relation to hospital visits for GI illness among children in Philadelphia (Schwartz, Levin, and Hodge 1997). In 2000 the same group published results of a similar study considering the elderly population in Philadelphia (Schwartz, Levin, and Goldstein 2000). Both studies found significant positive associations between turbidity and healthcare utilization for GI illness.

The overall goal of this dissertation is to address knowledge gaps by assessing the population impact on gastrointestinal illness attributable to multiple surrogates of drinking water quality using refined analytical techniques. Studies to date do not provide a clear understanding of the relative importance of the roles of the water quality at the treatment plant versus the water quality in the distribution system in contributing to subsequent gastrointestinal illness. This dissertation seeks to assess the role of drinking water in endemic gastrointestinal illness in Atlanta and to assess whether a role exists for the raw water source or the distribution system as a site of contamination leading to endemic GI illness.

Specifically, the following research questions are examined:

1. *Is incidence of gastrointestinal illness higher among people served by drinking water with longer estimated residence times?* Increased time within the distribution system (i.e., residence time) leads to an increased likelihood of contamination. The consideration of residence time accounts for the potential for microbial contamination throughout the entire distribution system. Previous studies focused only on measurements of indicator organisms taken at the plant or at a limited number of sampling sites within the distribution system to estimate total exposure to the microbial contamination in drinking water.
2. *Does incidence of gastrointestinal illness differ between populations served by different drinking water treatment plants?* Different treatment plants receive raw water from different sources and utilize different treatment processes. The age of the distribution systems served by different plants also varies. Therefore, differences in water quality may be more pronounced across different plants than temporally within each plant. While previous studies examined individual water systems and frequently considered only outbreak conditions, this study considers all six major water utilities serving the metro-Atlanta area during a time when no outbreaks were reported, providing estimates of the relative importance of system-specific factors that may contribute to endemic gastrointestinal illness.
3. *Are temporally varying measures of water quality related to daily incidence of gastrointestinal illness?* Previous time-series studies considered only filtered water turbidity, while this study considered two measures each for filtered and raw water turbidity. All results of the multiple tests required to adequately assess this relationship are presented, not only those that are statistically significant,

thereby reducing the potential for over-interpretation of significant results. This study examines turbidity from plants operated by all six of the major utilities providing Atlanta's drinking water, thereby increasing generalizability. This is a significant improvement over previous studies which examined only one utility.

This dissertation examines these three main research questions using information on emergency department (ED) visits from 41 of the 42 hospitals serving the 20-county Atlanta area and information regarding water quality from the six drinking water treatment utilities serving the five-county metro-Atlanta area from January 1, 1993 through December 31, 2004. The methods used to examine these research questions include logistic regression, Poisson regression, and time-series methods.

Atlanta provides an excellent setting for research to assess the role of drinking water in endemic GI illness, rather than epidemic levels of illness. Atlanta has high-quality drinking water, and there were no health-based EPA violations for any of the participating treatment plants during the study period. There were no drinking water-related outbreaks of GI illness in Atlanta during the study period, but endemic GI illness was present at low levels. The ED database is the largest known database of its kind and allows the studies to have high power to detect modest associations. The collaboration of the water utilities with the project greatly improves the availability and quality of relevant water quality data. This collaboration also provides the ability to consult with the utility personnel, enhancing the understanding of the data and our interpretation of the results. The analytical techniques use an *a priori* modeling strategy that reduces problems associated with multiple testing that were evident in other similar studies.

The results and conclusions of this research make several contributions, including adding to the existing body of literature on the magnitude of endemic waterborne disease, which is both limited and conflicting. This research has a special focus on the role of the distribution system and provides additional insight into the contribution of post-treatment contamination to endemic GI illness. Methodological concerns about previous similar studies are addressed where possible, providing a clearer interpretation of results. Ultimately, the results of this study should aid in informing decisions regarding the adequacy of current drinking water treatment and delivery.

CHAPTER 2

LITERATURE REVIEW: GASTROINTESTINAL ILLNESS

Enormous achievements have been made over the last 150 years in reducing the risk of morbidity and mortality from gastrointestinal illness. However, gastrointestinal illness remains an important public health issue. Worldwide, diarrheal illness is the third leading cause of death due to infectious disease and the seventh leading cause of death overall, causing four percent of all deaths and five percent of health loss due to disability (World Health Organization 1999, 2003, 2000). Although the majority of diarrheal illness occurs in developing countries, it is estimated that 211 million cases of gastrointestinal illness occur in the U.S. annually, resulting in an average of one episode per person each year (Mead et al. 1999).

Gastrointestinal illness, used here to describe diarrheal and other enteric diseases, is responsible for a large burden on the U.S. healthcare system. The hospitalization rate for adults 20 years of age and older is 2.7/1000/year, resulting in over 450,000 hospitalizations. The total number of deaths due to gastrointestinal illness in the U.S. has been steadily increasing since the early 1990s (Peterson and Calderon 2003). In 1996, over 8,000 deaths were attributed to specific enteric pathogens (Peterson and Calderon 2003). The true burden of gastrointestinal illness is likely much greater, given the inherent difficulty in identifying specific enteric pathogens. The cost of gastrointestinal illness, considering both the medical expenses and productivity loss, is estimated at over $23 billion a year (Garthright, Archer, and Kvenberg 1988). Although the causes of

gastrointestinal illness are varied and the implementation of prevention strategies complex, the costs associated with this type of illness warrant research on this topic.

Gastrointestinal Disease Processes

The gastrointestinal tract includes the esophagus, stomach, small intestine, colon, rectum, and anus. The surface of the gastrointestinal tract is composed of highly specialized epithelial cells that allow efficient absorption of nutrients, secretion of substances necessary for digestion, and excretion of waste products. The gastrointestinal tract must conduct these functions while protecting itself from invasion by pathogens. The mucosal surface itself and an extensive population of immune cells serve this protective function. Breaches of these defenses may disrupt the balance of the functions of the gastrointestinal tract and cause disease (Braunwald et al. 2003).

Diseases of the gastrointestinal tract manifest themselves through a variety of symptoms, including diarrhea, nausea, abdominal pain, bleeding, fever, constipation, and weight loss (Braunwald et al. 2003). Diarrhea, the most prominent of these symptoms, results from decreased absorption, increased secretion, or a change in motility of the gastrointestinal tract. Diarrhea is typically categorized as inflammatory or noninflammatory. Inflammatory diarrhea results from the disruption of the intestinal mucosal layer, either by direct invasion or cytotoxin release. This mucosal disruption leads to the emission of inflammatory cells, blood, and sera into the intestinal lumen. Noninflammatory diarrhea is caused by excessive secretion of ions and water stimulated by enterotoxins produced by infecting organisms (Friedman, McQuaid, and Grendell 2003).

Causes of Gastrointestinal Illness

Non-Infectious

Gastrointestinal illness is most frequently associated with infectious agents. However, it can also be caused by a variety of non-infectious conditions. Inflammatory bowel diseases, including ulcerative colitis and Crohn's disease, result in inflammatory diarrhea. Thyroid and other metabolic disorders often cause a variety of gastrointestinal symptoms. Medications are frequently associated with accompanying gastrointestinal symptoms, particularly noninflammatory diarrhea. Medications which may cause these symptoms include antibiotics, some cardiovascular medications, chemotherapy, and antimetabolites. Other conditions which may lead to common gastrointestinal symptoms include irritable bowel syndrome (IBS), malabsorption syndromes, allergies, and depression (Braunwald et al. 2003). Surveillance data regarding non-infectious gastrointestinal illness is not readily available. Therefore, the incidence of such disease is more difficult to estimate than for gastrointestinal illness with infectious etiology.

Infectious

Person-to-Person. Infectious agents of gastrointestinal illness can be transmitted through person-to-person contact, either directly or through an intermediary object (fomite). Fomites refer to surfaces or objects that can become contaminated with pathogenic organisms and serve as vehicles for subsequent transmission. Fomites become contaminated through direct contact with bodily secretions from an infected host or contact with other contaminated surfaces, including hands (Boone and Gerba 2007). Enteric viruses can remain infectous on surfaces for two months or longer (Boone and Gerba 2007).

Foodborne. In 2002, 76 million cases of illness in the U.S. were attributed to foodborne transmission. These illnesses resulted in 325,000 hospitalizations and 5,200 deaths (Mead et al. 1999). Foodborne illness is caused by a variety of agents, including bacteria, viruses, protozoa, and toxins. Food can become contaminated with these disease agents during production, processing, storage, or preparation. The role of foodborne disease is often difficult to determine, in part because many of the organisms that cause foodborne illness can also be transmitted via water (Mead et al. 1999).

Waterborne. Randomized-controlled trial data suggest that between 40 and 50 million cases of waterborne disease occur each year in the U.S. (2000; Payment, Franco et al. 1991). Both recreational water and drinking water are associated with disease transmission, and this transmission occurs on both an epidemic and endemic level. The primary disease agents of waterborne disease are similar to the infectious organisms associated with foodborne illness. Waterborne gastrointestinal illness will be the focus of the remainder of this chapter.

Organisms Causing Waterborne Gastrointestinal Illness

The infectious agents that cause waterborne gastrointestinal illness fall into three main categories: protozoa, viruses, and bacteria. Each organism has unique characteristics which allow it to be transmitted via water, causing delivery of microbially safe drinking water to be a complex and frequently incomplete process. A description of the most common waterborne infectious organisms highlights these characteristics.

Protozoa

Protozoa are the leading known cause of waterborne disease outbreaks in the U.S. since 1981 (Moe 2002). Their life cycle contains at least one stage in which the organism is environmentally stable, allowing for prolonged survival outside of host organisms. Protozoa are not, however, able to multiply in the environment outside the host. Protozoa vary in size, but typically measure approximately 10^{-6} meters (one micrometer) (Marshall et al. 1997). The cysts or oocysts of two common protozoa, *Cryptosporidium parvum* and *Giardia lamblia*, are widely distributed in surface waters from which drinking water is derived (Meinhardt, Casemore, and Miller 1996). One study found that 97 percent of raw water supplies contained cysts from at least one, and often both, of these organisms (LeChevallier, Norton, and Lee 1991). In addition, protozoa tend to be resistant to disinfection procedures, are able to infect humans with a relatively small infective dose, and are excreted by those who are infected for prolonged periods of time (Marshall et al. 1997; Meinhardt, Casemore, and Miller 1996; Moe 2002). These characteristics combine to pose a substantial risk for gastrointestinal illness through drinking and recreational water. Both of these organisms are zoonoses, meaning that they can be transmitted from animals to humans (Smith and Nichols 2006). Other notable protozoan organisms which cause waterborne illness are *Entamoeba histolytica,* Microsporidia, and *Cyclospora cayetanensis*.

Cryptosporidium parvum. Cryptosporidiosis has received increasing attention as a waterborne pathogen following the massive waterborne outbreak involving this organism in Milwaukee in 1993 (MacKenzie et al. 1994). The first documented human case of cryptosporidiosis, caused by the *C. parvum* organism, occurred in 1976 (Nime et

al. 1976). Illness occurs after ingestion of mature *C. parvum* oocysts, typically via food or water. After excystation of the oocysts in the small intestine, sporozites are released. These sporozites then parasitize the epithelial cells of the gastrointestinal tract. The incubation period for cryptosporidiosis is 5 to 28 days, with a mean of 7.2 days. Symptoms of cryptosporidiosis include diarrhea, abdominal pain, nausea, fever, and fatigue (Marshall et al. 1997). The disease is of particular concern for those with immune suppression, in whom the disease can be life-threatening (MacKenzie et al. 1995). Cryptosporidiosis has demonstrated a summer peak in incidence (Meinhardt, Casemore, and Miller 1996). While cryptosporidiosis is less commonly diagnosed than *G. lamblia*, it is responsible for more serious disease in terms of hospitalization and death (Mead et al. 1999). Since 1985, *C. parvum* has been the cause of more than 14 reported waterborne outbreaks of gastrointestinal illness (Gofti-Laroche et al. 2003).

Giardia lamblia. *G. lamblia* is the most frequently diagnosed intestinal parasite in the U.S. (Mead et al. 1999; Furness, Beach, and Roberts 2000). Although it is often asymptomatic, there are 20,000 to 28,000 giardiasis cases reported annually in the U.S. (Furness, Beach, and Roberts 2000). Reported cases reflect only a small portion of total cases, however, and it is estimated that between 2 and 2.5 million cases of giardiasis occur each year (Mead et al. 1999). Giardiasis, caused by infection with *G. lamblia*, was first reported in 1859, but there is evidence of its description as early as 1681. *G. lamblia* has a two-part life cycle, including an actively multiplying trophozite and an environmentally resistant cyst. Infection occurs when the cyst is ingested, typically via water, and excystation then occurs in the duodenum, releasing the trophozites. The trophozites attach to the duodenum or the proximal jejunal mucosa, where they replicate

asexually. The trophozites continue through the intestinal tract and, as they move through the colon, cyst formation occurs. The cysts are then shed by infected individuals. The incubation period for giardiasis is one to two weeks. The most notable symptom of the disease is explosive, watery, foul-smelling diarrhea. Other symptoms include nausea, anorexia, and abdominal gurgling (Marshall et al. 1997). There is a higher incidence of giardiasis in the summer months (Furness, Beach, and Roberts 2000).

Viruses

Viruses are small compared to other waterborne disease organisms, typically measuring approximately 20 - 100 nanometers (10^{-9} meters). Although viruses tend to be more susceptible to disinfection than protozoa, their size limits the ability of filtration techniques to remove the organisms from raw water. Successful disinfection depends upon many water treatment variables and viral characteristics. Viruses that are pathogenic in humans have no environmental source other than humans, and do not multiply outside of a human host. The infectious dose for viruses is low. Most gastrointestinal viruses have a short 12 to 48 hour incubation period, and disease duration is typically two to three days. In contrast to the symptoms of bacterial and protozoan pathogens, viruses tend to cause vomiting more often than diarrhea. The incidence of viral gastrointestinal illness is not easily quantified because of the difficulty in diagnosing viral infections (U.S. Centers for Disease Control and Prevention; Wheeler et al. 1999). However, while 75 percent of hospitalizations for gastrointestinal illness have no causative agent identified, among those cases for which an etiology is determined, 80 percent are caused by viruses (Jin et al. 1996). There are many viral agents that are responsible for waterborne disease, including hepatitis E virus, coxsackievirus A and B, adenovirus, and

astrovirus. The most prominent waterborne viral agents are norovirus, rotavirus, and Hepatitis A virus.

Norovirus. Noroviruses are in the family *Caliciriridae* and are the most common cause of non-bacterial acute gastrointestinal illness (Mead et al. 1999; Chin 2000). It is estimated that noroviruses cause up to 23 million cases of illness per year in the U.S. (Mead et al. 1999). Noroviruses are transmitted through water and food, particularly shellfish. They have a very low infective dose, and are frequently associated with outbreaks on college campuses, cruise ships, restaurants, child care centers, and nursing homes. The illness caused by the virus is usually self-limited, and all age groups are affected (Chin 2000).

Group A Rotavirus. Rotavirus is associated with approximately one-third of hospitalized cases of diarrhea in children less than five years of age, and it is believed that most children become infected with the virus by the age of three years (Chin 2000). Among U.S. children one month to four years of age hospitalized for gastrointestinal illness, rotavirus is the most frequently diagnosed pathogen (17 percent of all cases) (Parashar et al. 1998). Reporting of rotavirus is more complete than for many other gastrointestinal viruses for a number of reasons. Rotavirus infection is easily diagnosed using standard commercial assays (Glass et al. 2000). Further, the symptoms of rotavirus infection are more serious compared to many other viral agents, and because the disease occurs most often in younger children, caregivers may be more likely to seek medical care. Infections that occur in adults are more likely to be subclinical. Clinical disease in adults is seen primarily among those with immune suppression. In the U.S., the disease displays a marked seasonality, with peaks in incidence beginning in the cooler months in

the Southwest and moving across the country to the Northeast (Chin 2000; Ing et al. 1992).

Hepatitis A Virus (HAV). Serologic evidence suggests that 33 percent of the general population in the U.S. has been infected with HAV at one time. The virus is most often spread through contaminated water or food, the latter due primarily to handling by an infected food worker. The incubation period for HAV is long, with an average of 28 to 30 days. Infection with HAV typically leads to more serious illness in adults than in children, but the overall case-fatality rate is low. A vaccine for HAV is available (Chin 2000).

Bacteria

Most enteric bacteria are able to infect both humans and animals and they are able to multiply in the environment under favorable conditions. The infectious dose of bacteria required to cause human disease is generally larger than that required by protozoa or viruses (Moe 2002). There are many bacterial organisms associated with waterborne illness, including *Yersinia enterocolitica*, *Aeromonas* sp., *Helicobacter pylori*, and *Mycobacterium avium* complex. The organism which causes cholera, *Vibrio cholerae*, was once a major source of serious waterborne illness in the U.S., and it continues to be of substantial concern in many parts of the world. The waterborne bacterial pathogens of primary concern in the U.S. are *Shigella*, *Salmonella*, *Campylobacter*, and pathogenic *Escherichia coli*.

Shigella. Four different serotypes of *Shigella* organism cause shigellosis, also known as bacillary dysentery. *Shigella dysenteriae* 1 is associated with serious disease and reported fatality rates as high as 20 percent in hospitalized cases. *Shigella sonnei*, by

contrast, has a short clinical phase and almost no associated mortality, even among patients with immune suppression. *Shigella flexneri* and *Shigella boydii* can also cause disease. Humans are the only reservoir for these organisms. The minimum infectious dose for shigellosis is low, approximately 10 to 100 bacteria, and the incubation period is one to three days for most serotypes. Shigellosis typically affects the large and distal small intestines. Watery diarrhea, fever, nausea, bloody stool, and mucous stool are the most common symptoms of shigellosis, although asymptomatic infections often occur (Chin 2000). It has been estimated that 35,000 waterborne cases of shigellosis occur each year in the U.S. (Morris and Levin 1995).

Salmonella. There are many pathogenic strains of *Salmonella*, and the organism can be pathogenic in both animals and humans. Humans can become chronic carriers of the organism, but more commonly birds and other animals serve as the reservoir for this bacterium. *Salmonella* is primarily transmitted through food, but some outbreaks have been attributed to drinking water. The incubation period varies by subtype, but is typically 6 to 72 hours. Symptoms include headache, abdominal pain, diarrhea, nausea, and vomiting, and may last several days. For most serotypes, death is uncommon except for those in the very young and very old age groups and those with compromised immunity. An estimated 59,000 waterborne cases of salmonellosis occur in the U.S. each year (Morris and Levin 1995). *Salmonella typhi* and *Salmonella enterica* are typically categorized separately from the other serotypes of *Salmonella* because they cause typhoid fever and paratyphoid fever, respectively, which are more serious systemic diseases. Typhoid and paratyphoid are rare in the U.S., with less than 100 cases diagnosed a year, the majority of which are imported (Chin 2000).

Campylobacter. *Campylobacter* are estimated to cause 5 to 14 percent of diarrheal disease worldwide. The most prevalent serotype is *C. jejuni* and it has been found that the majority of raw poultry meat in the U.S. is contaminated with *C. jejuni* (Altekruse et al. 1999). *C. jejuni* can also be transmitted through water, as evidenced by its role as the causative agent in a few outbreaks of waterborne disease (Lee et al. 2002). The incubation period for this organism is two to five days, and symptoms include diarrhea, abdominal pain, malaise, fever, nausea, and vomiting, typically lasting two to five days. Children under five and young adults have the highest incidence of disease in developed countries. While outbreaks due to *C. jejuni* tend to occur in spring and fall, sporadic cases are more common in warmer months (Chin 2000). It is estimated that 320,000 cases of *Campylobacter* infection occur annually in the U.S. (Morris and Levin 1995).

Escherichia coli. Pathogenic *E. coli* have traditionally been considered agents of foodborne illness. However, recent outbreaks have demonstrated their capacity as an agent of waterborne illness as well. There are six major categories of pathogenic *E. coli*: enterohemorrhagic, enterotoxigenic, enteroinvasive, enteropathogenic, enteroaggregative, and diffuse adherent. The enterotoxigenic, enteroinvasive, enteropathogenic, and enteroaggregative categories are more common in developing countries, although the enteropathogenic strain was formerly common in the U.S. The diffuse adherent category is more recently recognized, and is not well-characterized. The category of most concern in the U.S. is the enterohemorrhagic *E. coli*, of which *E. coli* O157:H7 is a member. Cattle are the primary reservoir for this organism. The infectious dose for enterohemorrhagic *E. coli* relatively low, and the incubation period for disease is fairly

long, two to eight days. These bacteria contain a virulence gene that is involved in bacterial attachment to the intestinal mucosa. *E. coli* O157:H7 is of concern because of two particularly serious sequelae, hemolytic uremic syndrome (HUS) and thrombotic thrombocytopenic purpura (TTP) (Chin 2000). Each year in the U.S. an estimated 150,000 cases of *E. coli* are acquired through waterborne transmission (Morris and Levin 1995).

Susceptible Subpopulations

Gastrointestinal illness is usually self-limited, causing only temporary and sometimes mild morbidity. However, certain groups of individuals with reduced immune function are at higher risk of serious morbidity and mortality from gastrointestinal infections. These groups include the very young, the elderly, pregnant women, and those who have compromised immune systems. While any one of these groups may represent a relatively small segment of the population, together they comprise approximately 20 percent of the U.S. population. This proportion is likely to increase with the aging of the population and the increasing number of people living with compromised immune systems (Gerba, Rose, and Haas 1996).

Children Five Years Old and Younger

In 1991, the 16.5 million children age five and younger in the U.S. had between 21 and 37 million cases of diarrheal illness. These cases resulted in an estimated 2.1 to 3.7 million physician visits and 200,000 hospitalizations (Glass et al. 1991). It has been estimated that between 11 and 14 percent of all hospitalizations in children age 1 month to 4 years are for gastrointestinal symptoms (Glass et al. 1991; Jin et al. 1996; Parashar et

al. 1998). In the U.S., there are approximately 400 deaths each year in this age group due to diarrheal illness, accounting for ten percent of postneonatal preventable mortality (Glass et al. 1991). The mortality rates for gastrointestinal illness in children are highest in the southern states and among African-Americans, where the rates are 50 and 250 percent higher, respectively, than the average overall rate for children. Prematurity is a common co-morbidity among those children who die from gastrointestinal illness (Kilgore et al. 1995).

Elderly

The elderly are at even greater risk for serious complications of gastrointestinal illness than the young. The rate of hospitalization for gastrointestinal illness is almost three times higher among those 75 and older than for younger adults (age 20 to 74 years). Among adults over 20 years of age hospitalized for gastrointestinal illness, 53 percent of deaths are among those over 75 years of age. Given hospitalization for gastrointestinal illness, those over 75 are 33 times more likely to die than younger adults (Mounts et al. 1999).

Pregnant Women

Although no U.S. cases have yet been identified, several waterborne hepatitis E outbreaks in other countries have made clear the organisms disproportionate impact on pregnant women. Case fatality rates are between 10 and 20 percent among pregnant women, while these rates are 1 to 2 percent in the general population. Pregnant women also have the potential to pass disease on to the fetus (Gerba, Rose, and Haas 1996). A separate concern has been raised regarding chemical exposures in drinking water, including disinfection by-products (DBPs), because of evidence that such exposure may

be related to spontaneous abortion and low birth weight (Dodds et al. 2004; Office of Water 2000; Waller et al. 1998).

Immunocompromised

People with HIV and AIDS, patients receiving chemotherapy for cancer treatment, or patients receiving immune suppression therapy following organ or tissue transplantation are among those at higher risk for complications of infectious disease, including waterborne gastrointestinal illness (Gerba, Rose, and Haas 1996). A great deal of research has been conducted on the increased susceptibility to cryptosporidiosis experienced by people with HIV and AIDS, and cryptosproridiosis is considered an AIDS-defining condition (Castro et al. 1992). It is estimated that those with AIDS have twice the rate of cryptosporidiosis compared to those without AIDS (Perz, Ennever, and Le Blancq 1998). AIDS patients with CD4 counts less than 50 per microliter are at especially increased risk of mortality within one year of infection (Vakil et al. 1996). The median survival after infection for such a patient is 11 months (Sorvillo et al. 1998).

Routes of Waterborne Disease Transmission

Waterborne pathogens are transmitted to humans through two primary routes: recreational water exposure and drinking water exposure. These exposure routes are unique in their risk factors for transmitting pathogens. Recreational waters receive less treatment to remove pathogens than drinking water, but drinking water is ingested and therefore even small concentrations of pathogenic organisms have the potential to cause disease through this route.

Recreational Water

Recreational water, including lakes, rivers, swimming pools, and water parks, can become contaminated in a variety of ways. Natural bodies of water can be contaminated by sewage treatment malfunction or lack of sewage treatment, urban runoff, storm waters, faulty septic systems, and agricultural runoff (U.S. Environmental Protection Agency 2000). Contamination of treated recreational water (e.g. swimming pools) occurs primarily through fecal contaminants rinsing off of swimmer's bodies or fecal accidents. In either setting, the unintentional ingestion of contaminated water can result in gastrointestinal illness. Infants and children pose a particular risk to recreational water because they are more prone to fecal accidents and frequently enter the water wearing non-leakproof diapers (Lee et al. 2002). Microbial standards for recreational water are determined by state and local authorities. The U.S. Environmental Protection Agency (EPA) issues federal guidelines for these standards. Violations can lead to restricted activity or closure of a body of water or swimming facility (Lee et al. 2002; U.S. Environmental Protection Agency 2000). Determination of the standards for recreational water quality is controversial because frequently outbreaks are not associated with any recognized deterioration in water quality (Favero 1985).

Both outbreaks and sporadic cases of gastrointestinal illness are related to recreational water exposure. The number of reported gastrointestinal illness outbreaks associated with recreational water has steadily increased since the 1980s. Between 1993 and 2004, 281 outbreaks of recreational water-related gastrointestinal illness were reported to the U.S. Centers for Disease Control and Prevention (CDC), with a total of 20,298 associated cases (Barwick et al. 2000; Dziuban et al. 2006; Kramer et al. 1996;

Lee et al. 2002; Levy et al. 1998; Moore et al. 1993; Yoder et al. 2004). The majority of outbreaks occurred in treated water. The majority of these outbreaks were due to *Cryptosporidium parvum* infection. Chlorination, the primary method of recreational water treatment, is not as effective at destroying protozoan organisms as it is for other pathogenic organisms. Among those organisms that are more susceptible to chlorine, chlorination can be rendered less effective by heavy swimmer load, high levels of organic matter in the water, and UV light. The 14 freshwater recreational outbreaks that occurred during this time period were primarily due to bacterial organisms or an unidentified pathogen (Lee et al. 2002). Sporadic cases of gastrointestinal illness are not as easily linked to recreational water exposure as cases which occur during outbreaks. However, in non-epidemic time periods the risk of developing gastrointestinal illness has been estimated to be 1.4 to four times higher among those with recreational water exposure compared to those without such exposure (Colford et al. 2007; Seyfried et al. 1985; Wade et al. 2006).

Drinking Water

Pollution sources that affect natural recreational water also impact drinking water sources. The raw water from which drinking water is produced is often drawn from the same lakes and rivers used for recreation. Recreational activities themselves can contribute to pollution of these waters through the deposition of fecal material and the direct dumping of garbage. Drinking water is also susceptible to contamination throughout the treatment process, while in the distribution system (the pipes that carry water from the treatment plant to end-users), and upon reaching the end-user. The EPA, through the Safe Drinking Water Act and its amendments, regulates drinking water

quality up to the point of the end-user (U.S. Environmental Protection Agency 1999). Violations of these regulations can lead to fines and, in the worst instances, boil-water advisories for the customers of the violating water system. Despite regulations, drinking water has been associated with both endemic and epidemic gastrointestinal disease. There were 165 outbreaks of disease associated with drinking water between 1993 and 2004, resulting 415,855 cases of illness (Barwick et al. 2000; Blackburn et al. 2004; Kramer et al. 1996; Lee et al. 2002; Levy et al. 1998; Liang et al. 2006; Moore et al. 1993). These associations of drinking water exposure and GI illness will be discussed in greater detail in subsequent chapters.

CHAPTER 3

LITERATURE REVIEW: MICROBIAL DRINKING WATER QUALITY

Drinking Water Treatment and Distribution

Source Water

Drinking water in the U.S. originates from a variety of freshwater sources. The vast majority of these sources are classified as groundwater or surface water (Moeller 1992). While approximately 80 percent of community water systems are supplied by groundwater, a greater proportion of people are served by surface water (70 percent) due to the fact that the communities served by surface water are generally much larger than those served by groundwater (U.S. Environmental Protection Agency 2003). The differences between these two primary sources result in important distinctions in the treatment processes required to reduce microbial risk, the effectiveness of those processes, and the risk of transmission of waterborne disease.

Groundwater. Groundwater is generally found in the top one-half mile of the Earth's surface. There are an estimated 50,000 square miles of groundwater at any given time throughout the mainland U.S. (Council on Environmental Quality 1989; Moeller 1992). Groundwater can be an inexpensive source of drinking water when it is available at the point of use. Groundwater is obtained through wells or springs, and because it is presumed to be free of pathogens, treatment is often absent or abbreviated (Roberson 2003). However, there are many indications that groundwater is not as pure as was previously supposed (Moeller 1992). Pathogens and other contaminants can be introduced to groundwater through seepage of contaminated precipitation or surface

water, from improper disposal of waste materials through deep ground injection practices, and from septic tanks. Evidence of contamination is provided by the increasing proportion of drinking water-related outbreaks attributed to groundwater sources. Between 1993 and 2004, 72 percent of all drinking water-related outbreaks were linked to groundwater sources (Liang et al. 2006; Blackburn et al. 2004; Kramer et al. 1996; Levy et al. 1998; Barwick et al. 2000; Lee et al. 2002). The U.S. Environmental Protection Agency (EPA) recently established more stringent water treatment requirements for drinking water utilities using groundwater as their source (Lee et al. 2002; U.S. Environmental Protection Agency 2006). Another drawback of groundwater sources is that, once depleted, they are largely irreplaceable. Although 10 percent of precipitation that reaches land is eventually reabsorbed and serves to partially replenish groundwater sources, the demand for groundwater is rapidly depleting the supply at a faster rate than it can be replaced (Moeller 1992).

Surface Water. Approximately 20 percent of precipitation reaching land becomes surface water, filling lakes, streams, and rivers (Moeller 1992). Surface water is subject to many sources of microbial pollution, primarily through fecal contamination (Barrell, Hunter, and Nichols 2000; Craun and Calderon 2001). This contamination can originate from animal sources through agricultural run-off or from human sources through sewage discharge and municipal waste run-off (Ford 1999; Payment, Gamache, and Paquette 1988). Rivers are a particularly problematic source of drinking water because of the dynamic conditions they experience (Juranek and Mac Kenzie 1998). Rainfall contributes heavily to elevated microbial concentrations in surface waters, particularly rivers, through increased flow and washing of animal and human fecal matter

into the water. Specifically, *Giardia* cysts and *Cryptosporidium* oocysts are found at increased levels following rainfall (Meinhardt, Casemore, and Miller 1996). Rainfall also increases the overall particle content of the water, which can lead to subsequent stress and potential failure in the treatment system (Beaudeau et al. 1999; Atherholt et al. 1998). Periods of heavy rainfall following periods of drought pose a particular risk to surface water because an accumulation of particles and microorganisms will be deposited over a short time period, rather than being regularly deposited in smaller quantities (Atherton, Newman, and Casemore 1995). Approximately 70 percent of drinking water-related outbreaks have occurred following heavy rainfall events (Atherton, Newman, and Casemore 1995; Beaudeau et al. 1999; LeChevallier, Evans, and Seidler 1981; Smith et al. 1989; Curriero et al. 2001).

Some communities are able to partially protect their surface water supply by establishing protected watersheds, the water from which requires less treatment prior to distribution. However, not all watersheds can be adequately protected, and the surface water may still become contaminated such that it warrants full treatment (Moeller 1992). Securing an adequate supply is an issue for surface water, as well. A dramatic illustration of the high demand on surface water sources is the Colorado River, which now only rarely reaches the ocean because it is depleted prior to reaching the Colorado delta (Postel 2000). Despite the drawbacks of using surface water, it is the only source available to the majority of the U.S. population, and therefore the effectiveness of the treatment processes for this water source is of great importance (Moeller 1992).

Treatment Processes

There are many water treatment options available to drinking water utilities, but the processes utilized by the majority of systems are similar. Some utilities, because of the quality of their source water, are required to implement only a subset of the steps in the full treatment process. Although treatment is an important and necessary step in reducing microbial risk, no water treatment method is 100 percent effective, even during optimal operating conditions (Gale 1996).

The first step in drinking water treatment is pumping water from the source to a settling basin (Moeller 1992). The settling basin allows raw water turbidity, a measure indicating the concentration of particulate matter, to decrease naturally. This step is important because increased raw water turbidity impairs the ability of chlorine to disinfect efficiently (LeChevallier, Evans, and Seidler 1981). After settling, chemicals are added to the water which cause the formation of floc, collections of suspended matter bound by a coagulating agent. Alum ($Al_2(SO_4)_3$) or ferric chloride ($FeCl_3$) are the most commonly used coagulation chemicals, and they are typically added during rapid mixing of the water to promote the formation of floc. Flocculation is then further promoted by gentle mixing, which encourages the binding of floc particles to each other. These large floc particles are then allowed to settle at the bottom of the tank during the sedimentation step of the process. The settled floc is removed and the water proceeds to filtration (Moeller 1992).

The filtration step serves to remove any unsettled floc. There are multiple filtration methods, but rapid gravity filtration is most commonly used in the U.S. The filter consists of a two- to three-feet deep bed of media (anthracite coal and sand) through

which water is passed. Primary disinfection can occur before or after filtration. Disinfection is typically done through the use of chlorine, although there are other available disinfectants, including ozone, chloramines, and chlorine dioxide (Trussell 1999). Filtration is used in conjunction with disinfection because neither treatment is sufficient alone to fully treat contaminated surface water. Some pathogenic organisms are too small to be completely removed by filters, and others are resistant to chlorine and other disinfecting chemicals. Further, filters may clog and cease to effectively capture even larger microorganisms. Filters are regularly cleaned by backwashing to prevent clogging, but even this necessary maintenance has the potential to introduce contamination, particularly at the point when the filters are taken off-line or returned to service (Atherton, Newman, and Casemore 1995). Even with filtration and disinfection, some pathogenic organisms may survive water treatment. *Cryptosporidium* and *Giardia* have been found in treated water even from facilities with high removal efficiencies (LeChevallier, Norton, and Lee 1991). Treatment deficiencies were responsible for 293 reported drinking water-related outbreaks between 1971 and 2004, most notably in the massive 1993 Milwaukee cryptosporidiosis outbreak (Craun and Calderon 2001; Dykes et al. 1980; MacKenzie et al. 1994; Liang et al. 2006; Blackburn et al. 2004; Lee et al. 2002). Water utilities are also required to treat water such that it maintains a residual level of disinfectant throughout the distribution system (secondary disinfection) (Trussell 1999; U.S. Environmental Protection Agency 1989).

In addition to the filtration and disinfection treatment steps, measures can also be taken to improve the taste and odor, and minimize the corrosiveness of drinking water. Hardness, caused by dissolved calcium, can be removed by adding lime and sodium

carbonate to the water. Activated carbon (charcoal) can be added to the chemical coagulates or as a top layer of a filter to remove organic compounds (Craun and Calderon 2001; Moeller 1992; Council on Environmental Quality 1989). EPA issues guidelines and secondary regulations regarding water treatment practices that enhance the aesthetics of drinking water.

Distribution System

After treatment, water is pumped to the distribution system, the network of pipes through which water is delivered to end-users. In developed countries, the primary concern regarding drinking water safety in recent years is the degradation of water quality in the distribution system. Treated water can become contaminated in the distribution system through several mechanisms including biofilm growth and intrusion events.

Biofilms are composed of microorganisms, extracellular products, and inorganic and organic debris. Biofilms can form on the inner surface of distribution system pipes, particularly when the pipes are corroded, and can provide a protective environment from residual disinfectant for pathogenic microbes. Organic material in biofilms can react with chlorine and deplete the residual chlorine (Ford 1999; LeChevallier, Welch, and Smith 1996). Different organisms within biofilms may exchange genetic information and thereby adopt resistance or virulence factors. Biofilms can corrode pipes, leading to leaks or breaks, which can subsequently introduce additional contaminants to treated drinking water as it moves through the distribution system (Ford 1999). Warmer water temperatures encourage biofilm growth (LeChevallier, Welch, and Smith 1996). Regular pipe flushing, achieved through opening fire hydrants and other similar methods, can

reduce biofilm build-up, although disturbing biofilms can release pathogenic organisms and flushing itself may allow contaminants to enter the distribution system (LeChevallier, Welch, and Smith 1996). Although biofilms have the potential to contribute to deterioration of water quality, results from one randomized controlled trial suggested that biofilms may actually play a beneficial role in inactivating pathogens (Payment et al. 1997).

Contaminants can enter the distribution system through cross-connections, backflow, main breaks, or leaks. Cross-connections, through which sewage lines become connected with drinking water lines, pose a disturbing risk for contamination, but such occurrences are increasingly rare due to cross-connection control programs (Karim, Abbaszadegan, and LeChevallier 2003). Backflow of water from service connections, particularly from hazardous locations such as hospitals, dry cleaners, or mortuaries, poses a contamination risk, but backflow prevention devices can be installed to prevent this occurrence. However, only rarely do all service connections have such devices (Karim, Abbaszadegan, and LeChevallier 2003).

Main breaks pose a risk both for direct contamination and for subsequent contamination through the creation of pressure transients (Karim, Abbaszadegan, and LeChevallier 2003; LeChevallier 1999; LeChevallier et al. 2003). Constant high positive pressure is necessary to prevent pathogen intrusion through leaks in the distribution system pipes. Even among well-maintained systems, approximately ten percent of total drinking water production can be lost due to leaks (Kirmeyer et al. 2001). If a pressure transient, which can also be caused by sudden changes in demand, opening or closing of fire hydrants, and power outages, results in decreased positive pressure or negative

pressure, moisture surrounding the pipes can be drawn into the pipes (Karim, Abbaszadegan, and LeChevallier 2003; LeChevallier 1999; LeChevallier et al. 2003). Fecal contamination and culturable viruses have been detected in the ground surrounding distribution system pipes (LeChevallier et al. 2003). The level of soil contamination will largely depend on the proximity to sewer lines. Although there are regulations in place regarding allowable distances between sewer and drinking water pipes, pathogenic organisms are able to migrate over these distances, particularly with accompanying precipitation (Abu-Ashour et al. 1994; LeChevallier et al. 2003). Because these contamination events are likely to be small and localized, it is unlikely that they will be detected through routine surveillance (Karim, Abbaszadegan, and LeChevallier 2003).

Outbreaks attributable to distribution system deficiencies highlight the disease risks involved with this type of contamination. Between 1971 and 1998, 113 drinking waterborne disease outbreaks were attributed to deficiencies in the distribution system, resulting in 488 hospital visits and 13 deaths. Although most of these outbreaks had no etiologic agent identified, those organisms that were identified tended to be more chlorine-resistant and included *Giardia*, Norwalk-like viruses, *E. coli* O157:H7, *Campylobacter*, *Shigella*, and *Salmonella* (Craun and Calderon 2001). From 1999 through 2004, half (11/22) of all reported outbreaks in community water systems were due to distribution system deficiencies, resulting in 2002 additional cases (Liang et al. 2006; Blackburn et al. 2004; Lee et al. 2002).

Disinfection Issues

Disinfection is controversial because of the chemicals used and, in the case of secondary disinfection, because of uncertainty as to the effectiveness of the procedure in

reducing microbial risk. All of the chemicals used in disinfection produce by-products, with largely unknown associated risks (LeChevallier 1999; Trussell 1999). The by-products of chlorine are of particular concern because of the chemical's widespread use and epidemiologic evidence of potential links between exposure to chlorinated drinking water and cancer, as well as adverse pregnancy outcomes (Miller et al. 2004; Trussell 1999; LeChevallier 1999; Waller et al. 1998; Dodds et al. 2004; Cantor et al. 1987). Although declines in chlorine residual have been linked with adverse health outcomes, there is evidence that typical residual chlorine levels (0.2 to 2 mg/L) do not inactivate some bacteria and viruses, and chlorine at such low concentrations is completely ineffective in inactivating protozoa (Meinhardt, Casemore, and Miller 1996; Ford 1999; Egorov et al. 2002). Further, residual chlorine will inactivate coliforms accompanying a contamination event, which could mask distribution system contamination by removing this indicator (Payment 1999).

There are disinfection alternatives to chlorine, each with advantages and drawbacks. The use of chloramine as a disinfectant raises less concern about disinfection by-products (DBPs) and works well at controlling biological growth (biofilms) in distribution systems, but it does not work well as a primary disinfectant and it is not effective at inactivating viruses intruding into the distribution system (LeChevallier 1999; LeChevallier, Evans, and Seidler 1981; LeChevallier, Welch, and Smith 1996; Trussell 1999). Although chlorine dioxide has demonstrated some ability to inactivate the normally chlorine-resistant *Cryptosporidium* oocysts, high concentrations are required and concern over the toxic chloride ion produced as a by-product preclude its widespread use (Meinhardt, Casemore, and Miller 1996; Trussell 1999; Aggazzotti et al. 2003).

Oocysts have also shown susceptibility to high concentrations of ozone, and ozone is effective at treating trace levels of pesticides, herbicides, and other synthetic compounds (Chen et al. 1998; LeChevallier 1999; Meinhardt, Casemore, and Miller 1996). Additionally, DBPs are a much smaller concern with the use of ozone. However, ozone is expensive and is not able to provide a disinfectant residual (Moeller 1992). Ultraviolet (UV) light can be used as a means of primary disinfection, but it also produces no residual protection in the distribution system (LeChevallier 1999).

Indicators of Water Quality

The presence of pathogenic organisms in water cannot practically be measured by water treatment facilities on a continuous basis. The diversity of such pathogens is great, and the tests which indicate their presence tend to be complex, time-consuming, and expensive, if available at all. Further, such organisms are present infrequently and usually in very low concentrations when present (Christian and Pipes 1983). Therefore, water utilities utilize indicator organisms to determine both the quality of the source water they treat and the quality of the finished water they distribute. These indicators primarily serve to indicate fecal contamination.

There are several characteristics of an ideal indicator of microbial water contamination. It should be applicable to all types of water, including rivers, lakes, groundwater, and treated water. It should be present and absent corresponding to the pathogen's presence and absence. The concentration of the indicator should be correlated with the concentration of fecal contamination in the water. An ideal indicator organism is part of the intestinal flora of healthy humans, and thus not dependent on transmission

among the population. Most human pathogens are not able to multiply in the environment, and therefore an ideal indicator organism should similarly not multiply outside the human body (LeChevallier, Welch, and Smith 1996). Finally, in order to be applicable in water treatment facilities, an ideal indicator should be easy to detect accurately with simple, rapid, cost-effective tests (Ashbolt, Grabow, and Snozzi 2001).

Unfortunately there are no ideal indicators currently available to water utilities. Therefore, a variety of indicators are utilized which taken together are hoped to provide an indication of the amount of pathogen contamination present in the tested water. Some of these indicators are required to be measured by EPA. Others are available to provide additional information to water utilities regarding the microbial safety of their water.

Turbidity

Turbidity is a measure of the "cloudiness" of water. Stated more technically, it is an expression of the degree to which light is scattered by particulate matter present in the water (LeChevallier, Evans, and Seidler 1981). Particulate matter in water is often composed of clay, silt, organic matter, and microorganisms (Beaudeau 2003). Turbidity can be measured continuously, and measurement is rapid and inexpensive (LeChevallier, Evans, and Seidler 1981). Turbidity is measured using a metric called nephalometric turbidity units (NTU) and is measured with an instrument called a turbidimeter. The EPA currently requires that average daily turbidity levels must not exceed 0.3 NTU in 90 percent of daily samples taken within a given month and a level of 1 NTU must not be exceeded in any sample (U.S. Environmental Protection Agency 2001).

Turbidity removal and pathogen removal are correlated (Hendricks et al. 1998; LeChevallier and Norton; LeChevallier, Norton, and Lee 1991; Nieminski 1992; U.S.

Environmental Protection Agency 1999) and increased turbidity has been associated with outbreaks of gastrointestinal illness (MacKenzie et al. 1994) and endemic gastrointestinal illness (Aramini et al. 2000; Morris, Naumova, and Griffiths 1998; Schwartz, Levin, and Goldstein 2000; Schwartz, Levin, and Hodge 1997). However, there is little support for a direct physical relationship between turbidity and microbial contamination. Turbidity in treated water primarily serves to indicate an increased likelihood of water treatment failure. Increased organic material in source water may serve to increase populations of microbial pathogens in the water, and turbidity can be a measure of this potential. Elevated turbidity levels also indicate an increased burden on filtration and disinfection processes (Beaudeau 2003). The presence of particulate matter that causes water to be turbid makes disinfection less effective as well as making the measurement of coliforms, another water quality indicator, more difficult (LeChevallier, Evans, and Seidler 1981). The variety of water sources from which drinking water is produced result in varying background turbidity levels. It has therefore been suggested that changes in turbidity are more important for indicating contamination than the actual value of the measurement (Beaudeau et al. 1999; Beaudeau 2003). A related measurement, particle counts, taken using a particle size analyzer, indicates both the presence and size of particulate matter in water. These measurements are increasingly being used to augment turbidity measurements, but in practice the analyzers frequently malfunction (personal communication, Neal Spivey, Gwinnett County Department of Public Utilities).

Disinfectant Residual

A disinfectant residual, typically chlorine, is included with drinking water as it enters the distribution system in order to protect the microbial integrity of the treated

water until it reaches the end-user. The EPA, as part of the Surface Water Treatment Rule, requires that a defined number of samples from the distribution system be tested for their level of disinfectant residual (U.S. Environmental Protection Agency 1989). The number of tests required varies based on system size, with the largest systems being required to test more frequently and at more locations. A lack of detectable disinfectant residual indicates that water received down-line may not be protected from microbial organisms that can be introduced into the distribution system. A decreased or absent disinfectant residual also indicates that a contamination event may have occurred (LeChevallier 1999; Payment 1999; Trussell 1999). However, coliform bacteria can still be found in distribution systems that maintain a high disinfectant residual, and measurement of disinfectant residuals does not provide an indication of the levels of the many chlorine-resistant pathogenic organisms that may be present (Payment 1999; LeChevallier, Welch, and Smith 1996; Trussell 1999). Although disinfectant residuals are not direct indicators of contaminant levels, residual chlorine has been found to be inversely associated with gastrointestinal illness (Egorov et al. 2002).

Total Coliform Bacteria

Total coliform bacteria levels are measured as a more specific indicator of bacteriall contamination than turbidity. Coliform bacteria are so named because of their similarity in appearance to *E. coli*. Total coliforms are defined as gram-negative, non spore-forming, oxidase-negative, rod-shaped facultative anaerobic bacteria that ferment lactose to acid and gas within 24 to 48 hours at 36 degrees centigrade (Ashbolt, Grabow, and Snozzi 2001). Total coliforms include bacteria in the family *Enterobacteriaceae* and others, although the species of bacteria included in this group changes based on the media

on which it is cultured (Ashbolt, Grabow, and Snozzi 2001). The EPA requires that drinking water utilities test for coliforms in the distribution system. These tests typically determine presence versus absence of total coliforms. Although measurement of total colifoms to determine water quality dates back to the late 19th century, this measure is increasingly recognized as a problematic water quality indicator (Ashbolt, Grabow, and Snozzi 2001). Total coliforms are not limited to fecal bacteria, and there is environmental replication of many of the species included in this group (Edberg 1996; Ashbolt, Grabow, and Snozzi 2001). Further, total coliform levels are often not associated with the occurrence of waterborne disease (Hellard et al. 2001; MacKenzie et al. 1994; Payment, Franco, and Siemiatycki 1993). However, total coliform bacteria are easy to enumerate and are still recognized as a useful indicator of the efficacy of water treatment, particularly disinfection (LeChevallier, Welch, and Smith 1996; Ashbolt, Grabow, and Snozzi 2001; Edberg 1996).

Fecal Coliform Bacteria and E. coli

Fecal coliforms are a subgroup of total coliform bacteria which are thermotolerant. Fecal coliforms are intended to be more specific indicators of fecal contamination. However, even this group contains some bacteria which environmentally replicate and which are not specific to fecal contamination (Ashbolt, Grabow, and Snozzi 2001). Detection of fecal coliform bacteria involves more sophisticated culturing techniques and is not required by the EPA.

E. coli bacteria are included in the fecal coliform group, and are considered an even more specific indicator of fecal pollution than either total or fecal coliform groups. *E. coli* indicate not only the presence of fecal contamination, but the likely concentration

of other fecal bacteria (Ashbolt, Grabow, and Snozzi 2001). Testing for *E. coli* is not required by the EPA, but water utilities may test for its presence or for the more general fecal coliform group in order to better quantify microbial risk in their distribution system, particularly during instances of distribution system failures or specific health concerns.

Heterotrophic Plate Count

Heterotrophic organisms are those organisms that require organic carbon for growth. This group includes bacteria, yeasts, and mold. Heterotrophic bacteria are measured by plating water samples on culture media and counting the colonies which grow on the plate. There is no universally accepted definition of which bacteria are counted, and the exact bacterial species that are enumerated are defined by the media on which they are plated (Edberg 1996). Heterotrophic plate counts are related to the effectiveness of treatment methods, such as chlorination. It is not a measure used to directly quantify microbial risk of disease, because heterotrophic bacteria are not generally associated with levels of pathogenic organisms. However, elevated heterotrophic plate counts have been associated with drinking water-associated gastrointestinal illness (Payment, Franco et al. 1991; Payment, Franco, and Siemiatycki 1993). The EPA requires that heterotrophic plate counts be conducted if an absence of disinfectant residual is detected in the distribution system (U.S. Environmental Protection Agency 1989).

Bacteriophages

Bacteriophage are viruses that infect bacteria. They were initially used as an indicator of bacterial presence in water, but they are now considered more useful as models of human enteric viruses in water. Coliphage, bacteriophage that infect coliforms,

have specifically been identified for this modeling purpose (Ashbolt, Grabow, and Snozzi 2001). An indicator such as this is helpful due to the difficulty of isolating human viruses from water. Coliphage, like human viruses, are more persistent in water than coliforms (Havelaar, van Olphen, and Drost 1993). Also, coliphage, and human viruses, are more resistant to treatment processes than coliforms. Coliphage are still not ideal indicators of enteric virus behavior in water, however. Coliphage are excreted by animals and humans, while human viruses are specific to humans. Some coliphage have the ability to multiply in the environment, while enteric viruses are only able to replicate within the human host. Finally, humans excrete viruses only while infected, while coliphages can be persistently shed by humans. Coliphage alone are therefore not a direct measure of viral presence in water. The presence of coliphage may, however, indicate fecal contamination, which increases the likelihood of enteric viruses in the water (Ashbolt, Grabow, and Snozzi 2001). The EPA does not require that drinking water be tested for the presence of coliphage, or any other bacteriophage.

Fecal Streptococci

Fecal streptococci are gram-positive bacteria that are increasingly considered superior to fecal coliforms as indicators of fecal contamination in water (Barrell, Hunter, and Nichols 2000). One group of fecal streptococci, the enterococci, is considered a particularly good indicator of fecal contamination. The advantage of these bacteria, compared to coliforms, is that they are more likely to originate from fecal contamination and are able to persist in the environment without replication (Ashbolt, Grabow, and Snozzi 2001). The methods used to test for fecal streptococci are not currently developed

for widespread use by the water treatment industry, and there are currently no streptococci monitoring requirements designated by the EPA.

Pathogen Removal as Indicated through "Treatment Technique"

The emergence of protozoa as the predominant organism responsible for recent outbreaks of waterborne gastrointestinal disease raises concerns about the ability of current indicator organisms to signal the presence of this form of microbial contamination. Direct quantification of *Giardia* cysts and *Cryptosporidium* oocysts is time-consuming and difficult (Atherholt et al. 1998). Bacterial indicators have not demonstrated strong correlation with the presence of protozoa in water (Gofti-Laroche et al. 2003). Chlorine residual is not a good predictor of risk from protozoal contamination because these organisms are generally resistant to chlorine and other halogen disinfection agents (Juranek and Mac Kenzie 1998). Enteric viruses are also increasingly recognized for their difficult removal from drinking water. Some viruses are able to penetrate even very small filters and many are largely resistant to disinfection (LeChevallier 1999; Payment 1999; Payment, Franco, and Siemiatycki 1993). Therefore, the EPA regulates the levels of protozoa and viruses in drinking water through guidelines for disinfection and filtration techniques, deemed "treatment techniques" by EPA. Water utilities are required to remove 99 percent of *Cryptosporidium* oocysts, 99.9 percent of *Giardia* cysts, and 99.99 percent of viruses (U.S. Environmental Protection Agency 1998). Studies have shown, however, that even plants with high removal efficiencies can have protozoa present in finished water in sufficient quantities to cause illness (LeChevallier, Norton, and Lee 1991).

Clostridium perfringens

C. perfringens, a gram-positive bacteria, was originally considered an indicator of general fecal contamination, but its use as such an indicator received criticism because of its long persistence in the environment due to the formation of spores. However, this environmental persistence makes *C. perfringens* a good indicator organism for the presence of protozoa and some enteric viruses, that are also able to persist in the environment (Ashbolt, Grabow, and Snozzi 2001). There are currently no EPA regulations regarding *C. perfringens* testing in drinking water.

General Considerations for Measuring Water Quality

A general problem with measuring water quality is that contamination is unlikely to be evenly distributed throughout the water because contamination tends to take place through isolated incidents rather than as a continuous occurrence. There is evidence that spatial and temporal clustering of microbial organisms occurs (Christian and Pipes 1983; Gale 1996). Although water treatment decreases the overall number of pathogens in water, treatment may actually facilitate clustering among the organisms that are not removed through processes such as coagulation (Gale 1996). This phenomenon results in a greater proportion of water without any pathogenic organisms present, but higher concentrations of such organisms when contamination is present. Therefore, there is a low probability of detecting any pathogenic organisms in water samples, and although consumers are unlikely to ingest contaminated water, they will likely receive a larger dose of pathogen when contaminated water is consumed. Methods used to estimate the risk of gastrointestinal illness which do not account for the heterogeneous dispersal of organisms in water are likely to overestimate the risk of more infectious organisms and

underestimate the risk of less infectious organisms (Gale 1996). Water quality indicators are also susceptible to clustering, further complicating the interpretation of their measurement (Atherholt et al. 1998; Christian and Pipes 1983; Gale 1996).

In addition to promoting the clustering of organisms, water treatment processes have placed selective pressure on pathogenic organisms, promoting the development of a survival strategy referred to as the "viable but not culturable", or VNC, state. When an organism is in a VNC state they are not detectable by conventional testing methods and are less susceptible to treatment. Organisms cannot reproduce while in a VNC state, but they are ultimately still able to cause disease (Ford 1999; Ashbolt, Grabow, and Snozzi 2001). Another survival strategy that allows organisms to survive in drinking water is the use of protozoa as host organisms. The protozoa are less susceptible to treatment and less likely to be detected than free bacteria (Ford 1999).

Methods for more pathogen-specific, continuous monitoring of water quality are in development, including antibody tests, gene-sequence based methods, and biosensors (Ashbolt, Grabow, and Snozzi 2001). Until such methods are implemented by utilities on a wide-scale basis, proxy indicators of pathogenic organisms will continue as the standard for assessing the microbial quality of drinking water.

Drinking Water Regulations in the U.S.

The Safe Drinking Water Act

Until the latter part of the 20th century, regulations regarding drinking water quality in the U.S. were largely determined by state and local authorities, while the federal government provided only guidelines. The importance of maintaining a clean and

safe drinking water supply became more of a national priority in the late 1960s, largely due to concerns over chemical contamination (U.S. Environmental Protection Agency 1999). The Safe Drinking Water Act (SDWA) was enacted in 1974 to address this concern. The original SDWA created national interim standards for drinking water quality based on the 1963 Public Health Service guidelines. The SDWA also called for a comprehensive review of these regulations by the National Academy of Sciences. Between 1975 and 1979, interim standards were set for six synthetic organic chemicals, ten inorganic chemicals, turbidity, total coliform bacteria, radionucleotides, and total trihalomethanes (U.S. Environmental Protection Agency 1975, 1979). There have been several amendments to the SDWA, with the most significant occurring in 1986 and 1996.

The 1986 Amendments. The 1986 amendments set a 1989 deadline for the implementation of standards for 83 contaminants. Although this goal was not met, by 1992 standards had been issued for 76 contaminants. These standards are generally categorized into four basic rules: the Total Coliform Rule, the Surface Water Treatment Rule, the Chemical Rules, and the Lead and Copper Rule (U.S. Environmental Protection Agency 1999).

The Total Coliform Rule (TCR), enacted in 1989, requires that samples from drinking water distribution systems be tested for the presence of coliform bacteria on a monthly basis. The rule sets a maximum contaminant level (MCL) of less than five percent of monthly samples testing positive, and a maximum contaminant level goal (MCGL) of zero positive samples. A MCL is an enforceable standard that describes the highest level of a contaminant allowed by law. A MCGL is a non-enforceable guideline that describes the level of contaminant below which there is no known health risk.

MCLGs represent an ideal situation, while MCLs take available technology and monetary considerations into account (U.S. Environmental Protection Agency 1989).

The Surface Water Treatment Rule (SWTR), enacted in 1989, requires drinking water utilities that treat surface water supplies to remove or inactivate 99.9 percent of *Giardia* cysts and 99.99 percent of viruses through filtration and disinfection. The rule also requires a detectable disinfectant level to be maintained throughout the entire distribution system (U.S. Environmental Protection Agency 1989).

The Chemical Rules are a collection of standards that regulate both organic and inorganic chemicals that pose a health risk when ingested over long periods of time at levels consistently above the MCL. MCLGs are also set for these chemicals (U.S. Environmental Protection Agency 1999).

The Lead and Copper Rule (LCR), enacted in 1991, sets "action levels" for the levels of lead and copper in drinking water. An action level differs from a MCL in that the former serves as a trigger for requiring prevention or removal steps, rather than being a legal limit, as with MCLs. This rule requires that drinking water be sampled at the tap and tested for lead and copper levels. If more than ten percent of samples reach the action level, treatment steps must be initiated and the source water must be assessed. These chemicals receive particular attention because of the seriousness of both their short- and long-term health effects, including gastrointestinal symptoms, impairment of neurological and physical development in children, and damage to the kidney and liver (U.S. Environmental Protection Agency 1991).

The 1996 Amendments. The 1996 SDWA amendments focused on a broad range of issues and nearly doubled the text of the original act (Pontius 1997). In addition, the

complexity of the proposed standards increased substantially (Roberson 2003). These amendments included implementation of risk-based standard setting, provided greater flexibility to states to accommodate local circumstances, and increased public awareness regarding drinking water safety (U.S. Environmental Protection Agency 1999). Descriptions of the most important legislation arising from the amendments are described below.

In 1998, the Interim Enhanced Surface Water Treatment Rule (IESWTR) was enacted, and in 2001 it was revised to its final form (U.S. Environmental Protection Agency 1998, 2001). This rule, which applies only to water systems serving populations of 10,000 or more people, encompasses the provisions of the SWTR, and sets an additional MCL and removal requirement for *Cryptosporidium*. The rule also specifies more stringent turbidity standards and risk assessment through disinfection profiling and sanitary surveys. The Long Term 1 Enhanced Surface Water Treatment Rule (LT1ESWTR), enacted in 2002, extended the provisions of the IESWTR to water systems serving less than 10,000 people (U.S. Environmental Protection Agency 2002). The Long Term 2 Enhanced Surface Water Treatment Rule (LT2ESWTR) was finalized in 2006 (Pontius 2001, 2003; U.S. Environmental Protection Agency 2006). LT2ESWTR encompasses the previous ESWTRs, and additionally requires source water monitoring for *Cryptosporidium*, with extra treatment measures required for those water sources designated as high-risk. Disinfection profiling requirements are extended to ensure mutual compliance with concurrently released disinfectant by-product (DBP) standards.

Released in revised form in 2001, the Stage 1 Disinfectants and Disinfectant By-Products Rule (Stage 1 DBPR) established MCLs for disinfection by-products at

monitoring sites in the distribution system (U.S. Environmental Protection Agency 1998, 2001). The Stage 2 Disinfectants and Disinfectant By-products Rule (Stage 2 DBPR) was finalized concurrently with the LT2ESWTR (U.S. Environmental Protection Agency 2003; Pontius 2001; U.S. Environmental Protection Agency 2006). This rule extends the Stage 1 DBPR by specifying that areas at high-risk for DBP elevations be identified and specifically monitored.

The Information Collection Rule (ICR), enacted in 1996, stipulated a special 18-month monitoring period for large water systems serving more than 100,000 people (U.S. Environmental Protection Agency 1996). During this period, additional data were collected on microbial contaminants, with emphasis on *Cryptosporidium*, disinfection by-products, and the effectiveness of certain treatment technologies. These data were then applied to the formulation of the LT2SWTR. The Ground Water Rule (GWR), enacted in 2006, designates when ground water needs to be disinfected or otherwise treated using required sanitary surveys, hydrogeological assessments, and source water monitoring (U.S. Environmental Protection Agency 2000, 2006). The Filter Backwash Recycling Rule, enacted in 2001, sets forth provisions such that water systems do not compromise microbial control during backwash recycling (U.S. Environmental Protection Agency 2001).

Under the SDWA amendments, increased funding provisions were made for individual water systems for activities required under the amendments, such as source water assessments, to supplement state funding. The amendments also provided for programs in water system capacity development and outlined minimum guidelines for

training of operators of drinking water systems, for use by state water authorities (Office of Water 1999).

Several rules enacted under the 1996 SDWA amendments serve to increase public awareness regarding drinking water safety (Office of Water 1999). Consumer confidence reports are required to be sent by utilities annually to all customers detailing the source of their water supply and any regulatory infractions incurred in the past year, including the health consequences of the infraction (U.S. Environmental Protection Agency 1998). The amendments also improved public notification of water quality standards, provided for publicly-accessible databases regarding harmful contaminants and water quality, and provided for education about waterborne disease for healthcare providers and the public (Office of Water 1999).

Other Significant Water Legislation

There are several pieces of federal legislation, in addition to the SDWA, which significantly impact the management of the U.S. drinking water supply (U.S. Environmental Protection Agency 1999). The Clean Water Act of 1977 requires that states designate the uses allowed for a given body of water, such as for recreational activities or to supply public drinking water (1977). The states must then set standards that maintain water quality to safely support these designated uses. The Comprehensive Environmental Response, Compensation, and Liability Act protects source water from hazardous substances by controlling their release into the environment and providing emergency response in the case of accidental discharge (1980). The National Environmental Policy Act protects water sources from deleterious consequences of new development by requiring environmental impact assessments (1969). The Pollution

Prevention Act promotes pollution source reduction, including water contamination, over the less efficient processes of treatment and disposal (1990). The Toxic Substances Control Act calls for research regarding harmful chemicals, many of which have the potential to enter the water supply, and their effect on public health (1976).

CHAPTER 4

LITERATURE REVIEW: THE RELATIONSHIP BETWEEN DRINKING WATER AND GASTROINTESTINAL ILLNESS IN THE U.S.

Drinking Water-Related Disease Outbreaks

Compared to many other modes of spread of infectious disease, outbreaks of gastrointestinal (GI) illness related to drinking water can be fairly easy to recognize because drinking water systems often serve large populations. From 1993 to 2004, there were 165 outbreaks of GI illness due to drinking water exposure in the U.S. reported to the U.S. Centers for Disease Control and Prevention (CDC). These outbreaks resulted in over 415,855 cases of illness, over 4,500 hospitalizations, and at least 149 deaths. The disease agent was not identified in almost 30 percent of these outbreaks, and among the outbreaks of known etiology, bacteria, parasites, viruses, and chemicals were identified as the agents responsible (Liang et al. 2006; Blackburn et al. 2004; Kramer et al. 1996; Levy et al. 1998; Barwick et al. 2000; Lee et al. 2002). Outbreaks serve as a reminder of the disease risks from drinking water if it does not receive constant and adequate treatment.

The 1993 Milwaukee Outbreak

The largest waterborne outbreak in U.S. history occurred in Milwaukee from late March to early April of 1993. Using patient laboratory data and tests of ice frozen during the exposure period, *Cryptosporidium parvum* was identified as the etiologic agent. The source of the *Cryptosporidium* oocyst exposure was identified as public drinking water, specifically that distributed by one of the two drinking water treatment plants serving the

Milwaukee area. Although standard filtration and disinfection practices were followed by this plant, spikes in turbidity were noted prior to the onset of cases in the time period consistent with the incubation period for cryptosporidiosis. Analyses considering disease transmission models suggest that the outbreak was caused by a continual recontamination of drinking water by effluent containing pathogens shed by those who were ill (Eisenberg et al. 2005).

An estimated 403,000 people became infected with *C. parvum* during the outbreak, resulting in 4,400 hospitalizations, and over 100 deaths (Corso et al. 2003; Hoxie et al. 1997; MacKenzie et al. 1994). The elderly and those with AIDS experienced the highest rates of serious illness, and 85 percent of the outbreak-related mortality occurred among those with AIDS (Hoxie et al. 1997; Naumova et al. 2003). The total cost of the outbreak was estimated at $96.2 million, including medical costs and lost productivity (Corso et al. 2003). The magnitude of the outbreak brought increased attention to the issue of microbial risks from drinking water, particularly in regard to *C. parvum*. Many of the provisions of the 1996 amendments to the Safe Drinking Water Act were in response to the public concern prompted by this outbreak.

The 2000 Walkerton Outbreak

In May and June of 2000, a multi-bacterial drinking water outbreak occurred in Walkerton, Ontario, the first such outbreak recognized in Canada. *Campylobacter jejuni* and *E. coli* O157:H7 were identified as the primary etiologic agents responsible for the outbreak. The outbreak was recognized because of multiple pediatric cases of bloody diarrhea. Because the only common exposure among cases was living in or having visited Walkerton, drinking water was identified as the likely source. Walkerton's drinking water

was supplied by groundwater, and evidence of gross contamination was found in one of the wells. Although the well water was supposed to be chlorinated, there was evidence that the chlorination mechanism on the contaminated well was not functioning. Bacterial strains identical to the *C. jejuni* and *E. coli* O157:H7 strains identified from case patients were found in livestock feces from a farm adjacent to the contaminated well. Heavy rainfall in early May likely washed pathogens from the farm into the well (Bruce-Grey-Owen Sound Health Unit 2000).

Over 2,300 people were estimated to have become ill from this outbreak, resulting in 751 emergency department visits, 65 hospitalizations, and 7 deaths (O'Connor 2002). Although campylobacteriosis resulted in episodes of diarrhea and other GI symptoms, *E. coli* O157:H7 was responsible for more serious illness, as evidenced by the 27 cases who developed hemolytic uremic syndrome (HUS), 5 of whom died. Although the true distribution of infection was not known, among laboratory-confirmed cases there were more cases with *E. coli* infection only than *C. jejuni* infection only, and few patients were dually infected. This outbreak was the focus of enormous attention by the public and the government of Ontario, and numerous changes in the funding, governing structure, and provision of drinking water were made as a result (O'Connor 2002; Bruce-Grey-Owen Sound Health Unit 2000).

The 1987 Carrollton, Georgia Outbreak

The greater Atlanta-metro area (20-county) was the site of a large drinking water-related outbreak. In January and February of 1987, there was an outbreak of cryptosporidiosis in Carrollton, Georgia. An estimated 13,000 people became ill, 20 percent of the total population of Carrollton. *Cryptosporidium* oocysts were isolated from

samples of drinking water and confirmed that drinking water was the source of the outbreak. Infected cattle in the watershed and a sewage overflow were likely responsible for the contamination (Hayes et al. 1989). This outbreak was unique in that it was the first *Cryptosporidium* outbreak to be associated with drinking water that met all state and federal standards (Avery and Lemley 1996).

Randomized-Controlled Interventional Trials

The relationship between drinking water and outbreaks of GI illness is much easier to recognize than the role of drinking water in sporadic GI illness. There are many routes of transmission for GI illness in addition to drinking water, including foodborne and person-to-person transmission. Further, because of the limitations of monitoring microbial water quality, it is difficult to quantify actual pathogen exposure.

Randomized-controlled trials are a study design that has been employed to allow more accurate assessment of personal exposure and to reduce confounding. Although ethically, subjects cannot be randomized to drink water of known poor quality, study subjects can be randomized to regularly treated drinking water or water that has received extra treatment. This type of interventional trial has been conducted in several settings.

Laval, Canada

In 1988 and 1989, Payment et al. conducted a randomized-controlled trial in Laval, a suburb of Montreal, with subjects randomly assigned to either a home water treatment unit or normal tap water usage. They found that 35 percent of the GI illness among the tap water group could be attributed to their drinking water exposure (Payment,

Richardson et al. 1991). This result raised questions regarding the adequacy of the drinking water standards at that time, particularly for susceptible subpopulations.

In a second randomized-controlled trial conducted in Laval in 1993 and 1994, Payment et al. again found that a portion of endemic GI illness could be attributed to drinking water that met accepted standards of quality. In this study, between 14 and 40 percent of the observed case of GI illness were attributed to drinking water. This study particularly considered the role of water quality in the distribution system, and the results suggested that while additional disinfectant contact time within the distribution system may have been beneficial for water quality, overall the distribution system contributed to excess GI illness. This study did not specifically focus on susceptible subpopulations, although children were identified as those at greatest risk for GI illness (Payment et al. 1997).

The authors further suggested that the majority of the observed GI illness was likely attributable to enteric viruses, due to a short average duration of illness (Payment et al. 1997). Data from the first study found that a large proportion of GI illness among the study cohort was caused by enteric viruses, specifically noroviruses. However, the authors found that the observed difference in GI illness between study groups could not be attributed specifically to norovirus, as the rates of infection with these pathogens did not differ between the study groups (Payment, Franco, and Fout 1994).

Melbourne, Australia

A randomized-controlled trial conducted in Melbourne, Australia between 1997 and 1999 failed to find an association between the incidence of GI illness and drinking water. As in the Canadian trials, this study randomized households to home water

treatment devices and assessed rates of GI illness. The study design offered an improvement over the previous trials in that it utilized "sham" treatment devices in the control homes, and therefore blinding of exposure source was possible for both participants and researchers. The results of this study have limited generalizability to most North American water systems serving large populations. The surface water that served the study area was from a protected forest catchment, and although it was surface water, no farming or recreational activities or human habitation were allowed in the catchment area (Hellard et al. 2001).

Davenport, Iowa

The Water Evaluation Trial (WET) was a triple-blinded randomized-controlled interventional trial conducted in Davenport, Iowa from 2000 through 2002. The subjects were again randomized to receive either an active point-of-use water treatment device or a sham device. At the half-way point of the study period, the participants were switched to the alternative device. The surface water serving the study area had many sources of contamination, including agricultural and industrial runoff, and over 10 wastewater treatment plants discharged upstream of the drinking water intake of the treatment plant. The water was treated using advanced filtration techniques with disinfection, and the quality of the treated water was consistently high, even during a flood in the area which occurred during the study period (LeChevallier et al. 2004).

The results of the study did not show a signification difference in the rates of GI illness between those using the real in-home treatment devices and those using the sham devices (Yang 2003). Analyses of samples taken from water run through the treatment devices and regular tap water revealed that total coliforms were detected 6.5 times more

frequently, and at levels 20 times higher, for the treatment device samples than samples taken from tap water. This finding brings into question the effectiveness of home-filtration units, and clouds the conclusions of the study with regard to the impact of drinking water exposure on the occurrence of GI illness. The study results do support the conclusion that there may be no benefit from point-of-use water treatment units in preventing GI illness (LeChevallier et al. 2004). Another limitation of the study was the inclusion of few children, a group at increased risk for GI illness.

In preparation for the WET study, a pilot study that examined the feasibility of blinding to treatment device status was conducted in northern California (Colford Jr. et al. 2002). Both the pilot and full study found evidence of successful blinding (Colford Jr. et al. 2002; Colford et al. 2003; LeChevallier et al. 2004). The point estimate for the gastroenteritis incidence rate ratio found in the pilot study (IRR = 1.32) was consistent with that found in the Canadian trials, but because the pilot was not designed to test this hypothesis, it lacked the power to detect a significant association (Colford Jr. et al. 2002).

Time-Series Studies

Interventional trials may offer the "gold-standard" design for studies of drinking water and GI illness, but they are expensive and time-consuming. Further, they suffer from concerns regarding participant blinding and the effectiveness of the additional treatment measures in actually improving water quality. An observational study design that has been applied to examine this association is the time-series analysis. In this design, used widely for other environmental exposures such as air pollution, variation in water quality over time are related to variation in disease occurrence, typically quantified

through healthcare utilization. A strength of the time-series design is that the study population serves as its own control, and the only potentially confounding factors are those that vary over time. Time-series studies suffer from power issues, however, due to the bias to the null created by nondifferential misclassification of exposure and outcome measures. Determining the appropriate lag structure for relating drinking water exposure and healthcare utilization is difficult and testing a large number of potential lag structures can lead to an unacceptable risk for a type I error, whereby statistically significant results are observed due to chance rather than a true association. Time-series studies have been conducted in a variety of settings, as described below.

Milwaukee

Morris, et al. conducted a time-series analysis that examined turbidity and GI illness in Milwaukee. The study considered counts of physician-diagnosed GI illness as determined by ICD-9 codes for hospital admissions, emergency department (ED) visits, and outpatient visits at Medical College of Wisconsin facilities from January 1992 through April of 1993. This time period includes the large cryptosporidiosis outbreak, that had a recognized association with drinking water turbidity, and therefore analyses were conducted both including and excluding the outbreak period. ED visits and hospitalizations were considered separately, as were children (0 to 18 years) and adults (19 years and older). There were two drinking water treatment plants serving the Milwaukee area that treated surface water from a lake with multiple sources of contamination. Turbidity measurements from each of these plants were considered separately. The analysis utilized Poisson generalized linear models that controlled for seasonal and day-of-week trends, and the outcome count of GI illness was a two-week

moving average. The turbidity exposure was considered as a two-week moving average, with a one-week lag. During the outbreak, significant positive associations were found for all age groups and outcome groups when turbidity measured at the treatment plant that provided the contaminated water was considered. Excluding the outbreak, an increase of 0.5 NTU at one of the treatment plants was associated with a significant relative risk of 2.35 for ED visits for GI illness among children. There were no significant associations found among adults for this time period. The authors concluded that there may have been low-level microbial contamination present in the drinking water serving Milwaukee causing sporadic illness for up to a year preceding the outbreak (Morris et al. 1996).

In 1998, Morris et al. published results from a more sophisticated analysis of the Milwaukee data designed to identify the pathogen or pathogens causing illness. Time-series methodology similar to that used in the previous analysis was utilized to estimate the correlation between daily turbidity levels and counts of ED visits and hospitalizations for GI illness, considering various lag periods. Single-day lags were considered from 0 to 20 days; 21 total lags. The pre-outbreak and outbreak periods were considered separately, and children and adults were considered separately within these time periods. The subject's zip code of residence was used to assign a plant of service to each subject, and the turbidity reading from the assigned plant was used as the exposure measure. The strongest correlations for each analytical group were found at lags of 7 to 9 days, with much stronger correlations seen during the outbreak period than the pre-outbreak period. There were no tests conducted to determine the significance of the correlations. The authors concluded that the results offer support for sporadic drinking waterborne

cryptosporidiosis in Milwaukee in the year preceding the outbreak (Morris, Naumova, and Griffiths 1998).

The same group assessed the association between turbidity and GI illness during the same time period among the elderly using Medicare data. A time-series analytical method was used to assess the relationship between counts of ED visits and hospitalizations for GI illness and measures of turbidity of treated drinking water, at lags between 0 and 18 days. No association was found at any lag for the pre-outbreak period (Naumova et al. 2003).

Philadelphia

Schwartz, et al. published results of an investigation of treated water turbidity in relation to hospital visits among children in Philadelphia (Schwartz, Levin, and Hodge 1997). The investigators obtained data from the Children's Hospital of Philadelphia (CHOP) to determine daily counts of ED visits and hospitalizations for GI illness. ED data were available from July 1, 1992 through December 31, 1993 and hospital admission data were available for all of 1989 through 1993, except for a four month period in 1992 (March to June). ED and hospital admission data were considered separately. The investigators used two case definitions for GI illness. A more restrictive definition included those ED visits or hospital discharge diagnoses with primary ICD-9 codes 001-009.9 or 558.9. A broader definition included the restrictive definition, but also counted as cases those visits for which any of the secondary ICD-9 diagnostic codes were 001-009.9 or 558.9, if the primary code was considered conceivably related to GI illness (276, 691, 692, 787, or 789).

The three drinking water treatment plants that served the Philadelphia metropolitan area at the time of the study treated compromised surface water and provided 60 mgd, 100 mgd, and 200 mgd of water. These plants used conventional water treatment, including filtration and disinfection. The researchers obtained treated water turbidity data, measured at the plant, for the three plants from the Philadelphia Water Department and used the arithmetic average of each plant's mean daily turbidity level as the exposure measure. Turbidity data were available from 1989 through 1993. Patient zip code information was utilized to restrict cases to only residents served by the Philadelphia Water Department. In a sub-analysis, zip code was also used to assign a plant of service to each visit so that plant-specific models could be considered. Plants were considered separately because of potential bias caused by the fact that the service areas of the three plants differed in their distance from CHOP. Further, the plants utilized different raw water sources. Sub-analyses were also considered for 0 to 2 year olds and 3 to 15 year olds.

Generalized additive models were used to conduct Poisson regression that considered the association of the daily turbidity measures with the counts of daily ED visits and hospital admissions for GI illness. Lags of one to 14 days between the exposure measure and outcome count were considered. Long-term time trends, temperature, and day of week were controlled using Loess moving regression smoothers.

The researchers found an interquartile increase in turbidity (0.04 NTU) was significantly associated with a 7.2 percent increase in ED visits for GI illness four days later and a 6.7 percent increase 10 days later. The age-specific analyses showed a greater increase in ED visits with increased turbidity among the older age group, with a

significant 9.9 percent increase at a lag of four days. For the younger age group, a significant 5.9 percent increase in ED visits for GI illness was found with a 10-day lag. No results from models with other lag structures were significant for either age group. The plant-specific analyses indicated heterogeneity among plants. Results for one plant showed no significant associations, results for another plant showed significant 12.1 and 8.5 percent increases in ED visits for GI illness with lags of 1 and 10 days, respectively, and results for the remaining plant showed a significant 16.7 percent increase with a 7 to 9 day moving average lag.

Associations between turbidity levels and hospital admissions for GI illness were also found. The same turbidity change as that considered for ED visits (0.04 NTU) was significantly associated with an 8.7 percent increase in admissions for GI illness, with a lag of 8 days. The age-stratified analyses showed a 31.1 percent increase in hospital admissions for GI illness for a 6 to 7 day moving average increase in turbidity for the older age group and a 13.1 percent increase for the younger children after a lag of 13 days.

In 2000, Schwartz et al. published results of a similar study considering the association between treated water turbidity and hospital admissions for GI illness, this time considering the elderly Philadelphia population (Schwartz, Levin, and Goldstein 2000). The study utilized Medicare data to compile daily counts of hospital admissions for GI illness. A broad case definition was used, with admissions having a primary diagnosis of 001-009.9, 558.9, 276, 787, or 789 considered as a GI illness. Each of the three treatment plants was considered separately, with patient zip code used to identify a patient's plant of service. An analysis combining data across all plants was also

conducted. The exposure measure was again the mean daily turbidity level for each plant, with lags between 1 and 14 days considered. For the combined analysis, a weighted average of each plant's mean turbidity was used, with weighting based on the proportion of hospital admissions from each service area. Poisson regression with Loess regression smoothers was again used for the analyses, controlling for the effects of seasonal variation, long term time trends, temperature, and day of the week.

The results for one plant showed significant 9.1 and 15.1 percent increases in hospital admissions for GI illness with an interquartile increase in turbidity, using 4 to 6 day moving average and 11 day lags. The results from the other plants showed significant associations of turbidity lagged 9 and 10 days with GI illness hospital admission increases of 5.3 and 8.2 percent, respectively. The results of the combined analysis showed a 9 percent increase with a 9 to 11 day moving average lag. When stratified by age, the results for 75 year and older age group showed a significant 9.1 percent increase in hospital admissions for GI illness with an interquartile increase in turbidity lagged 9 to 11 days. This association was not significant for the 65 to 74 year old age group.

These studies received substantial criticism, particularly the earlier pediatric study (EPA review finds Philadelphia turbidity study seriously flawed 1998; Sinclair and Fairley 2000; Schwartz and Levin 1999). Concerns included the high potential for misclassification of exposure and disease. Cases of GI illness were assigned a turbidity exposure based on their zip code. This exposure does not account for water ingested outside of the home. Further, turbidity is a rough proxy for microbial contamination rather than a direct measure. The authors acknowledged both sources of error, but expected the results to create a bias to the null because the exposure misclassification

would be non-differential, and there were only small correlations found between turbidity levels and the values for other covariates (Schwartz and Levin 1999). The use of ICD-9 diagnostic codes in defining cases can lead to misclassification if the coding is not performed correctly. Additionally, the broader case definitions were also of concern due to the likely inclusion of many illnesses that were not related to waterborne disease. The authors also acknowledged these limitations, but made the argument that this outcome misclassification would result in only increased variance, rather than biased estimates of effect (Schwartz and Levin 1999; Schwartz, Levin, and Goldstein 2000).

Criticisms also included a lack of biological plausibility for such a small overall change in turbidity measured at the plant associated with a comparatively large increase in hospital admissions for GI illness (EPA review finds Philadelphia turbidity study seriously flawed 1998). As previously discussed, turbidity is only a rough indicator of microbial contamination. The daily levels of turbidity had only small variability, and consequently the interquartile increases used as the exposures were small. The daily fluctuations in pathogen level would thus have been exceedingly small, if associated with turbidity level at all. In the opinion of some in the drinking water community, these small changes in turbidity were within the margin of error of the turbidimeters used to derive the measurements (EPA review finds Philadelphia turbidity study seriously flawed 1998). Additionally, because the turbidity measurements were taken on treated water at the plant, no estimation of the impact of the distribution system could be assessed. Treated water at the plant is effectively at its cleanest point in the path between raw water and end user. Although the majority of contamination from the source water is removed during

treatment, this water can become re-contaminated in the distribution system through gross contamination events, loss of pressure, and biofilm build-up.

The analytical methods used in these studies were also critiqued. Of particular concern was the large number of regression models run in order to accommodate consideration of different lag structures. The only results that were presented were those for which a statistically significant result was produced. The consideration of so many models leads to a greater likelihood of a type I error. The authors attempted to address this problem in the latter article by displaying the direction of the association for each single day lag (1 to 14 days) for each plant and showing that the probability of finding as many positive associations as they did in the absence of a true association was very low (Schwartz, Levin, and Goldstein 2000). However, many of the results they reported were for lags that incorporate 2 and 3 day moving averages, which suggests that many more models were considered than the authors acknowledge. A more minor analytical concern regards their use of generalized additive models (GAMs) in the regression analyses. In the time since these studies were published, it has been reported that the software packages for GAMS underestimate the variance (Klein, Flanders, and Tolbert 2002).

Le Havre, France

A time-series study that examined anti-diarrheal medication (ADM) sales and measures of water quality was published by Beaudeau et al. in 1999. They considered the chlorine and turbidity levels of treated water as well as the raw turbidity level of water at two treatment plants in Le Havre, France from April 1993 to September 1996. The overall water quality during the study period was good, with consistently negative tests for fecal coliforms and streptococci. A statistically significant association was observed

between absences of chlorine residual in the distribution system and increases in anti-diarrheal drug sales, with a three to eight day lag. An association was also observed between increases in raw water turbidity and increases in ADM sales over the following three weeks. The authors noted that changes in turbidity had a greater association with ADM sales than the absolute level of turbidity. These turbidity changes occurred most frequently following periods of heavy rainfall, which occurred frequently in Le Havre during the study period. No association was found between treated water turbidity and ADM sales. The authors concluded that current regulations in this region do not provide complete protection from waterborne GI illness (Beaudeau et al. 1999).

Vancouver, Canada

Aramini et al. considered drinking water turbidity and healthcare utilization for GI illness, including hospitalizations, physician visits, and pediatric ED visits in Vancouver from 1992 to 1998. Using GAMs, they assessed this relationship controlling for seasonal and long-term time trends and considered one to 39 day lags. Both a Poisson model, considering counts of visits, and a case-control study, using those presenting with respiratory conditions as controls, were considered. They observed significant associations between turbidity and healthcare visits for GI illness in both models, with the strongest associations seen at lags of 3 to 6, 6 to 9, 12 to 16, and 19 to 21 days. Variations in water quality explained 0.8 to 2 percent of physician visits and 0.2 to 1.3 percent of hospitalizations for GI illness (Aramini et al. 2000).

Other Observational Studies

Cohort

A cohort study was conducted in Norway in 2003 and 2004 that considered as the exposure incidences of low water pressure, due to main breaks, construction accidents, or routine maintenance, and followed 1,165 households for the development of GI illness among members in the following week (Nygard et al. 2007). The researchers observed a risk ratio of 1.58 (95% CI: 1.2, 8.2) and estimated the attributable fraction of GI illness among those exposed to low water pressure was 37 percent. They also found that increased water consumption was associated with greater risk of GI illness among those exposed to low water pressure. The utilities participating in the study were asked to assess the risk of GI illness transmission associated with each low pressure event. None of the episodes were classified by the utilities as high-risk, and only a few were considered to impose even moderate risk.

Case-Control

A case-control study was conducted in England and Wales in 2001 and 2002 using the control group (N=427) of a study designed to examine risk factors for cryptosporidiosis (Hunter et al. 2005). Questionnaires that asked about the occurrence of diarrhea in the preceding two weeks were mailed to study subjects. Cases and controls were compared for a variety of exposures, and the variable most strongly associated with illness was self-report of water pressure loss at the home tap (odds ratio=12.5, 95% 3.5, 44.7).

A case-control study conducted in Vancouver, Canada specifically examined whether endemic giardiasis was related to drinking water (Mathias, Riben, and Osei

1992). There were 180 laboratory confirmed cases of giardiasis and 274 uninfected controls, both laboratory and neighborhood-based. The researchers found that having a young child in the house and travel outside the water district were associated with giardiasis, but that the amount of water consumed was not.

Cross-Sectional

A cross-sectional study was conducted in Cherepovets, Russia in 1998 and 1999 (N=100) that examined the contribution of distribution system water quality to endemic GI illness (Egorov et al. 2002). The researchers observed that as water traveled farther from the plant, the chlorine residual decreased. A decreased chlorine residual level was associated with higher incidence of self-reported GI illness. An interquartile decline in chlorine residual level (0.22 mg/l) was associated with a 1.42 relative risk of GI illness (95% CI: 1.05, 1.91).

In Israel, a cross-sectional analysis considered data regarding reported cases of salmonellosis, shigellosis, and hepatitis and their correlation with total coliform levels (Elkana, Gal, and Rishpon 1996). No association was found.

Ecologic

A study conducted in Sweden examined the relationship between 7,007 cases of *Campylobacter* infection and the average distribution system pipe length per person in different areas, a proxy for distribution system exposure, using Poisson regression (Nygard et al. 2004). The researchers found a rate ratio of 1.12 (95% CI: 1.08, 1.16) for GI illness with a 10 meter increase in average pipe length per person.

Conclusion

If there is a true association between drinking water that meets current quality standards and GI illness, the public health implications are far reaching. The vast majority of people in the U.S. rely on public drinking water and it is expected that this water is free of pathogens to the extent that it is highly unlikely that drinking water ingestion leads to illness. Those with immune suppression, due to age or illness, are particularly in need of clean water. Faced with the possibility that water may be insufficiently treated, it may be tempting to simply enact more stringent treatment regulations. However, the financial and technological burden that would be required by the water treatment facilities to meet more stringent regulations is great, and cannot be imposed without careful consideration. Further, recent research into the potential deleterious effects of disinfection by-products indicates the need for caution when considering increasing disinfection requirements for drinking water. In light of the balance necessary to ensure that drinking water is treated sufficiently to reduce pathogen contamination to safe levels while preventing over-exposure to potentially dangerous agents used in the treatment process, more research on this topic is warranted. Although previous studies have examined these issues, the collection of results from both the intervention trials and the time-series studies are conflicting in their conclusions, and many gaps in knowledge remain. The relative role of different water sources and treatment methods is not clear, and although the distribution system has been implicated in the deterioration of water quality, the impact of this deterioration has not been fully assessed. If an association between drinking water quality and GI illness does exist, water quality indicators that serve as sufficient proxies of pathogen presence need to be

identified. The impact of implementing new standards to improve water quality and reduce occurrence of waterborne disease needs to be investigated. The strengths and weaknesses of existing as well as future research must be considered when drawing conclusions, particularly those that may affect changes in public health practice, including legislation and recommendations to the public.

CHAPTER 5

PROJECT DESIGN AND ANALYSIS PLAN

Dissertation Goals

The overall goal of this dissertation was to address knowledge gaps by assessing the population impact of gastrointestinal (GI) illness attributable to multiple measures of drinking water quality using refined analytical techniques. Studies to date do not provide a clear understanding of the relative importance of the roles of the raw water source versus the water distribution system in contributing to microbial contamination and risk of subsequent GI illness. This dissertation sought to assess the role of drinking water in endemic illness in Atlanta and to assess whether a role exists for the raw water source or the distribution system as a site of contamination leading to endemic GI illness.

Specifically, the following research questions were examined:

1. *Is incidence of gastrointestinal illness higher among people served by drinking water with longer estimated residence times?* Increased time within the distribution system (i.e., residence time) leads to an increased likelihood of contamination. Water residence time can serve as an estimate of the potential for microbial contamination throughout the entire distribution system. Most previous studies focused only on measurements of indicator organisms taken at the plant or a limited number of sampling sites within the distribution system to estimate the level of microbial contamination at the point of ingestion.
2. *Does incidence of gastrointestinal illness differ between populations served by different drinking water treatment plants?* Different treatment plants receive raw

water from different sources and utilize different treatment processes. The age and complexity of the distribution systems served by different plants also vary. Therefore, variation in water quality may be more pronounced across different plants than temporally within a single plant. While previous studies examined individual water systems and frequently considered only outbreak conditions, this study considered all six major water utilities serving the metro-Atlanta area under non-outbreak conditions, and examined the relative importance of the various factors that may differ between water treatment plants.

3. *Are temporally varying measures of turbidity related to daily incidence of gastrointestinal illness?* Previous time-series analyses that examined only turbidity of filtered water measured at the treatment plant were replicated and improved upon by investigating both filtered and raw water turbidity. Primary analytical models were designated *a priori*, and all additional models were presented as secondary analyses to avoid emphasis on potentially spurious statistically significant results.

Research Project Design

The dissertation research project examined spatio-temporal relationships between drinking water quality and emergency department (ED) visits for GI illness. This research utilized data on emergency department visits from 28 Atlanta-area hospitals and water quality data from the ten drinking water treatment facilities serving the five-county metro-Atlanta area for the period January 1, 1993 through December 31, 2004 (Figure 5.1). The analyses modeled rates of ED visits for GI illness in the five-county Atlanta

area during this time period as a function of various measures of the likelihood for microbial contamination while controlling for potentially confounding variables.

Research Project Population

The five-county metro-Atlanta area includes Clayton, Cobb, DeKalb, Fulton, and Gwinnett counties. Subjects residing in one of these counties at the time of their ED visit were eligible for inclusion into the project. Residence in the project area was determined by zip code information available in the ED database. Demographic characteristics of the county populations are presented in Table 5.1.

Water Quality Information

The quality of drinking water in the Atlanta-metro area is generally very high. The water utilities that participated in this project had no health-based violations of U.S. Environmental Protection Agency (EPA) standards and only a total of four monitoring or reporting violations, none of which were considered serious, from 1993 through the 2004 (U.S. Environmental Protection Agency 2007) (Table 5.2). The overall goal of the project was to assess the role of drinking water in endemic GI illness, rather than in epidemics, in an area where compliance with current standards was routinely achieved. Therefore, Atlanta provided an excellent setting in which to conduct the research. Moreover, because of the large ED database, there was adequate power to assess subtle associations between water quality and GI illness.

Water quality data for this project were obtained from two main sources: 1) the drinking water treatment plants serving the five metro-Atlanta counties, 2) the Drinking Water Branch of the Georgia Department of Natural Resources (DNR).

Water Utilities

The five counties included in the project were served by a total of ten treatment plants, operated by six utilities (Table 5.3). Senior staff members at these six utilities, the Atlanta-Fulton County Water Resources Commission, the City of Atlanta Department of Water, the Cobb County-Marietta Water Authority, the Clayton County Water Authority, the DeKalb County Department of Public Works, and the Gwinnett County Water Production Division, collaborated with the project.

The utilities were asked to describe the water treatment process at each treatment plant under their operation, including changes since 1993, and to provide information that allowed designation of zip codes in their service area. The utilities were also asked to provide records of any regularly measured indicators of water quality, including but not limited to the measurements required by the Georgia DNR.

Georgia Department of Natural Resources

The Drinking Water Branch of the Georgia DNR requires the submission of monthly reports on various water quality parameters, including turbidity. The reports containing this information are publicly available, and when such data were not available directly from a utility, this information was obtained from the Georgia DNR.

Emergency Department Visit Information

SOPHIA

In 1998, the Study of Particles and Health in Atlanta (SOPHIA) began as part of a research effort examining air quality indicators and health outcomes. One component was a time-series investigation of ED visits designed to take advantage of unique air quality data being collected for the study. Participation was solicited from all of the EDs serving the 20-county Atlanta area. Thirty-one of 41 hospitals provided data on ED visits between January 1, 1993 and August 31, 2000. A second round of data collection was funded by the EPA and the National Institute of Environmental Health Sciences that allowed data to be collected through 2004. The results of the SOPHIA analyses involving air quality exposures and health outcomes, including cardiovascular events and pulmonary events, support several relationships between indicators of air quality and counts of ED visits (Metzger et al. 2002; Peel et al. 2005; Tolbert et al. 2002; Tolbert et al. 2000). Information contained in this ED database was not limited to cardiorespiratory outcomes, and the database was used in this dissertation project to examine GI illness outcomes. An EPA Science to Achieve Results (STAR) grant to fund the water quality portion of SOPHIA began in August of 2004. The project received approval and a HIPAA Waiver of Authorization from Emory University's Institutional Review Board.

Patient Population

Patients who visited any of the participating hospital EDs during the project period and resided in one of the 141 zip codes in the five-county area were eligible for inclusion in the project. Residence in an eligible zip code was determined by the patient's residential zip code recorded in the computerized billing data provided by the hospitals.

Changes in zip code designations during the project period were accounted for in exposure classification. Inclusion of eligible subjects into the project depended on the successful assignment of their residential zip code to a single treatment plant of service.

Health Outcome and Demographic Information

The 42 hospitals with EDs that served the greater (20-county) Atlanta metropolitan area during the project period were asked to participate in the study. For the first round of data collection, conducted from 1999 to 2000, 31 hospitals participated, four refused, and six could not provide useable computerized data. The final database contained data for at least part of the study period (January 1, 1993 to December 31, 2004) from 41 of the 42 hospitals that serve the area. The data provided by the hospitals included the following information: patient medical record number, unique visit number, date and time of admission, multiple International Classification of Diseases Ninth Revision (ICD-9) diagnostic codes, procedure codes, age, date of birth, gender, race, zip code of residence, residential street address, and financial class/payment method. All hospitals were not able to provide all of these data elements due to lack of information in their records or confidentiality concerns (e.g. date of birth, residential street address). Repeat visits by a patient within a single day were counted as a single visit.

The primary analyses utilized an *a priori* case definition (Table 5.4). Table 5.5 describes the number of total and GI illness-related ED visits included in the database to date, as well as the average visits per day for each year.

Covariate Information

Covariates relevant to the studies included in this project were those characteristics that may have varied spatially across zip codes or temporally over the study period. Some covariate information was provided from the ED database. Insurance status, which provides a measure of socio-economic status (SES), race, gender, and age were available from most, but not all, participating hospitals.

Census Information

Spatial covariate information not available from the ED database was obtained at the zip code level from publicly accessible U.S. Census files (U.S. Census Bureau 1990, 2000). Census data, including median income and percent minority, were used as indicators of SES. Although information was not available in the ED database regarding census tract of residence, the U.S. Census has developed a method of assigning census tracts to zip codes that allows approximate values of census information to be designated for each zip code (U.S. Census Bureau 2001).

Meteorological Information

The covariates that had the potential to confound the temporal analyses were those that varied over time and may have been associated with both the incidence of GI illness and water quality, primarily meteorological variables such as temperature and rainfall. Information regarding these variables was available in the air quality database used in other analyses utilizing the ED database. Measurements were conducted at Hartsfield-Jackson International Airport by the National Climatic Data Center (NCDC). Available data included minimum and maximum temperature and total precipitation. More localized precipitation data were available from some of the water utilities.

Epidemiologic Analyses

Descriptive Analysis

Descriptive analyses of the ED and water quality variables were conducted. Daily counts of ED visits were plotted over time, for both total visits and visits specifically for GI illness. The service areas of the participating drinking water treatment plants were mapped and linked with corresponding zip code information. Correlations between different water quality parameters were examined, in addition to correlations for the same parameters across geographic areas and across different time periods. Histograms, time plots, and graphical representations of water quality parameters were examined. Mapping provided a graphical description of the various spatially varying elements of interest.

Spatio-Temporal Analyses

The hypotheses described previously were examined through consideration of the following study aims:

- **Aim 1:** Assess the association between rates of ED visits for GI illness and estimated water residence times of drinking water serving the residential zip codes of the ED patients.
- **Aim 2:** Assess differences in rates of ED visits for GI illness among residents of the service areas of different drinking water treatment plants.
- **Aim 3:** Assess the temporal association between daily counts of ED visits for GI illness and drinking water quality, measured by raw and filtered water turbidity.
- **Aim 4:** Examine the impact of water quality on vulnerable subgroups (such as children and the elderly).

Poisson regression was used, where possible, to model the relationships between measured drinking water quality and counts of ED visits for GI illness described in Aims 1, 2, and 3. If the Poisson regression models failed to converge, due to model complexity, then logistic regression was used. Generalized estimating equations (GEE) with an autoregressive correlation structure were considered to accommodate autocorrelation (Liang and Zeger 1986; Zeger 1988) in the outcome variable. Variance estimates were scaled to account for Poisson overdispersion. The basic model had the following form:

$$\log(E(Y)) = [\text{offset}] + \alpha + \beta(\text{Exposure}) + \gamma(\text{Covariates})$$

The values of 'Y' and the vectors 'Exposure' and 'Covariates' were specified for each individual aim being investigated (Table 5.6). The offset term, used only in the analyses for Aims 1 and 2, was fundamentally the same in each analysis, and was the total number of non-injury ED visits corresponding to the time period and spatial aggregation (e.g., zip code) included in the outcome (Y) measure for each analysis. This offset was used to address the potential for differential ED utilization by those who live in different parts of the project area and for differential ED utilization over time. Non-injury visits were considered because people utilizing an ED for an injury may not have represented the underlying population from which the GI illness ED cases arose (U.S. Environmental Protection Agency 1996). The robustness of underlying assumptions was assessed by considering alternate formulations, such as including alternate numbers of knots to define the splines (Aim C), using different levels of aggregation, and using different lag structures.

Aim 1. The outcome variable for the study that assessed the association between estimated residence time and ED visits for GI illness was the proportion of ED visits for

GI illness out of all non-injury ED visits in each zip code for each year, season, hospital, age group, and Medicaid payment group. The exposure of interest in this study was residence time for each zip code, estimated using hydraulic models produced by the utility to simulate the flow of water through the distribution system and coded as a categorical variable. The cut-points for the residence time variable were defined at the 10^{th} and 90^{th} percentile of all zip code residence times, weighted by the number of non-injury ED visits in that zip code. The covariates that were considered in these models included measures of SES, race, and age derived for each zip code from census data. Individual-level information, derived from the ED database, on age and Medicaid payment status was also considered. Indicator variables for year were included to control for temporal trends. Hospital and distance from hospital to zip code centroid were included to control for the differential entry times of the hospitals into the database and for the different patterns of ED usage associated with increased distance from the hospital, respectively. Effect measure modification by year, age group, season, and Medicaid status were examined.

Aim 2. The outcome variable for the study that assessed the association between drinking water treatment plant and ED visits for GI illness was the count of ED visits for GI illness in each zip code for each year, season, hospital, age group, and Medicaid payment group. Aggregating ED counts by zip code instead of aggregating over plant service area was done to allow better control of confounding for zip codes within a service area that were heterogeneous for covariates. The exposure of interest in this study was the drinking water treatment plant of service, defined as dummy variables representing each of the treatment plants. It was not possible to assign some zip codes to

the service of a single plant either due to insufficient information or to mixing of water from more than one treatment plant. When these situations occurred, the zip code was excluded from the analysis. The covariates that were considered in this study included SES, race, age, year, residence time, and hospital. Year, age group, and season were considered as potential effect modifiers.

Aim 3. For the study that assessed the association between turbidity and ED visits for GI illness, individual models were considered for each treatment plant. The outcome variable was the count of ED visits for GI illness in each plant's service area on a given day. Several exposures were considered for this Aim: minimum and maximum raw water turbidity and maximum and mean filtered water turbidity. The plant-specific rate ratios were averaged and weighted by the inverse of the variance to create a summary measure.

The exposure was entered into the model using a 21-day unconstrained distributed lag structure (Almon 1965) and included the turbidity measurements on day t through day t-20. This lag structure was useful because of the multitude of factors that could influence the lag-time between the measurement of water quality and reporting to an ED for GI illness, including: varying water residence times, different incubation periods for the agents of waterborne illness (viruses, bacteria, and protozoa), different individual response times to pathogen exposure, and varying time periods between development of illness and presentation to an ED. Using a single *a priori* lag structure that incorporates the most plausible exposure window allowed fewer models to be run, and thereby reduced the likelihood of finding spurious, statistically significant results (type I error). Cubic splines were used to control for long-term time trends and seasonality. Knots were defined by season in the spring, summer, and autumn, and by month in the winter to

accommodate the rotavirus peaks in GI illness. Indicator variables were included to control for hospital entrance and exit from the database. Temperature and day-of-week were also included in the analytical model.

Aim 4. The impact of water quality on susceptible subgroups was considered within the context of each of the previously described studies. Whether the young and the elderly were more susceptible to serious GI illness given variations in the microbial quality of their drinking water was assessed using similar analyses as those described above, but restricted to only those age groups, and with the offset and covariates modified appropriately. Age groups were created from age information provided in the ED database.

Regression Diagnostics

Descriptive statistics for all exposure and covariate values and outcome counts were examined. Plausibility of values was assessed for all variables by examining the largest and smallest values. Utility personnel were consulted regarding ranges of plausible values for turbidity measures. Outlying observations that were considered implausible were deleted. Other extreme observations were retained in the primary analyses, and their impact was investigated through removal in secondary analyses. Plots of regression residuals against predicted outcome and against time were considered, both to identify outlying observations and to assess if time trends had been adequately accommodated. Autocorrelation was assessed by determining the correlation between values of a given variable on neighboring days. The deviance statistic was used to determine the goodness of fit of the regression models. This statistic was used to assess

overdispersion by considering the deviance divided by the degrees of freedom of the regression model. If overdispersion was present the model was scaled accordingly.

Power

The power estimates for the overall project and for analyses in which two of the important sub-populations were considered, young children (five years of age or younger) and elderly adults (65 years of age or older), were estimated using PASS 6.0 (PASS) (Table 5.7). The alpha-level was set at 0.05 and it was assumed that 25 percent of the variation in exposure was explained by the covariates. The logistic regression module was modified to simulate Poisson regression. This crude method of estimating power was compared to more refined techniques utilizing the actual study data in simulations and the results were found to be comparable.

The calculations indicate that the studies had sufficient power to detect a risk ratio as small as 1.007 between exposure groups dichotomized at the median, and potentially smaller. While the sub-analyses had less power, an odds ratio of 1.02 could be detected with sufficient power among the younger study group, and an odds ratio of 1.05 could be detected with sufficient power among the older group. These results suggested that the studies had sufficient power to detect the strength of associations expected in this type of environmental research. Schwartz et al. found increased risks of 1.09 to 1.32 associated with an interquartile increase in measured turbidity (Schwartz et al., 1997, 2000). These studies were criticized for detecting implausibly large associations. This project had the ability to detect much more subtle effects than those found in these previous studies examining the association between indicators of drinking water quality and incidence of

GI illness. These subtle effects were expected given the bias to null introduced by measurement error.

Limitations

Although this project was designed to significantly improve upon earlier studies of these topics, there were still many limitations of the data that affected the analyses and the interpretation of results. These limitations included the characteristics of the population included in the ED database, the difficulty of estimating individual-level environmental exposures, and the challenge of adequately controlling complex confounding factors.

Selection Factors Regarding ED Visits for GI Illness

The population that utilizes EDs is not likely to be representative of the general population. Those who utilize the ED for GI symptoms are likely to be of lower SES than the general population. The distribution of other characteristics, such as race or age, may also differ between subjects included in the ED database and those in the general population. There was information regarding these characteristics in the database and comparisons between the project population and the general population, using census data, were conducted to inform the generalizability of the study results.

The consideration of only GI illness ED cases missed many cases of GI illness in the community. However, the project captured the majority of ED visits for GI illness. Although cases of GI illness sufficiently severe to warrant an ED visit may not have been representative of all cases of GI illness, these more severe cases may be of greatest interest in terms of assessing the healthcare burden of drinking water-related illness.

People who experience the most severe symptoms are likely to include those most susceptible to disease, including the young, the elderly, and those who are immunocompromised (such as AIDS patients). These groups are of particular concern when considering the impact of low-level microbial contamination of drinking water.

One limitation of a retrospective study design that utilizes secondary data sources was the inability to consider other routes of transmission of GI illness. There are many transmission routes of GI illness in addition to drinking water, and there were GI illness cases included in the database that were not related to water exposure. However, because the incidence of these cases was not likely to be related to measures of water quality, a bias to the null is expected. One exception was outbreaks of GI disease that may coincidentally have been temporally or spatially related to water quality. For an outbreak to be detected in the database, it would have been of sufficient size that it would likely also have been detected by health authorities. Information regarding such outbreaks was obtained from the Georgia Division of Public Health.

Exposure Misclassification and Measurement Error

Several sources of exposure measurement error could have distorted the results when treatment plant- and distribution system-levels of measurement were used to characterize water pollution for this epidemiologic study. All of the studies used as exposures surrogates for the true exposure of interest - the likelihood of ingestion of microbially contaminated drinking water. All of the studies failed to account for water intake outside a subject's home. In addition, the degree to which individuals further treated their water at home, such as through boiling or home filtering, was not quantified. In the study addressing Aim 1 (impact of estimated water residence time on GI illness),

the residence time estimate based on distribution system modeling should have had less error than simply using the distance from the treatment plant to the end-user to estimate water age, but still there was a large amount of uncertainty associated with these estimates due to model assumptions and potential miscalibration. Among the sources of exposure misclassification in the study addressing Aim 2 (impact on illness of treatment plant) was the fact that all zip codes were not necessarily served by only one treatment plant. The study addressing Aim 3 (the time-series analysis) had several aspects that made it particularly prone to exposure mispecification. As discussed above, the amount of time water takes to travel through the distribution system differs for different geographical aggregations. The primary waterborne pathogens, viruses, bacteria, and protozoa, have different average incubation times. Further, individuals have different incubation times for a given organism. Individuals also have differential response times for reporting to an ED for GI symptoms. Considering a lag incorporating day 0 though day -20 may lessened the mispecification caused by this variability. Measurement error in the levels of turbidity must also be considered in this study.

Potential for Residual Confounding

The outcome counts and exposure variables considered in the studies for Aims 1 and 2 varied spatially. Therefore, the potential for confounding existed for characteristics that varied spatially across zip codes and may have been associated with both counts of ED visits for GI illness and water quality measures. Indicators of SES, race, and age were controlled in these models by incorporating zip code-specific data into informative variables. However, there may have been other spatially-varying zip code characteristics (e.g. child care center usage) that had the potential to confound the relationship between

water quality indicators and counts of ED visits for GI illness. Additionally, the variables derived to control for SES, race, and age distributions may not have controlled for all effects of these variables, and may have left residual confounding by these factors.

Time-series studies, such as that conducted for Aim 3, typically need only consider those factors that vary over time and might be associated with the exposure and outcome, because the nature of the study design is such that subjects effectively serve as their own controls. While the attempt was made to control for long-term time trends, seasonality, and meteorology in this study, there remained the potential for residual confounding by these factors. Cubic splines were utilized to accommodate long-term time trends in the models. When considering these temporal trends it was important to balance controlling for the effect of the confounding variables without concealing the temporal variation of the exposure of interest. This balance can be difficult to achieve and warranted careful consideration during model development.

Contributions

Atlanta provided an excellent setting for research designed to assess the role of drinking water in endemic GI illness, rather than in epidemic levels of illness. Atlanta is served by environmentally-challenged water sources, but the treatment that water receives is sufficiently effective that there were no health-based EPA violations during the project period. Further, there were no drinking water-related outbreaks of GI illness in Atlanta during the project period, but endemic GI illness was present at low levels. The ED database is the largest known database of its kind and provided sufficient power to detect modest associations. The collaboration of the water utilities with the project

greatly improved the availability and quality of relevant water quality data. This collaboration provided the ability to consult with the utility personnel, enhancing the interpretation of the results. The analytical techniques utilizing *a priori* models served to curb the problems associated with multiple testing seen in other similar studies.

The results and conclusions of this dissertation research make several contributions, including adding to the existing body of literature, which is both limited and conflicting. This research had a special focus on the role of the distribution system and provides insight into its role in endemic GI illness. The results may aid in informing the current revision of the Total Coliform Rule, which will focus on improving water quality in the distribution system. Some of the methodological concerns of previous similar studies were addressed in this study, providing a clearer interpretation of results. Ultimately, the results of this study should aid in informing decisions regarding the adequacy of current water treatment.

Figure 5.1: Location of drinking water treatment plants serving the five-county metro-Atlanta area, the hospitals included in the emergency department database, and the zip codes included in the studies

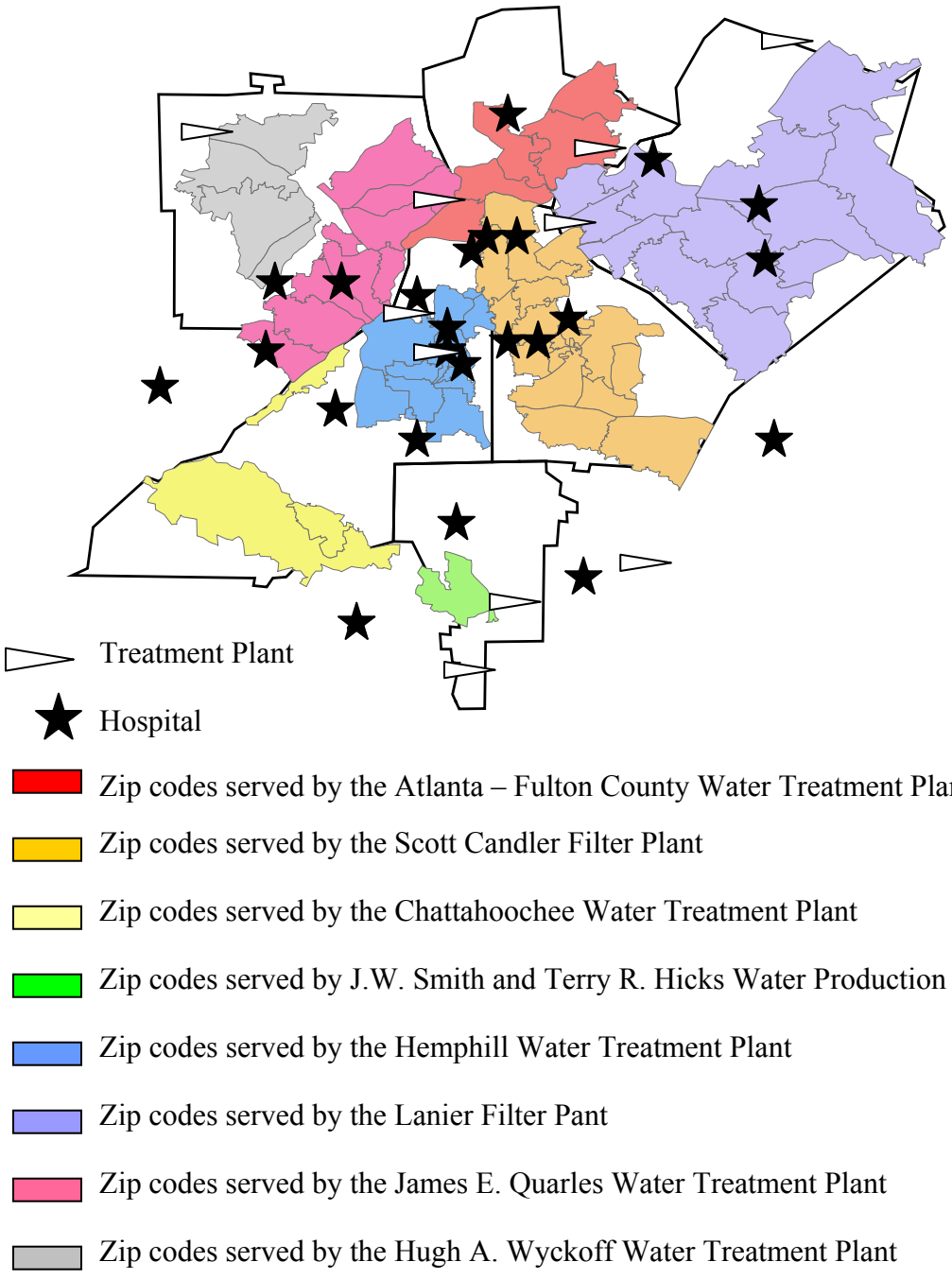


Table 5.1: Demographic characteristics of the five counties in the Atlanta-metro area, 1990 and 2000, from U.S. Census data (U.S. Census Bureau 1990, 2000)

	COUNTY									
	Clayton		Cobb		DeKalb		Fulton		Gwinnett	
CENSUS YEAR	**1990**	**2000**	**1990**	**2000**	**1990**	**2000**	**1990**	**2000**	**1990**	**2000**
Total Population	182,052	236,517	447,745	607,751	545,837	665,865	648,951	816,006	352,910	588,448
INDICATORS OF SES[a]										
Median Household Income	$33,472	$42,697	$41,297	$58,289	$35,721	$49,117	$29,978	$47,321	$43,518	$60,537
Percent High School Degree or Higher	77.2%	80.1%	85.9%	88.8%	83.9%	85.1%	77.8%	84.0%	86.7%	87.3%
Percent Bachelor's Degree or Higher	15.0%	16.6%	33.0%	39.8%	32.7%	36.3%	31.6%	41.4%	29.6%	34.1%
AGE DISTRIBUTION										
Less than 18 Years	27.9%	30.0%	25.3%	26.1%	23.7%	24.6%	24.2%	24.4%	28.0%	28.2%
18 – 24 Years	11.7%	10.4%	10.4%	9.0%	11.7%	10.9%	12.0%	11.0%	9.6%	8.7%
25 – 44 Years	37.3%	35.4%	40.4%	36.5%	38.1%	36.7%	36.4%	35.5%	42.2%	37.5%
45 – 64 Years	17.2%	18.4%	17.7%	21.5%	17.9%	19.7%	17.4%	20.7%	15.4%	20.3%
65 Years and Older	5.8%	5.9%	6.3%	6.9%	8.5%	8.0%	10.0%	3.5%	4.8%	5.4%
Median Age (Years)	N/A	30.2	N/A	33.2	N/A	32.3	N/A	32.7	N/A	32.5
RACIAL DISTRIBUTION										
White	72.4%	37.9%	87.5%	72.4%	53.6%	35.8%	47.8%	48.1%	90.9%	72.7%
Black	23.8%	51.6%	9.9%	18.8%	42.2%	54.2%	49.9%	44.6%	5.2%	13.3%
Asian	2.8%	4.6%	1.8%	3.1%	3.0%	4.0%	1.3%	3.0%	2.9%	7.2%
American Indian and Alaskan Native	0.3%	0.3%	0.2%	0.3%	0.2%	0.2%	0.2%	0.2%	0.2%	0.3%
Other	0.8%	3.5%	0.6%	3.6%	1.1%	3.5%	0.9%	2.6%	0.8%	4.3%
Two or More Races	N/A	2.1%	N/A	1.9%	N/A	2.1%	N/A	1.5%	N/A	2.2%
Hispanic (Any Race)	2.1%	7.5%	2.1%	7.7%	2.9%	7.9%	2.1%	5.9%	2.4%	10.9%

[a] SES = Socioeconomic Status

Table 5.2: EPA violations by drinking water utilities serving the five-county metro-Atlanta area, 1993 – 2004 (U.S. Environmental Protection Agency 2007)

UTILITY	VIOLATION	SDWA[a] RULE VIOLATED	SAMPLING PERIOD OF VIOLATION	DATE COMPLIANCE ACHIEVED	SIGNIFICANT VIOLATION
Clayton County Water Authority	CCR[b] Complete Failure to Report	CCR	July 1, 2001 – July 13, 2001	July 13, 2001	NO
City of Atlanta Bureau of Drinking Water	Filter Turbidity Reporting	IESWTR[c]	January 1, 2002 – January 31, 2002	July 9, 2002	NO
City of Atlanta Bureau of Drinking Water	Follow-up and Routine Tap Sampling	Lead and Copper Rule	October 1, 2000 – September 3, 2001	September 3, 2001	NO
City of Atlanta Bureau of Drinking Water	Follow-up and Routine Tap Sampling	Lead and Copper Rule	January 1, 1997 – September 22, 1997	September 22, 1997	NO

[a] SDWA = Safe Drinking Water Act
[b] CCR = Consumer Confidence Rule
[c] IESWTR = Interim Enhanced Surface Water Treatment Rule

Table 5.3: Characteristics of water utilities collaborating with the SOPHIA water quality project, 1993 – 2004

Name of Utility	**Utility Characteristics**
Atlanta-Fulton County Water Resources Commission	**Plant operated:** Atlanta-Fulton County Water Treatment Plant **Water production capacity:** 90 mgd **Connections:** ~100,000 65,000 distributed by Fulton County Water Services 35,000 distributed by the City of Atlanta **Population served:** ~350,000 **Raw water source:** Chattahoochee River **Treatment process:** Prechlorination; coagulation using alum, flocculation and sedimentation, dual media filtration, rechlorination
City of Atlanta Bureau of Drinking Water	**Plants operated:** Chattahoochee Water Treatment Plant; Hemphill Water Treatment Plant **Water production capacity:** 246.4 mgd Chattahoochee: 64.9 mgd Hemphill: 136.5 mgd **Connections:** 148,500 **Population served:** > 1,000,000 **Raw water source:** Chattahoochee River **Treatment process:** Coagulation using alum, flocculation and sedimentation, filtration, chlorination

Table 5.3, continued

Clayton County Water Authority	**Plants operated:** W.J. Hooper Water Treatment Plant; J.W. Smith Water Treatment Plant; Terry R. Hicks Water Treatment Plant **Water production capacity:** 42 mgd Hooper: 20 mgd Smith: 12 mgd Hicks: 10 mgd **Connections:** 75,000 **Population served:** 243,400 **Raw water source:** Little Cotton Indian Creek Reservoir, Blalock Reservoir, Shoal Creek Reservoir, Smith Reservoir Reservoirs served by rainwater, Flint River, Big Cotton Indian Creek, Little Cotton Indian Creek, Shoal Creek, Pates Creek Little Cotton Indian Creek **Treatment process:** Coagulation using alum; flocculation; conventional sedimentation at Smith and Hooper plants, upflow clarification pool at Hicks plant, dual media filtration, UV disinfection, chlorination
Cobb County-Marietta Water Authority	**Plants operated:** Hugh A. Wyckoff Water Treatment Plant; James E. Quarles Water Treatment Plant (2 plants on campus; older being phased out, newer being phased in) **Water production capacity:** 158 mgd Wyckoff: 72 mgd Quarles: 86 mgd **Connections:** 180,000 connections (through 14 wholesale customers) **Population served:** 650,000 **Raw water source:** Lake Alatoona (Wyckoff); Chattahoochee River (Quarles) **Treatment process:** coagulation using alum; addition of chlorine dioxide, flocculation with chlorination, sedimentation, dual media filtration, additional chlorination if necessary

Table 5.3, continued

DeKalb County Department of Public Works	**Plant operated:** Scott Candler Water Treatment Plant **Water production capacity:** 128 mgd **Connections:** 239,676 **Population served:** 576,000 **Raw water source:** Chattahoochee River **Treatment process:** coagulation using alum with chlorination; flocculation and sedimentation, dual media filtration, rechlorination
Gwinnett County Water Production Division	**Plant operated:** Lanier Filter Plant **Water production capacity:** 150 mgd **Connections:** 202,000 **Population served:** > 650,000 **Raw water source:** Lake Lanier, served by Chestatee and Chattahoochee Rivers **Treatment process:** standard coagulation with ferric salts and cationic polymer, flocculation, and filtration; use ozone and free chlorine for disinfection

Table 5.4: ICD-9[a] codes included in SOPHIA water quality study case definition and associated diagnoses

CASE DEFINITION	ICD-9 CODES INCLUDED IN CASE DEFINITION
A priori[b]	001-004, 005.0, 005.4, 005.89, 005.9, 006-007, 008.0, 008.42-008.44, 008.47, 008.49, 008.5, 008.6, 008.8, 009, 558.9[c], 787.01-787.03, 787.91
ASSOCIATED DIAGNOSES OF SELECTED ICD-9 DIAGNOSTIC CODES	
ICD-9 CODE	**ASSOCIATED DIAGNOSIS**
001	Cholera
002	Typhoid and paratyphoid fevers
003	Other *Salmonella* infections
004	Shigellosis
005	Other bacterial food poisoning
005.0	Staphylococcal food poisoning
005.4	Food poisoning due to *Vibrio vulnificus*
005.89	Other bacterial food poisoning
005.9	Unspecified food poisoning
006	Amebiasis
007	Other protozoal intestinal diseases
008	Intestinal infections due to other organisms
008.0	*E. coli*
008.42	*Pseudomonas*
008.43	*Campylobacter*
008.44	*Yersinia enterocolitica*
008.47	Other gram-negative bacteria
008.49	Other intestinal infection due to specified bacteria
008.5	Unspecified bacterial enteritis
008.6	Enteritis due to specified virus
008.8	Intestinal infection due to other organism not elsewhere classified
009	Ill-defined intestinal infections
558.9	Other and unspecified noninfectious gastroenteritis and colitis
787	Symptoms involving digestive system
787.01	Nausea with vomiting
787.02	Nausea alone
787.03	Vomiting alone
787.91	Diarrhea

[a] International Classification of Diseases, 9th Revision, Clinical Modification

[b] Defined using all available ICD-9 codes (primary and all secondary); the number of secondary codes available varies by hospital

[c] Although 558.9 is designated non-infectious, studies have indicated that it is often the code given to infectious gastrointestinal illness when the exact cause is not known (Schwartz, Levin, and Goldstein 2000; Gangarosa et al. 1992).

Table 5.5: Non-injury and gastrointestinal (GI) illness emergency department (ED) visits in the five county project area, yearly totals, from the SOPHIA database, 1993 – 2004

YEAR	TOTAL NON – INJURY ED VISITS	TOTAL GI ED VISITS (% OF NON-INJURY ED VISITS)
1993	56,421	4,461 (7.9)
1994	83,107	6,066 (7.3)
1995	113,814	9,599 (8.4)
1996	121,874	9,566 (7.8)
1997	194,639	15,751 (8.1)
1998	230,122	20,737 (9.0)
1999	251,208	21,424 (8.5)
2000	293,081	24,263 (8.3)
2001	313,948	27,179 (8.7)
2002	319,585	26,440 (8.3)
2003	344,930	30,826 (8.9)
2004	360,991	31,734 (8.8)
Total	2,683,720	228,046 (8.5)

Table 5.6: Components of analytical models

	STUDY AIM		
	1	**2**	**3**
Offset / Denominator	Count of non-injury ED visits by zip code, age category, season, hospital, and Medicaid payment group (Denominator)	Count of non-injury ED visits by zip code, age category, season, hospital, and Medicaid payment group (Offset)	None
Outcome (Y)	Count of ED visits for GI illness by zip code, age category, season, hospital, and Medicaid payment group	Count of ED visits for GI illness by zip code, age category, season, hospital, and Medicaid payment group	Daily count of ED visits for GI illness
Exposure (E)	Estimate of residence time, 3-level categorical variable based on 10^{th} and 90^{th} percentiles	Treatment plant serving zip code	Turbidity - Minimum and maximum raw water turbidity, average and maximum filtered water turbidity - 21-day unconstrained distributed lag incorporating day t through day t-20
Covariates (C)	- SES[a], derived from census data (zip code median income) and individual-level Medicaid payment status[b] - Racial distribution of zip code, derived from census data - Age category, derived from individual-level data from ED database[b] - Year[b] - Season[b] - Hospital - Distance from hospital to zip code centroid (3-level categorical variable)	- SES[a], derived from census data (zip code median income) and individual-level Medicaid payment status - Racial distribution of zip code, derived from census data - Age category, derived from individual-level data from ED database[b] - Year[b] - Season[b] - Hospital - Distance from hospital to zip code centroid (3-level categorical variable) - Distance from treatment plant to zip code centroid (3-level categorical variable)	- Long-term time trends (cubic spline) - Temperature - Hospital entry/exit dummy variables - Day-of-week

[a] SES = Socioeconomic status
[b] Considered as potential effect modifiers

Table 5.7: Power estimates for various levels of association among project population and important subsets

	Power		
Rate Ratio	**All (N=215,897 GI visits)**	**Young (N=44,188 GI visits)**	**Elderly (N=11,585 GI visits)**
1.007	0.859	0.337	0.157
1.02	1.0	0.967	0.568
1.05	1.0	1.0	0.998

Estimated using PASS 6.0 (PASS)

CHAPTER 6

DRINKING WATER RESIDENCE TIME AND EMERGENCY DEPARTMENT VISITS FOR GASTROINTESTINAL ILLNESS IN ATLANTA, 1996 – 2003

Tinker, Sarah C.[1]; Tolbert, Paige E.[1]; Moe, Christine L.[1]; Klein, Mitchel[1]; Flanders, W. Dana[1]; Uber, Jim[2]; Amirtharajah, Appiah[3]; Singer, Phillip[4]

[1] Rollins School of Public Health, Emory University, Atlanta, Georgia

[2] University of Cincinnati, Cincinnati, Ohio

[3] Georgia Institute of Technology, Atlanta, Georgia; CH2MHill

[4] University of North Carolina, Chapel Hill, North Carolina

Abstract

Over the last several years aging pipe systems used to distribute treated drinking water in the United States are playing an increasing role in waterborne disease outbreaks, and the potential for endemic disease transmission through this route is substantial. Water residence time, the time water takes to travel from the treatment plant to the end user, serves as an indicator of the potential for distribution system contamination. It can be estimated using hydraulic models, which incorporate pipe layout, size, and flow. The current study examined whether the average water residence time for a zip code was related to the proportion of emergency department (ED) visits for gastrointestinal (GI) illness among residents of that zip code. ED data were collected from all hospitals located in the five-county Atlanta, Georgia area from 1996 to 2003. Two hydraulic models, one covering 1996 to 1998 and another covering 1999 to 2003, were used to estimate water residence times for the service area of one of the largest water utilities in the city, which serves 650,000 people. People served by the utility had over one million total ED visits, 63,652 of them for GI illness. The relationship between water residence time and GI illness risk was assessed using unconditional logistic regression, controlling for potential confounding factors, including age, year, hospital, distance from zip code to hospital, and markers of SES. Models were stratified by the hydraulic model used to derive the residence time estimates. The odds ratios comparing GI illness risk among residents of zip codes with short average water travel times (10^{th} percentile or less) to risk among those in zip codes with intermediate average water travel times (11th to 89th percentile) suggested little or no association. However, the odds ratios for residents of zip codes in the top decile of average water residence times compared to residents of zip codes with

intermediate average water residence times suggested modestly increased risk of GI illness ED visits (odds ratio for hydraulic model 1=1.133, 95% confidence interval=1.054, 1.217; odds ratio for hydraulic model 2=1.045, 95% confidence interval=1.003, 1.089). Effect measure modification by year was also suggested. The results support a contribution of drinking water distribution system contamination to the burden of endemic GI illness.

Introduction

Community-based treatment of water intended for human consumption is one of the most important public health practices in the U.S. Waterborne disease outbreaks have decreased in frequency, size, and severity as the processes of source water protection and treatment have improved (U.S. Centers for Disease Control and Prevention 1999). Despite these advances, infectious disease can still be transmitted through drinking water in the U.S. An area of particular concern in recent years is post-treatment contamination of drinking water in the distribution system.

Aging pipe systems used to distribute treated drinking water are playing an increasing role in waterborne disease outbreaks in the U.S. Between 1971 and 1998, drinking water distribution system contamination was associated with 113 reported outbreaks, resulting in 21,000 cases of illness, 498 hospitalizations, and 13 deaths (Craun and Calderon 2001). From 1999 through 2004, half (11/22) of all reported outbreaks in community water systems were due to distribution system deficiencies, resulting in 202 additional cases (Blackburn et al. 2004; Lee et al. 2002; Liang et al. 2006; Nygard et al. 2007). While most reported distribution system outbreaks were caused by significant or continuing contamination, low-level or transient contamination is also likely to occur in distribution systems, and the potential for endemic disease transmission through this route is substantial. Results from randomized-controlled trials have suggested that the distribution system contributes to endemic drinking water-related gastrointestinal (GI) illness (Payment et al. 1997). These results have been supported by observational studies that found an association of increased distance from the plant to the end-user and increased incidence of GI illness (Egorov et al. 2002; Nygard et al. 2004) and also studies

that found an association between pressure-loss events in the distribution system and increased incidence of GI illness (Hunter et al. 2005; Nygard et al. 2004; Nygard et al. 2007).

Measuring the actual levels of pathogen contamination in the distribution system is not currently feasible, and therefore surrogates of the potential for contamination may be a useful alternative. Residence time, or water age, is the amount of time water spends in the distribution system between the treatment plant and the end user. The longer time that water spends in the distribution system, the more likely it is to encounter contamination. Residence time is estimated using hydraulic models, which incorporate pipe characteristics, operating rules, and pressures to simulate the way water flows through the distribution system pipes. For this study, we developed average residence time estimates for the service area of a large utility in the Atlanta, Georgia area. We examined the relationship of residence time with emergency department (ED) visits for GI illness among people served by this utility to assess the contribution of distribution system contamination in endemic GI illness in this population.

Methods

Emergency Department Visits

Information was available on visits from all of the hospitals operating within the five-county Atlanta area (23 hospitals) and from five hospitals located outside the study area that contributed a substantial number of visits by five-county residents. Data from one hospital had to be excluded from the analysis due to missing covariate information. The information provided by the hospitals for each ED visit included medical record

number, date of admission, International Classification of Diseases, 9th Revision (ICD-9), diagnosis codes, zip code of residence, payment method, and age or date of birth.

We defined cases of GI illness *a priori* using the primary and all available secondary ICD-9 diagnostic codes. This case definition included the following diagnoses: infectious GI illness (001-004, 005.0, 005.4, 005.89, 005.9, 006-007, 008.0, 008.42-008.44, 008.47, 008.49, 008.5, 008.6, 008.8, 009), non-infectious GI illness (558.9), and nausea and vomiting plausibly related to GI illness (787.01-787.03, 787.91). Non-infectious GI illness was included in the case definition because previous research has shown that many infectious cases of GI illness are misclassified into this diagnostic category (Gangarosa et al. 1992; Hoxie et al. 1997; Schwartz, Levin, and Hodge 1997).

Distribution System Residence Time

Residence time was estimated using the hydraulic models from one of the six major utilities serving Atlanta, which operates two treatment plants that share a single distribution system. The hydraulic models simulate the typical flow patterns of water through the distribution system, taking into account pipe layout and size, customer demand, and operating rules. The models allowed estimation of the average time water takes to flow from the treatment plant to specific pipe intersections, called nodes. For this analysis, we averaged the residence times for nodes in each zip code in the service area for each year from 1996 through 2003. Only zip codes served by the utility and with their centroid located within the boundary of the five-county Atlanta area were included in the analysis (19 zip codes). One zip code was excluded from the analysis due to missing covariate data. Two hydraulic models were considered. One described the distribution system from 1996 through 1998 and the other from 1999 through 2003.

Other Covariates

We obtained information regarding zip code level demographic characteristics, including median income and percent minority, from the 2000 U.S. Census (U.S. Census Bureau 2000).

Analytic Methods

All analyses were performed using SAS statistical software (SAS 2002 - 2003). We developed unconditional logistic regression models *a priori* to consider the association between residence time and ED visits for GI illness. Water residence time was considered as a three-level categorical variable, with breaks at the 10^{th} and 90^{th} percentile, with residence time estimates from the two hydraulic models considered separately (8.34 hours and 36.37 hours, respectively, for hydraulic model 1 and 11.18 hours and 51.78 hours, respectively, for hydraulic model 2). The models also included the following covariates to control for potential confounding: three indicators for a four-level categorical age variable (age 0 to 5, 6 to 18, 19 to 64, and 65+ years), indicator variables for year, hospital, and distance from residential zip code centroid to hospital; continuous variables for residential zip code median income and residential zip code percent minority; and Medicaid payment status (1/0). The models took the following form:

$$\text{logit}(P(Y_{z,a,y,s,h,m}/\text{denom}_{z,a,y,s,h,m})] = \alpha + \beta_1(\text{residence\_time\_10}_z) + \beta_2(\text{residence\_time\_90}_z) + \Sigma_{i=1\text{-}3}(\chi_i)(\text{age}_a) + \Sigma_{j=1\text{-}(2 \text{ or } 4)}(\delta_j)(\text{year}_y) + \Sigma_{k=1\text{-}3}(\varepsilon_k)(\text{season}_s) + \Sigma_{l=1\text{-}25}(\phi_l)(\text{hospital}_h) + \gamma(\text{dis\_h1}_{z,h}) + \eta(\text{dis\_h2}_{z,h}) + \iota(\text{median income}_z) + \varphi(\text{percent\_minority}_z) + \kappa(\text{Medicaid}_m) + \Sigma_{n=1\text{-}3}(\lambda_n)(\text{dis\_h1}_{z,h}*\text{age}_a) + \Sigma_{o=1\text{-}3}(\mu_o)(\text{dis\_h2}_{z,h}*\text{age}_a) + \Sigma_{p=1\text{-}25}(\nu_p)(\text{dis\_h1}_{z,h}*\text{hospital}_h) + \Sigma_{q=1\text{-}25}(o_q)(\text{dis\_h2}_{z,h}*\text{hospital}_h) + \Sigma_{r=1\text{-}3}(\pi_r)(\text{Medicaid}_m*\text{age}_a)$$

where Y indicates the number of ED visits for GI illness within each zip code (z), age category (a), year (y), season (s), hospital (h), and Medicaid payment group (m). This model estimates the probability of a person visiting an ED for a non-injury cause having a GI illness, given the estimated drinking water residence time of their zip code of residence, while controlling for potentially confounding factors. The variables dis_h1 and dis_h2 are indicators for a three-level categorization of the distance from the zip code centroid to the hospital. The denominator term indicates the number of non-injury ED visits in the same strata.

Models were run separately for 1996 through 1998 and 1999 through 2003, corresponding to the time period covered by each of the hydraulic models used to estimate water residence time. The analyses were stratified because the earlier hydraulic model was more skeletonized, with details of only the larger pipes, while the later hydraulic model incorporated details on all of the pipes. These differences in the hydraulic models led to consistently shorter water residence time estimates derived from the earlier model compared to those derived from the later model.

The comparison group considered in the logistic regression model was the number of non-injury ED visits in that stratum. We used this comparison group because information on ED visits was not available from all hospitals during the entire study period and because different hospitals had different proportions of ED visits from residents of different zip codes. We also considered the hospital at which the ED visit occurred as a covariate in the model because the proportion of non-injury ED visits for GI illness varied by hospital, more so than could plausibly be explained by drinking water exposure, which thereby created the potential for confounding. For example, data

from a hospital with a greater proportion of non-injury ED visits for GI illness than other hospitals could cause a spurious relationship if people living in zip codes with long or short water residence times utilized that hospital more than others.

We included year in the model because the water residence time estimates varied over time and we wanted to account for long term time trends in GI illness ED visits that may have been unrelated to drinking water. The age distributions of zip codes differed, and the incidence of GI illness, and use of an ED for this illness, likely differed by age group. Therefore, we controlled for age group in the analytical model.

Another spatially varying attribute considered in the model was socio-economic status (SES), which is plausibly related to both the incidence of GI illness and ED use. We attempted to control for SES in our model by including a variable indicating whether the patient paid for the ED visit with Medicaid and two contextual variables derived from U.S. Census data encoding the median income of the zip code and the percent of zip code residents that were of minority (non-Caucasian) race. Because fewer adults are Medicaid eligible, it serves as a better indicator of being low SES for children than adults. We therefore added an interaction term for age and Medicaid payment status in the model.

We considered the distance from zip code centroid to hospital (DIS_H1 and DIS_H2) as a potential confounder because results from preliminary descriptive analyses suggested that the rate of GI illness ED visits by people from a given zip code decreased as the distance between that zip code and the hospital increased. This association varied based on age group and hospital, so interaction terms for these variables were also included in the model.

The exposure, water residence time, was not estimated separately by season, and therefore season could not act as a confounding factor in the analysis. This variable was considered in the model in order to assess effect measure modification, because the association between drinking water residence time and GI illness ED visits could plausibly vary by season due to selective survival or transmission of the pathogens causing infectious GI illness.

We considered models stratified by year, age category, season, and Medicaid payment status. Likelihood ratio tests were used to assess effect measure modification.

Results

Twenty-seven hospitals provided data on 785,634 non-injury ED visits in the utility's service area and 63,652 (8.1%) of these visits were for GI illness. There were no drinking waterborne disease outbreaks in Atlanta during the study period, although there was a recreational water outbreak of pathogenic *E. coli* at a water park located in the study area (Gilbert and Blake 1998). Because of the small number of outbreak-related cases (26 confirmed) no indicator variable was used in the model to account for this outbreak. Thirty percent of the ED visits for GI illness were among children age five years or less, while only 17 percent of non-GI illness visits were among children in this age group (Table 6.1). The highest proportion of non-injury ED visits for GI illness occurred in winter months, while the highest proportion of non-injury ED visits for all other causes occurred during autumn months. Twenty-one percent of GI illness patients and 17 percent of non-injury, non-GI illness patients paid with Medicaid.

Average distribution system water residence times were estimated for each zip code in the utility's service area for each year from 1996 through 2003 (Table 6.2). The average residence time for estimates derived from the first hydraulic model, covering 1996 through 1998, was 24.7 hours (standard deviation=16.1 hours) and for estimates derived from the second hydraulic model, covering 1999 through 2003, was 32.8 hours (standard deviation=18.2 hours). The maximum residence time was 88.4 hours, and the minimum was 4.5 hours. A total of 272,782 nodes were used to develop the water age estimates. As expected, those zip codes closest to the treatment plants tended to have the shortest estimated residence times (Figure 6.1).

In the analytical models, the intermediate residence time category, encompassing zip codes falling in the 11^{th} to 89^{th} percentile of residence times estimated from each hydraulic model, was used as the referent. This choice of referent was based on the *a priori* hypothesis that short residence time has the potential for association with both decreased and increased risk of GI illness and long residence time has the potential to increase the risk of GI illness, as well as the stability of this referent category (50,936 GI illness ED visits). The water at the beginning of the distribution system has not had as much opportunity to encounter contamination within the distribution system, and therefore people receiving drinking water from the beginning of the distribution system may be at lower risk of GI illness. Alternatively, the water at the beginning of the distribution system has not had as much contact time with the disinfectant residual and if contamination from the raw water source was not eliminated at the plant, this water may have a higher risk for contamination than water that is exposed to this disinfectant for a longer amount of time. A long residence time represents a higher potential for

encountering distribution system contamination. The results of the adjusted models were consistent with the null when the odds of GI illness ED visits among people living in zip codes with short residence times was compared to that for people living in zip codes with intermediate residence times [Odds Ratio (OR) for hydraulic model 1 = 0.949, 95% Confidence Interval (CI) = 0.884, 1.020; OR for hydraulic model 2 = 1.010, 95% CI = 0.972, 1.049] (Figure 6.2). The results suggested a modest increase in risk of GI illness ED visits for people living in zip codes with the longest residence times compared to people living in zip codes with intermediate residence times (OR for hydraulic model 1 = 1.133, 95% CI = 1.054, 1.217; OR for hydraulic model 2 = 1.045, 95% CI = 1.003, 1.089).

Effect measure modification was assessed by year, age group, season, and Medicaid payment status using likelihood ratio tests. Year was the only covariate for which the test statistic was significant ($\alpha = 0.05$) (Figure 6.3). The positive risk ratio estimates observed in the overall adjusted analyses for the long residence time exposure were observed for all years but one in stratified analysis, although only three of the estimates were statistically significant ($\alpha = 0.05$), those for 1997, 1998, and 2002. One negative risk ratio estimate was observed for 1999, which was of borderline statistical significance.

Discussion

The results of the analysis suggest that people served by drinking water that has spent the greatest amount of time in the distribution system may be at increased risk of GI illness ED visits. These results support recent focus on the distribution system as a source

of waterborne illness (LeChevallier et al. 2003; Craun and Calderon 2001; Liang et al. 2006).

Studies of distribution system water quality have demonstrated that as distance from the plant increases, the level of bacterial contamination can steadily increase (Payment, Gamache, and Paquette 1988). There are many ways that transient distribution system contamination may occur. Low-pressure events, which can be caused by main breaks, sudden changes in demand, uncontrolled pump starting or stopping, opening and closing of fire hydrants, power failures, air valve slams, and flushing operations, decrease the normally positive pressure maintained in the pipes, thereby drawing water from outside the pipes inside through leaks (LeChevallier et al. 2003). Water systems commonly lose more than 10 percent of the water they produce through leaks in the distribution system (American Water Works Association Research Foundation 1992; Kirmeyer et al. 2001; LeChevallier et al. 2003). Studies have demonstrated that there are high levels of fecal contamination and human viruses in the soil and water surrounding distribution system pipes (Karim, Abbaszadegan, and LeChevallier 2003; LeChevallier et al. 2003). Although engineering standards call for a minimum separation of 0.5 to 3 meters for drinking water pipes and sewer lines, microbes can move several meters in short periods of time under certain conditions (Abu-Ashour et al. 1994; LeChevallier et al. 2003). The combination of negative pressure and proximity to contamination can result in pathogenic organisms entering the distribution system. While regulations in the U.S. require a disinfectant residual to be maintained in water traveling through the distribution system, this residual is not always present in all parts of the distribution system or may not be available in sufficient concentration to inactivate a large influx of

contaminants, particularly as distance from the plant increases. In addition, many waterborne disease organisms are somewhat resistant to disinfection. The transient nature of low pressure events likely results in a small total volume of contaminated water that is unlikely to be detected during routine sampling (Karim, Abbaszadegan, and LeChevallier 2003). A study in England found an odds ratio of 12.5 for experiencing a water pressure loss in the preceding two weeks and having a diarrheal episode (Hunter et al. 2005). Swedish investigators recently reported results of a cohort study in which instances of main breaks or other pressure loss events were associated with an increased risk of GI illness in the following week (Nygard et al. 2007).

Distribution system contamination can also result from biofilm disruption. Biofilms form when corrosion in the distribution system pipes produces turbercles, that increase the surface area of the pipe (LeChevallier, Welch, and Smith 1996). This process in turn promotes hydraulic mixing and transport of nutrients to the pipe surface and precipitation of organic compounds. The cracks and crevices in the pipe provide a place where bacteria and other organisms are protected from disinfection. In addition, biofilms accelerate further pipe corrosion, consume disinfectant residual, and promote the acquisition of resistance or virulence factors through organisms' proximity (Ford 1999). Biofilms can release pathogenic organisms into drinking water when flow disruptions occur, such as during routine flushing (Trussell 1999).

Few studies have examined the association between water quality in the drinking water distribution system and GI illness. Secondary data analysis from a randomized trial found no correlation between estimated drinking water residence time and incidence of GI illness (Payment et al. 1997). Correlations take into account only linear relationships,

however, and our results suggest an increased risk only for those served by water that had traveled the longest amount of time. A study in Russia found that as the distance from the plant increased, the chlorine residual decreased, and that a decreased chlorine residual was associated with higher rates of GI illness (Egorov et al. 2002). Secondary analyses of data from a different randomized-controlled trial also suggested higher rates of GI illness as distance from the plant increased (Payment, Franco, and Siemiatycki 1993).

There are several limitations of the analysis that must be considered when interpreting the results. Exposure misclassification likely occurred because of drinking water exposure outside the zip code of residence. One zip code was excluded because of missing covariate data. We examined the impact of excluding this zip code and determined it to be minimal compared to the importance of including the covariate information in the model (See Appendix A).

Our database captured only those GI illnesses severe enough to result in a visit to an ED, a small proportion of all cases of GI illness. However, these more severe cases may be those of greatest interest in terms of the public health impact of waterborne disease. The large size of our ED database provided enough power to detect modest associations. The consideration of only one utility's distribution system may limit generalizability of the results. Future analyses will incorporate water residence time estimates from other utilities in the Atlanta area.

While we attempted to control for factors that may confound the association between water residence time and ED visits for GI illness, residual confounding likely remains. Because residence time was assigned based on zip code, any zip code characteristics that might have been associated with GI illness had to be considered. A

difficult factor to control for in this analysis was SES. We attempted to account for SES using three variables: zip code median income, zip code percent minority, and whether or not the patient paid for the ED visit using Medicaid. There are likely other spatially-varying factors that we were not able to consider that may have influenced our results, including other risk factors for GI illness, such as day care attendance or food consumption habits.

There was also potential for confounding if an outbreak of GI illness due to a cause other than drinking water led to a cluster of cases in parts of the study area. An outbreak of pathogenic *E. coli* occurred at a water park located in the study area in 1998 (Gilbert and Blake 1998). However, there were only 26 confirmed cases and the park was located in a zip code classified in the referent category and therefore any clustering of cases in this area would be expected to bias in a negative direction. In sensitivity analyses we excluded the outbreak time period from the database and our conclusions were unchanged (See Appendix A).

In sensitivity analyses (See Appendix A), we considered generalized estimating equations (GEE) (Liang and Zeger 1986), which allow accommodation of correlation due to non-independent observations. Our primary analysis was conducted with the zip code as the unit of analysis and because several strata within zip code were considered the observations were not completely independent. We found that the point estimates from the GEE models were equal to those estimated using standard logistic regression and that the variances were only modestly underestimated by the *a priori* models. We were unable to run GEE models with all of the *a priori* covariates, as the models would not converge.

Therefore, the results from the regular logistic regression model are presented as the primary results.

An additional concern was the use of two hydraulic models to derive the residence time estimates. The two hydraulic models used to derive the estimates differed in their complexity. The hydraulic model covering 1996 to 1998 included only larger pipes, while the hydraulic model covering 1999 to 2003 included all pipes in the distribution system. As expected, the residence time estimates from the earlier hydraulic model were consistently shorter than those from the later hydraulic model (Table 6.2). The estimates in the later hydraulic model incorporated more complete system information, but as a consequence of that model's complexity, there was greater potential for model misspecification than with the simpler early model. Therefore, we considered the estimates produced from each hydraulic model separately and based our exposure classification on the percentile of water age from the collection of those estimated from each hydraulic model. Our goal was to assess a change in the risk of GI illness among people served by water at the beginning and at the end of the distribution system, and not to identify an absolute residence time at which water in the distribution system may impart a greater risk of GI illness.

The suggested effect measure modification by year may be explained by differences in the potential for distribution system contamination over time. This potential may be impacted by meteorological conditions, the frequency of pressure loss events, and infrastructure integrity, all of which may change from year to year. We must also consider that the degree of residual confounding may have differed by year, due to,

for example, a change in the underlying socioeconomic status of the zip codes included in the analysis.

The results of this study support the conclusion that a modest amount of GI illness may be transmitted through drinking water that spends the longest amount of time in the distribution system. These results also highlight the useful metrics that can be estimated using hydraulic models. Given the growing concern about the role of the drinking water distribution system in endemic GI illness, we demonstrate here that hydraulic models can be used to inform health studies and identify areas of vulnerability. This information can help water utilities achieve their goal of maintaining distribution system integrity and delivering microbially safe water to consumers.

References

Abu-Ashour, J., D. M. Joy, H. Lee, H. R. Whiteley, and S. Zelin. 1994. Transport of microorganisms through soil. *Water, Air, Soil Pollut* 75:141-58.

American Water Works Association Research Foundation. 1992. Water Industry Database: Utility Profiles: American Water Works Association.

Blackburn, B. G., G. F. Craun, J. S. Yoder, V. Hill, R. L. Calderon, N. Chen, S. H. Lee, D. A. Levy, and M. J. Beach. 2004. Surveillance for waterborne-disease outbreaks associated with drinking water--United States, 2001-2002. *MMWR Surveill Summ* 53 (8):23-45.

Craun, G. F., and R. L. Calderon. 2001. Waterborne disease outbreaks caused by distribution system deficiencies. *J Am Water Works Assoc* 93 (9):64-75.

Egorov, A., T. Ford, A. Tereschenko, N. Drizhd, I. Segedevich, and V. Fourman. 2002. Deterioration of drinking water quality in the distribution system and gastrointestinal morbidity in a Russian city. *Int J Environ Health Res* 12 (3):221-33.

Ford, T. E. 1999. Microbiological safety of drinking water: United States and global perspectives. *Environ Health Perspect* 107 (Suppl 1):191-206.

Gangarosa, R. E., R. I. Glass, J. F. Lew, and J. R. Boring. 1992. Hospitalizations involving gastroenteritis in the United States, 1985: the special burden of the disease among the elderly. *Am J Epidemiol* 135 (3):281-90.

Gilbert, L., and P. Blake. 1998. Outbreak of *Escherichia coli* O157:H7 infections associated with a water park. *Georgia Epidemiology Report* 14 (7):1-2.

Hoxie, N. J., J. P. Davis, J. M. Vergeront, R. D. Nashold, and K. A. Blair. 1997. Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *Am J Public Health* 87 (12):2032-5.

Hunter, P. R., R. M. Chalmers, S. Hughes, and Q. Syed. 2005. Self-reported diarrhea in a control group: a strong association with reporting of low-pressure events in tap water. *Clin Infect Dis* 40 (4):e32-4.

Karim, M. R., M. Abbaszadegan, and M. W. LeChevallier. 2003. Potential for pathogen intrusion during pressure transients. *J Am Water Works Assoc* 95 (5):134-46.

Kirmeyer, G. J., M. Friedman, K. Martel, D. Howie, M. W. LeChevallier, M. Abbaszadegan, M. R. Karim, J. E. Funk, and J. Harbour. 2001. Pathogen intrusion into the distibution system. Denver: American Water Works Association Research Foundation and American Water Works Association.

LeChevallier, M. W., R. W. Gullick, M. R. Karim, M. Friedman, and J. E. Funk. 2003. The potential for health risks from intrusion of contaminants into the distribution system from pressure transients. *J Water Health* 1 (1):3-14.

LeChevallier, M. W., N. J. Welch, and D. B. Smith. 1996. Full-scale studies of factors related to coliform regrowth in drinking water. *Appl Environ Microbiol* 62 (7):2201-11.

Lee, S. H., D. A. Levy, G. F. Craun, M. J. Beach, and R. L. Calderon. 2002. Surveillance for waterborne-disease outbreaks--United States, 1999-2000. *MMWR Surveill Summ* 51 (8):1-47.

Liang, J. L., E. J. Dziuban, G. F. Craun, V. Hill, M. R. Moore, R. J. Gelting, R. L. Calderon, M. J. Beach, and S. L. Roy. 2006. Surveillance for waterborne disease

and outbreaks associated with drinking water and water not intended for drinking--United States, 2003-2004. *MMWR Surveill Summ* 55 (12):31-65.

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

Nygard, K., Y. Andersson, J. A. Rottingen, A. Svensson, J. Lindback, T. Kistemann, and J. Giesecke. 2004. Association between environmental risk factors and *campylobacter* infections in Sweden. *Epidemiol Infect* 132 (2):317-25.

Nygard, K., E. Wahl, T. Krogh, O. A. Tveit, E. Bohleng, A. Tverdal, and P. Aavitsland. 2007. Breaks and maintenance work in the water distribution systems and gastrointestinal illness: a cohort study. *Int J Epidemiol* In press.

Payment, P., E. Franco, and J. Siemiatycki. 1993. Absence of relationship between health effects due to tap water consumption and drinking water quality parameters. *Water Sci Technol* 27 (3-4):137-43.

Payment, P., F. Gamache, and G. Paquette. 1988. Microbiological and virological analysis of water from two water filtration plants and their distribution systems. *Can J Microbiol* 34 (12):1304-9.

Payment, P., J. Siemiatycki, L. Richardson, G. Renaud, E. Franco, and M. Prevost. 1997. A prospective epidemiological study of gastrointestinal health effect due to the consumption of drinking water. *Int J Environ Health Res* 7:5-31.

SAS 9.1. SAS Institute, Inc., Cary, North Carolina.

Schwartz, J., R. Levin, and K. Hodge. 1997. Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* 8 (6):615-20.

Trussell, R. 1999. Safeguarding distribution system integrity. *J Am Water Works Assoc* 91:46-54.

U.S. Census Bureau. 2000. General Housing Characteristics: 2000. Census Summary File 1 (SF 1) 100-Percent Data. Washington, DC: American FactFinder, U.S. Census Bureau.

U.S. Centers for Disease Control and Prevention. 1999. From the U.S. Centers for Disease Control and Prevention. Control of infectious diseases, 1900-1999. *JAMA* 282 (11):1029-32.

Table 6.1: Descriptive statistics for emergency department visits for gastrointestinal (GI) illness and other non-injury causes, Atlanta[a], 1996 – 2003

	GI illness visits (% of all GI illness visits)	Non-GI illness visits (% of all non-GI illness visits)	All non-injury visits (% of all non-injury visits)
Year			
1996	2,345 (3.7)	31,263 (4.3)	33,608 (4.3)
1997	5,662 (8.9)	65,734 (9.1)	71,396 (9.1)
1998	7,758 (12.2)	83,326 (11.5)	91,084 (11.6)
1999	8,225 (12.9)	92,173 (12.8)	100,398 (12.8)
2000	9,254 (14.5)	107,900 (14.9)	117,154 (14.9)
2001	9,859 (15.5)	109,337 (15.1)	119,196 (15.2)
2002	9,368 (14.7)	110,678 (15.3)	120,046 (15.3)
2003	11,181 (17.6)	121,571 (16.8)	132,752 (16.9)
Age Category			
0 – 5 Years	19,336 (30.4)	123,953 (17.2)	143,289 (18.2)
6 – 18 Years	7,512 (11.8)	78,351 (10.9)	85,863 (10.9)
19 – 64 Years	32,095 (50.4)	428,639 (59.4)	460,734 (58.6)
65+ Years	4,709 (7.4)	91,039 (12.6)	95,748 (12.2)
Season			
Winter	18,503 (29.1)	184,011 (25.5)	202,514 (25.8)
Spring	16,682 (26.2)	176,898 (24.5)	193,580 (24.6)
Summer	13,315 (20.9)	173,021 (24.0)	186,336 (23.7)
Autumn	15,152 (23.8)	188,052 (26.0)	203,204 (25.9)
Medicaid Payment Status			
Paid with Medicaid	13,612 (21.4)	120,614 (16.7)	134,226 (17.1)
Did not pay with Medicaid	50,040 (78.6)	601,368 (83.3)	651,408 (82.9)
Total	63,652	721,982	785,634

[a] One county considered

Table 6.2: Descriptive statistics of estimated drinking water residence time[a] by year, summarized over zip codes, Atlanta[b], 1996 – 2003

	Short water residence time exposure category[c] Mean (Min – Max)	Intermediate water residence time category[c] Mean (Min – Max)	Long water residence time category[c] Mean (Min – Max)
Model 1: All years combined (1996-1998)	6.83 (4.54 – 8.28)	21.96 (8.34 – 36.37)	47.40 (37.61 – 68.46)
Model 2: All years combined (1999-2003)	10.12 (9.14 – 11.18)	33.43 (11.19 – 51.78)	74.41 (52.20 – 88.36)
1996	6.36 (4.54 – 8.28)	20.82 (8.47 – 35.07)	43.86 (38.39 – 68.23)
1997	6.63 (5.30 – 8.10)	20.66 (8.34 – 34.28)	45.99 (37.61 – 68.46)
1998	7.05 (5.29 – 8.00)	23.39 (9.11 – 36.37)	51.14 (39.31 – 63.38)
1999	10.03 (9.30 – 11.09)	31.24 (11.19 – 47.18)	66.45 (52.20 – 88.36)
2000	10.21 (9.14 – 11.04)	34.21 (19.08 – 51.78)	81.44[d]
2001	10.12 (9.31 – 11.18)	33.55 (11.19 – 48.93)	75.99 (66.17 – 83.75)
2002	10.38 (9.25 – 11.14)	35.46 (17.94 – 50.27)	74.43 (63.39 – 83.11)
2003	9.49[d]	32.84 (11.36 – 50.79)	79.62 (71.11 – 86.44)

[a] In hours
[b] One county considered
[c] Short water residence time defined as $\leq 10^{th}$ percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11^{th} to 89^{th} percentile; long = $\geq 90^{th}$ percentile
[d] Only one zip code in exposure category

Figure 6.1: Classification of water residence time by zip code by year

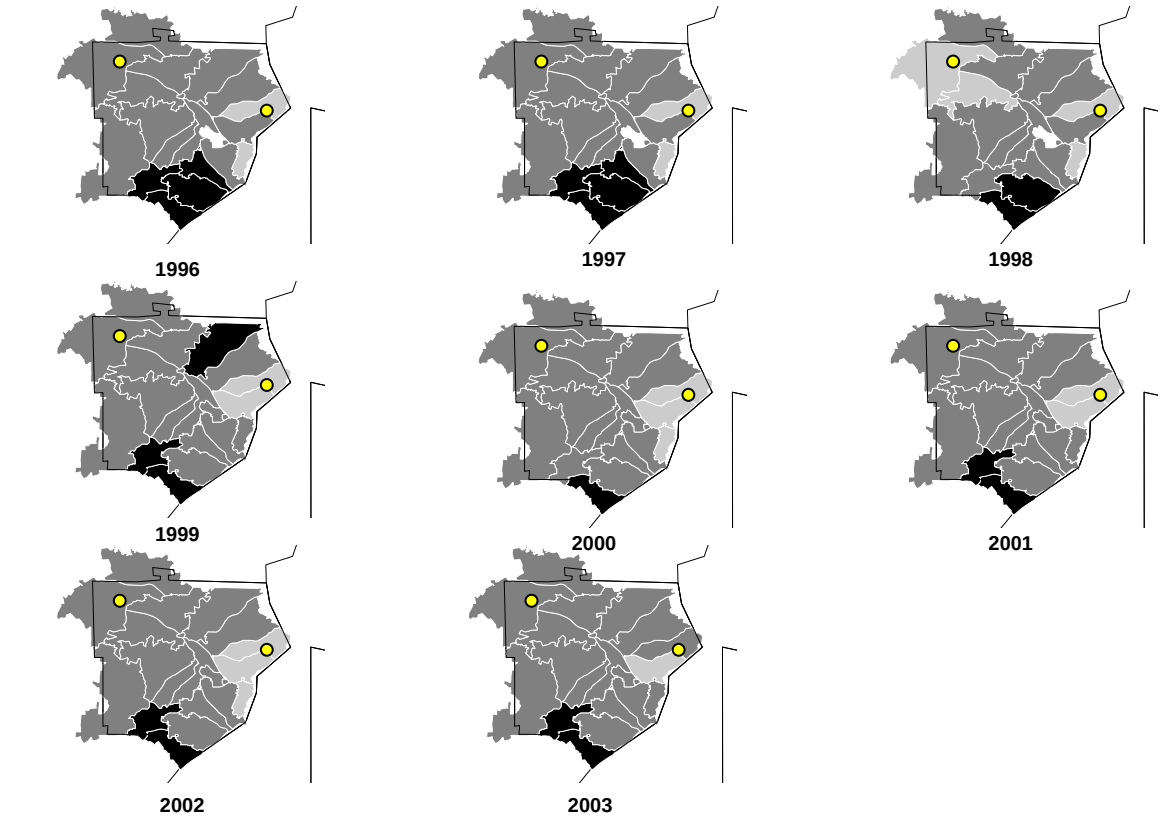
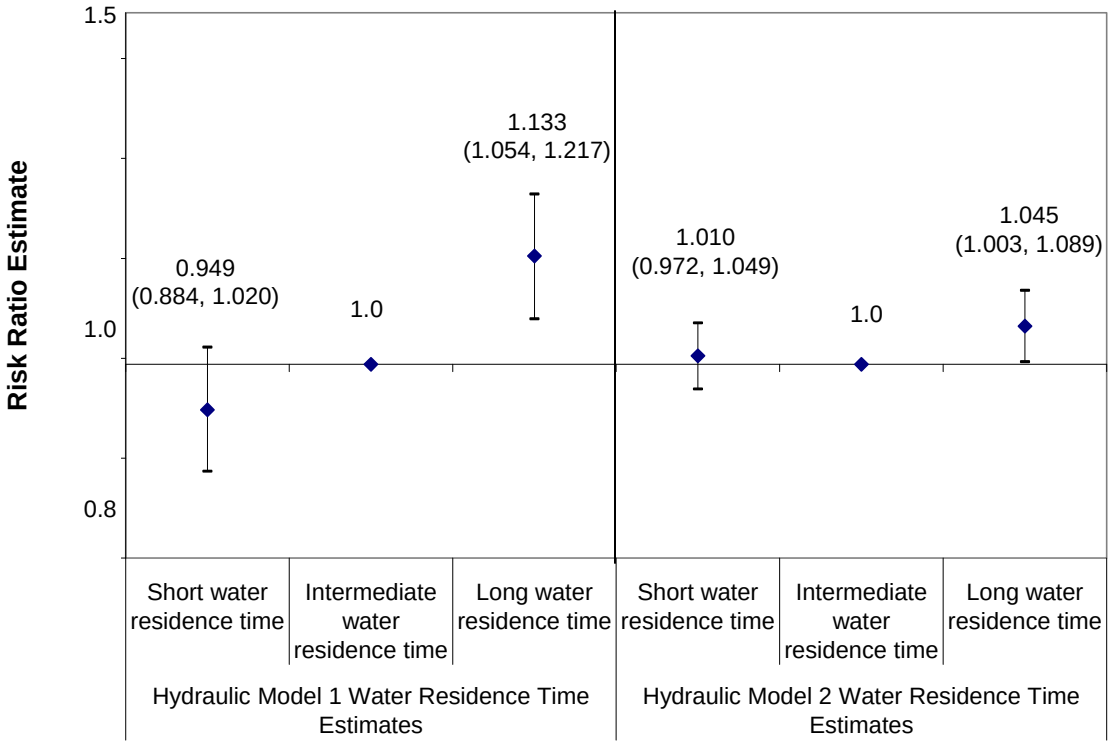
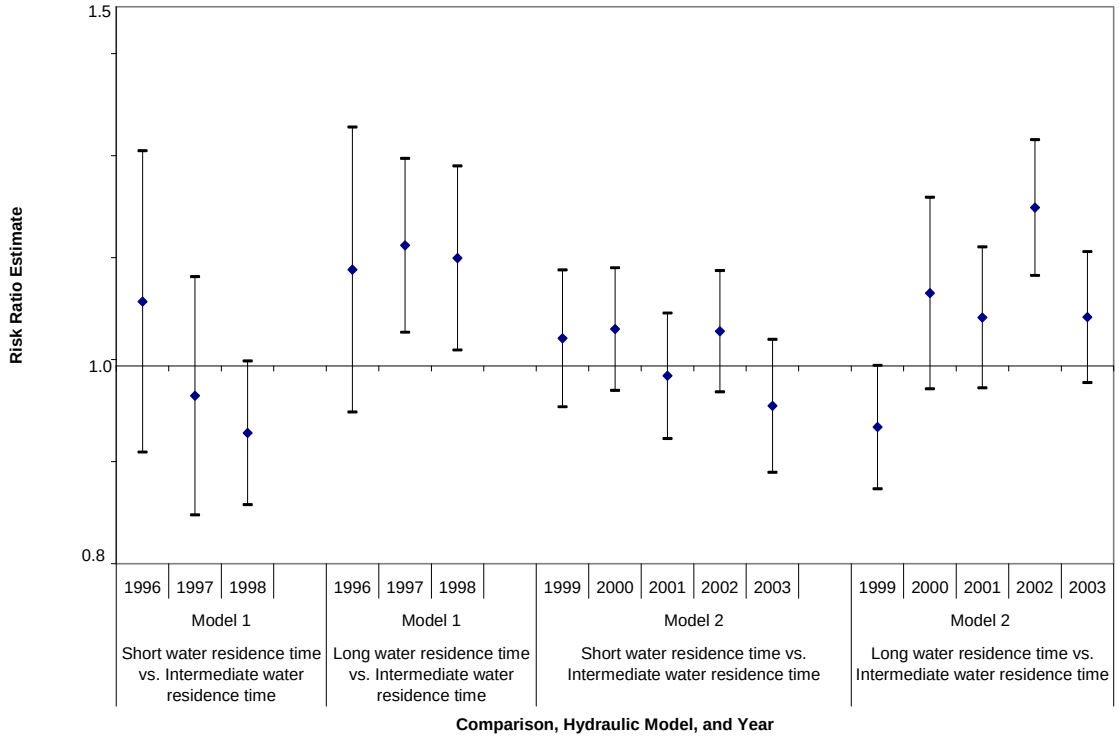


Figure 6.2: Risk ratio estimates[a] for gastrointestinal illness emergency department visits among people living in zip codes with short[b] drinking water residence times and long[b] drinking water residence times compared to intermediate[b] drinking water residence times, Atlanta[c], 1996 – 2003, stratified by hydraulic model[d] used to estimate residence time



[a] Error bars indicate 95 percent confidence intervals.
[b] Short water residence time defined as ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th to 89th percentile; long = ≥ 90th percentile
[c] One county considered
[d] Hydraulic model 1 covers 1996 – 1998, hydraulic model 2 covers 1999 - 2003

Figure 6.3: Risk ratio estimates[a] for gastrointestinal illness emergency department visits among people living in zip codes with short[b] drinking water residence times and long[b] drinking water residence times compared to intermediate[b] drinking water residence times, Atlanta[c], 1996 – 2003, stratified by year



[a] Error bars indicate 95 percent confidence intervals.
[b] Short water residence time defined as ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th to 89th percentile; long = ≥ 90th percentile
[c] One county considered

CHAPTER 7

DRINKING WATER TREATMENT PLANTS AND EMERGENCY DEPARTMENT VISITS FOR GASTROINTESTINAL ILLNESS IN ATLANTA, 1993 – 2004

Tinker, Sarah C.[1]; Tolbert, Paige E.[1]; Moe, Christine L.[1]; Klein, Mitchel[1]; Flanders, W. Dana[1]; Uber, Jim[2]; Amirtharajah, Appiah[3]; Singer, Phillip[4]

[1] Rollins School of Public Health, Emory University, Atlanta, Georgia

[2] University of Cincinnati, Cincinnati, Ohio

[3] Georgia Institute of Technology, Atlanta, Georgia; CH2MHill

[4] University of North Carolina, Chapel Hill, North Carolina

Abstract

As part of an ongoing study of drinking water and health, rates of emergency department visits for gastrointestinal illness were compared among people served by different drinking water treatment plants in Atlanta, Georgia from 1993 to 2004. We obtained information on almost four million emergency department (ED) visits from 28 hospitals in an area served by eight drinking water treatment plants, all of which use surface water. Using Poisson regression, we compared the rates of ED visits for gastrointestinal illness among residents in the treatment plant service areas, controlling for potential confounding factors including age and markers of socioeconomic status. We also examined treatment plant attributes, such as source water type and treatment method, that may contribute to differences in finished water quality and risk of gastrointestinal illness. Our results suggest there were modest differences in the rates of ED visits for gastrointestinal illness among those served by different treatment plants. People living in the service area of one particular plant had a consistently elevated rate of emergency department visits for gastrointestinal illness throughout the study period. This plant had several unique plant attributes, and therefore we were not able to isolate a particular plant attribute associated with this increase. However, this plant was the only one in the study area to utilize direct filtration, which is less effective at removing viral and bacterial contamination. Furthermore, season-specific analyses suggested that gastrointestinal illness rates for people living in this plant's service area were elevated only during cooler months, which is consistent with viral or bacterial transmission. Despite our efforts to control for confounding factors, the potential for residual confounding by spatial covariates in the results is high. However, these results suggest that drinking water from

one treatment plant was associated with a small increased risk of gastrointestinal illness, and support future research into the effectiveness of direct filtration methods.

Introduction

Drinking water treatment has advanced dramatically in the U.S. over the past century leading to a pronounced decrease in waterborne illness and mortality (U.S. Centers for Disease Control and Prevention 1999). Despite these advances, periodic outbreaks of gastrointestinal (GI) illness have been attributed to drinking water and an estimated 8.5 percent of endemic GI illness in the U.S. is caused by drinking water exposure (Messner et al. 2006). Each drinking water treatment facility in the U.S. faces unique challenges based on the characteristics of the source water, distribution system infrastructure, and budget. While many treatment guidelines have been developed for drinking water utilities, the implementations vary, as do the results of treatment processes. Therefore, drinking water quality, and in turn resulting GI illness, may vary across service areas of different plants. As part of a study of drinking water and health in Atlanta, we compared the rate of emergency department (ED) visits for GI illness among residents of the service areas of eight drinking water treatment plants, while controlling for potential confounding factors.

Methods

Emergency Department Visits

Information was available on visits to all of the hospitals operating within the five-county Atlanta area (23 hospitals) and from five hospitals located outside the study area that contributed a substantial number of visits by five-county residents. The information provided by the hospitals included medical record number, date of admission, International Classification of Diseases, 9th Revision (ICD-9), diagnosis

codes, zip code of residence, payment method, and age or date of birth. Data were available for 1993 through 2004, although not all hospitals were able to contribute data for the entire study period. Data from one hospital had to be excluded from the analysis due to missing covariate information.

We defined cases of GI illness *a priori* using the primary and all available secondary ICD-9 diagnostic codes. This case definition included the following diagnoses: infectious GI illness (001-004, 005.0, 005.4, 005.89, 005.9, 006-007, 008.0, 008.42-008.44, 008.47, 008.49, 008.5, 008.6, 008.8, 009), non-infectious GI illness (558.9), and nausea and vomiting plausibly related to GI illness (787.01-787.03, 787.91). Non-infectious GI illness was included in the case definition because previous research has shown that many infectious cases of GI illness are misclassified into this diagnostic category (Gangarosa et al. 1992; Hoxie et al. 1997; Schwartz, Levin, and Hodge 1997).

Service Area Identification

Ten major drinking water treatment plants, operated by six utilities, serve the study area (Table 7.1). These plants are served by a variety of surface water sources, including lakes, rivers, and reservoirs. The production capacity of the plants ranges from ten million gallons per day (mgd) to 150 mgd. The drinking water produced by these plants serves almost four million total customers. The treatment plants generally use conventional treatment methods, including coagulation, sedimentation, filtration, and disinfection. All of the plants use chlorine for disinfection. At one plant, which withdraws from a large lake, a direct filtration method is used to treat the water, where the usual treatment step of sedimentation is omitted. This plant uses ozone as an additional disinfectant. Three other plants use UV disinfection in addition to chlorination. These

plants withdraw from a reservoir that receives some water from land application systems (LAS). The LAS process begins with wastewater that is initially treated using conventional methods. This partially treated water is then applied to specific land areas that drain into the reservoir after natural filtration through the environment.

We determined the service area for each of the drinking water treatment plants through information provided by the utilities. Each zip code in the study area was assigned a treatment plant, when possible. In order to be included in the analysis, 80 percent of the zip code had to be served by a single drinking water treatment plant. Of the 140 zip codes in the study area, 80 (57%) were assigned to a drinking water treatment plant. Fifteen of these zip codes had to be excluded from analysis because of missing covariate information. The service area for two plants could not be separated because of water mixing in the distribution system, and these plants were considered together. One plant was not considered in the analysis because no zip codes could be assigned exclusively to its service.

Other Covariates

We obtained data regarding zip code level demographic characteristics, including median income and percent minority, from the 2000 U.S. Census (U.S. Census Bureau 2000).

Analytic Methods

All analyses were performed using SAS statistical software (SAS 2002 - 2003). We developed Poisson regression models *a priori* to consider the association between the treatment plant serving each zip code and GI illness ED visits among residents of that zip code. We compared the rate of GI illness ED visits among people living in each plant's

service area to the rates of GI illness among people living in all other plants' service areas combined. The models included the following covariates: three-indicators for a four-level categorical age variable (ages 0 to 5, 6 to 18, 19 to 64, and 65+ years); indicator variables for year, season, hospital, distance from zip code centroid to hospital, and distance from zip code centroid to plant; continuous variables for zip code median income and zip code percent minority; and Medicaid payment status (1/0). The model took the following form:

$$\log([E(Y_{z,a,y,s,h,m})] = \text{offset}_{z,a,y,s,h,m} + \alpha + \beta(\text{plant}_z) + \Sigma_{i=1\text{-}3}(\chi_i)(\text{age}_a) + \Sigma_{j=1\text{-}11}(\delta_j)(\text{year}_y) + \Sigma_{k=1\text{-}3}(\varepsilon_k)(\text{season}_s) + \Sigma_{l=1\text{-}25}(\phi_l)(\text{hospital}_h) + \gamma(\text{dis\_w1}_z) + \eta(\text{dis\_w2}_z) + \iota(\text{dis\_h1}_{z,h}) + \varphi(\text{dis\_h2}_{z,h}) + \chi(\text{median income}_z) + \mu(\text{percent minority}_z) + \nu(\text{Medicaid}_m) + \Sigma_{m=1\text{-}3}(\pi_m)(\text{dis\_h1}_{z,h}*\text{age}_a) + \Sigma_{n=1\text{-}3}(\theta_n)(\text{dis\_h2}_{z,h}*\text{age}_a) + \Sigma_{o=1\text{-}25}(\rho_o)(\text{dis\_h1}_{z,h}*\text{hospital}_h) + \Sigma_{p=1\text{-}25}(\sigma_p)(\text{dis\_h2}_{z,h}*\text{hospital}_h) + \Sigma_{q=1\text{-}25}\,\Sigma_{r=1\text{-}3}(\tau_q)(\text{age}_r*\text{hospital}_h)$$

with Y indicating the number of ED visits for GI illness within each zip code (z), age category (a), year (y), season (s), hospital (h), and Medicaid payment group (m). These strata were considered so that we could control for potential confounders and assess effect measure modification. The variables dis_h1 and dis_h2 are indicators for a three-level categorization of the distance from the zip code centroid to the hospital. Similarly, the variables dis_w1 and dis_w2 are indicators for a three-level categorization of the distance from the zip code centroid to the drinking water treatment plant serving that zip code.

An offset, representing the log of the total number of non-injury ED visits in the strata, was included in the model because information on ED visits was not available from all hospitals during the entire study period and because different hospitals had different proportions of ED visits from residents of different zip codes. When data from a

hospital became available in the database, the number of GI illness visits in the strata including that hospital also increased. The offset allowed us to model the proportion of ED visits for GI illness, rather than the absolute number. We also considered the hospital at which the ED visit occurred as a covariate in the model because the proportion of non-injury ED visits for GI illness varied by hospital, more so than could be plausibly explained by drinking water exposure, allowing for confounding if not controlled.

The age distribution of zip codes varied, and the incidence of GI illness, and use of an ED for this illness, differed by age group. Therefore, we controlled for this variable in the analytical model. We also assessed effect measure modification by age group since young children and the elderly may be more susceptible to GI illness given pathogen exposure through drinking water.

Control for socio-economic status (SES) was a concern as it is plausibly related to both the incidence of GI illness and ED usage and it varied spatially. Three variables for control of SES were included in our model. One variable indicated whether the patient paid for the ED visit with Medicaid. The other two variables, derived from U.S. Census data, encoded the median income of the zip code and percent of zip code residents that were of minority (non-Caucasian) race.

We considered the distance from zip code centroid to hospital as a potential confounding factor because results from preliminary descriptive analyses suggested that the rate of GI illness ED visits by people from a given zip code decreases as the distance between that zip code and the hospital increases. This association varied based on age group and hospital, so interaction terms for these variables were also included in the model.

Results from a separate analysis, that assessed the role of the distribution system in contributing to GI illness ED visits, suggested an association of the zip code average estimated water residence time of a zip code (the time taken by water to travel from the treatment plant to the end user) with GI illness ED visits among people living in that zip code. Unfortunately water residence time estimates were available for only two of the treatment plants. Therefore, we controlled for the distance from zip code centroid to treatment plant. The correlation between this distance estimate and the estimated water residence times was high among the zip codes in the service areas of the plants for which both estimates were available.

The exposure, drinking water treatment plant, did not change over time, and so neither season nor year could act as confounding factors in the analysis. These variables were considered in the model in order to assess effect measure modification, because the association between drinking water treatment plant and GI illness ED visits among people served by that treatment plant could plausibly have varied by season or by year.

Results

Twenty-eight hospitals provided data on 7,274,275 ED visits in the five-county Atlanta area. Once we restricted the database to non-injury visits from residents of zip codes for which a single treatment plant of service could be assigned and for which all covariate information was available, there remained 2,714,822 ED visits, 230,962 of them (8.5%) for GI illness. There were no reported drinking water-related disease outbreaks in Atlanta during the study period.

Before assessing if rates of GI illness differ among people served by different treatment plants, we assessed whether the treatment plant of service contributed to the predicted rate of GI illness ED visits. If the treatment plant of service, considered as an eight-level categorical variable, did not make a statistically significant contribution to the analytical model then chance was a more likely explanation for any statistically significant differences in rates of GI illness we may have observed among people served by different treatment plants. We conducted a likelihood ratio test and the results support the conclusion that the treatment plant of service did contribute to the risk of GI illness ED visits ($p < 0.0001$) in this population.

The results suggest that GI illness ED visits constituted a greater proportion of all non-injury ED visits among people served by Plant F compared to people served by other plants (Rate Ratio (RR): 1.054, 95% Confidence Interval (CI): 1.032 – 1.077) (Table 7.1). The results also suggest that people served by Plants A and E had a lower rate of GI illness ED visits compared to the other plants (RR: 0.982, 95% CI: 0.964 – 1.000; RR: 0.977, 95% CI: 0.958 – 0.996, respectively).

The results of models stratified by year suggest that the relative rates of GI illness ED visits among people served by different treatment plants were not constant over time (Figure 7.1). There were lower relative rates of GI illness ED visits among people served by Plant B for the first several years of the study, but the relative rates increased by the end of the study period. Conversely, the relative rate of GI illness ED visits steadily decreased for people served by Plant G over time. The rates of GI illness ED visits for people served by Plants D and E remained relatively constant throughout the study

period. While the relative rates of GI illness ED visits for people served by Plant F fluctuated over time, they were generally positive.

There was some variation in rate ratio estimates when models were stratified by age category (Figure 7.2). The strongest associations between treatment plant of service and GI illness ED visits were observed for young children (age 0 to 5 years) and the elderly (65+ years). The rate ratio estimates were consistently elevated for all age groups among people served by Plant F. The negative rate ratio estimates observed for people served by Plants A and E appeared to be restricted to young children. The season-specific analyses suggested that the elevated rate of GI illness ED visits among people served by Plant F was restricted to the autumn and winter (Figure 7.3).

Discussion

The overall results of our analysis show modest, but statistically significant, variations in rates of GI illness ED visits across the service areas of different treatment plants. While GI illness rates are expected to vary geographically, even in the absence of drinking water pathogen exposure, we attempted to control for factors which might explain this variation in order to isolate the association with drinking water treatment plant of service. Residual spatial confounding likely impacted our results; however, they are consistent with the possibility that some acute GI illness experienced by people residing in the service areas of certain drinking water treatment plants may be attributable to their drinking water.

In general, Atlanta's water sources are environmentally challenged and treatment is necessary to ensure a safe supply of drinking water for the metropolitan area. The

sources of water serving the city include two lakes, at least four rivers, including the Chattahoochee River, as well as creeks, reservoirs, and LAS. The treatment methods utilized by the plants are all in keeping with EPA regulations. The distribution systems delivering the water treated at the plants vary in age, complexity, and size. Each distribution system has a unique risk profile for recontamination of the treated water.

The results of these analyses suggest an elevated rate of GI illness ED visits for people served by Plant F compared to the other seven treatment plants. This treatment plant has several unique attributes, including the use of ozone as a disinfectant. However, one attribute of the plant may at least partially explain the excess proportion of GI illness observed among people in the service area. This plant is the only plant included in the study to use direct filtration, meaning that the usual step of sedimentation is omitted. The elevated rate ratios observed for people served by this treatment plant were restricted to the autumn and winter months; viral and bacterial transmission is heightened during these months in the U.S. due to greater survival of viral and bacterial organisms in colder temperatures (Flint 1987; Rzezutka and Cook 2004).

We attempted to identify other treatment plant attributes that might explain the relative rates of GI illness among people served by different treatment plants, particularly the negative association observed among people served by Plants A and E. However, because only eight plants were considered in the analysis and many traits were similar across plants, this assessment had limited power (See Appendix B).

Waterborne pathogens do not affect all age groups equally. In general, the young and elderly are more susceptible to severe disease because of underdeveloped or ailing immune systems and, in young children, naivety to the organisms, such as pathogenic *E.*

coli (Glass et al. 1991; Jin et al. 1996). Increased susceptibility may partially explain the stronger relative rates observed among children and the elderly. Another contributing factor may be the accuracy of the treatment plant assignment among these age groups compared to non-elderly adults. Young children and the elderly are more likely to drink water during the day produced by the same treatment plant as the one serving their residential zip code, due to school attendance and less mobility, while adults of working age are expected to travel more often outside their residential drinking water service area during the work week. The ED database only provided information on the zip code of residence and therefore we were unable to account for this type of misclassification in the analysis, which would likely bias our results toward the null.

We were not able to include all of the zip codes in the study area because of the requirement that the majority of the zip code be served by one treatment plant (See Appendix B). Even after excluding zip codes to promote exposure homogeneity, exposure misclassification likely still occurred, either because the subject resided in the small section of the zip code served by another plant or because of drinking water exposure outside the zip code of residence, as discussed above. We examined the impact of excluding 15 of the zip codes due to missing Census data, and determined that the impact was minimal compared to the importance of including the Census information in the model (See Appendix B).

While our database only captured GI illness severe enough to lead to an ED visit, a small proportion of all cases of infectious GI illness, these are arguably the cases of greatest interest. Our study had the power to detect modest associations due to the

tremendous size of the ED database. The participation of all of the drinking water utilities serving the study area allows for greater generalizability to other metropolitan areas.

Although we attempted to control for major confounding factors, residual confounding likely impacted our results. The potential for spatial confounding in this analysis was great because the exposure, drinking water treatment plant of service, was assigned based on zip code, and the rates of GI illness were expected to vary across zip codes due to many factors other than drinking water pathogen exposure. A particularly difficult factor to attempt to control for in this analysis was socioeconomic status (SES). We attempted to account for SES using three variables: zip code median income, zip code percent minority, and whether or not the patient paid for the ED visit using Medicaid. Zip code median income and percent minority, derived from 2000 U.S. Census data, described the community context from which the cases of GI illness arose, and these variables were associated with GI illness ED visits rates in preliminary analyses. However, these variables did not necessarily describe the SES of the individuals visiting the ED for GI illness. Medicaid payment status is an individual-level SES characteristic. However, it is also a crude marker of SES; more children qualify and utilize Medicaid to pay for healthcare than adults. Furthermore, people age 65 and older qualify for Medicare and would thus have no need to use Medicaid.

There are likely other spatially-varying factors that we were not able to consider that may be influencing our results. Outbreaks of GI illness, due to causes other than drinking water, also had the potential to confound these analyses. While there were no reported outbreaks of GI illness associated with drinking water during the study period, there were reported outbreaks of GI illness due to both food and recreational water

exposure. Small foodborne outbreaks were unlikely to impact our results because they resulted in few ED visits and were likely to occur with similar frequency throughout the study area. Larger outbreaks, such as a 2003 hepatitis A outbreak attributed to green onion exposure, affected all counties considered in the study area, because the food products associated with these outbreaks were widely distributed (Gabel and Wolthuis 2004). There were slightly more hepatitis A cases from the 2003 green onion outbreak residing in the service area of Plant A, but the results stratified by year (Figure 7.1) suggest they did not have a large impact on our results.

An outbreak of pathogenic *E. coli* infection at a water park located in the service areas of Plants G and H occurred in June 1998 (Gilbert and Blake 1998). Despite the attention this outbreak received, due largely to the severity of the illnesses (one child died), ultimately only 26 cases of infection were confirmed. The water park draws visitors from around the region, and while there may have been a higher proportion of visitors from the immediate area, many of the cases were from the surrounding counties, and 35 percent of the cases were visitors from outside of Georgia. The results of our study, stratified by year (Figure 7.1), suggest that this outbreak had little impact on the results. Although the rate ratio estimate for GI illness ED visits among people served by Plant G compared to people served by other plants suggested a positive association in 1998, the rate ratio estimate was consistent in magnitude with the rate ratio estimates of neighboring years and the 1998 rate ratio estimate was consistent with the overall downward trend of rate ratios over time observed among people served by this plant. A positive rate ratio for GI illness ED visits was also observed for people served by Plant H

compared to people served by other plants in 1998; however, the rate ratio estimates for neighboring years were even higher.

Another example of a spatial covariate with the potential to confound this analysis is day care usage. Day care attendance is a known risk factor for GI illness among children and their close contacts (Lu et al. 2004). Although we controlled for age group in the analysis, day care use among children is not necessarily equivalent across zip codes.

Although our results suggest modest rate ratios for GI illness ED visits among people served by different treatment plants, the size of the exposed population is large and therefore the excess number of GI illness ED visits potentially attributable to drinking water exposure is not trivial. The results of this analysis support further examination of the modest differences in acute GI illness incidence among people served by different treatment plants that may be attributable to their drinking water. Particular attention should be focused on plants using direct filtration.

References

Flint, K. P. 1987. The long-term survival of *Escherichia coli* in river water. *Appl Bacteriol* 63 (3):261-70.

Gabel, J. A., and J. S. Wolthuis. 2004. Multi-restaurant outbreak of hepatitis A associated with green onions in central and north Georgia, September - October 2003. *Georgia Epidemiology Report* 20 (10):1-3.

Gangarosa, R. E., R. I. Glass, J. F. Lew, and J. R. Boring. 1992. Hospitalizations involving gastroenteritis in the United States, 1985: the special burden of the disease among the elderly. *Am J Epidemiol* 135 (3):281-90.

Gilbert, L., and P. Blake. 1998. Outbreak of *Escherichia coli* O157:H7 infections associated with a water park. *Georgia Epidemiology Report* 14 (7):1-2.

Glass, R. I., J. F. Lew, R. E. Gangarosa, C. W. LeBaron, and M. S. Ho. 1991. Estimates of morbidity and mortality rates for diarrheal diseases in American children. *J Pediatr* 118 (4.2):S27-33.

Hoxie, N. J., J. P. Davis, J. M. Vergeront, R. D. Nashold, and K. A. Blair. 1997. Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *Am J Public Health* 87 (12):2032-5.

Jin, S., P. E. Kilgore, R. C. Holman, M. J. Clarke, E. J. Gangarosa, and R. I. Glass. 1996. Trends in hospitalizations for diarrhea in United States children from 1979 through 1992: estimates of the morbidity associated with rotavirus. *Pediatr Infect Dis J* 15 (5):397-404.

Lu, N., M. E. Samuels, L. Shi, S. L. Baker, S. H. Glover, and J. M. Sanders. 2004. Child day care risks of common infectious diseases revisited. *Child Care Health Dev* 30 (4):361-8.

Messner, M., S. Shaw, S. Regli, K. Rotert, V. Blank, and J. Soller. 2006. An approach for developing a national estimate of waterborne disease due to drinking water and a national estimate model application. *J Water Health* 4 (Suppl 2):201-40.

Rzezutka, A., and N. Cook. 2004. Survival of human enteric viruses in the environment and food. *FEMS Microbiol Rev* 28 (4):441-53.

SAS 9.1. SAS Institute, Inc., Cary, North Carolina.

Schwartz, J., R. Levin, and K. Hodge. 1997. Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* 8 (6):615-20.

U.S. Census Bureau. 2000. General Housing Characteristics: 2000. Census Summary File 1 (SF 1) 100-Percent Data. Washington, DC: American FactFinder, U.S. Census Bureau.

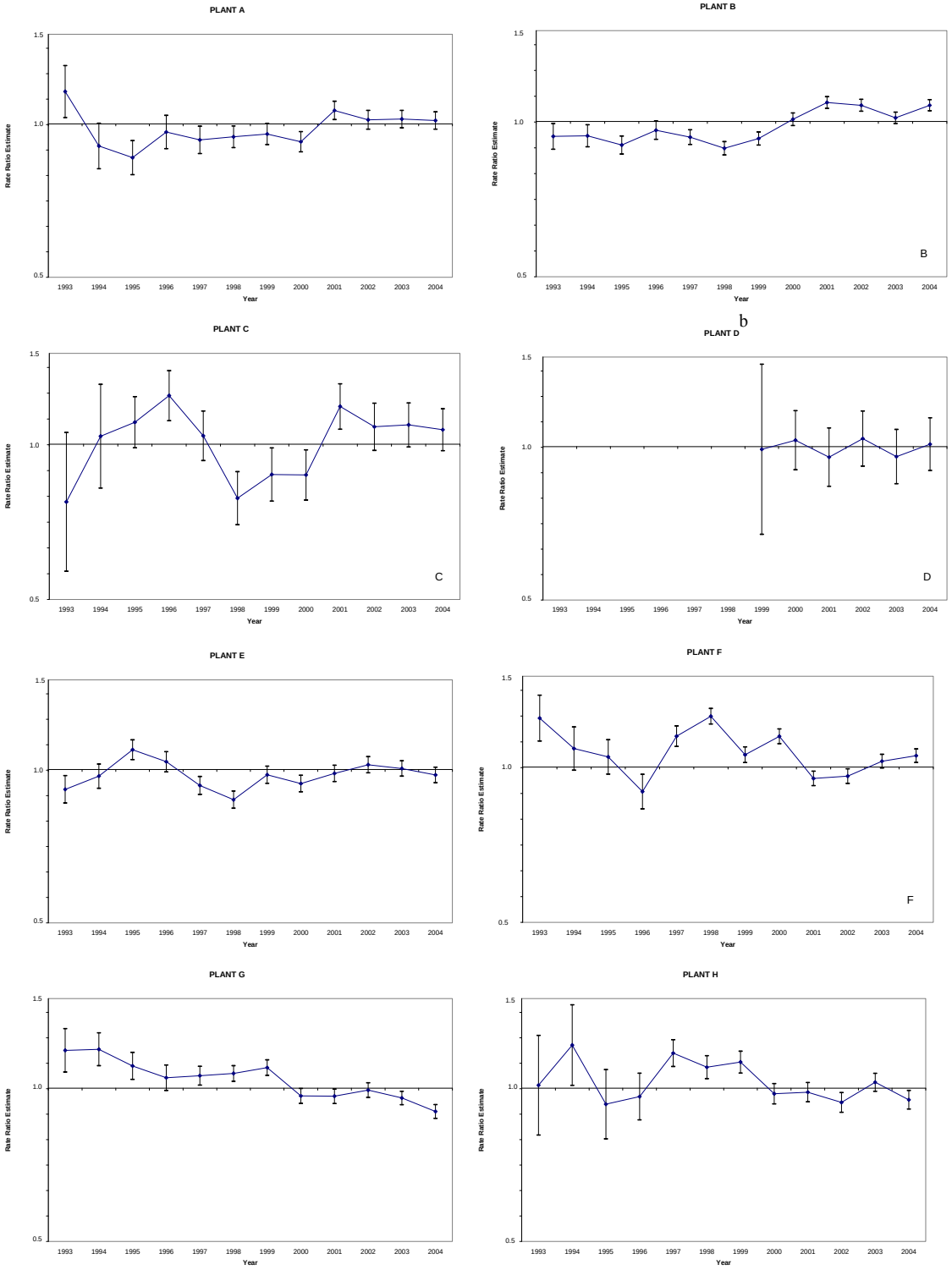
U.S. Centers for Disease Control and Prevention. 1999. From the U.S. Centers for Disease Control and Prevention. Control of infectious diseases, 1900-1999. *JAMA* 282 (11):1029-32.

Table 7.1: Adjusted[a] rate ratio estimates for emergency department visits for gastrointestinal illness among people served by one drinking water treatment plant compared to all other treatment plants and attributes of the drinking water treatment plants, five-county metro-Atlanta, 1993-2004

Plant	Rate Ratio[b]	Production Capacity	Source Water Type	Treatment Method[c]	Disinfectant
A	0.982 (0.964, 1.000)	90 mgd[d]	River	Standard	Chlorine
B	0.993 (0.980, 1.006)	128 mgd	River	Standard	Chlorine
C[e]	1.027 (0.985, 1.072)	65 mgd	River	Standard	Chlorine
D[f]	0.996 (0.904, 1.096)	20 mgd / 10 mgd	Creek-fed reservoir and LAS[g]	Standard	Chlorine, UV
E[e]	0.977 (0.958, 0.996)	137 mgd	River	Standard	Chlorine
F	1.054 (1.032, 1.077)	150 mgd	Lake	Direct filtration	Chlorine, ozone
G[e]	1.001 (0.986, 1.015)	86mgd	River	Standard	Chlorine
H[e]	1.015 (0.993, 1.036)	72 mgd	Lake	Standard	Chlorine

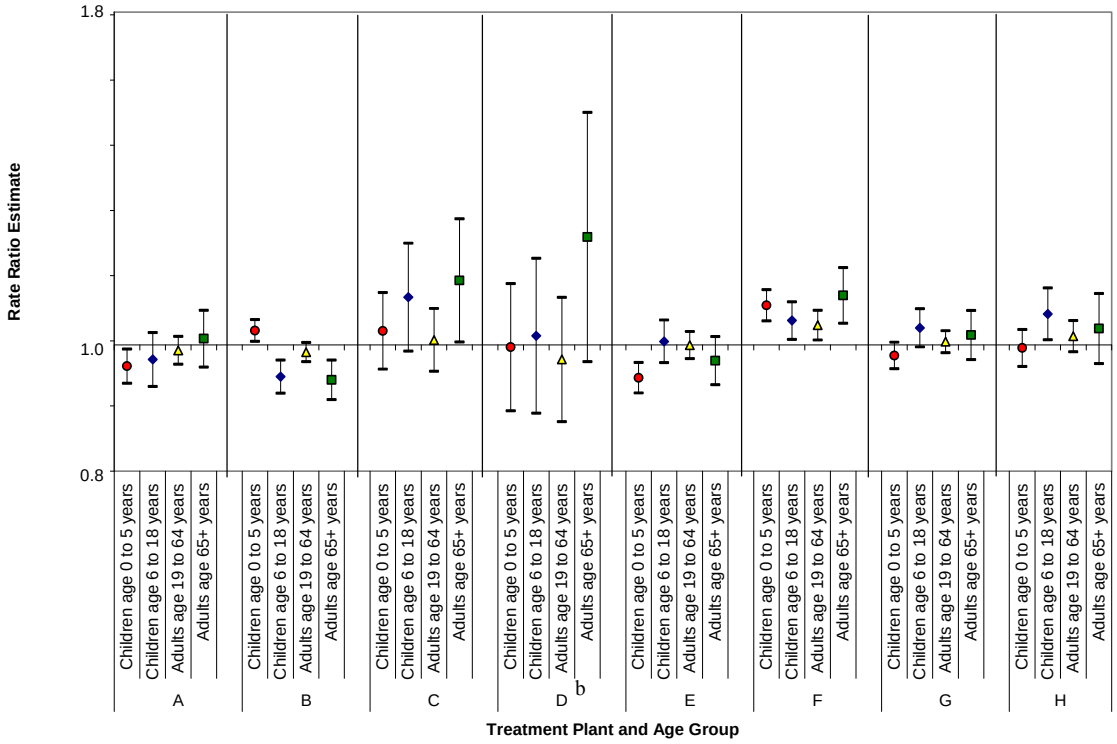
[a] Adjusted for age group, hospital, Medicaid payment status, zip code median income, zip code percent minority, distance from zip code to hospital, distance from zip code to treatment plant
[b] Comparing that plant to all other plants
[c] Standard treatment: Coagulation, flocculation, sedimentation, filtration, disinfection
[d] Million gallons per day
[e] Plants C and E are operated by the same utility; Plants G and H are operated by the same utility
[f] Two plants considered together because their water was mixed prior to distribution
[g] Land application system

Figure 7.1: Yearly rate ratio estimates[a] for emergency department visits for gastrointestinal illness among people served by one drinking water treatment plant compared to all other treatment plants, five-county metro-Atlanta, 1993-2004



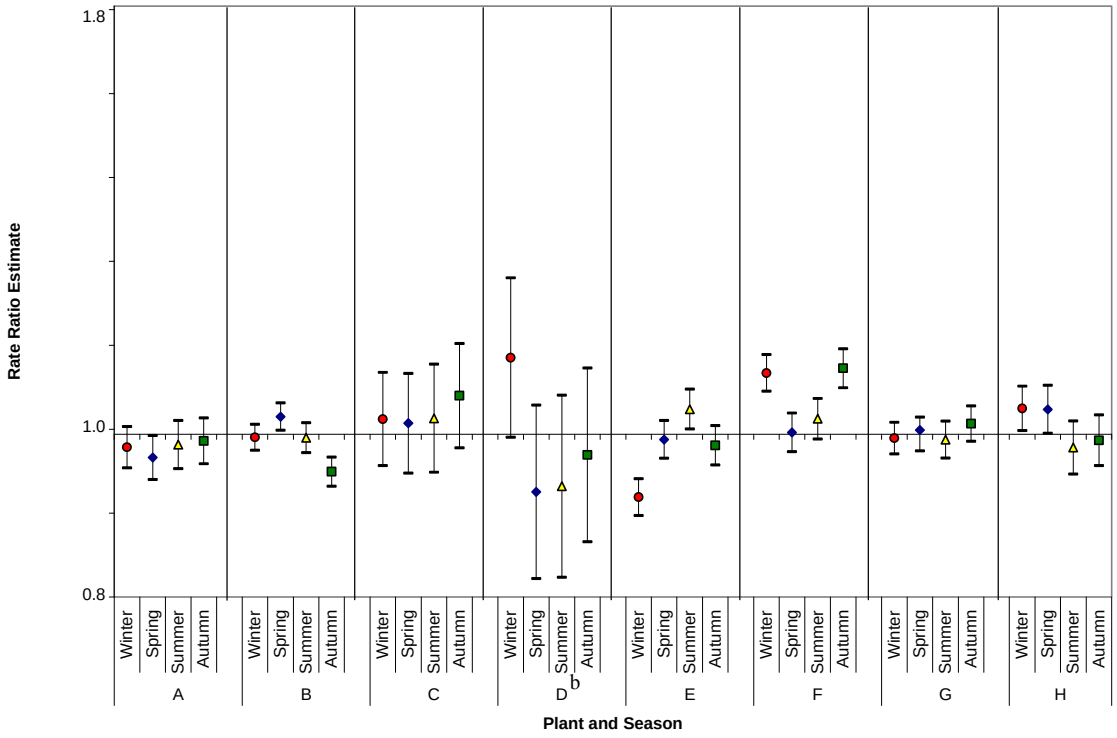
[a] Error bars indicate 95 percent confidence intervals.
[b] Plant D began operation in November 1999.

Figure 7.2: Age group-specific rate ratio estimates[a] for emergency department visits for gastrointestinal illness among people served by one drinking water treatment plant compared to all other treatment plants, five-county metro-Atlanta, 1993-2004



[a] Error bars indicate 95 percent confidence intervals.
[b] Plant D began operation in November 1999.

Figure 7.3: Season-specific rate ratio estimates[a] for emergency department visits for gastrointestinal illness among people served by one drinking water treatment plant compared to all other treatment plants, five-county metro-Atlanta, 1993-2004



[a] Error bars indicate 95 percent confidence intervals.
[b] Plant D began operation in November 1999.

CHAPTER 8

DRINKING WATER TURBIDITY AND EMERGENCY DEPARTMENT VISITS FOR GASTROINTESTINAL ILLNESS IN ATLANTA, 1993 – 2004

Tinker, Sarah C.[1]; Tolbert, Paige E.[1]; Moe, Christine L.[1]; Klein, Mitchell[1]; Flanders, W. Dana[1]; Uber, Jim[2]; Amirtharajah, Appiah[3]; Singer, Phillip[4]

[1] Rollins School of Public Health, Emory University, Atlanta, Georgia

[2] University of Cincinnati, Cincinnati, Ohio

[3] Georgia Institute of Technology, Atlanta, Georgia; CH2MHill

[4] University of North Carolina, Chapel Hill, North Carolina

Abstract

Despite major advances in drinking water treatment and delivery, infectious disease can still be transmitted through drinking water in the U.S. Using Poisson time-series methods, we examined the relationship between turbidity levels of raw and filtered surface water measured at eight major drinking water treatment plants in Atlanta, Georgia and emergency department (ED) visits for gastrointestinal (GI) illness during 1993-2004 among the population served by these plants. The median daily average and maximum filtered water turbidity levels were 0.06 and 0.09 nephelometric turbidity units (NTU), respectively, and the median daily minimum and maximum raw water turbidity levels were 4.2 and 6.2 NTU, respectively. We analyzed 4,179,340 ED visits, of which 240,925 were for GI illness. Overall, the results were consistent with no association of filtered water turbidity with ED visits for GI illness. We observed a modest association of a 10 NTU average increase in raw water turbidity over 21 days and increased GI illness ED visits. Results by plant and age group displayed heterogeneity. Our results suggest that source water quality modestly contributes to endemic GI illness in the study area. The association between turbidity and GI illness ED visits was only observed when raw water turbidity was considered, however; filtered water turbidity may not serve as a reliable indicator of modest pathogen risk at all treatment plants.

Introduction

Treated water quality in the United States is among the best in the world. Disease due to waterborne infectious organisms decreased dramatically during the 20th century with the implementation of drinking water treatment practices including filtration and disinfection (U.S. Centers for Disease Control and Prevention 1999). Despite the enormous resources dedicated to keeping the U.S. water supply safe, pathogenic organisms are present in source water used for drinking water, and previous research supports a role for drinking water in contributing to endemic gastrointestinal (GI) illness.

The first major randomized-controlled trials to examine this question, in which families were followed for development of GI illness after being randomized to receive conventionally treated drinking water or a home-unit providing additional treatment, found that up to 40% of GI illness could be attributed to drinking water exposure (Payment et al. 1991; Payment et al. 1997). Other randomized-controlled trials failed to find an association between drinking water exposure and GI illness (Hellard et al. 2001; Colford Jr. et al. 2005).

Evidence from observational studies also contributes to the collective knowledge regarding the burden of waterborne GI illness. One method which has been employed to examine this association is the time-series analysis, in which variation in water quality over time is examined in relation to variation in disease occurrence through quantified healthcare utilization. There were two studies using this design conducted in Milwaukee in response to the large 1993 *Cryptosporidium* outbreak (Morris et al. 1996; Morris, Naumova, and Griffiths 1998). Both found an association between turbidity, a measure of the cloudiness of the water and a rough proxy for microbial contamination, and hospital

utilization for GI illness, even before the outbreak period. In 1997 Schwartz, et al. published results of an investigation of drinking water turbidity in relation to hospital visits for GI illness among children in Philadelphia (Schwartz, Levin, and Hodge 1997). In 2000, the same group published results of a similar study considering the elderly population in Philadelphia (Schwartz, Levin, and Goldstein 2000). Both studies found significant positive associations between turbidity and healthcare utilization for GI illness.

In this study we utilized time-series methods to examine the association between raw and filtered water turbidity at ten drinking water treatment plants serving the five-county metropolitan Atlanta area and emergency department (ED) visits for GI illness over a 12 year period. Atlanta is a well-suited location in which to conduct this type of study because of the multiple utilities serving the area and the challenged raw water sources.

Methods

Environmental Information

There were ten major drinking water treatment plants, operated by six utilities, serving the study area. These plants used a variety of surface water sources, including lakes and rivers. The production capacity of the plants ranged from 10 million gallons per day (mgd) to 150 mgd. The drinking water produced by these plants served almost four million total customers. The treatment plants generally used conventional treatment methods, including coagulation, sedimentation, filtration, and disinfection. All of the plants used chlorine disinfection. One plant did not use sedimentation. This plant used

ozone as an additional disinfectant. Three plants used UV disinfection in addition to chlorination.

We obtained water quality data directly from the drinking water utilities. These data included daily summary measures of hourly turbidity measurements in nephelometric turbidity units (NTU) taken at the treatment plant. Daily values of average and maximum filtered water turbidity were available for July 1, 1993 through December 31, 2004, except for one plant for which maximum filtered turbidity data were available only from January 1, 2000 to December 31, 2004 and another plant that did not begin operation until November 1999. Daily values of minimum and maximum raw water turbidity were available for January 1, 2002 through December 31, 2004 from all plants except one, for which data availability began March 1, 2002.

We determined the service area of each of the drinking water treatment plants through information provided by the utilities. We assigned a treatment plant to each zip code in the study area, where possible. In order to be included in the analysis, 80 percent of the zip code had to be served by a single drinking water treatment plant. Of the 140 zip codes in the study area, we were able to assign 81 (58%) to a single drinking water treatment plant. We combined the data from two of the plants using a weighted average, based on water production, because their water was mixed prior to distribution. We excluded data from another plant because no zip codes could be assigned exclusively to its service.

Average temperature (average of the daily minimum and maximum) was measured at Hartsfield – Jackson Atlanta International Airport and obtained from the National Climatic Data Center Network.

Emergency Department Visits

Information was available on emergency department visits from all of the hospitals operating within the five-county Atlanta area (23 hospitals) and from five hospitals located outside the study area that contributed a substantial number of visits by five-county residents. The data provided by the hospitals included medical record number, date of admission, International Classification of Diseases, 9th Revision (ICD-9), diagnosis codes, zip code of residence, and age or date of birth.

The *a priori* case definition for GI illness used the primary and all available secondary ICD-9 diagnostic codes. The case definition included the following diagnoses: infectious GI illness (001-004, 005.0, 005.4, 005.89, 005.9, 006-007, 008.0, 008.42-008.44, 008.47, 008.49, 008.5, 008.6, 008.8, 009), non-infectious GI illness (558.9), and nausea and vomiting plausibly related to GI illness (787.01-787.03, 787.91). We included non-infectious GI illness in the case definition because previous research has shown that many infectious cases of GI illness are misclassified into this diagnostic category (Gangarosa et al. 1992; Hoxie et al. 1997; Schwartz, Levin, and Hodge 1997).

Analytic Methods

All data analyses were performed using SAS statistical software (SAS 2002 - 2003). We examined univariate statistics for each of the turbidity variables and the daily count of ED visits. We used Spearman rank correlation statistics to assess the correlation between turbidity measures on the same day.

We developed *a priori* analytical models to control for meteorology and long-term time trends. We ran separate Poisson generalized linear models (GLMs) for each

plant and calculated a weighted average of the plant-specific rate ratio estimates based on the inverse of the variance. The basic model had the following form:

$$\text{Log}[E(Y_t)] = \alpha + \Sigma_{i=0\text{-}20}\beta_i \textit{turbidity}_{t\text{-}i} + \Sigma_{k=1\text{-}7}(\chi_k \textit{day-of-week}_t) + \Sigma_{m=1\text{-}20}(\delta_m \textit{hospital}_t) + \varepsilon_1 \textit{temp}_{t\text{-}(t\text{-}6)} + \varepsilon_2(\textit{temp}_{t\text{-}(t\text{-}6)})^2 + \varepsilon_3(\textit{temp}_{t\text{-}(t\text{-}6)})^3 + \varepsilon_4 \textit{temp}_{(t\text{-}7)\text{-}(t\text{-}13)} + \varepsilon_5(\textit{temp}_{(t\text{-}7)\text{-}(t\text{-}13)})^2 + \varepsilon_6(\textit{temp}_{(t\text{-}7)\text{-}(t\text{-}13)})^3 + \varepsilon_7 \textit{temp}_{(t\text{-}14)\text{-}(t\text{-}20)} + \varepsilon_8(\textit{temp}_{(t\text{-}14)\text{-}(t\text{-}20)})^2 + \varepsilon_9(\textit{temp}_{(t\text{-}14)\text{-}(t\text{-}20)})^3 + g(\gamma_1, \ldots, \gamma_N; \textit{time})$$

where Y indicates the number of ED visits for GI illness in a given treatment plant's service area on day t. We considered each of the four available turbidity exposure variables, average and maximum filtered water turbidity and minimum and maximum raw water turbidity, in separate models using a 21-day unconstrained distributed lag encompassing turbidity on the same day as the ED visits and on the preceding 20 days. We also included indicator variables for day-of-week, with a separate category included for federal holidays (*day-of-week*), in the models. We included a variable to indicate hospital entry and exit (*hospital*) because not all hospitals were able to contribute ED data for the entire study period. We included temperature in the model using a moving average of the first (days t through t-6), second (days t-7 through t-13), and third (days t-14 through t-20) weeks preceding the day of the ED visit. We included quadratic and cubic terms to allow for greater flexibility in our control for average temperature. We controlled for long-term time trends using cubic splines with seasonal knots for spring, summer, and autumn, and monthly knots for winter [$g(\gamma_1, \ldots, \gamma_N; \textit{time})$]. Monthly knots were used in winter to better control for the potential confounding impact of GI illness spikes caused by rotavirus. The cubic splines were defined as follows:

$$g(\gamma_1,\ldots,\gamma_N;x) = \gamma_1 x + \gamma_2 x^2 + \gamma_3 x^3 + \sum_{j=4} \gamma_j w_j(x),$$

where $\gamma_1, \gamma_2, \ldots \gamma_N$ were parameters to be estimated and where $w_j(x) = (x-\tau_j)^3$ if $x \geq \tau_j$, and $w_j(x) = 0$ otherwise. The first and second derivatives of *g(x)* are continuous, allowing time trends to be modeled as a smooth function. To avoid collinearity, we linearly transformed the cubic spline terms by multiplying the design matrix by the eigenvectors of its variance-covariance matrix. We scaled all variance estimates to account for Poisson overdispersion. The rate ratios derived from the models described the change in the rate of GI illness ED visits for an average increment of 0.1 NTU for filtered water turbidity and 10 NTU for raw water turbidity over each 21-day period. We calculated 95 percent confidence intervals for all rate ratio estimates.

We conducted a number of secondary analyses, including age-specific models and models considering the turbidity exposure as a series of 3-day moving averages encompassing day t through day t-20. Additional sensitivity analyses are discussed in Appendix C.

This study received approval from the Emory University Institutional Review Board and was conducted in accordance with the Common Rule.

Results

There were no major EPA violations by any of the treatment plants during the study period. The mean filtered and raw water turbidity levels varied by treatment plant (Table 8.1). The mean average filtered water turbidity ranged from 0.03 NTU to 0.17 NTU and the mean maximum filtered water turbidity ranged from 0.04 NTU to 0.29 NTU. The highest hourly filtered water turbidity measurement observed during the study

period was 4.0 NTU. The mean minimum raw water turbidity during the study period ranged from 1.1 NTU to 16.3 NTU and the mean maximum raw water turbidity ranged from 1.5 NTU to 55.0 NTU. The highest hourly raw water turbidity measurement seen during the study period was 1,984 NTU. While filtered water turbidity measures varied little by season, raw water turbidity levels peaked in winter. The average and maximum filtered water turbidity measures were highly correlated ($r = 0.91$), as were the minimum and maximum raw water turbidity measures ($r = 0.95$). The filtered and raw water turbidity measures showed little correlation with each other (all $r < |0.1|$).

Twenty-eight hospitals provided data on 7,642,118 ED visits in the five-county Atlanta area. After restricting the database to those zip codes for which a single treatment plant of service could be assigned, there remained 4,179,340 ED visits, of which 240,925 (5.8%) were for GI illness. The average daily number of ED visits varied by treatment plant service area (Table 8.2). Counts of GI illness varied markedly by season, with large winter peaks, particularly among children. The highest number of GI illness ED visits occurred on Sundays and the lowest number occurred on Fridays. There were no reported drinking water-related disease outbreaks in Atlanta during the study period.

There was little correlation of the residuals for observations representing neighboring dates in the *a priori* models using GLM procedures (all $r < 0.07$) indicating the splines were accounting for any autocorrelation.

Filtered Water Turbidity

The rate ratio estimates suggested no association between changes in turbidity and counts of GI illness ED visits for either filtered water turbidity measure (Figure 8.1). The results from the plant-specific models displayed some heterogeneity (Figures 8.2 and

8.3). While the majority of the rate ratio estimates were consistent with the null, those for Plant G suggested a positive association of a 0.1 NTU average increase in filtered water turbidity over the previous 21 days with ED visits for GI illness. The rate ratio estimate when the average filtered water turbidity exposure was considered for Plant G was 1.68 (95% confidence interval (CI) = 1.26 – 2.24) and was 1.26 (95% CI = 1.07 – 1.47) when the maximum filtered turbidity exposure was considered.

Raw Water Turbidity

The rate ratio estimates when the raw water turbidity exposures were considered suggested a modest increase in the rate of GI illness ED visits with a 10 NTU average increase in turbidity over the 21-day lag period (Figure 8.1). The rate ratio estimate when minimum raw water turbidity was considered was 1.06 (95% CI = 1.04 – 1.08) and when maximum raw water turbidity was considered the rate ratio estimate was 1.02 (95% CI = 1.01 – 1.03). The results from the plant-specific models also displayed heterogeneity with the raw water turbidity exposure (Figures 8.4 and 8.5). Six of the eight point estimates, for both minimum and maximum raw water turbidity, were positive; those for Plants B, E, and G were statistically significant ($\alpha = 0.05$).

Secondary Analyses

The results of the age-specific models for which filtered water turbidity measures were considered were generally consistent with the null (Figure 8.6). The majority of the rate ratio estimates from the models for which raw water turbidity was considered were positive. The estimates for children age five years and younger were markedly stronger and were the only estimates that were statistically significant.

The results from the models in which turbidity was considered as a series of consecutive three-day moving averages, ranging from days t, t-1, and t-2 through days t-18, t-19, and t-20, were consistent with the *a priori* results (Figure 8.7). All rate ratio estimates from models for which filtered water turbidity was used as the exposure were consistent with the null. All of the rate ratio estimates from models for which raw water turbidity was used as the exposure were positive and the majority were statistically significant. The strongest associations were observed for lags incorporating turbidity measures approximately days 6 through 9 prior to the date of the ED visits. Similar results were observed when the maximum raw water turbidity measure was considered as the exposure, although the rate ratio estimates were attenuated.

Discussion

Drinking water utilities and regulatory agencies have long been concerned that turbidity is too crude a measure of pathogen load to be used as an indicator of drinking water quality (Trussel 2006). While turbidity removal is often correlated with pathogen removal (Hendricks et al. 1998; LeChevallier and Norton; LeChevallier, Norton, and Lee 1991; Nieminski 1992; U.S. Environmental Protection Agency 1999), there is often little or no association between actual turbidity levels and levels of pathogens in raw and filtered water (Edzwald and Kelley 1998; LeChevallier, Norton, and Lee 1991; Logsdon et al. 1985; Payment 1998; Payment and Hunter 2001). Our results support this concern. While we did observe an association between raw water turbidity and subsequent ED visits for GI illness, it is filtered water turbidity that is regulated by EPA standards, and we found no association between this measure and GI illness ED visits.

Microorganisms are generally too small to be detected by turbidimeters. The utility of turbidity as a metric of water quality stems from the assumption that as the concentration of suspended particles, such as clay, silt, and organic matter, increases in the water, so do the levels of microorganisms. How well levels of microorganisms may track with levels of larger particles is likely highly variable, depending on the conditions leading to the turbidity change. For example, a heavy rainfall following a period of drought may be more likely to carry pathogens that have concentrated on the shore into the water. Conversely, rainfall may serve only to dilute the concentration of pathogens in the raw water. A different phenomenon may be impacting the utility of turbidity as a water quality indicator in filtered water. The size and charge of a particle determine its likelihood of being extracted from the raw water during filtration. Clay, silt, and organic matter may be more likely to be removed during filtration due to their size and charge, whereas microorganisms may be less likely to be removed. Therefore, the turbidity level of filtered water may less accurately reflect the levels of microorganisms than the turbidity level of raw water.

The only other time-series study that considered both raw and filtered water turbidity was conducted in France and examined the association between turbidity and sales of over-the-counter anti-diarrheal medications (Beaudeau et al. 1999). The researchers found an association of raw water turbidity and anti-diarrheal drug sales over the following three weeks. They did not find an association between filtered water turbidity and anti-diarrheal drug sales. The results of this time-series study are consistent with those of our study. The results of these two studies considered together are made

stronger because very different outcome measures of GI illness, ED visits and anti-diarrheal drug sales, were each associated with raw water turbidity.

The overall results suggested no association between increases in average or maximum filtered water turbidity over 21 days and GI illness ED visits. When individual plants were considered, however, there was a positive association for Plant G. If this heterogeneity reflects drinking water quality, it could be explained by differences in raw water quality, treatment effectiveness, or distribution system integrity of this utility compared to the other utilities. Plant G had the 2nd highest levels of raw water turbidity of all the treatment plants considered in the study. This plant is served by a river and a waste-water treatment plant discharges into the river upstream of the raw water intake for the plant. We were unable to quantitatively assess specific attributes of this plant that might have contributed to this observation, as many traits were similar across plants.

We observed positive associations between both minimum and maximum raw water turbidity over 21 days and ED visits for GI illness. While positive associations were observed throughout the 21-day lag period, the strongest associations were observed for increased turbidity six to nine days prior to reporting to an ED for GI illness. This lag period is consistent with all three of the major types of organisms causing waterborne disease. Viruses have short incubation periods, 24 to 48 hours (Chin 2000), but coupled with water travel time in the distribution system, viral exposure could result in ED visits for GI illness six to nine days later. Protozoa typically have a longer incubation period, seven to 14 days (Chin 2000); however, a short travel time from the plant to the end user would be consistent with people ill from infection with protozoa to report to an ED for GI

illness between six and nine days later. The incubation periods for bacterial waterborne pathogens vary, but generally fall between those of viruses and protozoa.

As was observed for the overall results, the results of the age-specific models were generally consistent with the null when the filtered water turbidity exposure was considered. The statistically significant negative rate ratio observed for young children when the average filtered water turbidity measure was considered may have been spurious, as increased turbidity would not be expected to be associated with higher quality water. Positive statistically significant associations were observed only for children age five years and younger when the raw water turbidity exposure was considered. These results are consistent with other studies in which GI illness tends to more heavily impact the young, likely due to their underdeveloped and naïve immune systems (Glass et al. 1991; Jin et al. 1996).

We allowed a moderate level of exposure misclassification in our analyses by including zip codes that were served up to 20 percent by a treatment plant other than the one from which the turbidity exposure was taken. However, sensitivity analyses in which only zip codes served entirely by a single treatment plant were considered yielded similar results (see Appendix C). We were unable to assess the turbidity of drinking water that was not consumed from the tap at home, leading to the potential for further exposure misclassification.

Determining the exposure window for GI illness due to drinking water exposure is difficult given the many variables impacting the lag time between when turbidity is measured at the treatment plant and when someone presents for GI illness at an ED. This lag time encompasses the storage and travel time of the water from the treatment plant to

the home, the incubation period of the organism, and the time from illness onset until symptoms become sufficiently serious to visit an ED. Previous time-series studies assessed many models and considered multiple lag structures, leading to concerns about multiple comparisons that may result in the observation of spurious statistically significant associations. To mitigate this concern in the present study, our primary analysis used one *a priori* lag structure (a 21-day distributed lag) and reserved exploration of alternative lag structures for secondary analyses. In sensitivity analyses we examined the turbidity exposure as a series of three-day moving averages (Figure 8.7; plant-specific results are presented in Appendix C).

Our database captured only GI illness severe enough to warrant an ED visit, a small proportion of all cases of infectious GI illness. However, these more serious cases of disease may be those of greatest interest. Our study had the power to detect modest associations due to the large size of the ED database. The participation of all of the drinking water utilities allowed for greater generalizability to other metropolitan areas.

Despite the use of *a priori* models, we considered many models and there is always the possibility of spurious, statistically significant associations. Each of the four available turbidity measures considered, average and maximum filtered water turbidity and minimum and maximum raw water turbidity, has a different implication regarding an association with GI illness, and therefore we felt that all of these variables should be examined. The average filtered water turbidity and minimum raw water turbidity are more indicative of the overall quality of the water for a given day and high levels implicate a more long term source of contamination. Maximum filtered and raw water turbidity are indicative of potentially isolated incidents leading to spikes in turbidity.

Because we considered many models, the implications of the statistically significant results should be interpreted with particular caution.

While a strength of the time-series method is that factors that do not vary from day-to-day, such as age or socio-economic status, cannot act as confounding factors, confounding can still occur due to short- or long-term temporally varying factors. We attempted to control for covariates that may have been temporally associated with both turbidity levels and ED visits for GI illness by including variables indicating temperature and day of the week in the analytical model. There are other factors which may be related to both turbidity and GI illness we did not control for, such as other meteorological conditions. We did not control for rainfall in the *a priori* model because it is in the theoretical pathway of the association between turbidity and subsequent GI illness. We did examine the impact of rainfall in sensitivity analyses (see Appendix C).

We attempted to account for long-term time trends that could impact our analysis using a cubic spline with six knots per year. The addition of knots to the spline may have led to better control of long-term time trends, but we might also have lost the ability to detect a true association. Including fewer knots in the spline would mitigate that concern, but would result in an increased potential for residual confounding by long-term time trends. Sensitivity analyses in which alternate knot structures were considered are presented in Appendix C.

The results of our study suggest that drinking water pathogen exposure modestly contributes to endemic GI illness in this study area. The association between turbidity and GI illness was only observed when raw water turbidity was considered, however, and therefore filtered water turbidity may not serve as a reliable indicator of modest pathogen

risk at all drinking water treatment plants. The development of more refined indicators of water quality and health risk, such as standardized and reliable particle counters (Trussell 2006), should continue to be pursued.

References

Beaudeau, P., P. Payment, D. Bourderont, F. Mansotte, O. Boudhabay, B. Laubies, and J. Verdiere. 1999. A time series study of anti-diarrheal drug sales and tap-water quality. *Int J Environ Health Res* 9:293-311.

Chin, J. 2000. *Control of Communicable Diseases Manual*. 17 ed. Washington, DC: American Public Health Association.

Colford Jr., J. M., T. J. Wade, S. K. Sandhu, C. C. Wright, S. Lee, S. Shaw, K. Fox, S. Burns, A. Benker, M. A. Brookhart, M. van der Laan, and D. A. Levy. 2005. A randomized, controlled trial of in-home drinking water intervention to reduce gastrointestinal illness. *Am J Epidemiol* 161 (5):472-82.

Edzwald, J. K., and M. B. Kelley. 1998. Control of *Cryptosporididum*: from reservoris to clarifiers to filters. *Water Sci Technol* 37 (2):1-8.

Gangarosa, R. E., R. I. Glass, J. F. Lew, and J. R. Boring. 1992. Hospitalizations involving gastroenteritis in the United States, 1985: the special burden of the disease among the elderly. *Am J Epidemiol* 135 (3):281-90.

Glass, R. I., J. F. Lew, R. E. Gangarosa, C. W. LeBaron, and M. S. Ho. 1991. Estimates of morbidity and mortality rates for diarrheal diseases in American children. *J Pediatr* 118 (4.2):S27-33.

Hellard, M. E., M. I. Sinclair, A. B. Forbes, and C. K. Fairley. 2001. A randomized, blinded, controlled trial investigating the gastrointestinal health effects of drinking water quality. *Environ Health Perspect* 109 (8):773-8.

Hendricks, D. W., W. F. Clunie, W. L. Anderson, G. D. Sturbaum, D. A. Klein, T. Champlin, P. Krugens, C. M. Hancock, J. Hirsh, B. Mccourt, P. M. Wendling,

and G. Nordby. 1998. Biological paricle surrogates for filtration performance evaluation. Denver: American Water Works Association Research Foundation Report.

Hoxie, N. J., J. P. Davis, J. M. Vergeront, R. D. Nashold, and K. A. Blair. 1997. Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *Am J Public Health* 87 (12):2032-5.

Jin, S., P. E. Kilgore, R. C. Holman, M. J. Clarke, E. J. Gangarosa, and R. I. Glass. 1996. Trends in hospitalizations for diarrhea in United States children from 1979 through 1992: estimates of the morbidity associated with rotavirus. *Pediatr Infect Dis J* 15 (5):397-404.

LeChevallier, M. W., W. D. Norton, and R. G. Lee. 1991. *Giardia* and *Cryptosporidium* spp. in filtered drinking water supplies. *Appl Environ Microbiol* 57 (9):2617-21.

———. 1991. Occurrence of *Giardia* and *Cryptosporidium* spp. in surface water supplies. *Appl Environ Microbiol* 57 (9):2610-6.

LeChevallier, M. W., and W. D. Norton. Treatments to address source water concerns: Protozoa. In *Safety of Water Disinfection: Balancing Chemical and Microbial Risks*, edited by G. Craun. Washington, DC: ILSI Press.

Logsdon, G. S., V. C. Thurman, E. S. Frindt, and J. G. Stoecker. 1985. Evaluating sedimentation and various filter media for removal of *Giardia* cysts. *J Am Water Works Assoc* 77:61-6.

Morris, R. D., E. N. Naumova, and J. K. Griffiths. 1998. Did Milwaukee experience waterborne cryptosporidiosis before the large documented outbreak in 1993? *Epidemiology* 9 (3):264-70.

Morris, R. D., E. N. Naumova, R. Levin, and R. L. Munasinghe. 1996. Temporal variation in drinking water turbidity and diagnosed gastroenteritis in Milwaukee. *Am J Public Health* 86 (2):237-9.

Nieminski, E.C. 1992. *Giardia* and *Cryptosporidium* - Where do the cysts go. Paper read at American Water Works Association Water Quality Technology Conference.

Payment, P. 1998. Waterborne viruses and parasites: resistance to treatment and disinfection. In *OECD Workshop Molecular Methods for Safe Drinking Water*: EAWAG.

Payment, P., and P. R. Hunter. 2001. Endemic and epidemic infectious intestinal disease and its relationship to drinking water. In *Water Quality: Guidelines, Standards, and Health*, edited by L. Fewtrell and J. Bartram. London: IWA Publishing.

Payment, P., L. Richardson, J. Siemiatycki, R. Dewar, M. Edwardes, and E. Franco. 1991. A randomized trial to evaluate the risk of gastrointestinal disease due to consumption of drinking water meeting current microbiological standards. *Am J Public Health* 81 (6):703-8.

Payment, P., J. Siemiatycki, L. Richardson, G. Renaud, E. Franco, and M. Prevost. 1997. A prospective epidemiological study of gastrointestinal health effect due to the consumption of drinking water. *Int J Environ Health Res* 7:5-31.

SAS 9.1. SAS Institute, Inc., Cary, North Carolina.

Schwartz, J., R. Levin, and R. Goldstein. 2000. Drinking water turbidity and gastrointestinal illness in the elderly of Philadelphia. *J Epidemiol Community Health* 54 (1):45-51.

Schwartz, J., R. Levin, and K. Hodge. 1997. Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* 8 (6):615-20.

Trussell, R. 2006. Water treatment: the past 30 years. *J Am Water Works Assoc* 98 (3):100-9.

U.S. Centers for Disease Control and Prevention. 1999. From the U.S. Centers for Disease Control and Prevention. Control of infectious diseases, 1900-1999. *JAMA* 282 (11):1029-32.

U.S. Environmental Protection Agency. 1999. Importance of Turbidity. In *EPA Guidance Manual: Turbidity Provisions*, edited by U.S. Environmental Protection Agency.

Table 8.1: Distribution of daily environmental variables by plant service area, Atlanta, 1993 - 2004

Plant	*10%*	*25%*	*50%*	*75%*	*90%*	*Mean (Standard Deviation)*
Average Temperature (degrees Fahrenheit) [1]						
	42	51	64.5	76	80.5	62.8 (14.7)
Daily Average Filtered Water Turbidity (NTU) [2, a]						
A	0.01	0.02	0.03	0.04	0.05	0.03 (0.01)
B	0.03	0.04	0.06	0.07	0.08	0.06 (0.02)
C [b]	0.03	0.05	0.07	0.09	0.11	0.07 (0.04)
D [d]	0.06	0.06	0.07	0.08	0.10	0.08 (0.02)
E	0.03	0.04	0.07	0.13	0.20	0.10 (0.10)
F	0.06	0.09	0.15	0.23	0.29	0.17 (0.10)
G	0.04	0.05	0.05	0.07	0.08	0.06 (0.02)
H	0.05	0.06	0.07	0.09	0.12	0.08 (0.04)
All [3]	0.03	0.04	0.06	0.09	0.15	0.08 (0.07)
Daily Maximum Filtered Water Turbidity (NTU) [4, a]						
A	0.01	0.02	0.04	0.05	0.07	0.04 (0.03)
B	0.05	0.07	0.08	0.10	0.14	0.09 (0.06)
C [c]	0.09	0.09	0.11	0.12	0.15	0.11 (0.04)
D [d]	0.07	0.08	0.09	0.11	0.13	0.10 (0.03)
E	0.06	0.08	0.13	0.23	0.37	0.19 (0.19)
F	0.08	0.13	0.28	0.41	0.50	0.29 (0.19)
G	0.05	0.05	0.07	0.08	0.10	0.07 (0.04)
H	0.06	0.08	0.09	0.13	0.18	0.12 (0.07)
All [3]	0.04	0.06	0.09	0.13	0.27	0.13 (0.13)
Daily Minimum Raw Water Turbidity (NTU) [4, e]						
A	2.0	2.6	3.4	5.0	7.1	4.2 (2.8)
B	1.8	2.6	4.8	9.0	15.0	7.4 (7.9)
C	5.0	6.0	9.0	16.0	34.0	16.3 (22.9)
D [f]	1.2	1.6	2.2	4.5	7.6	3.5 (2.9)
E	3.0	4.0	6.0	11.0	18.0	8.7 (7.8)
F	0.7	0.9	1.1	1.2	1.4	1.1 (0.3)
G	3.3	4.2	6.1	10.2	20.7	11.0 (16.5)
H	1.6	2.1	3.5	12.0	20.0	7.9 (8.3)
All [3]	1.2	2.0	4.2	8.2	16.0	7.5 (12.1)
Daily Maximum Raw Water Turbidity (NTU) [4, e]						
A	2.8	3.6	5.0	7.4	10.9	6.5 (7.0)
B	3.0	4.5	7.7	15.2	31.0	13.5 (16.2)
C	11.0	14.5	26.0	53.0	126.0	55.0 (101.3)
D [f]	1.5	1.9	2.8	5.4	8.7	4.2 (3.3)
E	4.0	5.0	7.9	15.0	25.5	11.9 (10.8)
F	1.1	1.3	1.5	1.7	2.0	1.5 (0.4)
G	5.1	6.5	9.9	19.4	51.6	22.0 (34.3)
H	2.4	3.1	5.3	16.0	25.0	10.3 (10.0)
All [3]	1.6	3.0	6.1	14.3	31.1	15.7 (41.9)

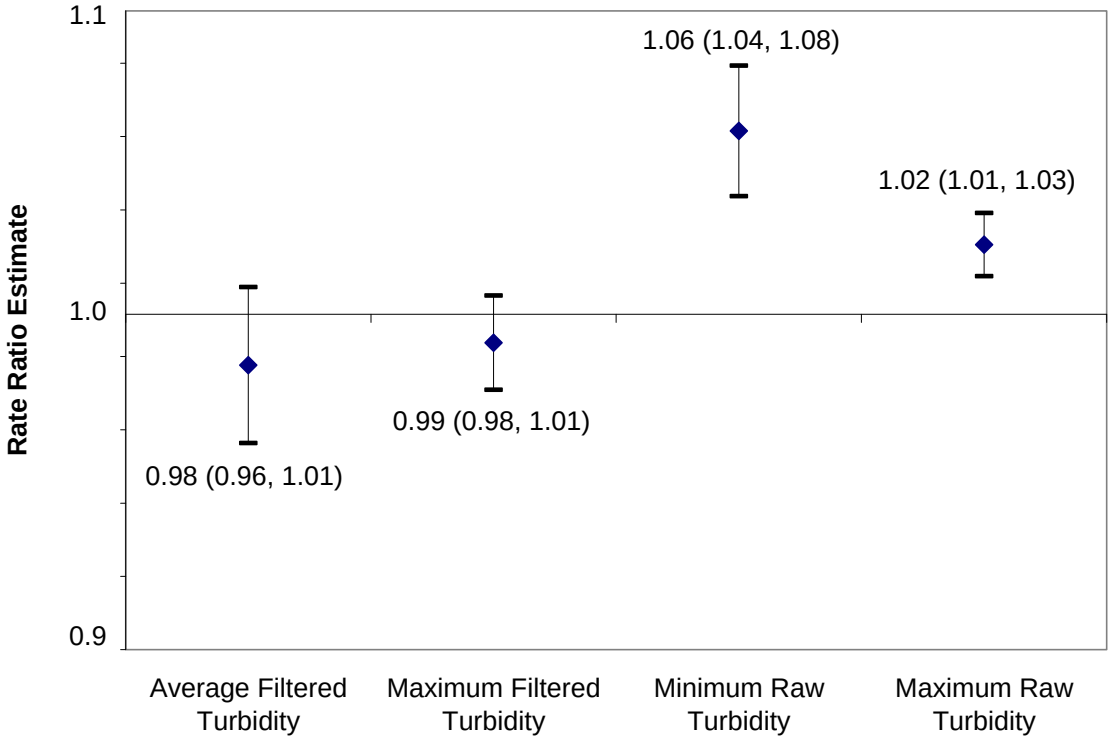
[1] Average of the daily minimum and maximum temperature measured at Hartsfield – Jackson International Airport
[2] Average turbidity indicates arithmetic average of hourly turbidity measures taken at the treatment plant in a 24-hour period.
[3] All plants combined.
[4] Minimum and maximum turbidity measures indicate the minimum and maximum hourly measurements taken at the treatment plant in a 24-hour period.
[a] Data available 7/1/1993 – 12/31/2004.
[b] Data available 1/1/1993 – 12/31/2004.
[c] Data available 1/1/2000 – 12/31/2004.
[d] Data available 11/1/1999 – 12/31/2004.
[e] Data available 1/1/2002 – 12/31/2004.
[f] Data available 3/1/2002 – 12/31/2004.

Table 8.2: Distribution of daily total and gastrointestinal illness emergency department visits by plant service area, Atlanta, 1993 – 2004

Plant	*10%*	*25%*	*50%*	*75%*	*90%*	*Mean (Standard Deviation)*
Total ED Visits						
A	26	43	76	96	110	71.2 (31.5)
B	195	241	299	366	405	298 (87.0)
C	3	11	17	22	26	16.1 (8.1)
D	14	17	22	26	31	22.0 (6.7)
E	110	142	174	343	388	220 (108.9)
F	33	54	206	322	366	194 (131.8)
G	45	70	176	210	229	153 (72.1)
H	4	11	75	102	115	66.2 (42.6)
All [a]	14	28	99	211	342	136 (121.5)
Total GI ED Visits						
A	1	2	4	6	8	4.3 (3.2)
B	7	11	16	23	30	17.8 (9.7)
C	0	0	1	1	2	0.8 (1.0)
D	0	0	1	2	3	1.2 (1.2)
E	4	6	10	16	24	12.0 (8.2)
F	1	3	11	18	24	11.8 (10.1)
G	1	4	8	12	16	8.3 (5.7)
H	0	1	3	6	8	3.7 (3.4)
All [a]	0	1	5	12	20	7.8 (9.5)

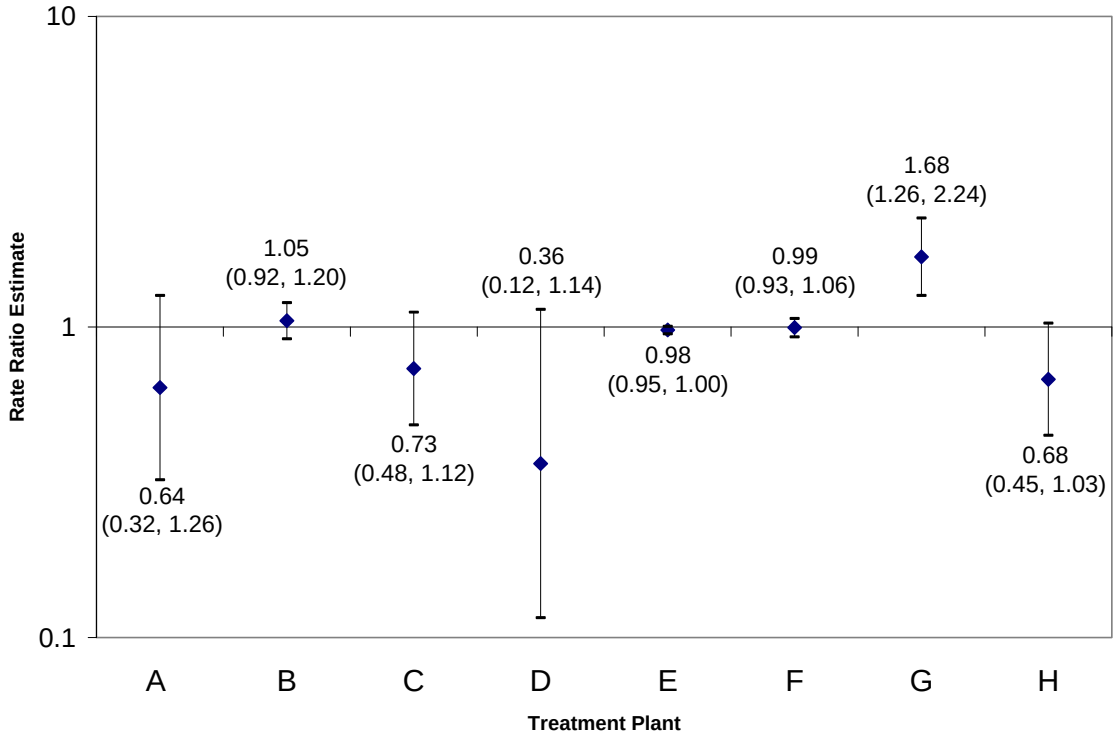
[a] All plants combined.

Figure 8.1: Rate ratio estimates[a] for a 0.1 NTU average increase in average[b,c] and maximum[c,d] filtered water turbidity and a 10 NTU average increase in minimum[c,e] and maximum[c,e] raw water turbidity over 21 days and emergency department visits for gastrointestinal illness, Atlanta, 1993 – 2004



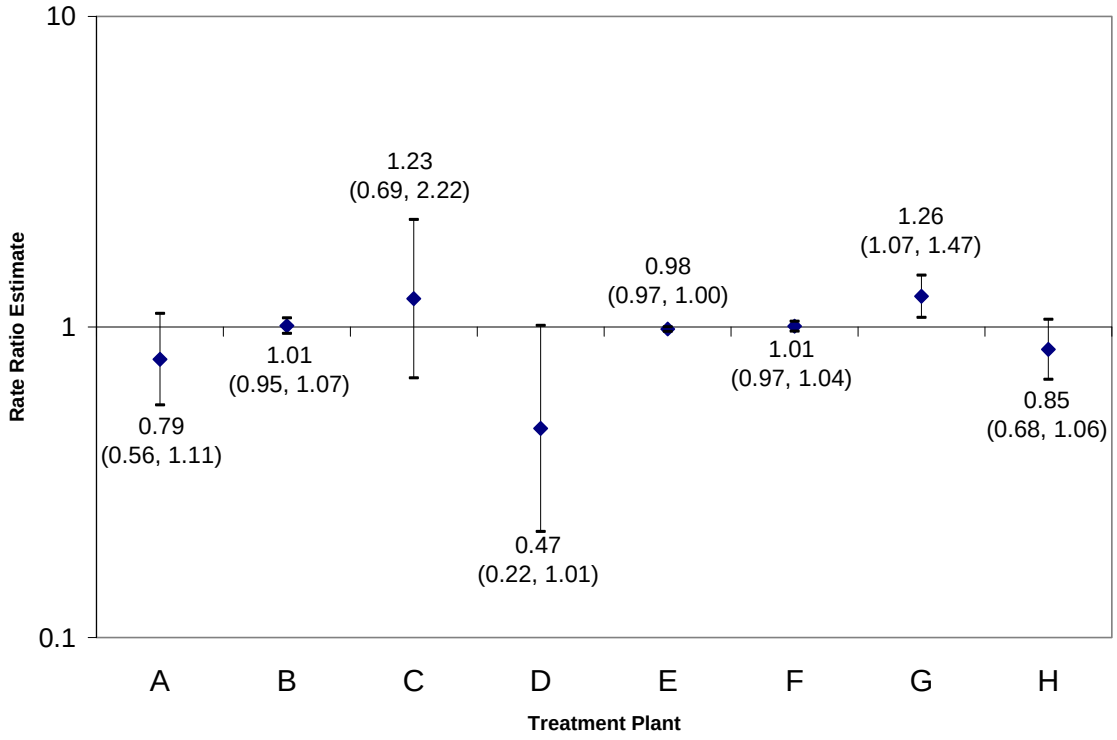
[a] Error bars indicate 95 percent confidence intervals.
[b] Average filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/1993 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[c] Average, minimum, and maximum turbidity based on hourly turbidity measures taken at the treatment plant in a 24-hour period.
[d] Maximum filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/2000 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[e] Minimum and maximum raw water turbidity data available 1/1/2002 – 12/31/2004, except Plant D (3/1/2002 – 12/31/20004).

Figure 8.2: Rate ratio estimates[a] for a 0.1 NTU average increase in average[b,c] filtered water turbidity over 21 days and emergency department visits for gastrointestinal illness, Atlanta, 1993 – 2004, stratified by drinking water treatment plant



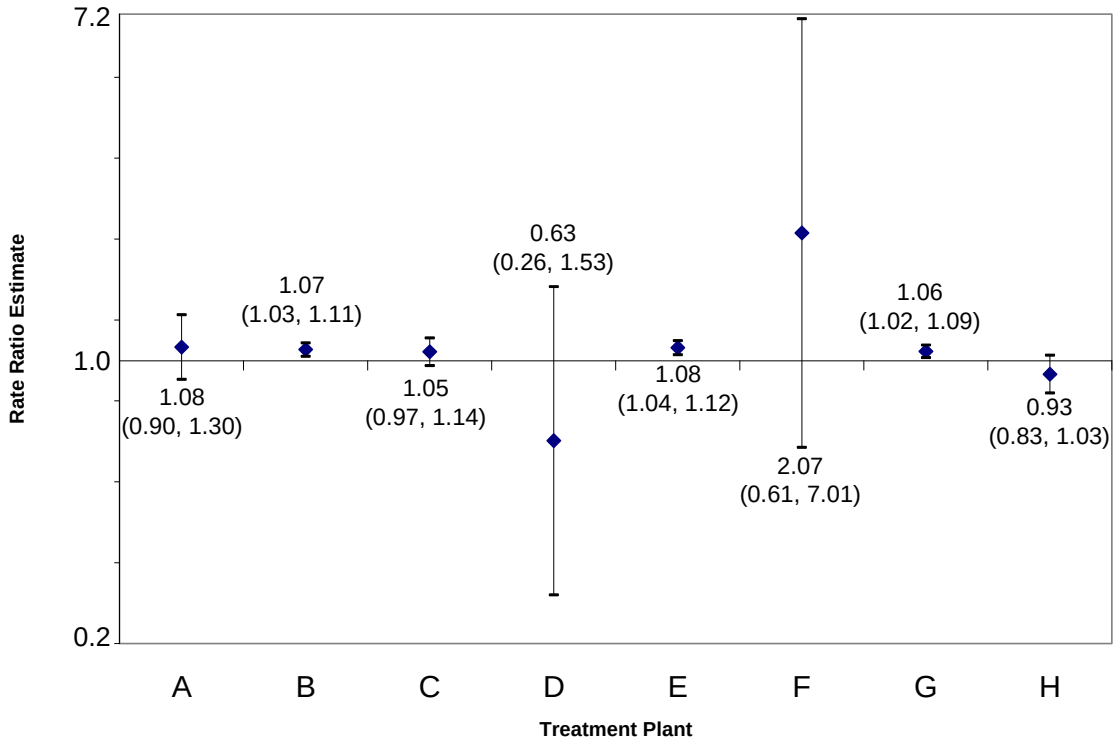
[a] Error bars indicate 95 percent confidence intervals.
[b] Average filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/1993 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[c] Average based on hourly turbidity measures taken at the treatment plant in a 24-hour period.

Figure 8.3: Rate ratio estimates[a] for a 0.1 NTU average increase in maximum[b,c] filtered water turbidity over 21 days and emergency department visits for gastrointestinal illness, Atlanta, 1993 – 2004, stratified by drinking water treatment plant



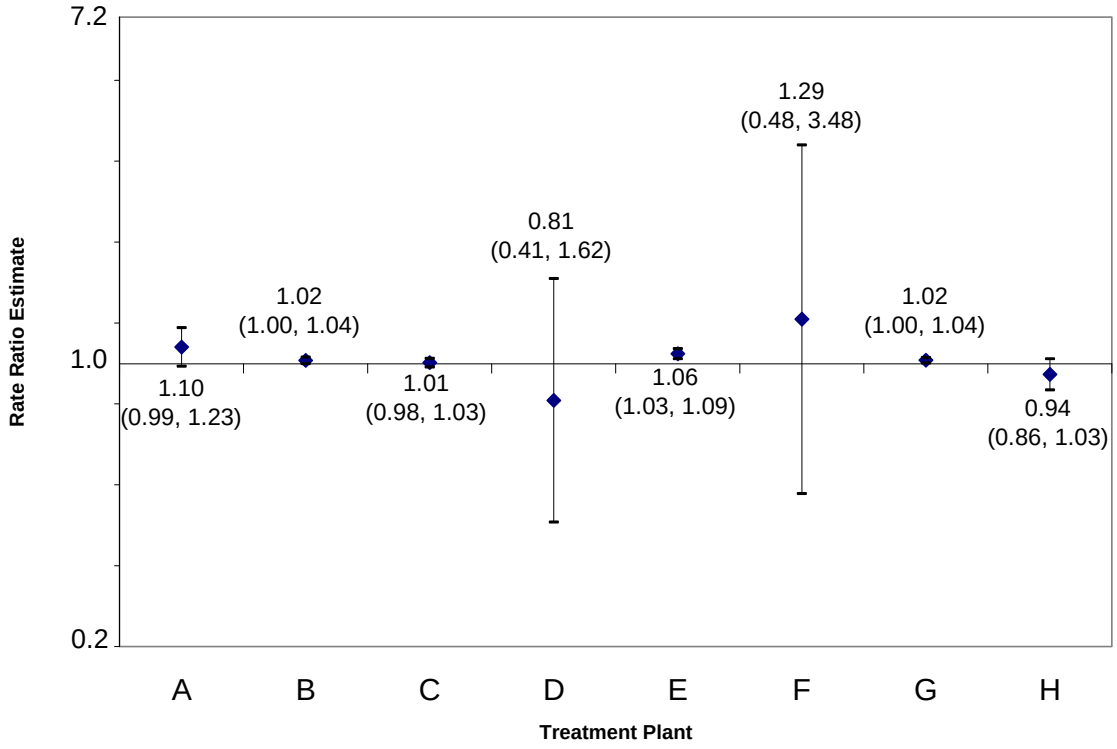
[a] Error bars indicate 95 percent confidence intervals.
[b] Maximum filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/2000 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[c] Maximum turbidity based on hourly turbidity measures taken at the treatment plant in a 24-hour period.

Figure 8.4: Rate ratio estimates[a] for a 10 NTU average increase in minimum[b,c] raw water turbidity over 21 days and emergency department visits for gastrointestinal illness, Atlanta, 1993 – 2004, stratified by drinking water treatment plant



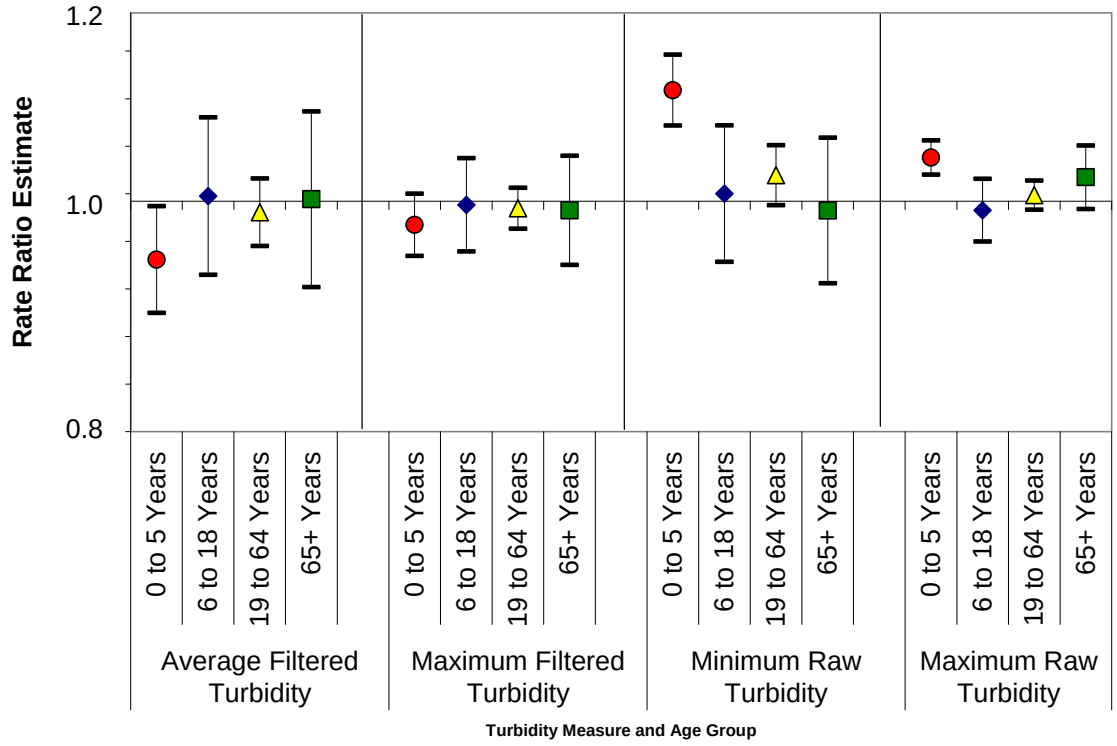
[a] Error bars indicate 95 percent confidence intervals.
[b] Minimum raw water turbidity data available 1/1/2002 – 12/31/2004, except Plant D (3/1/2002 – 12/31/20004).
[c] Minimum turbidity based on hourly turbidity measures taken at the treatment plant in a 24-hour period.

Figure 8.5: Rate ratio estimates[a] for a 10 NTU average increase in maximum[b,c] raw water turbidity over 21 days and emergency department visits for gastrointestinal illness, Atlanta, 1993 – 2004, stratified by drinking water treatment plant



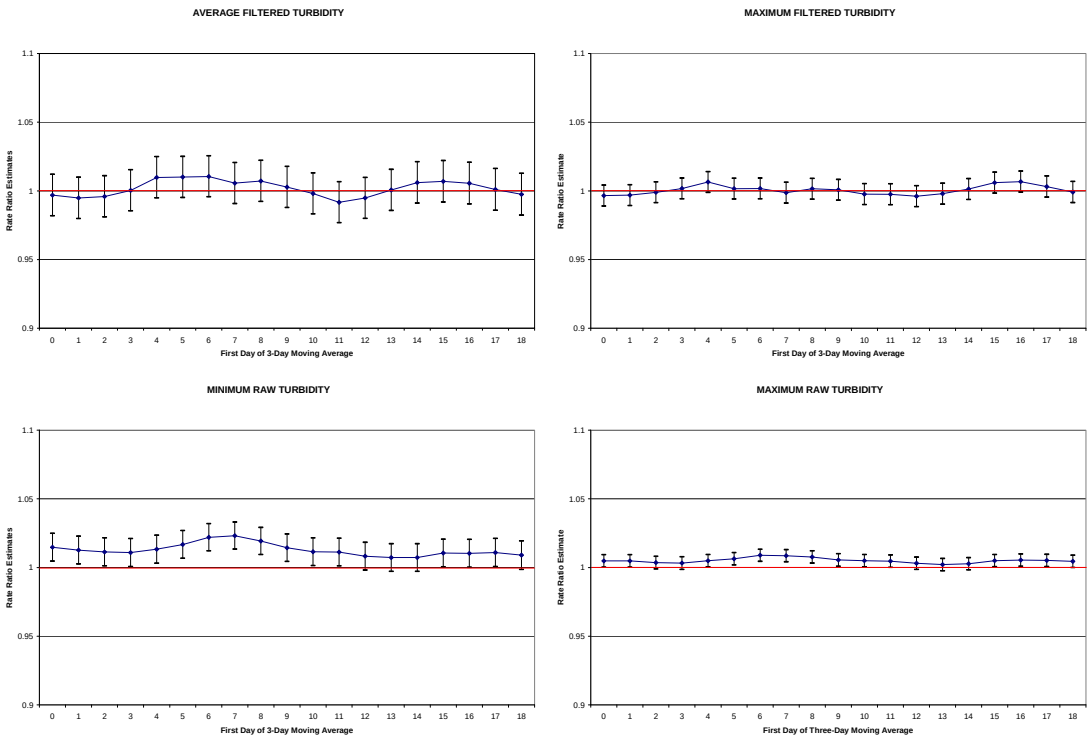
[a] Error bars indicate 95 percent confidence intervals.
[b] Maximum raw water turbidity data available 1/1/2002 – 12/31/2004, except Plant D (3/1/2002 – 12/31/20004).
[c] Maximum turbidity based on hourly turbidity measures taken at the treatment plant in a 24-hour period.

Figure 8.6: Rate ratio estimates[a] for a 0.1 NTU average increase in average[b,c] and maximum[c,d] filtered water turbidity and a 10 NTU average increase in minimum[c,e] and maximum[c,e] raw water turbidity over 21 days and emergency department visits for gastrointestinal illness, Atlanta, 1993 – 2004, stratified by age group



[a] Error bars indicate 95 percent confidence intervals.
[b] Average filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/1993 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[c] Average, minimum, and maximum turbidity based on hourly turbidity measures taken at the treatment plant in a 24-hour period.
[d] Maximum filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/2000 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[e] Minimum and maximum raw water turbidity data available 1/1/2002 – 12/31/2004, except Plant D (3/1/2002 – 12/31/20004).

Figure 8.7: Rate ratio estimates[a] for a three-day moving average increase of 0.1 NTU in average[b,c] and maximum[c,d] filtered water turbidity and 10 NTU in minimum[c,e] and maximum[c,e] raw water turbidity and emergency department visits for gastrointestinal illness, Atlanta, 1993 – 2004



[a] Error bars indicate 95 percent confidence intervals.
[b] Average filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/1993 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[c] Average, minimum, and maximum turbidity based on hourly turbidity measures taken at the treatment plant in a 24-hour period.
[d] Maximum filtered water turbidity data available 7/1/1993 – 12/31/2004, except Plant C (1/1/2000 – 12/31/20004) and Plant D (11/1/1999 – 12/31/2004).
[e] Minimum and maximum raw water turbidity data available 1/1/2002 – 12/31/2004, except Plant D (3/1/2002 – 12/31/20004).

CHAPTER 9

CONCLUSIONS

The contribution of drinking water to endemic gastrointestinal (GI) illness has been established by randomized-controlled trials and observational studies (Aramini et al. 2000; Beaudeau et al. 1999; Egorov et al. 2002; Hunter et al. 2005; Morris, Naumova, and Griffiths 1998; Morris et al. 1996; Nygard et al. 2004; Nygard et al. 2007; Payment, Richardson et al. 1991; Payment et al. 1997; Schwartz, Levin, and Goldstein 2000; Schwartz, Levin, and Hodge 1997). However, associations between drinking water type or quality and GI illness have not been observed in all studies (Colford Jr. et al. 2005; Hellard et al. 2001; Mathias, Riben, and Osei 1992), and therefore the magnitude of the GI disease burden in the U.S. attributable to this exposure remains unknown. The increasing role of drinking water distribution system contamination in outbreaks of GI illness has raised questions regarding the role of source water contamination not removed at the water treatment plant relative to post-treatment contamination occurring within the distribution system (Craun and Calderon 2001). Currently it is not feasible to directly measure pathogen levels in drinking water, and it is therefore important to understand the utility of indicators and surrogates used to assess the microbial quality of drinking water.

Three studies were conducted to examine the population impact of multiple surrogates of drinking water quality in Atlanta, Georgia. These studies made use of an extensive emergency department (ED) database compiled for the Study of Particles and Health in Atlanta (SOPHIA). This ED database contained information on more than 10 million visits made to 41 hospitals between 1993 and 2004. Information collected from

the hospitals included ICD-9 diagnostic codes, allowing designation of a GI illness outcome group, and residential zip code, which allowed the patients to be linked to their areas of residence.

The exposure of interest for the first of the studies was an estimate of the time taken by drinking water to travel from the treatment plant to the end user (water residence time). Water residence time estimates were derived from hydraulic models developed by one of the utilities serving the study area to simulate the flow of water through the distribution system pipes. The average residence time was estimated for each zip code in the utility's service area and categorical variables were created using cut-points at the 10^{th} and 90^{th} percentiles of all residence times. The results suggested that people served by water that has spent the longest amount of time in the distribution system were at increased risk of GI illness ED visits compared to those with short or intermediate residence times. These results support a role for contamination occurring within the distribution system in contributing to endemic GI illness because the longer water spends in the distribution system, the greater the opportunity for that water to become contaminated.

The second study examined the role of the drinking water treatment plant itself as a risk factor for GI illness, as source water quality and treatment methods differ across plants. The rate of GI illness ED visits among people served by a given plant was compared to the rate of GI illness ED visits experienced by those served by all other seven plants combined. The results suggested a moderately increased rate of GI illness among people served by one of the treatment plants. The strongest associations were observed among young children and elderly adults, which is consistent with previous

research suggesting greater susceptibility to waterborne pathogens among the youngest and older age groups (Gerba, Rose, and Haas 1996; Glass et al. 1991; Jin et al. 1996). The plant for which a positive rate ratio was observed used a unique treatment technique, direct filtration, which omits the sedimentation step of conventional drinking water treatment because of the low turbidity of its source water. This step may be particularly important for removal of viral and bacterial contamination. Viruses and bacteria have better survival in cooler temperatures (Flint 1987; Rzezutka and Cook 2004), and the positive associations with GI illness for people served by this plant were restricted to autumn and winter months.

The final study examined the association between turbidity, the primary indicator of drinking water quality used by utilities, and ED visits for GI illness using time-series methods. Four daily turbidity measures were considered, summarized from hourly measurements made by the utility: average and maximum filtered water turbidity and minimum and maximum raw water turbidity. The turbidity levels on the day of the ED visit and the 20 preceding days were considered in the same model using an unconstrained distributed lag structure. The results suggested that turbidity was related to GI illness ED visits, but this relationship was only observed when turbidity measured in raw water was considered. These results have important implications for the current U.S. Environmental Protection Agency guidelines regarding turbidity levels. If infectious organisms are transmitted through treated drinking water, as suggested by the results when the raw turbidity exposures were considered, these organisms were not successfully indicated by the filtered water turbidity measures in this study. The raw water turbidity measures describe the state of the source water and are not regulated. Filtered water

turbidity is used by EPA to serve as an indicator of the microbial quality of drinking water. The results of this study suggest that development of an improved water quality indicator may be needed to detect modest levels of pathogen contamination.

Caution must be exercised in interpreting the results, as alternative explanations exist. The analyses for the first two studies were particularly susceptible to confounding by spatial covariates, and while variables were included in the regression models in an attempt to accommodate this source of bias, residual spatial confounding likely remained. While the third study, for which a time-series analysis was conducted, was not as susceptible to spatial confounding as the other studies, this analysis was vulnerable to temporal confounding. Indicators for day-of-week and temperature and cubic splines were included in the analytical model to address these concerns, but residual confounding may have remained.

Roles for both the raw water source and the distribution system as sites of drinking water contamination are suggested by the results of the analyses. Filtered water turbidity, a primary water quality measure used by the utilities, did not appear to predict risk. Overall, the results of these studies suggest that a low level of GI illness in Atlanta may be attributable to drinking water exposure, particularly among young children and the elderly.

CHAPTER 10

REFERENCES

Abu-Ashour, J., D. M. Joy, H. Lee, H. R. Whiteley, and S. Zelin. 1994. Transport of microorganisms through soil. *Water, Air, Soil Pollut* 75:141-58.

Aggazzotti, G., L. Fabiani, M. Triassi, S. Sciacca, G. Ravera, S. Kanitz, F. Barbone, K. Sansebastiano, M. A. Battaglia, and V. Leoni. 2003. Exposure to chlorite and chlorate in drinking water and adverse pregnancy outcomes. *Epidemiology* 14 (5):S24.

Almon, S. 1965. The distributed lag between capital appropriations and expenditures. *Econometrica* 33:178-96.

Altekruse, S. F., N. J. Stern, P. I. Fields, and D. L. Swerdlow. 1999. *Campylobacter jejuni*--an emerging foodborne pathogen. *Emerg Infect Dis* 5 (1):28-35.

American Water Works Association Research Foundation. 1992. Water Industry Database: Utility Profiles: American Water Works Association.

Aramini, J., M. McLean, J. Wilson, J. Holt, R. Copes, B. Allen, and W. Sears. 2000. Drinking water quality and health-care utilization for gastrointestinal illness in greater Vancouver. *Can Commun Dis Rep* 26 (24):211-4.

Ashbolt, N. J., W. O. K. Grabow, and M. Snozzi. 2001. Indicators of microbial water quality. In *Water Quality: Guidelines, Standards, and Health*, edited by L. Fewtrell and J. Bartram. London, UK: World Heath Organization.

Atherholt, T. B., M. W. LeChevallier, W. D. Norton, and J. S. Rosen. 1998. Effect of rainfall on *Giardia* and *Cryptosporidium*. *J Am Water Works Assoc* 90:66-80.

Atherton, F., C. P. Newman, and D. P. Casemore. 1995. An outbreak of waterborne cryptosporidiosis associated with a public water supply in the UK. *Epidemiol Infect* 115 (1):123-31.

Avery, B. K., and A. Lemley. 1996. *Cryptosporidium*: A waterborne pathogen. Ithica, NY: Cornell University, U.S. Department of Agriculture.

Barrell, R. A., P. R. Hunter, and G. Nichols. 2000. Microbiological standards for water and their relationship to health risk. *Commun Dis Public Health* 3 (1):8-13.

Barwick, R. S., D. A. Levy, G. F. Craun, M. J. Beach, and R. L. Calderon. 2000. Surveillance for waterborne-disease outbreaks--United States, 1997-1998. *MMWR CDC Surveill Summ* 49 (4):1-21.

Beaudeau, P. 2003. Time Series Analyses. In *Drinking water and infectious disease: establishing the links*, edited by P. R. Hunter, M. Waite and E. Ronchi. London: CRC Press.

Beaudeau, P., P. Payment, D. Bourderont, F. Mansotte, O. Boudhabay, B. Laubies, and J. Verdiere. 1999. A time series study of anti-diarrheal drug sales and tap-water quality. *Int J Environ Health Res* 9:293-311.

Blackburn, B. G., G. F. Craun, J. S. Yoder, V. Hill, R. L. Calderon, N. Chen, S. H. Lee, D. A. Levy, and M. J. Beach. 2004. Surveillance for waterborne-disease outbreaks associated with drinking water--United States, 2001-2002. *MMWR Surveill Summ* 53 (8):23-45.

Boone, S. A., and C. P. Gerba. 2007. Significance of fomites in the spread of respiratory and enteric viral disease. *Appl Environ Microbiol* 73 (6):1687-96.

Braunwald, E., A. S. Fauci, K. J. Iselbacher, D. L. Kasper, S. L. Hauser, D. L. Longo, and J. L. Jameson. 2003. *Harrison's Principles of Internal Medicine*: McGraw-Hill.

Bruce-Grey-Owen Sound Health Unit. 2000. The investigative report of the Walkerton outbreak of waterborne gastroenteritis, May-June, 2000: Bruce-Grey-Owen Sound Health Unit.

Calderon, R. L., G. Craun, and D. A. Levy. 2006. Estimating the infectious disease risks associated with drinking water in the United States. *J Water Health* 4 (Suppl 2):1-2.

Cantor, K. P., R. Hoover, P. Hartge, T. J. Mason, D. T. Silverman, R. Altman, D. F. Austin, M. A. Child, C. R. Key, L. D. Marrett, M. H. Myers, A. S. Narayana, L. I. Levin, J. W. Sullivan, G. M. Swanson, D. B. Thomas, and D. W. West. 1987. Bladder cancer, drinking water source, and tap water consumption: a case-control study. *J Natl Cancer Inst* 79 (6):1269-79.

Castro, K. G., J. W. Ward, L. Slutsker, J. W. Buehler, H. W. Jaffee, and R. L. Berkelman. 1992. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR* 41 (RR-17).

Chen, J., E. N. Naumova, B. Matyas, S. Estes-Smargiassi, and R. D. Morris. 1998. The characteristics of reported cases of Cryptosporidiosis and Giardiasis in Massachusetts and their association with water source. *Epidemiology* 9 (4 Suppl):S49.

Chin, J. 2000. *Control of Communicable Diseases Manual*. 17 ed. Washington, DC: American Public Health Association.

Christian, R. R., and W. O. Pipes. 1983. Frequency distribution of coliforms in water distribution systems. *Appl Environ Microbiol* 45 (2):603-9.

Colford, J. M., Jr., T. J. Wade, K. C. Schiff, C. C. Wright, J. F. Griffith, S. K. Sandhu, S. Burns, M. 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 (1):27-35.

Colford , J., T. Wade, S. Sandhu, C. Wright, S. F. Burns, A. Benker, S. Lee, A. Brookhart, M. Van Der Laan, and D. Levy. 2003. A randomized, blinded, crossover trial of an in-home drinking water intervention to reduce gastrointestinal illness. *Epidemiology* 14 (5):S137.

Colford Jr., J. M., J. R. Rees, T. J. Wade, A. Khalakdina, J. F. Hilton, I. J. Ergas, S. Burns, A. Benker, C. Ma, C. Bowen, D. C. Mills, D. J. Vugia, D. D. Juranek, and D. A. Levy. 2002. Participant blinding and gastrointestinal illness in a randomized, controlled trial of an in-home drinking water intervention. *Emerg Infectect Dis* 8 (1):29-36.

Colford Jr., J. M., T. J. Wade, S. K. Sandhu, C. C. Wright, S. Lee, S. Shaw, K. Fox, S. Burns, A. Benker, M. A. Brookhart, M. van der Laan, and D. A. Levy. 2005. A randomized, controlled trial of in-home drinking water intervention to reduce gastrointestinal illness. *Am J Epidemiol* 161 (5):472-82.

Corso, P. S., M. H. Kramer, K. A. Blair, D. G. Addiss, J. P. Davis, and A. C. Haddix. 2003. Cost of illness in the 1993 waterborne *Cryptosporidium* outbreak, Milwaukee, Wisconsin. *Emerg Infect Dis* 9 (4):426-31.

Council on Environmental Quality. 1989. Environmental trends. Washington, DC: Office of the President.

Craun, G. F., and R. L. Calderon. 2001. Waterborne disease outbreaks caused by distribution system deficiencies. *J Am Water Works Assoc* 93 (9):64-75.

Current, W. L., and L. S. Garcia. 1991. Cryptosporidiosis. *Clin Microbiol Rev* 4:325-58.

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. *Am J Public Health* 91 (8):1194-9.

Dodds, L., W. King, A. C. Allen, A. Armson, D. B. Fell, and C. Nimrod. 2004. Trihalomethanes in public water supplies and risk of stillbirth. *Epidemiology* 15 (2):176-86.

Dodds, L., W. King, A. C. Allen, B. A. Armson, D. B. Fell, and C. Nimrod. 2004. Trihalomethanes in public water supplies and risk of stillbirth. *Epidemiology* 15 (2):179-86.

Dykes, A. C., D. D. Juranek, R. A. Lorenz, S. Sinclair, W. Jakubowski, and R. Davies. 1980. Municipal waterborne giardiasis: an epidemilogic investigation. Beavers implicated as a possible reservoir. *Ann Intern Med* 92 (2 Pt 1):165-70.

Dziuban, E. J., J. L. Liang, G. F. Craun, V. Hill, P. A. Yu, J. Painter, M. R. Moore, R. L. Calderon, S. L. Roy, and M. J. Beach. 2006. Surveillance for waterborne disease

and outbreaks associated with recreational water--United States, 2003-2004. *MMWR Surveill Summ* 55 (12):1-30.

Edberg, S. C. 1996. Assessing health risk in drinking water from naturally occurring microbes. *J Environ Health* 58 (6):18-24.

Edzwald, J. K., and M. B. Kelley. 1998. Control of *Cryptosporididum*: from reservoris to clarifiers to filters. *Water Sci Technol* 37 (2):1-8.

Egorov, A., T. Ford, A. Tereschenko, N. Drizhd, I. Segedevich, and V. Fourman. 2002. Deterioration of drinking water quality in the distribution system and gastrointestinal morbidity in a Russian city. *Int J Environ Health Res* 12 (3):221-33.

Eisenberg, J. N., X. Lei, A. H. Hubbard, M. A. Brookhart, and J. M. Colford, Jr. 2005. The role of disease transmission and conferred immunity in outbreaks: analysis of the 1993 *Cryptosporidium* outbreak in Milwaukee, Wisconsin. *Am J Epidemiol* 161 (1):62-72.

Elkana, Y., N. Gal, and S. Rishpon. 1996. Study of the association between waterborne diseases and microbial water quality in Israel. *Public Health Rev* 24 (1):49-63.

EPA review finds Philadelphia turbidity study seriously flawed. 1998. *Health Stream* (9).

Favero, M. S. 1985. Microbiologic indicators of health risks associated with swimming. *Am J Public Health* 75 (9):1051-4.

Flint, K. P. 1987. The long-term survival of *Escherichia coli* in river water. *Appl Bacteriol* 63 (3):261-70.

Ford, T. E. 1999. Microbiological safety of drinking water: United States and global perspectives. *Environ Health Perspect* 107 (Suppl 1):191-206.

Friedman, S. L., K. R. McQuaid, and J. H. Grendell. 2003. *Current Diagnosis and Treatment in Gastroenterology*: Lange Medical Books/McGraw-Hill.

Furness, B. W., M. J. Beach, and J. M. Roberts. 2000. Giardiasis surveillance--United States, 1992-1997. *MMWR CDC Surveill Summ* 49 (7):1-13.

Gabel, J. A., and J. S. Wolthuis. 2004. Multi-restaurant outbreak of hepatitis A associated with green onions in central and north Georgia, September - October 2003. *Georgia Epidemiology Report* 20 (10):1-3.

Gale, P. 1996. Developments in microbiological risk assessment models for drinking water--a short review. *J Appl Bacteriol* 81 (4):403-10.

Gangarosa, R. E., R. I. Glass, J. F. Lew, and J. R. Boring. 1992. Hospitalizations involving gastroenteritis in the United States, 1985: the special burden of the disease among the elderly. *Am J Epidemiol* 135 (3):281-90.

Garthright, W. E., D. L. Archer, and J. E. Kvenberg. 1988. Estimates of incidence and costs of intestinal infectious diseases in the United States. *Public Health Rep* 103 (2):107-15.

Gerba, C. P., J. B. Rose, and C. N. Haas. 1996. Sensitive populations: who is at the greatest risk? *Int J Food Microbiol* 30 (1-2):113-23.

Gilbert, L., and P. Blake. 1998. Outbreak of *Escherichia coli* O157:H7 infections associated with a water park. *Georgia Epidemiology Report* 14 (7):1-2.

Glass, R. I., S. Bresee, B. Jiang, J. Gentsch, T. Ando, R. Fankhauser, J. Noel, U. Parashar, B. Rosen, and S. S. Monroe. 2000. Gastroenteritis viruses: an overview. In *Symposium on Gastroenteritis Viruses*, edited by D. Chadwick and J. A. Goode. London.

Glass, R. I., J. F. Lew, R. E. Gangarosa, C. W. LeBaron, and M. S. Ho. 1991. Estimates of morbidity and mortality rates for diarrheal diseases in American children. *J Pediatr* 118 (4.2):S27-33.

Gofti-Laroche, L., D. Demanse, J. C. Joret, and D. Zmirou. 2003. Heath risks and parasitical quality of water. *J Am Water Works Assoc* 95 (5):162-72.

Havelaar, A. H., M. van Olphen, and Y. C. Drost. 1993. F-specific RNA bacteriophages are adequate model organisms for enteric viruses in fresh water. *Appl Environ Microbiol* 59 (9):2956-62.

Hayes, E. B., T. D. Matte, T. R. O'Brien, T. W. McKinley, G. S. Logsdon, J. B. Rose, B. L. Ungar, D. M. Word, P. F. Pinsky, M. L. Cummings, M. A. Wilson, E. G. Long, E. S. Hurwitz, and D. D. Juranek. 1989. Large community outbreak of cryptosporidiosis due to contamination of a filtered public water supply. *N Engl J Med* 320 (21):1372-6.

Hellard, M. E., M. I. Sinclair, A. B. Forbes, and C. K. Fairley. 2001. A randomized, blinded, controlled trial investigating the gastrointestinal health effects of drinking water quality. *Environ Health Perspect* 109 (8):773-8.

Hendricks, D. W., W. F. Clunie, W. L. Anderson, G. D. Sturbaum, D. A. Klein, T. Champlin, P. Krugens, C. M. Hancock, J. Hirsh, B. Mccourt, P. M. Wendling, and G. Nordby. 1998. Biological paricle surrogates for filtration performance evaluation. Denver: American Water Works Association Research Foundation Report.

Hoxie, N. J., J. P. Davis, J. M. Vergeront, R. D. Nashold, and K. A. Blair. 1997. Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *Am J Public Health* 87 (12):2032-5.

Hunter, P. R., R. M. Chalmers, S. Hughes, and Q. Syed. 2005. Self-reported diarrhea in a control group: a strong association with reporting of low-pressure events in tap water. *Clin Infect Dis* 40 (4):e32-4.

Ing, D., R. I. Glass, C. W. LeBaron, and J. F. Lew. 1992. Laboratory-based surveillance for rotavirus United States, January 1989-May 1991. *MMWR CDC Surveill Summ* 41 (3):47-56.

Jin, S., P. E. Kilgore, R. C. Holman, M. J. Clarke, E. J. Gangarosa, and R. I. Glass. 1996. Trends in hospitalizations for diarrhea in United States children from 1979 through 1992: estimates of the morbidity associated with rotavirus. *Pediatr Infect Dis J* 15 (5):397-404.

Juranek, D. D., and W. R. Mac Kenzie. 1998. Drinking water turbidity and gastrointestinal illness. *Epidemiology* 9 (3):228-31.

Karim, M. R., M. Abbaszadegan, and M. W. LeChevallier. 2003. Potential for pathogen intrusion during pressure transients. *J Am Water Works Assoc* 95 (5):134-46.

Kilgore, P. E., R. C. Holman, M. J. Clarke, and R. I. Glass. 1995. Trends of diarrheal disease--associated mortality in US children, 1968 through 1991. *JAMA* 274 (14):1143-8.

Kirmeyer, G. J., M. Friedman, K. Martel, D. Howie, M. W. LeChevallier, M. Abbaszadegan, M. R. Karim, J. E. Funk, and J. Harbour. 2001. Pathogen intrusion

into the distibution system. Denver: American Water Works Association Research Foundation and American Water Works Association.

Klein, M., W. D. Flanders, and P. E. Tolbert. 2002. Variances may be underestimated using available software for generalized additive models. *American Journal of Epidemiology* 155 (11):S106.

Kramer, M. H., B. L. Herwaldt, G. F. Craun, R. L. Calderon, and D. D. Juranek. 1996. Surveillance for waterborne-disease outbreaks--United States, 1993-1994. *MMWR CDC Surveill Summ* 45 (1):1-33.

LeChevallier, M. W., W. D. Norton, and R. G. Lee. 1991. *Giardia* and *Cryptosporidium* spp. in filtered drinking water supplies. *Appl Environ Microbiol* 57 (9):2617-21.

———. 1991. Occurrence of *Giardia* and *Cryptosporidium* spp. in surface water supplies. *Appl Environ Microbiol* 57 (9):2610-6.

LeChevallier, M. W. 1999. The case for maintaining a disinfectant residual. *J Am Water Works Assoc* 91:86-94.

LeChevallier, M. W., T. M. Evans, and R. J. Seidler. 1981. Effect of turbidity on chlorination efficiency and bacterial persistence in drinking water. *Appl Environ Microbiol* 42 (1):159-67.

LeChevallier, M. W., R. W. Gullick, M. R. Karim, M. Friedman, and J. E. Funk. 2003. The potential for health risks from intrusion of contaminants into the distribution system from pressure transients. *J Water Health* 1 (1):3-14.

LeChevallier, M. W., M. R. Karim, R. Aboytes, R. W. Gullick, J. Weihe, B. Earnhardt, J. Mohr, J. Starcevich, J. Case, J. S. Rosen, J. Sobrinho, J. L. Clancy, R. M. McCuin, J. E. Funk, and D. J. Wood. 2004. Profiling water quality parameters:

From source water to the household tap: American Water Works Association Research Foundation.

LeChevallier, M. W., and W. D. Norton. Treatments to address source water concerns: Protozoa. In *Safety of Water Disinfection: Balancing Chemical and Microbial Risks*, edited by G. Craun. Washington, DC: ILSI Press.

LeChevallier, M. W., N. J. Welch, and D. B. Smith. 1996. Full-scale studies of factors related to coliform regrowth in drinking water. *Appl Environ Microbiol* 62 (7):2201-11.

Lee, S. H., D. A. Levy, G. F. Craun, M. J. Beach, and R. L. Calderon. 2002. Surveillance for waterborne-disease outbreaks--United States, 1999-2000. *MMWR Surveill Summ* 51 (8):1-47.

Levy, D. A., M. S. Bens, G. F. Craun, R. L. Calderon, and B. L. Herwaldt. 1998. Surveillance for waterborne-disease outbreaks--United States, 1995-1996. *MMWR CDC Surveill Summ* 47 (5):1-34.

Liang, J. L., E. J. Dziuban, G. F. Craun, V. Hill, M. R. Moore, R. J. Gelting, R. L. Calderon, M. J. Beach, and S. L. Roy. 2006. Surveillance for waterborne disease and outbreaks associated with drinking water and water not intended for drinking--United States, 2003-2004. *MMWR Surveill Summ* 55 (12):31-65.

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

Logsdon, G. S., V. C. Thurman, E. S. Frindt, and J. G. Stoecker. 1985. Evaluating sedimentation and various filter media for removal of *Giardia* cysts. *J Am Water Works Assoc* 77:61-6.

Lu, N., M. E. Samuels, L. Shi, S. L. Baker, S. H. Glover, and J. M. Sanders. 2004. Child day care risks of common infectious diseases revisited. *Child Care Health Dev* 30 (4):361-8.

MacKenzie, W. R., N. J. Hoxie, M. E. Proctor, M. S. Gradus, K. A. Blair, D. E. Peterson, J. J. Kazmierczak, D. G. Addiss, K. R. Fox, J. B. Rose, and J. P. Davis. 1994. A massive outbreak in Milwaukee of cryptosporidium infection transmitted through the public water supply. *N Engl J Med* 331 (3):161-7.

MacKenzie, W. R., W. L. Schell, K. A. Blair, D. G. Addiss, D. E. Peterson, N. J. Hoxie, J. J. Kazmierczak, and J. P. Davis. 1995. Massive outbreak of waterborne cryptosporidium infection in Milwaukee, Wisconsin: recurrence of illness and risk of secondary transmission. *Clin Infect Dis* 21 (1):57-62.

Marshall, M. M., D. Naumovitz, Y. Ortega, and C. R. Sterling. 1997. Waterborne protozoan pathogens. *Clin Microbiol Rev* 10 (1):67-85.

Mathias, R. G., P. D. Riben, and W. D. Osei. 1992. Lack of an association between endemic giardiasis and a drinking water source. *Can J Public Health* 83 (5):382-4.

Mead, P. S., L. Slutsker, V. Dietz, L. F. McCaig, J. S. Bresee, C. Shapiro, P. M. Griffin, and R. V. Tauxe. 1999. Food-related illness and death in the United States. *Emerg Infect Dis* 5 (5):607-25.

Meinhardt, P. L., D. P. Casemore, and K. B. Miller. 1996. Epidemiologic aspects of human cryptosporidiosis and the role of waterborne transmission. *Epidemiol Rev* 18 (2):118-36.

Messner, M., S. Shaw, S. Regli, K. Rotert, V. Blank, and J. Soller. 2006. An approach for developing a national estimate of waterborne disease due to drinking water and a national estimate model application. *J Water Health* 4 (Suppl 2):201-40.

Metzger, K. B., P. E. Tolbert, M. Klein, J. Peel, W. D. Flanders, K. Todd, J. A. Mulholland, P. B. Ryan, H. Frumkin, and AIRES investigators. 2002. Particulate air pollution and cardiovascular emergency department visits in Atlanta, Georgia, 1998-2000. *Am J Epidemiol* 155 (11):S36.

Miller, D. P., J. A. Kay, K. Shea, N. Ziyadeh, C. Cali, C. Black, and A. M. Walker. 2004. Incidence of thrombotic thrombocytopenic purpura/hemolytic uremic syndrome. *Epidemiology* 15 (2):208-15.

Mills, C. J., R. J. Bull, K. P. Cantor, J. Reif, S. E. Hrudey, and P. Huston. 1998. Workshop report. Health risks of drinking water chlorination by-products: report of an expert working group. *Chronic Dis Can* 19 (3):91-102.

Moe, C. L. 2002. Waterborne transmission of infectious agents. In *Manual of Environmental Microbiology*, edited by C. Hurst, R. Crawford, G. Knudsen, M. McInerney and L. Stetzenbach: ASM Press.

Moeller, D. W. 1992. *Environmental Health*. Cambridge, MA: Harvard University Press.

Moore, A. C., B. L. Herwaldt, G. F. Craun, R. L. Calderon, A. K. Highsmith, and D. D. Juranek. 1993. Surveillance for waterborne disease outbreaks--United States, 1991-1992. *MMWR CDC Surveill Summ* 42 (5):1-22.

Morris, R. D., and R. Levin. 1995. Estimating the incidence of waterborne infectious disease related to drinking water in the United States. In *Assessing and Managing Health Risks from Drinking Water Contamination: Approaches and Applications.*,

edited by E. Reichard and G. Zapponi. Wallingford, UK: International Association of Hydrological Sciences.

Morris, R. D., E. N. Naumova, and J. K. Griffiths. 1998. Did Milwaukee experience waterborne cryptosporidiosis before the large documented outbreak in 1993? *Epidemiology* 9 (3):264-70.

Morris, R. D., E. N. Naumova, R. Levin, and R. L. Munasinghe. 1996. Temporal variation in drinking water turbidity and diagnosed gastroenteritis in Milwaukee. *Am J Public Health* 86 (2):237-9.

Mounts, A. W., R. C. Holman, M. J. Clarke, J. S. Bresee, and R. I. Glass. 1999. Trends in hospitalizations associated with gastroenteritis among adults in the United States, 1979-1995. *Epidemiol Infect* 123 (1):1-8.

Naumova, E. N., A. I. Egorov, R. D. Morris, and J. K. Griffiths. 2003. The elderly and waterborne *Cryptosporidium* infection: gastroenteritis hospitalizations before and during the 1993 Milwaukee outbreak. *Emerg Infect Dis* 9 (4):418-25.

Nieminski, E.C. 1992. *Giardia* and *Cryptosporidium* - Where do the cysts go. Paper read at American Water Works Association Water Quality Technology Conference.

Nime, F. A., J. D. Burek, D. L. Page, M. A. Holscher, and J. H. Yardley. 1976. Acute enterocolitis in a human being infected with the protozoan *Cryptosporidium*. *Gastroenterology* 70 (4):592-8.

Nygard, K., Y. Andersson, J. A. Rottingen, A. Svensson, J. Lindback, T. Kistemann, and J. Giesecke. 2004. Association between environmental risk factors and *campylobacter* infections in Sweden. *Epidemiol Infect* 132 (2):317-25.

Nygard, K., E. Wahl, T. Krogh, O. A. Tveit, E. Bohleng, A. Tverdal, and P. Aavitsland. 2007. Breaks and maintenance work in the water distribution systems and gastrointestinal illness: a cohort study. *Int J Epidemiol* In press.

O'Connor, D. R. 2002. Report of the Walkerton Inquiry: The Events of May 2000 and Related Issues. Toronto, Ontario: Ontario Ministry of the Attorney General.

Office of Water. 1999. 25 years of the Safe Drinking Water Act: history and trends: U.S. Environmental Protection Agency.

———. 2000. EPA studies on sensitive subpopulations and drinking water contaminants. Washington, DC: U.S. Environmental Protection Agency.

Parashar, U. D., R. C. Holman, M. J. Clarke, J. S. Bresee, and R. I. Glass. 1998. Hospitalizations associated with rotavirus diarrhea in the United States, 1993 through 1995: surveillance based on the new ICD-9-CM rotavirus-specific diagnostic code. *J Infect Dis* 177 (1):13-7.

Power Analysis and Sample Size 6.0. NCSS, Kaysville, UT.

Payment, P. 1998. Waterborne viruses and parasites: resistance to treatment and disinfection. In *OECD Workshop Molecular Methods for Safe Drinking Water*: EAWAG.

———. 1999. Poor efficacy of residual chlorine disinfectant in drinking water to inactivate waterborne pathogens in distribution systems. *Can J Microbiol* 45 (8):709-15.

Payment, P., and P. R. Hunter. 2001. Endemic and epidemic infectious intestinal disease and its relationship to drinking water. In *Water Quality: Guidelines, Standards, and Health*, edited by L. Fewtrell and J. Bartram. London: IWA Publishing.

Payment, P., E. Franco, and G. S. Fout. 1994. Incidence of Norwalk virus infections during a prospective epidemiological study of drinking water-related gastrointestinal illness. *Can J Microbiol* 40 (10):805-9.

Payment, P., E. Franco, L. Richardson, and J. Siemiatycki. 1991. Gastrointestinal health effects associated with the consumption of drinking water produced by point-of-use domestic reverse-osmosis filtration units. *Appl Environ Microbiol* 57 (4):945-8.

Payment, P., E. Franco, and J. Siemiatycki. 1993. Absence of relationship between health effects due to tap water consumption and drinking water quality parameters. *Water Sci Technol* 27 (3-4):137-43.

Payment, P., F. Gamache, and G. Paquette. 1988. Microbiological and virological analysis of water from two water filtration plants and their distribution systems. *Can J Microbiol* 34 (12):1304-9.

Payment, P., L. Richardson, J. Siemiatycki, R. Dewar, M. Edwardes, and E. Franco. 1991. A randomized trial to evaluate the risk of gastrointestinal disease due to consumption of drinking water meeting current microbiological standards. *Am J Public Health* 81 (6):703-8.

Payment, P., J. Siemiatycki, L. Richardson, G. Renaud, E. Franco, and M. Prevost. 1997. A prospective epidemiological study of gastrointestinal health effect due to the consumption of drinking water. *Int J Environ Health Res* 7:5-31.

Peel, J. L., P. E. Tolbert, M. Klein, K. B. Metzger, W. D. Flanders, K. Todd, J. A. Mulholland, P. B. Ryan, and H. Frumkin. 2005. Ambient air pollution and respiratory emergency department visits. *Epidemiology* 16 (2):164-74.

Perz, J. F., F. K. Ennever, and S. M. Le Blancq. 1998. *Cryptosporidium* in tap water: comparison of predicted risks with observed levels of disease. *Am J Epidemiol* 147 (3):289-301.

Peterson, C. A., and R. L. Calderon. 2003. Trends in enteric disease as a cause of death in the United States, 1989-1996. *Am J Epidemiol* 157 (1):58-65.

Physicians for Social Responsibility. 2000. Drinking water and disease: what health care providers should know: Physicians for Social Responsibility.

Pontius, F. W. 1997. Implementing the 1996 SDWA Amendments. *J Am Water Works Assoc* 89 (3):18-36.

———. 2001. Regulatory update for 2001 and beyond. *J Am Water Works Assoc* 93 (2):66-80.

———. 2003. Update of USEPA's drinking water regulations. *J Am Water Works Assoc* 95 (3):57-68.

Postel, S. L. 2000. Water and world population growth. *J Am Water Works Assoc* 92 (4):131-8.

Roberson, A. J. 2003. Complexities of the new drinking water regulations - everything you wanted to know but were afraid to ask. *J Am Water Works Assoc* 95 (3):48-56.

Rzezutka, A., and N. Cook. 2004. Survival of human enteric viruses in the environment and food. *FEMS Microbiol Rev* 28 (4):441-53.

SAS 9.1. SAS Institute, Inc., Cary, North Carolina.

Schwartz, J., and R. Levin. 1999. Drinking water turbidity and health. *Epidemiology* 10 (1):86-90.

Schwartz, J., R. Levin, and R. Goldstein. 2000. Drinking water turbidity and gastrointestinal illness in the elderly of Philadelphia. *J Epidemiol Community Health* 54 (1):45-51.

Schwartz, J., R. Levin, and K. Hodge. 1997. Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* 8 (6):615-20.

Seyfried, P. L., R. S. Tobin, N. E. Brown, and P. F. Ness. 1985. A prospective study of swimming-related illness. I. Swimming-associated health risk. *Am J Public Health* 75 (9):1068-70.

Sinclair, M. I., and C. K. Fairley. 2000. Drinking water and endemic gastrointestinal illness. *J Epidemiol Community Health* 54 (10):728.

Smith, H., and R. A. Nichols. 2006. Zoonotic protozoa - food for thought. *Parassitologia* 48 (1-2):101-4.

Smith, H. V., W. J. Patterson, R. Hardie, L. A. Greene, C. Benton, W. Tulloch, R. A. Gilmour, R. W. Girdwood, J. C. Sharp, and G. I. Forbes. 1989. An outbreak of waterborne cryptosporidiosis caused by post-treatment contamination. *Epidemiol Infect* 103 (3):703-15.

Sorvillo, F., G. Beall, P. A. Turner, V. L. Beer, A. A. Kovacs, P. Kraus, D. Masters, and P. R. Kerndt. 1998. Seasonality and factors associated with cryptosporidiosis among individuals with HIV infection. *Epidemiol Infect* 121 (1):197-204.

Tardiff, R. G., M. L. Carson, and M. E. Ginevan. 2006. Updated weight of evidence for an association between adverse reproductive and developmental effects and exposure to disinfection by-products. *Regul Toxicol Pharmacol* 45 (2):185-205.

Tolbert, P. E., M. Klein, K. B. Metzger, J. Peel, D. W. Flanders, K. Todd, J. A. Mulholland, P. B. Ryan, H. Frumpkin, and ARIES investigators. 2002. Particulate pollution and cardiorespiratory outcomes in Atlanta, August 1998 - August 2000 (AIRES/SOPHIA studies). Paper read at Air Quality III Conference, September 2002, at Arlington, VA.

Tolbert, P. E., M. Klein, K. B. Metzger, J. Peel, W. D. Flanders, K. Todd, J. A. Mulholland, P. B. Ryan, and H. Frumkin. 2000. Interim results of the study of particulates and health in Atlanta (SOPHIA). *J Expo Anal Environ Epidemiol* 10 (5):446-60.

Trussell, R. 1999. Safeguarding distribution system integrity. *J Am Water Works Assoc* 91:46-54.

———. 2006. Water treatment: the past 30 years. *J Am Water Works Assoc* 98 (3):100-9.

U.S. Census Bureau. 1990. 1990 Summary Tape File 1 (STF 1) - 100-Percent data. Washington, DC: American FactFinder, U.S. Census Bureau.

———. 2000. General Housing Characteristics: 2000. Census Summary File 1 (SF 1) 100-Percent Data. Washington, DC: American FactFinder, U.S. Census Bureau.

———. 2001. Census 2000 ZCTAs: Zip Code Tabulation Areas for Census Bureau Data Products: U.S. Census Bureau.

U.S. Centers for Disease Control and Prevention. 2007. *Viral Gastroenteritis* [Web site]. U.S. Centers for Disease Control and Prevention, August 3, 2006 [cited June 20 2007]. Available from http://www.cdc.gov/ncidod/dvrd/revb/gastro/faq.htm.

———. 1999. From the U.S. Centers for Disease Control and Prevention. Control of infectious diseases, 1900-1999. *JAMA* 282 (11):1029-32.

U.S. Congress. 1969. The National Environmental Policy Act. In *42 U.S.C. s/s/ 4321 et seq.*

———. 1976. The Toxic Substances Control Act. In *15 U.S.C. s/s 2601 et seq.*

———. 1977. Clean Water Act. In *33 U.S.C. 1251 et seq.*

———. 1980. The Comprehensive Environmental Response, Compensation, and Liability Act (Superfund). In *42 U.S.C. s/s 9601 et seq.*

———. 1990. The Pollution Prevention Act. In *42 U.S.C. 13101 and 13102, s/s/ et seq.*

U.S. Environmental Protection Agency. 1975. National Interim Primary Drinking Water Regulations. *Federal Register* 40 (248):59566.

———. 1979. Total Trihalomethanes National Interim Primary Drinking Water Regulation. *Federal Register* 44 (228):68624.

———. 1989. Filtration, Disinfection, Turbidity, *Giardia lamblia*, Viruses, *Legionella*, and Heterotrophic Bacteria. Final Rule. *Federal Register* 54 (124):27486.

———. 1989. Total Coliforms (Inclduing Fecal Coliforms and *E. coli*). Final Rule. *Federal Register* 54 (124):27544.

———. 1991. Lead and Copper. Final Rule. *Federal Register* 56 (110):26460.

———. 1996. Monitoring Requirements for Public Drinking Water Supplies; Final Rule. *Federal Register* 61 (94):24354.

———. 1996 Air Quality Criteria for Particulate Matter, Vol. III of III: U.S. Environmental Protection Agency.

———. 1998. Disinfectants and Disinfection By-products. Final Rule. *Federal Register* 63 (241):69390.

———. 1998. Interim Enhanced Surface Water Treatment. Final Rule. *Federal Register* 64 (241):69487.

———. 1998. National Primary Drinking Water Regulation: Consumer Confidence Reports; Final Rule. *Federal Register* 63 (160):44512.

———. 1999. Importance of Turbidity. In *EPA Guidance Manual: Turbidity Provisions*, edited by U.S. Environmental Protection Agency.

———. 1999. Understanding the Safe Drinking Water Act. Washington, DC: U.S. Environmental Protection Agency.

———. 2000. Ground Water Rule. Proposed Rule. *Federal Register* 65 (91):30194.

———. 2000. The quality of our nation's water: 305b water quality report: U.S. Environmental Protection Agency.

———. 2001. Filter Backwash Recycling. Final Rule. *Federal Register* 68 (111):31086.

———. 2001. Revisions to the Interim Enhanced Surface Water Treatment Rule (IESWTR), the Stage 1 Disinfectants and Disinfectant Byproducts Rule (Stage 1 DBPR), and Revisions to State Primacy Requirements to Implement the Safe Drinking Water Act (SDWA) Amendments; Final Rule. *Federal Register* 66 (10):3770.

———. 2002. Long Term 1 Enhanced Surface Water Treatment. Final Rule. *Federal Register* 67 (9):1812.

———. 2003. Stage 2 Disinfectants and Disinfection Byproducts Rule; National Primary and Secondary Drinking Water Regulations: Approval of Analytical Methods for Chemical Contaminants; Proposed Rule. *Federal Register* 68 (159):49548.

———. 2006. National Primary Drinking Water Regulations: Ground Water Rule; Final Rule. *Federal Register* 71 (216):65573.

———. 2006. National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rule. *Federal Register* 71 (3):653.

———. 2006. National Primary Drinking Water Regulations: Stage 2 Disinfectants and Disinfection Byproducts Rule; Final Rule. *Federal Register* 71 (2):387.

———. 2007. *Safe Drinking Water Information System (SDWIS)* [web site]. U.S. Environmental Protection Agency 2007 [cited 2007]. Available from http://www.epa.gov/enviro/html/sdwis/sdwis_ov.html.

Vakil, N. B., S. M. Schwartz, B. P. Buggy, C. F. Brummitt, M. Kherellah, D. M. Letzer, I. H. Gilson, and P. G. Jones. 1996. Biliary cryptosporidiosis in HIV-infected people after the waterborne outbreak of cryptosporidiosis in Milwaukee. *N Engl J Med* 334 (1):19-23.

Wade, T. J., R. L. Calderon, E. Sams, M. Beach, K. P. Brenner, A. H. Williams, and A. P. Dufour. 2006. Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness. *Environ Health Perspect* 114 (1):24-8.

Waller, K., S. H. Swan, G. DeLorenze, and B. Hopkins. 1998. Trihalomethanes in drinking water and spontaneous abortion. *Epidemiology* 9 (2):134-40.

Wheeler, J.G., S. Dinesh, J.M. Cowden, P.G. Wall, L.C. Rodrigues, D.S. Tompkins, M.J. Hudson, and P.J. Roderick. 1999. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *British Medical Journal* 318:1046-50.

World Health Organization. 1999. Geographical information systems (GIS): Mapping for epidmeiolgical surveillance. *Weekly Epidemiological Record* 74:281-5.

———. 2000. The world health report 2000. Geneva: World Health Organization.

———. 2003. Causes of death: global, regional, and country-specific estimates of death by cause, age, and sex. Geneva: World Health Organization.

Yang, S. 2003. Home-treated water no better than plain tap in preventing gastrointestinal illness, finds new study. *UC Berkeley News*, September 25, 2003.

Yoder, J. S., B. G. Blackburn, G. F. Craun, V. Hill, D. A. Levy, N. Chen, S. H. Lee, R. L. Calderon, and M. J. Beach. 2004. Surveillance for waterborne-disease outbreaks associated with recreational water--United States, 2001-2002. *MMWR Surveill Summ* 53 (8):1-22.

Zeger, S. L. 1988. A regression model for time series of counts. *Biometrika* 75 (4):621-9.

APPENDIX A

ADDITIONAL ANALYSIS OF DRINKING WATER RESIDENCE TIME AND EMERGENCY DEPARTMENT VISITS FOR GASTROINTESTINAL ILLNESS IN ATLANTA, 1996 – 2003

Objective

The objective of this appendix is to describe additional analyses conducted in the course of assessing the association of the estimated drinking water residence time and emergency department (ED) visits for gastrointestinal (GI) illness. Several additional analyses were considered:

- Use of a less refined measure of distribution system exposure than water residence time, Euclidean (straight line) distance from the zip code centroid to the treatment plant of service.
- Consideration of the water residence time exposure as a continuous variable
- Use of generalized estimating equations (GEE)
- Assessment of the impact of missing data
- Assessment of the impact of a high-profile recreational waterborne outbreak in the study area during the study period
- Consideration of the chlorine residual levels present in the distribution system and their relation to the estimated water residence time

Additional Methods, Results, and Conclusions

Euclidean distance exposure

The model in which the Eucldiean distance exposure was considered was the same as that described for the model in which the water residence time exposure was considered (described in Chapter 6). The distance exposure was considered as a three-level categorical variable with cut-points at the 10^{th} and 90^{th} percentile, analogous to the categorization used for the *a priori* water residence time exposure. The correlation between the average water residence time and the Eucldiean distance was high, with a Spearman correlation coefficient of 0.78 ($p<0.0001$). The same time period, 1996 to 2003, was considered for the Euclidean distance sensitivity analysis as for the *a priori* water residence time analysis.

The results from the models in which the Euclidean distance exposure was considered were similar to those observed from the models in which water residence time was considered, although neither of the comparisons, risk of ED visits for GI illness among people living in zip codes with short or long compared to intermediate residence distances, were statistically significant ($\alpha = 0.05$) (Figure A.1). The risk ratio estimate for GI illness ED visits among people living in zip codes near to the treatment plant compared to those living at an intermediate distance was 0.993 (95% confidence interval (CI): 0.933, 1.057) and the risk ratio estimate comparing those living far from the plant to those living at an intermediate distance was 1.061 (95% CI: 0.997, 1.130). Although the results were similar, the conclusions drawn from these results based on statistical significance would differ from those suggested by the results of the analysis in which the water residence time exposure was considered.

An examination of the results of the two analyses suggests that the impact of using the Euclidean distance exposure was not as great as might be concluded from significance testing. The risk ratio estimates from the models in which water residence time estimated from Hydraulic Model (HM) 2 was considered were more similar to the results observed from the models in which Euclidean distance was considered than models in which the water residence time estimates derived from HM1 were considered. These results suggest that Euclidean distance may be an adequate proxy for water residence time in terms of assessing the risk of GI illness ED visits from drinking water distribution system exposure. While water residence time is a more refined indicator of distribution system exposure, deriving these estimates can be difficult and time-consuming, if possible at all based on the availability of hydraulic models. Estimating the Euclidean distance between a location, here zip code centroid, and the treatment plant serving that location requires only the latitude and longitude of the location and the treatment plant, and is therefore much easier to estimate.

Water residence time as a continuous variable

The water residence time variable was considered in the *a priori* analytical models as a categorical variable, with cut-points at the 10^{th} and 90^{th} percentile, because our goal was to examine whether people living at the extremes of the distribution system were at increased risk. In a sensitivity analysis, we also considered the water residence time variable continuously. The models contained the same covariates as that described for the *a priori* models and they were stratified on the hydraulic model from which the water residence time estimates were derived.

The results of the models in which water residence time estimates from the earlier hydraulic model (HM1) were considered as a continuous variable suggest a significant association with a continuous one day (24 hour) increase in water residence time and risk of GI illness (RR = 1.091; 95% CI: 1.035, 1.149). These results were not consistent with those observed when the water residence time estimates from the later hydraulic model (HM2) were considered (RR = 1.017; 95% CI: 0.998, 1.036).

We also considered the continuous water residence time variable using a cubic spline. The use of a spline allows the linear relationship between GI illness ED visits and water residence time to vary for different areas in the distribution of water residence times. We added knots at the 10^{th}, 50^{th}, and 90^{th} percentiles. The results of this analysis suggest that the categorization scheme we used for the *a priori* water residence time exposure was an appropriate dose response (Figures A.2 and A.3). There was little change in the risk of GI illness ED visits in the intermediate category, defined as the 11^{th} to 89^{th} percentile of water residence times. The increase in risk for GI illness appears to be largely isolated to those zip codes served by drinking water with the longest residence time, over the 90^{th} percentile. The magnitude of the increase appears stronger when water residence time estimates from the earlier hydraulic model were considered compared to those from the later hydraulic model.

Using GEE models

As discussed in Chapter 6, the unit of analysis for this study was the zip code and because several strata within zip code were considered, there was the potential for the observations in the database to have some dependence. The logistic regression model using generalized linear model (GLM) estimating procedures used in the *a priori* analysis

assumes independent observations, and lack of independence can lead to underestimation (or in some cases overestimation) of the variance. In this sensitivity analysis we considered generalized estimating equations (Liang and Zeger 1986), which allowed us to accommodate the potential correlation of the observations in the database. We were unable to run GEE models including all of the variables designated *a priori*; the models would not converge. Therefore we ran both GLM and GEE models excluding the interaction variables for distance from hospital to zip code centroid and hospital.

The results of this analysis suggest that there was little impact on results from autocorrelation of the data (Figure A.4). The point estimates from the GEE and GLM models were essentially equal. The confidence interval was tighter for the comparison of short water residence time to intermediate water residence time for estimates from HM2 and the confidence interval was wider for the comparison of long water residence time to intermediate water residence time for estimates from HM2. The risk ratio estimates from both the GLM and GEE estimates were slightly more positive than those observed from the *a priori* model.

Impact of missing data

Missing Census data. One of the zip codes located in the service area of the utility considered in this study had to be excluded because of missing Census information. This zip code was assigned to an Air Force base and, likely because of medical services available on the base, very few ED visits were from residents of this zip code. To assess the impact of the exclusion of this zip code, results from three models were considered: the *a priori* model that included variables using Census data and

excluded the zip code, a model that excluded the Census data and included the excluded zip code, and a model that excluded both the Census data and the zip code.

The results suggest that excluding the Census variables had a greater impact on results than excluding the single zip code for which Census data were missing (Figure A.5). The risk ratio estimates from the *a priori* model and the model from which both the Census variables and the zip code were excluded differed, and the only difference in these models was the exclusion of the Census variables. By contrast, when the models in which neither the Census variables nor data from the zip code were included were compared to the models in which the zip code was included and the Census variables were therefore necessarily excluded, the risk ratio point estimates remained virtually the same.

Missing Medicaid payment information. Data from one of the 28 hospitals contributing ED data had to be excluded due to missing information regarding Medicaid payment status (whether the patient paid for the ED visit using Medicaid). Using a similar method to that described previously to assess the impact of the excluded zip code, three models were considered: the *a priori* model that excluded visits from the hospital and included the variable indicating Medicaid payment status, a model that excluded ED visits from the hospital and the Medicaid payment variable, and a model that excluded the Medicaid payment variable and included data on ED visits from the excluded hospital.

The results of this sensitivity analysis again suggest that the inclusion of the *a priori* variable, Medicaid payment status, was more important than the inclusion of data from the one hospital in the ED database for which this variable was not available (Figure

A.6). The results varied little between the three models. However, based on statistical significance ($\alpha = 0.05$), the conclusion for the association between living in a zip code with a long average water residence time compared to an intermediate water residence time in terms of risk for GI illness ED visits would be different for water residence time estimates derived from HM2 if the Medicaid variable were excluded.

Impact of recreational water outbreak

In 1998 there was a highly publicized outbreak of pathogenic *E. coli* O157:H7 at a recreational water park in the study area (Gilbert and Blake 1998). The source of the outbreak was traced to a contaminated children's pool. There were 25 confirmed cases and two deaths. Because this outbreak could have resulted in unreported cases of GI illness and because people affected by this outbreak in the study area may not have been homogeneous in their spatial distribution, we considered a sensitivity analysis in which the time period surrounding the outbreak (summer of 1998) was excluded.

The results of the sensitivity analysis suggest that there was little impact of the outbreak on our results (Figure A.7). The risk ratio estimates for the comparison of the risk of GI illness ED visits among people living in zip codes with short average water residence times compared to people living in zip codes with intermediate average water residence times were similar. When the outbreak period was removed from the database, the risk ratio estimate for the comparison of GI illness risk among people living in zip codes with long average water residence times compared to intermediate water residence times was even stronger than the *a priori* point estimate. This result may be attributable to the fact that the zip code in which the water park was located, and therefore theoretically the zip code likely to have the highest number of GI illness cases resulting

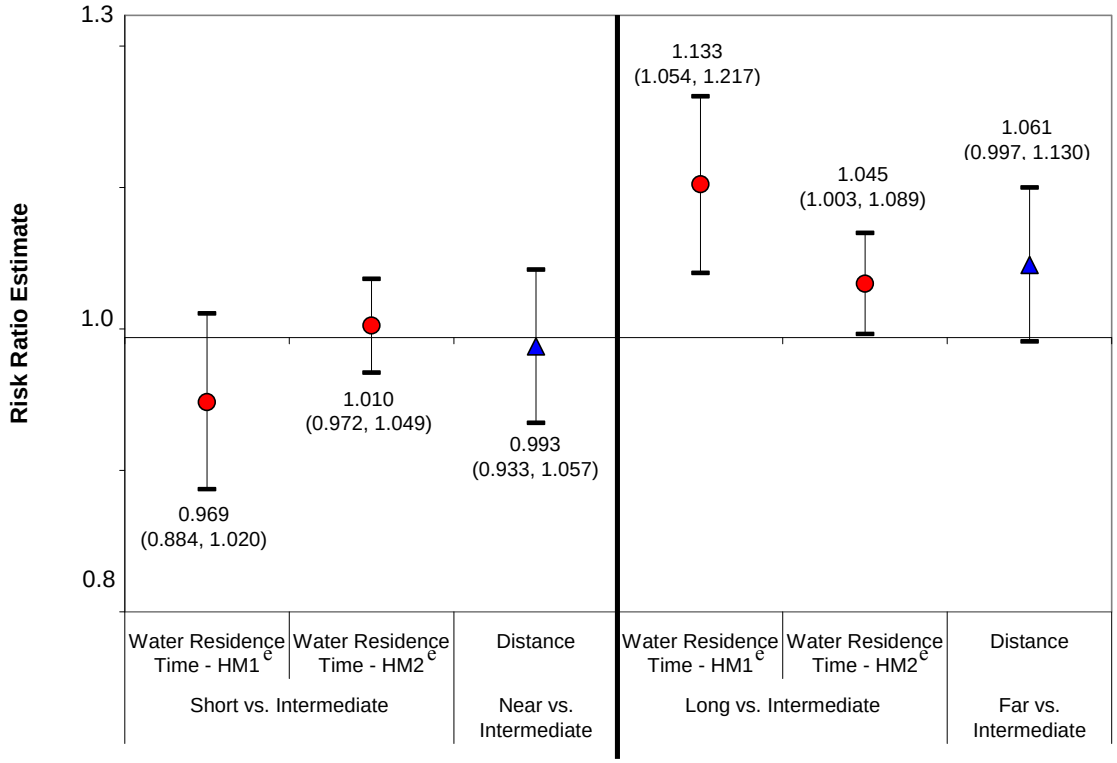
from the outbreak, was classified in the referent category, which could have resulted in a bias in the negative direction.

Chlorine residual levels in the distribution system

We have not yet conducted formal analyses regarding the association between chlorine residual levels in the distribution system, another indicator of waterborne pathogen exposure risk, and risk of GI illness. However, chlorine residuals are likely to be associated with water residence times and we briefly explored the relationship between these two indicators of distribution system contamination.

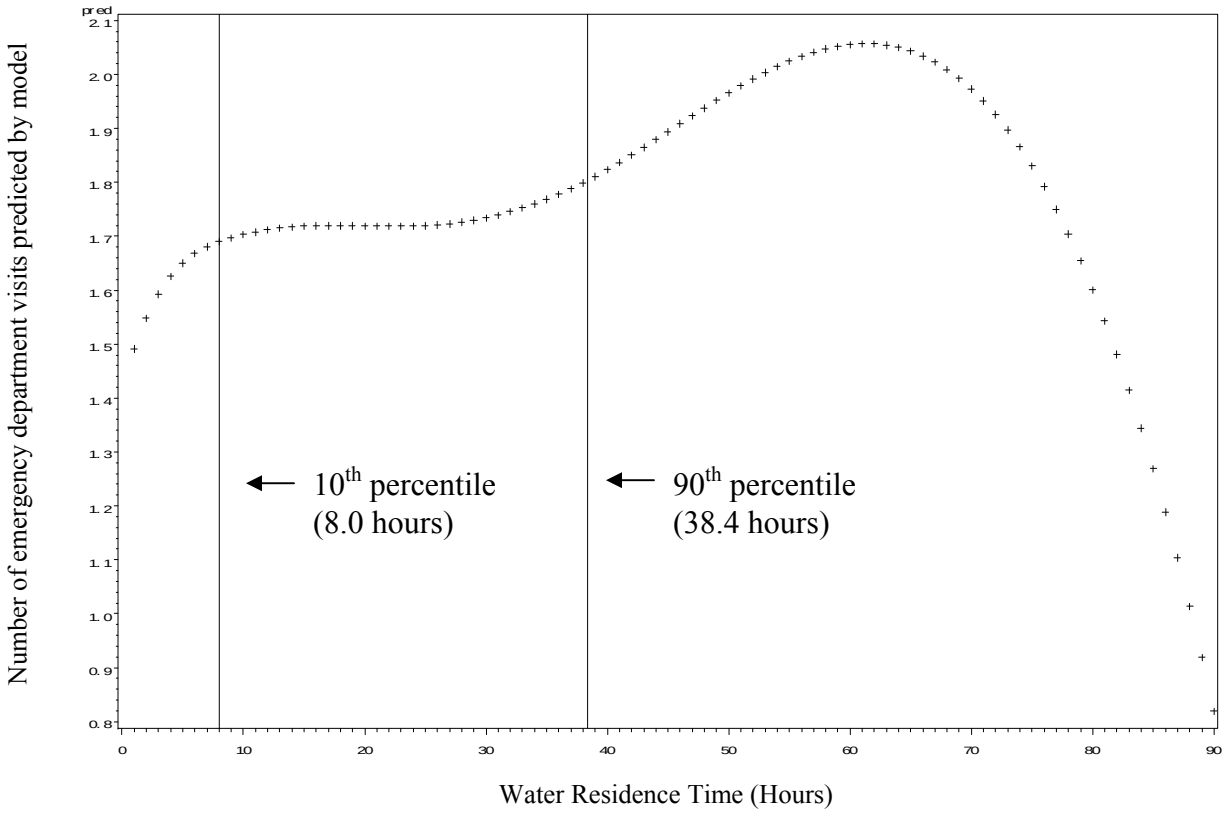
All utilities are required by the Georgia Department of Natural Resources to collect samples from the distribution system each month, with the number of sampling sites based on the size of the distribution system. We obtained these records and assigned each chlorine residual measurement to a zip code. The chlorine residual levels averaged over each zip code have an inverse correlation with water residence time (Spearman correlation coefficient = -0.66, $p<0.0001$). This inverse relationship is expected given the tendency of chlorine residual level to decrease as water travels through the distribution system. Examination of the average chlorine residual levels in each zip code for each year suggest that the relative levels of chlorine residual did not vary much across zip codes over time, but that overall the actual level of chlorine decreased over time (Figure A.8).

Figure A.1: Risk ratio estimates[a] for gastrointestinal illness emergency department visits among people living in zip codes with short[b] drinking water residence times and long[b] drinking water residence times compared to intermediate[b] water residence times, compared to the risk ratio estimates for people living in zip codes near[c] the drinking water treatment plant and far[c] from the drinking water treatment plant compared to people living at an intermediate[c] distance from the treatment plant, Atlanta[d], 1996 – 2003



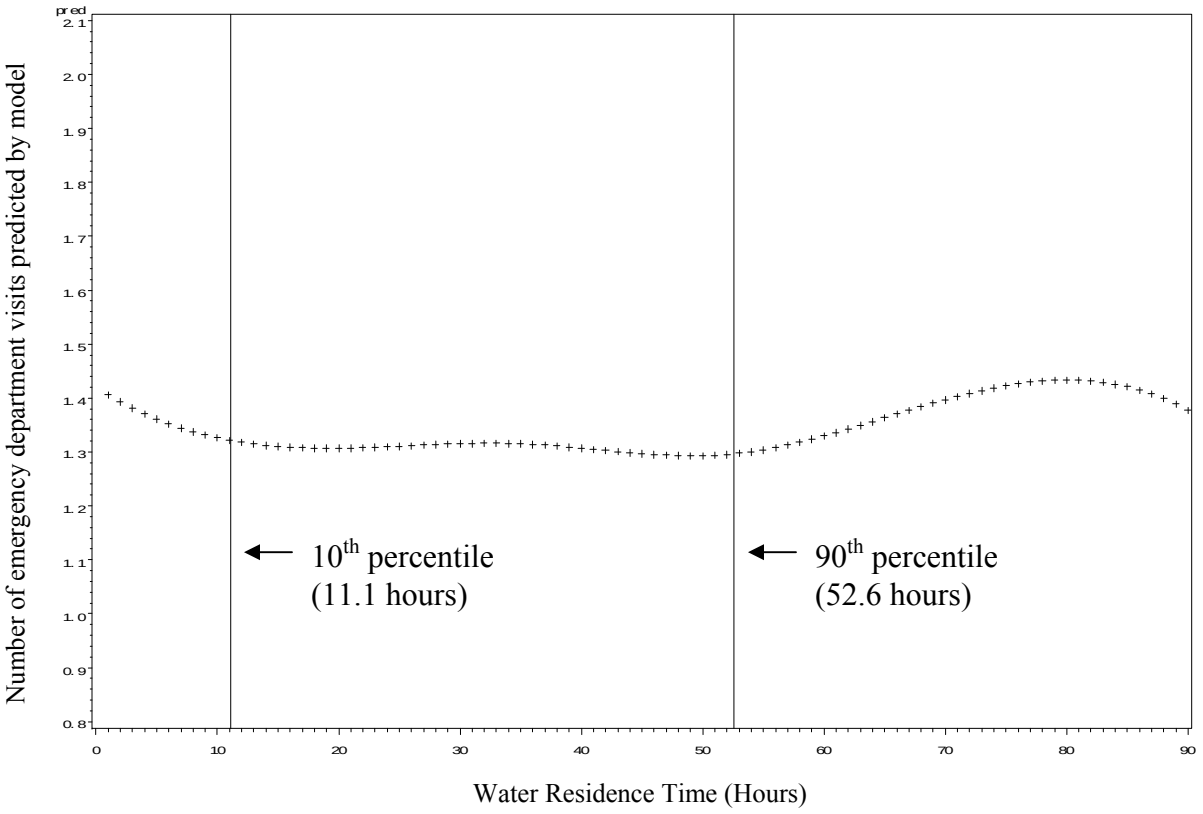
[a] Error bars indicate 95 percent confidence intervals.
[b] Short = ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th – 89th percentile; long = ≥ 90th percentile
[c] Euclidean distance from zip code centroid to drinking water treatment plant serving that zip code ; near = ≤ 10th percentile of all Euclidean distance estimates; intermediate = 11th – 89th percentile; far = ≥ 90th percentile
[d] One county considered
[e] HM = hydraulic model from which the water residence time estimates were derived; HM1 includes 1996 – 1998, HM2 includes 1999 – 2003

Figure A.2: Plot of the number of emergency department visits for gastrointestinal illness predicted from a regression model in which the water residence time estimates were included as a cubic spline with knots at the 10th, 50th, and 90th percentile of all water residence times, Atlanta[a], water residence time estimates derived from hydraulic model 1 only (1996 – 1998)



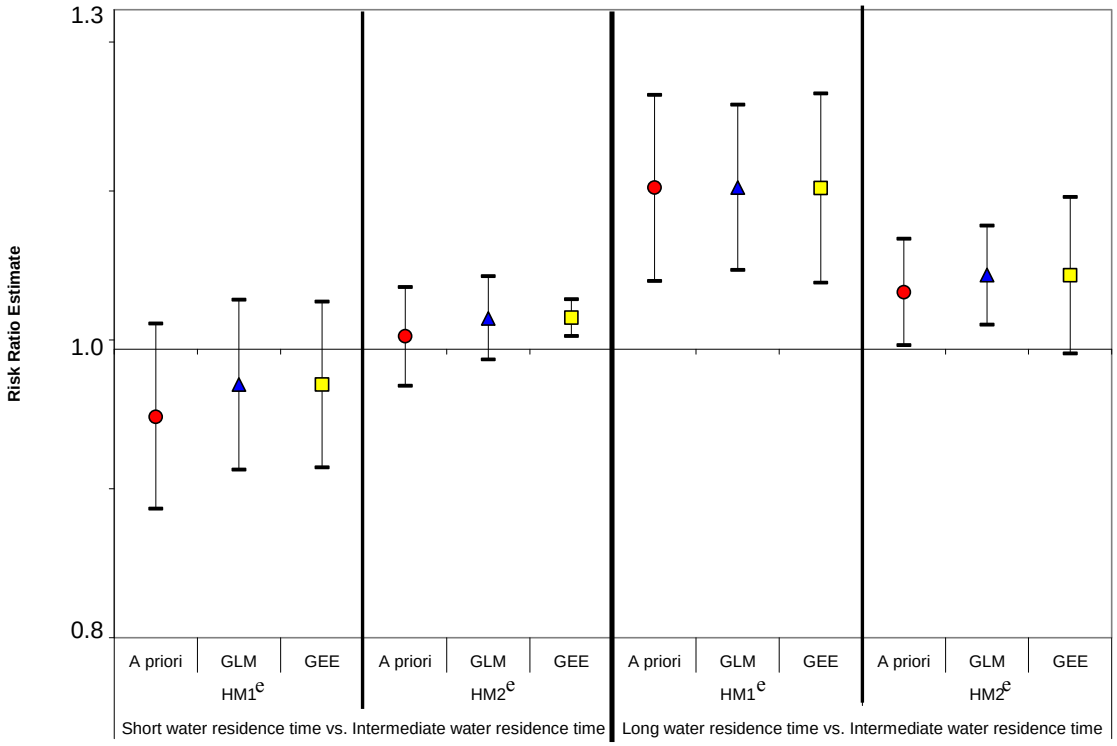
[a] One county considered

Figure A.3: Plot of the number of emergency department visits for gastrointestinal illness predicted from a regression model in which the water residence time estimates were included as a cubic spline with knots at the 10th, 50th, and 90th percentile of all water residence times, Atlanta[a], water residence time estimates derived from hydraulic model 2 only (1999 – 2003)



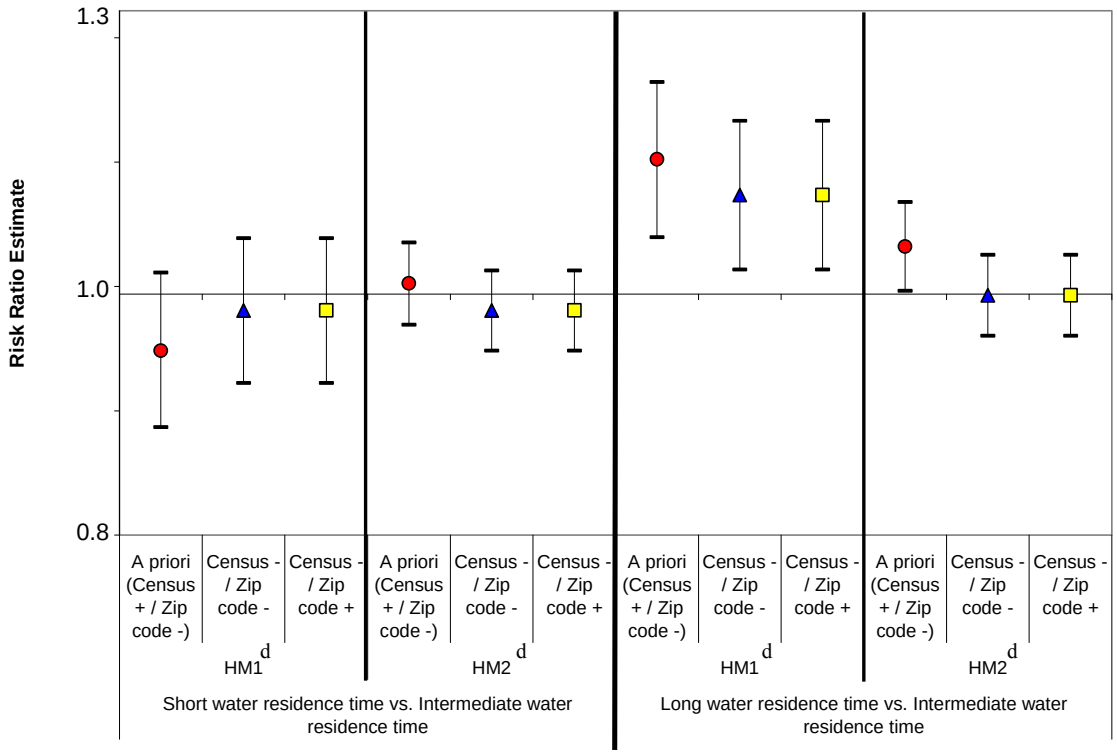
[a] One county considered

Figure A.4: Risk ratio estimates[a] for gastrointestinal illness emergency department visits among people living in zip codes with short[b] drinking water residence times and long[b] drinking water residence times compared to intermediate[b] water residence times, Atlanta[c], 1996 – 2003, assessing impact of using alternate[d] generalized linear models (GLM) and generalized estimating equations (GEE) compared to *a priori*[d] models



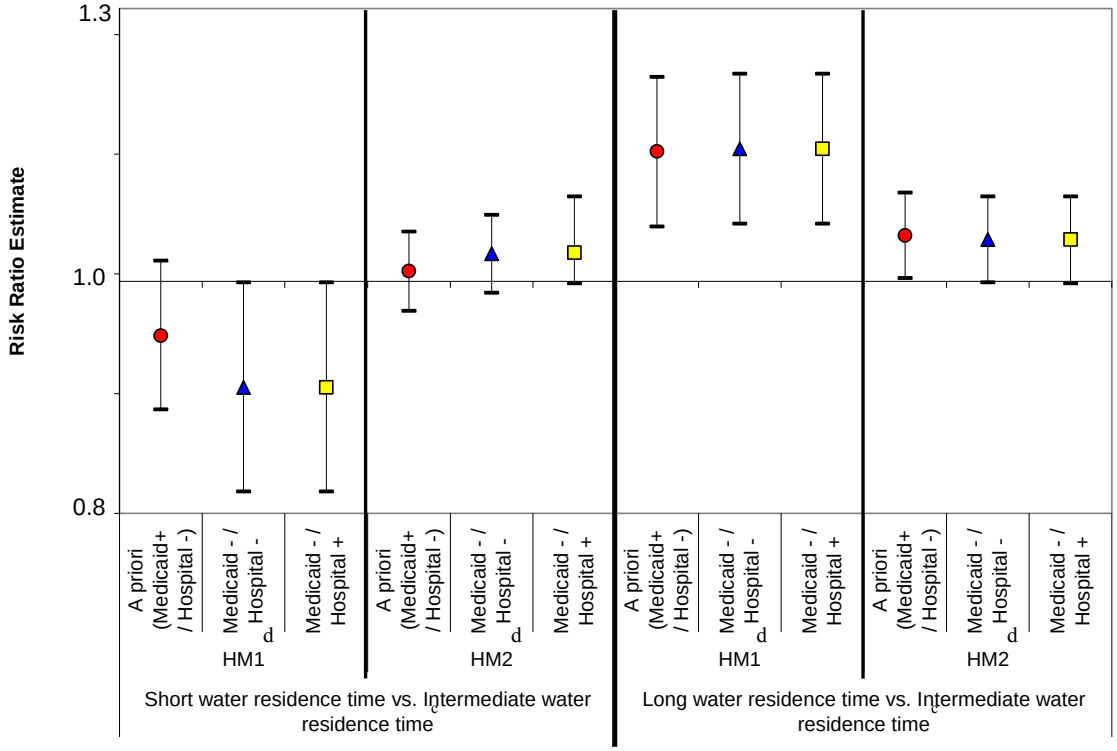
[a] Error bars indicate 95 percent confidence intervals.
[b] Short = ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th – 89th percentile; long = ≥ 90th percentile
[c] One county considered
[d] *A priori* model is a GLM with the following variables: age category, year, hospital, distance from zip code to hospital, zip code median income, zip code percent minority, Medicaid payment status, season, interaction terms for age category with distance to hospital, hospital and distance to hospital, and Medicaid status and age category; alternate GLM and GEE models include all of these variables except the interaction terms for hospital with distance from zip code to hospital.
[e] HM = hydraulic model from which the water residence time estimates were derived; HM1 includes 1996 – 1998; HM2 includes 1999 – 2003

Figure A.5: Risk ratio estimates[a] for gastrointestinal illness emergency department visits among people living in zip codes with short[b] drinking water residence times and long[b] drinking water residence times compared to intermediate[b] water residence times, Atlanta[c], 1996 – 2003, assessing impact of Census variables and missing zip code



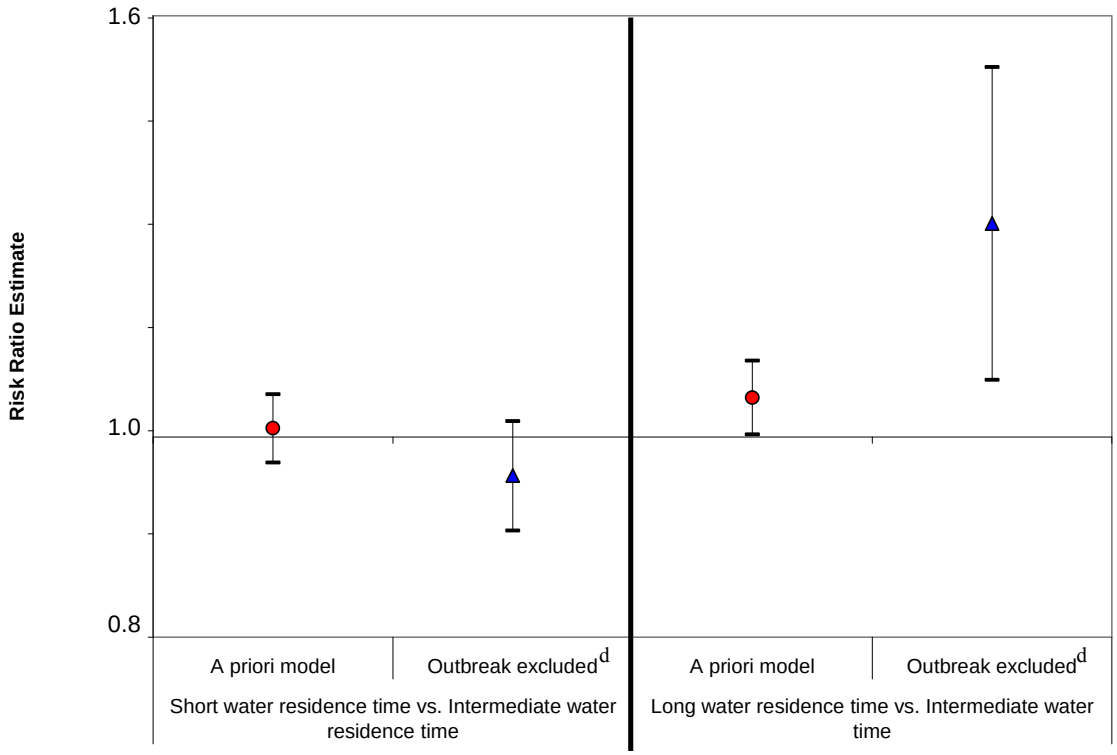
[a] Error bars indicate 95 percent confidence intervals.
[b] Short = ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th – 89th percentile; long = ≥ 90th percentile
[c] One county considered
[d] HM = hydraulic model from which the water residence time estimates were derived; HM1 includes 1996 – 1998; HM2 includes 1999 – 2003

Figure A.6: Risk ratio estimates[a] for gastrointestinal illness emergency department visits among people living in zip codes with short[b] drinking water residence times and long[b] drinking water residence times compared to intermediate[b] water residence times, Atlanta[c], 1996 – 2003, assessing impact of Medicaid payment variable and exclusion of data from one hospital



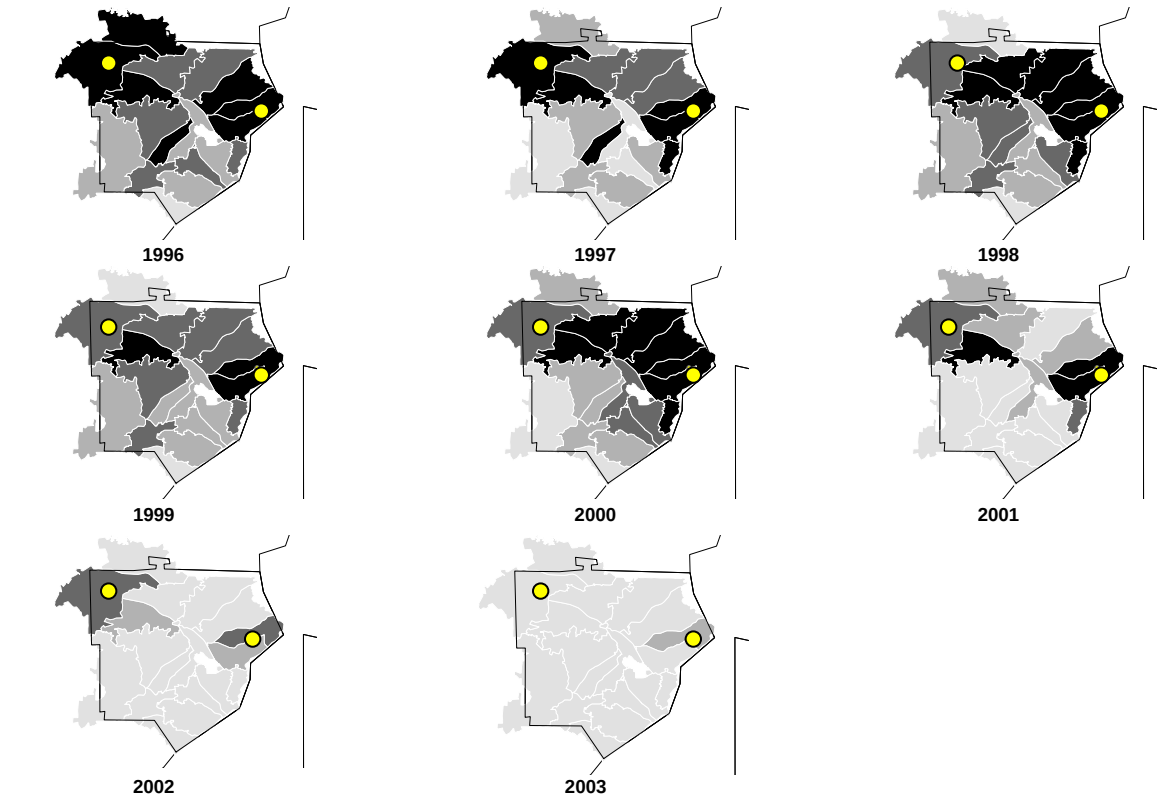
[a] Error bars indicate 95 percent confidence intervals.
[b] Short = ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th – 89th percentile; long = ≥ 90th percentile
[c] One county considered
[d] HM = hydraulic model from which the water residence time estimates were derived; HM1 includes 1996 – 1998, HM2 includes 1999 – 2003

Figure A.7: Risk ratio estimates[a] for gastrointestinal illness emergency department visits among people living in zip codes with short[b] drinking water residence times and long[b] drinking water residence times compared to intermediate[b] water residence times, Atlanta[c], assessing impact of recreational waterborne outbreak (1996 – 1998 only)



[a] Error bars indicate 95 percent confidence intervals.
[b] Short = ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th – 89th percentile; long = ≥ 90th percentile
[c] One county considered
[d] Outbreak and surrounding time period excluded: June 1, 1998 – August 31, 1998

Figure A.8: Chlorine residual levels measured in drinking water distribution system, averaged over time and zip code, Atlanta[a], 1996 – 2003



[a] One county considered

APPENDIX B

ADDITIONAL ANALYSIS OF DRINKING WATER TREATMENT PLANTS AND EMERGENCY DEPARTMENT VISITS FOR GASTROINTESTINAL ILLNESS IN ATLANTA, 1993 – 2004

Objective

The objective of this appendix is to describe additional analyses conducted in the course of assessing the association of the drinking water treatment plant of service and emergency department (ED) visits for gastrointestinal (GI) illness. Several additional analyses were conducted:

- Formal examination of treatment plant attributes
- Inclusion of only zip codes served entirely by one treatment plant
- Assessment of the impact of missing data

Additional Methods, Results, and Conclusions

Treatment Plant Attributes

An objective of the analysis examining the GI illness rates among people served by different drinking water treatment plants was to identify candidate plant attributes that might be at least partially responsible for producing water of better or lesser quality, resulting in decreased or increased rates of GI illness, respectively. We were unable to conduct such an analysis, however, because the majority of the plants had similar attributes and those that did not tended to differ from the other plants in multiple ways.

The attributes we considered in this analysis were source water type, treatment method, and disinfection method. There were two plants served by lakes, five plants served by rivers, and one plant served by water reclamation, creeks, and reservoirs. One of the plants served by a lake was the only plant that used a direct filtration treatment method and the only plant that used ozone, in addition to chlorine, in disinfection. The plant that was served by water reclamation, creeks, and reservoirs was the only plant that used UV disinfection, in addition to chlorination. All of the plants served by rivers used standard treatment methods and disinfected with chlorine.

Future analyses will consider additional treatment plant attributes, such as the coagulants used in treatment and assessment of watershed susceptibility to both point and non-point sources of pollution.

Restricting database to zip codes served 100% by one plant

ED visits were included in the study if the patient resided in a zip code served at least 80 percent by a single drinking water treatment plant. This inclusion criterion introduced exposure misclassification because up to 20 percent of a zip code's residents may have been served by a different treatment plant than the one to which they were assigned. In a sensitivity analysis, we considered only those zip codes for which 100 percent of the zip code was served by a single plant. All ED visits from residents of zip codes served by Plants C, D, and E were excluded from the analysis because none of the zip codes in their service areas were served entirely by a single plant.

It is difficult to directly compare the results of the *a priori* analysis and this sensitivity analysis because of the three treatment plants that were excluded. Entirely removing ED visits from residents of some of the plants' service areas changes the relationship between

the remaining treatment plants because each plant is compared to all other plants. The weighted average of all of the rate ratios must sum to 1.0, and therefore if people served by one plant have a higher rate of ED visits for GI illness, there must be lower rates in the service areas of one or more of the other plants. It is difficult to distinguish what may be a true difference because of the inclusion of only zip codes served completely by a single plant and what may be due to the exclusion of the three treatment plants. The results did not meaningfully differ among people served by Plants G and H (Figure B.1). People served by Plant F had a significantly higher rate of ED visits for GI illness compared to other plants, in both the *a priori* and sensitivity analyses. In the *a priori* analysis, the point estimate for GI illness ED visits among people served by Plant B was slightly negative and was not statistically significant; when only zip codes served completely by a single plant were considered, the point estimate became much more strongly negative and was statistically significant. Conversely, while in the *a priori* analyses a statistically significant negative rate ratio estimate for GI illness ED visits was observed for people served by Plant A, in this sensitivity analysis, the results suggested an increased rate of GI illness among people served by Plant A compared to others. It is difficult to discern whether the differences between the sensitivity analysis and the *a priori* analysis are due to improved exposure classification of zip codes, the inclusion of only 35 percent of the database, or the loss of three of the treatment plants considered in the *a priori* analysis.

Impact of missing data

Missing Census data. In addition to the zip codes excluded from the *a priori* analysis because less than 80 percent of the population was served by a single drinking water treatment plant, an additional 15 zip codes were excluded from the analysis because of

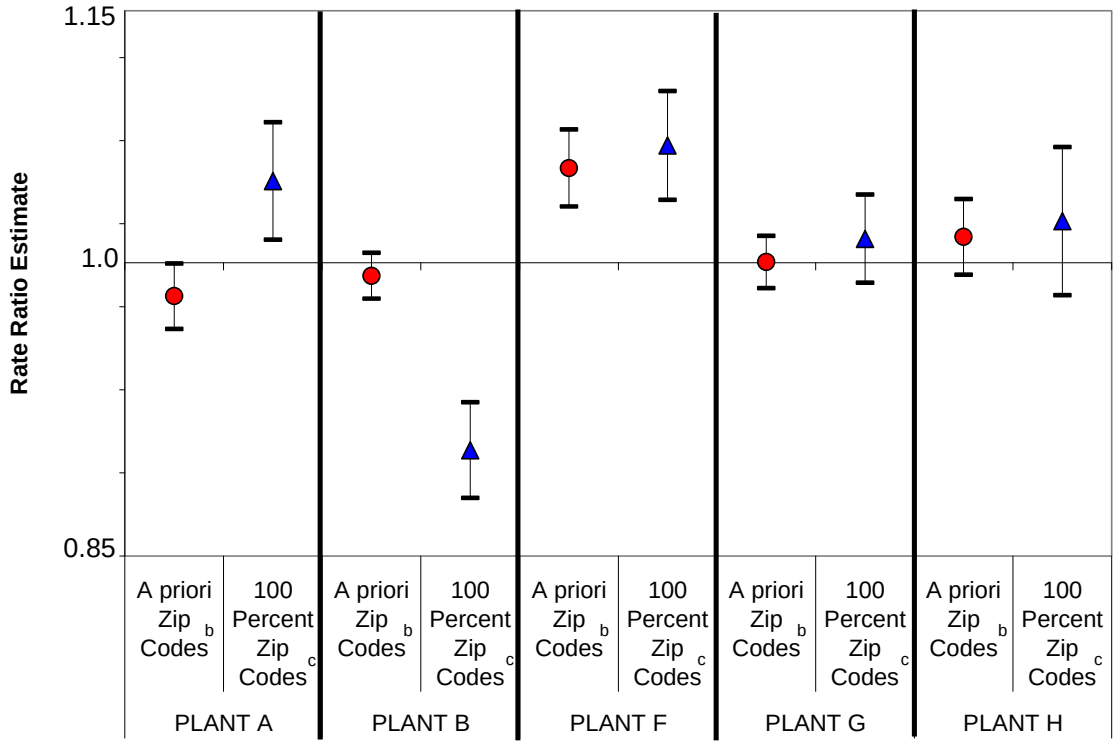
missing Census information. Census information was missing if the zip code boundary or number assignment changed since the beginning of the study and was not traceable to a current zip code. To assess the impact of eliminating these zip codes, we considered results from three models: the *a priori* model, which did not include the excluded zip codes, but did include variables derived from the Census; a model including the excluded zip codes and excluding variables derived from the Census; and a model excluding the excluded zip codes and excluding variables derived from the Census.

The results of this analysis (Table B.2) suggest that it was more important to include the Census variables than to include ED visits from zip codes with missing Census variables. The results from the models excluding the Census variables are quite similar. This suggests that ED visits from the excluded zip codes did not make much of an impact, as the only difference between these models is the inclusion of the excluded zip codes. The results of the models excluding both zip codes and Census data were very different from those observed from the *a priori* model. These results suggest that including the Census variables in the model was important; the only differences between the models that excluded both zip codes and Census variables and the *a priori* models were the Census variables and the results were in some cases meaningfully different.

Missing Medicaid payment information. Data from one of the 28 hospitals contributing ED data had to be excluded from the analysis due to missing information regarding Medicaid payment status (whether the patient paid for the ED visit using Medicaid). Using a similar method as that described previously for excluded zip codes, three models were considered to assess the impact of excluding data from this hospital: the *a priori* model, which excludes the hospital and includes a Medicaid payment

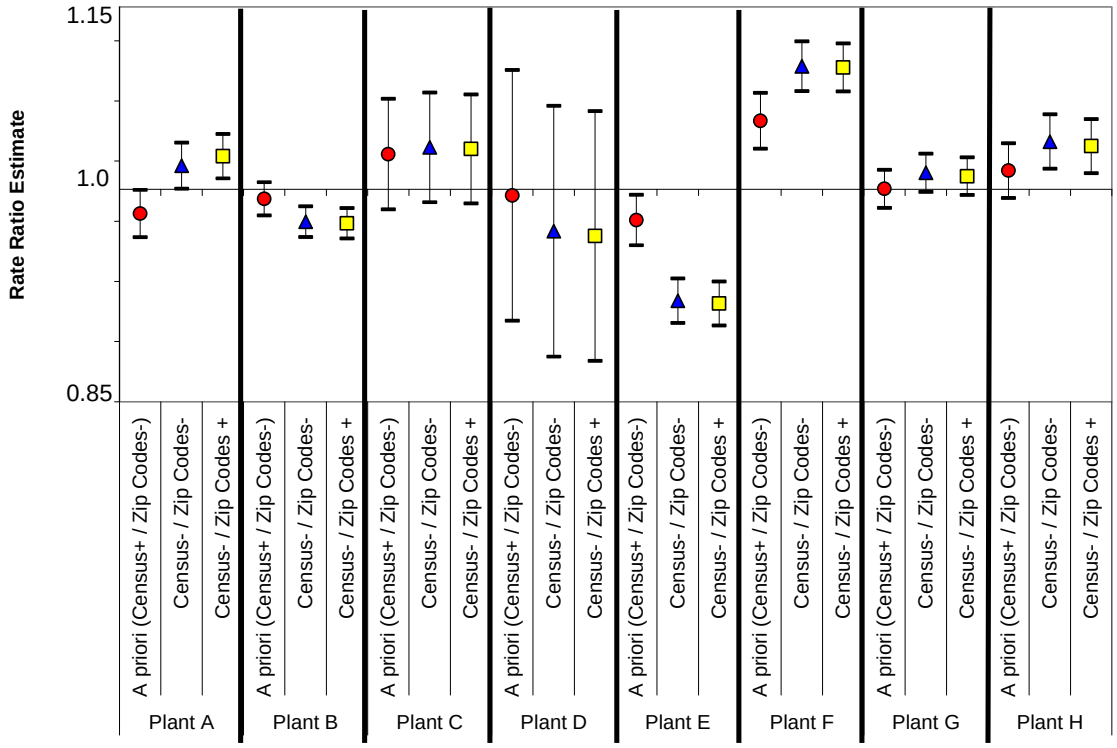
variable; a model that includes the hospital and excludes the Medicaid payment variable; and a model which excludes both the hospital and the Medicaid payment variable. The results of these three models did not differ substantially and the conclusions drawn regarding the relative rates of GI illness among people served by different treatment plants were the same regardless of whether the hospital visits were excluded (Figure B.3).

Figure B.1: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness among people served by one drinking water treatment plant compared to all other treatment plants, five-county metro-Atlanta, 1993 – 2004, assessing impact of zip code treatment plant assignment



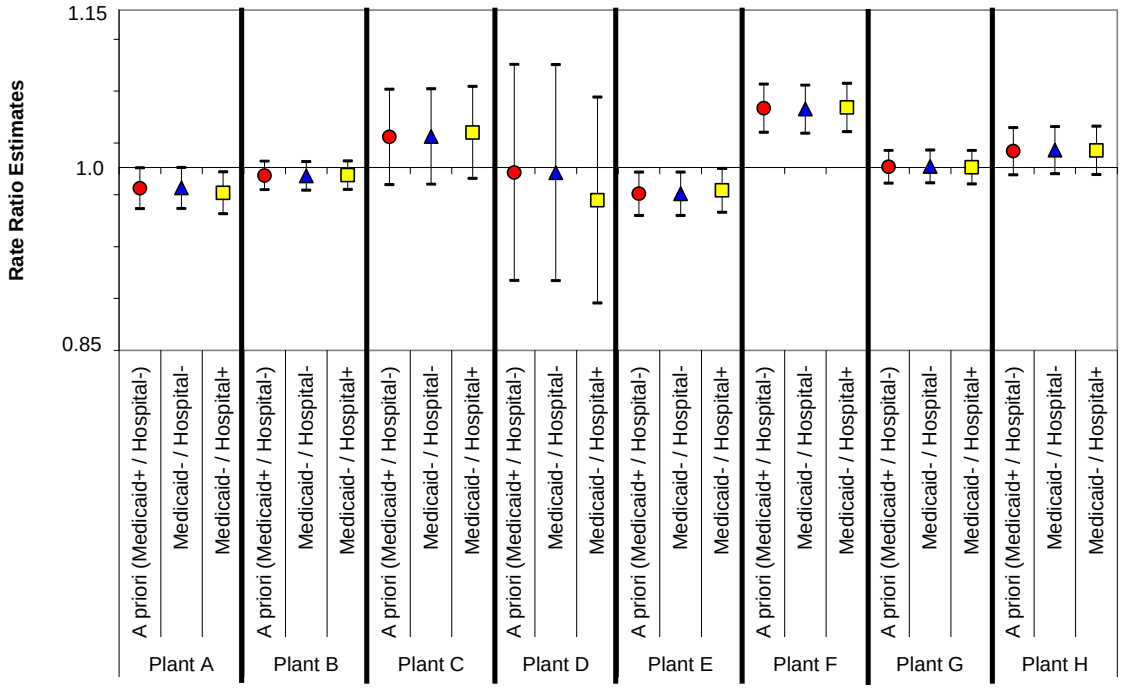
[a] Error bars indicate 95 percent confidence intervals.
[b] Zip codes served at least 80 percent by a single drinking water treatment plant
[c] Zip codes served entirely by a single drinking water treatment plant

Figure B.2: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness among people served by one drinking water treatment plant compared to all other treatment plants, five-county metro-Atlanta, 1993 – 2004, assessing impact of zip codes excluded due to missing Census data



[a] Error bars indicate 95 percent confidence intervals.

Figure B.3: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness among people served by one drinking water treatment plant compared to all other treatment plants, five-county metro-Atlanta, 1993 – 2004, assessing impact of hospital exclusion due to missing Medicaid payment information



[a] Error bars indicate 95 percent confidence intervals.

APPENDIX C

ADDITIONAL ANALYSIS OF DRINKING WATER TURBIDITY AND EMERGENCY DEPARTMENT VISITS FOR GASTROINTESTINAL ILLNESS IN ATLANTA, 1993 – 2004

Objective

The objective of this appendix is to describe additional analyses conducted in the course of assessing the association of drinking water turbidity and emergency department (ED) visits for gastrointestinal (GI) illness. Several additional analyses were conducted:

- Increase and decrease in the frequency of knots in the cubic spline used to control for long-term time trends
- Use of more specific or more sensitive case definitions for GI illness
- Control of rainfall in the analytical models
- Consideration of the turbidity exposure as a three-day moving average, stratified by drinking water treatment plant
- Elimination of visits from the database that occurred during peak rotavirus activity
- Inclusion of only zip codes served entirely by the same treatment plant

Additional Methods and Results

Alternate cubic spline knot structure

The results from the analytical models in which monthly knots were used in the cubic spline, included in the model to control for long-term time trends, were similar to

those observed using the *a priori* knot structure (monthly knots in winter and seasonal knots otherwise) (Figure C.1). The rate ratio point estimates were lower when the monthly knots were used in the cubic spline than the rate ratio estimates from the *a priori* models. The conclusions regarding the association between turbidity and ED visits for GI illness, based on statistical significance ($\alpha = 0.05$), were the same for all turbidity measures except for maximum raw water turbidity. When this exposure was considered in the *a priori* models a statistically significant, positive, rate ratio was observed. While still positive, the rate ratio estimate from the model in which the monthly-knot spline was used was not statistically significant.

The results from the models in which seasonal knots were designated for the cubic spline were quite similar to those observed from the *a priori* model. The rate ratio estimates from the former model were slightly higher than those from the latter model. However, the same conclusions based on statistical significance were reached based on the results from the models including the seasonal-knot spline as the models including the *a priori* spline.

Alternate Case Definitions

The *a priori* case definition included all ED visits for confirmed infectious GI pathogens, along with more general diagnoses consistent with infectious GI illness. Two alternate case definitions were considered: a more specific case definition that included only those cases of GI illness for which a microbial pathogen was confirmed and a more broad definition that included even more symptoms that are compatible with GI illness than were included in our *a priori* definition. The latter definition matched that used by

Schwartz et al. (Schwartz, Levin, and Goldstein 2000) in their time-series analysis of the elderly population in Philadelphia.

None of the rate ratio estimates from models in which the infectious-only case definition were used were statistically significant, but the direction of the point estimate was consistent with that observed from analyses in which the *a priori* case definition was considered for all four turbidity measures (Figure C.2). The results from the models in which the broad case definition was used were consistent with those from the *a priori* model. The rate ratio estimates for GI illness ED visits with a 0.1 NTU increase in average or maximum filtered water turbidity were consistent with the null, and those for a 10 NTU increase in minimum or maximum raw water turbidity were positive and statistically significant.

Controlling for rain

Rainfall was added to the analytical model using a 21-day moving average, incorporating the total precipitation recorded at Hartsfield-Jackson Atlanta International Airport during the days corresponding to the turbidity exposure. This variable was not included in the *a priori* model because turbidity is likely to be in the pathway between rain and GI illness. Turbidity levels were highly correlated with rain levels, with rainfall events being one of the major factors contributing to turbidity. Therefore, controlling for precipitation would yield a loss of statistical power is assessing the true association between turbidity and subsequent GI illness. However, because turbidity is known to be associated with rainfall, and because rainfall could theoretically be independently correlated with other causes of GI illness, we controlled for this variable in a sensitivity analysis.

The point estimate for the average filtered water turbidity exposure when rainfall was included in the model was positive and the point estimate from the *a priori* model was negative, however both were close to the null with confidence intervals including the null value of 1.0 (Figure C.3). The rate ratio estimates from the *a priori* model and the model including rainfall were almost identical when the maximum filtered water turbidity exposure was considered. The addition of rain in the analytical model in which minimum raw water turbidity was considered attenuated the rate ratio estimate compared to the result from the *a priori* model; however, the confidence interval still excluded the null and therefore a similar conclusion was suggested regarding the presence of a modest direct relationship between minimum raw water turbidity and ED visits for GI illness. The rate ratio estimate for the model including rainfall and considering maximum raw water turbidity as the exposure was also attenuated compared to the *a priori* estimate; however, in this case the 95 percent confidence limits included the null value of 1.0. Based on statistical significance testing, this estimate does not support a relationship between maximum raw water turbidity and ED visits for GI illness, while the results from the *a priori* model considering this exposure did.

Turbidity Considered as a Three-Day Moving Average: Results by Treatment Plant

Three-day moving averages were considered for lags ranging from day 0 (the day of the ED visit) to day -20 (20 days prior to the ED visit). All consecutive combinations were considered, from the moving average of turbidity on day 0, day -1, and day -2 through the moving average of turbidity on day -18, day -19, and day -20.

The results of the models considering average filtered water turbidity as a series of three-day moving averages suggested very little association between turbidity and ED

visits for GI illness at most plants (Figure C.4), which was consistent with the results observed for this exposure in the *a priori* analysis. There was a modest association of average filtered water turbidity with GI illness ED visits among people served by Plant B at lags greater than one week. Although there was a positive rate ratio estimate for the association between GI illness ED visits among people served by Plant B and average filtered water turbidity considered in a 21-day distributed lag (the *a priori* lag structure), this estimate was not statistically significant as was observed for some of the lags periods when the three-day moving average was considered. Average filtered water turbidity was associated with ED visits for GI illness among people served by Plant G lagged up to two weeks, which is consistent with the significant positive rate ratio estimate observed from the plant-specific *a priori* results for people served by Plant G. The same pattern of results as described for models in which the average filtered water turbidity exposure was considered was observed when the maximum filtered water turbidity measure was considered, but the point estimates tended to be attenuated for the latter exposure (Figure C.5).

Associations were suggested between minimum raw water turbidity and ED visits for GI illness among people served by most of the treatment plants (Figure C.6), which was consistent with the overall positive association observed for this exposure in the *a priori* analysis. In general these associations were modest and positive. Minimum raw water turbidity was associated with ED visits for GI illness, lagged approximately 9 to 13 days, among people served by Plant A. Modest associations with minimum raw water turbidity were suggested with ED visits for GI illness among people served by Plant B for lags between approximately two and 12 days. Minimum raw water turbidity was

associated with ED visits for GI illness among people served by Plant E lagged up to one week. Modest associations were observed between minimum raw water turbidity and ED visits for GI illness among people served by Plant G at lags of approximately 5 to 11 days. People served by Plant H displayed a decreased rate of GI illness ED visits with increased minimum raw water turbidity for most lag periods. The results from the models for which maximum raw water turbidity were considered as the exposure were similar to those observed when minimum raw water turbidity was considered (Figure C.7).

Eliminating rotavirus peaks

Rotavirus displays a marked seasonality, with large peaks in incidence in the winter. These peaks were identified for each of the treatment plant service areas and these time periods were removed from the database in order to examine whether the associations observed in our *a priori* analyses could be biased due to the influence of the rotavirus peaks, which could not plausibly be explained in their entirety by drinking water exposure.

The results from the models considering both filtered water turbidity measures did not differ between the *a priori* models and the models using the database from which the days corresponding to the rotavirus peaks had been removed (Figure C.8). The rate ratio estimates from the models excluding rotavirus peaks were attenuated for both of the raw water turbidity measures when compared to the *a priori* results, and were no longer statistically significant.

Restricting analyses to zip codes served 100% by one plant

ED visits were included in the study if the patient resided in a zip code served at least 80 percent by a single drinking water treatment plant. This inclusion criterion

introduced exposure misclassification because up to 20 percent of a zip code's residents may have been served by a different treatment plant than the one to which they were assigned. In order to assess what impact this criterion had on our results, we examined models in which only zip codes entirely served by a single plant were considered. This analytical strategy required the elimination of all observations for people served by Plants C, D, and E, as none of the zip codes served by these plants was served completely by a single plant.

Considering only those zip codes that were entirely served by a single plant did not alter the conclusions of the analysis (Figure C.9). Although the point estimate for the association of average filtered water turbidity with ED visits for GI illness was positive in this sensitivity analysis and the estimate was negative in the *a priori* analysis, neither estimate was statistically significant. The rate ratios estimates from the models in which raw water turbidity was considered were slightly attenuated compared to the *a priori* models, but the confidence intervals excluded the null.

Additional Conclusions

The results of the sensitivity analyses in which the association of drinking water turbidity and ED visits for GI illness was considered largely support the findings of the *a priori* analysis presented in Chapter 8. A particular concern of time-series analyses is the appropriate control for long-term temporal trends. Too little control can lead to residual confounding and too much control can obscure a true association. An additional concern in this analysis was that the exposure window was relatively long for a time-series study, three weeks, and therefore placement of the knots in the cubic spline near a three week

window had the potential to completely obscure any association. The results were largely unaffected by changing the placement of the knots in the cubic spline used to control for long-term time trends.

The results of the models in which alternate case definitions were considered were as expected. The rate ratio estimates from models in which the more broad case definition was considered had tighter confidence intervals, due to increased power from the inclusion of additional ED visits counted as being for GI illness, but the results were attenuated compared to the models in which the *a priori* case definition was used. The broad case definition likely included many more cases of GI illness that were not due to drinking water exposure and these cases would be unlikely to be related to turbidity, resulting in non-differential misclassification and bias to the null. The models for which only confirmed infectious cases of illness were counted had low power due to the small number of GI illness ED visits for which a laboratory test was conducted. Therefore, the confidence intervals for the estimates from these models were wide and power was likely insufficient to draw conclusions.

The addition of precipitation to the analytical model led to an attenuation of the rate ratio estimates compared to the *a priori* model when the raw water turbidity exposures were considered. These results suggest that some confounding may have affected the results of the *a priori* models. However, the same conclusion, that a 10 NTU average increase in raw water turbidity over 21 days is associated with a modest increase in ED visits for GI illness, is reached from the results of both analyses. Another possible explanation for the attenuation of results observed in the sensitivity analysis, aside confounding, is that a loss of power led to more random error. Turbidity is in one of the

theoretical pathways linking rainfall to GI illness, and control of rainfall in a model to assess the relationship between turbidity and GI illness may result in less variation in the turbidity exposure and thereby less power to detect an association.

The results from the models in which the turbidity exposure was considered as a series of three-day moving averages support the general conclusions of the *a priori* analysis, but provide additional information regarding the lag periods for which the strongest associations were present. Because the different waterborne pathogens, viruses, bacteria, and protozoa, have different incubation times, identifying the lag times between turbidity and subsequent GI illness may suggest certain organisms are present in the water more often than others. Longer lag periods are more consistent with bacterial or protozoal contamination because these organisms can have longer incubation periods. Shorter lag periods are more suggestive of viral organisms, although some bacterial organisms also have short incubation periods.

Considering turbidity in a series of three-day moving averages also allowed us to see associations that may have been obscured by the *a priori* 21-day distributed lag. In the *a priori* analysis the only statistically significant association observed with the filtered water turbidity exposure was among people served by Plant G. When turbidity was considered in three-day moving averages, a modest association was also suggested among people served by Plant B with lags greater than one week.

The results from the models in which raw water turbidity was considered as a three-day moving average were consistent with those observed from the *a priori* models, with the exception of the inverse association observed for most lag periods among people served by Plant H. This association was unexpected, given that increased turbidity is

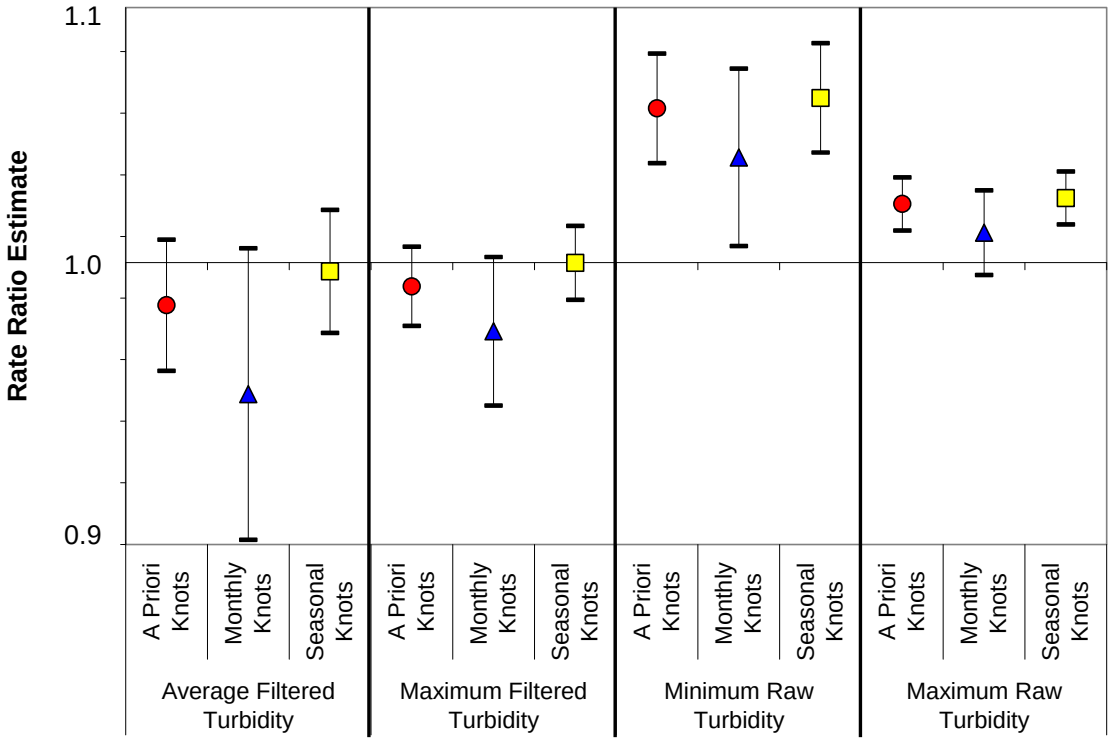
generally assumed to have a direct relationship with pathogen contamination in the water. This result may be explained, however, if the treatment plant responded to the increased raw water turbidity by increasing the treatment efforts used on the water and thereby ended up producing water of higher quality than usual, resulting in fewer pathogens in the water consumed by the customers.

When the time periods corresponding to the peaks in rotavirus were removed, the results from the models in which the raw water turbidity exposure was considered were attenuated and no longer statistically significant. While these results may suggest that the rotavirus peaks were confounding the relationship between turbidity and GI illness ED visits, there are other scenarios that are consistent with the results. The large peaks in rotavirus observed in winter were unlikely to be completely explained by drinking water, but a proportion of these cases could have been contracted via drinking water exposure. Further, viral and bacterial pathogens have better survival in water during cold weather. By eliminating a large portion of the winter days included in the analysis, we may have eliminated the time period during which the association between turbidity and subsequent GI illness was the strongest.

Our *a priori* inclusion criterion for zip codes in the analysis, that they be at least 80 percent served by a single treatment plant, allowed misclassification of the turbidity exposure among up to 20 percent of the ED visits from people living in a given zip code. The impact of this misclassification appears to be minimal, however, given the results from the sensitivity analysis in which only zip codes completely served by a single treatment plant were included. The *a priori* criterion were used because the impact of excluding zip codes not served 100 percent by a single plant was minimal and we felt it

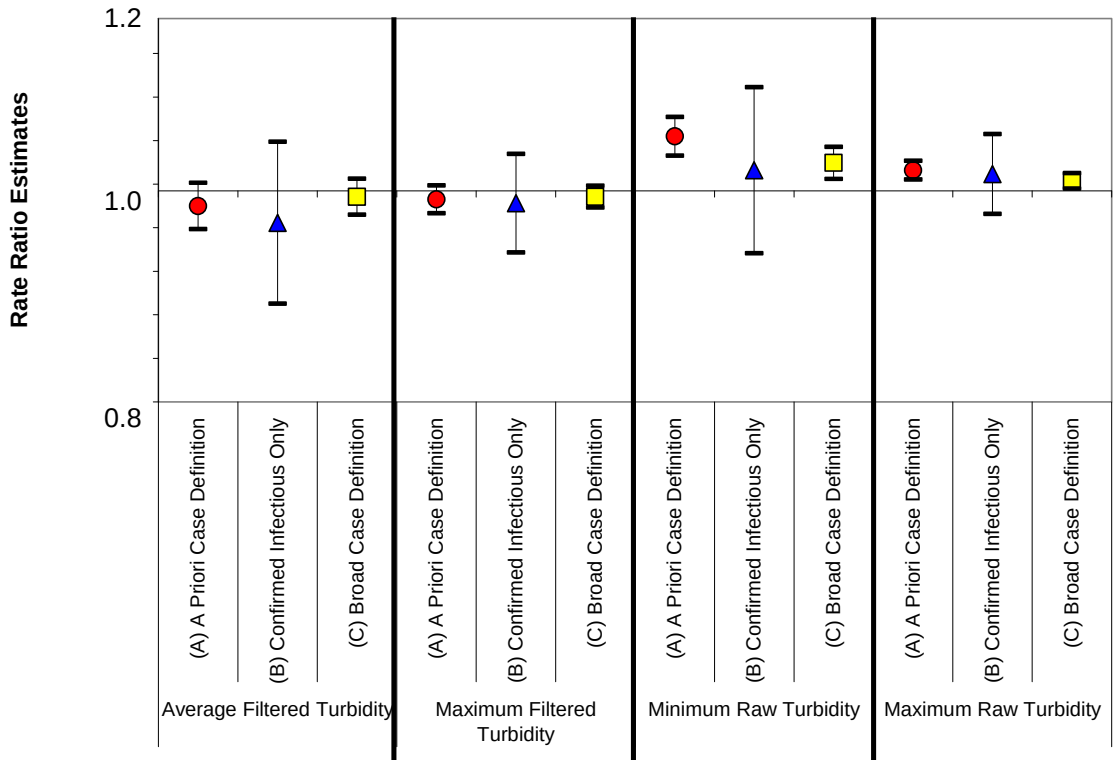
was more important to maximize the generalizability of our results than to use strict inclusion criteria.

Figure C.1: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a 0.1 NTU average increase over 21 days in average and maximum filtered water turbidity and a 10 NTU average increase over 21 days in minimum and maximum raw water turbidity, Atlanta, 1993 – 2004, assessing alternate knot placement[b] in cubic spline



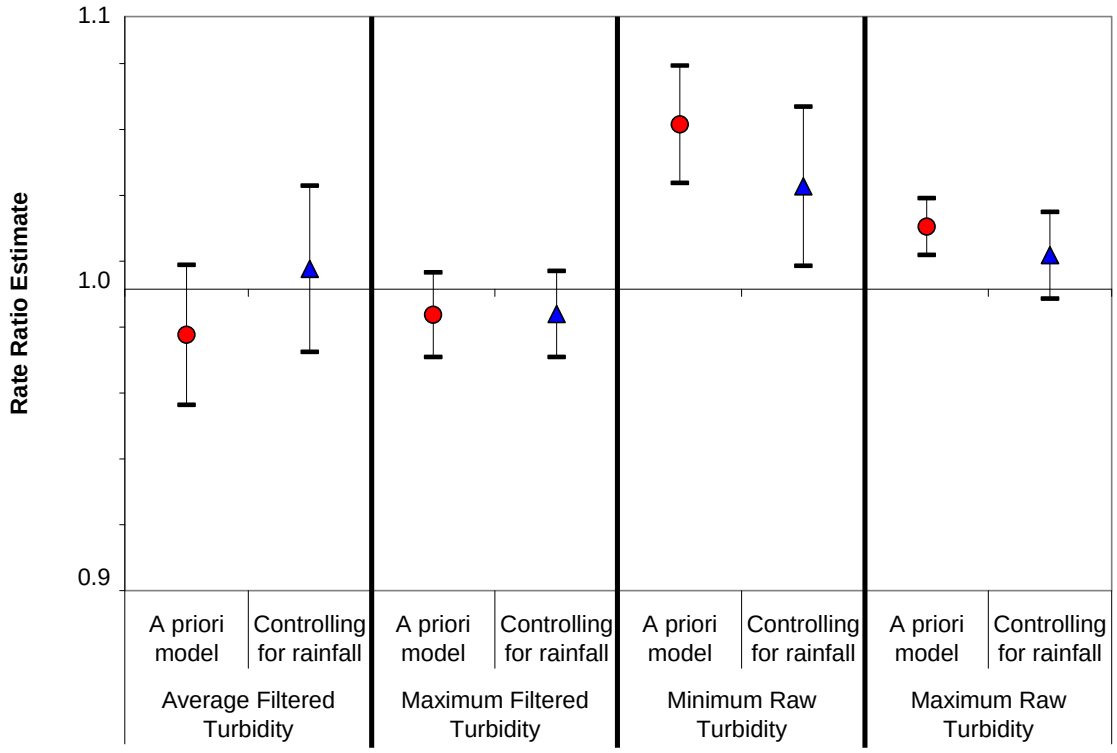
[a] Error bars indicate 95 percent confidence intervals
[b] *A priori* knots are seasonal during spring, summer, and autumn; monthly during winter

Figure C.2: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a 0.1 NTU average increase over 21 days in average and maximum filtered water turbidity and a 10 NTU average increase over 21 days in minimum and maximum raw water turbidity, Atlanta, 1993 – 2004, assessing alternate case definitions for gastrointestinal illness



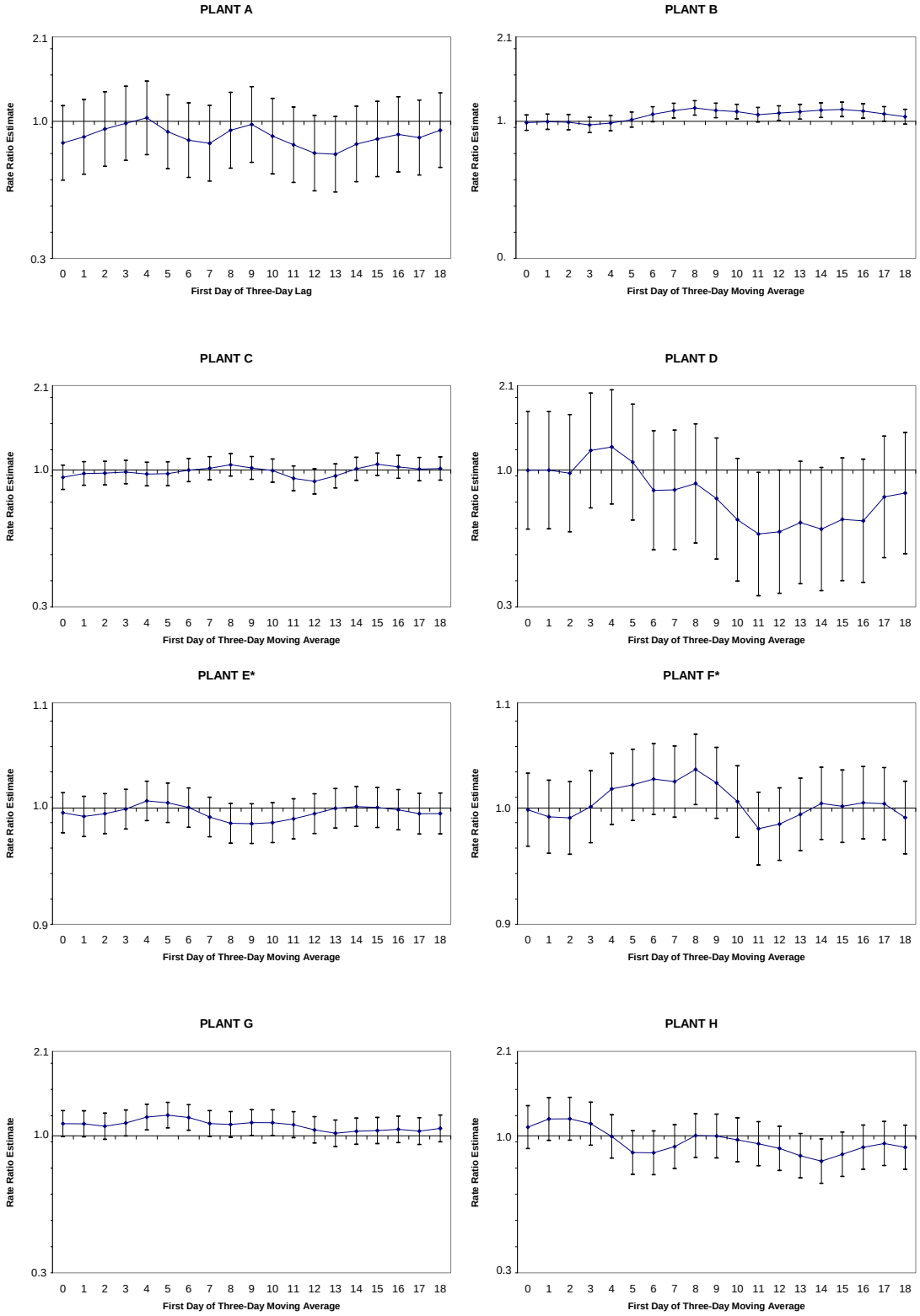
[a] Error bars indicate 95 percent confidence interval
(A) *A priori* Case Definition: ICD-9 codes 001 – 001, 005.0, 005.4, 005.89, 005.9, 006 – 006, 008.0, 008.42 – 008.44, 008.47, 008.49, 008.5, 008.6, 008.8, 009, 558.9, 787.01 – 787.03, 787.91
(B) Confirmed Infectious Only Case Definition: ICD-9 codes 001 – 001, 005.0, 005.4, 005.89, 005.9, 006 – 006, 008.0, 008.42 – 008.44, 008.47, 008.49, 008.5, 008.6, 008.8, 009
(C) Broad Case Definition: ICD-9 codes 001 – 009.9, 276, 787, 789, 558.9

Figure C.3: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a 0.1 NTU average increase over 21 days in average and maximum filtered water turbidity and a 10 NTU average increase over 21 days in minimum and maximum raw water turbidity, Atlanta, 1993 – 2004, assessing addition of rainfall[b] to the analytical model



[a] Error bars indicate 95 percent confidence intervals
[b] Rainfall measured at Hartsfield-Jackson Atlanta International Airport and included in the model as a 21-day moving average

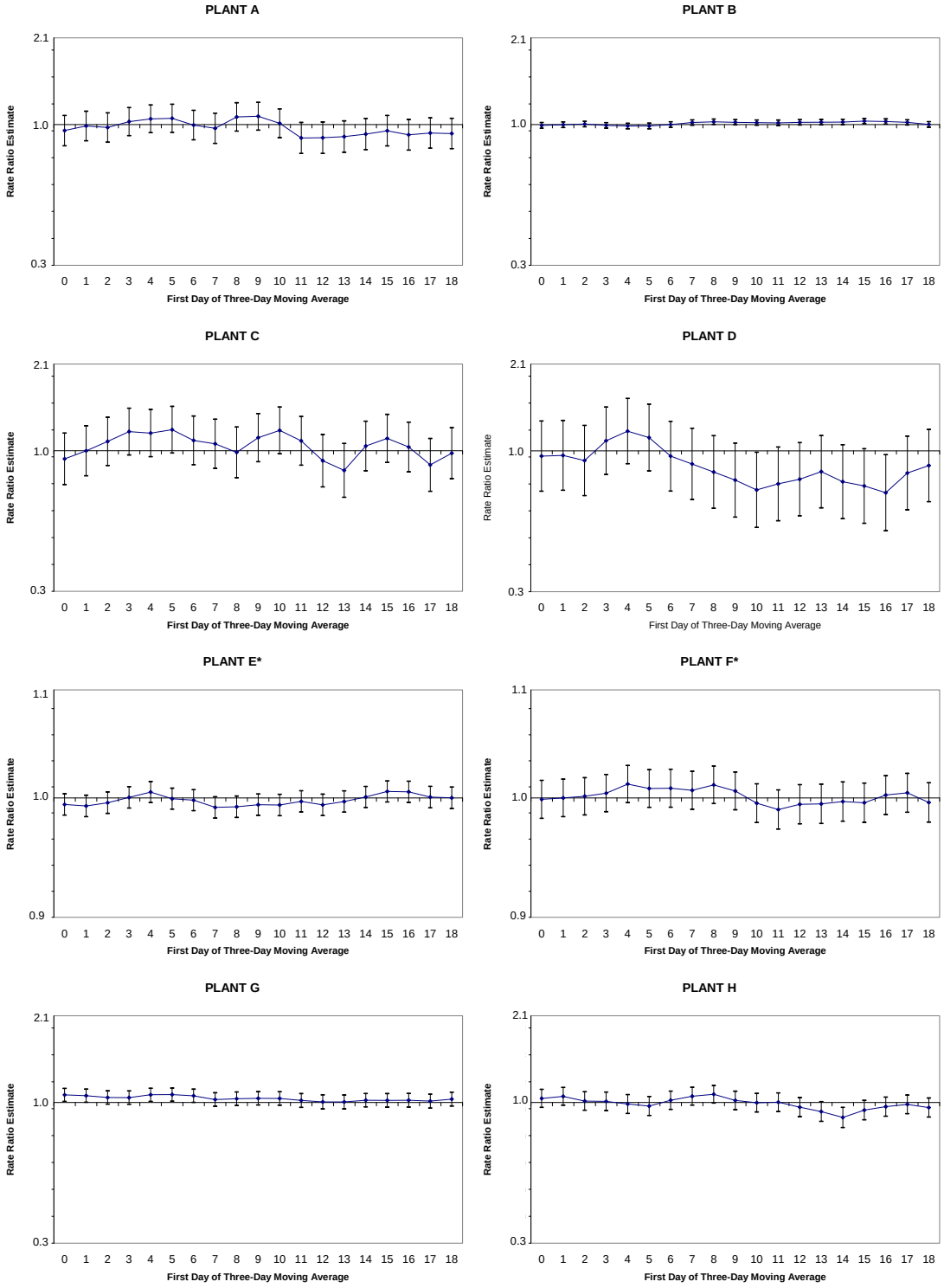
Figure C.4: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a three-day moving average increase of 0.1 NTU in average filtered water turbidity, Atlanta, 1993 – 2004, by plant



[a] Error bars indicate 95 percent confidence intervals

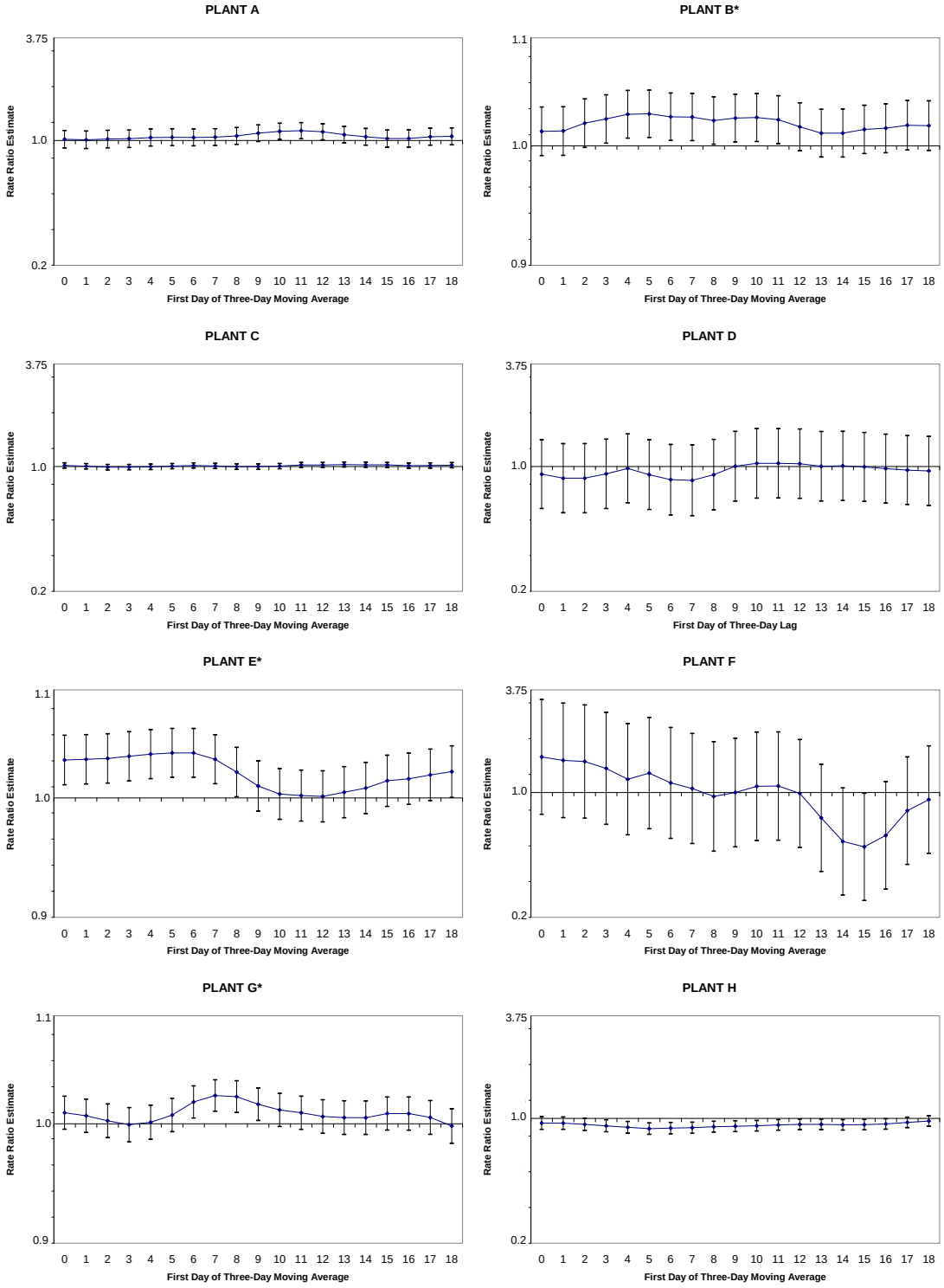
* Note alternate axis.

Figure C.5: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a three-day moving average increase of 0.1 NTU in maximum filtered water turbidity, Atlanta, 1993 – 2004, by plant



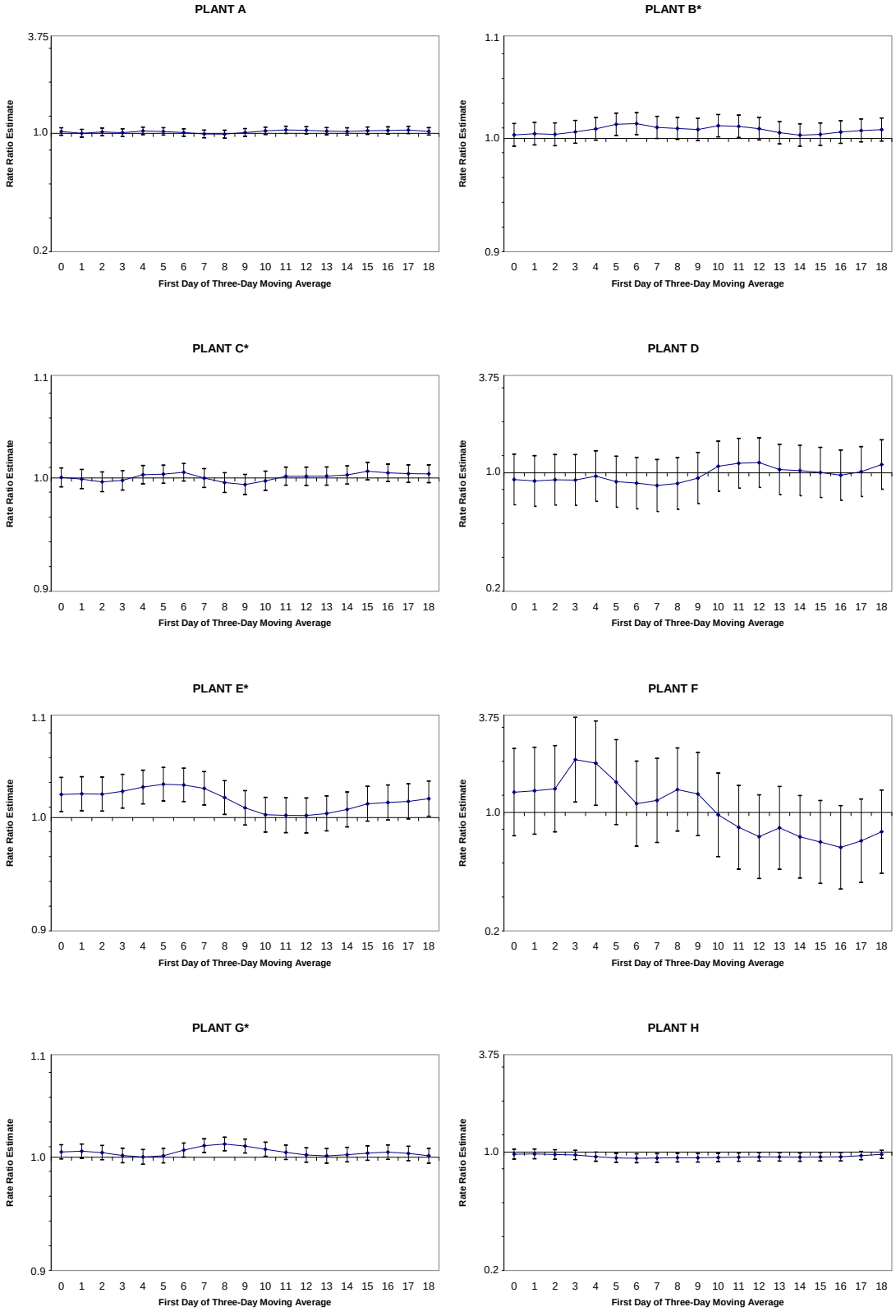
[a] Error bars indicate 95 percent confidence intervals
* Note alternate axis.

Figure C.6: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a three-day moving average increase of 10 NTU in minimum raw water turbidity, Atlanta, 1993 – 2004, by plant



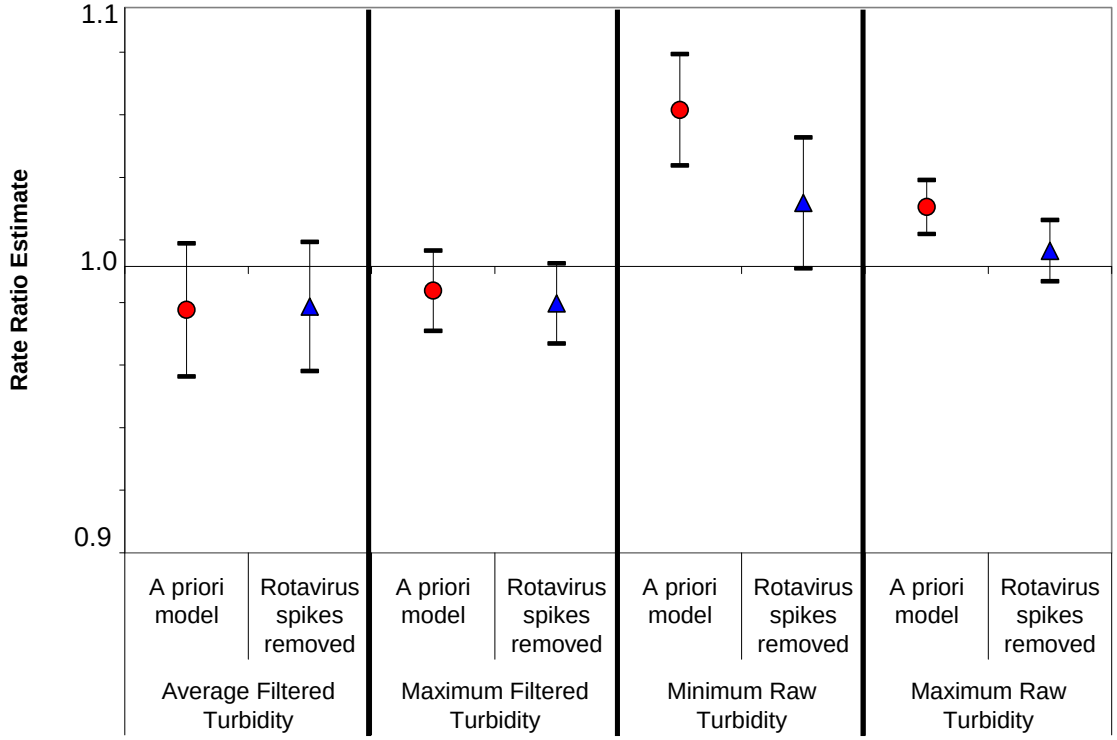
[a] Error bars indicate 95 percent confidence intervals
* Note alternate axis.

Figure C.7: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a three-day moving average increase of 10 NTU in maximum raw water turbidity, Atlanta, 1993 – 2004, by plant



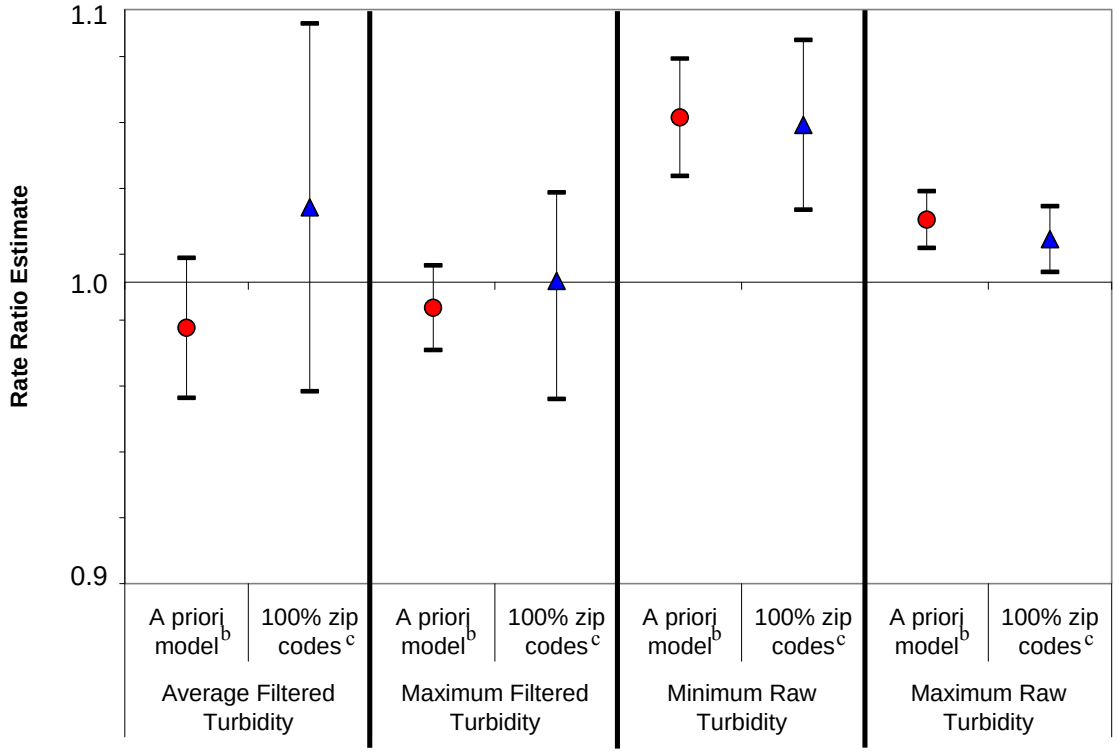
[a] Error bars indicate 95 percent confidence intervals
* Note alternate axis.

Figure C.8: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a 0.1 NTU average increase over 21 days in average and maximum filtered water turbidity and a 10 NTU average increase over 21 days in minimum and maximum raw water turbidity, Atlanta, 1993 – 2004, assessing removal of time periods corresponding to peaks in emergency department visits for rotavirus



[a] Error bars indicate 95 percent confidence intervals

Figure C.9: Rate ratio estimates[a] for emergency department visits for gastrointestinal illness with a 0.1 NTU average increase over 21 days in average and maximum filtered water turbidity and a 10 NTU average increase over 21 days in minimum and maximum raw water turbidity, Atlanta, 1993 – 2004, assessing impact of zip code treatment plant assignment



[a] Error bars indicate 95 percent confidence intervals
[b] *A priori* turbidity exposure assignment based on zip code being at least 80 percent served by a single treatment plant
[c] Database includes only those zip codes served in their entirety by a single treatment plant.