GIARDIASIS AND CRYPOSPORIDIOSIS IN THE URBAN-RURAL SPECTRUM

By

Rebecca L. Ives

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Fisheries and Wildlife

2011

ABSTRACT

GIARDIASIS AND CRYPOSPORIDIOSIS IN THE URBAN-RURAL SPECTRUM

By

Rebecca L. Ives

This study explored occurrence of the environmentally transmissible forms of *Cryptosporidium* spp. and *Giardia* spp. and disease incidence patterns in rural and urban land cover types. The risks of infection and illness caused by *Cryptosporidium* spp. and *Giardia* spp. in Combined Sewer Overflow (CSO) discharge receiving waters were also examined. Significant differences in patterns of giardiasis and cryptosporidiosis were found between urban and rural zip codes. The correlation of population with both diseases suggests that population levels are important factors of transmission in the study area. Although differences in the pattern of disease occurrence between urban and rural areas were found, no evidence of a difference was found in post-hoc testing of the occurrence of the environmentally transmissible forms of the parasites between the surface waters of the Grand River and River Raisin watersheds.

Health risks due to *Cryptosporidium* spp. and *Giardia* spp. in the Grand River at the point of CSO discharge and at recreational sites downstream were assessed. A hypothetical swimming scenario at the discharge point represents moderate daily health risk to a swimmer while multiple day recreation at the discharge point over the recreation season represents a high health risk. Despite CSO discharges, recreational use of the Grand River is predicted to meet the recommended freshwater recreational criteria of 8 illnesses per 1000 swimmers at least 92.4% of the time for both cryptosporidiosis and giardiasis at the three recreational sites examined.

ACKNOWLEDGEMENTS

This study was funded in part by the National Oceanic and Atmospheric Administration (NOAA) grant #89043. The author acknowledges S. Devin McLennan for assistance in plug flow modeling and personnel at the City of Grand Rapids Wastewater Plant for providing access to the MARB facility and sampling assistance.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter One: Introduction

1.1 Significance of *Cryptosporidium* and *Giardia* to Public Health

Cryptosporidium and *Giardia* are parasitic protozoa that infect the gastrointestinal tract of animals and humans. These parasite infections may progress to clinical diseases, called cryptosporidiosis and giardiasis, respectively. *Cryptosporidium* and *Giardia* are generally considered zoonotic pathogens, meaning that the life cycle of the organism can involve a human and a non-human animal host. This capability leads to reservoirs of active infections outside the human population, with the potential for pathogen transmission between human and animal populations. Both parasites are obligate parasites; therefore they cannot reproduce outside of the host. However, both *Giardia* and *Cryptosporidium* produce environmentally resistant stages, called cysts and oocysts, respectively. These resistant stages are shed in the feces of infected individuals and allow the organism to survive in the environment outside of the host's body. This environmental resistance increases the likelihood that a suitable host will encounter the organism and possibly become infected. There are several routes of transmission of *Giardia* and *Cryptosporidium*; the routes of interest in this study involve waterborne transmission. Transmission may involve recreational contact with contaminated waterbodies or transport through water that is then ingested in some manner; such as contaminated drinking water and contaminated water used for irrigation of crops or washing of foods that are eaten raw.

In the human population, *Cryptosporidium* infection occurs in both the immunocompetent and immunocompromized, in both adults and children, and in both the developing and industrialized world (18). The species of *Cryptosporidium* that cause

the majority of disease cases in humans are *C. parvum* and *C. hominis*. Less commonly, disease cases in humans are also caused by *C. meleagridis, C. felis, C. canis, C. suis,* and *C. baileyi,* as well as the cervine genotype of *Cryptosporidium* (18). Among the immunocompetent, infection most commonly presents as gastrointestinal illness with acute diarrhea after an incubation period of 3-14 days (18). Other clinical presentations include persistent diarrhea as well as asymptomatic infection (18). The site of infection is the intestinal tract among the immunocompetent, with diarrheal symptoms usually lasting 6-14 days and the disease is generally self-limiting (18). Among the immunocompromised population, diarrheal symptoms can be much more severe, with daily stool outputs between 2 to 17L (18). Infection in the immunocommpromised population may be persistent or chronic and may develop into infection at extraintestinal sites, such as the biliary tract (18). *Cryptosporidium* infection is associated with increased morbidity and mortality in the immunocompromised population (18).

The infective stage of *Cryptosporidium*, the oocyst, is shed with the feces from an infected individual. The peak intensity of oocyst shedding coincides with the peak intensity of symptoms, during this period an infected individual sheds an average concentration of 10^6 oocyst per gram of feces (4,56). The duration of oocyst shedding varies and may continue after diarrheal symptoms have cleared. In a study of 33 immunocompetent individuals (31 children), Baxby *et al* found that the majority (85%) shed oocysts after diarrhea had ceased (4). Within this study, when measured from the start of diarrheal symptoms, 100% of the infected individuals shed oocysts for 1 week, 82% for 2 weeks, 42% for 3 weeks, and 21% were still shedding oocysts after 4 weeks (4,56)

Like crypotosporidiosis, giardiasis is a common protozoal infection worldwide, infecting both children and adults in both the industrialized and developing world (23). The prevalence of *Giardia* in stool specimens submitted for parasite analysis is between 2% -5% in industrialized countries and higher (20% - 30%) in developing nations (41). Of the five recognized *Giardia* species, only *Giardia lamblia* (syn. *G. duodenalis, G. intestinalis*) has been found to infect humans (1). Like *Cryptosporidium*, *Giardia* infect the intestinal tract, primarily the duodenum and upper jejunum (41). *Giardia* infections may be asymptomatic, acute, or chronic and infections are generally self-limiting in the immunocompetent population (23). Asymptomatic infections, in which infected individuals either have transient diarrhea that is unnoticed or no symptoms, may be the most common (23). The incubation time between initial infection and shedding of *Giardia* cysts is 12-19 days and symptoms generally occur between 6-15 days after infection, but may appear between 1-75 days after infection (5,30,43,56). Clinical symptoms of acute *Giardia*sis are predominantly diarrhea, which is characteristically watery during the initial onset of symptoms (23). Some *Giardia* infections will become chronic with intermittent diarrhea. Weight loss may also occur in chronic infections, up to 10-20% of body weight (23). Complications of *Giardia*sis include nutrient deficiencies, which are generally reversed after the infection is cleared (either spontaneous or with treatment) (23). However, nutrient deficiencies in children with chronic infection may negatively affect growth and development (23).

1.2 Environmental Transmission of *Cryptosporidium* and *Giardia*

The environmentally transmissible forms of *Cryptosporidium* and *Giardia* are the oocyst and the cyst, respectively. Since both *Cryptosporidium* and *Giardia* are obligate parasites, the original source of the oocysts and cysts in the environment is a fecal contamination event. Individuals shed these oocysts and cysts during the patent period of *Cryptosporidium* and *Giardia* infections and, as previously discussed, this shedding period may be of long duration in some cases. High numbers of *Cryptosporidium* oocysts and *Giardia* cysts may be shed by infected individuals. In infected humans, for example, 10^8 *Cryptosporidium* oocysts per gram and/or 10^6 *Giardia* cysts per gram of feces may be shed (4,46,56).

Cryptosporidium oocysts and *Giardia* cysts have several qualities that allow them to be effectively transmitted through the environment. Oocysts and cysts are environmentally resistant, protecting the organism from adverse conditions outside the host. Both *Cryptosporidium* oocysts and *Giardia* cysts are resistant to chlorine, one of the most commonly used disinfectants in water treatment, although *Giardia* is less resistant than *Cryptosporidium*. Oocysts and cysts are immediately infectious, no maturation is required before they can initiate infection in a new host. Additionally, low doses of oocysts and cysts are capable of initiating infection. In humans, *Giardia lamblia* infections may result from the ingestion of 10 cysts and doses of 10 to 25 cysts in feeding studies caused infection in 36.4% of 22 volunteers (43). Feeding studies using several *C. parvum* isolates have demonstrated that there is a range of infectivity among isolates (38). However, the ID_{50} (the dose at which 50% of a population is infected) of some isolates is as low as 9 ingested oocysts and infection may potentially be initiated

by a single oocyst (38). Also of importance, the waterborne transmission of *Cryptosporidium* and *Giardia* from a single contamination event can potentially infect large numbers of people (23).

1.3 Sources of Environmental Contamination

Various species of *Cryptosporidium* and *Giardia* infect humans, domestic livestock, pets, and wildlife. This leads to multiple potential sources of contamination by these parasites in the environment (18,49), although not all species may be medically important in terms of human disease. Human sources of *Cryptosporidium* and *Giardia* contamination may include improperly treated sewage, discharges of untreated sewage via sanitary sewer overflows or combined sewer overflows (CSOs), land application of biosolids and septage, and leaking sewer or septic systems. Animal sources may also contribute to environmental contamination in a variety of ways. Runoff from domestic livestock operations or fields fertilized using animal manure may carry *Cryptosporidium* and *Giardia* into waterbodies. Defecation by pet animals in the environment and defecation of wildlife also may provide sources of these parasites. However, although genotyping studies of *Cryptosporidium* and *Giardia* have demonstrated that both host-adapted and zoonotic strains are present in wildlife (1,22), most strains are host-adapted (22). Feng *et al* conclude that while wildlife can contribute to *Cryptosporidium* contamination in the water, wildlife sources may not have major public health significance (22). Studies on *Cryptosporidium* and *Giardia* concentrations have also found higher concentrations in domestic animals than in wildlife (17,26). Fecal contamination of waterways from human sources and domestic animals, therefore, may be of greater concern than contamination by wildlife.

Cattle, sheep, pigs, and horses are susceptible to *Cryptosporidium* and *Giardia* infections. Cattle are of particular interest due to high prevalence of *Cryptosporidium* among some herds and high intensity shedding patterns. In studies examining cattle, the percentage of animals infected with *Cryptosporidium* range from 0.17% to 69.2% (3,58). In studies examining cattle, the percentage of animals infected with *Giardia* range from 11% to 73% (39,40,58). In a shedding study of calves, individual animals have been documented as shedding up to 26 million *Cryptosporidium* oocysts per gram of feces and 4.2 million *Giardia* cysts per gram of feces (58). No molecular typing of *Cryptosporidium* or *Giardia* was performed in the shedding study, although *Cryptosporidium* oocysts were morphologically similar to *C. parvum*, and therefore assumed to be pathogenic to humans. In the study of shedding patterns in calves, *Cryptosporidium* shedding was found to peak earlier and with higher maximum individual concentrations per gram of feces than *Giardia* shedding. However, once the calves reached four weeks of age, *Giardia* shedding intensity was four orders of magnitude greater than the *Cryptosporidium* shedding intensity. The earliest date of shedding was four days for both *Giardia* and *Cryptosporidium*. All sampled animals had at least one *Cryptosporidium* shedding episode and all sampled animals had at least one *Giardia* shedding episode (58).

The parasites produce infections that lead to disease states in livestock, however, asymptomatic animals have also been found to shed *Cryptosporidium* and *Giardia*. In a 1997 study of asymptomatic sheep, pigs, cattle, and horses by Olsen *et al*, differences in the population age and shedding were found between *Cryptosporidium* and *Giardia* (40). In cattle, detection of *Cryptosporidium* tended to be more frequent in young animals (<6

months of age) than older animals, whereas in sheep, pigs and horses, detection tended to be more frequent in older animals(>6 months of age) than younger animals. With *Giardia*, detection in both cattle and sheep tended to be more frequent in young animals than older animals. Detection of *Giardia* in pigs and horses tended to be more frequent in older animals than younger animals. Olsen *et al* found that the difference in prevalence between younger and older animals tended to be significant (P=0.05) except in the case of sheep, in which no significant difference was found between older and younger animals (40). Another study of asymptomatic and symptomatic cattle found no statistical difference in prevalence of *Cryptosporidium* infection among diarrheic (63.3% prevalence) and non-diarrheic (69.2% prevalence) animals (58).

The shedding of parasites by asymptomatic animals presents a challenge to reduction of environmental loading of the parasites by management practices on the farm. If only symptomatic animals shed the parasites, animals presenting with symptoms could conceivably be isolated from the rest of the herd and manure from these isolated animals more heavily managed to reduce environmental loading. In a seven-state study, estimates of environmental loading of *Cryptosporidium* by feedlot steers ranged from 2.8×10^4 to 1.4×10^5 oocysts/steer per day for the arithmetic mean estimates and 9.1×10^3 to 3.7×10^4 oocysts/steer per day for the geometric mean estimates (3).

In a 2005 review of zoonotic transmission pathways by Hunter *et al*, studies indicated that *C. hominis* is transmitted only between humans but domestic livestock, especially cattle, are reservoirs of *C. parvum* (28). These studies also indicated that zoonotic disease transmission results from direct contact with cattle and indirectly through contamination of drinking water (28). In the case of giardiasis, however,

evidence from studies does not support zoonotic transmission as a major transmission route (28).

In regards to the parasites *Cryptosporidium* and *Giardia*, one of the challenges to management of animal manures for protection of human and animal health is the stability of these parasites. In a study of *C. parvum* inoculated into various animal manures/slurries, estimates of the amount of time for a 1 log decrease in infectivity (as measured by DAPI and PI vital staining) ranged from three months to almost 1 year (29). Even these long time periods may underestimate of the time required for inactivation, as the vital staining methods are not as sensitive to low numbers of infective (oo)cysts as animal infectivity or, in the case of *Cryptosporidium*, cell culture infectivity. The long time periods required for inactivation through storage present logistical challenges to manure management, as storage space may be limited. This may lead to spreading of manures that still contain infectious (oo)cysts onto agricultural lands in order to clear storage space for more recently deposited fecal material.

1.4 *Cryptosporidiosis* and *Giardiasis:* Reporting of Human Disease Cases

Human cases of cryptosporidiosis and giardiasis are currently on the list of nationally notifiable diseases compiled by the United States Center for Disease Control (CDC) and public health officials (10). According to the CDC,

> "A notifiable disease is one for which regular, frequent, and timely information on individual cases is considered necessary for the prevention and control of the disease." (16)

National reporting for cryptosporidiosis began in 1995 (2,972 cases, 27 states) (15). National reporting for giardiasis officially began in 2002 (21,206 cases, 46 states) (11),

however, beginning in 1992 there are earlier cases reported from a smaller subset of states which have been summarized by the CDC (12). Disease reporting at the state and local levels is mandatory, however state reporting to the CDC is voluntary, therefore the data used in summary reports generated by the CDC may be incomplete (10). Table 1.1 summarizes the reported cases of cryptosporidiosis and giardiasis from onset of national reporting to 2008, the most recently published surveillance data. Cryptosporidiosis in the United States has been trending upwards and recent data (2005-2008) show the highest number of reported cases. Cryptosporidiosis in Michigan has also been increasing. However, the cryptosporidiosis trend in Michigan has been fairly stable over the period from 2001-2008. Giardiasis in the United States and Michigan peaked in 1996 and 1995, respectively. In the last five years of published data, the giardiasis trends in both the United States and Michigan have been fairly stable.

The reported data for both cryptosporidiosis and giardiasis are likely to be underestimates of the disease burden in the United States. For example, extrapolating from the estimates of salmonellosis reporting (1% - 5%), in 2002 the true disease burden in the Unites States could have been between 60,320 - 301,600 cases of cryptosporidiosis and 424,120 -2,120,600 cases of giardiasis (reported cases: 3,016 and 21,206, respectively) (11).

Nationally reported surveillance data may contain both confirmed cases and probable cases of giardiasis and cryptosporidiosis, depending on how a jurisdiction (i.e. State) reports the data. The CDC has used four definitions for a laboratory confirmed case of cryptosporidiosis from 1995 through 2011 as methods have developed. The criteria used for these definitions are shown in Table 1.2. A case must meet one of the required

criteria (specific to definition publication interval) to be defined as a confirmed case of cryptosporidiosis. The current (2011) CDC definition of a laboratory confirmed case of cryptosporidiosis is "the detection of *Cryptosporidium* organisms or DNA in stool, intestinal fluid, tissue samples, biopsy specimens, or other biological sample" which includes direct fluorescent antibody tests but excludes other immunodiagnostic tests (6). The CDC has defined a laboratory confirmed case of giardiasis as "the detection of *Giardia* organisms, antigen, or DNA in stool, intestinal fluid, tissue samples, biopsy specimens or other biological sample"(7). A probable case is a case that is epidemiologically linked with a confirmed case with symptoms that match with the clinical description of disease (9). The epidemiological linkage of a probable case with a confirmed case is necessary, due to the similarity of cryptosporidiosis and giardiasis symptoms with many other gastrointestinal diseases.

Table 1.1 Reported Cases of Cryptosporidiosis and Giardiasis in the United States and Michigan

Disease	Cryptosporidosis				Giardiasis			
Year	United States		Michigan	Reference	United States		Michigan	Reference
	Reported cases	Number of reporting states	Reported cases		Reported cases	Number of reporting states	Reported cases	
1992	*	*	*		12,793	22**	1,333	(12)
1993	*	*	*		19,565	36**	1,199	(12)
1994	*	*	*		26,346	42**	1,370	(12)
1995	2,972	27	*	(15)	25,571	43**	1,435	(12)
1996	2,426	42	*	(15)	27,778	45**	1,290	(12)
1997	2,566	40	46	(14)	25,389	43**	1,212	(12)
1998	3,793	42	39	(15)	24,204	42	1,172	(11)
1999	2,769	46	52	(11)	23,245	43	1,166	(11)
2000	3,128	46	97	(11)	21,772	44	1,135	(11)
2001	3,785	49	187	(11)	19,659	42	1,003	(11)
2002	3,016	50	135	(11)	21,206	46	923	(11)
2003	3,506	49	152	(9)	19,718	45	781	(9)
2004	3,911	49	157	(9)	20,655	45	718	(9)
2005	8,269	50	112	(9)	19,789	45	783	(9)
2006	6,479	50	145	(8)	18,958	45	715	(8)
2007	11,657	50	215	(8)	19,421	45	620	(8)
2008	10,500	50	284	(8)	18,913	45	611	(8)

* No Data Available
** Reports do not distinguish between zero cases and non-reporting of data

Table 1.2 CDC Criteria for Cryptosporidiosis Case Confirmation

Required Criteria (X)	Definition Publication Year			
	1995	1998	2009	2011
Microscopic demonstration of *Cryptosporidium oocysts* in stool	X	X	X	X
Microscopic demonstration of *Cryptosporidium* in intestinal fluid or small-bowel biopsy specimens	X	X	X	X
Demonstration of *Cryptosporidium* antigen in stool by immunodiagnostic test	X	X	X	
Demonstration of *Cryptosporidium* antigen in intestinal fluid by immunodiagnostic test			X	
Demonstration of *Cryptosporidium* nucleic acid by PCR in stool, intestinal fluid, or tissue samples or biopsy specimens		X	X	X
Demonstration of *Cryptosporidium* reproductive stages in tissue preparations		X		X

The Michigan Public Health Code (333.5111 of the Michigan Compiled Laws, Public Health Code, Act No. 368 of the Public Acts of 1978) requires that certain conditions and agents of infection, including *Giardia lamblia* and *Cryptosporidium*, be reported to the local health authorities. All physicians and laboratories are required to

report cases of these infections or conditions. Other health professionals are authorized to report, but are not required to do so. Reports of individual infection are required to contain the following information: the patient's full name, residential address, telephone number, date of birth (or age), and gender. Reports must also include the specific laboratory test used for diagnosis, date performed, test results, the name and address of the reporting clinical laboratory, and the name, address, and telephone number of the ordering person. This surveillance data can be used for identifying outbreaks, tracking incidence of disease over time, and other epidemiological uses to protect the public health.

1.5 Epidemiological Use of Geographic Information Systems (GIS)

Spatial statistical methods and GIS have been used to investigate *Cryptosporidium* and *Giardia epidemiology* in a few studies. Odoi *et al* used GIS to explore the role of livestock density / manure application in giardiasis rates in Canada but didn't find strong evidence that these were important epidemiological factors (37). Pollock *et al* examined the association of cryptosporidiosis with a number of variables including human population density, livestock density (cattle, sheep, farmed deer, goat, and pig), ratio of farms to human inhabitants and two orthogonal axes (West-East and South-North) in Scotland (42). Pollock *et al* found increased rates of *C. parvum* infection in areas with lower human population densities, in areas with higher ratios of farms to humans, and higher ratios of private water supplies to human populations, all of which can be considered rural indicators (42). In a GIS case study in England and Wales, Lake *et al* found higher rates of cryptosporidiosis in rural areas (defined by housing density), areas

with more agricultural manure application, and areas with poorly treated water supplies (31). These studies demonstrate the usefulness of GIS in epidemiological investigations.

1.6 Research Objectives

Much of the research on *Cryptosporidium* and *Giardia* infection to date has been focused on individuals or groups of individuals in order to determine the parasitic life cycles, the symptoms of the diseases they produce, host specificity, resistance to disinfectants, and medical interventions. Environmental research on *Cryptosporidium* and *Giardia* tends to focus on occurrence in water and disinfection by water treatment. Few studies are available regarding land cover and occurrence of *Cryptosporidium* spp. and *Giardia* spp. or risk of infection/illness by *Cryptosporidium* spp. and *Giardia* spp in recreational waters.

This work focused on parasite occurrence patterns in rural and urban land cover types, as well as recreational health risks. The specific objectives were:

1. Establish a prototype GIS database for land cover attributes, demographic attributes, incidence of cryptosporidiosis, and incidence of giardiasis in Hillsdale, Lenawee, Ottawa, and Kent Counties.
2. Examine patterns between urban, rural/agricultural land cover and incidence of giardiasis and cryptosporidiosis
3. Assess patterns between urban, rural/agricultural land cover and occurrence of *Cryptosporidium* spp. and *Giardia* spp. in surface waters.
4. Assess the health risks due to *Cryptosporidium* spp. and *Giardia* spp. in combined sewer overflow (CSO) receiving waters

Chapter two describes the creation of the prototype GIS database (research objective one), environmental sampling, sample processing procedures, and data analysis methods. In chapter three, research objective two is addressed, exploring the patterns between disease and land cover type. In chapter three, the null hypothesis that disease levels are the same in urban and rural areas was tested. Chapter four is a descriptive analysis of the *Cryptosporidium* spp. and *Giardia* spp occurrence in two watershed segments with differences in urban and rural/agricultural land cover. In chapter four, the null hypothesis that parasite occurrence values was the same in the Grand River watershed and in the River Raisin watershed was tested. Chapter five addresses objective four in a health risk assessment of *Cryptosporidium* spp. and *Giardia* spp. from exposure via recreation in combined sewer overflow (CSO) receiving waters.

Chapter Two: Material and Methods

2.1 Creation of ArcGIS Database

To investigate patterns between land cover type and the incidence of giardiasis and cryptosporidiosis, a database was constructed in ArcGIS (ArcGIS 9.2, ESRI). The database was populated with publicly available information from the Michigan Geographic Data Library including census tracts from the counties of interest (Hillsdale, Lenawee, Ottawa, and Kent) and the 2001 land cover raster image of Michigan's Lower Peninsula (30m x 30m cell size). The database also included information obtained from the United States Census Bureau on the five digit ZIP code tabulation areas (ZCTAs) for the year 2000 (51). Information on giardiasis and cryptosporidiosis cases was obtained from the Michigan Department of Community Health. Case information over the time span of January 2000 to December 2008 was summed by ZIP code and included in the database. Data on additional attributes for each ZIP code in the study area were obtained from the 2000 United States Census. These attributes were:

1. Population
2. Percentage of individuals below poverty level
3. Median household income
4. Median age
5. Mean travel time to work
6. Percentage of Caucasian/white individuals

A case information table with cases of cryptosporidiosis and cases of giardiasis summed by ZIP code was created based on information from the Michigan Department

of Community Health (34). Using the "Union" tool in ArcGIS, the census tract information for Kent, Ottawa, Lenawee, and Hillsdale counties was combined into an output file. Using the "Extract by mask" tool with the census tract output file and the raster file for the 2001 land cover of Michigan's Lower Peninsula, an output file containing the land cover of Kent, Ottawa, Lenawee, and Hillsdale counties was created.

Land cover was reclassified to allow hypothesis testing of the differences between urban and rural land cover. The land cover of the counties was reclassified following the reclassification scheme shown in Table 2.1 to produce a file of Kent, Ottawa, Lenawee, and Hillsdale counties with land cover classified as "rural", "urban", and "other". The decision tree with the criteria used to reclassify all land cover categories other than the "Parks / Golf courses" is shown in Figure 2.1. The "Parks / Golf courses" is an ambiguous category as the parks could range from city parks well connected to sewer infrastructure and drinking water distribution networks located in high population density areas to State parks with limited sewer and drinking water access. In this case, "Parks / Golf courses" was reclassified as "urban" since this category is associated with human activity rather than agricultural livestock.

Figure 2.1 Land Cover Reclassification Decision Tree

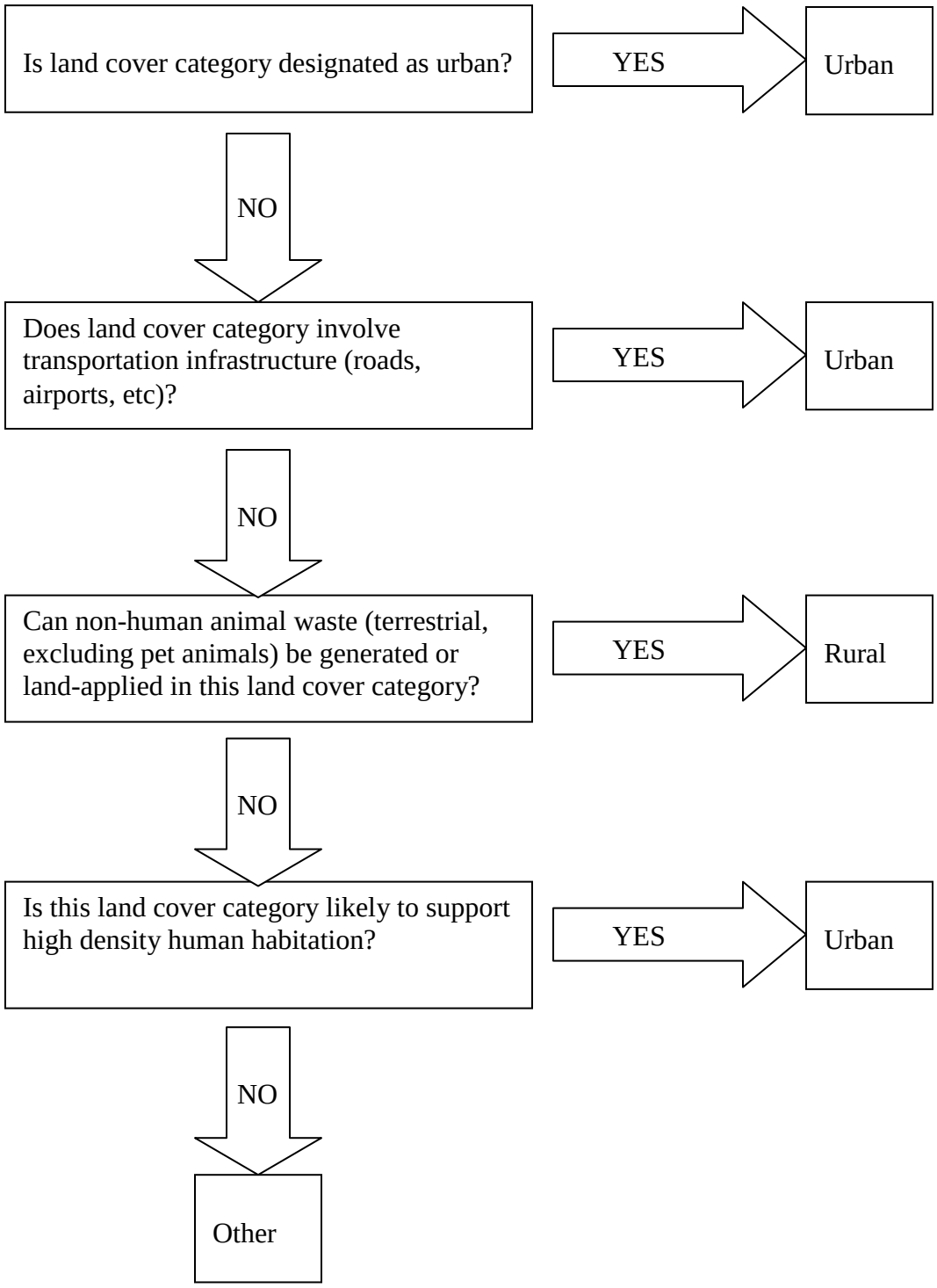


Table 2.1 Land Cover Reclassification Scheme

2001 Land cover category	Reclassified category
Non-vegetated Farmland	Rural
Forage Crops/Non-tilled Herbaceous Agriculture	Rural
Orchards/Vineyard/Nursery	Rural
Herbaceous Openland	Rural
Upland Shrub / Low density Trees	Rural
Oak Association	Rural
Aspen Association	Rural
Other Upland Deciduous	Rural
Mixed Upland Deciduous	Rural
Pines	Rural
Other Upland Conifers	Rural
Mixed Upland Conifers	Rural
Upland Mixed Forest	Rural
Lowland Deciduous Forest	Rural
Lowland Coniferous Forest	Rural
Lowland Mixed Forest	Rural
Lowland Shrub	Rural
Emergent Wetland	Rural
Mixed Non-Forest Wetland	Rural
Low Intensity Urban	Urban
High Intensity Urban	Urban
Airport	Urban
Road / Parking Lot	Urban
Parks / Golf Courses	Urban
Floating Aquatic	Other
Sand / Soil	Other
Water	Other
Exposed Rock	Other
Mud Flats	Other
Other Bare / Sparsely vegetated	Other

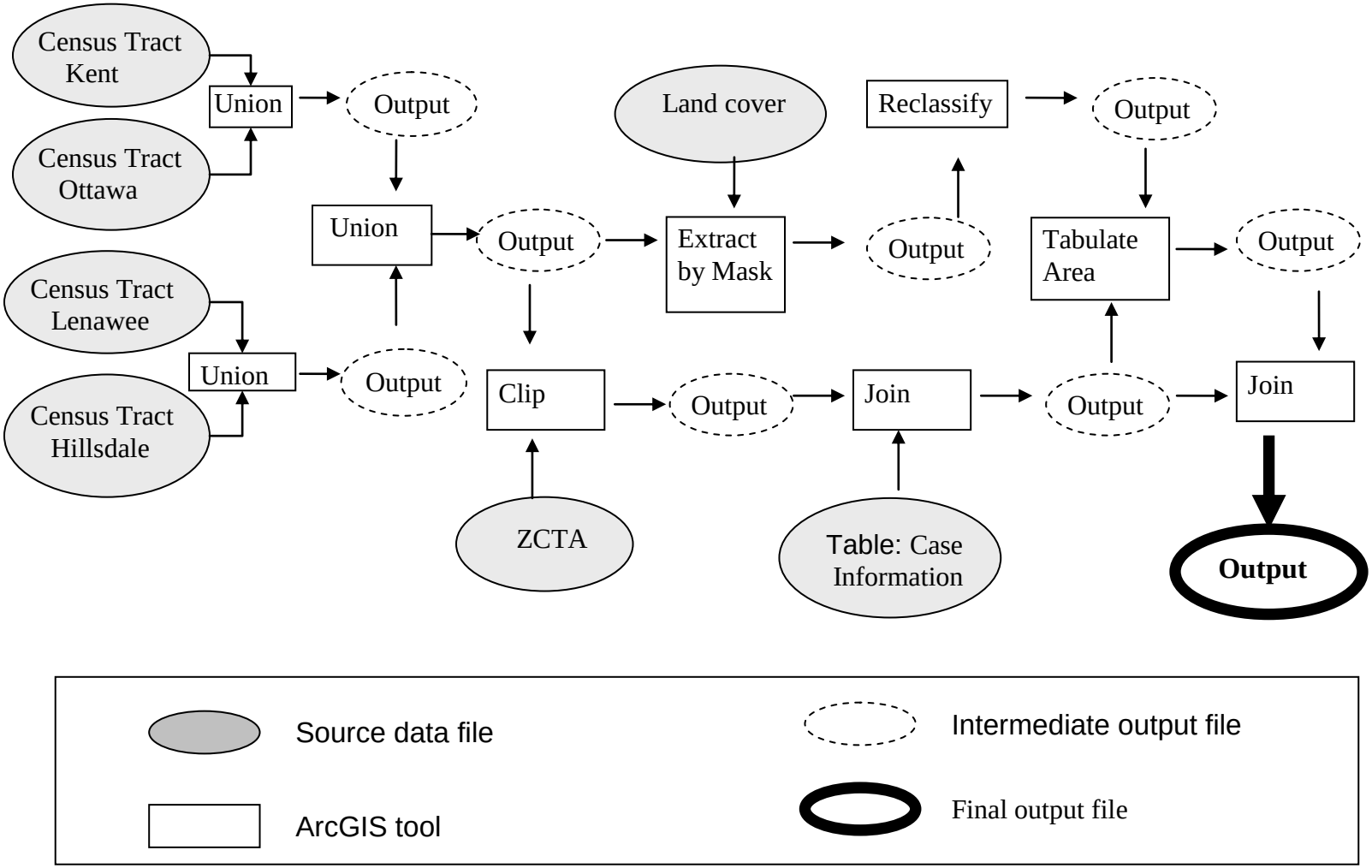
The reclassified raster file was then converted to a vector file using the spatial analyst tools. The census tract output file was also used with the "Clip" tool and the Census 2000 5-Digit ZIP Code Tabulation Areas (ZCTA file) to produce an output file with both the census tract information and the ZIP code for Kent, Ottawa, Lenawee, and Hillsdale counties. Using the "Join attributes from a table" tool, the case information table was added to the output. The join was based on the ZCTA field. The spatial analyst tool "Tabulate area" was used with the vector land cover file and the output containing the case information to produce a field with information on the percentage of "rural", "urban", and "other" land cover for each ZIP code in the counties of interest. Using the "Join" tool, this field was added to the database. The end product was a database that has records of ZIP codes, percentage of "rural" area, percentage of "urban" area, percentage of "other" area, and numbers of cases of giaridasis and numbers of cases of cryptosporidiosis during the period from January 2000 to December 2008. ZIP codes that were ≥50% rural by area were designated as rural ZIP codes. ZIP codes that were ≥50% urban by area were designated as urban ZIP codes. Information on data files used in the creation of the ArcGIS database is shown in Table 2.2. Figure 2.2 shows the data model of the analysis process.

Using the database, ZIP codes that overlapped into counties other than Kent, Ottawa, Lenawee, and Hillsdale counties were identified. Data from these ZIP codes were not used in statistical analysis of disease occurrence and land cover type.

Table 2.2 ArcGIS Database Source Files

Data Type	Coordinate System		Age of Dataset	Reference
	Geographic	Projection		
Case information by ZIP code	N/A	N/A	01/01/2000 - 12/31/2008	34
2001 Land cover Lower Peninsula (Raster)	North American Datum of 1983	Michigan GeoRef	Constructed from image dates 1997 - 2001	36
River Raisin Watershed	North American Datum of 1983	Michigan GeoRef	1998	36
Grand River Watershed	North American Datum of 1983	Michigan GeoRef	1998	36
Census 2000 5-Digit ZIP Code Tabulation Areas (ZCTAs)	North American Datum of 1983	Michigan GeoRef	1/1/2000	51
Census 2000 Tracts Hillsdale county	North American Datum of 1983	Michigan GeoRef	5/1/2009	36
Census 2000 Tracts Ottawa county	North American Datum of 1983	Michigan GeoRef	5/1/2009	36
Census 2000 Tracts Kent county	North American Datum of 1983	Michigan GeoRef	5/1/2009	36
Census 2000 Tracts Lenawee county	North American Datum of 1983	Michigan GeoRef	5/1/2009	36

Figure 2.2 ArcGIS Analysis Process Data Model



2.2 Statistical Analysis – Land Cover and the Incidence of Cryptosporidiosis and Giardiasis

There were 102 ZIP codes identified in the Kent, Ottawa, Lenawee, and Hillsdale counties and 54 of these ZIP codes were excluded from the statistical analysis. Two of the excluded ZIP codes had neither >50% urban area or >50% rural area, five ZIP codes were generic 3 digit ZIP codes which had no population associated with them (i.e. areas located alongside interstate highways), two ZIP codes had no area associated with them in the ArcGIS database, and 45 of the ZIP codes extended beyond the borders of Hillsdale, Lenawee, Kent, and Ottawa county. Thus, 48 ZIP codes of the 102 ZIP codes in Kent, Ottawa, Lenawee, and Hillsdale counties were analyzed.

2.2.1 Statistical Comparison of Urban and Rural Demographic Attributes

Information on study area attributes for each of the 48 ZIP codes was obtained from either the constructed ArcGIS database or the 2000 United States Census (51) for statistical comparison. These attributes were:

1. Population
2. Area
3. Population density
4. Percentage of individuals below poverty level
5. Median household income
6. Median age
7. Mean travel time to work
8. Percentage of Caucasian/white individuals

These attributes were compared between the urban and rural ZIP codes and examined for statistically significant differences. Datasets for statistical analysis were not normally distributed. The "Median household income" and "Percentage of individuals below poverty level" attributes could be transformed to produce normally distributed datasets using a log transformation. For these attributes, t-tests were used to examine the datasets for statistically significant differences.

No transformation was found that produced normally distributed datasets for the other demographic attributes. The non-parametric Mann-Whitney Rank Sum Test was used to examine these datasets for statistical differences.

2.2.2 Statistical Comparison of Disease Occurrence

Statistical comparisons of classified ZIP codes:

Statistical analysis of urban cryptosporidiosis, rural cryptosporidiosis, urban giardiasis, and rural giardiasis was performed using ZIP codes that had been classified as either urban or rural. Four treatments of the four datasets were analyzed:

1. Analysis of case data by ZIP code for each disease as extracted from the ArcGIS database.
2. Analysis of case data by ZIP code for each disease with number of cases divided by the area of the ZIP code.
3. Analysis of case data by ZIP code for each disease with number of cases divided by the population of the ZIP code.
4. Analysis of case data by ZIP code for each disease with number of cases divided by the population density of the ZIP code.

Datasets for statistical analysis were not normally distributed and no transformation was found that produced normally distributed datasets. Therefore, statistical analysis using non-parametric tests were used. For each analysis, a Kruskal-Wallis One Way Analysis of Variance (ANOVA) on Ranks was performed (SigmaPlot 11, Systat Software, Inc), followed by a Dunn's test to evaluate multiple pairwise differences.

Correlation assessment:

ZIP codes (non-classified into urban or rural) were examined for correlation between the cases of giardias or cases of cryptosporidiosis and nine other factors. These factors were:

1. Percentage of urban area
2. Percentage of rural area
3. Population
4. Population density
5. Percentage of Caucasian/white individuals
6. Percentage of individuals below poverty level
7. Median household income
8. Median age
9. Mean travel time to work

The percentage of urban area and percentage of rural area of ZIP codes in the study area (non-classified into urban or rural) were examined for correlation between seven other factors. These factors were:

1. Population
2. Population density
3. Percentage of Caucasian/white individuals
4. Percentage of individuals below poverty level
5. Median household income
6. Median age
7. Mean travel time to work

The statistical test used for all correlation assessments was Spearman correlation on ranks.

2.3 Sample Collection – Grand River Watershed

2.3.1 Recreational Areas

The Grand River (Michigan, USA) is 420 km long with a 14,431 km^2 drainage area and an average discharge of 108 m^3/s (50). Within the lower Grand River watershed, two sites along the Grand River, Deer Creek Park and Riverside Park, were monitored for *Cryptosporidium* and *Giardia*. North Beach Park, a Lake Michigan beach to the north of the mouth of the Grand River was also monitored for *Cryptosporidium* and *Giardia*. Deer Creek Park, Riverside Park, and North Beach Park are 29.2 km, 38.6 km, 64.9 km downstream from the City of Grand Rapids Wastewater Treatment Plant, respectively (see Figure 2.3a). Surface water samples were collected from the banks of

the Grand River at Riverside Park and Deer Creek Park at wadeable depths (approximately 30 cm). At North Beach Park, samples were collected from wadeable depths (approximately 30 cm). One 20L grab sample was collected per site during each sampling event. Between April 2005 and August 2006, 21 samples were collected from Deer Creek Park and 19 samples each were collected from Riverside Park and North Beach Park. Riverside Park, Deer Creek Park, and North Beach Park are all publicly maintained recreational areas.

2.3.2 Waste Treatment Systems

Samples were collected from the City of Grand Rapids, Michigan sewage treatment system. Samples were collected from multiple points in the Grand Rapids sewage treatment system including the Grand Rapids Market Ave Retention Basin (MARB), and combined sewer overflow (CSO) weirs during CSO events. Sampling locations are shown in Figure 2.3a and 2.3b.

Figure 2.3a shows the recreational area sampling locations of North Beach Park, Riverside Park, and Deer Creek Park within the Grand River watershed, Michigan USA. Figure 2.3b shows the sampling locations (within City of Grand Rapids city limits) of the Market Avenue Retention Basin (MARB), combined sewer weir at the Ionia and Stevens intersection, and #2 Goodrich Grand River sampling sites within the Grand River watershed, Michigan USA.

Composite samples were collected from the Grand Rapids Market Ave Retention Basin (MARB, see Figure 2.3b) during eight rainfall events between March 2008 and August 2008. The MARB retains a combination of stormwater and sewage during storm events when the treatment capacity of the Grand Rapids sewage treatment plant is

exceeded. MARB helps reduce the occurrence of CSO discharges, however, CSO discharges still occur in this system. MARB samples provide an estimate of *Cryptosporidium* spp. and *Giardia* spp. loads that may be discharged from a combined sewer. An autosampler collects the MARB influent after passage through a bar screen, but prior to entry into the MARB catchment area. This sampling location allows flows from individual rain events to be sampled without compositing with flows from previous rain events held in the MARB catchment area. The autosampler drew a 1.6L sample every 5 minutes until a total volume of 10L was collected.

Samples were collected from one CSO event in February 2006 and one event in July 2006 from a weir at the Ionia and Stevens intersection in the City of Grand Rapids (see Figure 2.3b). During the February 2006 event, 1.443 million gallons were released from the Ionia and Stevens outfall. The July 2006 event occurred during a storm that caused a total overflow of 25.331 million gallons from the Grand Rapids combined sewer system, with 1.443 million gallons released from the Ionia and Stevens CSO outfall. The Grand Rapids CSO weir locations are accessed through manholes in city streets. Installation and powering an autosampler is difficult at these locations, so samples were collected by Grand Rapids Sewer treatment personnel.

Additionally, one sample was collected during the July 2006 CSO event at the #2 Goodrich site (see Figure 2.3b) which is a Grand River sampling site 1.6 km downstream of the discharge point for the Ionia and Stevens CSO. The #2 Goodrich site was sampled to determine if *Cryptosporidium* and *Giardia* concentrations in the river could remain, elevated downstream of a CSO discharge. During each event 20L grab samples were collected.

CSO events are often of short duration, which makes it difficult for personnel to reach the site in time to collect overflow. Entering and exiting manholes located in street intersections during storm events present fall hazards and traffic hazards. To improve safety of sampling and to collect samples that allow estimation of parasite loads that enter surface waters, the majority of combined stormwater/sewage samples were collected at the Grand Rapids Market Ave Retention Basin (MARB).

Figure 2.3 Grand River Watershed Sampling Locations

Figure 2.3a

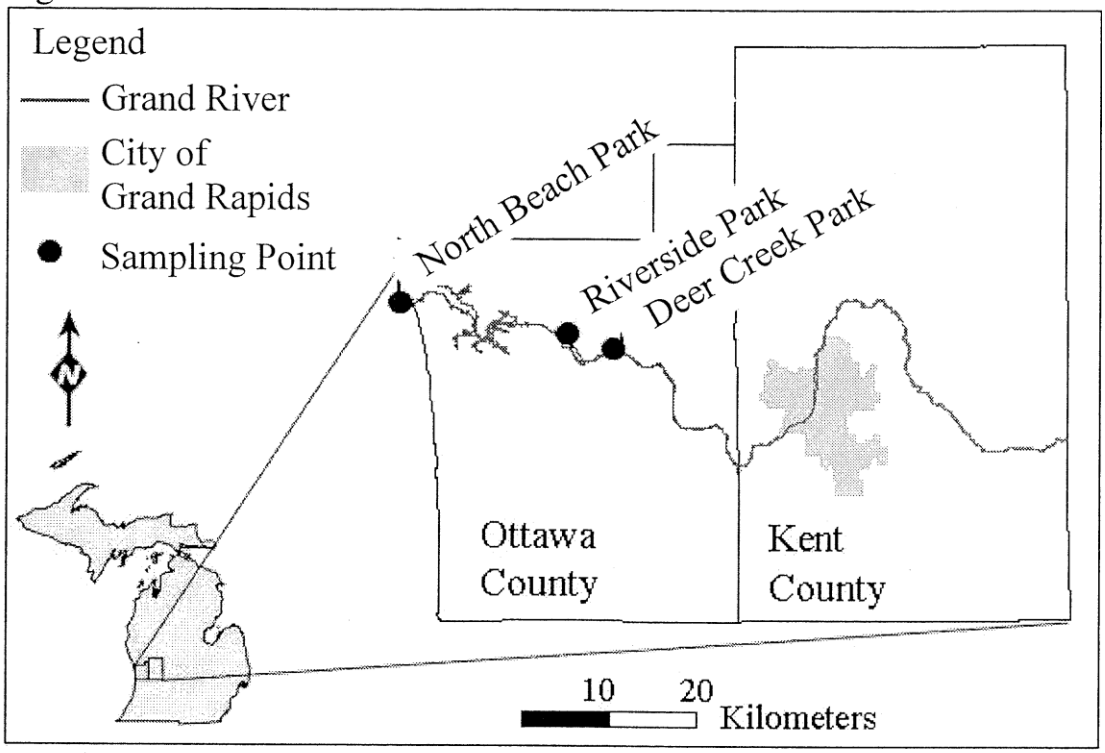
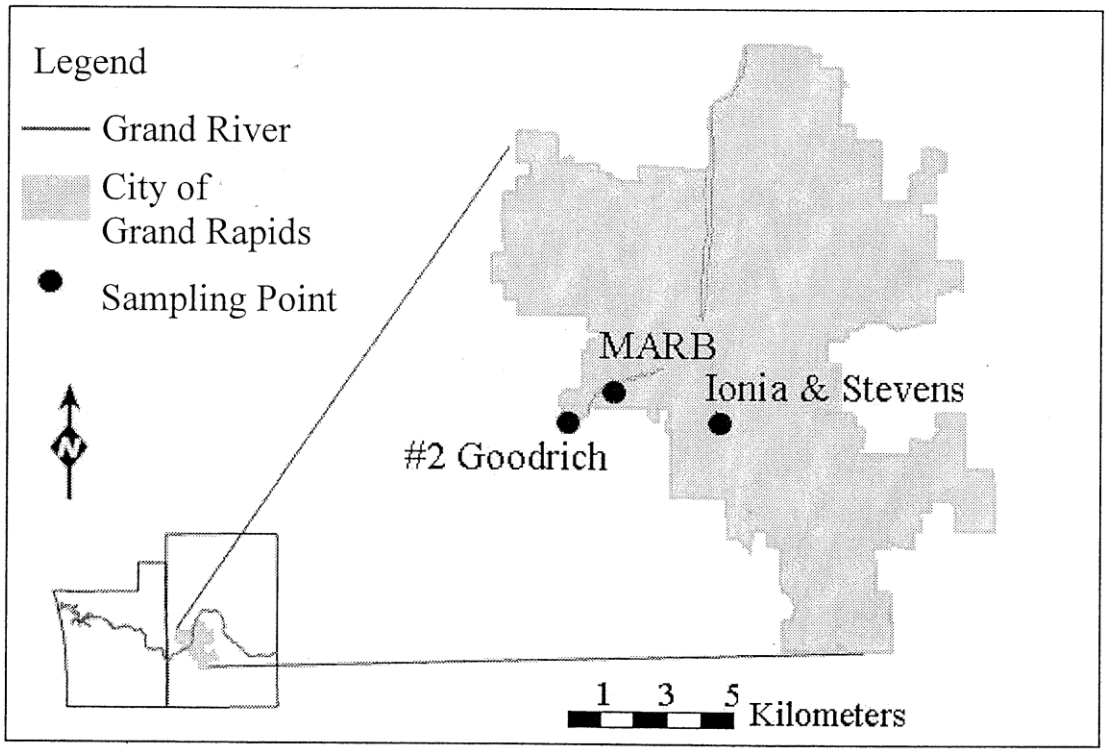


Figure 2.3b



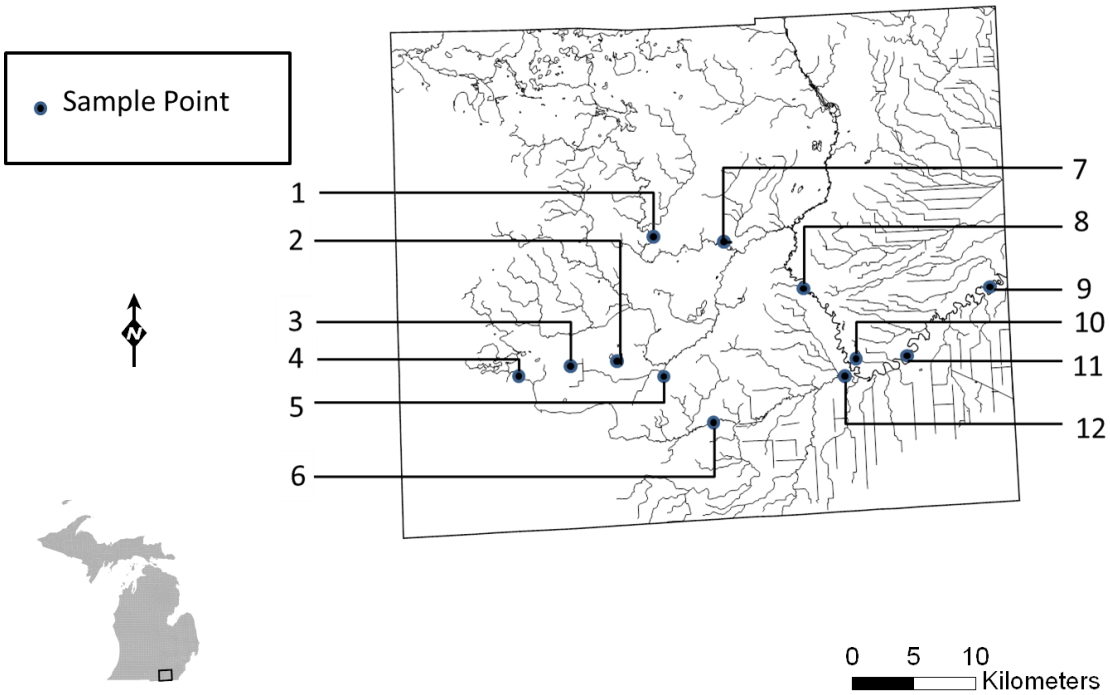
2.4 Sample Collection – River Raisin Watershed

The River Raisin (Michigan, USA) is 216 km long with a 2,776 km^2 drainage area (20). Within the Lenawee County, Michigan portion of the River Raisin watershed the main branch of the River Raisin, four tributary creeks, and one field drainage structure were sampled between June 2004 and February 2005. The locations of the sampling sites are shown in Figure 2.4.

Five sites in the main branch River Raisin were sampled; the main branch River Raisin at Crockett Rd, the main branch River Raisin at Deerfield Rd, and the main branch River Raisin at the drinking water plants of the City of Adrian, City of Deerfield, and the City of Blissfield. Two sites each were sampled from the Stoney Creek and Bear Creek tributaries and one site was sampled from the Black Creek and Wolf Creek tributaries.

The River Raisin surface waters provide drinking water to the cities of Adrian, Blissfield, and Deerfield. Rice Lake Drain is a drainage ditch adjacent to a Confined Animal Feeding Operation (CAFO). Four of the monitored sites in the River Raisin watershed are downstream of CSO outfalls. These sites include the main branch River Raisin at Deerfield Rd, the main branch River Raisin at Crockett Rd, the intake to the Deerfield water treatment plant, and the intake to the Blissfield water treatment plant. A total of 39 samples from 12 sites in the River Raisin watershed were collected. Samples were collected as 20L grab samples from each site.

Figure 2.4 River Raisin Watershed Sampling Locations



1. Wolf Creek at Forrister Rd – 3 samples
2. Rice Lake Drain – 5 samples
3. Stoney Creek at Senecca Rd – 2 samples
4. Bear Creek at Medina Rd – 1 sample
5. Stoney Creek at Gorman Rd – 2 samples
6. Bear Creek at Morse Rd – 1 sample
7. Main Branch River Raisin at Adrian Water Works (raw influent) – 5 samples
8. Main Branch River Raisin at Deerfield Rd – 2 samples
9. Main Branch River Raisin at Deerfield Water Works (raw influent) – 5 samples
10. Main Branch River Raisin at Crockett Rd – 2 samples
11. Main Branch River Raisin at Blissfield Water Works (raw influent) – 6 samples
12. Black Creek at Crockett Rd – 5 samples

2.5 Sample Processing

After collection, samples were placed on ice and transported to the laboratory for analysis. Samples that met the holding period criteria of ≤4 days were analyzed using U.S. Environmental Protection Agency (U.S.EPA) Method 1623, which is a standard method designed to meet survey and monitoring requirements of the EPA (54). The method has been validated for use in surface waters since February 1999, and has the capability of identifying the genera *Cryptosporidium* and *Giardia*. The method cannot determine species, host species of origin, or viability/infectivity of (oo)cysts.

In brief, parasite detection was performed by filtering water through Envirochek™ HV filters (Pall Gelman Laboratories, Ann Arbor, MI, USA). After filtration, the filter was eluted and the eluate concentrated by centrifugation according to U.S.EPA Method 1623. Since high turbidity samples may clog the filter after a low volume of sample has passed through the filter, for samples with turbidity >35 NTU the filtration step was omitted and the sample directly concentrated by centrifugation. After a concentrated pellet was obtained, the pellet was resuspended and divided into subsamples based on the original pellet size size so that each subsample had 0.5ml of pellet per 10mL. Parasites were separated from the resuspended materials using the Dynal Immunomagnetic Separation Technique (IMS) (Dynabeads® CG-combo Kit, Dynal Biotech, Inc., Lake Success, NY, USA) in Leighton tubes. Modifications of the 1623 protocol included a second hydrochloric acid wash step and neutralization of the IMS concentrate within a microcentrifuge tube rather than on a glass slide. Modifications were primarily made to reduce the concentration of sample debris that was not removed

by IMS, as sample debris interferes with microscopic analysis. Quality control tests were performed and acceptance criteria were met for all modifications as required in U.S.EPA Method 1623 (54). The (oo)cyst suspension was placed on slides and allowed to dry before samples were fixed with methanol and stained. Samples were stained with DAPI to help visualize nucleic acid content then stained using an immunofluorescent assay (IFA) method. This method uses monoclonal antibodies (EasyStain™, Biotechnology Frontiers, Australia) tagged with fluorescein isothiocyanate to specifically stain the (oo)cyst walls. Microscopic examination of the slides after IFA resulted in total counts of oocysts and cysts in the sample. Positive staining controls consisted of slides with purified *Cryptosporidium* and *Giardia* (EasyStain™, Biotechnology Frontiers, Australia). Negative staining controls consisted of slides prepared with phosphate buffered saline in place of the sample. These control slides were fixed, stained, and read with each set of samples processed.

Matrix specific alterations, primarily to address differences in particulate matter, are described below. When particulate matter resulted in >0.5mL of concentrated pellet following centrifugation, the pellet was divided into subsamples as required by U.S.EPA Method 1623. In this study, equivalent volumes less than 10L are due to amounts of particulate matter that resulted in subsampling.

MARB samples

Samples were mixed on a stirring plate using a magnetic stir bar. Subsample volumes of 75mL to 250mL were withdrawn and centrifuged to produce a concentrated 0.5ml pellet. The sample was remixed between withdrawal of additional subsample volumes. The pellet was processed as described above. Equivalent volumes of 75ml to

250ml were examined on slides. A total of 27 subsamples from 8 storm events were analyzed.

CSO samples

Samples were analyzed as described above. Equivalent volumes of 1L to 11.19L were examined on slides. A total of 2 samples were analyzed from the Ionia and Stevens CSO discharge.

Grand River watershed samples

Samples were analyzed as described above. A total of 22, 19, and 19 samples were analyzed for Deer Creek Park, Riverside Park, and North Beach Park, respectively. Equivalent volumes of 5.54L to 13.3L were examined on slides from Deer Creek Park, Riverside Park, and North Beach Park. An equivalent volume of 1 liter from the #2 Goodrich sample was analyzed.

River Raisin watershed samples

Samples were analyzed as described above. Equivalent volumes of 0.66L to 11.05L were examined on slides.

2.6 Recovery Efficiency

Ongoing Precision and Recovery analyses

Recovery efficiencies in laboratory reagent water were assessed by seeding with a known concentration of *Cryptosporidium* and *Giardia.* River Raisin watershed samples were seeded with Colorseed^{TM} (Biotechnology Frontiers, Australia). MARB and Grand River watershed samples were seeded with EasySeed^{TM} (Biotechnology Frontiers,

Australia). These ongoing precision and recovery (OPR) samples were processed as described above. When sample matrices required centrifugation the OPR samples were also processed by centrifugation. After processing, counts of *Cryptosporidium* and *Giardia* were compared to the number of seeded organisms and a method blank of laboratory reagent water containing no seeded *Cryptosporidium* and *Giardia* to calculate the method's efficiency. Acceptance criteria for *Cryptosporidium* and *Giardia* recovery were 13-111% and 15-118%, respectively. In model simulations, recovery efficiency was not used to extrapolate values for model simulation. A total of 23 OPR samples were performed during sampling of the Grand River watershed. A total of four OPR samples were performed during sampling of the River Raisin Watershed. A total of five OPR samples were performed during sampling of MARB.

Matrix Spikes

To determine recovery efficiencies in sample matrices, duplicate samples were seeded with a known concentration of *Cryptosporidium* and *Giardia.* River Raisin watershed samples were seeded with Colorseed™ (Biotechnology Frontiers, Australia). MARB and Grand River watershed samples were seeded with EasySeed™ (Biotechnology Frontiers, Australia). These matrix spike samples were concentrated and processed as described above. After processing, counts of *Cryptosporidium* and *Giardia* were compared to the number of seeded organisms and the number of naturally occurring *Cryptosporidium* and *Giardia* in the associated field sample to calculate the method's efficiency in the environmental matrices. Recovery efficiencies were examined in six MARB samples, two North Beach Park samples, two Riverside Park samples, two Deer Creek Park samples, and 14 River Raisin watershed samples. At least one matrix

spike was performed with every lot of samples analyzed from the River Raisin Watershed. Recovery efficiencies were not examined for CSO samples.

2.7 Statistical Analysis: Grand River and River Raisin Watershed Comparison

2.7.1 Meaningful Comparisons

Statistical analyses were conducted to identify meaningful comparisons within and between sites. In this study, meaningful comparisons are those which compare:

1. *Giardia* versus *Cryptosporidium* within the same site.
2. *Cryptosporidium* versus *Cryptosporidium* between multiple sites
3. *Giardia* versus *Giardia* between multiple sites.

In this study, statistical differences in other comparisons will not be discussed. Detection limits were analyzed to check for bias that could be introduced due to differences in analysis sensitivity. Recovery efficiencies were analyzed to check for differences in method performance.

2.7.2 Test of Normality and Equal Variance Assumptions

Probability distributions were assessed for normality using the Shapiro-Wilk test ($\alpha = 0.05$) and normal distributions were assessed for equal variance using the Levine Median test ($\alpha = 0.05$) in SigmaPlot (SigmaPlot v11.0, Systat Software, Inc).

The probability distributions of *Cryptosporidium* and *Giardia* occurrence data and detection limits all failed the normality test. Data was adjusted by adding one to each data point and then transformation of the data to produce normal distributions was attempted. Natural logarithm, $\log_{10}$, square root, square, and exponential

transformations failed the normality test. Using the reciprocal transformation, only the *Cryptosporidium* dataset for the River Raisin category downstream of the CSO outfall passed the normality test. As most occurrence and detection limit datasets violate the assumption of normality, all further statistical analyses on these datasets were conducted using non-parametric tests.

Cryptosporidium and *Giardia* recovery efficiency in sample matrices (matrix spike analyses) follow normal distribution patterns. However, all matrix spike recovery efficiency comparisons involving the River Raisin watershed samples failed the equal variance test. Therefore, all statistical tests involving matrix spike recovery efficiency in the River Raisin watershed were conducted using non-parametric statistical tests. All other matrix spike efficiency comparisons were conducted with parametric statistical tests. *Cryptosporidium* and *Giardia* recovery efficiency in laboratory reagent water (OPR analyses) follow normal probability distributions with equal variance. All statistical analyses on these datasets were conducted using parametric tests.

2.7.3 Statistical Analysis of Grand River Watershed Recreational Sites

Kruskal-Wallis one way ANOVA on ranks was used to compare parasite occurrence datasets from North Beach Park, Riverside Park, and Deer Creek Park. A pairwise multiple comparison using Dunn's method was performed to examine the groups that differed from the others. Kruskal-Wallis one way ANOVA on ranks was used to compare detection limit datasets from North Beach Park, Riverside Park, and Deer Creek Park. A t-test was used to compare recovery efficiency of *Giardia* versus recovery efficiency of *Cryptosporidium* in matrix spike samples. A t-test was used to

compare recovery efficiency of *Giardia* versus recovery efficiency of *Cryptosporidium* in OPR samples.

2.7.4 Statistical Analysis of Grand River Watershed Recreational Sites and MARB

Kruskal-Wallis one way ANOVA on ranks was used to compare parasite occurrence from MARB, North Beach Park, Riverside Park, and Deer Creek Park datasets. A pairwise multiple comparison using Dunn's method was performed to examine the groups that differed from the others. Kruskal-Wallis one way ANOVA on ranks was used to compare detection limit datasets from North Beach Park, Riverside Park, Deer Creek Park, and MARB. A pairwise multiple comparison using Dunn's method was performed to examine the groups that differed from the others. Statistical differences between detection limits at the #2 Goodrich site (single sample) and North Beach Park, Deer Creek Park, and Riverside Park were not examined. One way ANOVA was used to compare parasite recovery efficiency in matrix spike samples from MARB and the combined North Beach Park, Riverside Park, and Deer Creek Park recovery efficiency dataset.

One way ANOVA was used to compare parasite recovery efficiency in OPR samples from MARB and the combined North Beach Park, Riverside Park, and Deer Creek Park recovery efficiency dataset ($p=0.005$).

2.7.5 Statistical Analysis of River Raisin Watershed

Kruskal-Wallis one way ANOVA on ranks was used to compare parasite occurrence datasets from sites upstream of CSO outfalls against sites downstream of CSO outfalls. Parasite occurrence datasets from upstream and downstream sites were

pooled and *Cryptosporidium* occurrence was compared to *Giardia* occurrence using the Mann-Whitney rank sum test. Detection limit datasets from upstream and downstream sites were compared using the Mann Whitney rank sum test. Kruskal-Wallis one way ANOVA on ranks was used to compare parasite recovery efficiency in matrix spike samples between sites upstream of CSO outfalls and sites downstream of CSO outfalls. A pairwise multiple comparison using Dunn's method was performed to examine the groups that differed from the others. Parasite recovery efficiency in matrix spike sample datasets from upstream and downstream sites were pooled and compared using the Mann-Whitney rank sum test. Recovery efficiencies of *Giardia* and *Cryptosporidium* in OPR samples were compared using a t-test.

2.7.4 Statistical Analysis of River Raisin vs Grand River Watershed

Deer Creek Park and Riverside Park parasite occurrence datasets were pooled for comparison with the River Raisin watershed sites. Upstream and downstream sites in the River Raisin watershed were pooled for comparison with the Grand River watershed sites.

Kruskal-Wallis one way ANOVA on ranks was used to compare parasite occurrence in datasets from North Beach Park, from the combined Deer Creek Park- Riverside Park dataset, and from the total River Raisin dataset. A pairwise multiple comparison using Dunn's method was performed to examine the groups that differed from the others.

Kruskal-Wallis one way ANOVA on ranks was used to compare detection limits in datasets from North Beach Park, from the combined Deer Creek Park- Riverside Park dataset, and from the total River Raisin dataset. A pairwise multiple comparison using Dunn's method was performed to examine the groups that differed from the others.

Kruskal-Wallis one way ANOVA on ranks was used to compare parasite recovery efficiency in matrix spike samples from the total Grand River dataset (North Beach Park, Deer Creek Park, and Riverside Park) and from the total River Raisin dataset A pairwise multiple comparison using Dunn's method was performed to examine the groups that differed from the others.

One way ANOVA was used to compare parasite recovery efficiency in OPR samples from the total Grand River dataset (North Beach Park, Deer Creek Park, and Riverside Park) and from the total River Raisin dataset. A pairwise multiple comparison using the Holm Sidak method was performed to examine the groups that differed from the others.

2.8 Risk Simulation and Statistical Analysis

A scenario for risk assessment involving children under the age of 16 swimming in recreational areas was developed and analyzed using a Monte Carlo method. This risk assessment focused on ingestion of water contaminated with *Cryptosporidium* and *Giardia* and the probability of infection and illness. Probability distributions for exposure factors, including the amount of time spent swimming and the volume ingested during swimming, were developed using Crystal Ball™ (version 7.3.1, Oracle). Probability distributions were also calculated for the fraction of ingested organisms that survive ingestion to initiate infection (based on exponential dose response parameter). Concentrations of parasites from samples collected from the three recreational areas (Riverside Park, Deer Creek Park, North Beach Park) and the Market Avenue Retention Basin were fit to distributions, which were ranked according to the Kolmogorov-Smirnov goodness-of-fit test. The highest ranking distribution was chosen as the best fit. In distribution fitting, non-detects were replaced by the detection limit, producing the

most conservative (i.e. most protective) concentration estimates. All detected parasites were assumed to be viable and capable of initiating infection in humans. Data on duration and volume of CSO events reported by the Grand Rapids Wastewater Treatment Plant (35) and the mean volumetric flow of the Grand River from January 1, 2002 to December 31, 2006 at the USGS gauging station located Grand Rapids (52) were also fit to distributions using the Kolmogorov-Smirnov goodness-of-fit test. All distributions were truncated at a minimum of zero. Table 5.1 contains the distributions and distribution parameters.

Equation 1 is the daily ingested number of organisms (d_{SITE}) at each recreational site calculated as a function of the concentration of organisms at a site (C_{SITE}), ingestion volume during swimming (I), the amount of time spent swimming (T_{SWIM}) and a constant to convert units of months to days. d_{SITE} was calculated for each of the three recreational sites.

(1) $$d_{SITE} = 0.033 \cdot C_{SITE} \cdot I \cdot T_{SWIM}$$

Equation 2 is the daily ingested number of organisms (d_{RIVER}) in the river at the CSO discharge point calculated as a function of the concentration of organisms in the river at the CSO discharge point (C_{RIVER}, see Equation 5), ingestion volume during swimming (I), the amount of time spent swimming (T_{SWIM}) and a constant to convert units of months to days.

(2) $$d_{RIVER} = 0.033 \cdot C_{River} \cdot I \cdot T_{SWIM}$$

Equation 3 is the daily probability, or risk, of infection (P_I) due to ingestion of a volume of contaminated water where r = fraction of ingested organisms that survive to initiate infection (45). The value of r for each organism is shown in Table 5.1. In equation 3, d = d_{SITE} when calculating the probability of infection at each recreational site and d = d_{RIVER} when calculating the probability of infection at the CSO discharge point.

(3) $$P_I = 1 - \exp(-r \cdot d)$$

Equation 4 is the annual risk of infection due to ingestion (P_{ANNUAL}). P_{ANNUAL} is a function of the daily risk where N is the number of days of exposure during a recreational season to the health hazard (25), in this case, parasites.

(4) $$P_{ANNUAL} = 1 - (1 - P_I)^N$$

Equation 4 was used to calculate daily risk by setting N = 1. Seasonal risk of infection was calculated by making the assumption that recreation occurs over a period of 90 days in the summer (N = 90).

Equation 5 is the concentration of parasites in the river at the CSO discharge point (C_{RIVER}).

(5) $$C_{RIVER} = \frac{C_{CSO} \cdot V_{CSO}}{T_{CSO}(V_{CSO} \cdot d^{-1} + Q_R)}$$

Where C_{RIVER} = Parasite concentration in the river

C_{CSO} = Parasite concentration in the MARB influent (estimate of CSO discharge concentration)

V_{CSO} = Volume of CSO discharge

T_{CSO} = Duration of the CSO discharge

Q_R = Volumetric flow rate of the river

The risk of daily illness and seasonal illness was calculated as a function of the daily probability of infection multiplied by a morbidity ratio (see Table 5.1). The morbidity ratio for *Giardia* was based on an indirect assessment of outbreak attack rates compared with infectivity dose-response relationships (45). The morbidity ratio for *Cryptosporidium* was based on dose-response studies in healthy adults (25).

Using Crystal Ball™ (version 7.3.1, Oracle), Monte Carlo analysis was conducted by running 10,000 simulations based on values randomly generated from the parameter distributions. Each of these randomly generated values was then used in the calculations above to produce an estimated distribution of risk. Probability distributions for daily risk of illness and seasonal risk of illness were compared with the EPA

recommended illness rate water quality criteria for recreational waters of 0.8% (8 illnesses in 1,000 swimmers) (21).

One way analysis of variance was performed on the simulation outputs using the mean, standard deviation, and simulation run size of 10,000 (SigmaPlot v11.0, Systat Software, Inc). The Holm-Sidak method for multiple pairwise comparison was used to identify significant differences in risk of infection at Deer Creek Park, Riverside Park, and at the point of CSO discharge to the river. The Holm-Sidak method for multiple pairwise comparison was also used to identify significant differences in risk of illness at Deer Creek Park, Riverside Park, and at the point of CSO discharge to the river.

Chapter Three: Land Cover and the Incidence of Cryptosporidiosis and Giardiasis

In chapter three, the patterns between disease and land cover type were explored using the prototype GIS database. The null hypothesis that disease levels are the same in urban and rural areas was tested. Cryptosporidiosis and giardiasis were the diseases chosen for this study.

3.1 Results

Figure 3.1 shows the land cover classifications of the study area. Table 3.1 summarizes attributes of the study area by rural and urban designation, including the number of ZIP codes, area, population, population density. The majority of ZIP codes in the study area were designated as rural ZIP codes. In the total study area, the rural area and rural population were greater than the urban area and urban population. However, the median population value was greater in the urban ZIP codes than in the rural ZIP codes. The urban population density of the study area was greater than the rural population density.

Table 3.1 also summarizes demographic information about the study area, including the percentage of the population that is Caucasian/white, median age, mean travel time to work, median age, median household income, and the percentage of individuals living below the poverty level. When compared to urban ZIP codes, the rural ZIP codes had a higher median percentage of Caucasian/white individuals in the population, the median population age was older, the mean travel time to work was longer, and the median

household income was higher. Additionally, in the rural ZIP codes, a lower percentage of individuals were living below the poverty level than in the urban ZIP codes.

Results of the statistical comparison of the study area attributes are shown in Table 3.1. Each attribute investigated showed a significant difference between the urban and rural ZIP codes, with p values ranging from <0.001 to 0.031

Figure 3.1 Land Cover Classifications of Study Area

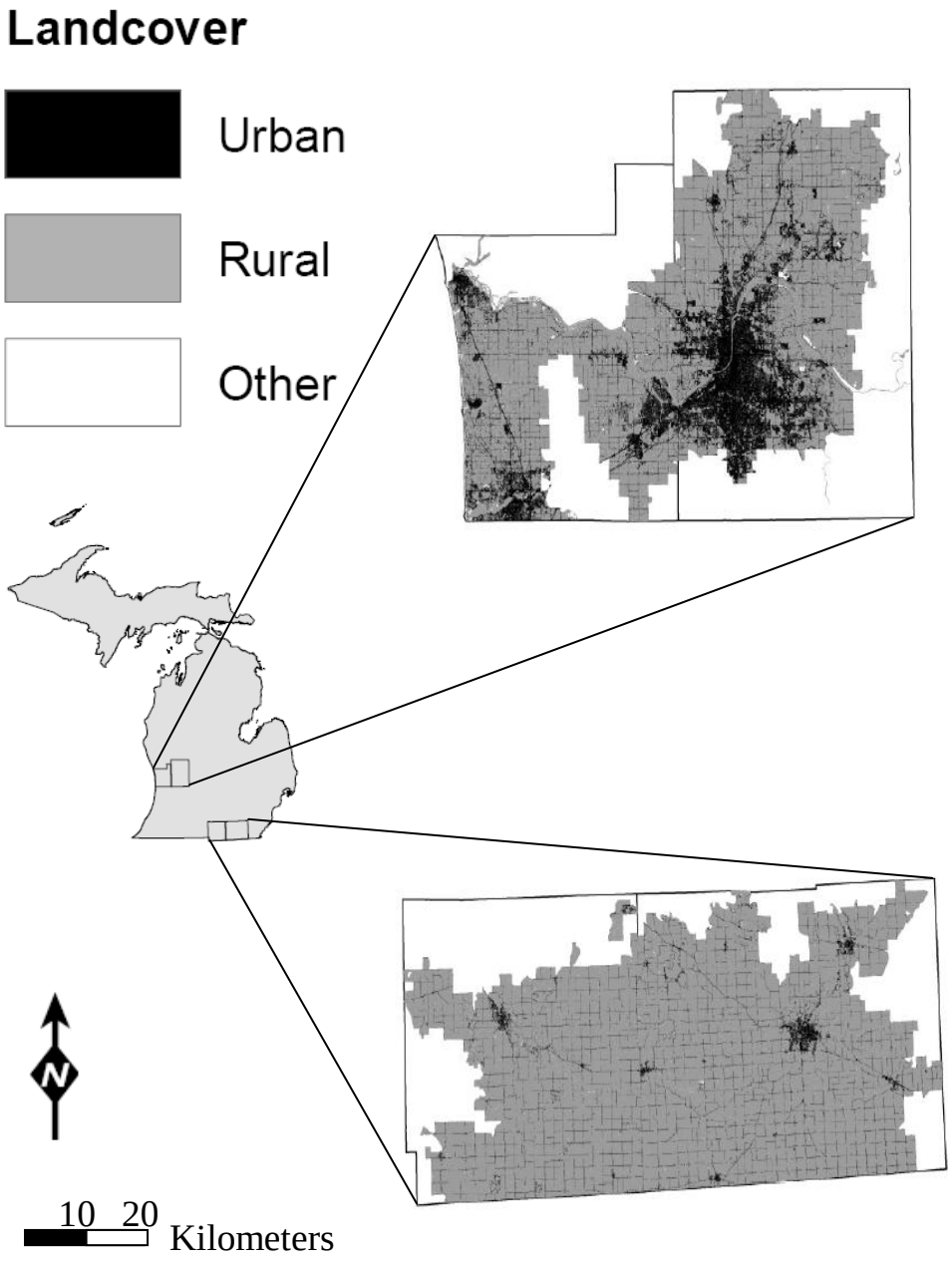


Table 3.1 Attributes of Study Area

Attributes	ZIP code designation	ZIP code median value	Total study area	t – Test Urban vs Rural	Mann-Whitney Rank Sum Test Urban vs Rural
Number of ZIP codes analyzed			48		
Number of ZIP codes	Urban		8		
	Rural		40		
Population*	Urban	**36,821.5**	310,302		**Significant difference (P<0.001)**
	Rural	6,875	475,667		
Area (km^2)	Urban	25.380	211.00		**Significant difference (P<0.001)**
	Rural	**89.818**	4529.0		
Population density (People/km^2)	Urban	**1383**	1469		**Significant difference (P<0.001)**
	Rural	61	105		
Percentage individuals below the poverty level*	Urban	**10.05**		**Significant difference (P<0.001)**	
	Rural	5.75			
Median household income (US dollars)*	Urban	$40,127		**Significant difference (P=0.031)**	
	Rural	**$46,270**			
Median age (Years) *	Urban	32.2			**Significant difference (P<0.001)**
	Rural	**35.95**			
Mean travel time to work (minutes) *	Urban	19.5			**Significant difference (P=0.001)**
	Rural	**23.9**			
% Caucasian/ white*	Urban	79.4			**Significant difference (P=0.001)**
	Rural	**97.0**			

*Source: 2000 US Census (51)

A comparison of the median number of disease cases per ZIP code and total disease cases by urban or rural designation over the eight year study period is shown in Table 3.2. When comparing locations with the greatest amount of disease incidence, there was a difference in disease pattern between the median case per ZIP code and the total number of cases. The dataset of disease cases represents disease incidence, which is the number of cases in a given amount of time. In this dataset, the total number of both giardiasis and cryptosporidiosis cases were greatest in the rural areas. However, the ZIP code median values of giardiasis and cryptosporidiosis case incidence were greatest in the urban ZIP codes. Once the incidence datasets were normalized (by population, area, or population density), the patterns of disease occurrence between the "total number of cases" and the "median number of cases per ZIP code" categories were in agreement.

When the disease cases were divided by population values, the overall number of cryptosporidiosis disease cases was greatest in the rural ZIP codes. Conversely, for giardiasis, the greatest numbers of disease cases were in the urban ZIP codes. When the disease cases were divided by area, the numbers of disease cases were greatest in the urban ZIP codes for both cryptosporidiosis and giardiasis. When the disease cases were divided by population density, the numbers of disease cases was greatest in the rural ZIP codes for both cryptosporidiosis and giardiasis.

Results of the statistical comparison of disease between urban and rural designated ZIP codes are shown in Table 3.3. All examined treatments resulted in significant differences between the datasets examined. Results from the Dunn's test to evaluate multiple pairwise differences are also shown in Table 3.3. The source of the statistical difference differed between the four treatments of the dataset. The rural giardiasis vs

urban cryptosporidiosis and urban giardiasis vs rural cryptosporidiosis comparisons are not shown in Table 3.3 as these comparisons are not particularly meaningful in elucidating disease transmission.

Results of the correlation assessment are shown in Table 3.4. No significant correlations with either the median household income or the percentage of individuals below poverty level were found ($p>0.15$). Three strong correlations were found. Strong positive correlations exist between giardiasis and population ($r = 0.918$, $p<0.001$) and between the percentage of urban area and population density ($r = 0.978$, $p<0.001$). A strong inverse relationship exists between the percentage of rural area and population density ($r = -0.977$, $p<0.001$). Five moderately strong positive correlations and four moderately strong inverse correlations were found in this assessment. The moderately strong positive correlations consisted of giardiasis and population density ($r =0.799$, $p<0.001$), giardiasis and percentage urban area ($r=0.79$, $p<0.001$), cryptosporidiosis and population ($r = 0.793$, $p<0.001$), percentage urban area and population ($r = 0.754$, $p<0.001$), and percentage rural area and percentage of Caucasian / white individuals ($r = 0.763$, $p<0.001$). The moderately strong inverse correlations consisted of giardiasis and percentage rural area ($r = -0.788$, $p<0.001$), the percentage of urban area and percentage of Caucasian / white individuals ($r = -0.777$, $p<0.001$), the percentage of rural area and population ($r = -0.752$, $p<0.001$), and giardiasis and mean travel time to work ($r = -0.75$, $p<0.001$).

The relationship with population produced the highest correlation coefficients for cases of either giardiasis or cryptosporidiosis, although the correlation with giardiasis is

strong and the correlation with cryptosporidiosis is only moderate. For both giardiasis and cryptosporidiosis, the next strongest relationship was with population density.

Table 3.2 Comparison of Disease Cases by Urban or Rural Designation.

Disease	ZIP code designation	Treatment of dataset	Median number of cases per ZIP code	Total number of cases
Cryptosporidiosis	Urban	Disease Cases	5.0	39
		$\frac{\text{Disease Cases}}{\text{Population}}$ [a]	0.000136	0.00013
		$\frac{\text{Disease Cases}}{\text{Area}}$	0.184	0.185
		$\frac{\text{Disease Cases}}{\text{Population density}}$	0.00330	0.027
Cryptosporidiosis	Rural	Disease Cases	1.5	113
		$\frac{\text{Disease Cases}}{\text{Population}}$ [a]	0.000184	0.00024
		$\frac{\text{Disease Cases}}{\text{Area}}$	0.0145	0.025
		$\frac{\text{Disease Cases}}{\text{Population density}}$	0.0156	1.076
Giardiasis	Urban	Disease Cases	53.0	405
		$\frac{\text{Disease Cases}}{\text{Population}}$ [a]	0.00117	0.00131
		$\frac{\text{Disease Cases}}{\text{Area}}$	1.96	1.92
		$\frac{\text{Disease Cases}}{\text{Population density}}$	0.0340	0.276
Giardiasis	Rural	Disease Cases	5.0	467
		$\frac{\text{Disease Cases}}{\text{Population}}$ [a]	0.000788	0.00098
		$\frac{\text{Disease Cases}}{\text{Area}}$	0.0392	0.100
		$\frac{\text{Disease Cases}}{\text{Population density}}$	0.0729	4.448

[a] Calculated as the number of cases per capita

Table 3.3 Statistical Comparison of Disease Occurrence.

Treatment of dataset	ANOVA on Ranks	Multiple Pairwise Occurrence (Dunn's Test)[a]			
		Cryptosporidiosis Urban vs Rural	Giardiasis Urban vs Rural	Urban *Cryptosporidiosis* vs *Giardiasis*	Rural *Cryptosporidiosis* vs *Giardiasis*
None (incidence)	**Significant difference (P<0.001)**	No significant difference	**Significant difference (P<0.05)**	No significant difference	No significant difference
Population normalized	**Significant difference (P<0.001)**	No significant difference	No significant difference	**Significant difference (P<0.05)**	No significant difference
Area normalized (disease intensity)	**Significant difference (P<0.001)**	**Significant difference (P<0.05)**	**Significant difference (P<0.05)**	No significant difference	No significant difference
Population density normalized	**Significant difference (P=0.002)**	No significant difference	No significant difference	No significant difference	**Significant difference (P<0.05)**

[a] Results of rural giardiasis vs urban cryptosporidiosis and urban giardiasis vs rural cryptosporidiosis comparisons not shown.

Table 3.4 Correlation Assessment.

Factor	Giardiasis	Cryptosporidiosis	%Urban	% Rural
%Urban				
r	**0.79**	0.595	--	--
p value	**<0.001**	<0.001	--	--
% Rural				
r	**-0.788**	-0.591	--	--
p value	**<0.001**	<0.001	--	--
Population				
r	**0.918**	**0.793**	**0.754**	**-0.752**
p value	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Population density				
r	**0.799**	0.609	**0.978**	**-0.977**
p value	**<0.001**	<0.001	**<0.001**	**<0.001**
% Caucasian / white				
r	-0.694	-0.554	**-0.777**	**0.763**
p value	<0.001	<0.001	**<0.001**	**<0.001**
Median age				
r	-0.638	-0.49	-0.554	0.557
p value	<0.001	<0.001	<0.001	<0.001
Mean Travel Time to Work				
r	**-0.75**	-0.663	-0.65	0.648
p value	**<0.001**	<0.001	<0.001	<0.001
Median Household Income				
r	0.207	0.121	0.0939	-0.0843
p value	0.157	0.413	0.524	0.567
% Individuals Below the Poverty Level				
r	0.115	-0.0525	0.177	-0.19
p value	0.434	0.722	0.227	0.194

3.2 Discussion

In this study, after assigning ZIP codes a rural or urban designation, the study area attributes were examined for differences that might produce significant differences between disease in urban and rural areas. Significant differences were found in all attributes examined. Three of these attributes (population, area, and population density) were used to normalize the cases of disease. The attributes of population and area both had higher values in rural areas than in urban areas. In contrast, the population density was higher in urban areas than in rural areas. The three normalized datasets and the non-normalized dataset were used to test the null hypothesis that disease levels in urban and rural areas are the same.

Based on the statistically significant differences between disease occurrence in ZIP codes designated as urban and those designated as rural ($p<0.001$), the null hypothesis was rejected in favor of the research hypothesis that there were significant differences between patterns of disease in urban and rural areas. Statistically significant differences were present between disease occurrence in ZIP codes designated as urban and those designated as rural ($p<0.001$). This statistical difference was maintained even when disease occurrence was normalized by different attributes present in the study area (ie population, area, or population density of the ZIP codes) although the cause(s) of this difference changed as the datasets were normalized. For example, in the non-normalized dataset, the cause of the significant difference between urban disease and rural disease was due to the difference between urban giardiasis and rural giardiasis. However, when the dataset was population normalized the cause of the significant difference between

urban disease and rural disease was the difference between urban cryptosporidiosis and urban giardiasis.

As there was no obvious pattern in the pairwise comparisons in the normalized or non-normalized datasets that suggested particular transmission routes, the disease cases in all ZIP codes were examined for correlations with the study area attributes. For the correlation assessment, instead of using the urban and rural designations for the ZIP codes, the percentages of urban area and rural area for each ZIP code were used. Of the resulting correlations, the direct correlation with population was strongest for both giardiasis and cryptosporidiosis. The strong direct correlation of population with giardiasis ($r=0.918$, $p<0.001$) and the moderate direct correlation of population with cryptosporidiosis ($r=0.793$, $p<0.001$) suggest that population levels were important factors of transmission of these parasitic diseases in the study area. Since the percentage of urban area had a direct positive relationship with the population levels ($r = 0.754$, moderate) and population density ($r = 0.978$, strong), positive relationships between the percentage of urban area and disease cases were logical. The correlation between the percentage of urban area and giardiasis is stronger than the correlation between the percentage of urban area and cryptosporidiosis, however. The correlation of disease with population, particularly for giardiasis, suggests that higher disease transmission is mediated in some way by larger population levels. This may involve direct person to person transmission, transmission through food or water contaminated with feces from an infected human, or transmission through the environment (ie contaminated recreational water).

Livestock density and manure land application were not included in the prototype database, although these may be important factors in disease transmission. Information on livestock density is only available at the county scale. Including this information into the prototype database would require all comparisons to be done at the county scale, potentially obscuring patterns that occur at more local levels. However, there are some studies in the scientific literature that have examined associations between cryptosporidosis / giardiasis and livestock density, manure application, and other livestock related factors. These studies provide insight into which attributes would be desirable to include in an expanded GIS database.

Previous studies of cryptosporidiosis and giardiasis in England, Scotland, and Wales have demonstrated association with rural areas in some studies and associations with urban areas in others. During the 2001 foot and mouth disease outbreak in livestock in England and Wales, in which control measures were implemented including restriction of access to farms, limiting movement of livestock for trade and between pasturage, and culling of affected herds and flocks, there was a corresponding decrease in reported cryptosporidiosis in humans ranging from 35-63% (47) throughout England and Wales with declines of 81.8% in northwest England (23). During the interval of foot and mouth disease control measures, the proportion of *C. parvum* cases decreased compared to case incidence in 2000, suggesting a decrease in human infection due to a decrease in exposure to pathogen reservoirs in livestock (47). No significant reduction in giardiasis cases was observed during this interval, which may indicate a difference between *Giardia* and *Cryptosporidium* transmission routes and reservoirs of infection (47). Studies of area based cryptospordiosis rates in England and Wales using housing density

to define rural areas demonstrated higher illness rates in rural areas than urban areas, higher rates in areas with more agricultural manure application, and higher rates in areas with inadequate drinking water treatment (31, 32). A study of spatial epidemiology of sporadic cases of human cryptosporidiosis in Scotland found increased rates of *C. parvum* infection in areas with lower human density, a higher ratio of farms to humans, and a higher ratio of private water supplies to the human population, indicating an association of *C. parvum* infection with rural areas (42). Unlike *C. parvum*, *C. hominis* was reported more often in the more heavily populated areas of south Scotland, associating this genotype more strongly with urban areas (42). In a case control study of cryptosporidiosis in the United Kingdom, the urban-rural gradient was not found to be a significant variable in the full model of disease aetiology when both *C. parvum* and *C. hominis* cases were included. However, when *C. hominis* cases were excluded, cryptosporidiosis was negatively associated with urban areas and when *C. parvum* cases were excluded, cryptosporidiosis was positively associated with urban areas, indicating genotype specific transmission associated with the geographical classifications (32). The differences in the association of cryptosporidiosis with urban or rural areas in the scientific literature may therefore be partially due to the *Cryptosporidium* genotype causing the infections.

In a spatial investigation of giardiasis in Canada that explored associations with livestock density and land application of manure with disease patterns, low correlation coeffients between giardiasis rates and cattle density (r=0.11) and between giardiasis rates and land application (r=0.09) were observed when all geographic regions in the study area were included (37). However, these correlation coefficients were higher in

certain regions of the study area when these areas were independently examined (37). These results suggest that livestock density and land application of manure can contribute to transmission of *Giardia*, but that other factors may be more important.

In the current study, the major caveat to the conclusion that significant differences in the patterns of giardiasis and cryptosporidiosis exist between urban and rural ZIP codes is the small number of ZIP codes that are designated as urban in the study area. Since this method of examining health data appears promising, expanding the study area to larger geographic regions of the United States (ie Midwest) in future work is recommended. Expanding the geographic region would also allow for agricultural census data to be incorporated into the study design by allowing analysis at the county level, which is the minimum scale of the agricultural census data. Previous studies from the scientific literature suggest that livestock density, animal transport frequency, the number of farm workers/visitors, manure application rate, ratio of farms to humans, and the *Cryptosporidium* genotype would be informative attributes in an expanded GIS database. Utilizing GIS as a tool to integrate factors from the rural environment with factors from the urbanized environment has the potential to be extremely useful to public health agencies in targeting funds to reduce disease transmission in communities.

3.3 Conclusion

The prototype GIS database was a useful tool for testing hypotheses regarding disease incidence and land cover type. The relatively small number of urban zip codes versus the number of rural zip codes was the major caveat in hypothesis testing. During the development of the prototype database, data gaps in agricultural information and data

scale issues were identified. Expansion of the study area, identified data gaps and scale issues will be addressed in the next iteration of the GIS database.

Significant differences in the patterns of giardiasis and cryptosporidiosis exist between urban and rural ZIP codes in the study area. The strong direct correlation of population with giardiasis and the moderate direct correlation of population with cryptosporidiosis suggest that population levels are important factors in determining disease incidence in the study area. Giardiasis was moderately correlated with the percentage of urban area and cryptosporidiosis was weakly correlated with the percentage of urban area.

Chapter 4. Grand River and River Raisin Watershed Comparison: *Cryptosporidium* and *Giardia* Occurrence

In this study, the association of land cover with the parasites *Cryptosporidium* and *Giardia* was further examined by comparing parasite concentrations in surface water at the watershed scale. Two watersheds with differences in the percentage of urban and rural land cover were compared to see if occurrence patterns in these watersheds were associated with land cover types. The metrics of this comparison differs from the comparison in Chapter 3 in that the concentration of parasites in surface water rather than disease is the factor of interest and the scale of the comparison changed from the ZIP code scale to the watershed scale. Surface waters of the Grand River watershed and River Raisin watershed were first individually examined for parasite occurrence and then the parasite concentrations of each watershed compared against one another. As the experimental design for these watersheds was not originally planned for this inter-watershed comparison (post-hoc testing), detection limits and recovery efficiency were considered alongside parasite occurrence. Differences in these factors could introduce bias into comparisons of parasite occurrence.

The following datasets from the Grand River watershed in Ottawa County, MI were assessed: North Beach Park, Riverside Park, Deer Creek Park, and combined datasets of Riverside Park and Deer Creek Park. Occurrence data from the River Raisin watershed in Lenawee County, MI (12 sampling sites) were divided into 3 categories: upstream of Combined Sewer Overflow (CSO) outfalls, downstream of CSO outfalls, and total data. Locations of sampling sites, sample collection and processing methods are described in Chapter 2.

Figure 4.2 shows the location and land cover classifications of the watersheds in the counties of interest. Land cover classification methods are described in Chapter 2. The Grand River watershed in Ottawa County was 86.0% rural and 11.9% urban. The River Raisin watershed in Lenawee County was 94.1% rural and 5.5% urban.

4.1 Parasite Occurrence Results

Cryptosporidium and *Giardia* were detected at all of the monitored sites during at least one sampling event. Among the Grand River recreational sites, *Cryptosporidium* and *Giardia* percentage occurrence was highest at Riverside Park, 73.7% and 89.5%, respectively, and lowest at North Beach Park, 26.3% and 26.3%, respectively (Table 4.1). At Deer Creek Park, *Cryptosporidium* was detected in 59.1% of the samples and *Giardia* in 36.4%. The Market Avenue Retention Basin (MARB) samples were positive 62.5% and 100% of the time for *Cryptosporidium* and *Giardia*, respectively. *Cryptosporidium* occurred at a frequency equal or greater than occurrence of *Giardia* in North Beach Park and Deer Creek Park samples. In Riverside Park and MARB samples, however, the percentage of *Giardia* occurrence is higher than the percentage of *Cryptosporidium* occurrence.

Among the River Raisin watershed sites, *Cryptosporidium* was detected at least 50% of the time (Table 4.1). *Giardia* was detected in the River Raisin watershed less frequently than *Cryptosporidium*. *Cryptosporidium* percentage occurrence was slightly higher in the River Raisin watershed sites downstream from CSO outfalls (53.3%) than in the River Raisin watershed sites upstream from CSO outfalls (50%). Conversely, *Giardia* percentage occurrence was higher in the River Raisin watershed sites upstream

from CSO outfalls (29.2%) than in the River Raisin watershed sites downstream from CSO outfalls (13.3%). When evaluating the combined Grand River recreational sites (Deer Creek Park and Riverside Park) and all sites in the River Raisin watershed, *Cryptosporidium* was detected more frequently than *Giardia* in the aggregated datasets.

Table 4.1. Water Samples Positive for *Cryptosporidium* and *Giardia*

Watershed	Site	Number of sampling locations	Total number of sample events	*Cryptosporidium* Positive Events (%)	*Giardia* Positive Events (%)
Grand River watershed	North Beach Park	1	19	5 (26.3)	5 (26.3)
	Riverside Park	1	19	14 (73.7)	17 (89.5)
	Deer Creek Park	1	22	13 (59.1)	8 (36.4)
	Grand River Sites (Deer Creek & Riverside Park)	2	41	27 (65.9)	25 (61.0)
	MARB	1	8[d]	5 (62.5)	8 (100)
River Raisin watershed	River Raisin[a]	8	24	12 (50.0)	7 (29.2)
	River Raisin[b]	4	15	8 (53.3)	2 (13.3)
	River Raisin[c]	13	39	20 (51.3)	9 (23.1)

Cryptosporidium and *Giardia* occurrence values are presented as positive events and as percentage.

[a] River Raisin watershed sites upstream of CSO outfalls

[b] River Raisin watershed sites downstream of CSO outfalls

[c] All River Raisin watershed sites

[d] Rain dependent sampling

Figure 4.1 and Table 4.2 show the geometric mean of parasite concentrations at the monitored sites. While the parasites were detected frequently, the concentrations were consistently low except at MARB. When MARB samples were excluded, the geometric mean *Cryptosporidium* concentrations found in the River Raisin sites were higher than those in the Grand River watershed. There was no consistent pattern to *Giardia* occurrence between the River Raisin and Grand River sites, however. The parasite concentrations from two CSO discharge events (sampled directly from the combined sewer weir at the Ionia and Stevens intersection and the #2 Goodrich site) were highly variable (Table 4.3). *Cryptosporidium* concentrations from these events were below detectable levels and *Giardia* ranged from a low concentration similar to the Grand River and River Raisin samples on the 2/16/2006 event, to a concentration that fell between that of the Grand and River Raisin concentration and the MARB concentration during the 7/18/2006 event.

Of the routinely monitored sites, MARB samples had the highest geometric mean concentrations for both *Cryptosporidium* and *Giardia,* 6 oocysts L^{-1} and 1350 cysts L^{-1}, respectively. The next highest geometric mean parasite concentration detected, the *Cryptosporidium* concentration found in the River Raisin sites upstream from CSO outfalls (0.615 oocysts L^{-1}), was approximately 10 fold lower than the MARB *Cryptosporidium* concentration. After MARB, in decreasing order of *Giardia* concentration were Riverside Park (0.29 cysts L^{-1}), River Raisin sites (0.2 cysts L^{-1}, all site combinations), and Deer Creek Park and North Beach Park (0.1 cysts L^{-1}, each). Within the Grand River watershed, the highest geometric mean for *Cryptosporidium*

concentrations was found in Deer Creek Park (0.27 oocysts L^{-1}), while the highest geometric mean for *Giardia* was detected in water samples from Riverside Park (0.29 cysts L^{-1}; range 0.1 to 0.452). For *Cryptosporidium*, the combined Deer Creek Park and Riverside Park geometric mean (0.25 cysts L^{-1}) was less than the overall geometric mean from all River Raisin watershed sites (615 cysts L^{-1}). However, the geometric mean for *Giardia* from the combined Deer Creek Park and Riverside Park samples (0.20 cysts L^{-1}) was similar to the overall geometric mean from all the River Raisin watershed sites (0.2 cysts L^{-1}) . For all river sites, the geometric mean value was below 1 (oo)cyst L^{-1}, although the *Cryptosporidium* detection in the River Raisin watershed was variable, producing an upper 95% confidence limit of 17.2 oocysts L^{-1} and 10.5 oocysts L^{-1} for River Raisin sites upstream of CSO outfalls and all the River Raisin watershed sites, respectively. This variability is largely due to the Rice Lake Drain sampling location in which *Cryptosporidium* concentrations ranged from 0.3 to 599 oocysts L^{-1}.

Figure 4.1 Geometric Mean of Parasite Occurrence Versus Site

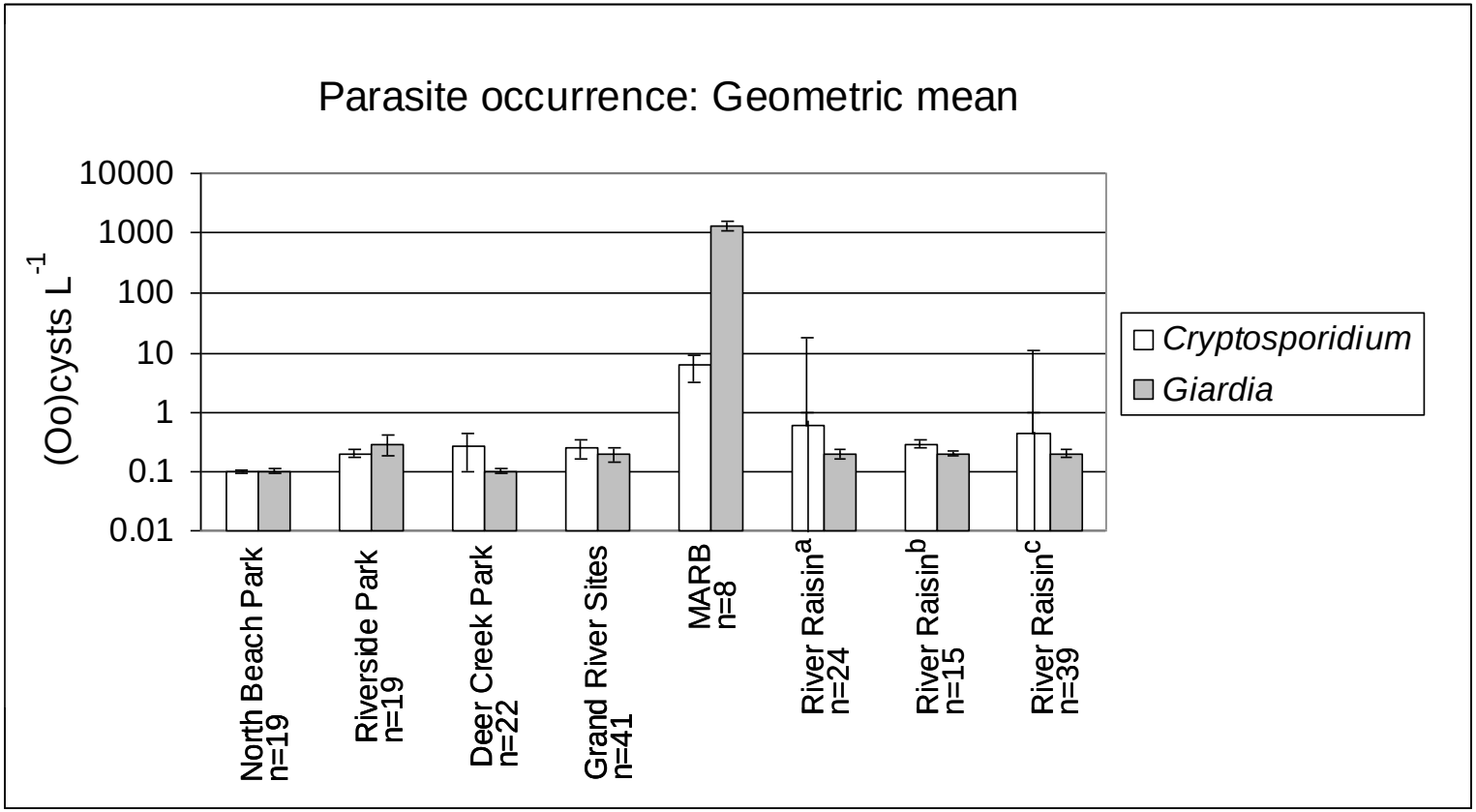


Figure 4.2 Geometric mean of parasite occurrence versus site. Error bars indicate the 95% confidence interval. Detection limits were used to calculate concentrations in samples where parasites were not detected.

[a] River Raisin watershed sites upstream of CSO outfalls

[b] River Raisin watershed sites downstream of CSO outfalls

[c] All River Raisin watershed sites

Table 4.2 Geometric Mean of Parasite Occurrence by Sampling Location

Watershed	Sampling Location	Number of Sampling Events	Geometric Mean ± 95% Confidence Interval	
			Cryptosporidium (oocysts L^{-1})	*Giardia* (cysts L^{-1})
Grand River watershed	North Beach Park	19	0.1 ± 0.008	0.1 ± 0.01
	Riverside Park	19	0.2 ± 0.03	0.29 ± 0.11
	Deer Creek Park	22	0.27 ± 0.17	0.1 ± 0.02
	Grand River Sites	41	0.25 ± 0.091	0.20 ±0.055
	MARB	8	6 ± 3	1350 ± 274
River Raisin watershed	River Raisin[a]	24	0.615 ± 17.2	0.2 ± 0.04
	River Raisin[b]	15	0.3 ± 0.04	0.2 ± 0.02
	River Raisin[c]	39	0.440 ± 10.5	0.2 ± 0.03

Table 4.2 Geometric mean of parasite occurrence by sampling location. Detection limits were used to calculate concentrations in samples where parasites were not detected.

[a] River Raisin watershed sites upstream of CSO outfalls

[b] River Raisin watershed sites downstream of CSO outfalls

[c] All River Raisin watershed sites

Table 4.3. Parasite Concentrations in CSO Samples

Site	Date	*Cryptosporidium* (oocyst L^{-1})	*Giardia* (cyst L^{-1})
Ionia and Stevens	2/16/2006	<0.09	0.09
Ionia and Stevens	7/18/2006	<1	184
#2 Goodrich	7/18/2006	<1.0	102

4.2 Recovery Efficiency Results

For North Beach Park matrix spike samples, the mean percentage and standard deviation of *Cryptosporidium* and *Giardia* recovery was 72.2 ± 2.1% and 67.3± 28.8%, respectively (Table 3.1). For Deer Creek Park matrix spike samples, the mean percentage and standard deviation of *Cryptosporidium* and *Giardia* recovery was 66.2 ±16.9% and 42.5± 25.2%, respectively. For the single Riverside Park matrix spike sample, *Cryptosporidium* and *Giardia* recovery was 73.0% and 43.4%, respectively. All the Grand River watershed matrix spike and OPR sample recovery percentages were acceptable according to the criteria (OPR criteria) used to assess *Cryptosporidium* and *Giaridia* recovery in laboratory reagent water (Table 4.4 and Table 4.5).

Recoveries dropped for MARB matrix spike samples, the percentage of *Cryptosporidium* recovery ranged from 11% to 73% with a mean recovery of 33.7% and a standard deviation of 22.6 (Table 4.4). The percentage of *Giardia* recovery for MARB samples, ranged from 9.5% to 79.5% with a mean recovery of 42.1% and a standard deviation of 24.9. Two of the seeded MARB samples had recovery efficiencies below

the criteria used to assess recovery in laboratory reagent water (OPR criteria). These low recoveries are most likely due to matrix interference as the recoveries from OPR samples were acceptable (Table 4.5).

For all River Raisin watershed matrix spike samples, the mean percentage and standard deviation of *Cryptosporidium* and *Giardia* recovery was 49.6± 24.2% and 9.7± 10.6%, respectively (Table 4.4). The percentage of *Cryptosporidium* recovery ranged from 16.5% to 90.4%. All *Cryptosporidium* recovery percentages in matrix spike samples were acceptable according to the criteria used to assess *Cryptosporidium* recovery in laboratory reagent water (OPR criteria). The percentage of *Giardia* recovery ranged from 0% to 34%. Only two of the River Raisin watershed matrix spike samples met the criteria for acceptable *Giardia* recovery in laboratory reagent water (OPR criteria). The recoveries from all *Giardia* OPR samples were acceptable (Table 4.5), indicating probable matrix interference with *Giardia* recovery in the matrix spike samples.

Table 4.4. *Cryptosporidium* and *Giardia* Recovery in Matrix Spike Samples

Location	*Cryptosporidium* Recovery %		*Giardia* Recovery %	
	Individual recovery %	Mean ± standard deviation	Individual recovery %	Mean ± standard deviation
River Raisin Watershed All Sites		49.6 ± 24.2		9.7 ± 10.6
Upstream of CSO	39.1	49.4 ± 27.7	3.8	4.7 ± 5.1
	36.5		12.4	
	26.15		4.8	
	73.45		13.5	
	62.1		0	
	16.5		2.75	
	90.4		5.3	
	80.3		0	
	20.0		0	
Downstream of CSO	54.4	50.0 ± 19.2	34	18.7 ± 12.4
	60.1		8.17	
	24.2		30.3	
	37.8		9.1	
	73.3		11.8	
Grand River Watershed Recreational Sites		69.9 ± 9.20		52.6 ± 23.3
North Beach Park	73.7	72.2 ± 2.1	46.9	67.3± 28.8
	70.7		87.6	
Deer Creek Park	78.1	66.2 ±16.9	24.7	42.5± 25.2
	54.2		60.3	
Riverside Park	73.0	NA	43.4	NA
MARB	11.0	33.7 ± 22.6	9.50	42.1± 24.9
	14.0[a]		46.0[a]	
	53.0		40.5	
	73.0		12.0	
	32.0		59.0	
	35.0		48.0	
	18.0		80.0	

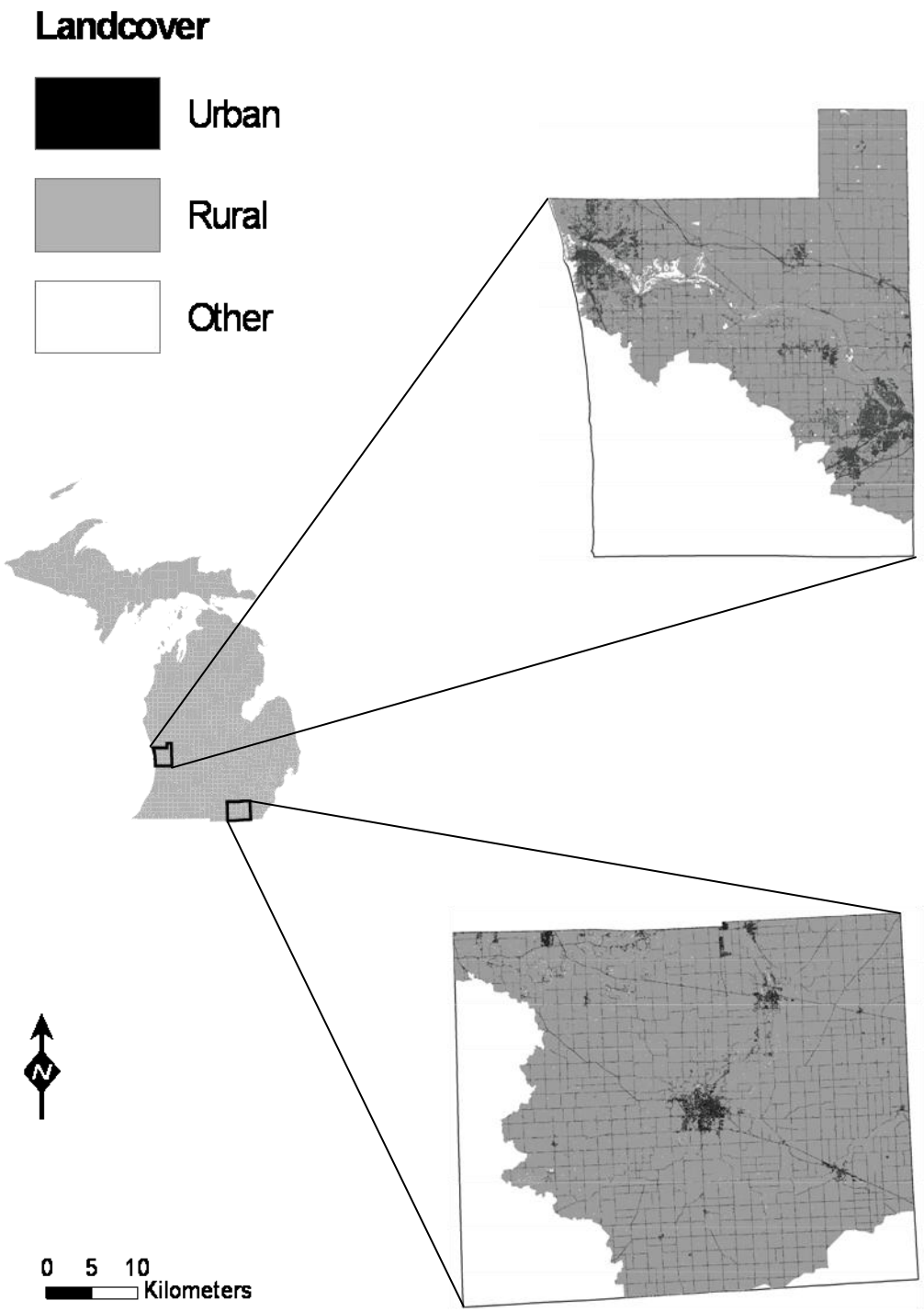
[a] Matrix spike duplicate

Table 4.5. *Cryptosporidium* and *Giardia* Recovery in Ongoing Precision and Recovery (OPR) Samples

Location	*Cryptosporidium* Recovery %		*Giardia* Recovery %	
	Individual recovery %	Mean ± standard deviation	Individual recovery %	Mean ± standard deviation
River Raisin Watershed All Sites	49	41.3 ± 7.8	55	42.1 ± 10.5
	30.6		35.4	
	44		46.05	
	41.6		31.8	
Grand River Watershed Recreational Sites	56	55.2 ± 11.7	44	43.9 ± 15.7
	51		22.8	
	80.8		34	
	40.4		29	
	40.4		32	
	48.5		33.7	
	56.6		46.9	
	37.4		42.9	
	49.5		44.9	
	64.6		41.8	
	63.6		23.5	
	37.4		34.7	
	63.6		16.3	
	68.7		51.0	
	69.7		31.6	
	60.6		72.4	
	60.6		65.3	
	50.5		41.8	
	70.7		73.5	
	45.5		51	
	47.5		62.2	
	46.1		56.1	
	60.6		59.2	
MARB	44.9	47.3 ± 13.6	55.6	67.0 ± 17.3
	36.7		97.0	
	37.0		55.5	
	48.0		62.0	
	70.0		65.0	

Note: Acceptance criteria for *Cryptosporidium* and *Giardia* recovery were 13-111% and 15-118%, respectively.

Figure 4.2 Watershed Land Cover in Ottawa and Lenawee County, MI



4.2 Statistical Analysis Results: Grand River Watershed Recreational Sites

Comparison of parasite occurrence datasets from North Beach Park, Riverside Park, and Deer Creek Park identified significant differences in these datasets. *Cryptosporidium* occurrence at North Beach Park was significantly less than *Cryptosporidium* occurrence at either Deer Creek Park or Riverside Park ($p<0.05$). North Beach Park *Giardia* occurrence was significantly less than *Giardia* occurrence at Riverside Park ($p<0.05$). No other meaningful comparison of parasite occurrence was found to be statistically significant.

Comparison of parasite detection limit datasets from North Beach Park, Riverside Park, and Deer Creek Park found no significant difference between these sites ($p=0.700$).

No significant difference between *Giardia* and *Cryptosporidium* recovery efficiency in matrix spike samples was found ($p=0.160$), however the power of this test was low (power = 0.173). Therefore, there could be significant differences between these recovery efficiencies that this test cannot detect.

When recovery efficiency in OPR samples was examined, significantly less *Giardia* was recovered compared to *Cryptosporidium* ($p=0.008$)

4.3 Statistical Analysis Results: Grand River Watershed Recreational Sites and MARB

There was a statistically significant difference between parasite occurrence from MARB, North Beach Park, Riverside Park, and Deer Creek Park datasets ($p<0.001$). *Cryptosporidium* occurrence in MARB was significantly greater than *Cryptosporidium* occurrence in both Deer Creek Park ($p<0.05$) and North Beach Park ($p<0.05$) datasets.

MARB *Giardia* occurrence was significantly greater than *Giardia* occurrence in both Deer Creek Park ($p<0.05$) and North Beach Park ($p<0.05$) datasets. No other meaningful comparison of parasite occurrence was found to be statistically significant.

Detection limit datasets from MARB, North Beach Park, Riverside Park, and Deer Creek Park were significantly different ($p<0.001$). The detection limits from MARB were significantly higher than those of the North Beach Park, Riverside Park, and Deer Creek Park detection limits.

No significant differences between recovery efficiencies in matrix spike samples from MARB and the combined North Beach Park, Riverside Park, and Deer Creek Park were found ($p=0.055$), however the power of this test was low (power = 0.429). Therefore, there could be significant differences between these recovery efficiencies that this test cannot detect.

There were significant differences between parasite recovery efficiency in OPR samples from MARB and the combined North Beach Park, Riverside Park, and Deer Creek Park recovery efficiency dataset ($p=0.005$). Significantly fewer *Giardia* than *Cryptosporidium* were recovered in the combined North Beach Park, Riverside Park, and Deer Creek Park dataset. Significantly less *Giardia* were recovered in MARB dataset than in the combined North Beach Park, Riverside Park, and Deer Creek Park dataset.

4.4 Statistical Analysis Results: River Raisin Watershed

No significant difference ($p=0.282$) in parasite occurrence was found between sites upstream of CSO outfalls and sites downstream of CSO outfalls. No significant differences were found when detection limits between upstream and downstream sites were compared ($p=0.444$). No significant difference between *Cryptosporidium* and

Giardia concentration was found (p=0.082) when parasite occurrence datasets from upstream and downstream sites were pooled.

There were significant differences in parasite recovery efficiency in matrix spike samples between sites upstream of CSO outfalls and sites downstream of CSO outfalls ($p<0.001$). The recovery of *Cryptosporidium* was significantly greater than recovery of *Giardia* in sites upstream from CSO outfalls ($p<0.05$). No other meaningful comparison of parasite recovery in matrix spike samples was found to be statistically significant. When parasite recovery efficiency in matrix spike sample datasets from upstream and downstream sites were pooled and compared *Cryptosporidium* recovery was significantly greater than *Giardia* recovery ($p<0.001$).

No significant differences were found when recovery efficiencies of *Giardia* and *Cryptosporidium* in OPR samples were compared (p=0.911), however the power of this test was low (power = 0.05). Therefore, there could be significant differences between these recovery efficiencies that this test cannot detect.

4.5 Statistical Analysis Results: River Raisin vs Grand River Watershed

As no significant differences were found between parasite occurrence between Deer Creek Park and Riverside Park, these datasets were pooled for comparison with the River Raisin watershed sites. As no significant differences were found between upstream and downstream sites in the River Raisin watershed, these datasets were pooled for comparison with the Grand River watershed sites.

There were significant differences between parasite occurrence in datasets from North Beach Park, from the combined Deer Creek Park-Riverside Park dataset, and from

the total River Raisin dataset ($p<0.001$). North Beach Park *Cryptosporidium* occurrence was significantly less than *Cryptosporidium* occurrence in the combined Deer Creek Park-Riverside Park dataset, and significantly less than *Cryptosporidium* occurrence in the total River Raisin dataset. North Beach Park *Giardia* occurrence was also significantly less than *Giardia* occurrence in the Deer Creek Park-Riverside Park dataset, and significantly less than *Giardia* occurrence in the total River Raisin dataset.

There were significant differences between detection limits in datasets from North Beach Park, from the combined Deer Creek Park and Riverside Park dataset, and from the total River Raisin dataset ($p<0.001$). The detection limits in the total River Raisin dataset were significantly higher than detection limits in the North Beach Park dataset and significantly higher than detection limits in the combined Deer Creek Park-Riverside Park dataset.

There were significant differences between parasite recovery efficiency in matrix spike samples from the total Grand River dataset (North Beach Park, Deer Creek Park, and Riverside Park) and from the total River Raisin dataset ($p<0.001$). *Cryptosporidium* recovery in matrix spike samples was significantly higher than *Giardia* recovery in the total River Raisin dataset. *Giardia* recovery in the total Grand River dataset was significantly higher than *Giardia* recovery in the total River Raisin dataset.

There were significant differences between parasite recovery efficiency in OPR samples from the total Grand River dataset (North Beach Park, Deer Creek Park, and Riverside Park) and from the total River Raisin dataset ($p<0.001$). *Cryptosporidium* recovery in OPR samples was significantly higher than *Giardia* recovery in the total Grand River dataset.

4.6 Discussion

4.6.1 River Raisin Watershed

During the time period that sampling was conducted in the River Raisin watershed, four CSO discharges occurred (35). However, the lack of significant differences between sites upstream and downstream of CSO outfalls suggests that CSO discharges were not significantly impacting parasite occurrence in the watershed during the sampling period.

Matrix spike recovery efficiency was significantly greater for *Cryptosporidium* than *Giardia* in upstream sites, which may explain why geometric mean occurrence of *Cryptosporidium* in sites upstream of CSO outfalls was higher (although not significantly) than in sites downstream of COS outfalls.

Pooled upstream and downstream sites did not have significantly different concentrations of *Cryptosporidium* and *Giardia*, however the pooled matrix spike recovery efficiencies were significantly different, indicating a potential bias against *Giardia*. As no significant differences in detection limits or in recovery efficiency of OPR samples was found, these factors are not introducing bias into intra-watershed comparisons.

4.6.2 Grand River Watershed

The Grand River watershed recreational sampling sites included both river and lake sampling locations. Comparisons of parasite occurrence between the river and lake locations could be affected by dilution effects, as the *Cryptosporidium* occurrence at North Beach Park was significantly lower than *Cryptosporidium* occurrence at either Deer Creek Park or Riverside Park. However, *Giardia* occurrence was not significantly

different between North Beach Park and Deer Creek Park, while it was significantly higher at Riverside Park than at North Beach Park. This pattern suggests that there may be an additional source of *Giardia* contamination between Deer Creek Park and Riverside Park. As no significant differences in detection limits or in recovery efficiency of matrix spike samples was found, these factors are not considered to be introducing bias into intra-watershed comparisons. However, significantly less *Giardia* than *Cryptosporidium* was recovered in OPR samples, indicating a method bias against detection of *Giardia.*

The Grand River watershed study area also contains the Market Avenue Retention Basin (MARB) and other CSO outfalls that may impact the recreational water quality of the Grand River recreation areas during CSO discharges. *Cryptosporidium* occurrence in MARB was not significantly different from *Cryptosporidium* occurrence in Riverside Park and *Giardia* occurrence in MARB was not significantly different from *Giardia* occurrence in Riverside Park. This information, combined with the indication of a contamination source located between Deer Creek Park and Riverside Park suggests that this suspected source is a frequently discharging CSO outfall, a sewage treatment plant discharge with low physical removal of *Cryptosporidium* and *Giardia*, or an illicit discharge of untreated wastewater.

The detection limits in MARB samples were significantly higher than those of the recreational Grand River sites. Since *Giardia* was detected in 100% of MARB samples, this detection limit difference is not a problem when comparing *Giardia* occurrence values. However, the limit of detection may introduce bias when comparing *Cryptosporidium* values as the geometric mean is calculated using the detection limit

when no parasite is detected. This difference could produce artificially higher *Cryptosporidium* values in MARB samples.

Additionally, there were significant differences between parasite recovery efficiency in OPR samples from MARB and the combined North Beach Park, Riverside Park, and Deer Creek Park recovery efficiency dataset. This could introduce bias against *Giardia* occurrence values in the Grand River recreational sites as they are compared to MARB occurrence values. As no significant differences in recovery efficiency of matrix spike samples was found, this factor is not considered to be introducing bias into intra-watershed comparisons.

4.6.3 River Raisin vs Grand River Watershed

When the inter-watershed comparison of the River Raisin and Grand River watershed was conducted, parasite occurrence from North Beach Park, the combined Deer Creek Park- Riverside Park dataset, and the total River Raisin dataset were compared. Both *Cryptosporidium* and *Giardia* occurrence in North Beach Park were significantly different from *Cryptosporidium* and *Giardia* occurrence in the River Raisin sites. Both *Cryptosporidium* and *Giardia* occurrence in North Beach Park were significantly different from *Cryptosporidium* and *Giardia* occurrence in the combined Deer Creek Park- Riverside Park dataset. However, no significant differences between the combined Deer Creek Park- Riverside Park dataset, and the total River Raisin dataset were found. As previously discussed, the difference between parasite occurrence at North Beach Park and the combined Deer Creek Park- Riverside Park dataset may be due to dilution effects. To address the question of whether parasite occurrence patterns in the watersheds were associated with differences in land cover type, the most

informative comparison is therefore that between the combined Deer Creek Park-Riverside Park dataset and the total River Raisin dataset. No significant difference in parasite occurrence patterns between the combined Deer Creek Park- Riverside Park dataset and the total River Raisin dataset was detected; therefore we fail to reject the null hypothesis that the parasite occurrence in both watersheds is the same.

Potential biases exist that may affect this conclusion due to differences in detection limits and matrix spike recoveries. The detection limits in the total River Raisin dataset are significantly higher than detection limits in the combined Deer Creek Park-Riverside Park dataset which could artificially increase the parasite occurrence values in the River Raisin dataset. In matrix spike samples, *Giardia* recovery in the total Grand River dataset was significantly higher than *Giardia* recovery in the total River Raisin dataset, which could produce artificially low *Giardia* occurrence values in the River Raisin dataset. Significant differences were also found between OPR recovery efficiencies, however these differences were between *Cryptosporidium* and *Giardia* recovery in an intra-watershed comparison of the Grand River watershed and do not affect the inter-watershed comparison of the Grand River and River Raisin watershed.

Although there were differences in the percentage of urban and rural land cover between the Grand River watershed in Ottawa County and the River Raisin watershed in Lenawee County, the parasite occurrence in both watersheds was statistically similar. Parasite occurrence was not shown to vary with differences in land cover, unlike parasitic disease incidence (Chapter 3). The difference in patterns between parasite occurrence and parasitic disease could be a result of scale. Parasitic disease patterns were examined by ZIP code which is a finer scale than the watershed and may resolve

patterns better. At coarser scales, to detect statistically significant differences in parasite occurrence, the differences in the land cover percentages may need to be more extreme than those found between the Grand River watershed in Ottawa County and the River Raisin watershed in Lenawee County.

The third caveat in the inter-watershed comparison involves potential biases due to differences in weather patterns between the watersheds. The original experimental design for data collection in the River Raisin watershed was based on sampling after rainfall to examine parasite concentrations after storm events. The River Raisin dataset is therefore less representative of the yearly parasite concentration in surface water than the Grand River dataset, which was collected at regular intervals over a longer time period. The differences in seasonal weather patterns and differences between yearly weather patterns may introduce biases into the comparison of parasite concentrations between the River Raisin and Grand River watersheds, although this is beyond the scope of the study.

4.7 Conclusion

Although differences in the pattern of disease occurrence between urban and rural areas were found in Chapter 3, no evidence was found in the post-hoc testing of parasite occurrence between the Grand River and River Raisin watersheds that suggests a difference in parasite occurrence between urban and rural areas. Caveats to this conclusion include differences in detection limits and matrix spike recoveries. The data used to drawn this conclusion may also be affected by weather patterns during sampling, as the Grand River watershed and River Raisin watershed were sampled in different years and over different seasons, although this was beyond the scope of the study.

Chapter 5: Recreational Health Risks of Combined Sewer Overflows

Combined sewers are sewers that are connected to storm drains. During dry weather, they carry sewage to the sewage treatment plant. During rain events, or during seasonal thaws of ice and snow, water from storm events enters these sewers as well. Combined sewers are designed with overflow points (or outfalls) where combined sewage can be discharged, untreated, into the environment. CSO discharges occur when either the capacity of the sewage treatment plant to treat incoming sewage or when the capacity of the sewage pipe network to transport sewage is exceeded. Discharges of untreated sewage into recreational waters can lead to pathogen exposure, infection, and illness in the human population. In this study, health risks caused by exposure to *Cryptosporidium* and *Giardia* in children under the age of 16 swimming in recreational areas that receive inputs from CSOs were investigated.

5.1 Risk Assessment Results

Table 5.1 contains the distributions and distribution parameters used in the Monte Carlo analysis of the health risks. The predicted concentration of *Cryptosporidium* and *Giardia* at the CSO discharge point (C_{RIVER}) is shown in Table 5.2.

Table 5.1. Risk assessment distribution parameters.

Exposure Factor	Distribution	Distribution Parameter	Unit	Source
Cryptosporidium concentration at recreational area[a]				
North Beach Park	Weibull	Location = 0.07519 Scale = 0.04541 Shape = 0.94238	organism · L^{-1}	This study
Deer Creek Park	Weibull	Location = 0.07874 Scale = 0.29437 Shape = 0.52793		
Riverside Park	Beta	Minimum = 0.09858 Maximum = 0.84573 Alpha = 0.37251 Beta = 1.1563		
Giardia concentration at recreational area[a]				
North Beach Park	Maximum Extreme	Likeliest = 0.10853 Scale = 0.03819	organism · L^{-1}	This study
Deer Creek Park	Pareto	Location = 0.07653 Shape = 1.5981		
Riverside Park	Lognormal	Mean = 0.41996 Standard Deviation = 0.44272		

Table 5.1 (cont'd).

Exposure Factor	Distribution	Distribution Parameter	Unit	Source
Cryptosporidium concentration in MARB influent[a]	Triangular	Minimum = 0.77 Likeliest = 13.33 Maximum = 30.56	organism $\cdot L^{-1}$	This study
Giardia concentration in MARB influent[a]	Beta	Minimum=478.15 Maximum=5,220.23 Alpha=0.50131 Beta=1.57452	organism $\cdot L^{-1}$	This study
Duration of CSO discharge (374 events)	Weibull	Location = -1.8981×10^{-72} Scale = 91.461 Shape = 0.52722	minute	This study
Volumetric flow rate of Grand River	Gamma	Location = 363.57 Scale = 1126.5 Shape = 1.2229	$L \cdot minute^{-1}$	This study

Table 5.1 (cont'd).

Exposure Factor	Distribution	Distribution Parameter	Unit	Source
Ingestion volume during swimming (I)	Triangular	Minimum = 0 Likeliest = 0.000833 Maximum = 0.00166	$L \cdot minute^{-1}$	(53)
Amount of time spent swimming (T_{SWIM})	Triangular	Minimum = 0 Likeliest = 137.6 Maximum = 181	$minute \cdot month^{-1}$	(53)
r Fraction of ingested *Cryptosporidium* that survive to initiate infection	Constant	0.00419	Unitless	(25)
r Fraction of ingested *Giardia* that survive to initiate infection	Constant	0.01982	Unitless	(45)
Morbidity Ratio : *Cryptosporidium*	Triangular	Minimum = 0.19 Likeliest = 0.39 Maximum = 0.62	Unitless	(25)
Morbidity Ratio: *Giardia*	Normal	Mean = 0.505 Standard Deviation = 0.1902	Unitless	(45)

Table 5.1 (cont'd).

Exposure Factor	Distribution	Distribution Parameter	Unit	Source
Volume of CSO discharge at outfall [b]				
Grand River	Beta	Minimum = -61738914.76 Maximum = 62078291.44 Alpha = 100 Beta = 100	L	(35), This study
Grand River via Coldbrook Drain	Beta	Minimum = -894803.55 Maximum = 902736.90 Alpha = 100 Beta = 100		
Grand River via Plaster Creek	Logistic	Mean = 5051.00 Scale = 97680.66		
Grand River via Market Avenue Retention Basin	Logistic	Mean = 5061.39 Scale = 696103.24		

[a] Based on empirical values with non-detects represented using the detection limit. Use produces the most conservative (ie most protective) estimate of risk

[b] Based on time interval from January 1, 2002 to December 31, 2006

Table 5.2. Predicted Concentration Distributions of *Cryptosporidium* and *Giardia* at the CSO Discharge Point (C_{RIVER}).

Summary statistics of distribution	*Cryptosporidium* (oocyst L^{-1})	*Giardia* (cyst L^{-1})
Mean	13.96	1,502.22
Median	13.66	1,048.94
Standard Deviation	6.03	1,127.78
Variance	36.34	1,271,881.27
Minimum	1.28	95.07
Maximum	30.18	5,097.08
Lower (5%) confidence limit	4.33	459.35
Upper (95%) confidence limit	24.93	3,915.89

The estimated health risks to children under the age of sixteen from exposure to the parasites *Cryptosporidium* and *Giardia* while swimming are shown in Table 5.3. The mean daily risks of *Cryptosporidium* infection at North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were 1.50×10^{-6}, 7.41×10^{-6}, 3.61×10^{-6}, and 1.78×10^{-4}, respectively. The mean daily risks of *Giardia* infections at North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were 1.58×10^{-6}, 2.48×10^{-6}, 5.18×10^{-6}, and 1.80×10^{-2}, respectively. The mean seasonal risks of *Cryptosporidium* infection at North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were 1.35×10^{-4}, 6.66×10^{-4}, 3.24×10^{-4}, and 1.58×10^{-2}, respectively. The mean seasonal risks of *Giardia* infection at

North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were 1.42×10^{-4}, 2.23×10^{-4}, 4.66×10^{-4}, and 6.31×10^{-1}, respectively.

The mean daily risks of cryptosporidiosis (illness) at North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were about 40% of the infectivity and were estimated at 6.02×10^{-7}, 2.98×10^{-6}, 1.43×10^{-6}, and 7.03×10^{-5}, respectively. The mean daily risks of giardiasis at North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were 7.90×10^{-7}, 1.29×10^{-6}, 2.62×10^{-6}, and 9.39×10^{-3}, respectively. Again seasonal risks went up by about 90 times. The mean seasonal risks of cryptosporidiosis at North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were 5.42×10^{-5}, 2.68×10^{-4}, 1.29×10^{-4}, and 6.24×10^{-3}, respectively. The mean seasonal risks of giardiasis at North Beach Park, Deer Creek Park, Riverside Park, and at the CSO point of discharge were 7.11×10^{-5}, 1.16×10^{-4}, 2.36×10^{-4}, and 3.25×10^{-1}, respectively. Additional summary statistics including the mean, median, standard deviation, variance, minimum, maximum, and upper and lower confidence limits of the risk distributions for daily and seasonal risk of infection and daily and seasonal risk of illness are shown in Table 5.3.

Giardia risks were significantly higher than *Cryptosporidium* risks at the CSO discharge site. ($P<0.001$, $\alpha = 0.05$). At the recreational sites the risks were not statistically different between *Giardia* and *Cryptosporidium* ($P \geq 0.694$, $\alpha = 0.05$). The daily risk of infection, daily risk of illness, seasonal risk of infection, and seasonal risk of illness associated with *Giardia* at the CSO discharge point were all significantly

higher than the respective risk associated with *Giardia* at the recreational sites ($P<0.001$, $\alpha = 0.05$). The seasonal risk of infection and seasonal risk of illness associated with *Cryptosporidium* at the CSO discharge point were both significantly higher than the respective risk associated with *Cryptosporidium* at the recreational sites ($P<0.001$, $\alpha = 0.05$).

Using the Monte Carlo analysis, the probability of meeting recreational water quality criteria for illness rates was computed and the results are shown in Table 5.4. These are based on the probability distributions produced from 10,000 trials. Table 5.4 shows the likelihood of meeting the EPA recommended freshwater recreational criteria of 8 illnesses in 1000 swimmers (0.8%).

Using the estimate of daily risk, the water quality of all recreational sites met the freshwater recreational criteria in 99.9% of cases for both cryptosporidiosis and giardiasis. The water quality at the point of the CSO discharge was also predicted to meet the recreational criteria for cryptosporidiosis in 99.9% of cases. However, the water quality at the point of CSO discharge was predicted to meet the freshwater recreational criteria for giardiasis in only 62.7% of cases.

Using the estimate of seasonal risk, the water quality of North Beach Park and Riverside Park met the recommended freshwater recreational criteria 99.9% of the time for both cryptosporidiosis and giardiasis. The water quality of Deer Creek Park met the freshwater recreational criteria 92.4% of the time for cryptosporidiosis and 99.9% of the time for giardiasis. The water quality at the point of CSO discharge met the recreational criteria in 73.0% of cases for cryptosporidiosis and in 0.46% of cases for giardiasis.

Table 5.3. Estimated Risk of Health Outcomes for Children Swimming at Recreational Sites and at Point of CSO Discharge in the Grand River Watershed

North Beach Park *Cryptosporidium*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	1.50×10^{-6}	1.35×10^{-4}	6.02×10^{-7}	5.42×10^{-5}
Median	1.29×10^{-6}	1.16×10^{-4}	4.85×10^{-7}	4.36×10^{-5}
Standard Deviation	1.06×10^{-6}	9.51×10^{-5}	4.67×10^{-7}	4.20×10^{-5}
Variance	1.12×10^{-12}	9.04×10^{-9}	2.18×10^{-13}	1.76×10^{-9}
Minimum	5.97×10^{-8}	5.38×10^{-6}	2.23×10^{-8}	2.01×10^{-6}
Maximum	8.49×10^{-6}	7.64×10^{-4}	3.80×10^{-6}	3.42×10^{-4}
Lower (5%) confidence limit	2.78×10^{-7}	2.50×10^{-5}	9.92×10^{-8}	8.93×10^{-6}
Upper (95%) confidence limit	3.52×10^{-6}	3.17×10^{-4}	1.50×10^{-6}	1.35×10^{-4}
North Beach Park *Giardia*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	1.58×10^{-6}	1.42×10^{-4}	7.90×10^{-7}	7.11×10^{-5}
Median	1.33×10^{-6}	1.19×10^{-4}	6.27×10^{-7}	5.64×10^{-5}
Standard Deviation	1.14×10^{-6}	1.02×10^{-4}	6.43×10^{-7}	5.79×10^{-5}
Variance	1.29×10^{-12}	1.05×10^{-8}	4.14×10^{-13}	3.35×10^{-9}
Minimum	2.26×10^{-9}	2.03×10^{-7}	1.78×10^{-9}	1.60×10^{-7}
Maximum	7.79×10^{-6}	7.00×10^{-4}	4.05×10^{-6}	3.65×10^{-4}
Lower (5%) confidence limit	2.60×10^{-7}	2.34×10^{-5}	1.01×10^{-7}	9.11×10^{-6}
Upper (95%) confidence limit	3.68×10^{-6}	3.31×10^{-4}	2.03×10^{-6}	1.83×10^{-4}

Table 5.3. (cont'd)

Deer Creek Park *Cryptosporidium*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	7.41×10^{-6}	6.66×10^{-4}	2.98×10^{-6}	2.68×10^{-4}
Median	2.31×10^{-6}	2.08×10^{-4}	9.13×10^{-7}	8.21×10^{-5}
Standard Deviation	1.58×10^{-5}	1.42×10^{-3}	6.40×10^{-6}	5.73×10^{-4}
Variance	2.51×10^{-10}	2.01×10^{-6}	4.10×10^{-11}	3.29×10^{-7}
Minimum	7.85×10^{-8}	7.06×10^{-6}	3.76×10^{-8}	3.39×10^{-6}
Maximum	2.18×10^{-4}	1.94×10^{-2}	7.56×10^{-5}	6.75×10^{-3}
Lower (5%) confidence limit	3.74×10^{-7}	3.37×10^{-5}	1.42×10^{-7}	1.28×10^{-5}
Upper (95%) confidence limit	3.23×10^{-5}	2.90×10^{-3}	1.31×10^{-5}	1.18×10^{-3}
Deer Creek Park *Giardia*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	2.48×10^{-6}	2.23×10^{-4}	1.29×10^{-6}	1.16×10^{-4}
Median	1.48×10^{-6}	1.33×10^{-4}	6.78×10^{-7}	6.10×10^{-5}
Standard Deviation	4.70×10^{-6}	4.22×10^{-4}	2.98×10^{-6}	2.68×10^{-4}
Variance	2.21×10^{-11}	1.78×10^{-7}	8.90×10^{-12}	7.17×10^{-8}
Minimum	2.20×10^{-8}	1.98×10^{-6}	4.96×10^{-9}	4.46×10^{-7}
Maximum	9.99×10^{-5}	8.95×10^{-3}	6.47×10^{-5}	5.80×10^{-3}
Lower (5%) confidence limit	2.84×10^{-7}	2.55×10^{-5}	1.11×10^{-7}	9.98×10^{-6}
Upper (95%) confidence limit	7.65×10^{-6}	6.89×10^{-4}	4.08×10^{-6}	3.67×10^{-4}

Table 5.3 (cont'd).

Riverside Park *Cryptosporidium*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	3.61×10^{-6}	3.24×10^{-4}	1.43×10^{-6}	1.29×10^{-4}
Median	2.43×10^{-6}	2.19×10^{-4}	9.37×10^{-7}	8.44×10^{-5}
Standard Deviation	3.51×10^{-6}	3.16×10^{-4}	1.45×10^{-6}	1.31×10^{-4}
Variance	1.23×10^{-11}	9.99×10^{-8}	2.11×10^{-12}	1.71×10^{-8}
Minimum	8.45×10^{-8}	7.60×10^{-6}	3.28×10^{-8}	2.95×10^{-6}
Maximum	2.08×10^{-5}	1.87×10^{-3}	9.47×10^{-6}	8.51×10^{-4}
Lower (5%) confidence limit	4.38×10^{-7}	3.94×10^{-5}	1.57×10^{-7}	1.42×10^{-5}
Upper (95%) confidence limit	1.06×10^{-5}	9.51×10^{-4}	4.31×10^{-6}	3.88×10^{-4}
Riverside Park *Giardia*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	5.18×10^{-6}	4.66×10^{-4}	2.62×10^{-6}	2.36×10^{-4}
Median	3.27×10^{-6}	2.94×10^{-4}	1.54×10^{-6}	1.39×10^{-4}
Standard Deviation	5.84×10^{-6}	5.25×10^{-4}	3.29×10^{-6}	2.96×10^{-4}
Variance	3.41×10^{-11}	2.76×10^{-7}	1.09×10^{-11}	8.77×10^{-8}
Minimum	3.31×10^{-8}	2.97×10^{-6}	1.78×10^{-8}	1.60×10^{-6}
Maximum	5.09×10^{-5}	4.57×10^{-3}	3.03×10^{-5}	2.72×10^{-3}
Lower (5%) confidence limit	4.05×10^{-7}	3.64×10^{-5}	1.64×10^{-7}	1.47×10^{-5}
Upper (95%) confidence limit	1.62×10^{-5}	1.46×10^{-3}	8.85×10^{-6}	7.96×10^{-4}

Table 5.3 (cont'd).

CSO Point of Discharge: *Cryptosporidium*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	1.78×10^{-4}	1.58×10^{-2}	7.03×10^{-5}	6.24×10^{-3}
Median	1.43×10^{-4}	1.27×10^{-2}	5.56×10^{-5}	4.97×10^{-3}
Standard Deviation	1.34×10^{-4}	1.18×10^{-2}	5.61×10^{-5}	4.93×10^{-3}
Variance	1.80×10^{-8}	1.40×10^{-4}	3.14×10^{-9}	2.44×10^{-5}
Minimum	3.97×10^{-7}	3.57×10^{-5}	1.62×10^{-7}	1.46×10^{-5}
Maximum	7.24×10^{-4}	6.31×10^{-2}	3.70×10^{-4}	3.23×10^{-2}
Lower (5%) confidence limit	2.34×10^{-5}	2.10×10^{-3}	9.74×10^{-6}	8.76×10^{-4}
Upper (95%) confidence limit	4.45×10^{-4}	3.93×10^{-2}	1.85×10^{-4}	1.63×10^{-2}
CSO Point of Discharge: *Giardia*				
Summary statistics of risk distribution	Daily Risk of Infection	Seasonal Risk of Infection	Daily Risk of Illness	Seasonal Risk of Illness
Mean	1.80×10^{-2}	6.31×10^{-1}	9.39×10^{-3}	3.25×10^{-1}
Median	1.18×10^{-2}	6.57×10^{-1}	5.60×10^{-3}	3.05×10^{-1}
Standard Deviation	1.82×10^{-2}	2.70×10^{-1}	1.09×10^{-2}	1.95×10^{-1}
Variance	3.33×10^{-4}	7.28×10^{-2}	1.19×10^{-4}	3.80×10^{-2}
Minimum	2.84×10^{-5}	2.56×10^{-3}	2.07×10^{-5}	1.86×10^{-3}
Maximum	1.26×10^{-1}	1.00×10^{0}	8.82×10^{-2}	9.90×10^{-1}
Lower (5%) confidence limit	1.91×10^{-3}	1.58×10^{-1}	7.43×10^{-4}	5.65×10^{-2}
Upper (95%) confidence limit	5.48×10^{-2}	9.94×10^{-1}	3.16×10^{-2}	6.87×10^{-1}

Note: Non-detects in the dataset assigned the value of the detection limit for simulation purposes

Table 5.4. Likelihood of Meeting Recreational Criteria of 8 illnesses in 1,000 swimmers (0.8%).

Site	Daily Risk	Seasonal Risk
North Beach Park *Cryptosporidium*	>99.9	>99.9
North Beach Park *Giardia*	>99.9	>99.9
Riverside Park *Cryptosporidium*	>99.9	>99.9
Riverside Park *Giardia*	>99.9	>99.9
Deer Creek Park *Cryptosporidium*	>99.9	92.4
Deer Creek Park *Giardia*	>99.9	>99.9
CSO discharge point *Cryptosporidium*	>99.9	73.0
CSO discharge point *Giardia*	62.7	0.460

Note: Non-detects in the dataset assigned the value of the detection limit for simulation purposes

5.2 Discussion

It has long been known that untreated sewage can contribute numerous enteric pathogens at fairly significant concentrations to receiving waters and that this poses a risk to recreational users if they are exposed to these discharges. EPA has estimated that between 3,448 to 5,576 cases of illness annually at state recognized beaches in the United States and U.S. territories are attributable to recreational exposure to water contaminated by combined sewer overflows and sanitary sewer overflows (55). Epidemiological studies performed at Great Lake beaches have shown increased illness rates in children associated with sewage impacted beaches (56). However, health surveillance as a passive system is unable to detect all the potential illnesses, particularly as cases may intermittently occur over long time periods. Therefore, an alternative approach has been to use water quality monitoring and quantitative risk assessment methods to examine potential public health impacts and risk reduction strategies.

Great expenditures are being made to control combined sewer overflows to the Grand River in Michigan. The Grand River discharges to Lake Michigan and has the potential to affect a number of recreational beaches, including North Beach. This study was undertaken to address parasite contamination of the Grand River associated with CSOs and to estimate the potential risk to swimmers. Previous studies on wastewater have shown varying concentrations of *Cryptosporidium* oocysts (averaging about 1 oocyst L^{-1}) and higher concentrations of *Giardia* ranging from 10 to 100 cysts L^{-1} (33, 44). Overall, in this investigation low concentrations of *Cryptosporidium* were found in samples of combined stormwater and sewage from the Market Avenue Retention Basin. This may be due to low levels of infection in the human population in the sanitary sewer service area or dilution effects of stormwater. Concentrations of *Giardia* were much

higher, suggesting that *Giardia* infection is more prevalent in the area than *Cryptosporidium*. Detection of higher *Giardia* concentrations than *Cryptosporidium* concentrations agrees with the findings of other sewage and CSO studies (2, 55).

The MARB system was designed to capture the majority of combined stormwater and sewage for eventual return to the sewage treatment plant. This study demonstrates that the river is currently being protected from loading of billions of (oo)cysts during most storm events for which the majority of the CSO flows are captured by the MARB. However, using highly conservative estimates of risk, the remaining flows of combined sewage and storm water can still produce risk levels of cryptosporidiosis and giardiasis that exceed the EPA recommended freshwater recreational water quality criteria for illness at the point of discharge. Due to the higher concentrations of *Giardia* found in the combined sewage/stormwater, larger fraction that survives to initiate infection, and larger morbidity ratio, exceeding the recommended criteria at the point of CSO discharge is more likely for *Giardia* than for *Cryptosporidium*. Multiple day exposure to the water at the point of discharge is required to produce a cryptosporidiosis risk level that exceeds the water quality criteria. The daily risk of giardiasis at the CSO discharge point is predicted to exceed the recommended freshwater recreational water quality criteria for illness approximately 37% of the time. These risk estimates are highly conservative due to the assumption that all detected (oo)cysts are viable and capable of initiating infection in humans.

Although no public recreational facilities are located at the point of CSO discharge to the Grand River, a hypothetical swimming scenario at this point represents moderate daily health risk to a swimmer. Using conservative risk estimates developed in this

study, multiple day recreation at the discharge point over the recreation season represents a high health risk to a swimmer. At the point of CSO discharge, the river is of sufficient size to allow boating, water skiing, and the use of personal watercraft. Therefore, recreational exposure is possible even in the absence of public recreational facilities at the point of discharge. The elevated *Giardia* concentrations at the #2 Goodrich sampling location suggests that parasite concentrations released in CSO discharge may increase health risks at locations greater that 1.6 km from the CSO discharge point.

Parasite concentrations from the Market Avenue Retention Basin may overestimate the concentration of parasites entering the river. Comparisons between the predicted concentrations at the CSO discharge points (C_{RIVER}, Equation 5) and the concentrations detected at during CSO discharge show the predicted concentrations are higher than the detected parasite concentrations. This discrepancy may be due to different ratios of sewage and stormwater that may be discharged at an outfall location compared to that entering the retention basin. These ratios are likely to be affected by the amount of precipitation entering the storm drain in the area of the outfall and the number of sewer connections served by that line. The influent into the retention basin may represent a higher number of sanitary service connections than the CSO outfalls. Donovan *et al* found *Giardia* concentrations at 1,860 cysts per liter when sampling from a New Jersey CSO discharge and 798 cysts per liter 10 feet downstream of the CSO discharge (19). These *Giardia* values fall within the 95% confidence interval predicted at the CSO discharge point in the current study (459.35 to 3,915.89 cyst L^{-1}, Table 5.2). No *Cryptosporidium* was detected by Donovan *et al* at either location (19). A study of

CSO discharge from two CSO events at three CSO outfall sites in Atlanta, GA and Louisville, KY found that geometric means ranged from 59 to 100 *Cryptosporidium* oocysts per 100L and 200 to 30,000 *Giardia* cysts per 100L (2). These ranges are below the 95% confidence interval predicted in the current study for both *Cryptosporidium* (4.33 to 24.93 oocyst L^{-1}, Table 5.2) and *Giardia* (459.35 to 3,915.89 cyst L^{-1}, Table 5.2). A study of parasite concentration in CSO discharge by Gibson *et al* (24) reported geometric mean and median values of *Giardia* that fell below the 95% confidence interval of predicted concentration at the CSO discharge point in the present study. Gibson *et al* (24) also reported arithmetic mean, geometric mean, and median values of *Cryptosporidium* that exceed 95% confidence interval of predicted concentration at the CSO discharge point. States *et al* (48) reported geometric mean values for *Giardia* that fell below the 95% confidence interval and geometric mean values for *Cryptosporidium* that fell within the 95% confidence interval predicted at the CSO discharge point. When comparing parasite concentrations in CSO discharges from different geographic areas it is possible that value discrepancies may also result from different levels of infection in the populations.

In this study, the low health risks associated with *Cryptosporidium and Giardia* at the recreational areas suggest that the impacts of CSO discharges at these sites are being reduced by the retention basin, dilution, and other environmental transport processes. *Cryptosporidium* concentrations in the retention basin samples were about 20 times higher than *Cryptosporidium* concentrations at Deer Creek Park and Riverside Park downstream from CSO discharge points, but *Giardia* cyst concentrations were approximately 2000 times greater in MARB samples than in the river. The concentration

of (oo)cysts associated with routine sewage effluent discharges, non-point sources and CSOs may be reduced in the water column via dilution and sedimentation. The Grand River's currents, mixing zones, and flow rates may have an impact on the concentrations of parasites found downstream of the CSO discharge point and may influence the likelihood that parasites will be carried downstream in the water column versus sedimentation. The factors of sedimentation versus dilution by tributary streamflow are difficult to separate due to the presence of only one stream flow gage in this area of the Lower Grand River Watershed. Sedimentation may remove parasites from the water column and either temporarily or permanently sequester them in the sediments. Tributary streams add additional volumes to the river, potentially diluting contaminant concentrations further. Both of these processes may reduce the exposure of an individual to these pathogens, with an accompanying decrease in risk of infection or illness.

5.3 Conclusion

Despite the occasional untreated sewage discharges, recreational use of the Grand River met the EPA recommended freshwater recreational criteria of 8 illnesses per 1000 swimmers at least 92.4% of the time for both cryptosporidiosis and giardiasis at the three recreational sites examined. Reporting of CSO discharges can provide further protection of public health, especially when reporting is done in real time. This study demonstrates the value of monitoring and a quantitative microbial risk approach for examining specific pathogens and recreational waters.

APPENDICES

APPENDIX A: Statistical analysis of study area demographics

Table A1 Demographics: Population

Demographic data: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Rural	A	40	0	6875.000	2165.500	20032.500
Urban	B	8	0	36821.500	32915.500	40132.000
Mann-Whitney U Statistic= 16.000						
T = 340.000	n(small)= 8	n(large)= 40				
P <0.001						
The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P <0.001)						

Table A2 Demographics: Area (km^2)

Demographic data: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Rural	A	40	0	89.818	60.660	156.283
Urban	B	8	0	25.380	18.941	31.648
Mann-Whitney U Statistic= 25.000						
T = 61.000	n(small)= 8	n(large)= 40				
P <0.001						
The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P <0.001)						

Table A3 Demographics: Population Density (people/km^2)

Demographic data: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Rural	A	40	0	61.187	26.107	196.894
Urban	B	8	0	1382.606	1268.094	1806.801
Mann-Whitney U Statistic= 0.000						
T = 356.000	n(small)= 8	n(large)= 40				
P <0.001						
The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P <0.001)						

Table A4 Demographics: Percentage of Individuals Below the Poverty Level

Demographic data: **t – Test**						
	Group	**N**	**Missing**	**Mean**	**Std Dev**	**SEM**
Rural	A	40	0	0.755	0.213	0.0337
Urban	B	8	0	1.050	0.195	0.0689
Difference -0.295						
t = -3.620 with 46 degrees of freedom.						
P <0.001						
95 percent confidence interval for difference of means: -0.459 to -0.131						
The difference in the mean values of the two groups is greater than would be expected by chance; there is a statistically significant difference between the input groups (P <0.001).						
Power of performed test with alpha = 0.050: 0.943						

Table A5 Demographics: Median household income (US dollars)

Demographic data: **t - Test**						
	Group	**N**	**Missing**	**Mean**	**Std Dev**	**SEM**
Rural	A	40	0	4.677	0.0823	0.0130
Urban	B	8	0	4.607	0.0718	0.0254
Difference 0.0697						
t = 2.228 with 46 degrees of freedom.						
P = 0.031						
95 percent confidence interval for difference of means: 0.00675 to 0.133						
The difference in the mean values of the two groups is greater than would be expected by chance; there is a statistically significant difference between the input groups (P = 0.031).						
Power of performed test with alpha = 0.050: 0.486						

Table A6 Demographics: Median Age (Years)

Demographic data: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Rural	A	40	0	35.950	34.400	37.450
Urban	B	8	0	32.200	30.250	33.300
Mann-Whitney U Statistic= 41.500						
T = 77.500	n(small)= 8	n(large)= 40				
P =0.001						
The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P =0.001)						

Table A7 Demographics: Mean Travel Time to Work (Minutes)

Demographic data: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Rural	A	40	0	23.900	20.500	29.000
Urban	B	8	0	19.450	19.300	19.900
Mann-Whitney U Statistic= 42.000						
T = 78.000	n(small)= 8	n(large)= 40				
P =0.001						
The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P =0.001)						

Table A8 Demographics: % Caucasian

Demographic data: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Rural	A	40	0	96.950	94.600	97.650
Urban	B	8	0	79.400	68.900	84.200
Mann-Whitney U Statistic= 10.000						
T = 46.000	n(small)= 8	n(large)= 40				
P <0.001						
The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P <0.001)						

APPENDIX B: Statistical comparison of disease between urban and rural areas

Table B1 Disease Comparison: Non-normalized

Disease Comparison Data: Population Normalized: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Urban giardiasis	A	8	0	53.000	34.000	65.500
Urban cryptosporidiosis	B	8	0	5.000	4.000	5.500
Rural giardiasis	C	40	0	5.000	0.500	19.000
Rural cryptosporidiosis	D	40	0	1.500	0.000	4.000
H = 29.106 with 3 degrees of freedom.						
P < 0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P < 0.001)						

Disease Comparison Data: Population Normalized: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
A vs D	55.575	5.151	Yes
A vs C	38.750	3.592	Yes
A vs B	35.375	2.540	No
B vs D	20.200	1.872	No
B vs C	3.375	0.313	Do Not Test
C vs D	16.825	2.701	Do Not Test
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

Table B2 Disease Comparison: Population Normalized

Disease Comparison Data: Population Normalized: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Urban giardiasis	A	8	0	0.00117	0.000950	0.00172
Urban cryptosporidiosis	B	8	0	0.000136	0.000113	0.000154
Rural giardiasis	C	40	0	0.000788	0.000164	0.00118
Rural cryptosporidiosis	D	40	0	0.000184	0.000	0.000347
H = 27.502 with 3 degrees of freedom						
P < 0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P < 0.001)						

Disease Comparison Data: Population Normalized: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
A vs B	49.875	3.581	Yes
A vs D	46.112	4.274	Yes
A vs C	24.013	2.226	No
C vs B	25.862	2.397	No
C vs D	22.100	3.548	Do Not Test
D vs B	3.763	0.349	Do Not Test
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

Table B3 Disease Comparison: Area Normalized

Disease Comparison Data: Area Normalized: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Urban giardiasis	A	8	0	1.961	1.406	2.522
Urban cryptosporidiosis	B	8	0	0.184	0.167	0.250
Rural giardiasis	C	40	0	0.0392	0.00440	0.172
Rural cryptosporidiosis	D	40	0	0.0145	0.000	0.0360
H = 38.089 with 3 degrees of freedom.						
P < 0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P < 0.001)						

Disease Comparison Data: Area Normalized: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
A vs D	58.612	5.433	Yes
A vs C	43.263	4.010	Yes
A vs B	18.625	1.337	No
B vs D	39.987	3.706	Yes
B vs C	24.638	2.284	No
C vs D	15.350	2.464	No
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

Table B4 Disease Comparison: Population Density Normalized

Disease Comparison Data: Population Density Normalized: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Urban giardiasis	A	8	0	0.0340	0.0230	0.0435
Urban cryptosporidiosis	B	8	0	0.00330	0.00252	0.00422
Rural giardiasis	C	40	0	0.0729	0.0165	0.146
Rural cryptosporidiosis	D	40	0	0.0156	0.000	0.0438
H = 15.228 with 3 degrees of freedom.						
P = 0.002						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = 0.002)						

Disease Comparison Data: Population Density Normalized: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
C vs B	33.438	3.099	Yes
C vs D	19.425	3.118	Yes
C vs A	9.688	0.898	No
A vs B	23.750	1.705	No
A vs D	9.737	0.903	Do Not Test
D vs B	14.013	1.299	Do Not Test
C vs B	33.438	3.099	Yes
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

APPENDIX C: Statistical comparison of occurrence data

Table C1 Occurrence Data: River Raisin Categorized

Occurrence data: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium occurrence: River Raisin sites downstream of CSO outfalls	A	15	0	0.194	0.149	0.489
Giardia occurrence: River Raisin sites downstream of CSO outfalls	B	15	0	0.157	0.126	0.193
Cryptosporidium occurrence: River Raisin sites upstream of CSO outfalls	C	24	0	0.262	0.131	2.629
Giardia occurrence: River Raisin sites upstream of CSO outfalls	D	24	0	0.191	0.131	0.316
H = 3.815 with 3 degrees of freedom.						
P< 0.282						
The differences in the median values among the treatment groups are not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference (P = 0.282)						

Table C2 Occurrence Data: River Raisin Pooled

Occurrence data: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium occurrence: all River Raisin samples	A	39	0	0.240	0.132	0.635
Giardia occurrence: all River Raisin samples	B	39	0	0.161	0.128	0.276
Mann-Whitney U Statistic= 586.000						
T = 1715.000	n(small)= 39	n(large)= 39				
P = 0.082						
The difference in the median values between the two groups is not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference (P = 0.082)						

Table C3 Occurrence Data: Grand River

Occurrence Data: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium occurrence: North Beach Park	A	19	0	0.1000	0.0905	0.133
Giardia recovery : North Beach Park	B	19	0	0.109	0.0936	0.136
Cryptosporidium occurrence: Riverside Park	C	19	0	0.181	0.137	0.379
Giardia recovery : Riverside Park	D	19	0	0.280	0.152	0.434
Cryptosporidium occurrence: Deer Creek Park	E	22	0	0.176	0.1000	0.510
Giardia recovery : Deer Creek Park	F	22	0	0.112	0.1000	0.200
H = 25.884 with 5 degrees of freedom						
P < 0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)						

C3 (cont'd).

Occurrence Data: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
D vs A	44.789	3.969	Yes
D vs B	38.316	3.395	Yes
D vs F	30.713	2.819	No
D vs E	12.258	1.125	Do Not Test
D vs C	8.579	0.760	Do Not Test
C vs A	36.211	3.209	Yes
C vs B	29.737	2.635	No
C vs F	22.134	2.032	Do Not Test
C vs E	3.679	0.338	Do Not Test
E vs A	32.531	2.986	Yes
E vs B	26.057	2.392	Do Not Test
E vs F	18.455	1.760	Do Not Test
F vs A	14.077	1.292	No
F vs B	7.603	0.698	Do Not Test
B vs A	6.474	0.574	Do Not Test
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

Table C4 Occurrence Data: Grand River and MARB

Occurrence Data: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium occurrence: Market Avenue Retention Basin (MARB)	A	8	0	5.000	3.333	18.333
Giardia recovery : Market Avenue Retention Basin (MARB)	B	8	0	1541.667	732.400	2040.000
Cryptosporidium occurrence: North Beach Park	C	19	0	0.1000	0.0905	0.133
Giardia recovery : North Beach Park	D	19	0	0.109	0.0936	0.136
Cryptosporidium occurrence: Riverside Park	E	19	0	0.181	0.137	0.379
Giardia recovery : Riverside Park	F	19	0	0.280	0.152	0.434
Cryptosporidium occurrence: Deer Creek Park	G	22	0	0.176	0.1000	0.510
Giardia recovery : Deer Creek Park	H	22	0	0.112	0.1000	0.200
H = 61.849 with 7 degrees of freedom.						
P < 0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)						

Table C4 (cont'd).

Occurrence Data: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
B vs C	94.395	5.684	Yes
B vs D	87.921	5.294	Yes
B vs H	80.318	4.937	Yes
B vs G	61.545	3.783	Yes
B vs E	58.184	3.504	Yes
B vs F	49.526	2.982	No
B vs A	9.063	0.460	Do Not Test
A vs C	85.332	5.138	Yes
A vs D	78.859	4.748	Yes
A vs H	71.256	4.380	Yes
A vs G	52.483	3.226	Yes
A vs E	49.122	2.958	No
A vs F	40.464	2.437	Do Not Test
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

Table C4 (cont'd).

Occurrence Data: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
F vs C	44.868	3.510	Yes
F vs D	38.395	3.003	No
F vs H	30.792	2.495	Do Not Test
F vs G	12.019	0.974	Do Not Test
F vs E	8.658	0.677	Do Not Test
E vs C	36.211	2.832	No
E vs D	29.737	2.326	Do Not Test
E vs H	22.134	1.794	Do Not Test
E vs G	3.361	0.272	Do Not Test
G vs C	32.849	2.662	Do Not Test
G vs D	26.376	2.137	Do Not Test
G vs H	18.773	1.580	Do Not Test
H vs C	14.077	1.141	Do Not Test
H vs D	7.603	0.616	Do Not Test
D vs C	6.474	0.506	Do Not Test
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

Table C5 Occurrence Data: River Raisin (Pooled) and Grand River

Occurrence Data: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium occurrence: North Beach Park	A	19	0	0.1000	0.0905	0.133
Giardia recovery : North Beach Park	B	19	0	0.109	0.0936	0.136
Cryptosporidium occurrence: Deer Creek Park and Riverside Park pooled samples	C	41	0	0.181	0.1000	0.500
Giardia recovery : Deer Creek Park and Riverside Park pooled samples	D	41	0	0.143	0.1000	0.315
Cryptosporidium occurrence: all River Raisin samples	E	39	0	0.240	0.132	0.635
Giardia occurrence: all River Raisin samples	F	39	0	0.161	0.128	0.276
H = 30.060 with 5 degrees of freedom.						
P < 0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)						

Table C5 (cont'd).

Occurrence Data: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
E vs A	72.244	4.506	Yes
E vs B	62.455	3.896	Yes
E vs D	25.779	2.011	No
E vs F	19.397	1.495	Do Not Test
E vs C	15.133	1.181	Do Not Test
C vs A	57.112	3.591	Yes
C vs B	47.322	2.976	Yes
C vs D	10.646	0.841	Do Not Test
C vs F	4.265	0.333	Do Not Test
F vs A	52.847	3.296	Yes
F vs B	43.057	2.686	No
F vs D	6.381	0.498	Do Not Test
D vs A	46.465	2.922	No
D vs B	36.676	2.306	Do Not Test
B vs A	9.789	0.527	Do Not Test
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

APPENDIX D: Statistical analysis of detection limits

Table D1 Detection Limits: River Raisin Categorized

Detection limits: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
River Raisin sites downstream of CSO outfalls	A	15	0	0.157	0.126	0.193
River Raisin sites upstream of CSO outfalls	B	24	0	0.168	0.119	0.316
Mann-Whitney U Statistic= 153.000						
T = 273.000	n(small)= 15	n(large)= 24				
P = 0.444						
The difference in the median values between the two groups is not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference (P = 0.444)						

Table D2 Detection Limits: Grand River

Detection limits: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
North Beach Park	A	19	0	0.1000	0.0915	0.129
Riverside Park	B	19	0	0.1000	0.0948	0.133
Deer Creek Park	C	22	0	0.1000	0.0944	0.129
H = 0.715 with 2 degrees of freedom						
P = 0.700						
The differences in the median values among the treatment groups are not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference (P = 0.700)						

Table D3 Detection Limits: Grand River and MARB

Detection limits: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
North Beach Park	A	19	0	0.1000	0.0915	0.129
Riverside Park	B	19	0	0.1000	0.0948	0.133
Deer Creek Park	C	22	0	0.1000	0.0944	0.129
MARB	D	8	0	5.000	3.333	18.333
H = 21.495 with 3 degrees of freedom.						
P <0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)						
Detection limits: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**						
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**			
D vs A	36.500	4.380	Yes			
D vs C	33.795	4.140	Yes			
D vs B	31.737	3.808	Yes			
B vs A	4.763	0.742	No			
B vs C	2.059	0.332	Do Not Test			
C vs A	2.705	0.437	Do Not Test			
Note: The multiple comparisons on ranks do not include an adjustment for ties.						

Table D4 Detection Limits: Grand River and River Raisin

Detection limits: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Deer Creek Park and Riverside Park	A	41	0	0.1000	0.0947	0.130
River Raisin total sites	B	43	0	0.161	0.128	0.241
North Beach Park	C	19	0	0.1000	0.0915	0.129
H = 34.233 with 2 degrees of freedom.						
P <0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P <0.001)						
Detection limits: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**						
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**			
B vs C	37.729	4.584	Yes			
B vs A	33.402	5.122	Yes			
A vs C	4.327	0.522	No			
Note: The multiple comparisons on ranks do not include an adjustment for ties.						

APPENDIX E: Statistical analysis of matrix spike recovery efficiency

Table E1 Matrix Spike: River Raisin Categorized Data

Matrix Spike Sample Recovery Efficiency: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium recovery: River Raisin sites downstream of CSO outfalls	A	5	0	54.400	34.400	63.4
Giardia recovery: River Raisin sites downstream of CSO outfalls	B	5	0	11.800	8.867	31.225
Cryptosporidium recovery: River Raisin sites upstream of CSO outfalls	C	9	0	39.100	24.612	75.163
Giardia recovery: River Raisin sites upstream of CSO outfalls	D	9	0	3.800	0.000	7.075
H =19.276 with 3 degrees of freedom						
P <0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P <0.001)						

Matrix Spike Sample Recovery Efficiency: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
A vs D	15.333	3.342	Yes
A vs B	8.600	1.653	No
A vs C	0.111	0.0242	Do Not Test
C vs D	15.222	3.926	Yes
C vs B	8.489	1.850	Do Not Test
B vs D	6.733	1.468	No
Note: The multiple comparisons on ranks do not include an adjustment for ties.			

Table E2 Matrix Spike: River Raisin Pooled Data

Matrix Spike Sample Recovery Efficiency: **Mann-Whitney Rank Sum Test**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium recovery : all River Raisin associated matrix spike samples	A	14	0	46.750	26.150	73.300
Giardia recovery : all River Raisin associated matrix spike samples	B	14	0	6.735	2.750	12.400
Mann-Whitney U Statistic= 8.000						
T = 293.000	n(small)= 14	n(large)= 14				
P <0.001						
The difference in the median values between the two groups is greater than would be expected by chance; there is a statistically significant difference (P = <0.001)						

Table E3 Matrix Spike: Grand River

Matrix Spike Sample Recovery Efficiency: **t - Test**						
	Group	**N**	**Missing**	**Mean**	**Std Dev**	**SEM**
Cryptosporidium recovery : all recreational Grand River site associated matrix spike samples	A	5	0	69.940	9.198	4.113
Giardia recovery : all recreational Grand River site associated matrix spike samples	B	5	0	52.580	23.344	10.440
Difference 17.360						
t = 1.547 with 8 degrees of freedom.						
P = 0.160						
95 percent confidence interval for difference of means: -8.516 to 43.236						
The difference in the mean values of the two groups is not great enough to reject the possibility that the difference is due to random sampling variability. There is not a statistically significant difference between the input groups (P = 0.160).						
Power of performed test with alpha = 0.050: 0.173						
The power of the performed test (0.173) is below the desired power of 0.800. Less than desired power indicates you are less likely to detect a difference when one actually exists. Negative results should be interpreted cautiously.						

Table E4 Matrix Spike: Grand River and MARB

Matrix Spike Sample Recovery Efficiency: **One Way Analysis of Variance**						
	Group Name	**N**	**Missing**	**Mean**	**Std Dev**	**SEM**
Cryptosporidium recovery : all MARB associated matrix spike samples	A	7	0	33.714	22.625	8.552
Giardia recovery : all MARB associated matrix spike samples	B	7	0	42.071	24.852	9.393
Cryptosporidium recovery : all recreational Grand River site associated matrix spike samples	C	5	0	69.940	9.198	4.113
Giardia recovery : all recreational Grand River site associated matrix spike samples	D	5	0	52.580	23.344	10.440

Source of Variation	**DF**	**SS**	**MS**	**F**
Between Groups	3	4183.007	1394.336	3.000
Residual	20	9295.383	464.769	
Total	23	13478.390		

P = 0.055

The differences in the mean values among the treatment groups are not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference (P = 0.055).

Power of performed test with alpha = 0.050: 0.429

The power of the performed test (0.429) is below the desired power of 0.800.
Less than desired power indicates you are less likely to detect a difference when one actually exists. Negative results should be interpreted cautiously.

Table E5 Matrix Spike: Grand River and River Raisin

Matrix Spike Sample Recovery Efficiency: **Kruskal-Wallis One Way Analysis of Variance on Ranks**						
	Group	**N**	**Missing**	**Median**	**25%**	**75%**
Cryptosporidium recovery : all recreational Grand River site associated matrix spike samples	A	5	0	73.000	66.575	74.800
Giardia recovery : all recreational Grand River site associated matrix spike samples	B	5	0	46.900	38.725	67.125
Cryptosporidium recovery : all River Raisin associated matrix spike samples	C	14	0	46.750	26.150	73.300
Giardia recovery : all River Raisin associated matrix spike samples	D	14	0	6.735	2.750	12.400
H = 24.144 with 3 degrees of freedom.						
P<0.001						
The differences in the median values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0.001)						

Table E5 (cont'd).

Matrix Spike Sample Recovery Efficiency: **All Pairwise Multiple Comparison Procedures (Dunn's Method)**			
Comparison:	**Difference of Ranks**	**Q**	**P<0.05**
A vs D	22.786	3.936	Yes
A vs C	6.500	1.123	No
A vs B	5.400	0.768	Do Not Test
B vs D	17.386	3.003	Yes
B vs C	1.100	0.190	Do Not Test
C vs D	16.286	3.877	Yes

Note: The multiple comparisons on ranks do not include an adjustment for ties.
A result of "Do Not Test" occurs for a comparison when no significant difference is found between two means that enclose that comparison. Note that not testing the enclosed means is a procedural rule, and a result of Do Not Test should be treated as if there is no significant difference between the means, even though one may appear to exist.

APPENDIX F: Statistical analysis of ongoing precision and recovery (OPR) efficiency

Table F1 OPR: River Raisin

OPR: **t - Test**						
	Group	**N**	**Missing**	**Mean**	**Std Dev**	**SEM**
Cryptosporidium recovery : all River Raisin site associated OPR samples	A	4	0	41.300	7.771	3.885
Giardia recovery : all River Raisin site associated OPR samples	B	4	0	42.063	10.535	5.268
Difference -0.763						
t = -0.116 with 6 degrees of freedom.						
P = 0.911						
95 percent confidence interval for difference of means: -16.779 to 15.254						
The difference in the mean values of the two groups is not great enough to reject the possibility that the difference is due to random sampling variability. There is not a statistically significant difference between the input groups (P = 0.911).						
Power of performed test with alpha = 0.050: 0.050						
The power of the performed test (0.050) is below the desired power of 0.800. Less than desired power indicates you are less likely to detect a difference when one actually exists. Negative results should be interpreted cautiously.						

Table F2 OPR: Grand River

OPR: **t - Test**						
	Group	**N**	**Missing**	**Mean**	**Std Dev**	**SEM**
Cryptosporidium recovery : all recreational Grand River site associated OPR samples	A	23	0	55.230	11.741	2.448
Giardia recovery : all recreational Grand River site associated OPR samples	B	23	0	43.939	15.681	3.270
Difference 11.291						
t = 2.764 with 44 degrees of freedom.						
P = 0.008						
95 percent confidence interval for difference of means: 3.059 to 19.523						
The difference in the mean values of the two groups is greater than would be expected by chance; there is a statistically significant difference between the input groups (P = 0.008).						
Power of performed test with alpha = 0.050: 0.720						

Table F3 OPR: Grand River and MARB

OPR: **One Way Analysis of Variance**

	Group Name	N	Missing	Mean	Std Dev	SEM
Cryptosporidium recovery : all MARB associated OPR samples	A	5	0	47.320	13.601	6.083
Giardia recovery : all MARB associated OPR samples	B	5	0	67.020	17.257	7.718
Cryptosporidium recovery : all recreational Grand River site associated OPR samples	C	23	0	55.230	11.741	2.448
Giardia recovery : all recreational Grand River site associated OPR samples	D	23	0	43.939	15.681	3.270

Source of Variation	DF	SS	MS	F
Between Groups	3	2909.014	969.671	4.861
Residual	52	10373.439	199.489	
Total	55	13282.454		

P = 0.005

The differences in the mean values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = 0.005).

Power of performed test with alpha = 0.050: 0.799

Table F3 (cont'd).

OPR: **All Pairwise Multiple Comparison Procedures (Tukey Test):**						
Comparison	**Diff of Means**	**p**	**q**	**P**	**P<0.050**	
B vs. D	23.081	4	4.684	0.009	Yes	
B vs. A	19.700	4	3.119	0.135	No	
B vs. C	11.790	4	2.392	0.338	Do Not Test	
C vs. D	11.291	4	3.834	0.044	Yes	
C vs. A	7.910	4	1.605	0.670	Do Not Test	
A vs. D	3.381	4	0.686	0.962	No	
Note: A result of "Do Not Test" occurs for a comparison when no significant difference is found between two means that enclose that comparison. Note that not testing the enclosed means is a procedural rule, and a result of Do Not Test should be treated as if there is no significant difference between the means, even though one may appear to exist.						

Table F4 OPR: Grand River and River Raisin

OPR: **One Way Analysis of Variance**

	Group Name	N	Missing	Mean	Std Dev	SEM
Cryptosporidium recovery : all recreational Grand River site associated OPR samples	A	23	0	55.230	11.741	2.448
Giardia recovery : all recreational Grand River site associated OPR samples	B	23	0	43.939	15.681	3.270
Cryptosporidium recovery : all River Raisin site associated OPR samples	C	4	0	41.300	7.771	3.885
Giardia recovery : all River Raisin site associated OPR samples	D	4	0	42.063	10.535	5.268

Source of Variation	DF	SS	MS	F
Between Groups	3	1893.032	631.011	3.523
Residual	50	8956.390	179.128	
Total	53	10849.422		

P = 0.021

The differences in the mean values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = 0.021).

Power of performed test with alpha = 0.050: 0.585

Table F4 (cont'd).

OPR**: All Pairwise Multiple Comparison Procedures (Holm-Sidak method)**						
Comparison	**Diff of Means**	**t**	**Unadjusted P**	**Critical Level**	**P<0.050**	
A vs. B	11.291	2.861	0.006	0.009	Yes	
A vs. C	13.930	1.921	0.060	0.010	No	
A vs. D	13.168	1.816	0.075	0.013	No	
B vs. C	2.639	0.364	0.717	0.017	No	
B vs. D	1.877	0.259	0.797	0.025	No	
D vs. C	0.763	0.0806	0.936	0.050	No	
Note: The multiple comparisons on ranks do not include an adjustment for ties.						

REFERENCES

REFERENCES

1. Applebee, A.J. Thompson, R.C., Olson, M.E., 2005. *Giardia* and *Cryptosporidium* in mammalian wildlife – current status and future needs. Trends in Parasitology 21: 370-376.

2. Arnone, R.D., Walling. J.P., 2006. Evaluating *Cryptosporidium* and *Giardia* concentrations in combined sewer overflow. Journal of Water and Health 4(2):157-65.

3. Atwill, E.R., Pereira, M.D., Alonso, L.H., Elmi, C., Epperson, W.B., Smith, R., Riggs, W., Carpenter, L.V., Dargatz, D.A., Hoar, B., 2006. Environmental Load of *Cryptosporidium parvum* Oocysts from Cattle Manure in Feedlots from the Central and Western United States. Journal of Environmental Quality 35:200-206.

4. Baxby, D., Hart, C.A., Blundell, N., 1985. Shedding of oocysts by immunocompetent individuals with cryptosporidiosis. Journal of Hygiene 95:703–709.

5. Brodsky, R.E., Spencer, H.C., Schultz M.G., 1974. Giardiasis in American travelers to the Soviet Union. Journal of Infectious Diseases 130:319–323.

6. Centers for Disease Control and Prevention, 2011. Cryptosporidiosis (*Cryptosporidium* spp.) case definition. Atlanta, GA: US Department of Health and Human Services, CDC. Available at http://www.cdc.gov/osels/ph_surveillance/nndss/casedef/cryptosporidiosis_current.htm. Accessed May 5, 2011.

7. Centers for Disease Control and Prevention, 2011. Giardiasis (*Giardia lamblia*) case definition. Atlanta, GA: US Department of Health and Human Services, CDC. Available at http://www.cdc.gov/osels/ph_surveillance/nndss/print/giardiasis_current.htm. Accessed May 5, 2011.

8. Centers for Disease Control and Prevention, 2010. Cryptosporidiosis Surveillance – United States 2006-2008 and Giardiasis Surveillance – United States 2006-2008. In: Surveillance Summaries, June 11, 2010. MMWR 59 (No. SS-6).

9. Centers for Disease Control and Prevention, 2007. Cryptosporidiosis Surveillance – United States 2003-2005 and Giardiasis Surveillance – United States 2003-2005. In: Surveillance Summaries, September 7, 2007. MMWR 56 (No. SS-7).

10. Centers for Disease Control and Prevention, 2006. Summary of notifiable diseases - United States, 2006. MMWR 55(No. 53).

11. Centers for Disease Control and Prevention, 2005. Cryptosporidiosis Surveillance – United States 1999-2002 and Giardiasis Surveillance – United States 1998-2002. In: Surveillance Summaries, January 28, 2005. MMWR 54 (No. SS-1).

12. Centers for Disease Control and Prevention, 2000. Giardiasis Surveillance --- United States, 1992—1997. In: Surveillance Summaries August 11 2000. MMWR 49 (No SS-07).

13. Centers for Disease Control and Prevention, 1998. Summary of notifiable diseases, United States, 1998. MMWR 47(53): 5.

14. Centers for Disease Control and Prevention, 1998. Summary of notifiable diseases, United States, 1997. MMWR 46(54): 5.

15. Centers for Disease Control and Prevention, 1996. Summary of notifiable diseases, United States, 1996. MMWR 45(53):vii.

16. Centers for Disease Control and Prevention, 1993. Summary of notifiable diseases - United States, 1993. Published October 21, 1994, for MMWR 42(53).

17. Cox, P., Griffith, M., Angles, M., Deere, D., Ferguson, C., 2005. Concentrations of Pathogens and Indicators in Animal Feces in the Sydney Watershed. Applied and Environmental Microbiology 71: 5929-5934

18. *Cryptosporidium* and Cryptosporidiosis, 2nd edition. Fayer, R and Xiao, L., ed. 2008 CRC Press. Boca Raton, Fl.

19. Donovan, E., Unice, K., Roberts, J.D., Harris ,M., Finley, B., 2008. Risk of gastrointestinal disease associated with exposure to pathogens in the water of the lower Passaic River. Applied and Environmental Microbiology 74: 994-1003.

20. Erickson, D.L., 1995. Rural land use and land cover change: Implications for local planning in the River Raisin watershed. Land Use Policy. 12(3): 223-236

21. Federal Register, 2004. Water quality standards for coastal and Great Lakes recreation waters; final rule. Federal Register 69(220): 67217-67243.

22. Feng, Y., Alderisio, K.A., Yang, W., Blancero, L.A., Kuhne, W.G., Nadareski, C.A., Reid, M., Xiao, L., 2007. *Cryptosporidium* Genotypes in Wildlife from a New York Watershed. Applied and Environmental Microbiology 73: 6475-6483.

23. *Giardia*: From Molecules to Disease. Thompson, R.C.A, Reynoldson, J.A., Lymbery, A.J., ed. 1994. CAB International, United Kingdom.

24. Gibson, C.J., Stadterman, K.L., States, S., Sykora, J., 1998. Combined sewer overflows: a source of *Cryptosporidium* and *Giardia*? Water Science and Technology 38(12): 67-72.

25. Haas, C.N., Crockett, C.S., Rose, J.B., Gerba, C.P., Fazil, A.M., 1996. Assessing the risk posed by oocysts in drinking water. American Water Works Association Journal 88(9): 131-136.

26. Heitman, T.L., Frederick, L.M., Viste, J.R., Guselle, N.J., Morgan, U.M., Thompson, R.C.A., Olson, M.E., 2002. Prevalence of *Giardia* and *Cryptosporidium* and characterization of *Cryptosporidium* spp. isolated from wildlife, human, and agricultural sources in the North Saskatchewan River Basin in Alberta, Canada. Canadian Journal of Microbiology 48:530-541.

27. Hunter, P.R. Thompson, R.C., 2005. The zoonotic transmission of *Giardia* and *Cryptosporidium*. International Journal for Parasitology. 35(11-12):1181-90.

28. Hunter, P.R., Chalmers, R.M., Syed, Q., Hughes, L.S., Woodhouse, S., Swift, L., 2003. Foot and Mouth Disease and Cryptosporidiosis: Possible interaction between two emerging infectious diseases. Emerging Infectious Diseases 9(1):109-112.

29. Hutchison, M.L., Walters, L. D., Moore, A., Avery, S.M., 2005. Declines of zoonotic agents in liquid livestock wastes stored in batches on-farm. Journal of Applied Microbiology 99:58-65.

30. Jokipii, A.M.M., Hemila, M., Jokipii, L., 1985. Prospective study of acquisition of *Cryptosporidium, Giardia lamblia,* and gastrointestinal illness. *Lancet*, 2(8453):487–489.

31. Lake, I.R., Nichols, G., Harrison, F.C.D., Bentham, G., Kovats, R.S., Grundy, C., Hunter, P.R., 2009. Using infectious intestinal disease surveillance data to explore illness aetiology; a cryptosporidiosis case study. Health and Place 15(1):333-339.

32. Lake, I.R., Harrison, F.C.D., Chalmers, R.M., Bentham, G., Nichols, G., Hunter, P.R., Kovats, R.S., Grundy, C., 2007. Case-control study of environmental and social factors influencing cryptosporidiosis. European journal of epidemiology 22(11): 805-811.

33. Levine, A.D., Harwood, V.J., Farrah, S.R., Scott, T.M., Rose, J.B., 2008. Pathogen and indicator organism reduction through secondary effluent filtration: implications for reclaimed water production. Water Environment Research 80(7): 596-608.

34. Michigan Department of Community Health. Personal communication.

35. Michigan Geographic Data Library. http://www.mcgi.state.mi.us/mgdl/ Accessed January 2010

36. Michigan Department of Environmental Quality (MDEQ). Online Services: Combined Sewer Overflow and Sanitary Sewer Overflow System. http://www.deq.state.mi.us/csosso Accessed October 2008.

37. Odoi, A., Martin, S.W., Michel, P., Middleton, D., Holt, J., Wilson, J., 2004. Investigation of clusters of giardiasis using GIS and a spatial scan statistic. International Journal of Health Geographics 3(11).

38. Okhuysen, P.C., Chappell, C.L., Crabb, J.H., Sterling, C.R., DuPont, H.L., 1999. Virulence of three distinct *Cryptosporidium parvum* isolates for healthy adults. The Journal of Infectious Diseases 180(4): 1275-1281.

39. Olson, M.E., Guselle, N.J., O'Handley, R.M., Swift, M.L., McAllister, T.A., Jelinski, M.D., Morck.D.W., 1997. *Cryptosporidium* in dairy calves in British Columbia. The Canadian Veterinary Journal 38(11):703-706.

40. Olson, M.E., Thorlakson, C.L., Deselliers, L., Morck.D.W., McAllister, T.A., 1997. *Giardia* and *Cryptosporidium* in Canadian farm animals. Veterinary Parasitology 68: 375-381.

41. Ortega, Y.R., Adam, R.D., 1997. *Giardia*: Overview and Update. Clinical Infectious Diseases 25:545-549.

42. Pollock, K.G.J., Ternent, H.E., Mellor, D.J., Chalmers, R.M., Smith, H.V., Ramsay, C.N., Innocent, G.T., 2010. Spatial and Temporal Epidemiology of Sporadic Human Cryptosporidiosis in Scotland. Zoonoses and Public Health 57: 487-492.

43. Rendtorff, R.C., 1954. The experimental transmission of human intestinal protozoan parasites: *Giardia intestinalis* cysts given in capsules. American Journal of Hygiene 59: 209–220.

44. Rose, J.B., 2007. Water reclamation, reuse and public health. Water Science & Technology 55(1-2): 275–282.

45. Rose, J.B., Haas, C.N., Regli, S., 1991. Risk assessment and control of waterborne giardiasis. American Journal of Public Health. 81(6): 709-13

46. Roxstrom-Lindquist, K.,Palm, Reiner, D., Ringqvist, D.E., Svard, .G., 2006. *Giardia* immunity – an update. Trends in Parasitology 22(1): 26-31.

47. Smerdon, W.J., Nichols, T., Chalmers, R.M., Heine, H., Reacher, M.H., 2003. Foot and mouth disease in livestock and reduced cryptosporidiosis in humans, England and Wales. Emerging infectious diseases. 9(1): 22-28.

48. States, S., Stadterman, K.L., Ammon, L., Vogel, P., Baldizar, J., Wright, D., Conley, L., Sykora, J., 1997. Protozoa in river water: sources, occurrence, and treatment. Journal of the American Water Works Association 89(9):74-87.

49. Thompson, R.C.A., Hopkins, R.M., Homan, W.L., 2000. Nomenclature and genetic groupings of *Giardia* infecting mammals. Parasitology Today 16: 210-213

50. U.S. Army Corps of Engineers (USACE), 1972. Comprehensive Water Resource Study of the Grand River Basin Michigan.

51. U.S. Census Bureau. http://www.census.gov/ Access date: January 2010.

52. U.S. Department of the Interior, U.S. Geological Survey. USGS Surface-Water Data for Michigan. http://waterdata.usgs.gov/mi/nwis/sw? Site 04119000 Access date: November 2008

53. U.S. Environmental Protection Agency (U.S.EPA), 2008. Child-specific exposure factors handbook. National Center for Environmental Assessment, Washington, DC; EPA/600/RP-060/096F.

54. U.S. Environmental Protection Agency (U.S.EPA), 2005. Method 1623: *Cryptosporidium* and *Giardia* in water by filtration/IMS/FA. United States Environmental Protection Agency. Office of Water. Washington, D.C. EPA 815-R-05-002

55. U.S. Environmental Protection Agency (U.S.EPA), 2004. Report to Congress. Impacts and control of CSOs and SSOs. EPA 833-R-04-001. U.S. Environmental Protection Agency, Washington, D.C.

56. Wade, T.J., Calderon, R.L., Brenner, K.P., Sams, E., Beach, M., Haugland, R., Wymer, L., Dufour, A.P., 2008. High sensitivity of children to swimming-associated gastrointestinal illness. Epidemiology 19(3): 375–383.

57. World Health Organization. Guidelines for Drinking-Water Quality, 2nd Edition. 1996. Addendum: Microbiological agents in drinking water. Protozoan Parasites (Cryptosporidium, Giardia, Cyclospora). World Health Organization, Geneva.

58. Xiao, L., Herd, R.P., 1994. Infection patterns of *Cryptosporidium* and *Giardia* in Calves. Parasitology Today 55:254-262.

59. Zisan, E., Alabay, B.M., Fidanci, H., Düzgün, A., Cerci, H., 1998. Prevalence of *Cryptosporidium* spp. infection and its relation to other enteric pathogens (*Escherichia coli* K99 and rotavirus) in cattle in Ankara, Turkey. Turkish Journal of Veterinary and Animal Sciences 22: 453-457.