

Distribution Agreement

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

Signature:

Leigh M. Tyndall

Date

The Seasonality and Climatic Drivers of Cryptosporidiosis

By

Leigh M. Tyndall

Master of Public Health

Global Environmental Health

Karen Levy, PhD, MPH, MSc

Committee Chair

Paige Tolbert, PhD, MSPH

Committee Member

The Seasonality and Climatic Drivers of Cryptosporidiosis

By

Leigh M. Tyndall

Bachelor of Science

University of North Carolina at Chapel Hill

2011

Thesis Committee Chair: Karen Levy, PhD, MSPH, MSc

An abstract of

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

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health

2014

Abstract

The Seasonality and Climatic Drivers of Cryptosporidiosis

By Leigh M Tyndall

Purpose: There is much uncertainty about the relationship between climate and diarrheal disease in the scientific literature, due to a lack of studies that target this question, a lack of studies of the relationships between climate and individual pathogens, and also true heterogeneity of effect. This study attempts to address these factors through an analysis of cryptosporidiosis specifically, and its relationship with temperature (°C) and rainfall (mm), testing for heterogeneity both within and between datasets.

Methods: All US cryptosporidiosis cases reported monthly between 1997-2011 were obtained from the National Notifiable Disease Surveillance System (NNDSS). These data were analyzed with monthly temperature and precipitation data, using generalized linear model and generalized estimating equation regression analyses to calculate incidence rate ratios for each state, nine climate regions, and for the US as a whole. Heterogeneity of results was assessed using the I^2 statistic. A systematic review of the literature was also performed, searching for studies worldwide that presented at least one full year of monthly data on cryptosporidiosis incidence. These data were extracted, matched with climate data for the same periods, and analyzed separately. The results were compared to the NNDSS analysis.

Results: There is an overall positive relationship between temperature and cryptosporidiosis in the US—for every 1 °C increase in temperature, cryptosporidiosis case incidence increases by 2.51%. This is supported by the global literature review which reports a 2.96% increase in cryptosporidiosis for every 1 °C increase in temperature worldwide. There is much variability in the relationship between precipitation and cryptosporidiosis in the US, which may be due to local geographic and temporal factors. There was no significant heterogeneity in results between states, but considerable heterogeneity between climate regions.

Conclusion: In general, there is a positive relationship between cryptosporidiosis and temperature, shown both in the US and worldwide. The relationship between cryptosporidiosis and precipitation is not as clear and is likely due to factors not considered in this study. The relationship between these climatic variables and cryptosporidiosis cases was remarkably consistent across states and between the US and global analyses. This suggests the temperature-disease relationship is robust to varying conditions.

The Seasonality and Climatic Drivers of Cryptosporidiosis

By

Leigh M. Tyndall

Bachelor of Science

University of North Carolina at Chapel Hill

2011

Thesis Committee Chair: Karen Levy, PhD, MSPH, MSc

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

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health

2014

Acknowledgements

Thanks first and foremost must go to Sharia Ahmed, not only for her analysis of the literature review datasets, but also for her patience, encouragement, and many hours over Skype explaining the intricacies of meta-analysis and modeling. Special thanks also must go to Dr. Karen Levy, my fantastic advisor, who has had incredible patience with me and been very supportive throughout this project.

I would like to thank Michele Hlavsa and Jonathan Yoder at the Centers for Disease Control for their help, insight, and encouragement. I would also like to thank the NNDSS staff at CDC for allowing usage of the NNDSS data, and Dr. Jyotsna Jagai for graciously allowing me access to her dataset for the literature review portion of this thesis.

A very special thank you must go to my dear parents, my sister Diane, and my fiancé, Samuel, for loving me, encouraging me, and praying me through this whirlwind of a Masters degree. I could not have done this without your love, support, and sense of humor.

TABLE OF CONTENTS

INTRODUCTION	1
Background.....	1
Purpose and Motivation for Study	3
Literature Review	4
Cryptosporidiosis	4
<i>Cryptosporidium</i> Parasite.....	4
Cryptosporidiosis Epidemiology.....	4
Clinical Symptoms of Cryptosporidiosis	5
Cryptosporidiosis Treatment and Prevention.....	5
Risk factors for Cryptosporidiosis	6
Seasonality	7
In Temperate Regions.....	7
In Tropical Regions.....	9
Climatic Drivers	10
Temperature and Rainfall as Climatic Drivers.....	10
Infectious Diarrheal Disease.....	11
Cryptosporidiosis and Climatic Drivers.....	14
National Surveillance Data	17
National Notifiable Disease Surveillance System.....	18
Methods.....	20
Hypothesis	20
Outcome Variables	20
Systematic Review	20
National Notifiable Diseases Surveillance System	22
Predictor Variables.....	24
Climate-related Variables.....	24
Demographic Variables	25
Methods of Analysis.....	26
Summarizing Data and Seasonality Analysis	26
Statistical Analysis.....	27
RESULTS:.....	29
Systematic Review	29
National Notifiable Disease Surveillance System Data.....	30
Season Strength.....	30
Seasonality across US Climate Regions.....	31
Generalized Linear Modeling	32
United States and US Climate Regions	32
Heterogeneity between States and Climate Regions.....	34
Generalized Estimating Equations Modeling	35
DISCUSSION AND RECOMMENDATIONS:	37
Season Strength and Seasonality.....	37
Temperature	39

Precipitation.....	42
Well Use	44
Heterogeneity within NNDSS Data	45
Comparison between NNDSS and Meta-analysis.....	45
Pooled Analysis GEE Model.....	45
Temperature	47
Precipitation.....	48
Limitations.....	50
Recommendations	53
Conclusions.....	57
References	59
Tables	76
Table 1: Characteristics of States and Regions included in study.....	76
Table 2: Regression results for overall Generalized Estimating Equations analysis of NNDSS data	78
Table 3: Regression results for overall Generalized Estimating Equations analysis of global literature review data, developed countries in the northern hemisphere and all studies	79
Figures.....	80
Figure 1.	80
Figure 2.	81
Figure 3.	82
Figure 4a.	83
Figure 4b.	84
Figure 5 a and b.	85
Figure 6 a and b.	86
Figure 7 a and b.	87
Appendices.....	88
Appendix A. Detailed search protocol.....	88
Appendix B. PRISMA flow diagram.....	91
Appendix C. Registration Information and Data Use Restrictions Agreement.....	92
Appendix D. Map of NCDC climate regions	94
Appendix E. IRB Declaration Letter	95
Appendix F. Meta-data table for studies	96
Appendix G. Plot of Cryptosporidiosis Season Strength against Temperature and Precipitation Season Strengths	97
Appendix H. Correlations between demographic variables and plots	98
Appendix I. Supplementary Maps of Temperature and Precipitation Season Strength	100

INTRODUCTION

Background

Infectious diarrheal disease is a significant concern worldwide, with an estimated 1.7 to 4 billion cases and 2.2 million deaths annually.^{1,2} A vast majority of these cases occur in developing countries among children under the age of five and in immunocompromised populations. Diarrheal disease is also a significant cause of morbidity in developed countries such as the United States of America (US).²

The exact burden of cryptosporidiosis worldwide is unknown, but a recent review of parasitic waterborne disease outbreaks worldwide found that 60% of the outbreaks were due to *Cryptosporidium*.³ In the US alone it is one of the top causes of recreational water illness and waterborne disease outbreaks, with average yearly attributable hospitalization costs of \$45,770,572.⁴ This number does not include outpatient healthcare costs, the considerable number of individuals who do not seek care for gastro-intestinal (GI) illness, nor loss of economic productivity, which was estimated to be as high as \$64 million in the case of one outbreak, in Wisconsin in 1993.⁵ In the recent Global Enteric Multicenter Study (GEMS), which followed 22,568 children in 7 countries in Sub-Saharan Africa and south Asia, *Cryptosporidium* species rank in the top four etiologies of moderate-to-severe diarrhea in children under five.⁶

With climate change becoming an increasing concern, it is important to investigate the physical drivers that impact infectious diarrheal disease burden, namely ambient temperature and rainfall. The National Institute of Environmental Health Sciences (NIEHS) suggest, in their *A Human Health Perspective on Climate Change Report*, that research quantifying the relationship between change in temperature and precipitation and change in disease is necessary for adequate preventative measures to be taken.⁷

There is currently much uncertainty in the scientific literature concerning how climatic drivers affect diarrheal disease incidence. A study attempting to project change in temperature and the accompanying change in diarrheal disease found more uncertainties associated with diarrheal disease risk than climate projections. The authors highlight the lack of empirical data in this area of research, question current assumptions of a linear relationship, and argue for developing more than one model in such projects to reduce overall uncertainty.⁸

The uncertainty in the literature is due in part to lack of studies that examine climate-diarrheal disease relationships, but also to a lack of studies that examine these relationships for specific pathogens. Diarrheal disease seasonality depends on the etiological agent. Viral diarrheal disease, such as Norovirus and Rotavirus, peak during the cold winter months, whereas bacterial and parasitic diarrheal diseases have the opposite pattern, generally peaking during the warm summer months.⁹ Additionally, most diarrheal diseases have a single seasonal peak. Cryptosporidiosis is somewhat unique in that it displays a bi-modal seasonality in many, but not all countries and climate zones.^{10,11} Given these differences between pathogens, pathogen-specific studies are key to understanding climate-disease relationships.

Of course, there is also true heterogeneity within and between studies that causes uncertainty in the results. This study attempts to address all of these causes of uncertainty by performing a pathogen-specific analysis of the relationship between diarrheal disease incidence, temperature and rainfall, and specifically examining heterogeneity in effect for one dataset, and between two different datasets.

Purpose and Motivation for Study

Research correlating cryptosporidiosis case incidence rates directly to degree change in temperature and millimeter (mm) change in rainfall is lacking in the currently published literature. While there is one published meta-analysis of cryptosporidiosis seasonality and its relationship to temperature and rainfall globally, the authors normalized data reported in various formats (prevalence, incidence, etc.) to build their model, and did not make direct degree and mm associations with case incidence.¹⁰ A thorough literature search yielded no studies attempting to compare seasonality results across different datasets. Similarly, no articles were found that compared models of surveillance data to models using data from a literature review. Several studies were found, however, comparing different models types built with the same data, but these comparisons used data from a single city and not over multiple climate regions.¹²⁻¹⁴ Overall, there is a lack of consensus regarding the relationships between *Cryptosporidium* incidence and climatological factors, including temperature and precipitation. This project seeks to address this gap in the literature by exploring cryptosporidiosis seasonality and its relationship with temperature and rainfall across a variety of climate regions in the US, and by comparing results for the US to those found using data from a global literature review. Results from this study could provide vital information for generating projections of cryptosporidiosis seasonality and incidence for various climate scenarios.

Literature Review

Cryptosporidiosis

Cryptosporidium Parasite

The infectious agent of cryptosporidiosis, the coccidian parasite *Cryptosporidium*, has been a known cause of infectious diarrheal disease in humans since 1976.¹⁵ It can be found globally, making this disease and its associated morbidity and mortality ubiquitous.^{16, 17} The parasite *Cryptosporidium* has around 20 known species (spp.), of which two most commonly affect humans: *Cryptosporidium hominis*, which is most often spread anthropomorphically, and *Cryptosporidium parvum*, typically found in young calves, which has a zoonotic transmission.¹⁸ Other species found in human hosts include *C. meleagridis*, *C. felis*, *C. canis*, *C. muris*, and *C. andersoni*, indicating that *Cryptosporidium* is not entirely host-specific. It is only within the last decade that species typing has become possible with molecular typing techniques such as Polymerase Chain Reactions (PCR).¹⁷ Species other than *C. hominis* and *C. parvum* are most often seen in HIV-positive patients or in children from developing countries.¹⁹

Cryptosporidiosis Epidemiology

Cryptosporidium is spread largely via fecal-oral transmission routes, through contact with human or animal feces. It is highly resistant to chlorine due to its robust oocyst form. Coupled with a low infectious dose of 9-10 oocysts, a high density of up to 10^9 oocysts excreted per gram of stool, and persistent oocyst shedding up to two months post symptom resolution, this pathogen creates a difficult public health problem globally.¹⁷ Additionally, the GEMS study found that between 30-40% of cryptosporidiosis cases are asymptomatic, which makes controlling transmission more difficult.⁶ Children under five are most often the victims of the disease, though people of all ages are susceptible.^{20, 21}

The two species of *Cryptosporidium* that most often infect humans appear to have slightly different epidemiologies, transmission patterns, and risk factors, though their clinical manifestations are the same.

Clinical Symptoms of Cryptosporidiosis

Symptomatically, cryptosporidiosis patients can present with prolonged diarrhea (over two weeks), vomiting, nausea, abdominal cramps, anorexia, and low fever.²² In an immunocompetent host, cryptosporidial diarrhea typically resolves itself within a few weeks and is fairly harmless. Immunocompromised individuals, however, are at a high risk of becoming malnourished, of wasting, and of death due to this opportunistic pathogen, making it a disease of high morbidity and mortality in that population.²³ Young children in developing countries, particularly those already malnourished, are also at risk. For this population, diarrhea can last longer than 14 days, and can result in severe malnutrition, growth faltering, and developmental delays, particularly if they are infected during the first year of life.^{20, 24, 25}

Cryptosporidiosis Treatment and Prevention

Treatment for cryptosporidiosis is largely symptomatic, prescribing fluid intake for dehydration and painkillers for pain management.²² A single FDA-approved drug, nitazoxanide, is available in the United States for children over one year of age and adults. Nitazoxanide has been shown to be the most effective drug in treating cryptosporidiosis in immunocompetent populations.^{26, 27} It is not commonly used in developing countries, however, and has shown no effect on the immunocompromised population, leaving those who suffer the highest morbidity and mortality from cryptosporidiosis with no recourse.²⁸

Since no vaccine for cryptosporidiosis exists to date, current preventative measures are very general and similar to preventative measures for all waterborne disease. The notable exception is chlorine usage, which is highly ineffective against all *Cryptosporidium spp.* Several household water treatments are effective and cheap ways to neutralize the parasite in water used for drinking and cooking. These can range from using filters, especially those with pore diameters of less than 1 micron,²⁹ to solar disinfection (SODIS),³⁰ to simply boiling water for one minute.²²

In more high-tech situations preventative measures include improving water and sanitation infrastructure, complex water filtration systems and more recently, UV light sterilization in drinking water treatment plants. In regards to recreational water, the main source of cryptosporidiosis cases in the US, CDC recommends: excluding anyone who has had diarrhea within the past 2 weeks from the pool, showering before swimming, washing children thoroughly with soap and water after they use the bathroom or their diaper is changed, checking diapers often and changing them in the bathroom and not by the pool.¹⁶

Risk factors for Cryptosporidiosis

Cryptosporidiosis is often studied in outbreak settings, particularly in developed countries where diarrheal disease is not endemic. Therefore a large portion of the literature on cryptosporidiosis risk factors only address outbreaks, a small portion of total cases. Some outbreak-associated risk factors are exposure to drinking water, recreational water, calves, and unpasteurized apple cider.³¹⁻³³ Outbreaks are often related to unpredictable, localized point source contaminations of food or water that can be easily identified and eliminated. In contrast, endemic cryptosporidiosis appears more dependent on long-term environmental and social influences that are not easily managed. Several risk factors have been shown to be associated with sporadic cryptosporidiosis, namely drinking un-boiled water, international

travel, contact with cattle, swimming in fresh water, and contact with children suffering from diarrhea.^{34,35} While these associations have been made, the epidemiology of sporadic cryptosporidiosis and the driving forces behind its seasonality are not yet fully elucidated.

Seasonality

In Temperate Regions

The 'temperate' climates of the earth are between 23.5° and 66.5° latitude in both the northern and southern hemisphere. There are generally four seasons during the year in the temperate region: spring, summer, fall, and winter.

The previously published global systematic review and meta-analysis of cryptosporidiosis found a slight spring peak in mid-latitude climates, and a slight fall peak in cold temperate climates.¹⁰ A systematic review of studies conducted specifically on developed countries in temperate regions demonstrated an overall dual-peak seasonality for cryptosporidiosis over all regions analyzed, and also separately in the United Kingdom (UK), Canada, and Oceania. These researchers did not find a dual peak in the US or mainland Europe, but this may have been due to a lack of studies in those areas that met their selection criteria.¹¹

Two studies from the US found a single summer/autumn peak without the additional spring peak. Naumova *et al.* report that their Massachusetts state data failed to find a secondary seasonal peak in the spring. They did, however, observe the strong late summer/autumn peak, noting that it was stronger in children than in adults.³⁶ Naumova *et al.* suggest this pattern is due to their passive, and therefore under-reporting, surveillance system and patient population of more than 60% children, who are most likely to engage in swimming activities and poor hygiene than adults.³⁷ Another study performed in one Massachusetts hospital found the same summer/autumn peak, and though they report cases

in February and March, their dataset is not continuous or large enough to be able to identify the slight increase as a second ‘peak’.³⁸ In contrast, a similar state-wide study on cryptosporidiosis in Oregon did display dual peak seasonality, with a primary peak in late summer/autumn, and a secondary peak in early spring (February and March).³⁹

Studies in the UK report peaks in both the late summer/autumn time period and also strong spring peaks. Two studies, one performed in Liverpool, England and one in Galway, Ireland, both found a single strong peak in the spring months. Both studies suggest that this is due to the rural and largely agricultural nature of the studies’ catchment areas.^{40, 41} Other studies from the UK show the more common dual peak seasonality.⁴²⁻⁴⁶ One study from Scotland that genotyped its cases attributes the dual peak seasonality to the two separate species (*C. parvum* – spring, *C. hominis* – fall) that make up most of cryptosporidiosis in humans.⁴⁷ A study in Korea presents a similar dual peak, in the spring and fall.⁴⁸

In New Zealand (NZ), cryptosporidiosis seasonality is characterized by a large peak in the spring months, or August through October, and a peak in the autumn months of February through April that is about one third as large as the spring peak.⁴⁹ This pattern is intriguing, as the NZ seasonality matches cryptosporidiosis seasonality in temperate areas of the Northern hemisphere where month is concerned, but the seasons are reversed. In the UK the autumn peak is 20% larger than the spring peak.⁵⁰ In both locations the spring peak appears to be associated with calving season, when there is a sudden increase in young calves. This increase in calves effectively increases the number of carriers of the zoonotically-transmitted *C. parvum*, which infects humans via direct contact with the calves or in agricultural run-off from manure.^{51, 52} Similarly, the autumn peak in both locations is associated with anthroponotically-transmitted *C. hominis*, and closely matches the swimming

season and warmer temperatures, which lends support for recreational water use as a potential means of transmission.^{37, 51}

A study performed in South Africa found a seasonal increase in the late summer and early autumn months of January through March, during which time it is also fairly rainy.⁵³ Another study performed on South African populations found similar results,⁵⁴ although a different study found case incidence to be highest in the warm dry months of December to March.²³ This discrepancy may be due to a difference in study location and therefore climate: Fripp *et al.* and Steel *et al.* both used patients from the same hospital near Pretoria, while Nel *et al.* performed the study on the Western Cape, on the opposite side of the country. Nel *et al.* also used a population of with high HIV prevalence, which could significantly affect their results.

In Tropical Regions

The ‘tropical’ climates of the world are between the latitudes of 23.5° N and 23.5° S. Seasonality in this region is typically defined by two seasons, ‘wet’ and ‘dry’.

The previously published global systematic review and meta-analysis of cryptosporidiosis found that cryptosporidiosis seasonality in moist tropical climates was highly varied. The authors found that arid regions, most of which are also in the ‘tropics’, had no seasonal peak in cryptosporidiosis rate.¹⁰

In a Guatemala study the highest incidence of cryptosporidiosis among infants was at the end of the dry season, between February and May of the calendar year. Cruz *et al.* suggest this is due partly to the lack of water supplies and also to the increase of dust, which aerosolizes *Cryptosporidium* oocysts, leading to respiratory infections.⁵⁵ Two studies in Bangladesh, where cryptosporidiosis peaks April through July, support this theory of part of

the cryptosporidiosis burden being due to airborne infection. In Bangladesh, April through July are hot and humid months of the year.^{56, 57}

A study conducted in Uganda indicates that in some tropical countries, the seasonal peak of cryptosporidiosis is during months of the wet, or rainy, season.⁵⁸ In Nigeria as well, Cryptosporidiosis prevalence was found to be highly associated with the rainy season, particularly in HIV+ populations.⁵⁹ This seasonal peak during the rainy season is also found in Burkina-Faso,⁶⁰ and in Guinea-Bissau, where peak cryptosporidiosis occurs at the start of the rainy season.⁶¹ Results from studies in Brazil,⁶² The Gambia,⁶³ Zambia,⁶⁴ Malawi,⁶⁵ Mexico,⁶⁶ Costa Rica,⁶⁷ India^{68, 69} all show similar findings.

A study in Kenya found the dual-peak *Cryptosporidium* seasonality, which is similar to that seen in temperate regions. In the case of Kenya, however, the two peaks were found in during the dry seasons after the rainy periods. The larger peak was found during the hottest and driest time of the year, after the short rainy season. The smaller peak was found during the cold, dry season after the long rainy season.⁷⁰ This pattern is unusual, but the study only spanned two years, making it difficult discern whether or not this pattern is typical or due to unusual weather during those two years.

Climatic Drivers

Temperature and Rainfall as Climatic Drivers

The global surface temperature has risen 0.8 °C in the past century, with 0.6 °C of that in the past thirty years.⁷¹ Temperatures are expected to increase further, which will lead to an increase in the amount of energy in the atmosphere, causing more storms of increasing severity and more extreme weather events.⁷² Floods and droughts have become more prevalent and hurricanes as well as winter storms have become more powerful and devastating. Sea levels are also rising, having risen 8 inches in the past century, with an

expected rise of 1-4 feet in the next century.⁷³ These changes will have an impact on human health, including waterborne diarrheal diseases such as cryptosporidiosis.

In light of these environmental changes, there have been several studies that explore the relationship between infectious diarrheal disease and the aforementioned climatic drivers.

Infectious Diarrheal Disease

Temperature

A World Health Organization (WHO) study associating degree temperature change with percent increase of diarrhea put forth a conservative estimate of a 5% increase in diarrhea per 1 °C increase in temperature.⁷⁴ Another study projecting how diarrheal disease might increase with projected temperature increase found a predicted 8-11% increase in diarrheal disease depending on geographic location⁸. In general the literature consensus seems to point to a significant association between increases in temperature and increases diarrheal disease,⁷⁵⁻⁷⁹ though at least one study found significant correlation, but not significant predictive power.⁸⁰ One exception to the rise of diarrheal diseases with rising temperatures is the set of viral diseases, such as Norovirus and rotavirus, which tends to show a significant inverse relationship with temperature that echoes the viruses' winter seasonalities.^{9, 81, 82}

Precipitation

The literature is less unanimous on the association between diarrheal disease and precipitation, and can vary even within the same climate region. For example, several studies performed in the tropics, where there are wet and dry seasons, explored the relationship between diarrheal disease incidence and rainfall, each providing different results. A study of bacterial diarrheal diseases in Vietnam showed a weak positive association between rainfall

and disease incidence, with a strong difference in precipitation between ‘high’ and ‘low’ disease periods.⁸³ Similarly a longitudinal study in Botswana revealed no relationship between precipitation and diarrheal disease incidence, but a weak positive association with the previous month’s precipitation after controlling for season.⁸⁴ In Thailand there was no direct effect of precipitation on disease incidence, though the study does suggest that rainfall may impact various points on the fecal-oral contamination route. These points included change in water availability, change in water use patterns, and contamination of food.⁸¹ In Ecuador, the relationship between diarrheal disease incidence and heavy rainfall events differed based on pre-existing rainfall conditions. If the weeks preceding the rainfall event were dry, the extreme rainfall was positively associated with diarrheal disease. If the weeks preceding the rainfall were wet, diarrheal disease was negatively associated with disease incidence.⁸⁵

Members of the Ecology and Health Research Centre in Wellington, NZ attempted to incorporate disease and precipitation patterns over a larger geographic and temporal area in their 2001 study of diarrheal disease in the Pacific Islands. They found that rainfall below the 50th percentile was positively associated with an increase in diarrheal disease, while the relationship between diarrheal disease and rainfall above the 50th percentile was modified by a one-month lag. Diarrheal disease increased with high rainfall during the same month but decreased with high rainfall in the previous month.⁷⁶ A study of non-cholera diarrheal disease performed in Bangladesh found similar results. Diarrheal disease increased by approximately 5% for every 10 mm of rainfall above the average monthly threshold of 52 mm, but also increased by approximately 4% for every 10 mm of rainfall below an average monthly threshold of 52 mm. However, Hashizume *et al.* also found that when these estimates were adjusted by water level of a nearby river, nearly all of the association between rainfall and diarrheal disease incidence disappeared.⁷⁵

The literature review failed to find studies performed on the incidence of diarrheal disease in general in relationship to rainfall in temperate regions. This lack of research is potentially due to the fact that many countries in temperate regions are developed and therefore have generally low incidence of diarrheal disease. A global study, which included temperate, tropical, and arid climate regions found that diarrheal disease decreased by 4% when average monthly precipitation increased by 10 mm.⁸⁶ In this study, Lloyd *et al.* hypothesize that this decrease is due to low levels of precipitation in the long term causing water scarcity, and leading to poor hygiene behavior. Given the lack of consensus in the literature, it appears the relationship between rainfall and diarrheal disease is highly modified by surrounding geographic and temporal conditions.^{7, 85}

Extreme Precipitation

An association often explored along with temperature and rainfall is the association between diarrheal disease and extreme weather events such as El Niño or flooding.

A longitudinal study of the relationship of the effect of extreme precipitation on waterborne disease outbreaks in the US showed a significant association. Outbreaks caused by contaminated surface water was found to be correlated with extreme precipitation in the month prior to the outbreak, and outbreaks caused by contaminated ground water was found to be associated with extreme precipitation two months prior.⁸⁷ Comparable studies performed on waterborne outbreaks caused by, or associated with extreme rainfall events in Canada, showed a similar relationship between extreme precipitation and diarrheal disease outbreaks.^{88, 89}

In terms of flooding, two studies in Bangladesh found a sharp increase in the number of patients presenting with diarrhea during flood conditions, with levels gradually reducing to normal by four weeks after flood conditions ended.^{90, 91}

Cryptosporidiosis and Climatic Drivers

Temperature

Studies similar to those discussed in the previous section have attempted to explore the relationship between temperature and rainfall and cryptosporidiosis specifically. Jagai *et al.* reported in their global systematic review and meta-analysis that cryptosporidiosis incidence was strongly related to temperature in all climate regions: moist, tropical regions, arid and semiarid regions, mid-latitude regions, and cold temperate regions. However, temperature was only a significant predictor of cryptosporidiosis in tropical, mid-latitude, and cold temperate regions.¹⁰

One study in NZ conducted a time series analysis of the relationship between disease incidence, temperature, rainfall, and the impact of the El Niño cycle. They found cryptosporidiosis case incidence to be positively associated with the average monthly temperature one month prior.⁹² This finding is similar to other research performed in NZ, which found that both current and previous month temperatures were associated with rate of cryptosporidiosis.⁴⁹ This study, performed by Lake *et al.* on data from 1997-2005, only found such an association during the summer and autumn months, and not during the spring. A contradictory study of NZ data from 1997-2006 found that an increase in temperature was connected with a decrease in cryptosporidiosis cases.⁹³ The authors of this contradictory study, Britton, *et al.* attribute this difference to the utilization of different geographic resolutions of the data: Britton *et al.* analyzed data at the census area units, while Lake *et al.* analyzed data at the national level.

Studies from Australia have also found a strong relationship between temperature and cryptosporidiosis case incidence. One study in Brisbane estimated that an increase of 1 °C could lead to an increase of up to 50 additional cases per year.¹² The other two studies

built several different models and found similar strong, positive relationship between temperature and cryptosporidiosis incidence.^{13, 14}

In the Northern hemisphere, a study of cryptosporidiosis cases from Massachusetts, US found that case incidence was strongly associated with the prior month's ambient temperature. Naumova *et al.* attribute this association to recreational water use and other outdoor activities, such as camping, which encourage poor hygiene and keep people in confined spaces like tents.³⁷ In contrast, a study done in Korea found no correlation between temperature and case incidence.⁴⁸ A tri-location study in India found a positive relationship between temperature and case incidence in one city, Delhi, but not in the other two cities of Vellore and Trichy.⁹⁴

Precipitation

Similar to diarrheal disease in general, the relationship with precipitation and cryptosporidiosis is vague. The global systematic review and meta-analysis reports the following for the relationship between precipitation and cryptosporidiosis: in moist tropical areas, cryptosporidiosis was only correlated with one of two annual precipitation peaks, and is marginally associated with precipitation in cold, temperate regions. Cryptosporidiosis was only accurately predicted by precipitation in the moist tropical climate regions.¹⁰

Studies in NZ found no relationship between cryptosporidiosis case incidence and precipitation. Lake *et al.* hypothesize this lack of relationship is due to the inability of a singular precipitation measurement to be representative of the entire country.⁴⁹ Lal *et al.* suggests the relationship between precipitation and cryptosporidiosis is likely modified by drinking water quality, something they did not account for in their study.⁹² Another study in NZ utilized a 'rainfall to evaporation ratio', and found that for every unit increase of the ratio (every increase in rainfall or decrease in evaporation rate), cryptosporidiosis increased

by 1%. It is difficult to discern whether or not this information is comparable to Lake *et al.* or Lal *et al.*⁹³

A set of studies based on a dataset from Brisbane, Australia gave contradictory results. One set of models did not find a relationship between precipitation and case incidence,¹² while another set of models found an inverse association, indicating that cryptosporidiosis decreased with increasing rainfall.¹³ A third set of models built with this dataset found that precipitation created a weak interaction in the relationship between socio-economic status and cryptosporidiosis incidence.¹⁴

Turning to the northern hemisphere, a study in the northwestern part of England found a positive association between cryptosporidiosis and rainfall, with an increase of 27% when the rainfall of the previous week was over the 75th percentile.⁴³ Similarly, a South Korean study found a weak association between monthly rainfall and monthly case prevalence.⁴⁸

In the subtropical regions of Africa and South America, it can be assumed that there is exists some relationship between cryptosporidiosis and rainfall given that many studies indicate rainy season seasonal peaks.^{58, 60-67} However, since these studies did not examine direct associations between incidence and rainfall, it is difficult to quantify or qualify that relationship.

In India the relationship between precipitation and cryptosporidiosis is unclear. One study in Kolkata found a positive relationship between case incidence and precipitation.⁶⁹ Three other studies in India, one in Kolkata,⁹⁵ one in Varanasi,⁶⁸ and one in Vellore,⁹⁶ found an association between cryptosporidiosis and the rainy season, which implies some relationship with rainfall. However, another study, that took place in three locations (Delhi, Trichy, and Vellore), found no relationship between precipitation and case incidence in any

of the cities.⁹⁴ This difference in relationship could potentially be due to differing climates in each study's site, although the opposite relationship between case incidence and rainfall in the two studies in Vellore would be interesting to explore.

Extreme Precipitation

The largest cryptosporidiosis outbreak documented in the United States was an outbreak preceded by extreme precipitation, where the excessive turbidity of the water overwhelmed the local drinking water treatment plant.⁹⁷ One study in Peru have linked diarrheal disease in general and cryptosporidiosis specifically with the 1997-1998 El Niño weather event. Bennett et al found a 55% increase in general diarrheal incidence during an El Niño event, and a 52% increase in cryptosporidial diarrhea incidence specifically.⁹⁸

With the exception of a few studies,^{10, 76} most of the studies on the relationships between cryptosporidiosis case incidence, temperature and rainfall were performed in a single geographic location. This study attempts to encompass a much larger and more varied group of climate locations, both around the world and across the US. This study seeks to explore the relationship between cryptosporidiosis seasonality and temperature and rainfall only. Data pertaining to extreme weather events such as El Niño, drought, or flooding were not searched for explicitly, nor were they included in analysis.

National Surveillance Data

According to the WHO, public health surveillance data is the “continuous, systematic collection, analysis and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice”.⁹⁹ Many countries are now capable of collecting infectious disease case data on a national level, including Scotland, Germany, England, France, whose data was included as part of the systematic literature review. Data can be collected at a higher level as well: a surveillance center sponsored by the

EU and run by Swedes collects data from multiple European countries (<http://www.ecdc.europa.eu/>). This data centralization allows for analysis on a larger scale than studies collecting data from one or more point locations, creating a broader picture of disease patterns. In practice, disease incidence reports are collected on a local level first—it is the obligation of the physician or laboratory staff that discovers the case to notify their local health authority. The quality and quantity of data is dependent on whether or not the data is actively or passively collected. Active surveillance systems seek out case data in healthcare settings such as hospitals and are more resource-intensive. Passive systems do not require state or local entities to report case data the cases come to them, only collecting cases that seek healthcare and get reported. In general, active systems yield more cases than passive systems, which are more likely to under report cases.

National Notifiable Disease Surveillance System

The US national surveillance data was chosen for this study since the US has a large geographic area that is easily subdivided and a wide variety of average temperatures and rainfalls that are representative of the many of the world's climates. This variety of climates and the size of the US make the National Notifiable Diseases Surveillance System (NNDSS) dataset an appropriate comparison to a systematic-review based meta-analysis. Currently all 50 states, New York City, and Washington D.C. report cryptosporidiosis cases to the Centers for Disease Control and Prevention (CDC).¹⁰⁰ In the US each state is responsible for its own surveillance system rules and requirements, which vary based on the budget and time constraints of each state. State-level reporting is required, but reporting to the Federal level is voluntary, making the NNDSS data 'passively' collected data, even if the cases are actively collected on the state level.¹⁰¹ It should also be noted that it is not currently required to test

for *Cryptosporidium* in stool samples submitted for routine examination, making it even more likely that cases are being underreported.¹⁶

Methods

Hypothesis

We hypothesized that there is a correlation between change in overall cryptosporidiosis incidence and change in temperature and rainfall, while controlling for latitude, demographics, and season. Additionally, the effects of data type and data collection methodology on our primary hypothesis of interest were explored. Secondly, we hypothesized that the two datasets, the literature review data and surveillance system data, would produce similar models of the effect that climate factors have on cryptosporidiosis incidence.

Outcome Variables

Systematic Review

Relevant studies were obtained through a literature search using the PubMed database, using search string ‘cryptosporidium or cryptosporidiosis’ coupled with the following search terms: ‘ambient temperature’, ‘climate’, ‘rain’, ‘relative humidity’, ‘season*’, and ‘weather’. ‘Season*’ was used to search all variants of season (seasons, seasonality, etc.) simultaneously. There were no language restrictions on the search. All PubMed articles were imported into EndNote X6 (Thomson Reuters; New York, NY) and duplicate articles were removed. Two reviewers separately examined the title and abstract of each article and studies were excluded based on the following exclusion criteria:

- Non-human studies
- Studies of fewer than 25 cases
- Studies on biology or lab methods
- Studies that only included HIV+ patients

- Studies that did not include a full year (12 months) of data
- Studies on outbreak data only

Examination of the reference lists of key cryptosporidiosis literature reviews¹⁰²⁻¹⁰⁹ provided additional studies, and a dataset from a previously published meta-analysis on cryptosporidiosis seasonality was obtained with permission of the author.¹⁰ These additional two sources were also scanned for duplicates and assessed based on the above exclusion criteria. All papers with titles and abstracts appearing to meet criteria were downloaded for full review. A more detailed description of the search protocol is provided in Appendix A.

In order to maintain data consistency and properly correlate number of cases to degree temperature and rainfall, only studies presenting monthly count data were used. For articles that were not in English, appropriate text and figure/table titles were translated and data was extracted if the paper met the inclusion criteria. For articles that did not present monthly case numbers but presented graphs or tables with monthly prevalences or incidences, the monthly case count was back-calculated if possible. Authors were contacted to request monthly case numbers for each year for studies that did not display monthly case numbers of human cryptosporidiosis cases for each year of the study, or for which back-calculation of case numbers was not possible, if the studies appeared to meet all other criteria. Authors were also contacted if their published graph was not well enough resolved for extraction. Outbreak and immunocompromised cases were included in the analysis if they came from a surveillance dataset where outbreak or immune status was recorded, but not specifically selected for. Datasets used in multiple studies were only added to the final dataset once. Any studies that used FoodNet Data or data from the NNDSS were also excluded to prevent overlap with the comparison NNDSS dataset.

The flow chart of inclusion/exclusion is shown in Appendix B. Meta-data were collected from the included studies, and monthly case numbers were extracted from tables and graphs using Plot Digitizer.¹¹⁰

The predictor variables analyzed with the above literature review outcome data were obtained and analyzed by a second analyst, and the results are included in this study as a comparison to the NNDSS data. For details pertaining to predictor variable compilation and data analysis please refer to the methodology described by Ahmed *et al.*⁹

National Notifiable Diseases Surveillance System

The cryptosporidiosis case data from the NNDSS data was requested from the Division of Integrated Surveillance Systems and Services (DISSS) in the National Center for Public Health Informatics (NCPHI) at the CDC using a Registration Information and Data Use Restrictions Agreement (RIDURA) form (included in Appendix C). Access was given to the public aggregate monthly count files of total counts of confirmed and probable cryptosporidiosis by state for each year available. More geographically detailed data was not available for access.

Cryptosporidiosis has been a national notifiable disease since 1995, when 24 states began reporting cases voluntarily to the CDC.¹¹¹ Cryptosporidiosis data was available from 1997-2011. The 2012 data was unavailable for public analysis, as the CDC had not published it at the time of this analysis. Data from 1995 and 1996 were also unavailable, as they remain unpublished. Cryptosporidiosis cases are initially collected by healthcare providers and laboratories, which are required to report cases to the state and local level within a number of days, depending on each state. Currently, all 50 states and two metropolitan areas (Washington DC and New York City (NYC)) voluntarily report cryptosporidiosis cases to the CDC.

While reporting systems can vary from state to state, the CDC, collaborating with the Council of State and Territorial Epidemiologists (CSTE), first published a set of case definitions for notifiable diseases in 1990, to make the reporting process more uniform and comparable across state surveillance systems. Cryptosporidiosis was first added to the list in the 1997 report, where probable cases were defined as ‘a clinically compatible case that is epidemiologically linked to a confirmed case’ and confirmed cases were defined with laboratory confirmation. The clinical symptoms have been described previously in the introduction of this paper. Laboratory criteria include evidence of *Cryptosporidium* oocysts or antigens in stool samples or biopsy specimens. This definition was updated in 1998 to include laboratory-confirmed asymptomatic cases and identification of antigens by polymerase chain reaction (PCR) techniques, in 2009 to include evidence of *Cryptosporidium* nucleic acid and to require reporting of molecular characterization and species when available.^{112,113} In 2011, laboratory criteria for diagnosis was delineated into probable and confirmed when the rapid immunological assays were found to produce false positives 44% of the time.¹¹⁴ Thereafter, a case that was diagnosed by rapid assay was considered probable unless further testing was done using an established laboratory method (direct fluorescent antibody testing (DFA), PCR, or enzyme-linked immunosorbent assay (ELISA)).¹¹⁵

Given the passive nature of the NNDSS, and the fact that cryptosporidiosis is not routinely tested for in stool samples, it is highly likely that the dataset under represents the actual number of cases. Hawaii and Alaska, for example, only report 7 and 15 cases total and were therefore excluded from analysis. In addition, PCR techniques have only been used recently to identify *Cryptosporidium* species and species is not recorded in the NNDSS, therefore, it is impossible to analyze the distinct epidemiologies of each parasite species.

As NYC data is recorded separately from New York State in the NNDSS, NYC cases were added to the New York State cases for easier analysis. The other metropolitan area to collect cases is the District of Columbia (D.C.), which from here forward shall be referred to as a 'state'. Months of data that were available prior to when a state declared cryptosporidiosis nationally notifiable were deleted so as not to confuse between a 0 case count from non-reporting, and a 0 case count from a month where there were no cases. Additionally, the NNDSS case data was \log_{10} -transformed to meet the 'variance = mean' assumption of the Poisson distribution regression analysis.

Predictor Variables

Climate-related Variables

The primary predictors of interest in this study were monthly average temperature, in degrees Celsius ($^{\circ}\text{C}$), and monthly average precipitation in millimeters (mm). These weather data were obtained from the National Climatic Data Center's (NCDC) Climate Data Online monthly summaries database. These data are averaged from the Global Historical Climatology Network-Daily (GHCN-D) data collected from weather stations across the world¹¹⁶. Monthly temperature and precipitation data was obtained from the city of largest population in each state, as determined by the 2010 census. For cities with data available from several weather stations, the station that encompassed the complete time series (1997-2011) was used. Weather station latitude was also downloaded from the source to be used as a predictor.

The NCDC presents data in the format of tenths of degrees Celsius for temperature and tenths of millimeters for rainfall. Temperature and rainfall data were divided by 10 to obtain results in whole degrees and whole millimeters. In order to observe general patterns beyond the climatologically arbitrary state boundaries and to put the data into a regional

context, each state was assigned to one of nine US Climate Regions, as defined by the NCDC.¹¹⁷ A map of the US Climate Regions is available in Appendix D. These Climate Regions only cover the contiguous US, therefore Alaska and Hawaii were assigned to their own separate climate regions for this study.

The climates of states that span many degrees in latitude would be inadequately described by a measure of temperature or rainfall from only one city. Therefore, Alaska, California, Florida, and Texas, were excluded from regression analysis, though summary weather statistics and season strength were still calculated for these states.

Demographic Variables

To account for various demographics and suspected cryptosporidiosis risk factors, secondary covariates were obtained from three different sources: the American Community Survey (ACS), an annual survey performed by the U.S. Census Bureau to record various economic and social statistics about the American people that are no longer included in the decennial census; the 1990 decennial census, the last census to collect water and sanitation data at the household level; and the 2003 Rural-Urban Continuum Codes (RUCCs), a system used by the Office of Management and Budget (OMB) to define metropolitan (metro), urban, and rural counties by size and adjacency of urban/metro area.

Statewide percentages of families and people whose income in the past 12 months was below the poverty level were obtained from the 2010 ACS five-year estimates.¹¹⁸ Statewide percentages of total well use (drilled and dug wells combined) were obtained from the 1990 decennial census.¹¹⁹ Statewide percentages of urban counties were calculated using RUCC data,¹²⁰ by counting the number of counties in the state considered 'urban' (counties with urban populations of 20,000 or more and adjacent to or part of a metropolitan area), and dividing by the total number of counties in the state. The 'urban' classification was used

instead of the OMB's 'metro' classification in order to include counties that are not considered metro but have large populations and highly developed infrastructure indicative of non-rural areas.

Methods of Analysis

The general methods of analysis used in similar studies of infectious diarrheal disease agents such as Rotavirus and Norovirus have been described previously^{9, 82} and was performed separately for the literature review dataset. This analysis was adapted and expanded for the NNDSS data and performed using SAS version 9.3 (SAS Institute Inc.; Cary, NC). These methods are described below.

Summarizing Data and Seasonality Analysis

The monthly case, temperature, and precipitation data were averaged over all years to create the monthly mean for each month for each state, and for each climate region. These monthly means were averaged to obtain yearly average temperature estimates, or summed to obtain yearly average precipitation or case count estimates for each state and climate region.

Monthly temperature, precipitation, and case averages were normalized by peak values, allowing for comparison of an average year across regions. The normalized monthly averages of cases, temperature, and rainfall were then plotted by month to obtain a long-term seasonal curve for each climate region.

Season strengths, defined as the peak-to-mean ratio, of cryptosporidiosis, temperature, and precipitation were calculated using the ratio of monthly averages to their respective averages per month. A heat map of cryptosporidiosis season strength was created using ArcGIS version 10.1 (Environmental Systems Research Institute; Redlands, CA) in order to visualize spatial variability in season strength. Season strengths of cryptosporidiosis

were plotted against the season strengths of temperature and rainfall as an initial test of association.

Correlation between each of the demographic variables was assessed in order to prevent multicollinearity in the model.

Statistical Analysis

Generalized log-linear Poisson regression models (GLM) were used to test the association between monthly cryptosporidiosis incidence and temperature and precipitation. Cryptosporidiosis case counts were transformed by log base-10 in order to meet the Poisson assumption of an equal mean and variance. For each state, two sets of GLM models were created: one presenting case counts as a function of temperature, and one presenting case counts as a function of precipitation. Incidence rate ratios (IRRs) from these models were plotted in a forest plot organized by climate region in order to visualize heterogeneity across states. The same modeling process and forest plot graphing was performed at the climate region level in order to assess heterogeneity across climate regions.

Heterogeneity between states and between climate regions was assessed visually by examining the forest plots for overlap of confidence intervals, and statistically with I^2 statistics. These statistics were assessed at $\alpha=0.05$ using a method described by Neyeloff *et al.* and modified for this study.¹²¹

All state data were also pooled to assess the overall relationship between cryptosporidiosis and temperature and precipitation for the United States. This was performed with a generalized estimating equation (GEE) model, using a Poisson distribution and the auto-regressive correlation structure and clustering by state. Cryptosporidiosis cases were initially modeled as a function of both monthly temperature and one month lagged monthly temperature, as well as precipitation and one month lagged precipitation. The IRR,

Wald Statistic, and p -value were used to evaluate which of the lagged data (0 or 1-month) had the largest effect on cryptosporidiosis incidence. Cryptosporidiosis cases were also modeled univariately as a function of each of the three demographic covariates: percent well use, percent poverty, and percent urban. This variable selection was done to avoid collinearity among the demographic variables and to include the variable with the largest effect on cryptosporidiosis incidence in the model. A backwards selection model was used to select the final model, starting with all variables of interest (lagged temperature, precipitation, latitude, percent well use, percent urban, percent poverty) and removing variables with p -values of greater than 0.05 until only significant variables remained in the model.

Upon review of the GEE modeling results, the well use variable was divided into deciles, and the same regression techniques were used to obtain IRRs for temperature and precipitation in each well use decile. These IRRs were also plotted in forest plots in order to examine patterns, and heterogeneity was assessed using the same methodology previously mentioned.

The forest plots of GLM models and the results of the final GEE model were compared with similar models built with the literature review data to observe the differences across datasets. Two models were built with the literature review data: one using all studies and one using only studies from developed countries in the northern hemisphere for a more appropriate comparison to the US.

Permission has been granted to utilize all data not publically available. Institutional Review Board (IRB) approval was declared unnecessary for this study, as it failed to meet the definition of human research. The Letter of Declaration from the IRB stating the above can be found in Appendix E.

RESULTS:

Systematic Review

The literature review search produced 970 articles, 353 of which were duplicates. Of the 617 remaining articles, 535 were excluded based on their title and abstract not meeting the exclusion criteria. Examination of the reference lists of eight cryptosporidiosis literature reviews provided an additional 113 studies. The dataset from the published meta-analysis on cryptosporidiosis seasonality provided 66 additional studies, all of which were subject to the same reviewing process as the articles from PubMed and the literature review reference lists.¹⁰ Of all 718 articles screened by title and abstract, 183 articles were downloaded for full review. Of these 183, 70 were excluded for the following reasons: 18 contained fewer than 25 cases, 12 collected data from too large of an area to accurately match temperature and rainfall (low geospatial resolution), 3 were water quality studies, 3 only studied HIV+ patients, 9 were not about cryptosporidiosis or addressed modeling and not cases, 18 did not span 12 months, 3 exclusively studied outbreak data, and 4 were reviews or letters to the editor. Of the 61 authors contacted to provide useable data, 10 authors were unable to share data due to it being unavailable to them, 23 authors could not be contacted by email, and 28 authors did not respond to email contact.

The final dataset included the 42 studies with at least one of the following characteristics: the study met all necessary inclusion criteria, the authors shared original data that met criteria,^{36, 39, 48, 94, 98, 122-126} or the authors shared an alternate data source from which we were able to extract data.^{49, 127}

There were 46,522 cases of cryptosporidiosis reported from 1982-2012, spanning six continents, and coming from a variety of sources including hospitals, labs, community-based studies, and countrywide surveillance systems. Study designs included cross-sectional studies,

cohort studies, case-control studies, and surveillance studies. Methods of detection varied as well, including several different staining techniques, ELISA, and PCR. Further meta-data describing these studies are shown in the Appendix F.

National Notifiable Disease Surveillance System Data

There were 49,045 cases of cryptosporidiosis with an associated symptom onset date reported to the CDC through the NNDSS between 1997 and 2011. Total counts of cryptosporidiosis cases in each state range from seven cases (Hawaii) to 6,369 cases (Wisconsin). For each climate region, total numbers ranged from 291 cases in the West climate region, including Nevada and California, to 11,013 cases in the Upper Midwest climate region, which includes Iowa, Michigan, Minnesota, and Wisconsin. For the entire country there was an average of 3,270 cases per year.

Descriptive statistics and characteristics are displayed in Table 1, in which the states are organized by climate region. Average yearly temperature ranged from 5.87 °C in North Dakota to 23.10 °C in Arizona. Average yearly rainfall ranged from 125 mm in Nevada to 1,495 mm in Louisiana. Latitudes and names of the cities where the NCDC weather stations are located and demographic variables for each state and climate region are also displayed in Table 1.

Season Strength

The distribution of season strength, or peak-to-mean ratio, across the US is displayed visually in Figure 1 and numerically in Table 1. Season strength varies across the states, with higher peak-to-mean ratios seen in the middle and western parts of the country, and lower in the eastern regions. Overall, the US has cryptosporidiosis season strength of 2.72. Connecticut, Kansas, and Missouri showed the greatest season strength of the states, with peak-to-mean ratios of 4.89, 4.10 and 3.96 respectively. The lowest season strength was seen

in Georgia, Maryland, and Washington with peak to mean ratios of 1.57, 1.62, and 1.64 respectively. Of climate regions, the South showed the strongest season strength at 3.41, and the Southeast states showed the weakest season strength at 1.94. Plots of cryptosporidiosis peak-to-mean ratio against temperature and precipitation peak-to-mean ratio showed no definitive pattern. These plots are available in the Appendix G.

Seasonality across US Climate Regions

The average monthly case distribution of cryptosporidiosis in the United States is shown in Figure 2, along with average monthly temperature and precipitation distributions. Cryptosporidiosis displays clear summer seasonality in the US as a whole. A large peak begins in June and last until October, reaching its height in August. The cryptosporidiosis case peak falls one month after the temperature peak (July), and two months after the precipitation peak (June) (Figure 2).

In Figure 3 the distribution of average cases, temperature, and precipitation per month is displayed for each US climate region. The large peak of cases in the summer/early autumn months (July-October) is evident in all climate regions shown in Figure 3. These peaks reach maximum height in August, similar to the countrywide peak in Figure 2. The West, Northwest, and Northern Rockies and Plains regions all show a second, distinct, much smaller, seasonal peak during the spring months.

Temperature seasonality shows consistent summer-peak seasonality across the US, with highest temperatures in July and August in all regions. Northern regions show slightly colder temperatures on average, and southern regions slightly warmer temperatures on average, as expected. The seasonal cryptosporidiosis case curves seem to be positively related to the temperature curves in all regions. This is supported by the GLM relationship values displayed in Table 1.

Precipitation seasonality is more varied across the regions. The Ohio Valley, Upper Midwest, South, the Northern Rockies and Plains, and Southwest show a seasonal peak in the late spring/early summer months of April – June. The Northwest and West regions have seasonal precipitation peaks in the winter months of November- February. The Northeast region shows little variation in precipitation across the year, and the Southeast region shows peak precipitation in late summer, early autumn, similar to the temperature curve. The relationship between cryptosporidiosis and precipitation is similarly variable across climate regions. The West and Northwest regions seem to show a negative relationship between precipitation and cryptosporidiosis, while the Southeast seems to show a positive association. Most of the remaining regions seem to have a positive relationship with the precipitation of two months prior. These visual relationships are supported by those found through GLM modeling that are reported in Table 1, even though not all are statistically significant.

Generalized Linear Modeling

United States and US Climate Regions

The log-linear Poisson regression analyses for each state show a clear positive relationship between temperature and cryptosporidiosis case incidence in all states, with 27 of the 46 states analyzed (59%) showing a statistically significant relationship between temperature and cryptosporidiosis incidence. IRRs for the contiguous US range from 1.01 in Georgia (95% CI 0.99, 1.03), Kentucky (95% CI 0.96, 1.06), and Washington, D.C. (95% CI 0.95, 1.08) to 1.11 (95% CI 0.98,1.25) in Connecticut. Temperature IRRs with 95% confidence intervals for each state, grouped by climate region, are displayed in Figure 4a.

The same individual GLM analyses of each state show weak, highly variable relationships between cryptosporidiosis and precipitation. Of the 46 states analyzed, 34 show a positive relationship between cryptosporidiosis and precipitation, of which two (6%) are

statistically significant. Twelve states show negative relationships between cryptosporidiosis and precipitation, none of which are significant. One state, Tennessee, shows no relationship between cryptosporidiosis incidence and precipitation (IRR=1.000, 95% CI 0.997, 1.004). No positive IRRs exceeded the 1.010 (95% CI 0.999, 1.021) value of Wyoming, and none of the negative IRRs exceeded the 0.994 (95% CI 0.973, 1.016) of Nevada. Precipitation IRRs with 95% confidence intervals for each state, grouped by climate region, are displayed in Figure 4b.

The log-linear Poisson regression analyses on data aggregated to the US climate region level show similar results as the analyses on the state level data. All nine climate regions show a statistically significant positive relationship between cryptosporidiosis and temperature, with IRRs ranging from 1.01 (95% CI 1.001-1.020) in the Ohio Valley, the Southwest, and the Southeast to 1.04 (95% CI 1.004-1.083) in the West. The temperature IRRs for each climate region are shown in Figure 5a.

Five of the nine climate regions show a positive relationship between cryptosporidiosis incidence and precipitation, of which three (Northeast, Upper Midwest, Upper Rockies and Plains) are statistically significant. Of the four climate regions showing negative relationships between cryptosporidiosis incidence and precipitation (Ohio Valley, South, Northwest, and West), none are statistically significant. Positive precipitation IRRs range from 1.0004 (95% CI 0.999, 1.002) in the Southwest to 1.005 (95% CI 1.004, 1.007) in the Northern Rockies and Plains. Negative precipitation IRRs range from 0.994 (95% CI 0.973, 1.016) in the West to 0.9995 in the South and the Ohio Valley (95% CI 0.998, 1.001 for both). The precipitation IRRs for each climate region are shown in Figure 5b.

Heterogeneity between States and Climate Regions

Upon visual examination, all confidence intervals of the state-level temperature IRRs overlapped the other IRR confidence intervals in their climate region and in other climate regions. Visually, therefore, there is little evidence of homogeneity among the state temperature IRRs within or between climate regions. The associated I^2 was negative ($I^2 = -93.6$), and therefore set to zero, which also indicates a lack of heterogeneity between states.¹²⁸

The confidence intervals for precipitation were very narrow. Shown at the same scale as the temperature estimates, it was very difficult to assess heterogeneity between precipitation confidence intervals visually. However, the negative I^2 statistic ($I^2 = -37.8$) was set to zero, indicating no heterogeneity between states.

When the data is aggregated to climate region, all temperature IRR confidence intervals are overlapped by the confidence intervals of at least four other IRRs, but not by all IRRs. This indicates some visual heterogeneity among the data when aggregated to climate region. Statistically, the heterogeneity analysis yielded an I^2 value of 68.5. The I^2 informs that 68.5% of the total variation across the climate regions is due to heterogeneity between them.

Similarly to the temperature IRRs, all climate region precipitation IRR confidence intervals are overlapped by at least two other IRR confidence intervals. However, like the state-level precipitation IRRs, heterogeneity is difficult to assess visually due to the narrowness of most of the confidence intervals. Statistical analysis found heterogeneity to be responsible for 82.1% of the total variation across climate regions ($I^2 = 82.1$).

Temperature and precipitation IRRs were also calculated for each percentage well use decile. These IRRs are displayed in Figure 6, a and b. The temperature IRRs show some visual heterogeneity between well use deciles, but there is no obvious pattern. However, the I^2 value indicates that 74.1% of the variation between well use deciles is accounted for by

heterogeneity. The precipitation IRRs also show visual heterogeneity, which accounts for 87.0% of the variation between well deciles.

Generalized Estimating Equations Modeling

In the single variable, unadjusted GEE models, the one-month lagged temperature (unadjusted IRR 1.024, 95% CI 1.021-1.027) was found to have a marginally stronger relationship with cryptosporidiosis incidence than zero-month lagged temperature (unadjusted IRR 1.022, 95% CI 1.020-1.025). Precipitation was also significant in the single variable models, with an unadjusted IRR of 1.0002 (95% CI 1.0000, 1.0004). Latitude, precipitation, temperature, and lagged temperature all show a significant association with cryptosporidiosis when modeled univariably. Unadjusted Incidence Rate Ratios for all variables are shown in Table 2.

There was no correlation found between the demographic variables, percent well use, percent urban counties, and percent individuals under poverty. Correlation output and plots are shown in Appendix H.

The final, adjusted model included lagged temperature, precipitation, latitude, percent total well use, percent urban counties, and percent individuals below the Federal Poverty Line. Analysis of NNDSS data across all states, shows a 2.5% increase in \log_{10} cryptosporidiosis incidence in the United States for every 1 °C increase in temperature (IRR 1.025, 95% CI 1.22-1.028). A very weak positive association was found between cryptosporidiosis incidence and precipitation, but this was not statistically significant (IRR 1.0001, 95% CI 0.9999-1.0003). A statistically insignificant 3.9% increase in \log_{10} cryptosporidiosis incidence was found to occur for every degree increase in latitude (IRR 1.039, 95% CI 0.988-1.093). None of the other variables showed significant relationships with cryptosporidiosis incidence. Of interest, however, is the relationship between well use

and cryptosporidiosis. While not significant, the pooled analysis model predicted an increase of over 232% in \log_{10} cryptosporidiosis cases for every 1% increase in well use (IRR 2.32 95% CI 0.34-16.04, $p=0.39$).

DISCUSSION AND RECOMMENDATIONS:

This study of NNDSS data shows clear late summer seasonality for cryptosporidiosis in the US. There are several findings of note from this analysis. First, seasonality differed across climate regions, with some regions showing the dual peak seasonality described in the literature, and others showing only a single seasonal peak. Season strength also differs across the states. Second, the relationship between temperature and cryptosporidiosis is positive across all states individually and overall for the entire country, with cryptosporidiosis incidence increasing on average by 2.5% for every 1 °C increase in temperature. Third, the relationship between precipitation and cryptosporidiosis is weak overall, and not statistically significant in most states. Fourth, temperature and precipitation IRRs show little heterogeneity between states, but do show heterogeneity across climate regions that accounts for 68% of variation among temperature IRRs and 82% of variation among precipitation IRRs. Finally, these relationships are similar to the relationships found in the analysis of the global literature review data.

Season Strength and Seasonality

The season strength of cryptosporidiosis varies across the states, likely due to varying distributions of risk factors and climate factors. Visual comparison to maps of temperature and precipitation peak-to-mean ratios across the US did not show any obvious similarities in patterns (Temperature and Precipitation Season Strength maps are available in Appendix I). There are several states that have unusually high season strengths compared to the states that surround them, namely Connecticut (4.89), Delaware (3.43), and Ohio (3.37). Connecticut and Delaware's season strengths could be spurious results driven by low numbers of reported cases (27 and 35 total cases, respectively). However Ohio, which has about 250

cases yearly, has statistically relevant season strength significantly higher than the surrounding states.

The US's single seasonal peak spans the months of June to October (Figure 2). Aside from a slight increase in normalized cases in March and April, there is no evidence of *Cryptosporidium*'s dual peak seasonality when looking at the US as a whole. These observations parallel the results from Lal *et al.*'s systematic review.¹¹ Considering that the large summer/autumn peak aligns almost exactly with the swim season, it could be that most of the sporadic cases of *Cryptosporidium* in the US are due to anthroponotic transmission of *C. hominis* through recreational water during swim season. This hypothesis is supported by data from the CDC's Molecular Epidemiology Lab looking at outbreaks of *Cryptosporidium* over 10 years, from 1993 to 2003. Several cases from each outbreak were analyzed to see which species might have produced the outbreak. Of the 20 US outbreaks listed, only five had a majority of *C. parvum* cases, indicating that most outbreaks of cryptosporidiosis are due to *C. hominis*, which is typically transmitted anthroponotically through recreational water venues.¹²⁹

When the NNDSS data was broken into climate regions (Figure 3), all of the regions showed the same strong summer seasonality as the entire US did in Figure 2. The West, Northwest, and Northern Rockies and Plains regions were the only regions to show the second, smaller peak in the spring. This dual seasonality in the northwestern states of the US is consistent with cryptosporidiosis seasonality results from a study in Oregon performed in the late 1980s.³⁹ The Northeast region did not show dual peak seasonality, which echoes the single, summer peak seasonality Naumova *et al.* found in Massachusetts.³⁶ The seasonal peak in the West is the only one that has such high percentages of cases in July and September. The West climate region produces a more logarithmic-shaped curve from May to August,

which is unique among all climate regions, though what factors cause this strange seasonal shape are unknown.

Temperature

The positive relationship between temperature and cryptosporidiosis is an important finding in this study. In the NNDSS data, for every 1 °C increase of temperature, the \log_{10} of cryptosporidiosis incidence increased by 2.5%. This is markedly similar to the results obtained from the pooled analysis of literature review studies from countries in the northern hemisphere, which found a 2.6% increase in cryptosporidiosis incidence for every 1 °C increase of temperature (Levy *et al.* unpublished data). In the pooled analysis of all global literature review data, for every 1 °C increase of temperature, cryptosporidiosis incidence increased by 2.96%. The lagged-temperature parameter estimates of our models are significantly smaller than those of a previously published cryptosporidiosis meta-analysis by Jagai *et al.*¹⁰ They do not present IRRs, therefore only model parameter estimates could be compared. For a closer comparison with the NNDSS model, we looked at Jagai *et al.*'s models that included both lagged temperature and an adjustment for latitude (Model 4). Categorizing studies from the US into the humid mid-latitude climate (C) and the cold temperate climate (D), Jagai *et al.* found lagged temperature parameter estimates of 0.379 and -0.220, respectively, though neither was significant. The NNDSS model in our study found the lagged-temperature parameter estimate of 0.025, an order of magnitude smaller than Jagai *et al.*'s estimate values. Similarly, model of northern hemisphere studies produced an estimate of 0.026. The relationship between temperature and cryptosporidiosis is the same, but the magnitude of effect is much smaller in the NNDSS and northern hemisphere literature review models. Lagged temperature had a slightly higher effect than non-lagged

temperature in our NNDSS model and Jagai *et al.*'s climate C model, but not in either of the literature review models.

The pooled analysis of the 'all studies' literature review data estimated an increase in cryptosporidiosis cases of 2.96% with every 1 °C increase in temperature, from a parameter estimate of 0.029 (Levy *et al.* unpublished data). For Jagai *et al.*'s 'all studies' model, comparable to our study's global literature review model, the temperature estimate is 0.167, again an order of magnitude higher than our estimate. A potential reason for this difference may be that this review only included studies with case counts for individual months, whereas Jagai *et al.* used any study with 12 months of data regardless of whether the numbers were prevalences or incidences, aggregate or averaged.¹⁰

The individual GLM analyses on studies from the global literature review found results similar to the NNDSS, particularly in the studies that took place in developed countries in the northern hemisphere. Temperature IRRs with their confidence intervals for each study are presented in Figure 7a (Levy *et al.* unpublished data). Some studies included in the systematic review were excluded from GLM analyses because there was no weather data available for the study location. Globally, there is a trend of increasing dependency on temperature with increasing proximity to the equator, according to the temperature IRRs. The relationship between temperature and cryptosporidiosis incidence increases in magnitude in the studies that are from areas in or geographically near to the tropics. However, the widening confidence intervals indicate that uncertainty in these estimates also increases in or near the tropics.

Studies from Brisbane also found higher increases in case incidence, from 9% to 17% for every 1 °C increase in temperature, depending on what type of model was used.^{13,14} A longitudinal study from NZ found a parameter estimate of 0.057 for both temperature

and lagged temperature.⁴⁹ This estimate from NZ is more similar to the NNDSS and global literature review parameter estimates. Other studies have found a positive relationship with one month lagged temperature and cryptosporidiosis,^{37, 92} but few others thus far have attempted to quantify this relationship.

While the literature disagrees on the magnitude of temperature's effect on cryptosporidiosis incidence, almost all studies find that cryptosporidiosis incidence will increase with rising temperatures. Based on their review of climate variability on cryptosporidiosis, Lal *et al.* suggest that this increase could be due to the lengthening of the transmission season and introduction of pathogens into areas that had previously been too cold.¹³⁰ One notable exception is Britton *et al.*, who found a 2% decrease in cryptosporidiosis incidence with every 1 °C increase in temperature,⁹³ which others attributed to the influence of local factors.¹³⁰ Overall, the positive relationship between temperature and cryptosporidiosis seems to be a stable phenomenon worldwide.

An increase in case incidence as small as 2.5% could substantially increase the burden of cryptosporidiosis across a population as large as the United States', where the disease already creates high monetary strain on the healthcare system.⁴ For several reasons it may be assumed that the NNDSS under-reports cases: many people are not willing or able to seek care, the doctors of those who do seek care may not send a stool sample to a lab for testing, and of those doctors that do send a sample, many will not specify a test for *Cryptosporidium*. The NNDSS, therefore, only records the tip of the proverbial iceberg of actual cryptosporidiosis incidence. These same considerations apply to an increase of 2.96% across the world, especially in developing countries where there is already such a high burden of disease due to cryptosporidiosis.

Precipitation

Precipitation was found to have a significant relationship with cryptosporidiosis in the unadjusted pooled analysis NNDSS model, which showed a very small increase in cryptosporidiosis incidence of 0.02% for each 1 mm increase in precipitation. When placed in the final adjusted model, however, precipitation was no longer significant (IRR=1.0001, 95% CI 0.993-1.007). One month lagged precipitation showed no significance either in the unadjusted model or the adjusted model (Table 2). This was different from the comparable ‘northern hemisphere’ model that found lagged precipitation to be significant (IRR=1.007, 95%CI 1.003-1.011), but the same as the ‘all studies’ literature review model that found neither precipitation or lagged precipitation to be significant (IRR=1.004, 95% CI 1.000-1.006). (Levy *et al.* unpublished data).

Similar to the temperature estimates, our precipitation estimates are significantly smaller than those found by Jagai *et al.* in their meta-analysis. In their non-lagged model adjusted for latitude, Jagai *et al.* found precipitation parameter estimates of 0.168 and -0.280 for climate regions C and D respectively.¹⁰ The parameter estimate for precipitation in the NNDSS was 0.0001, several orders of magnitude smaller. The parameter estimates for lagged precipitation in the literature review models were similarly small at 0.007 (northern hemisphere) and 0.004 (all studies). Jagai *et al.*’s parameter estimate for the lagged ‘all studies’ model was much larger, at 0.091. Both the northern hemisphere literature-review analysis and Jagai *et al.* found the lagged-precipitation estimate to be statistically significant, in contrast to the NNDSS and ‘all studies’ literature review results.

Precipitation had very few significant relationships with cryptosporidiosis in the GLM models as well, with only Iowa and North Dakota showing a significant positive relationship between cryptosporidiosis incidence and precipitation, and no definitive pattern

in the states across climate regions. The literature review GLM forest plots in Figure 7b show a similar lack of association (Levy *et al.* unpublished data). The seven studies that do show significant positive relationships between precipitation and cryptosporidiosis incidence are Scotland (Study 1, IRR=1.39), Germany (Study 6, IRR=1.09), Kuwait (Study 24, IRR=1.26), Pakistan (Study 26, IRR=1.06), Bangladesh (Study 27, IRR=1.02), India (Study 28, IRR=1.05), and South Africa (Study 39, IRR= 1.04). Given this wide range of latitudes, and climate regions, it appears that the relationship between precipitation and cryptosporidiosis is likely not strong, and modified by other factors. In terms of NNDSS data aggregated to climate region, only the Northeast, the Upper Midwest, and the Northern Rockies and Plains showed significant but weak positive relationships.

Other studies have found significant positive relationships between cryptosporidiosis and precipitation. One study in the North-West of England found a 27% increase in the weekly rate of cryptosporidiosis if the preceding week's rainfall had been in the 75th percentile.⁴³ A similar study in New Zealand found only a 2% increase in cryptosporidiosis incidence with increasing rainfall to evaporation ratio.⁹³ Neither of these attempted to correlate case incidence with mm of rainfall. However, both of these studies had a much higher geographic resolution in their data than our study did and also used specific exposure metrics, such as rainfall to evaporation ratio, that were more appropriate than total or average rainfall. These two factors may have enhanced their ability to find significant associations between precipitation and cryptosporidiosis.

It should be noted that the rainfall measurements used in this study are in volume (mm) and not in frequency. There is a difference between a climate of frequent, low volume rainfall such as the Northwest region (yearly average of 782 mm) and a climate with infrequent high volume rainfall such as the Southeast region (yearly average of 1218 mm). It

is possible that the relationship between cryptosporidiosis and precipitation depends less on the average and more on the frequency of rainfall or environmental conditions prior to heavy rainfall.⁸⁵ This is another possible explanation as to why our analysis of precipitation, found null results.

The seasonality over time for each climate region, shown in Figure 3, provides another possible explanation as to why the literature conclusion on the relationship of cryptosporidiosis with precipitation is so varied. The seasonality of precipitation is highly variable across climate regions, dependent on many factors, including latitude, local geography, and variation in gulf stream strength and position, to name a few. This variability motivates the present study—our results show that the relationship between cryptosporidiosis seasonality and precipitation seasonality varies by climate region, implying that a single study from one location is cannot necessarily be generalized, and perhaps that data needs to be of a higher geographic resolution for the most accurate results. It is interesting however, than the seasonality curves show stronger regional differences in precipitation than the GLM models.

Well Use

While not statistically significant, the adjusted IRR for well use was significantly higher than any other variable in the pooled analysis model, including temperature (IRR = 2.51, 95% CI 0.356 - 17.635). It has been suggested that use of private water supplies, such as wells, are a risk factor for cryptosporidiosis.^{47, 131, 132} However, when the temperature and precipitation IRRs were aggregated and presented by well use (Figure 6 a & b), there was no discernible pattern apparent, although it did appear that the states that represented the 70th-80th percentile in well use (Connecticut, Delaware, South Carolina, and West Virginia) seemed to be driving the temperature-cryptosporidiosis incidence relationship, (IRR= 1.07,

95%CI 1.04-1.09). The Cochran's Q values of 33.9 and 68.3, for temperature and precipitation respectively, indicate considerable heterogeneity between results, though there is no trend among these differences.

Heterogeneity within NNDSS Data

The nature of surveillance data collection across the US makes it a good candidate for comparison to a meta-analysis. Each state reports cases differently depending upon the amount of time and resources they are willing to invest. Similarly, studies in a meta-analysis have differing designs and methods of data collection. Therefore, both can be tested for heterogeneity between the different data sources (states and studies).

The I^2 was ideal for analyzing heterogeneity statistically between states as it can be used regardless of outcome data type or effect measure, and can calculate the effect of any heterogeneity in the analysis. In the literature, an I^2 value below 25% is considered low heterogeneity, between 50% and 75% is considered moderate heterogeneity, and over 75% is considered a high amount of heterogeneity.¹²⁸ Therefore, the I^2 value for temperature IRRs between climate regions of 68.49% is considered 'moderate' heterogeneity, and the I^2 value of 82.10 for the precipitation IRRs between climate regions would be considered high. I^2 values for both temperature and precipitation IRRs of states were rounded up to 0 and indicate no heterogeneity at all between states. This is also visually apparent in the forest plots, particularly in the temperature plots (Figure 4a and 4b).

Comparison between NNDSS and Meta-analysis

Pooled Analysis GEE Model

Both the NNDSS analysis and literature review meta-analysis found an overall positive relationship between temperature and cryptosporidiosis, of approximately the same magnitude: The NNDSS analysis showed a 2.5% increase in cases and the literature review

showed a 2.6% increase in cases in developed countries in the northern hemisphere, and a 3.0% increase in cases in all studies, with 1 °C increase. The similarity between the NNDSS model estimate and the estimate from the northern hemisphere countries in the literature review indicate generalizability of results across the two datasets. The increases in effect of temperature on cryptosporidiosis incidence in the ‘all studies’ model suggests that temperature may have more of an impact on disease incidence in countries that do not fall into the ‘developed countries in the northern hemisphere’ category. This result is supported by the GLM forest plots of the literature review data shown in Figure 7a (Levy *et al.* unpublished data).

Both pooled analysis models from the literature review found that zero-month-lagged temperature was more strongly associated with cryptosporidiosis incidence than one-month-lagged temperature. As the NNDSS lagged-temperature effect was only marginally larger than the non-lagged temperature effect, this was not cause for concern. Adjusted IRRs for each variable in both literature review models are shown in Table 3.

The overall relationship with precipitation differed between models. Precipitation was not significant in the NNDSS model, lagged or not, and found only a very small association of 0.02% increase with one mm increase of precipitation. The global literature review models found the relationship between one-month-lagged precipitation and cryptosporidiosis significant, with a 0.38% increase of cryptosporidiosis with one mm increase in precipitation in the prior month in the ‘all studies’ model and an increase of 0.7% per one mm precipitation increase in the prior month in the ‘northern hemisphere’ model. This difference could be attributed to the different climate regions included in each analysis. The US and the ‘northern hemisphere’ models do cover a variety of temperate climate regions, but cannot legitimately be compared to tropical climates, whereas the ‘all studies’

literature review model did include studies from tropical areas. The larger impact of precipitation on northern hemisphere countries than both the 'all studies' and the NNDSS results is puzzling. Regardless, all precipitation associations are very weak.

Temperature

In terms of individual effect sizes estimated using GLM, non-US studies in the literature review analysis tended to present IRRs of a much larger effect size than the NNDSS IRRs. However, the US studies from the literature review showed similar magnitudes in the effect of temperature on cryptosporidiosis incidence (IRR range 1.01-1.09) to the NNDSS analysis (IRR range 1.01-1.11).

Results from two studies of state-level surveillance data from the literature review provide an interesting comparison to the NNDSS data results. The temperature IRRs from Naumova *et al.*'s data from Boston, Lowell, and Worcester, Massachusetts were 1.03 (95% CI 0.99-1.06), 1.03 (95% CI 0.96-1.11), and 1.09 (95% CI 1.05-1.13), respectively. This is almost exactly the same as the NNDSS temperature IRR for Massachusetts (IRR=1.04, 95% CI 0.99-1.09) in both magnitude and significance. The results from Oregon in both the literature review and the NNDSS data were similar as well. Oregon (Skeels *et al.*: IRR=1.09, 95% CI 1.01-1.18; NNDSS: IRR=1.04, 95% CI 0.9925-1.0862). These similarities legitimize the comparison of the NNDSS data to data from other studies performed in the US, as far as the relationship between temperature and cryptosporidiosis is concerned.

Comparing the NNDSS state GLM estimates with other developed countries in the northern hemisphere reveals that most studies from such countries also show a positive relationship between temperature and cryptosporidiosis. Corbett-Feeny *et al.*, is the only study in that category that found a negative relationship, with cryptosporidiosis incidence

decreasing 2% with every 1 °C increase in temperature. This result was statistically insignificant, however.

Other countries that present temperature IRRs of less than one are Kuwait (both studies), and Costa Rica, both of which have different climates than the US and sit at latitudes closer to the equator. This difference, along with the trend of increasing dependency on temperature with proximity to the equator, suggests comparison between temperate and tropical climates may not be legitimate. Both New Zealand studies also report IRRs of less than one. These New Zealand IRRs, while insignificant, echo Britton *et al.*, who also found a negative relationship between temperature and cryptosporidiosis incidence in New Zealand.⁹³

Precipitation

In general, precipitation IRRs from the literature review show smaller effect sizes than the temperature IRRs, like they do in the NNDSS analysis. Many more precipitation IRRs than temperature IRRs are insignificant, again, similar to the NNDSS results. The only statistically significant relationships are the positive ones, again like the NNDSS results. In the US studies specifically, the range of IRRs is 0.94 to 1.03. The literature review IRRs show a much greater effect than the NNDSS IRRs, which hover around the null, but none of the literature review IRRs for US studies are significant.

The comparison between the studies in Oregon and Massachusetts to their respective NNDSS data found results similar to those found for the relationship between cryptosporidiosis and temperature. Both datasets show a negative relationship between precipitation and cryptosporidiosis in Oregon (Skeels *et al.* IRR=0.94, 95%CI 0.87-1.01 and NNDSS IRR=0.996, 95% CI 0.991-1.001), but neither are statistically significant. The NNDSS results are closer to the null. In Massachusetts, the NNDSS IRR is closer to the null

at 1.001 (95% CI 0.997-1.004) than the values from Naumova *et al.*'s data (Boston IRR=1.01, 95% CI 0.97-1.05; Lowell IRR=0.94, 95% CI 0.85-1.03; Worcester IRR=1.03, 95% CI 0.95-1.12). There is disagreement among the estimates from the Naumova data as to the relationship between cryptosporidiosis and precipitation, but all of the IRRs are statistically insignificant.

The precipitation IRRs from the literature review studies, like the precipitation IRRs in the NNDSS analysis, have much smaller effect sizes than the temperature IRRs and are more varied. With the exception of Study 1 in Scotland (IRR 1.39, 95% CI 1.27-1.52) and Study 24 in Kuwait (IRR=1.26, 95% CI 1.07, 1.47) all precipitation IRRs fall between England's IRR of 0.89 (Study 3, 95% CI 0.79, 1.01) and Spain's IRR of 1.10 (Study 20, 95% CI 1.00, 1.20). Developed countries in the northern hemisphere made up 14 of the studies, eight of these studies showed a positive relationship (57.1%) and six showed a negative relationship (42.8%). These are a slightly more even distributions of negative and positive relationships than was seen in the NNDSS data, where 74% of the states showed a positive relationship, and 26% showed a negative relationship.

The stronger effect sizes in the literature review data are possibly due to having fewer zero counts in the literature review data. The literature review specifically included studies that had continuous monthly data for one year or more. The NNDSS data, however, had many months where zero cases were reported for a particular state. This is a function of the data being passively collected. The excess zeros and highly right-skewed counts per month in the NNDSS caused the variance to be unequal to the mean. This unmet assumption was accounted for by \log_{10} transforming the NNDSS data, which may have biased IRRs closer to the null. Another possible explanation is that the weather data for the literature review was localized to the city in which the study took place, or one very close by. The NNDSS data

was aggregated to the state level, but monthly temperature and precipitation was only representative of a single city in that state and not the state as a whole.

Limitations

Limitations are inherent in any observational data collection process, and the data in this study are no exception.

Any literature review that does not include ‘grey literature’ has a publication bias against insignificant results. This literature review only included published articles, and is therefore at risk of excluding data from null results, which might change the outcome of this study. However, as the data collected was cryptosporidiosis case counts and not odds ratios or rate ratios, the potential of excluding truly null results is low. The ‘case-count per month’ inclusion criteria excluded many viable studies from our analysis, including reports with monthly averages and prevalence data.

The NNDSS data was not available at a higher geographic resolution than state level, affecting the representativeness of the temperature and precipitation data, which were taken from a single city in the state. We attempted to reduce this source of bias by supporting our state level data with results from data aggregated and averaged to climate regions defined by the NCDC as being historically climatically consistent.¹¹⁷ Additionally, the NNDSS case definitions of probable and confirmed have changed from 1997 to 2011, meaning that, data reporting may not have been consistent over time. A recent change in case definition, stating that any case diagnosed by rapid immunological assay test is only probable and not confirmed, is also not necessarily relevant as we did not differentiate between probable and confirmed cases in our models. The variable defining outbreak status is also questionable. This study only excluded cases known to be related to outbreaks, which left the sporadic and

‘outbreak status unknown’ cases. As there is no way of knowing, without molecular subtyping, whether the supposedly sporadic cryptosporidiosis cases are related to other cases, it is possible that some of the cases used in this dataset were associated with an undetected outbreak. The weather data is from a single city in each state and does not necessarily represent the weather of that state as a whole. This limitation was lessened by the exclusion of states with ranges of latitudes too wide for weather data from one location to be representative of the entire state. It is likely, however, that some misrepresentation remains. However, the GLM and GEE modeling results still show a significant positive relationship with temperature in all states and climate regions. These results indicate that the relationship between temperature and case incidence is robust, as the relationship exists despite these limitations. Lastly, the results from the NNDSS specifically can only be generalized to other developed countries in temperate climate regions and cannot be reliably compared to studies conducted in tropical developing countries, though the results found from the literature review analysis were comparable.

A recent review of cryptosporidiosis showed that prevalence estimates vary widely across different study designs, different methods of detection, and population immune status.¹⁷ These are important considerations for both datasets in this study. The literature review studies range across continents and time, therefore, the study designs and the methods of detection are highly varied. Study designs include hospital-based case-control studies, community-based cohort studies, and countrywide laboratory surveillance, among others. The NNDSS receives data from the individual and independently run state health departments, whose methods of data collection differ according to funding availability and state priorities.¹³³ Methods of detection mentioned in the studies include microscopy with a variety of stains (Giemsa, modified Kinyoun acid-fast, and safranin-methylene-blue stains

among others), PCR, and rapid immunoassays. In many of the surveillance datasets, including the NNDSS data, method of detection was not presented. This limits the ability to compare across states in a similar way that the different study designs incorporated into a meta-analysis limits cross-study comparison. Immune status is a key determiner of cryptosporidiosis frequency and severity. However, immune status is not reported in NNDSS,¹⁰⁰ nor is it specified in many of the literature review studies. This lack of important data makes it difficult to see the impact of cryptosporidiosis on the general population and creates some uncertainty in the results. This lack of generalizability/comparability was lessened by the inclusion of known immune-compromised cases from literature review data in which immune status was collected but not specifically selected for.

Perhaps most importantly, many of these studies predate the discovery that cryptosporidiosis is caused by several genetically different, yet phenotypically indistinguishable species, not just one species (*C. parvum*). Of the studies performed after this discovery, very few presented monthly case counts by species, choosing instead to present total numbers of cryptosporidiosis. The scientific community is now realizing that not only are the species different genetically, they are also different epidemiologically, which leads to a dual peak seasonality⁵⁰⁻⁵², and difficulty in fitting models that probe both sets of risk factors. This same limitation is true of the NNDSS data. Genotype is not yet captured in the surveillance system, and therefore the models built with the NNDSS data and with the literature review data likely only capture risk factors for the larger summer/fall month peak, and even then the models may not fit as well. There is a major need in this area specifically for studies and data that take *Cryptosporidium* species into account when modeling in order to better understand the dynamic relationships between cryptosporidiosis and climatological factors.

Recommendations

The greatest need in *Cryptosporidium* research at this time is molecular epidemiology and species genotyping. Given the different and complex transmission patterns and seasonalities of each species, knowledge of species type is key for modeling risk factors and understanding sporadic cryptosporidiosis epidemiology. This study would have benefited from species-specific data, particularly in interpreting the dual peak seasonality of cryptosporidiosis and in fitting a model to describe *Cryptosporidium*'s complex relationships with climatological factors. The lack of species-specification in this data could have contributed to some of the heterogeneity in the results. Species type is also important for developing appropriate preventative measures and risk communication tools.

Molecular typing will also increase the ability to detect and differentiate between outbreaks, an ability that is currently restricted due to the similarities in *Cryptosporidium* spp. morphology and lack of testing. *Cryptosporidium* species genotyping is performed in the US through the Molecular Epidemiology lab in the Waterborne Disease Prevention Branch of the CDC, though not on a large scale. However, the CDC is currently developing a molecular tracking surveillance system called CryptoNet, the first parasite infection tracking system of its kind, in order to capture species-level information throughout the US. Setting up systems similar to CryptoNet or the United Kingdom's *Cryptosporidium* Reference Unit, in other countries around the world would be highly beneficial to this cause, although, it would be difficult to find resources for such systems in developing countries. Countries with laboratories that currently perform molecular analysis, however, would need no additional resources to perform the required molecular typing.¹³⁴

Tantamount to species typing is including testing for *Cryptosporidium* in routine stool examinations, something that is currently not done in the US. At present *Cryptosporidium* is

only tested for by request, leading to a high level of under-reporting in the NNDSS.¹⁶

Knowing the actual burden of disease in the US would both increase awareness of risk and help epidemiologists and public health practitioners understand how to mitigate risks of infection.

Until *Cryptosporidium* is routinely tested for and molecular genotyping is well established, the following recommendations may be useful in further research:

1. A proxy variable could be substituted for *Cryptosporidium* species, segmenting the data, and building separate models for both categories. *C. hominis* is transmitted by human-to-human contact, typically in chlorinated recreational water facilities, which would be more highly utilized in population dense urban centers. *C. parvum* is usually transmitted by young calves, which are much more prevalent in rural communities. A study in Scotland found that increasing population density was protective against *C. parvum*, and *C. hominis* was prevalent much more frequently in densely populated areas.⁴⁷ Therefore, a variable that differentiates between urban and rural areas might be a good approximation of species type.
2. Different predictor variables could be used to more fully describe risk factors. This study would benefit from data on the number of recreational water sources, or cattle density first, known risk factors for *C. hominis* and *C. parvum*, respectively.^{47, 135} It is reasonable to conjecture that the secondary spring seasonal peak may have been more pronounced in areas of high cattle density, if data were available for this analysis. This inference is supported by the fact that Naumova *et al.*'s Massachusetts data did not have a double peak,³⁶ being largely urban (85.7%) but Oregon, with 44.4% urban counties, did display dual peak seasonality in Skeels, *et al.* Skeels *et al.* also state the plausibility of contamination

of Oregon surface water by dairy cattle.³⁹ The seasonal distributions of cryptosporidiosis in the Northeast (no second peak) and Northwest (distinct second peak) climate regions also support this (Figure 3).

3. Use of weather data that is more representative of the state as a whole instead of one city would be beneficial.
4. Separating the NNDSS data into a finer geographic resolution, such as county or census block instead of state, would also provide more accurate associations between cases and climate factors. Alternatively, separating data into groups that are classified according to climate rather than political boundaries, such as US EPA climate zone designations, would likely accomplish the same goal.
5. Lastly, it would be interesting to look at cryptosporidiosis data from Australia, a similarly large and developed country in the Southern hemisphere, to see if the relationships between cryptosporidiosis, temperature and rainfall parallel observations in the US and other developed countries in the Northern hemisphere.

In terms of elucidating the complex relationships between climatic drivers and infectious diarrheal disease, this study has several implications for future research:

1. The relationship between temperature and cryptosporidiosis is robust across the United States, and other countries in temperate regions. However, studies performed in countries closer to the equator saw an increase in the effect of temperature on cryptosporidiosis incidence. This implies that the relationship between temperature and cryptosporidiosis is affected by latitude, and whether or not the study is done in a temperate or tropical climate. Considering that developed countries tend to be in temperate climates and developing countries

tend to be in tropical climates, future research must also evaluate these climate-disease relationships with an eye to populations most at risk.

2. The relationship between precipitation and cryptosporidiosis is highly variable across all states and climate regions, and also across all studies. This suggests that simply examining precipitation in volumetric measures such as mm will not explain the relationship. Other confounders and effect modifiers, such as local geography, demographics, or weather conditions must be taken into account to fully understand the interaction between precipitation and diarrheal disease. Future research would benefit from studies that take a more detailed longitudinal view of precipitation patterns.
3. Both datasets found the same relationships between temperature and precipitation and cryptosporidiosis incidence when looking at United States data. This indicates that climate, more so than data collection methods drives the relationship between cryptosporidiosis and temperature and rainfall.

There is also a pressing need for practical measures to be taken to prevent the spread of cryptosporidiosis and other infectious diarrheal diseases. Given *Cryptosporidium*'s positive relationship with temperature and the fact that global temperature appears to be increasing, the burden of disease will likely become larger over time.⁷ Preparations must be made in order to reduce cryptosporidiosis incidence, including but not limited to better methods of oocyst inactivation, structural barriers that break the connection between agricultural runoff or fecal waste and drinking water sources, as well as education on appropriate sanitation and hygiene. Such prevention methods are useful not only to prevent cryptosporidiosis but also to prevent diarrheal disease in general, reducing much morbidity and mortality with a few practical solutions.

Conclusions

Current estimates state that approximately 5% of global disease burden and approximately 4% of environmentally-attributed global mortality is due to diarrheal disease.¹³⁶ The NIEHS has called for studies such as the present one, and others that seek to understand how waterborne disease will change with changing climate conditions.⁷ More studies are needed that directly investigate the relationship between specific pathogens and temperature and precipitation in a quantifiable and easily communicated manner. This present study has explored some of the uncertainty in the relationship between climatic drivers and cryptosporidiosis specifically, but uncertainty and lack of precision still remain. This uncertainty is due to lack of specific pathogen research in general, but this study especially demonstrates the need of species identification even between *Cryptosporidium spp.*

Our data suggests that the potential impacts of climate change, particularly the impacts of increasing temperature, are important and preventative measures must be considered. Cryptosporidiosis is one of the top four etiologies of moderate-to-severe diarrheal disease in children under five in developing countries and is the top cause of recreational water outbreaks in the US.^{6,134} The robust, positive relationship with temperature that was shown across all states in the US and most studies in the literature review, is cause for concern. The influence of precipitation on cryptosporidiosis in particular and on diarrheal disease in general likely varies depending on location and precipitation extremity, frequency and prior conditions. More research is needed concerning the impact of these factors on the infectious diarrheal disease-precipitation relationship. As precipitation events become more extreme, more frequent, and more unpredictable, it would behoove governments and communities to take precautionary measures to protect themselves and their people against diarrheal diseases spread by water. Implications are graver for

immunocompromised populations and children under five, who bear the majority of the morbidity and mortality, given *Cryptosporidium*'s opportunistic nature. Education about appropriate hygiene and sanitation is necessary in all countries, along with structural barriers to break the cycle of disease caused by fecal-oral contamination. At the very least, the US healthcare and public health infrastructure, and healthcare systems around the world, must adapt to accommodate large increases in absolute numbers of cases of *Cryptosporidium*-attributable diarrheal disease that would follow the suspected global rise in temperature.

References

1. World Health Organization. Global Water Supply and Sanitation Assessment. Geneva, Switzerland; 2000.
2. World Health Organization. Diarrhoeal Disease Fact Sheet. 2013 [cited; Available from: <http://www.who.int/mediacentre/factsheets/fs330/en/>]
3. Baldursson S, Karanis P. Waterborne transmission of protozoan parasites: Review of worldwide outbreaks – An update 2004–2010. *Water Research*. 2011; **45**(20): 6603-14.
4. Collier SA, Stockman LJ, Hicks LA, Garrison LE, Zhou FJ, Beach MJ. Direct healthcare costs of selected diseases primarily or partially transmitted by water. *Epidemiology and infection*. 2012; **140**(11): 2003-13.
5. Corso PS, Kramer MH, Blair KA, Addiss DG, Davis JP, Haddix AC. Cost of illness in the 1993 waterborne *Cryptosporidium* outbreak, Milwaukee, Wisconsin. *Emerging infectious diseases*. 2003; **9**(4): 426-31.
6. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet*. 2013; **382**(9888): 209-22.
7. Portier C, Thigpen Tart K, Carter S, al e. A Human Health Perspective on Climate Change: A Report Outlining the Research Needs on the Human Health Effects of Climate Change. *Environmental health perspectives*. 2010.
8. Kolstad EW, Johansson KA. Uncertainties associated with quantifying climate change impacts on human health: a case study for diarrhea. *Environmental health perspectives*. 2011; **119**(3): 299-305.

9. Ahmed SM, Lopman BA, Levy K. A systematic review and meta-analysis of the global seasonality of norovirus. *PloS one*. 2013; **8**(10): e75922.
10. Jagai JS, Castronovo DA, Monchak J, Naumova EN. Seasonality of cryptosporidiosis: A meta-analysis approach. *Environmental research*. 2009; **109**(4): 465-78.
11. Lal A, Hales S, French N, Baker MG. Seasonality in human zoonotic enteric diseases: a systematic review. *PloS one*. 2012; **7**(4): e31883.
12. Hu W, Tong S, Mengersen K, Connell D. Weather variability and the incidence of cryptosporidiosis: comparison of time series poisson regression and SARIMA models. *Annals of epidemiology*. 2007; **17**(9): 679-88.
13. Hu W, Mengersen K, Fu SY, Tong S. The use of ZIP and CART to model cryptosporidiosis in relation to climatic variables. *International journal of biometeorology*. 2010; **54**(4): 433-40.
14. Hu W, O'Leary RA, Mengersen K, Low Choy S. Bayesian classification and regression trees for predicting incidence of cryptosporidiosis. *PloS one*. 2011; **6**(8): e23903.
15. Nime FA, Burek JD, Page DL, Holscher MA, Yardley JH. Acute enterocolitis in a human being infected with the protozoan *Cryptosporidium*. *Gastroenterology*. 1976; **70**(4): 592-8.
16. Centers for Disease Control and Prevention. *Cryptosporidium* ("Crypto"). *Parasites* 2013 16 Jan 2013 [cited; Available from: <http://www.cdc.gov/parasites/crypto/>]
17. Shirley DA, Moonah SN, Kotloff KL. Burden of disease from cryptosporidiosis. *Current opinion in infectious diseases*. 2012; **25**(5): 555-63.
18. Sopwith W, Osborn K, Chalmers R, Regan M. The changing epidemiology of cryptosporidiosis in North West England. *Epidemiology and infection*. 2005; **133**(5): 785-93.

19. Dillingham RA, Lima AA, Guerrant RL. Cryptosporidiosis: epidemiology and impact. *Microbes and infection / Institut Pasteur*. 2002; **4**(10): 1059-66.
20. Lima AA, Guerrant RL. Persistent diarrhea in children: epidemiology, risk factors, pathophysiology, nutritional impact, and management. *Epidemiologic reviews*. 1992; **14**: 222-42.
21. Yoder JS, Beach MJ. Cryptosporidium surveillance and risk factors in the United States. *Experimental parasitology*. 2010; **124**(1): 31-9.
22. Savioli L. Cryptosporidiosis. In: Heymann D, editor. *Control of Communicable Diseases*. 18th ed. Washington, D.C.: American Public Health Association; 2004.
23. Nel ED, Rabie H, Goodway J, Cotton MF. A retrospective study of cryptosporidial diarrhea in a region with high HIV prevalence. *Journal of tropical pediatrics*. 2011; **57**(4): 289-92.
24. Checkley W, Epstein LD, Gilman RH, Black RE, Cabrera L, Sterling CR. Effects of *Cryptosporidium parvum* infection in Peruvian children: growth faltering and subsequent catch-up growth. *American journal of epidemiology*. 1998; **148**(5): 497-506.
25. Guerrant DI, Moore SR, Lima AA, Patrick PD, Schorling JB, Guerrant RL. Association of early childhood diarrhea and cryptosporidiosis with impaired physical fitness and cognitive function four-seven years later in a poor urban community in northeast Brazil. *The American journal of tropical medicine and hygiene*. 1999; **61**(5): 707-13.
26. Hussien SM, Abdella OH, Abu-Hashim AH, Aboshiesha GA, Taha MA, El-Shemy AS, et al. Comparative study between the effect of nitazoxanide and paromomycine in treatment of cryptosporidiosis in hospitalized children. *Journal of the Egyptian Society of Parasitology*. 2013; **43**(2): 463-70.

27. Rossignol JF, Kabil SM, el-Gohary Y, Younis AM. Effect of nitazoxanide in diarrhea and enteritis caused by *Cryptosporidium* species. *Clinical gastroenterology and hepatology* : the official clinical practice journal of the American Gastroenterological Association. 2006; **4**(3): 320-4.
28. Abubakar I, Aliyu SH, Arumugam C, Usman NK, Hunter PR. Treatment of cryptosporidiosis in immunocompromised individuals: systematic review and meta-analysis. *British journal of clinical pharmacology*. 2007; **63**(4): 387-93.
29. Addiss DG, Pond RS, Remshak M, Juranek DD, Stokes S, Davis JP. Reduction of risk of watery diarrhea with point-of-use water filters during a massive outbreak of waterborne *Cryptosporidium* infection in Milwaukee, Wisconsin, 1993. *The American journal of tropical medicine and hygiene*. 1996; **54**(6): 549-53.
30. Gomez-Couso H, Fontan-Sainz M, Navntoft C, Fernandez-Ibanez P, Ares-Mazas E. Comparison of different solar reactors for household disinfection of drinking water in developing countries: evaluation of their efficacy in relation to the waterborne enteropathogen *Cryptosporidium parvum*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2012; **106**(11): 645-52.
31. Blackburn BG, Mazurek JM, Hlavsa M, Park J, Tillapaw M, Parrish M, et al. Cryptosporidiosis associated with ozonated apple cider. *Emerging infectious diseases*. 2006; **12**(4): 684-6.
32. Cho EJ, Yang JY, Lee ES, Kim SC, Cha SY, Kim ST, et al. A waterborne outbreak and detection of cryptosporidium oocysts in drinking water of an older high-rise apartment complex in seoul. *The Korean journal of parasitology*. 2013; **51**(4): 461-6.

33. Hopkins J, Hague H, Hudgin G, Ross L, Moore D. An outbreak of *Cryptosporidium* at a recreational water park in Niagara Region, Canada. *Journal of environmental health*. 2013; **75**(9): 28-33.
34. Roy SL, DeLong SM, Stenzel SA, Shiferaw B, Roberts JM, Khalakdina A, et al. Risk factors for sporadic cryptosporidiosis among immunocompetent persons in the United States from 1999 to 2001. *Journal of clinical microbiology*. 2004; **42**(7): 2944-51.
35. Gualberto FA, Heller L. Endemic *Cryptosporidium* infection and drinking water source: a systematic review and meta-analyses. *Water science and technology : a journal of the International Association on Water Pollution Research*. 2006; **54**(3): 231-8.
36. Naumova EN, Chen JT, Griffiths JK, Matyas BT, Estes-Smargiassi SA, Morris RD. Use of passive surveillance data to study temporal and spatial variation in the incidence of giardiasis and cryptosporidiosis. *Public health reports (Washington, DC : 1974)*. 2000; **115**(5): 436-47.
37. Naumova EN, Jagai JS, Matyas B, DeMaria A, Jr., MacNeill IB, Griffiths JK. Seasonality in six enterically transmitted diseases and ambient temperature. *Epidemiology and infection*. 2007; **135**(2): 281-92.
38. Wolfson J, Richter J, Waldron M, Weber D, MacCarthy D, Hopkins C. Cryptosporidiosis in Immunocompetent Patients. *New England Journal of Medicine*. 1985; **312**(20): 1278-82.
39. Skeels MR, Sokolow R, Hubbard CV, Andrus JK, Baisch J. *Cryptosporidium* infection in Oregon public health clinic patients 1985-88: the value of statewide laboratory surveillance. *American journal of public health*. 1990; **80**(3): 305-8.
40. Baxby D, Hart CA. The incidence of cryptosporidiosis: a two-year prospective survey in a children's hospital. *The Journal of hygiene*. 1986; **96**(1): 107-11.

41. Corbett-Feeney G. Cryptosporidium among children with acute diarrhoea in the west of Ireland. *The Journal of infection*. 1987; **14**(1): 79-84.
42. Thomson MA, Benson JW, Wright PA. Two year study of cryptosporidium infection. *Archives of disease in childhood*. 1987; **62**(6): 559-63.
43. Naumova EN, Christodouleas J, Hunter PR, Syed Q. Effect of precipitation on seasonal variability in cryptosporidiosis recorded by the North West England surveillance system in 1990-1999. *Journal of water and health*. 2005; **3**(2): 185-96.
44. Chalmers RM, Smith R, Elwin K, Clifton-Hadley FA, Giles M. Epidemiology of anthroponotic and zoonotic human cryptosporidiosis in England and Wales, 2004-2006. *Epidemiology and infection*. 2011; **139**(5): 700-12.
45. Chalmers RM, Elwin K, Thomas AL, Guy EC, Mason B. Long-term Cryptosporidium typing reveals the aetiology and species-specific epidemiology of human cryptosporidiosis in England and Wales, 2000 to 2003. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*. 2009; **14**(2): 19086.
46. Goh S, Reacher M, Casemore DP, Verlander NQ, Chalmers R, Knowles M, et al. Sporadic cryptosporidiosis, North Cumbria, England, 1996-2000. *Emerging infectious diseases*. 2004; **10**(6): 1007-15.
47. Pollock KG, Ternent HE, Mellor DJ, Chalmers RM, Smith HV, Ramsay CN, et al. Spatial and temporal epidemiology of sporadic human cryptosporidiosis in Scotland. *Zoonoses and public health*. 2010; **57**(7-8): 487-92.
48. Chai JY, Kim NY, Guk SM, Park YK, Seo M, Han ET, et al. High prevalence and seasonality of cryptosporidiosis in a small rural village occupied predominantly by aged people in the Republic of Korea. *The American journal of tropical medicine and hygiene*. 2001; **65**(5): 518-22.

49. Lake IR, Pearce J, Savill M. The seasonality of human cryptosporidiosis in New Zealand. *Epidemiology and infection*. 2007; **136**(10): 1383-7.
50. Lake IR, Bentham G, Kovats RS, Nichols GL. Effects of weather and river flow on cryptosporidiosis. *Journal of water and health*. 2005; **3**(4): 469-74.
51. Learmonth JJ, Ionas G, Ebbett KA, Kwan ES. Genetic characterization and transmission cycles of *Cryptosporidium* species isolated from humans in New Zealand. *Applied and environmental microbiology*. 2004; **70**(7): 3973-8.
52. Learmonth J, Ionas G, Pita A, Cowie R. Seasonal shift in *Cryptosporidium parvum* transmission cycles in New Zealand. *The Journal of eukaryotic microbiology*. 2001; **Suppl**: 34S-5S.
53. Fripp PJ, Bothma MT, Crewe-Brown HH. Four years of cryptosporidiosis at GaRankuwa Hospital. *The Journal of infection*. 1991; **23**(1): 93-100.
54. Steele AD, Gove E, Meewes PJ. Cryptosporidiosis in white patients in South Africa. *The Journal of infection*. 1989; **19**(3): 281-5.
55. Cruz JR, Cano F, Caceres P, Chew F, Pareja G. Infection and diarrhea caused by *Cryptosporidium* sp. among Guatemalan infants. *Journal of clinical microbiology*. 1988; **26**(1): 88-91.
56. Shahid NS, Rahman AS, Sanyal SC. *Cryptosporidium* as a pathogen for diarrhoea in Bangladesh. *Tropical and geographical medicine*. 1987; **39**(3): 265-70.
57. Shahid NS, Rahman AS, Anderson BC, Mata LJ, Sanyal SC. Cryptosporidiosis in Bangladesh. *British medical journal (Clinical research ed)*. 1985; **290**(6462): 114-5.
58. Tumwine JK, Kekitiinwa A, Nabukeera N, Akiyoshi DE, Rich SM, Widmer G, et al. *Cryptosporidium parvum* in children with diarrhea in Mulago Hospital, Kampala, Uganda. *The American journal of tropical medicine and hygiene*. 2003; **68**(6): 710-5.

59. Akinbo FO, Okaka CE, Omoregie R. Seasonal Variation of Intestinal Parasitic Infections Among HIV-Positive Patients in Benin City, Nigeria. *Ethiopian journal of health sciences*. 2011; **21**(3): 191-4.
60. Nacro B, Bonkougou P, Nagalo K, Tall FR, Curtis V. [Clinical profile of cryptosporidiosis in a pediatric hospital environment in Burkina Faso]. *Medecine tropicale : revue du Corps de sante colonial*. 1998; **58**(1): 47-50.
61. Molbak K, Hojlyng N, Gottschau A, Sa JC, Ingholt L, da Silva AP, et al. Cryptosporidiosis in infancy and childhood mortality in Guinea Bissau, west Africa. *BMJ (Clinical research ed)*. 1993; **307**(6901): 417-20.
62. Newman RD, Sears CL, Moore SR, Nataro JP, Wuhib T, Agnew DA, et al. Longitudinal study of *Cryptosporidium* infection in children in northeastern Brazil. *The Journal of infectious diseases*. 1999; **180**(1): 167-75.
63. Adegbola RA, Demba E, De Veer G, Todd J. *Cryptosporidium* infection in Gambian children less than 5 years of age. *The Journal of tropical medicine and hygiene*. 1994; **97**(2): 103-7.
64. Nchito M, Kelly P, Sianongo S, Luo NP, Feldman R, Farthing M, et al. Cryptosporidiosis in urban Zambian children: an analysis of risk factors. *The American journal of tropical medicine and hygiene*. 1998; **59**(3): 435-7.
65. Morse TD, Nichols RA, Grimason AM, Campbell BM, Tembo KC, Smith HV. Incidence of cryptosporidiosis species in paediatric patients in Malawi. *Epidemiology and infection*. 2007; **135**(8): 1307-15.
66. Javier Enriquez F, Avila CR, Ignacio Santos J, Tanaka-Kido J, Vallejo O, Sterling CR. *Cryptosporidium* infections in Mexican children: clinical, nutritional, enteropathogenic, and

- diagnostic evaluations. *The American journal of tropical medicine and hygiene*. 1997; **56**(3): 254-7.
67. Mata L, Bolanos H, Pizarro D, Vives M. Cryptosporidiosis in children from some highland Costa Rican rural and urban areas. *The American journal of tropical medicine and hygiene*. 1984; **33**(1): 24-9.
68. Nath G, Choudhury A, Shukla BN, Singh TB, Reddy DC. Significance of *Cryptosporidium* in acute diarrhoea in North-Eastern India. *Journal of medical microbiology*. 1999; **48**(6): 523-6.
69. Das P, Roy SS, MitraDhar K, Dutta P, Bhattacharya MK, Sen A, et al. Molecular characterization of *Cryptosporidium* spp. from children in Kolkata, India. *Journal of clinical microbiology*. 2006; **44**(11): 4246-9.
70. Gatei W, Wamae CN, Mbae C, Waruru A, Mulinge E, Waithera T, et al. Cryptosporidiosis: prevalence, genotype analysis, and symptoms associated with infections in children in Kenya. *The American journal of tropical medicine and hygiene*. 2006; **75**(1): 78-82.
71. Hansen J, Sato M, Ruedy R, Lo K, Lea DW, Medina-Elizade M. Global temperature change. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; **103**(39): 14288-93.
72. Stocker T, Qin D, Plattner G, Tignor M, Allen S, Boschung J, et al., editors. *Climate Change 2013: The Physical Science Basis*. Cambridge, UK and New York, NY, USA: Cambridge University Press; 2013.
73. National Climate Assessment and Development Advisory Committee. NCADAC Draft Climate Assessment Report. online: U.S. Global Change Research Program; 2013.

74. World Health Organization. Climate change and human health: risks and responses. Geneva, Switzerland: WHO,; 2004.
75. Hashizume M, Armstrong B, Hajat S, Wagatsuma Y, Faruque AS, Hayashi T, et al. Association between climate variability and hospital visits for non-cholera diarrhoea in Bangladesh: effects and vulnerable groups. *International journal of epidemiology*. 2007; **36**(5): 1030-7.
76. Singh RB, Hales S, de Wet N, Raj R, Hearnden M, Weinstein P. The influence of climate variation and change on diarrheal disease in the Pacific Islands. *Environmental health perspectives*. 2001; **109**(2): 155-9.
77. Onozuka D, Hashizume M, Hagihara A. Effects of weather variability on infectious gastroenteritis. *Epidemiology and infection*. 2010; **138**(2): 236-43.
78. Lama JR, Seas CR, Leon-Barua R, Gotuzzo E, Sack RB. Environmental temperature, cholera, and acute diarrhoea in adults in Lima, Peru. *Journal of health, population, and nutrition*. 2004; **22**(4): 399-403.
79. Checkley W, Epstein LD, Gilman RH, Figueroa D, Cama RI, Patz JA, et al. Effect of El Nino and ambient temperature on hospital admissions for diarrhoeal diseases in Peruvian children. *Lancet*. 2000; **355**(9202): 442-50.
80. Bhandari GP, Gurung S, Dhimal M, Bhusal CL. Climate change and occurrence of diarrheal diseases: evolving facts from Nepal. *Journal of Nepal Health Research Council*. 2012; **10**(22): 181-6.
81. Pinfold JV, Horan NJ, Mara DD. Seasonal effects on the reported incidence of acute diarrhoeal disease in northeast Thailand. *International journal of epidemiology*. 1991; **20**(3): 777-86.

82. Levy K, Hubbard AE, Eisenberg JN. Seasonality of rotavirus disease in the tropics: a systematic review and meta-analysis. *International journal of epidemiology*. 2009; **38**(6): 1487-96.
83. Kelly-Hope LA, Alonso WJ, Thiem VD, Canh do G, Anh DD, Lee H, et al. Temporal trends and climatic factors associated with bacterial enteric diseases in Vietnam, 1991-2001. *Environmental health perspectives*. 2008; **116**(1): 7-12.
84. Alexander KA, Carzolio M, Goodin D, Vance E. Climate change is likely to worsen the public health threat of diarrheal disease in Botswana. *International journal of environmental research and public health*. 2013; **10**(4): 1202-30.
85. Carlton EJ, Eisenberg JN, Goldstick J, Cevallos W, Trostle J, Levy K. Heavy rainfall events and diarrhea incidence: the role of social and environmental factors. *American journal of epidemiology*. 2014; **179**(3): 344-52.
86. Lloyd SJ, Kovats RS, Armstrong BG. Global diarrhoea morbidity, weather and climate. *Climate Research*. 2007; **34**(2): 119-27.
87. Curriero FC, Patz JA, Rose JB, Lele S. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994. *American journal of public health*. 2001; **91**(8): 1194-9.
88. Auld H, MacIver D, Klaassen J. Heavy rainfall and waterborne disease outbreaks: the Walkerton example. *Journal of toxicology and environmental health Part A*. 2004; **67**(20-22): 1879-87.
89. Thomas KM, Charron DF, Waltner-Toews D, Schuster C, Maarouf AR, Holt JD. A role of high impact weather events in waterborne disease outbreaks in Canada, 1975 - 2001. *International journal of environmental health research*. 2006; **16**(3): 167-80.

90. Hashizume M, Wagatsuma Y, Faruque AS, Hayashi T, Hunter PR, Armstrong B, et al. Factors determining vulnerability to diarrhoea during and after severe floods in Bangladesh. *Journal of water and health*. 2008; **6**(3): 323-32.
91. Schwartz BS, Harris JB, Khan AI, Larocque RC, Sack DA, Malek MA, et al. Diarrheal epidemics in Dhaka, Bangladesh, during three consecutive floods: 1988, 1998, and 2004. *The American journal of tropical medicine and hygiene*. 2006; **74**(6): 1067-73.
92. Lal A, Ikeda T, French N, Baker MG, Hales S. Climate variability, weather and enteric disease incidence in new zealand: time series analysis. *PloS one*. 2013; **8**(12): e83484.
93. Britton E, Hales S, Venugopal K, Baker MG. The impact of climate variability and change on cryptosporidiosis and giardiasis rates in New Zealand. *Journal of water and health*. 2010; **8**(3): 561-71.
94. Ajjampur SS, Liakath FB, Kannan A, Rajendran P, Sarkar R, Moses PD, et al. Multisite study of cryptosporidiosis in children with diarrhea in India. *Journal of clinical microbiology*. 2010; **48**(6): 2075-81.
95. Pal S, Bhattacharya SK, Das P, Chaudhuri P, Dutta P, De SP, et al. Occurrence and significance of *Cryptosporidium* infection in Calcutta. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1989; **83**(4): 520-1.
96. Mathan MM, Venkatesan S, George R, Mathew M, Mathan VI. *Cryptosporidium* and diarrhoea in southern Indian children. *Lancet*. 1985; **2**(8465): 1172-5.
97. Mac Kenzie WR, Hoxie NJ, Proctor ME, Gradus MS, Blair KA, Peterson DE, et al. A Massive Outbreak in Milwaukee of *Cryptosporidium* Infection Transmitted through the Public Water Supply. *New England Journal of Medicine*. 1994; **331**(3): 161-7.

98. Bennett A, Epstein LD, Gilman RH, Cama V, Bern C, Cabrera L, et al. Effects of the 1997-1998 El Nino episode on community rates of diarrhea. *American journal of public health*. 2012; **102**(7): e63-9.
99. World Health Organization. Public health surveillance. *Health Topics* 2014 [cited; Available from: http://www.who.int/topics/public_health_surveillance/en/
100. Yoder JS, Wallace RM, Collier SA, Beach MJ, Hlavsa MC. Cryptosporidiosis surveillance--United States, 2009-2010. *Morbidity and mortality weekly report Surveillance summaries* (Washington, DC : 2002). 2012; **61**(5): 1-12.
101. Centers for Disease Control and Prevention. National Notifiable Diseases Surveillance System. 2013 [cited; Available from: <http://wwwn.cdc.gov/nndss/default.aspx>
102. Areeshi MY, Beeching NJ, Hart CA. Cryptosporidiosis in Saudi Arabia and neighboring countries. *Annals of Saudi medicine*. 2007; **27**(5): 325-32.
103. Fraser D. Epidemiology of Giardia lamblia and Cryptosporidium infections in childhood. *Israel journal of medical sciences*. 1994; **30**(5-6): 356-61.
104. Ghenghesh KS, Ghanghish K, El-Mohammady H, Franka E. Cryptosporidium in countries of the Arab world: the past decade (2002-2011). *The Libyan journal of medicine*. 2012; **7**.
105. Guerrant RL, Hughes JM, Lima NL, Crane J. Diarrhea in developed and developing countries: magnitude, special settings, and etiologies. *Reviews of infectious diseases*. 1990; **12** **Suppl 1**: S41-50.
106. Mor SM, Tzipori S. Cryptosporidiosis in children in Sub-Saharan Africa: a lingering challenge. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2008; **47**(7): 915-21.

107. Reinthaler FF. Epidemiology of cryptosporidiosis in children in tropical countries. *Journal of hygiene, epidemiology, microbiology, and immunology*. 1989; **33**(4 Suppl): 505-13.
108. Robertson LJ, Chalmers RM. Foodborne cryptosporidiosis: is there really more in Nordic countries? *Trends in parasitology*. 2013; **29**(1): 3-9.
109. Snyder JD, Merson MH. The magnitude of the global problem of acute diarrhoeal disease: a review of active surveillance data. *Bulletin of the World Health Organization*. 1982; **60**(4): 605-13.
110. Huwaldt J. Plot Digitizer. In: SourceForge.net, editor. Boston, MA: Free Software Foundation; 2011.
111. Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 1995; 1995.
112. Centers for Disease Control and Prevention. Case definitions for infectious conditions under public health surveillance; 1997.
113. Council of State and Territorial Epidemiologists. Public health reporting and national notification for cryptosporidiosis; 2009.
114. Robinson TJ, Cebelinski EA, Taylor C, Smith KE. Evaluation of the positive predictive value of rapid assays used by clinical laboratories in Minnesota for the diagnosis of cryptosporidiosis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2010; **50**(8): e53-5.
115. Council of State and Territorial Epidemiologists. Update to cryptosporidiosis case definition; 2010.
116. National Climatic Data Center. GHCND (Global Historical Climatology Network)-Monthly Summaries,. In: National Oceanic and Atmospheric Administration, editor. Asheville, NC.

117. Karl TR, Koss WJ. Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983. Asheville, NC: National Climatic Data Centet; 1984.
118. American Community Survey Office. American Community Survey Documentation. U.S. Census Bureau; 2010.
119. United States Census Bureau HaHESD. Source of Water. In: United States Census Bureau, editor. Washington D.C.; 1990.
120. Economic Research Service. 2003 Rural-Urban Continuum Codes. In: United States Department of Agriculture, editor.; 2003.
121. Neyeloff JL, Fuchs SC, Moreira LB. Meta-analyses and Forest plots using a microsoft excel spreadsheet: step-by-step guide focusing on descriptive data analysis. BMC research notes. 2012; **5**: 52.
122. Huh JW, Moon SG, Lim YH. A survey of intestinal protozoan infections among gastroenteritis patients during a 3-year period (2004-2006) in Gyeonggi-do (province), South Korea. The Korean journal of parasitology. 2009; **47**(3): 303-5.
123. Merino FJ, Perez MT, Ladron de Guevara C, Villasante PA, Beltran M, Velasco AC. [Clinical and epidemiologic characteristics of diarrhea caused by *Cryptosporidium*]. Enfermedades infecciosas y microbiologia clinica. 1989; **7**(2): 72-6.
124. Perch M, Sodemann M, Jakobsen MS, Valentiner-Branth P, Steinsland H, Fischer TK, et al. Seven years' experience with *Cryptosporidium parvum* in Guinea-Bissau, West Africa. Annals of tropical paediatrics. 2001; **21**(4): 313-8.
125. Pintar KD, Pollari F, Waltner-Toews D, Charron DF, McEwen SA, Fazil A, et al. A modified case-control study of cryptosporidiosis (using non-*Cryptosporidium*-infected enteric cases as controls) in a community setting. Epidemiology and infection. 2009; **137**(12): 1789-99.

126. Vrbova L, Johnson K, Whitfield Y, Middleton D. A descriptive study of reportable gastrointestinal illnesses in Ontario, Canada, from 2007 to 2009. *BMC public health*. 2012; **12**: 970.
127. Terletskaia-Ladwig E, Eggers M, Enders M, Regnath T. [Epidemiological aspects of gastrointestinal infections]. *Deutsche medizinische Wochenschrift (1946)*. 2011; **136**(3): 69-75.
128. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ (Clinical research ed)*. 2003; **327**(7414): 557-60.
129. Xiao L. Subtyping *Cryptosporidium* spp. in Supporting Investigations of Waterborne and Foodborne Outbreaks of Cryptosporidiosis. 2009: Association of Public Health Laboratories; 2009.
130. Lal A, Baker MG, Hales S, French NP. Potential effects of global environmental changes on cryptosporidiosis and giardiasis transmission. *Trends in parasitology*. 2013; **29**(2): 83-90.
131. Tollestrup K, Frost FJ, Kunde TR, Yates MV, Jackson S. *Cryptosporidium* infection, onsite wastewater systems and private wells in the arid Southwest. *Journal of water and health*. 2014; **12**(1): 161-72.
132. Hunter PR, de Sylor MA, Risebro HL, Nichols GL, Kay D, Hartemann P. Quantitative microbial risk assessment of cryptosporidiosis and giardiasis from very small private water supplies. *Risk analysis : an official publication of the Society for Risk Analysis*. 2011; **31**(2): 228-36.
133. Centers for Disease Control and Prevention. *What Data Users Should Know About the National Notifiable Diseases Surveillance System*. Atlanta, GA: Centers for Disease Control and Prevention,; 2001.

134. Centers for Disease Control and Prevention. CryptoNet: Molecular-based Tracking to Better Understand U.S. Cryptosporidiosis Transmission Fact Sheet. 2013 [cited; Available from: http://www.cdc.gov/parasites/crypto/pdf/cryptonet_fact_sheet508c.pdf
135. Jagai JS, Griffiths JK, Kirshen PH, Webb P, Naumova EN. Patterns of protozoan infections: spatiotemporal associations with cattle density. *EcoHealth*. 2010; **7**(1): 33-46.
136. Mathers C, al e. The Global Burden of Disease: 2004 Update. Geneva, Switzerland: World Health Organization,; 2008.

Tables

Table 1: Characteristics of States and Regions included in study

Tabulated monthly and yearly temperature and rainfall data, season strength, urban and economic metrics, and well usage. Urban counties were defined as having 20,000 or more residents and being adjacent to or part of a metropolitan area. Poverty was defined as percentage of individuals under the Federal Poverty Line, and well use was estimated by the 1990 US decennial census. The Temp and Rain columns represent GLM-calculated associations between cryptosporidiosis incidence rates, and temperature and precipitation, respectively. A '+' indicates a positive relationship between the two and a '-' indicates a negative relationship. Statistically significant relationships ($p < 0.05$) are shown in bold and marked with '*'.

Climate Region	State	City (largest by population as of 2010 US Census)	Latitude of City (min, sec)	Range of latitudes†	Average Yearly Temperature (C)	Average Monthly Temperature Range	Average yearly rainfall (mm)	Average Monthly Precipitation Range	Total # crypto cases	Average yearly crypto cases	Season Strength	Percent Urban	Percent Poverty	Percent Well Use	Temp	Rain	
Northeast	CT	Bridgeport	41° 9' N	40° 58'- 42° 3' N	11.50	-7.01 - 26.87	1177.77	48.58- 141.65	8,322	46.23	2.35	100.0	10	21.90	+	+	
	DE	Wilmington	39° 46' N	38° 27'-39° 50' N	12.74	0.34 - 24.7	1258.21	74.41 - 136.76	35	2.33	3.43	100.0	11.5	23.60	+	+	
	DC	Washington D.C.	38° 56' N	NA	13.12	1.17 - 24.74	1066.11	65.72 - 118.11	42	2.80	1.71	100.0	18.5	0.10	+	-	
	ME	Portland	43° 38' N	43° 4'-47° 28' N	8.28	-5.09 - 20.75	1234.56	76.68 - 141.65	371	24.73	1.88	37.5	13.3	41.90	+	+	
	MD	Baltimore	39° 10' N	37° 53'-39° 43' N	13.26	1.15 - 25.17	1133.30	71.73 - 129.85	371	24.73	1.62	79.2	9.4	16.70	+	+	
	MA	Boston	42° 21' N	41° 10'-42° 53' N	11.08	-1.37 - 23.19	1125.59	79.73 - 112.34	842	56.13	2.58	85.7	11	8.10	+	+	
	NH	Manchester	43° 11' N	42° 40'-45° 18' N	8.21	-6.07 - 21.15	1120.16	72.1 - 112.17	343	22.87	2.31	60.0	8.4	37.50	+	+	
	NJ	Newark	40° 40' N	38° 55'-41° 21' 23" N	13.11	0.43 - 25.43	1206.89	71.63 - 123.81	358	23.87	2.61	100.0	9.9	10.20	+	+	
	NY	New York City	40° 47' N	40° 29' 40"-45° 0' 42" N	13.04	0.71 - 24.86	1317.67	81.89 - 134.1	2,606	173.73	2.47	74.2	14.9	11.40	+	+	
	PA	Philadelphia	39° 58' N	39° 43'- 42° 2' N	14.46	1.51 - 26.87	1303.03	78.01 - 140.45	2,713	180.87	2.35	68.7	13.1	19.80	+	+	
	RI	Providence	41° 43' N	41° 18'- 42° 1' N	11.12	-1.31 - 23.3	1243.99	77.73 - 139.54	73	4.87	2.79	100.0	13.2	11.10	+	-	
	VT	Burlington	44° 28' N	42° 44'- 45° 0' 43" N	8.01	-7.01 - 21.45	988.85	48.58 - 111.25	541	36.07	2.28	28.6	11.6	36.80	+	+	
Southeast	AL§	Birmingham	33° 41' N	30° 13'- 35° N	16.57	5.94 - 26.49	1432.77	74.82 - 153.21	715	47.67	2.05	49.3	18.1	12.00	+	-	
	FL*	Jacksonville	30° 29' N	24° 30' -31° N	20.26	11.83 - 27.7	1287.81	35.85 - 204.45	2,903	193.53	2.24	68.7	15.6	13.00	NA	NA	
	GA	Atlanta	33° 38' N	30° - 35° N	17.05	6.76 - 26.67	1185.90	74.22 - 127.75	2,311	154.07	1.57	48.4	17.4	18.10	+	+	
	NC§	Charlotte	35° 13' N	34° -36° 21' N	15.95	5.43 - 25.99	1048.93	66.83 - 111.67	874	58.27	2.00	57.0	16.8	32.40	+	+	
	SC	Columbia	33° 58' N	32° 4' 30"-35° 12' N	19.39	9.08 - 28.86	1109.65	58.86 - 132.06	320	21.33	1.99	56.5	17.6	22.40	+	-	
	VA‡	Virginia Beach	36° 54' N	36° 31'-39° 37' N	16.01	5.23 - 26.45	1243.16	70.15 - 164.55	656	43.73	1.88	63.0	11.1	21.60	+	+	
					13.05	-4.27 - 27.63	1137.57	30.65 - 170.41	10,245	97.57	3.00					+	-
Ohio Valley	IL	Chicago	41° 59' N	36° 58'-42° 30' N	10.24	-4.27 - 23.63	960.63	45.97 - 127.37	2,195	146.33	2.43	44.1	13.7	9.80	+	+	
	IN	Indianapolis	39° 43' N	37° 47'-41° 46' N	12.13	-1.85 - 24.31	1098.55	67.67 - 130.04	1,548	103.20	1.91	58.7	14.7	25.10	+	+	
	KY	Louisville	38° 10' N	36° 30' - 39° 9' N	14.69	1.79 - 26.23	1229.01	79.16 - 137.47	95	6.33	1.64	31.7	18.6	13.70	+	+	
	MO	Kansas City	39° 7' N	36° -40° 35' N	13.81	-0.7 - 27.39	1040.89	30.65 - 170.41	2,101	140.07	3.96	32.2	15	15.30	+	+	
	OH	Columbus	39° 54' N	38° 27'-41° 58' N	11.61	-1.44 - 23.52	1012.21	52.14 - 104.59	3,739	249.27	3.37	68.2	15.4	16.20	+	+	
	TN	Memphis	35° 7' N	35° -36° 41' N	16.45	4.51 - 27.63	1440.61	90.97 - 155.02	415	27.67	2.43	47.4	17.3	12.10	+	+	
	WV	Charleston	38° 18' N	37° 10'-40° 40' N	12.43	0.62 - 23.4	1181.12	72.68 - 149.47	152	10.13	2.37	41.8	17.6	23.60	+	+	
Upper Midwest	IA	Des Moines	41° 32' N	40° 36'-43° 30' N	10.66	-5.29 - 24.81	902.75	23.39 - 152.59	1,903	126.87	3.20	23.2	12.2	18.40	+	+	
	MI	Detroit	42° 13' N	41° 41'-48° 15' N	10.39	-3.52 - 23.28	879.86	50.11 - 102.12	1,359	90.60	1.89	37.3	16.3	29.10	+	+	
	MN	Minneapolis	44° 52' N	43° 34'-49° 23' N	8.37	-8.93 - 23.86	769.36	20.053 - 118.44	1,382	92.13	2.60	29.9	11.2	26.20	+	+	
	WI	Milwaukee	43° 4' N	42° 30' - 47° 3' N	9.46	-5.25 - 23.16	845.59	36.19 - 123.93	6,369	424.60	3.37	44.4	12.5	32.80	+	+	
					18.09	0.41 - 29.55	1196.18	23.58 - 165.43	5,316	59.07	3.41					+	-
South	AR	Little Rock	34° 50' N	33° -36° 30' N	17.49	5.73 - 28.56	1269.07	70.28 - 131.07	107	7.13	2.92	29.3	18.7	17.50	+	+	
	KS	Wichita	37° 38' N	37° - 40° N	14.19	0.41 - 27.65	892.95	23.58 - 144.42	407	27.13	4.10	19.0	13.2	10.00	+	+	
	LA	New Orleans	29° 59' N	29° - 33° N	21.11	12.25 - 28.87	1494.57	92.65 - 197.44	332	22.13	2.28	57.8	18.7	10.70	+	-	
	MS	Jackson	32° 19' N	30° 13'-35° N	18.43	8.29 - 27.9	1357.91	93.84 - 131.46	40	2.67	3.30	23.2	22.3	12.10	+	-	
	OK‡	Oklahoma City	35° 23' N	33° 35' - 37° N	16.27	3.93 - 28.37	867.71	37.03 - 122.53	766	51.07	2.69	28.6	16.6	12.60	+	+	
	TX*	Houston	29° 58' N	25° 50' -36° 30' N	21.09	11.92 - 29.55	1294.89	74.83 - 165.43	3,664	244.27	3.88	35.8	17.4	8.10	NA	NA	
					8.38	-12.78 - 25.49	572.91	6.24 - 122.67	2,503	33.37	2.83					+	+
Northern Rockies and Plains	MT	Billings	45° 48' N	44° 26'- 49° N	9.04	-2.79 - 23.51	344.31	11.39 - 57.56	373	24.87	2.93	7.1	14.8	30.30	+	+	
	NE	Omaha	41° 18' N	40° - 43° N	10.93	-4.78 - 25.49	790.92	17.56 - 122.11	890	59.33	3.60	10.8	12.4	16.80	+	+	
	ND	Fargo	46° 55' N	45° 55'- 49° N	5.87	-12.79 - 21.85	659.81	16.57 - 122.67	286	19.07	2.43	7.5	12.1	19.20	+	+	

Climate Region	State	City (largest by population as of 2010 US Census)	Latitude of City (min, sec)	Range of latitudes†	Average Yearly Temperature (C)	Average Monthly Temperature Range	Average yearly rainfall (mm)	Average Monthly Precipitation Range	Total # crypto cases	Average yearly crypto cases	Season Strength	Percent Urban	Percent Poverty	Percent Well Use	Temp	Rain
<i>Southwest</i>	SD	Sioux Falls	43° 34' N	42° 29' 30"-45° 56' N	7.96	-8.52 - 23.25	690.39	15.41 - 101.62	851	56.73	2.26	10.6	13.8	16.70	+	+
	WY	Cheyenne	41° 8' N	41° - 45° N	8.10	-2.33 - 21.57	379.12	6.24 - 58.61	103	6.87	1.75	13.0	11	20.00	+	+
					14.87	-1.01 - 34.59	292.30	0.49 - 57.29	1,950	32.50	2.63				+	+
	AZ	Phoenix	33° 35' N	31° 20' -37° N	23.10	11.91 - 34.59	200.71	0.49 - 37.13	239	15.93	2.66	73.3	17.2	4.70	+	+
	CO	Denver	39° 56' N	37° -41° N	10.35	-1.27 - 23.73	337.99	7.93 - 52.06	771	51.40	2.88	28.1	12.9	8.10	+	+
<i>Northwest</i>	NM	Albuquerque	35° 2' N	31° 20' -37° N	14.39	2.57 - 26.23	225.88	7.71 - 41.71	658	43.87	2.53	27.3	19.5	15.40	+	+
	UT‡	Salt Lake City	40° 46' N	37° - 42° N	11.64	-1.01 - 26.78	404.61	11.85 - 57.29	282	18.80	2.68	41.4	12.1	3.10	+	-
					11.18	-1.01 - 22.85	782.05	6.31 - 172.26	1,626	36.13	2.07				+	-
	ID‡	Boise	43° 43' N	42° -49° N	9.90	-1.01 - 22.85	459.49	6.31 - 68.54	626	41.73	2.68	31.8	15.1	27.20	+	-
<i>West</i>	OR	Portland	45° 35' N	42° -46° 15' N	12.41	4.88 - 20.93	911.66	11.15 - 140.96	304	20.27	2.21	44.4	15.5	17.10	+	-
	WA**	Seattle	47° 27' N	45° 32' - 49° N	11.22	4.92 - 18.77	974.99	15.27 - 172.26	696	46.40	1.64	59.0	12.9	13.00	+	-
					18.45	7.72 - 32.99	220.29	0.073 - 93.45	291	9.70	2.02				+	-
	CA*	Los Angeles	33° 56' N	32° 30' -42° N	17.13	13.8 - 20.91	315.49	0.073 - 93.45	121	8.07	2.58	72.4	15.3	4.20	NA	NA
<i>Total</i>	NV	Las Vegas	36° 2' N	35° - 42° N	19.77	7.72 - 32.99	125.09	0.47 - 27.82	170	11.33	2.47	29.4	14.2	7.10	+	-
		United States Total		24° 30' -49° 23' N	13.34	1.28 - 25.14	955.09	62.59 - 98.33	49,045	3,269.67	2.72					

*states excluded because they span too many latitudes to be accurately represented by weather data from one city

†latitude ranges from: <http://www.netstate.com/>

Data for most states were available from 1997-2011. Exceptions include: †1998-2011; §1999-2011; ‡1997, 1999-2011, and **2001-2011

Table 2: Regression results for overall Generalized Estimating Equations analysis of NNDSS data

Unadjusted IRRs show increase in cryptosporidiosis incidence for a one-unit increase for each variable with no other variables in the model. Adjusted IRRs show increase in cryptosporidiosis incidence for a one-unit increase for each variable, controlling for all other variables in the model. The one-unit increase for temperature and lagged temperature is 1 °C. The one-unit increase for precipitation and lagged precipitation is 1 mm. The one unit increase for percent well use, percent urban, and percent below poverty is 1%. Data is clustered by state to account for association between cases in each state. Temperature and lagged precipitation are included at the bottom of the table for comparison purposes, but were not included in the final model.

<i>Parameter</i>	Unadjusted Incidence Rate Ratio			Adjusted Incidence Rate Ratios		
	<i>IRR</i>	<i>95% Confidence Limits</i>	<i>p-value</i>	<i>IRR</i>	<i>95% Confidence Limits</i>	<i>p-value</i>
<i>Lagged Temperature</i>	1.024	1.021-1.027	<.0001	1.025	1.022-1.028	<.0001
<i>Precipitation</i>	1.0002	1.0000-1.0004	0.0471	1.0001	0.9999-1.0003	0.3238
<i>Latitude</i>	1.029	0.990-1.069	0.144	1.0391	0.988-1.093	0.1347
<i>Percent Total Well Use</i>	3.431	0.512-22.983	0.2038	2.3216	0.336-16.037	0.393
<i>Percent Urban Counties</i>	0.999	0.992-1.006	0.7815	0.999	0.932-1.071	0.9768
<i>Percent of Individuals below Poverty Line</i>	0.978	0.924-1.035	0.4384	0.999	0.992-1.007	0.8059
<i>Temperature</i>	1.022	1.0197-1.0247	<.0001	<i>Temperature was not included in the final model</i>		
<i>Lagged Precipitation</i>	1.0001	0.9999-1.0003	0.4857	<i>Lagged Precipitation was not included in the final model</i>		

Table 3: Regression results for overall Generalized Estimating Equations analysis of global literature review data, developed countries in the northern hemisphere and all studies

Adjusted IRRs show increase in cryptosporidiosis incidence for a one-unit increase for each variable, controlling for all other variables in the model. The one-unit increase for temperature and lagged temperature is 1 °C. The one-unit increase for precipitation and lagged precipitation is 1 mm. Data is clustered by study to account for association between cases from each study location.

<i>Parameter</i>	Developed countries, Northern hemisphere			All studies/Global		
	<i>IRR</i>	<i>95% Confidence Limits</i>	<i>p-value</i>	<i>IRR</i>	<i>95% Confidence Limits</i>	<i>p-value</i>
<i>Temperature</i>	1.026	1.018-1.035	<0.0001	1.030	1.021-1.038	<0.0001
<i>Lagged precipitation</i>	1.007	1.003-1.012	0.00059	1.003	1.000-1.006	0.0649
<i>Population Density</i>	1.004	1.000-1.009	0.07562	1.003	1.000-1.005	0.0288

Figures

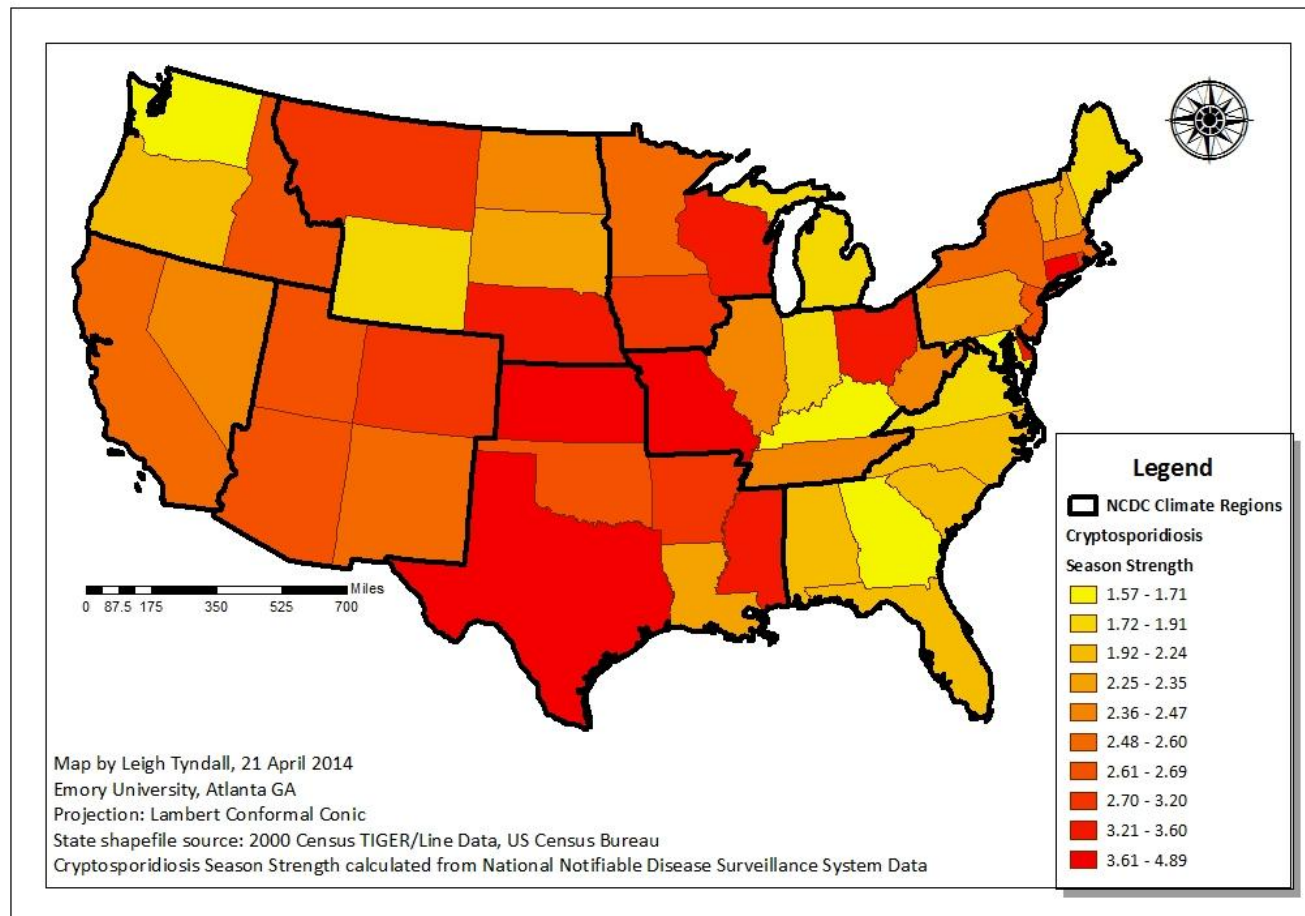


Figure 1.

Cryptosporidiosis Season Strength in the United States: Season strength is defined as the peak to mean ratio, which was calculated using the month with the highest average number of cases over the average number of cases per month for each state and each climate region. NCDC Climate regions are outlined in bold and labeled.

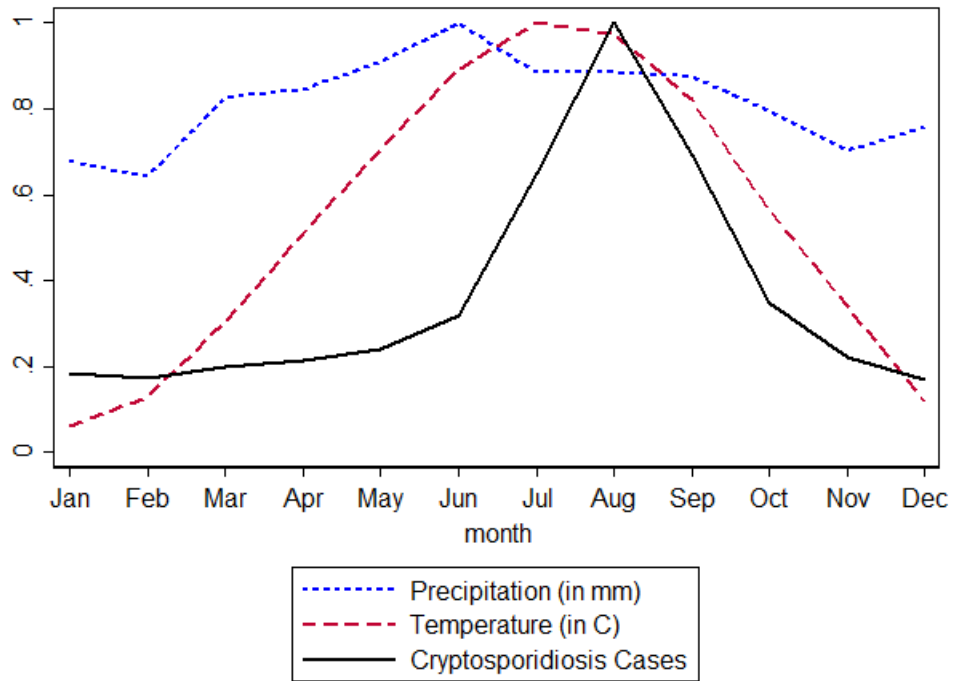


Figure 2.

Average normalized proportion of cryptosporidiosis cases, temperature, and rainfall by month in the United States, 1997-2011.

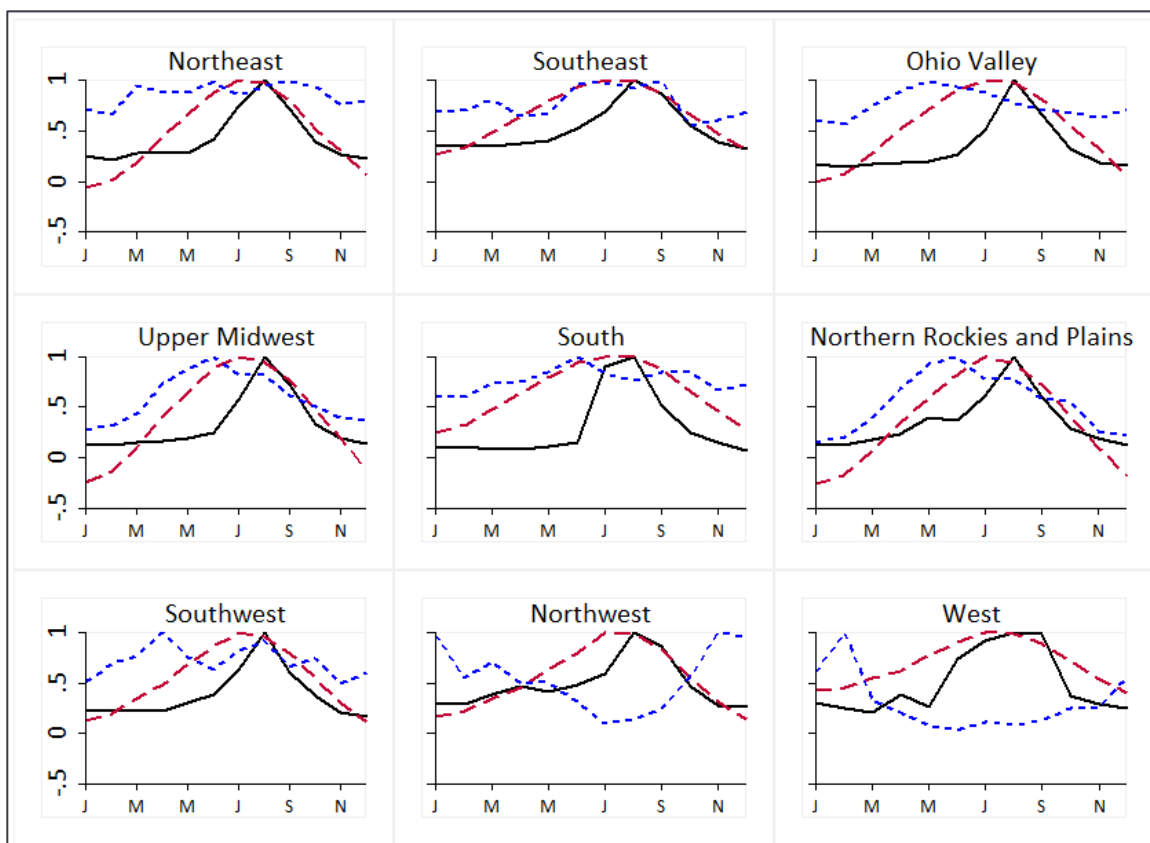


Figure 3.

Average normalized proportion of cryptosporidiosis cases, temperature, and rainfall by month for each US climate region, 1997-2011.

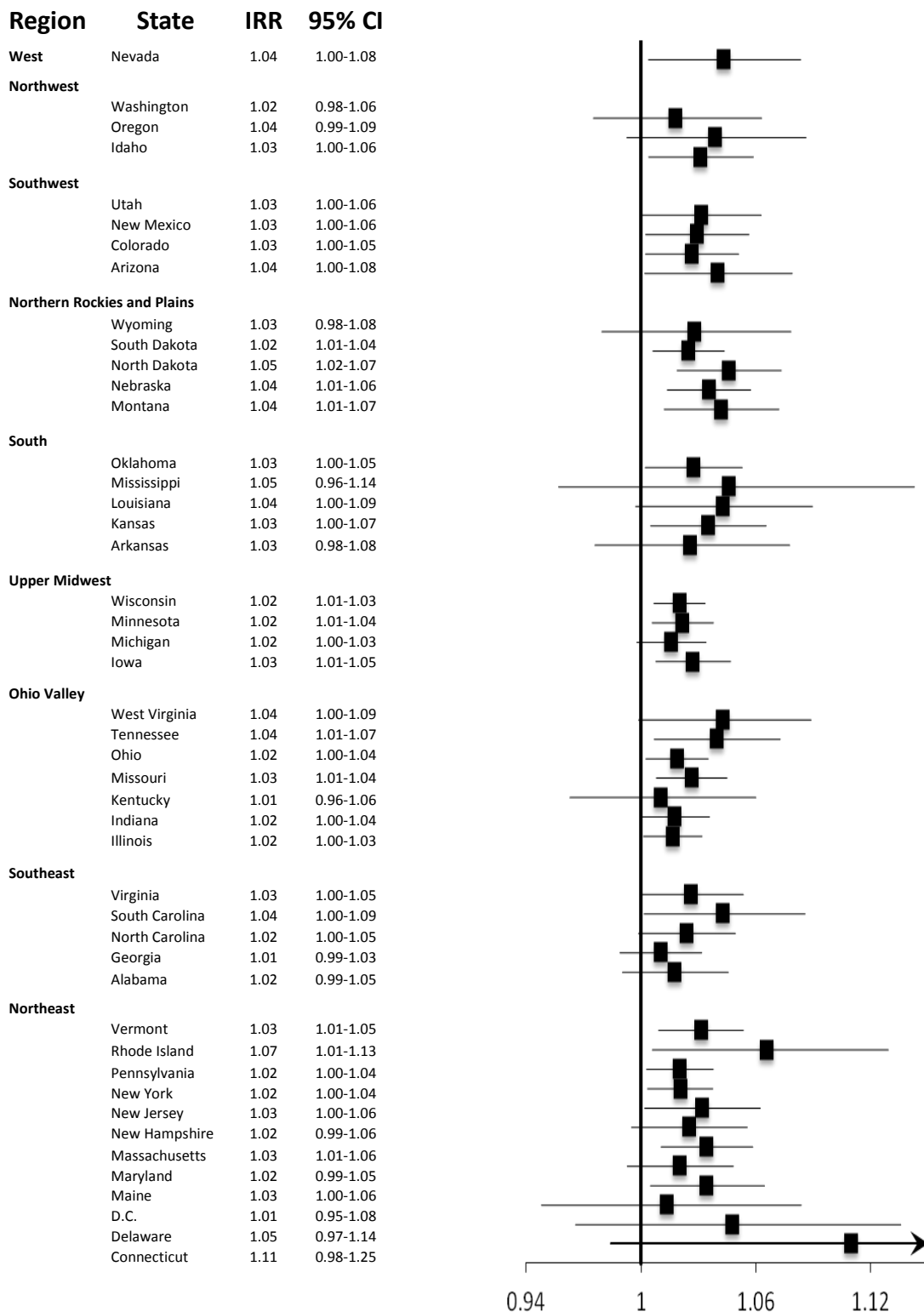


Figure 4a.

Results of Individual State analysis. a. Temperature: Incidence Rate Ratios in each state for a 1 °C change in temperature are shown with their 95% Confidence intervals, displayed by climate region.

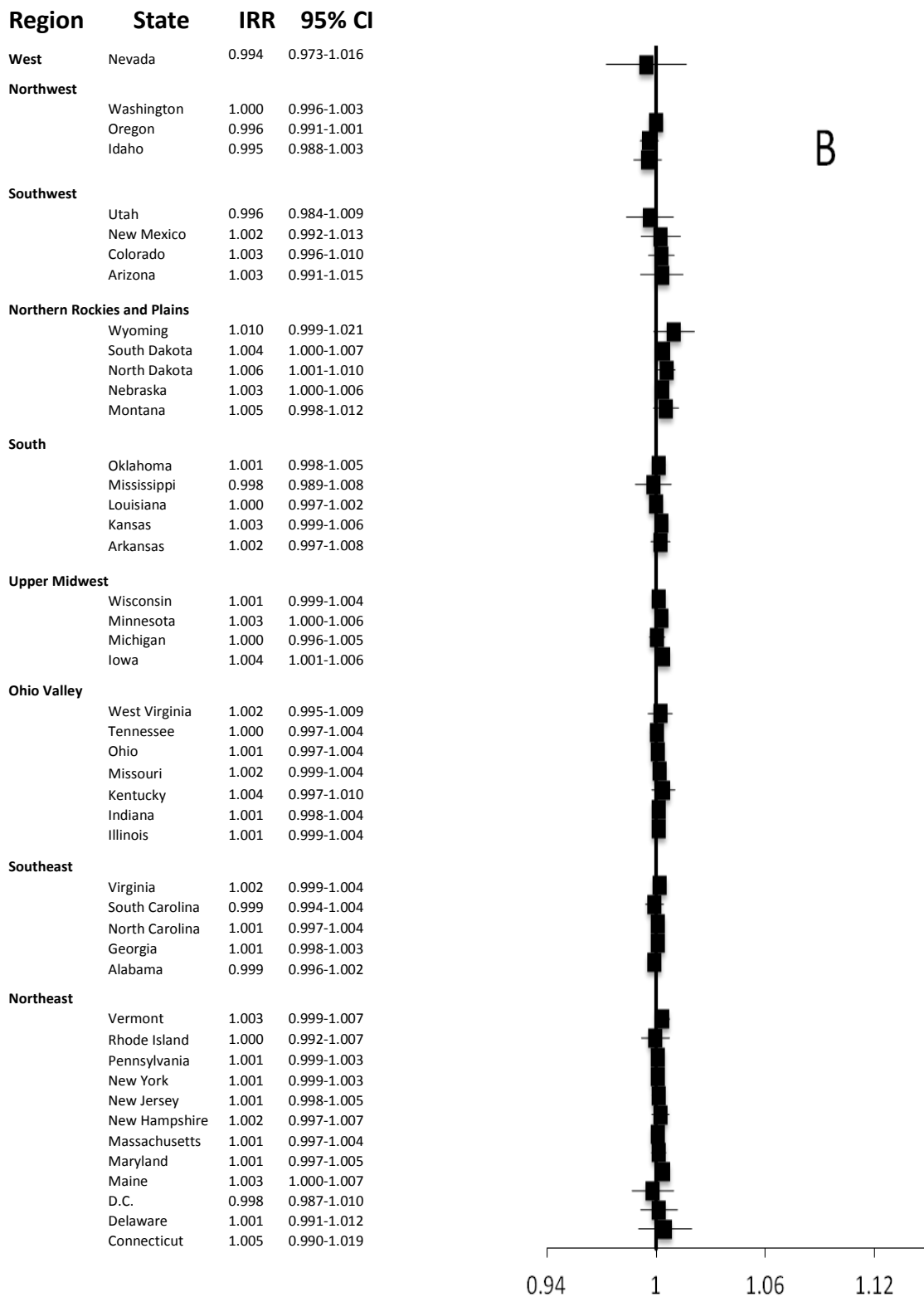


Figure 4b. Results of Individual State analysis. b. Precipitation: Incidence Rate Ratios in each state for a 1 mm change in rainfall are shown with their 95% Confidence Intervals, displayed by climate region.

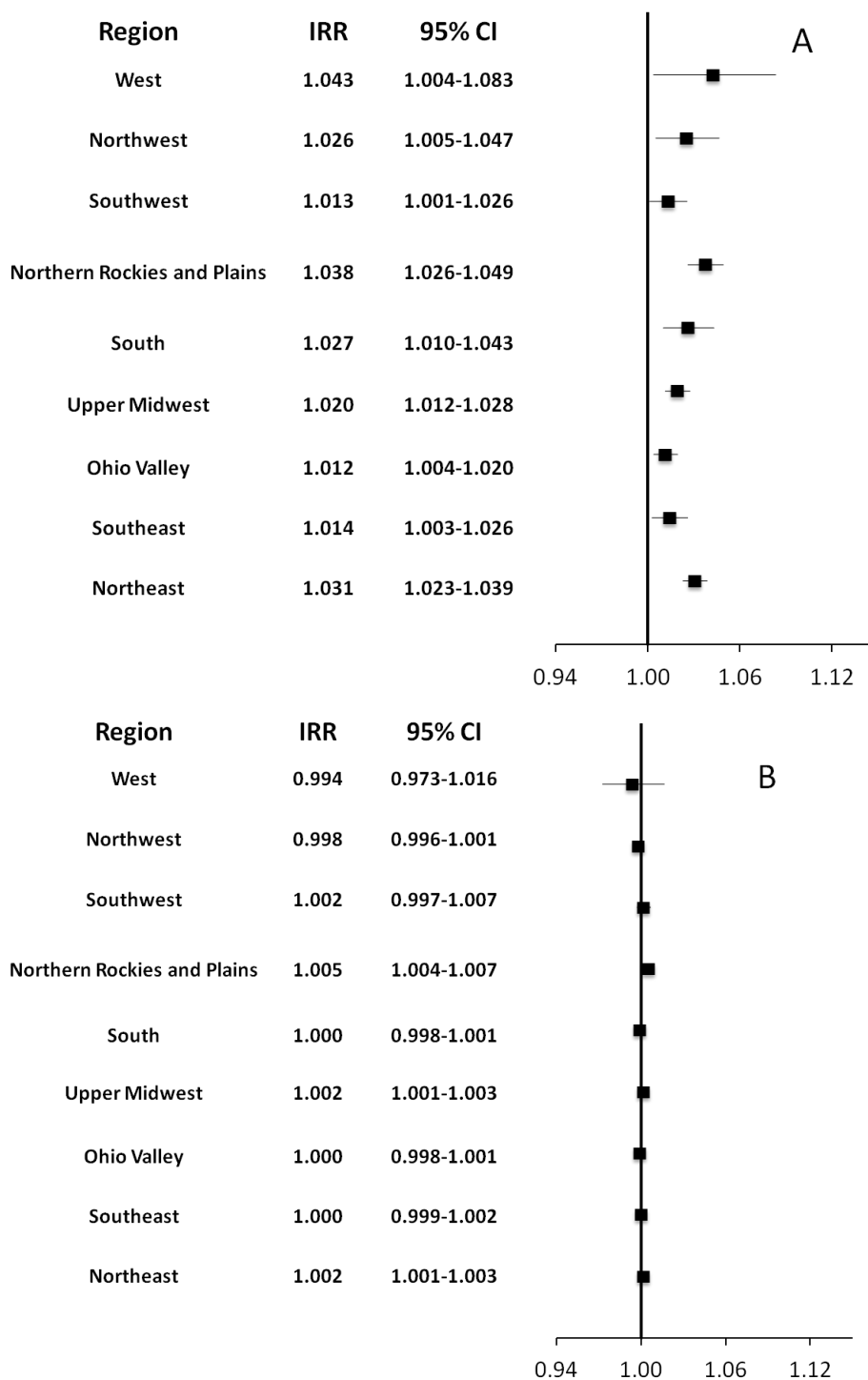


Figure 5 a and b.

Results of Climate Region analysis. a. Temperature: Incidence Rate Ratios in each climate region for a 1 °C change in temperature are shown with their 95% Confidence Intervals, displayed by climate region. b. Precipitation: Incidence Rate Ratios in each climate region for a 1 mm change in rainfall are shown with their 95% Confidence Intervals, displayed by climate region.

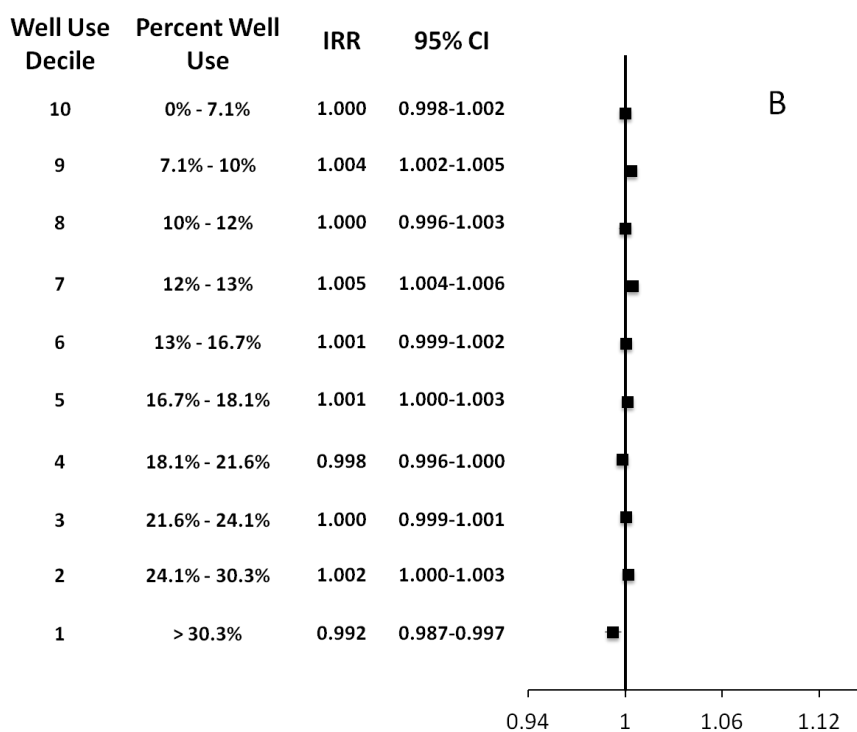
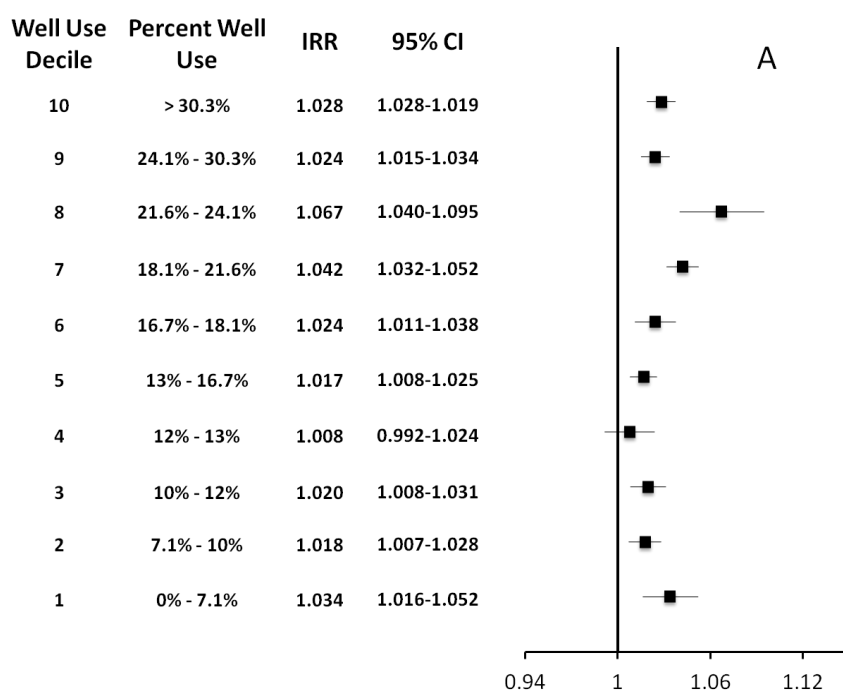


Figure 6 a and b.

Results of Well Use analysis. a. Temperature: Incidence Rate Ratios in each decile of percentage well use for a 1 °C change in temperature are shown with their 95% Confidence Intervals, displayed by well use decile. b. Precipitation: Incidence Rate Ratios in each decile of percentage well use for a 1 mm change in rainfall are shown with their 95% Confidence Intervals, displayed by well use deciles.

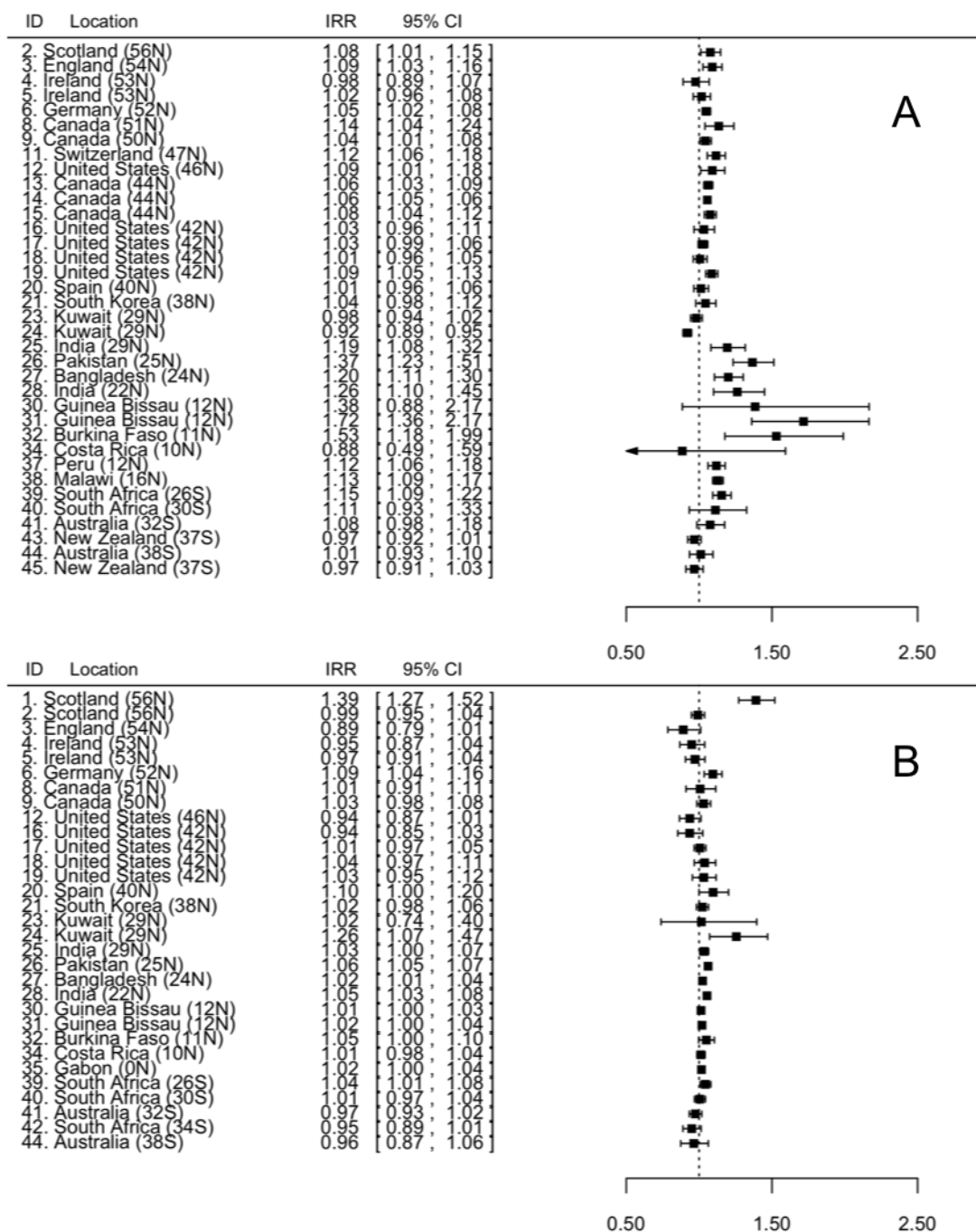


Figure 7 a and b.

Results of Analysis of Literature Review Dataset (Levy *et al.* unpublished data). Temperature: Incidence Rate Ratios for each study for a 1 °C change in temperature are shown with their 95% Confidence Intervals, displayed by latitude. b. Precipitation: Incidence Rate Ratios for each study for a 1 mm change in rainfall are shown with their 95% Confidence Intervals, displayed by latitude.

Appendices

Appendix A. Detailed search protocol

Creating the Complete (“No Deletes”) Pathogen Library

- 1) Go into EndNote > File > Save As > Title “Cryptosporidium_No_Deletes” > Hit Enter
- 2) Right-click on “My Groups” on left-hand side > Create Group > Title “[Climate-Related Term].” Create a group for each of these terms.
 - a) Climate-related term list:
 - i) Season* (wild-card to encompass all variations on ‘season’ and ‘seasonality’)
 - ii) Rain
 - iii) Ambient Temperature
 - iv) Relative Humidity
 - v) Climate
 - vi) Weather
- 3) Access PubMed through Emory Libraries website
- 4) In the “Search” bar at the top of the page, type in the following string: ‘cryptosporidium” OR “cryptosporidiosis” AND [climate-related term]’
- 5) Under the ‘Send to’ link on the right hand side of the page, choose ‘Citation manager’ as the destination
 - a) Select the number of citations to send. This is based on how many citations the search found.
 - i) If there are fewer than 200 citations, set ‘200’ as number to send and click create file
 - ii) If there are more than 200 citations, set ‘200’ as number to send and click create file. Repeat the process, but this time make sure the ‘start from citation’ number is updated to 201, 401, etc
 - b) Clicking ‘create file’ will download all the citation information in one file.
- 6) Open the file with EndNote and the citations will import
- 7) Click on the “Unfiled” or ‘Imported” group at the top of the left column. Highlight all references and drag into the [climate-related term] group. Note date, time, and number of hits for each search.
- 8) Repeat for each climate-related term
- 9) Click on “All References” at the top of the left column. Go to References > Find duplicates > press CANCEL > all duplicates will be highlighted > right-click and move duplicate references to Trash. Note number of duplicates and number of “All References” after deleting duplicates.
- 10) After deleting duplicate > Right-click on trash > Empty Trash

Creating the “Deletes” Library

- 1) Go to the “Cryptosporidium_No_Deletes” library. File > Save a Copy > rename it “Cryptosporidium_Deletes”
- 2) Right click on “My Groups” on the left-hand side > Delete Group Set
 - a) This removes all studies from their groups and puts them back in ‘Unfiled’
- 3) Right click on “My Groups” > Create New Group > title “Based on Time”
- 4) Repeat step 3: title “Based on Title/Abstract”

- 5) Begin excluding irrelevant references from the main library based on examining titles and abstracts ONLY at this point
 - a) Reject based on Title/Abstract if the subject matter is not appropriate
 - i) water quality studies
 - ii) animal studies
 - iii) lab studies
 - iv) studies on HIV+/immunocompromised populations only
 - v) no mention of cryptosporidiosis
 - vi) outbreak-only data
 - vii) review
 - b) Reject based on Time if study:
 - i) Has less than one full year of data (12 months)
 - ii) Does not report weekly or monthly incidence of diarrheal disease/pathogen incidence
- 6) Click and drag each reference that is being excluded into the appropriate group based on the reason for exclusion.

Creating the Combined Library

- 1) File > New > Library. Then File > Save As > “Cryptosporidium_Compare”
- 2) Right-click on “My Groups” > Create New Group > create groups with the following titles:
 - a) “Agree-Include”
 - b) “X includes, Y does not”
 - c) “Y includes, X does not”
 - d) “Pending/Ask PI”
- 3) Compare individuals’ initial libraries and place the references into the appropriate category
 - a) create a new library called “Common Papers b/w X & Y” and copy all references from the “Common Papers b/w X & Y” group into this library
 - b) Copy all of the references from individual X’s library into the “Common Papers b/w X & Y” library
 - c) Search for duplicates, delete these, and file the remaining references into the “Differences b/w X & Common Papers” group in the “Cryptosporidium_Combined” library
 - d) Delete duplicates in the “Cryptosporidium_Combined” library
 - e) Repeat steps a-d for individual Y
- 4) Both partners should go through the two “Differences” groups and decide to include or exclude the references. Move included references into ‘Agree’, create a new group for excluded references (‘Exclude’), and move references you disagree on into the “Conflict” group.
- 5) Go through the “Conflict” library with the PI, and move references into ‘Include’ or ‘Exclude’ as appropriate.

PDF Search

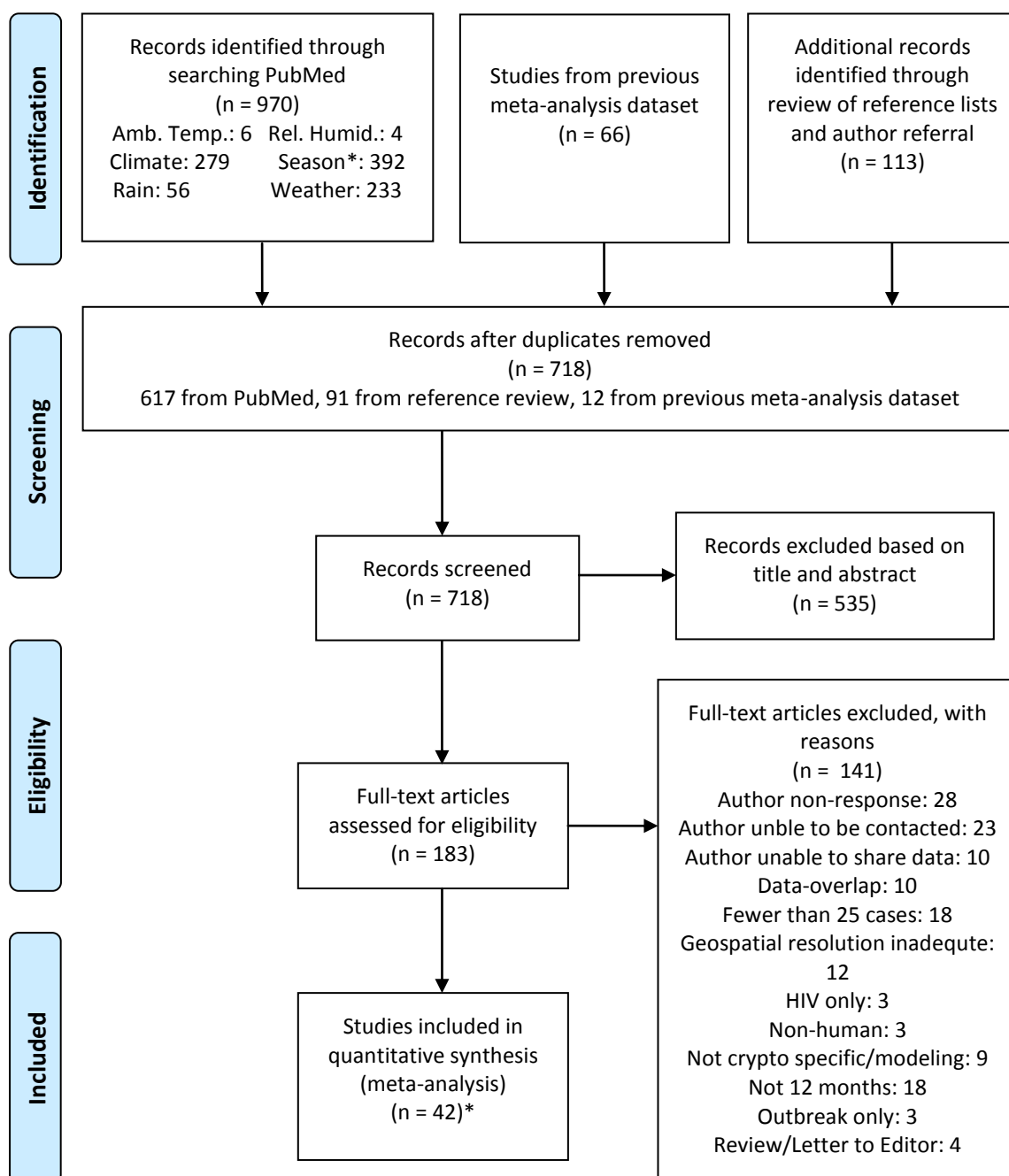
- 1) File > New Library. Then File > Save As > title “Cryptosporidium_PDF_Search”
- 2) Copy citations for all articles agreed upon for full review into this library.
- 3) When a pdf is found, save the document in a separate folder, and also attach it to the reference in EndNote.

- 4) Select all citations agreed upon for full review. Right click > Find Full Text > Find Full Text. This will occasionally bring up the pdfs, more often will bring up a URL, and most often will be unable to find the pdf.
- 5) Search through Pubmed through the Emory library system. This works best on campus, logged into the University network, and logged into the library system.
- 6) When articles are not available online, manually search the Emory library system using reference information from the EndNote library.
- 7) For articles that are only available in print, go to the library and make photocopies of the relevant articles. Scan these copies, save them as pdf's, and upload them to the appropriate Endnote reference.
- 8) For articles that are not available at Emory, go to <http://www.library.emory.edu/uhtbin/nph-illiad> > sign in using Emory username and password > under Create New Request, click on "Copy of Article." Fill out the request. Be sure to use the full name of the journal, not the abbreviation

PDF Exclusions

- 1) In the library "Cryptosporidium_PDF_Search_X" create the following group sets and groups:
 - a) Set "Exclude"
 - i) Not_full_year
 - ii) Not_monthly_data
 - iii) No_data
 - iv) Outbreak (depending on pathogen)
 - v) Under 25 cases
 - vi) HIV/immunocompromised populations
 - vii) Non-human studies
 - b) Set "Use"
 - i) To Extract
 - ii) Maybe
 - iii) Have Extracted
 - iv) Can't Extract
- 2) Go through each pdf. Place excluded references in the appropriate group based on why it was excluded, and place all included references in the "To Extract" group
- 3) Go to the separate folder of all pdf's. Create the following sub folders:
 - a) Cryptosporidium_Yes
 - b) Cryptosporidium_No
 - c) Cryptosporidium_Maybe
 - d) Cryptosporidium_Can't Extract
- 4) Transfer all pdf's into the appropriate subfolder.
 - a) All references from the "Exclude" set group get placed in the "Cryptosporidium_No" folder
 - b) All references from the "To Extract" group get placed in the "Cryptosporidium_Yes" folder
 - c) All references from the "Maybe" group get placed in the "Cryptosporidium_Maybe" folder
- 5) Send the "Cryptosporidium_Maybe" folder to the PI for final review and shift references and pdf's to appropriate groups and subfolders

Appendix B. PRISMA flow diagram



Appendix C. Registration Information and Data Use Restrictions Agreement

Data Release Guidelines for the

Page 1 of 3

Data Release Guidelines for the
National Notifiable Diseases Surveillance System
January 31, 2000
Updated, November 2007

I Introduction

National Notifiable Diseases Surveillance System (NNDSS) data are collected and reported voluntarily to CDC by U.S. states and territories. The Council of State and Territorial Epidemiologists (CSTE) represents the collective voice of the U.S. states and territories with regard to the development of surveillance systems for conditions/diseases that have public health importance, the development of surveillance data confidentiality protections, and other important issues. A fundamental premise in effective public health practice is that surveillance data should be useful for the control and prevention of disease and disability while at the same time protecting confidentiality. Recent developments in telecommunications and computerization have greatly enhanced our ability to compile and share data. In 1996, CDC and CSTE jointly negotiated a policy for the release of data from CDC's NNDSS so as to facilitate the public health community's access to NNDSS data while at the same time preserving data confidentiality. This document describes the 1996 policy that CDC and CSTE developed for the release of CDC's NNDSS data. (Based on needs expressed by CSTE in 1996, the Indian Health Service has special status in terms of access to the NNDSS data [refer to section II item D])

Since 1991, all 50 states, two cities, and four United States Territories have transmitted NNDSS data through the National Electronic Telecommunications System for Surveillance (NETSS). The National Center for Public Health Informatics (NCPHI), Division of Integrated Surveillance Systems and Services (DISSS) administers and maintains the NNDSS, NEDSS and NETSS at CDC. NETSS allows for the collection and transmission of case-specific data; prior to 1991, only aggregate case counts were reported to CDC by most states. CDC updates and NNDSS data sets annually when final data are available from the states. While acquired immunodeficiency syndrome (AIDS), tuberculosis (TB), sexually-transmitted diseases (STD), and human immunodeficiency virus (HIV) infection are included in the list of nationally notifiable conditions, data on these conditions are reported separately by state programs with responsibility for these diseases to the respective CDC program. Cases of AIDS and HIV infection are reported directly to the AIDS/HIV Surveillance Program, while STD and TB are reported both to the responsible CDC program and to DISSS/NCPHI through NETSS (which sometimes leads to differences in case counts). A revised data release policy exclusively for AIDS (available elsewhere) was approved by the CSTE Executive Committee in 1995.

II Data release specifications of NNDSS data, and referrals to other CDC programs for AIDS/HIV, STD and TB data requests

A To be made available as a limited use data set upon completion of the approval process (refer to section III below on access to the NNDSS data base)

A line-listed case-specific data file of all NNDSS cases^[1] (final data only, of conditions designated as nationally notifiable) reported to CDC through NETSS or NEDSS will be made available in the format of a limited use data set. Since the NNDSS data may vary in the variables reported and in completeness by state and by disease/condition, CDC will provide suitable documentation to ensure that users understand the limitations of the surveillance data. The variables in the data file may include reporting state; disease/condition; year; event month; age grouped as <1 yr, 1-4 yrs, and by 5 year intervals thereafter; sex; race; and ethnicity. Data by county or any geographic unit smaller than state, cannot be released. New York City (encompassing a 5 county area) and Washington, D.C. report data directly through NETSS and their NNDSS data are handled as if they were states. If the total number of cases in a state in a year for a given disease/condition is ≤ 3 , then race and ethnicity (but no other variables of those listed above) will be suppressed for all cases in that state in that year.

B To be made available to the public as summary or aggregate count files

The following NNDSS data¹ (finalized data only, of conditions designated as nationally notifiable) may be made available depending on the date that reporting began for the disease/condition:

- Total counts of a particular disease/condition by state by month by year.
- Total counts of a particular disease/condition by county by year (but not month).

C Data requests from agencies, institutions, or persons outside CDC, including other federal agencies:

There will be no release of data in formats other than those described in section II items A and B above, unless the format is more restrictive than described in section II items A and B. Requests for data must be made using the form (attached) "Registration Information and Data Use Restrictions Agreement". Any agency, institution, or person (including other federal agencies) seeking from CDC more detailed data than available in the data sets described in section II items A and B above will be directed by CDC to each state so that data requestors can negotiate data release and obtain data directly from individual states.

D Use of NNDSS data by CDC programs and the Epidemiology Branch of the Indian Health Service

Provisional as well as final NNDSS data^[2] in the format of case reports, detailed line-listed data files, and extended case records, may be released to CDC programs and the Epidemiology Branch of the Indian Health Service (IHS). CDC and IHS data requestors must complete the form "Registration Information and Data Use Restrictions Agreement" (attached) before data access can be granted. This agreement is intended to ensure that NNDSS data are used in a manner consistent with the jointly developed CDC/CSTE data release and use policy. In addition, the agreement is intended to prevent CDC and IHS programs or legal agents of these organizations from subsequently releasing case reports, extended records, or detailed line-listed NNDSS data to other users outside CDC and IHS (including Tribal Nations) in a format inconsistent with this policy. Provisional NNDSS data, other than that which is already publicly available (see section IV, item 1 below), cannot be published or released to the public under any conditions. Finalized NNDSS data must not be published or released in a form showing frequencies or cross-tabulations with more detailed information than could be obtained from the data files described in section II items A and B above.

E Referrals for access to AIDS/HIV, STD, and TB data:

Individuals interested in accessing STD data should contact CDC's Division of STD Prevention, Statistics and Data Management Branch by calling (404) 639-8356.

Individuals interested in accessing TB data should call (404) 639-5309 to contact Valerie Robison, DDS, MPH, PhD, Chief, Surveillance Team, Surveillance, Epidemiology and Outbreak Branch, Division of Tuberculosis Elimination.

Individuals interested in accessing HIV/AIDS data should call (404) 639-2050 to contact the Research and Dissemination Team, HIV Incidence and Case Surveillance Branch, Division of HIV/AIDS Prevention.

III Access to the data base

The CDC's NCPHI/DISSS NNDSS team is charged with the responsibility of maintaining the confidentiality and security of the NNDSS surveillance data base. Access to the data base for individuals outside the NCPHI/DISSS can be granted in accordance with the CDC and CSTE data use restrictions policy on an as needed basis, after completing the form "Registration Information and Data Use Restrictions Agreement" (attached) and subsequently receiving approval from the NNDSS team leader in NCPHI/DISSS. Also included in this document is a summary of the NNDSS conditions/diseases and corresponding years for which these conditions/diseases are included in the NNDSS historical data files (refer to Addendum A). This information is provided to assist individuals in formulating their data requests.

IV Alternative methods of access to the NNDSS surveillance data base

1. CDC's Morbidity and Mortality Weekly Report (MMWR) includes a collection of tables presenting provisional NNDSS case counts in the United States and the results of epidemiologic investigations of these and other cases. The report includes either national or state-specific provisional data depending on the disease/condition, by week of report and provides comparison data from the previous year if data are available. These data tables are available at the following CDC MMWR Internet site: http://www.cdc.gov/mmwr/mmwr_wk.html
2. The MMWR Summary of Notifiable Diseases, United States is published each year and has final NNDSS data, either at the national- or state-specific level

<http://isd-v-ncph-nnd/NNDSS/FSI%20Documents/EPO%20NNDSS%20Data%20Releas...> 10/31/2013

2/4

Data Release Guidelines for the

Registration Information and Data Use Restrictions Agreement (RIDURA) Form
for the
National Notifiable Diseases Surveillance System (NNDSS)

A Registration Information

NOTE: Access to NNDSS data is limited to the "Requestor" who reviewed the NNDSS data release guidelines, completed this RIDURA form, and received approval from NCPHI/DISS Surveillance Systems Operations Team. If additional people (including those within a specific CDC program or external to CDC) need data access, they also need to review the NNDSS data release guidelines, complete a RIDURA form, and receive approval from NCPHI/DISS's NNDSS team to access the data.

Requestor's Name: LEIGH TYNDAL
 Requestor's Affiliation: EMORY UNIVERSITY (STUDENT INTER) AT NCE210/DFWED/WATERBORNE DISEASE PREVENTION BRANCH
 Mailing Address: 1659 BRIDDLIET ROAD
ATLANTA, GA 30309
 Telephone number: 404.924.0550
 Fax number: _____
 E-mail address: L.M.TYNDAL@EMORY.EDU
 Mainframe User ID (For CDC data users only): XGQZ
 Instructions for transferring the data set to CDC data requestor: _____

 CDC Data Requestor's Supervisor's Name and Signature.

B Data Use Restrictions Agreement

I have read and agree to follow the guidelines listed below. These guidelines were jointly developed by CDC and the Council of State and Territorial Epidemiologists. They prohibit the knowing disclosure of any information that could be used directly or indirectly to identify individuals. In addition, the guidelines represent a balance between the potential for inadvertent disclosure and the need for CDC to be responsive to information requests having legitimate public health application.

In accepting access to the NNDSS data, I agree to the following:

1. I am permitted to release final but not provisional national and regional tabulations from the NNDSS case data base (excluding AIDS/HIV, STD, and TB case data) in either narrative or tabular format.
2. I am permitted to include in presentations, slides, interviews, and publications final NNDSS data in tabular or narrative format that reports information on the following variables: disease/condition, year, event month, age grouped as <1 yr, 1-4 yrs, and by 5 year intervals thereafter; sex; race; ethnicity; and state. If the total number of cases in a state in a year for a given disease/condition is #3, then race and ethnicity at a minimum will be suppressed for all cases in that state in that year.
1. I will not disclose or otherwise make public NNDSS county level data or data on any geographic unit smaller than state from case-specific NNDSS data files.
2. I understand that release of data not specifically permitted by this agreement is prohibited unless an exemption in writing is first obtained from the Chief Surveillance Systems Operations Team, NCPHI/DISS/CDC.
5. I agree that access to NNDSS data is limited to the "Requestor" named within this RIDURA form. I agree to refer all requests for access to NNDSS data sets to the Surveillance Systems Operations Team.
6. I agree to use the data in the NNDSS data sets for statistical reporting and analysis only.
7. I agree to make no disclosure or use of the identity of a person discovered inadvertently and will advise the NCPHI/DISS/NNDSS team of any such discovery.
1. I agree not to deliberately combine NNDSS data sets or alternatively, combine an NNDSS data set with a non-NNDSS data set for the purpose of matching records to identify individuals.
2. I also agree to the following security practices:
 - a) I will password protect the NNDSS data file(s) I receive. In addition, any temporary or permanent analysis files, such as those produced by SAS or other statistical package, will be password protected as well.
 - b) I will treat all data at my desk site confidentially and will not give other persons access to this data.
 - c) I will keep all hard copies of data runs containing small cells (i.e., #3) locked in my desk when not in use, shredding them when they are no longer necessary to my analysis. In addition, I will review all printed or electronic output and delete or blackout any direct or indirect identifiers and any small cell counts (#3) that involve race or ethnicity.
 - d) I will not produce any copies of the data file(s), even for back-up purposes.
 - e) I am responsible for obtaining IRB review of projects when appropriate.
10. Attach an addendum to this data use restrictions agreement briefly stating the purpose of investigation and providing an outline of the proposed analysis, including names of NNDSS variables to be used as well as conditions/diseases and time period for which data are needed.
3. CDC staff who write an MMWR article based on NNDSS data analyses should send a courtesy copy of the completed MMWR article to the Chief, Surveillance Systems Operations Team at the same time you transmit a copy of the completed paper to the MMWR office.

Signature: [Signature] Date: 8 Nov 2013

Return your completed form to

National Center for Public Health Informatics
 Division of Integrated Surveillance Systems and Services
 Centers for Disease Control and Prevention
 Mail Stop E-91, NNDSS Administrator
 1600 Clifton Road
 Atlanta, Georgia 30333
 Fax: (404) 498-6450

^[1] Excludes AIDS/HIV, STD, and TB data.

^[2] Provisional data are reported to CDC each week throughout the year and are subject to continual updating as new information becomes available to surveillance staff. At the end of a calendar year, NNDSS reporting sites finalize reports of cases for the year and reconcile their state reports with reports sent to CDC throughout the year. The resulting finalized data are published in the MMWR *Summary of Notifiable Diseases, United States*.

3/4

Appendix D. Map of NCDC climate regions

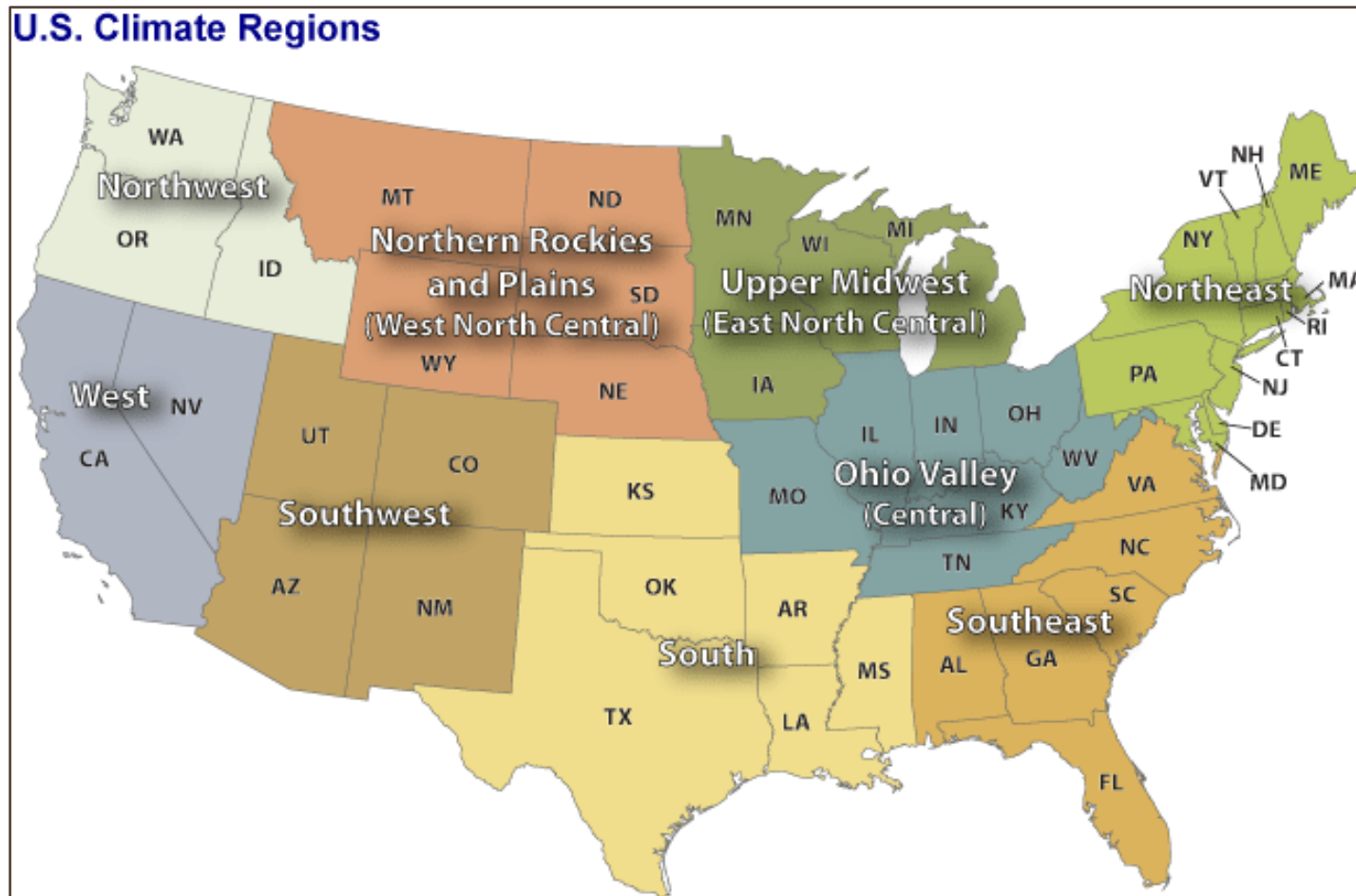


Image obtained from: National Oceanic and Atmospheric Administration, National Climatic Data Center:

Thomas R. Karl and Walter James Koss, 1984: *"Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983."* Historical Climatology Series 4-3, National Climatic Data Center, Asheville, NC, 38 pp.
<http://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php>

Appendix E. IRB Declaration Letter



EMORY
UNIVERSITY

Institutional Review Board

January 7, 2014

Leigh Tyndall
Rollins School of Public Health
Atlanta, GA 30322

RE: Determination: No IRB Review Required
IRB00071405
The Global Seasonality of Cryptosporidiosis as Compared to Cryptosporidiosis in the United States
Investigator: Leigh Tyndall

Dear Investigator:

Based on a review of the materials you have provided for this study, we have determined that it does not require IRB review because it does not meet the definition of a study involving "human subjects" as set forth in Emory policies and procedures and federal rules, if applicable. Specifically, in this project, you will perform analysis using an NNDSS file containing aggregate cryptosporidiosis data categorized by state. These data are publicly available and are devoid of personal identifiers.

HHS regulations define *human subject* at 45 CFR 46.102(f) as follows:

Human subject means a living individual about whom an investigator (whether professional or student) conducting research obtains:

- (1) data through intervention or interaction with the individual, or
- (2) identifiable private information.

This determination could be affected by substantive changes in the study design, subject populations, or identifiability of data. If the project changes in any substantive way, please contact our office for clarification.

Thank you for consulting the IRB.

Sincerely,

A handwritten signature in blue ink, appearing to read "S Roberts".

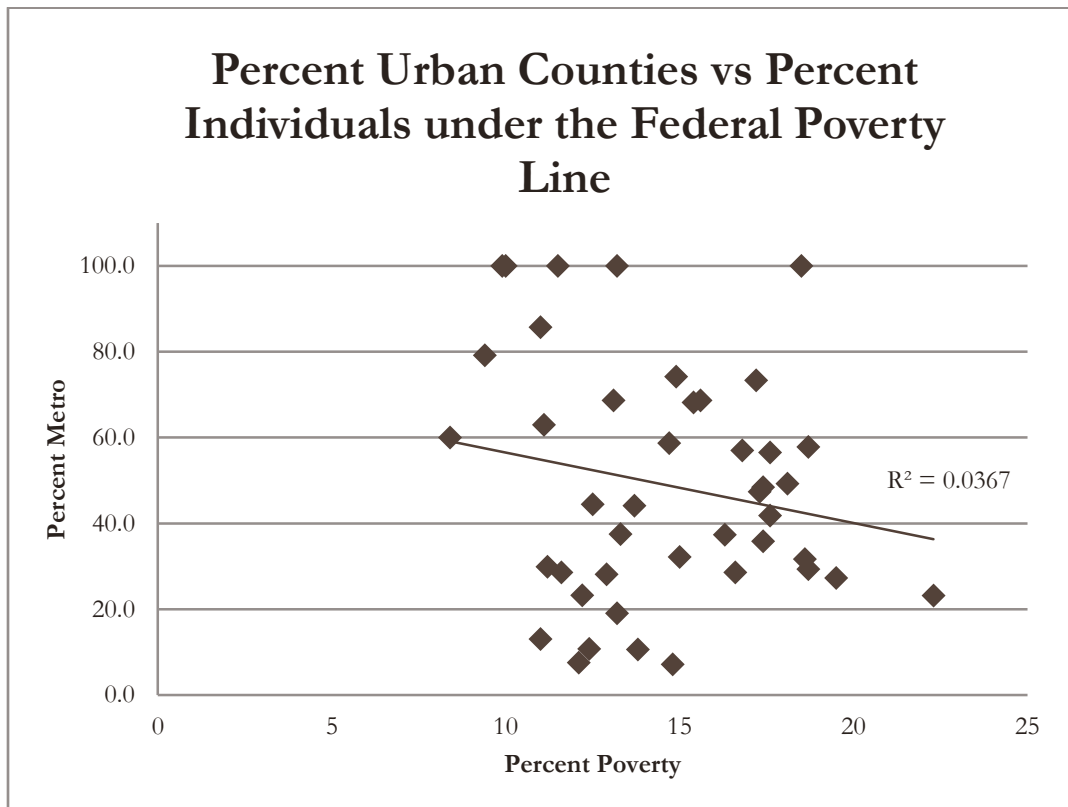
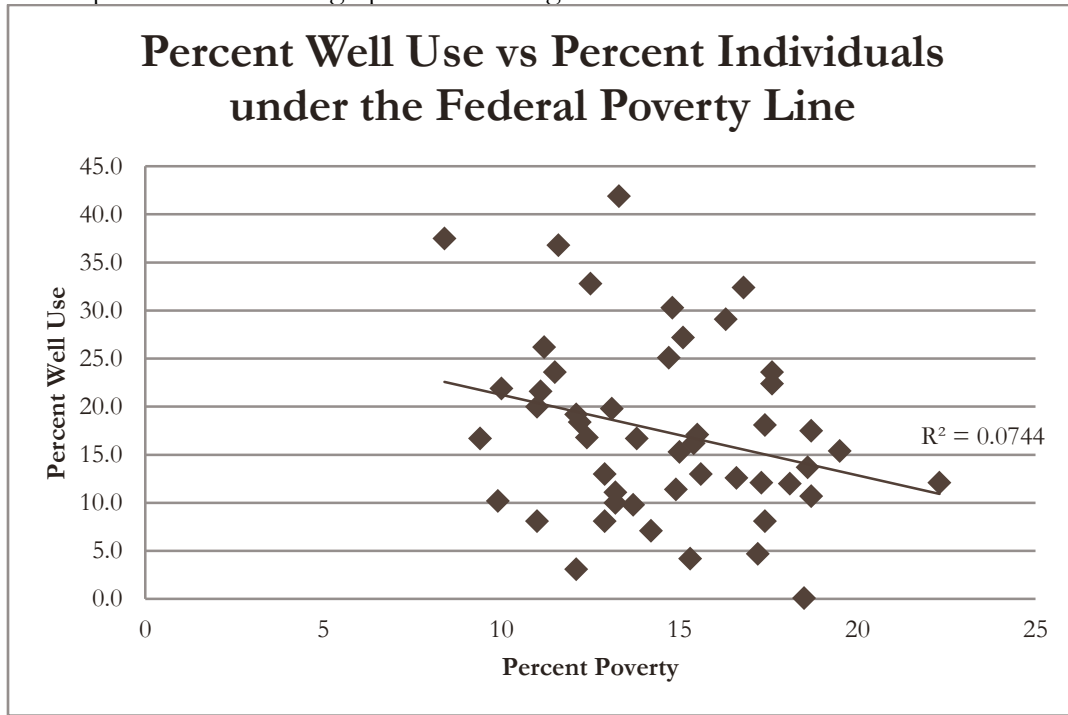
Sam Roberts, BA CIP
Senior Research Protocol Analyst

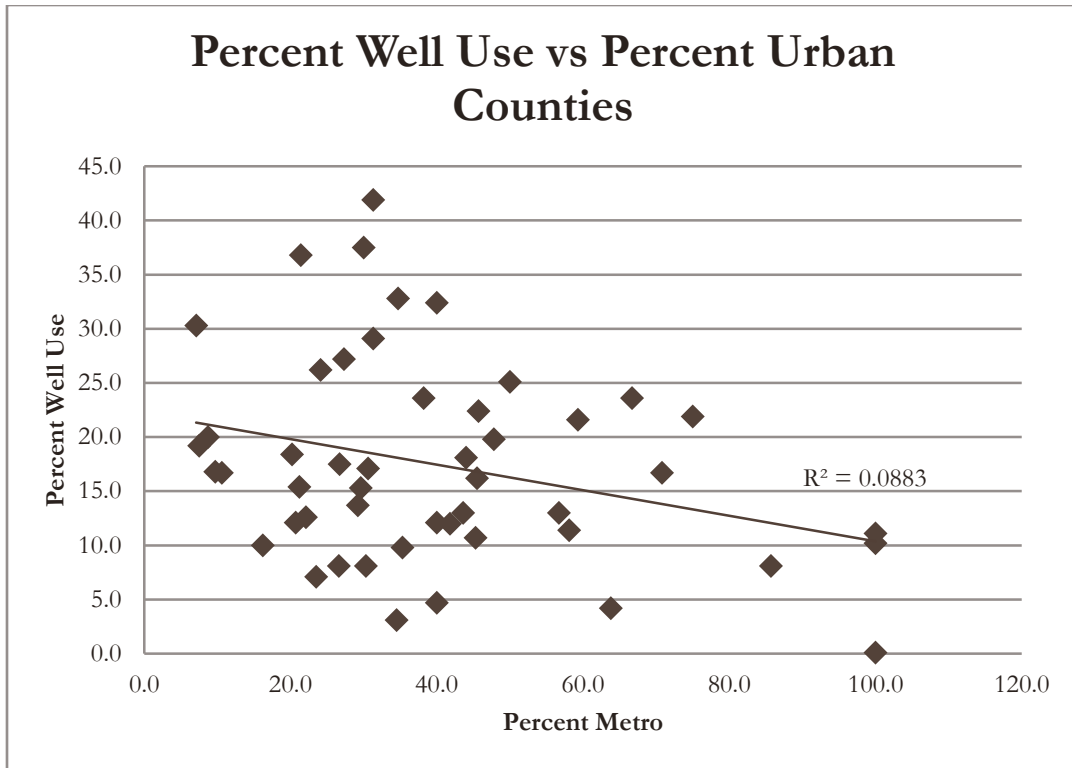
Appendix F. Meta-data table for studies

Study ID	Author	City	Country	Latitude	Zone	Altitude (m)	Avg. yearly rainfall (mm)	Avg. yearly temp (C)	Period of Study	Total Months of Study	Total # crypto cases	Ages included (years)	Study setting	Temp	Rain
34	Achi	Canton de Puriscal, San Jose, Heredia	Costa Rica	09 59N	tropics	920	1790	20	01/82-12/84	24	56	< 2.5	Community and hospital based	-	+
25	Ajampur-Delhi	Delhi	India	28 35N	tropics	215	700	25.2	12/05-12/08	37	34	< 5	Hospital based	++	+
33	Ajampur-Trichy	Trichy	India	10 49N	tropics	73	886.4	26.9	12/05-12/08	37	16	< 5	Hospital based		
29	Ajampur-Vellore	Vellore	India	12 55N	tropics	252	945	26.3	12/05-12/08	37	20	< 5	Hospital based		
10	ANOFEL	NA - (data from Paris)	France	48 44N	N	96	637.4	12.4	01/06-12/09	48	407	all	Country level surveillance		
41	Assadamongkol	Perth	Australia	31 56S	S	20	800	18	01/86-12/88	36	100	< 5	Hospital based	++	-
37	Bennett	Lima	Peru	12 00S	tropics	13	---	20	06/95-08/98	39	133	< 12	Community based	++	+
44	Biggs	Melbourne	Australia	37 48S	S	33	680	14	11/84-10/86	24	55	all	Community based	+	-
22	Chai	Hwasun-gun - data from Chungnim	Korea	35 01N	N	74	1477	12	11/96-10/97	12	77	all	Community based		
7	Chalmers	NA - (data from London)	UK	51 09N	N	61	750	10.3	01/00-12/06	84	12,208	all	Country level surveillance		
4	Corbett-Feeney	NA - (data from Dublin)	Ireland	53 26N	N	85	758	9.8	02/85-02/86	24	41	< 12	Lab based study	-	-
23	Daoud	Kuwait City	Kuwait	29 13N	N	54	---	26	01/88-06/89	18	35	< 8	Hospital based	-	+
28	Das	Kolkata	India	22 34N	tropics	12	1585.7	26.7	01/03-12/04	24	40	< 5	Hospital based	++	++
35	Duong	Libreville	Gabon	00 27N	tropics	14	2100	26	10/89-09/90	12	70	< 2	Community based	++	++
39	Frapp	Pretoria	South Africa	25 44S	S	1299	700	18	10/85-09/89	48	289	all, most < 3	Hospital based	++	++
5	Garvey	NA - (data from Dublin)	Ireland	53 26N	N	85	758	9.8	01/04-12/06	36	1369	all	Country level surveillance	++	-*
3	Goh	Allerdale and Copeland (North Cumbria) - data from Keswick	England	54 37N	N	81	1521	9.4	03/96-02/00	48	152	all	Community based	++	-*
26	Haider	Karachi	Pakistan	24 54N	tropics	21	210	26	01/07-12/07	12	37	all	Lab based	++	++
1	HPS_2008-2011(2015?)	NA - (data from Glasgow)	Scotland	55 52N	N	7	901.7	9.6	01/06-12/10	60	378	all	Country level surveillance	++	++
21	Huh	Gyeonggi-do Province - data from Uijongbu	South Korea	37 43N	N	71	111	12	01/04-12/06	36	29	all	Hospital based	++	+
24	Iqbal	Kuwait City	Kuwait	29 13N	N	54	---	26	09/95-08/97	24	51	3 months-13 years	Hospital based	-*	++
36	Katsumata	Surabaya	Indonesia	07 13S	tropics	3	1470	28	08/92-07/93	12	41	mostly < 5	Hospital and Community based		
43	Lake (ESR data substituted)	NA - (data from Auckland)	New Zealand	37 01S	S	6	1150	15.5	01/96-12/12	204	13310	all	Country level surveillance	-*	
8	Laupland	Calgary Health Region	Canada	51 06N	N	1083	418.8	4.4	05/99-04/02	36	173	all	Lab based	++	+
45	Learmonth	Wellington, Hamilton, Hawkes Bay, Southland, Bay of Plenty, Hutt Valley (data from Wellington)	New Zealand	41 17S	S	128	1220	12.9	08/00-08/03	49	6056	all	Lab based	-*	
11	Mai Nguyen	Bern	Switzerland	46 55N	N	509	1028	7.9	01/86-12/86	12	31	all	Hospital based	++	
13	Majowicz	Ontario - data from Toronto	Canada	43 40N	N	172	852.9	7.7	01/96-12/97	24	451	all	Province level surveillance	++	
9	Mann	Winnipeg, Manitoba	Canada	49 54N	N	238	510	2	10/83-10/84	11	39	all	Lab based	++	+
20	Merino	Madrid	Spain	40 27N	N	581	450	14.1	11/85-02/88	28	41	all	Hospital based	+	+
40	Miller	Durban	South Africa	29 58S	S	7	1050	21	08/85-07/86	12	29	< 2	Hospital based	+	+
30	Molbak	Bissau	Guinea Bissau	11 53N	S	35	---	27	04/87-03/90	36	239	< 4	Community based	++	++
32	Nacro	Bobo Dioulasso	Burkina Faso	11 10N	tropics	459	1088	27	01/95-12/95	12	72	< 3	Hospital based	++	++
17	Naumova-Boston	Boston	United States	42 22N	N	1	1071	10.7	01/93-12/02	48	102	all	State level surveillance	++	+
16	Naumova-Lowell	Lowell	United States	42 39N	N	33	1101	9.4	01/93-12/02	48	18	all	State level surveillance	+	-
19	Naumova-Worcester	Worcester	United States	42 16N	N	299	1061	8.9	01/93-12/02	48	58	all	State level surveillance	++	+
42	Nel	Cape Town	South Africa	33 59S	S	42	580	17	06/04-05/05	24	63	< 18	Hospital based	+	-
38	Peng	Blantyre	Malawi	15 40S	tropics	777	880	22	08/97-03/99	20	69	< 2.5	Hospital based	+	
31	Perch	Bissau	Guinea Bissau	11 53N	S	35	---	27	01/91-12/97	84	351	< 5	Community based	++	++
15	Pintar	Waterloo region, Ontario (data from Waterloo)	Canada	43 29N	N	328	940	---	04/05-12/07	33	36	all	Region level surveillance	++	+
27	Shahid	Dhaka	Bangladesh	23 46N	tropics	9	2148	26.1	01/84-12/84	12	71	all	Hospital Based	++	++
2	Shepherd	Paisley (data from Erskine)	Scotland	55 52N	N	8	1165	7.8	01/86-12/87	24	83	all	Lab based	++	-
12	Skeels	Oregon (data from Portland)	United States	45 36N	N	6	914.4	12.5	01/85-12/88	48	32	all	Lab based	++	-*
6	Terletskaia (RK data substituted)	NA - (data from Berlin)	Germany	52 23N	N	47	566	9	01/01-12/06	72	9044	all	Country level surveillance	++	++
14	Vrbova	Ontario - (data from Toronto)	Canada	43 40N	N	172	852.9	7.7	01/07-12/12	36	473	all	Province level surveillance	++	+
18	Wolfson	Boston, MA	United States	42 22N	N	1	6.4	10.7	02/83-01/84	12	43	all	Lab based	+	+

Appendix H. Correlations between demographic variables and plots

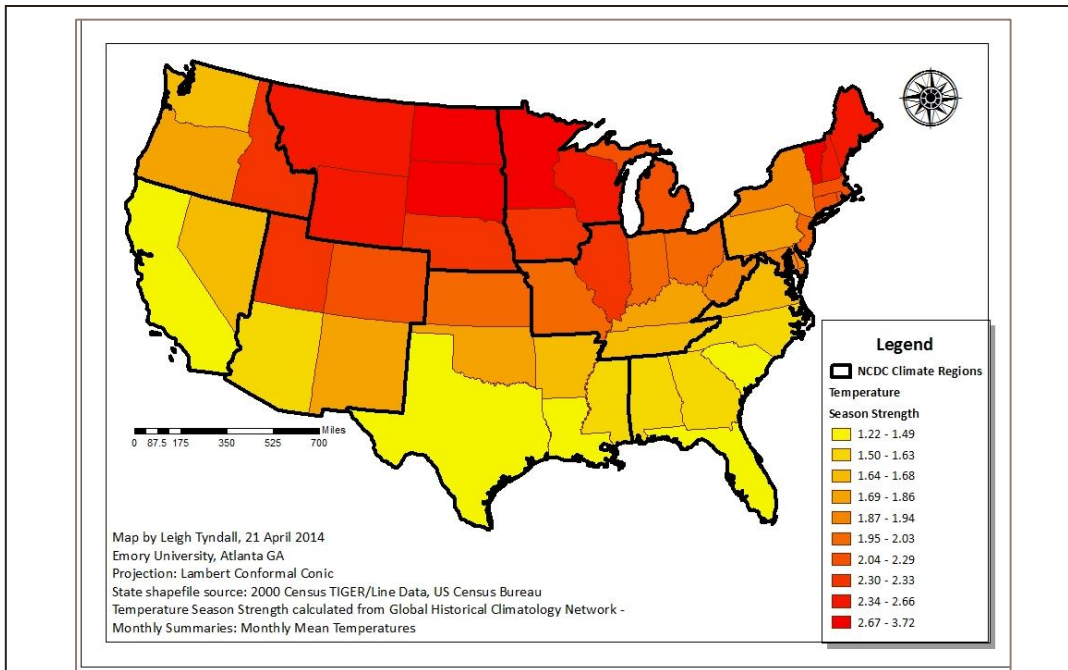
Scatterplots of each demographic variable against the others:



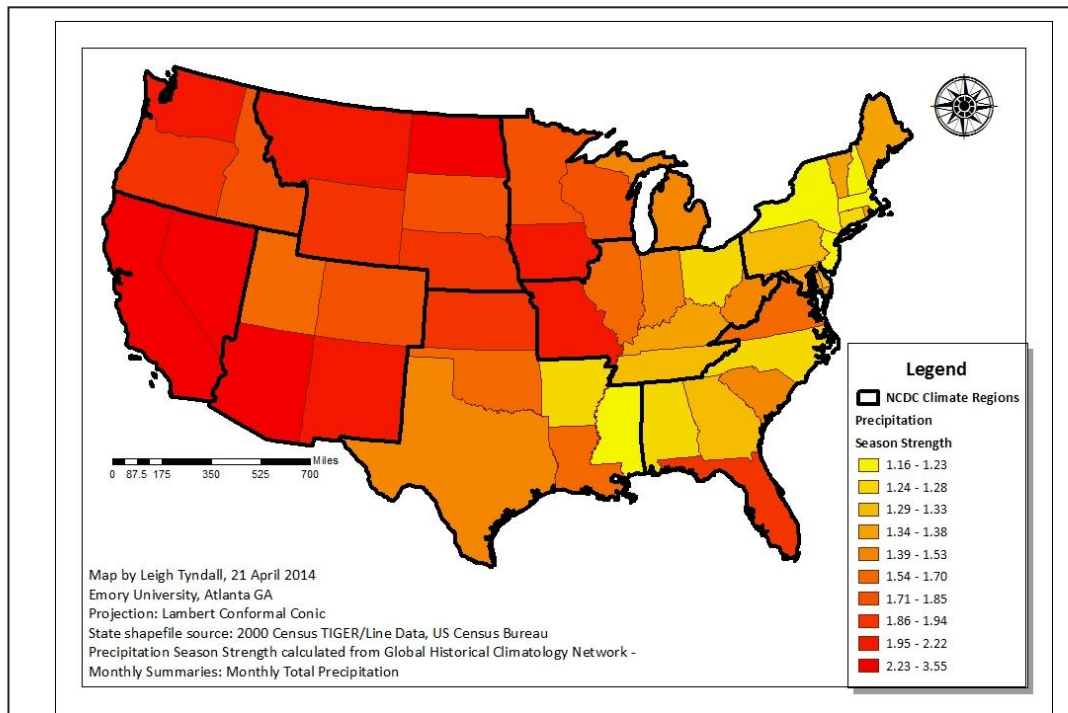


	<i>Percent Well Use</i>	<i>Percent Individuals Below FPL</i>	<i>Percent Urban</i>
<i>Percent Well Use</i>	1.0	-0.23201	-0.26677
		<0.0001	<0.0001
<i>Percent Individuals Below FPL</i>	-0.23201	1.0	
	<0.0001		
<i>Percent Urban</i>	-0.26677	-0.08640	1.0
	<0.0001	<0.0001	<0.0001

Appendix I. Supplementary Maps of Temperature and Precipitation Season Strength



Temperature Season Strength in the United States. Season strength is defined as the peak to mean ratio, which was calculated using the month with the highest average temperature over the average temperature per month. NCDC Climate regions are outlined in bold and labeled.



Precipitation Season Strength in the United States. Season strength is defined as the peak to mean ratio, which was calculated using the month with the highest average precipitation over the average precipitation per month. NCDC Climate regions are outlined in bold and labeled.