Quantifying the effect of extreme and seasonal floods on waterborne infectious disease in the United States

Victoria Devereux Lynch

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2022

Abstract

Quantifying the effect of extreme and seasonal floods on waterborne infectious disease in the United States

Victoria Devereux Lynch

The severity of flood events is predicted to increase as a consequence of climate change and may lead to a higher burden of waterborne infectious diseases in the United States. Contaminated floodwater transports bacterial, protozoal, and viral pathogens that typically cause moderate intestinal or respiratory disease, but can also lead to more serious disseminated infections among immunocompromised, young, and older people. Hydroclimatology and drinking water infrastructure influence the transmission of disease, but their roles are not well-understood and may vary by pathogen-type or geographic region. Specific outbreaks of waterborne disease have been attributed to major floods and cases have been positively associated with some meteorological variables, but the association between infections and flooding has not been systematically examined.

In this dissertation, we examine the association between seasonal and extreme floods and parasitic and bacterial infections using multiple flood-indicator variables and exposure definitions. In Chapter 2, we use multimodel inference and generalized linear mixed models to determine the effect of seasonal meteorology on hospitalizations across the US. We found that hospitalization rates were generally higher in rural areas and in places that relied on groundwater for drinking water sources. Soil moisture, precipitation, and runoff were associated with significant increases in hospitalizations for Legionnaires' disease, Cryptosporidiosis, and Campylobacteriosis, respectively. In Chapter 3, we use 23 years of weekly case data to examine

the effect of cyclonic storms on six waterborne infections in a conditional quasi-Poisson statistical model. Storm exposure was defined separately for distinct storm hazards, namely wind speed and cumulative rainfall, and effects were examined over 3 weeks post-storm. We found that exposure to storm-related rainfall was associated with immediate and lagged increases in cases. In Chapter 4, we use a nonparametric bootstrap to examine the effect of anomalous meteorological conditions, i.e. extremes unrelated to cyclonic storms, on Legionnaires' disease hospitalizations. We also assess the effect of exposure to specific cyclonic storms in a GLMM framework and compare these approaches. Extreme precipitation and months with cyclonic storms were positively associated with Legionnaires' disease hospitalizations. Determining the effect of flooding on Legionnaires' disease is particularly important as it causes severe illness and has steadily increased in incidence for 20 years.

An objective of this dissertation was to develop a framework for examining flood-disease dynamics in the context of hydrometeorological and infrastructure-related factors that may influence transmission. We demonstrated that drinking water source, rurality, and geography may play an important role in these dynamics; the analyses also underscored, however, the urgent need for more extensive epidemiological surveillance and water quality data. Climate change will likely place a considerable strain on aging water infrastructure in the US. A nuanced understanding of flood-disease dynamics is central to mitigating these effects.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike Information Criterion
CAFO	Concentrated animal feeding operation
CAP	Community-acquired pneumonia
CBSA	Core Based Statistical Area
CDC	Centers for Disease Control and Prevention
CSO	Combined sewer overflow
CWS	Community water system
GLMM	Generalized linear mixed model
GWR	Groundwater rule
HCUP	Healthcare Cost and Utilization Project
HSA	Hospital service area
EPA	Environmental Protection Agency
MCLG	Maximum contaminant level goal
MMI	Multimodel Inference
MSA	Metropolitan Statistical Area
NIS	National Inpatient Sample
NLDAS-2	NASA/NOAA North American Land Data Assimilation System 2
NNDSS	National Notifiable Disease Surveillance System
SDWA	Safe Drinking Water Act
SWTR	Surface water treatment rule
USGS	United States Geological Survey
WWTP	Wastewater treatment plant

Acknowledgments

I am so grateful to Jeff Shaman for his support and guidance over the last five years. Despite his incredible workload throughout the pandemic, when his own research was quite literally of global importance, he always made time for me. Thank you, Jeff, for taking me on as a student and for your mentorship. Thank you to my committee members Joan Casey, Tiffany Sanchez, Patrick Kachur, and Julie Herbstman for your thoughtful feedback and patience throughout this process. It has been a true pleasure getting to know you all and I am so grateful for your support.

Thank you to the Shaman lab group members for fostering such a warm and genuinely fun environment. In particular, thank you to Haruka Morita for keeping all of us organized and to Sasi Kandula for helping me get started with the HCUP dataset. Bernice Ramos-Perez, thank you so much for your tireless work to help us navigate grant submissions, financial issues, and departmental concerns, the doctoral program would not run without you.

The best part of my experience at Columbia has been getting to know and work with the remarkable students in EHS. Sebastian Rowland, Ahlam Abuawad, and Stephen Lewandowski, thank you for being wonderful cohort-mates; I am so excited to see what you brilliant people do next. I am deeply grateful to Jenni Shearston, Nicole Comfort, Vivian Do, Nina Flores, and Jocelyn Dicent for our work on the doctoral student sub-committee. I am inspired by your commitment to making our department a more equitable place. Finally, thank you to Misbath Daouda, Emma Gorin, and Alex Heaney for your friendship and encouragement.

I am fortunate to have the most wonderful friends who have provided invaluable support throughout this process. I could not have completed my dissertation research without the constant reassurance and love of Madeline Gilmore and Alison Footman; my roommate Maddie

was incredibly patient while I turned our living room in my office for the last two years and Alison has encouraged and inspired me since we started our MPH program together seven years ago. My doctoral research was rather solitary, a situation compounded by a pandemic that required physical isolation. Thankfully, I had the greatest quarantine pod support system that anyone could ask for: Annalee Leggett, Dan Oricchio, Saam Ayria, and Scott Reeves, I am so grateful for your friendship and ability to find joy – and humor – in any circumstance. To my best friend Charlotte Simms, thank you for your incredible belief in me and for helping me keep things in perspective.

Finally, I am profoundly grateful for the love and support of Robert, Robin, and Ann Lynch. My parents sacrificed a great deal for my education and are my biggest advocates. They have dedicated their lives to public service and have provided me with an amazing model for how to pursue meaningful work. Ann, thank you for being my lifelong partner; nothing makes me prouder than being your sister and our relationship is my greatest source of joy.

Dedication

I dedicate this dissertation to Susan Ann Keefe, an extraordinary scholar whose work was motivated by joyful curiosity and a profound love for others.

Chapter 1

1.1 Overview

An estimated 7,150,000 cases of waterborne infectious disease occur annually in the United States through contact with contaminated environmental, recreational, and drinking waters [1]. Floods are likely an important driver of disease transmission, but their effect has not been systematically examined across multiple pathogens and flood-indicator variables. A challenge to establishing flood-disease dynamics is that distinct flood types should be measured by different hydrometeorological variables; river floods, for example, are best characterized by soil moisture or runoff whereas precipitation is the appropriate metric for flash floods [2, 3]. The effects of flooding are also not uniform among pathogens but vary based on biology and predominant reservoir type, which influence their ability to enter and persist in the environment [4-7]. Further complicating flood-disease dynamics are the roles of hydroclimatology, land use, and drinking water infrastructure. Hydroclimatology determines where and when floods occur while land use influences the sources of contamination; flooding near agricultural land can result in different types of pathogenic contamination compared to coastal floods driven by cyclonic storms [8, 9]. Drinking water systems, meanwhile, affect the likelihood of transmission because effective treatment removes waterborne pathogens regardless of contamination events [10]. In practice, however, the quality of drinking water treatment varies by water source and location as rural areas are particularly vulnerable to inadequate water systems [11]. Quantifying the effects of flooding on multiple waterborne pathogens thus requires an analytical framework that accounts for the variability of flood types, contamination sources, and infrastructure.

1.2. Waterborne pathogens

Thirteen bacterial, protozoal, and viral pathogens are responsible for most waterborne infectious disease in the US and cause health conditions ranging from ear infections to life-threatening pneumonia or septicemia [1]. This work focuses on the 9 waterborne pathogens that lead to gastrointestinal or respiratory illness. While the pathogens are distinct, they share characteristics that demonstrate why waterborne disease is difficult to eradicate and how flooding may affect their transmission. Prime among these characteristics is the ability to exploit a wide range of animal and environmental reservoirs [12-14]. The enteric bacterial and parasitic pathogens infect wild and domestic animals including, crucially, livestock and birds [15, 16]; the biofilm-forming bacteria are natural inhabitants of soil and a diversity of environmental waters [17]. This array of hosts and reservoirs provides numerous opportunities for water contamination and infections. Transmission occurs after exposure to pathogenic drinking, environmental, or recreational water.

These waterborne diseases are also uniformly most severe among vulnerable groups [1]. Mild to moderate infections are common within the general population [1], but severe disease is typically limited to immunocompromised, young (under 5), or older people [18, 19]. The parasite *Cryptosporidium* emerged in the 1980s as an important infection in AIDS patients [14] and has manifest fatality rates ranging from 52% to 68% among immunocompromised people [18]. Cryptosporidiosis in young and malnourished children increases the risk of stunted growth and can lead to chronic infections [20]. Children are also more susceptible to Shiga-toxin producing *E. coli* (STEC) infections, which have a high risk of hospitalization and kidney failure [21]. Younger and older people may be particularly susceptible to enteric infections as they have less stable microbiomes that enable pathogenic bacteria to outcompete commensal species in the

gastrointestinal tract [22]. Biofilm-forming bacteria, meanwhile, are opportunistic pathogens that primarily cause severe illness in elderly hospitalized patients [17]. All of the pathogens are alike in that they require a very low infectious dose for infection [16, 23, 24]; as few as 30 oocysts are required for infection with *Cryptosporidium*, whereas an infected person excretes 10^9 in a single bowel movement [25].

Despite these shared traits, pathogen biology substantially varies among and within the enteric bacterial, parasitic, and biofilm-forming pathogen groups. Waterborne diseases encompass pathogens from several phyla and have distinct morphologies and infection mechanisms (Table 1.1). Some of these differences are related to pathogen ability to persist in the environment and may provide insight into how floods affect transmission. Variations in size may further explain differences in transmission dynamics because size is one of the factors that determines whether a pathogen remains suspended in water or deposits into sediment [23]. Several of the infections are caused by multiple species or serotypes within a particular pathogen's genus (Table 1.1), and this information can also indicate the source of the contamination as these different strains are typically specific to certain reservoirs [26, 27].

The proportion of infections attributable to waterborne transmission, rather than foodborne or person-to-person transmission, most clearly distinguishes pathogen groups. The biofilm-forming bacteria are almost exclusively waterborne whereas a small fraction of the enteric bacteria is thought to be transmitted via water (Table 1.2); these estimates mask, however, the role of contaminated water in foodborne transmission [28]. Agricultural runoff and irrigation water are often highly pathogenic and can contaminate crops [28, 29] leading to foodborne outbreaks that are at least partially, or indirectly, driven by water. The origin of most

foodborne cases is unknown and water may play a critical and under-recognized role in foodborne cases [16].

The clinical differences among pathogens demonstrate that the effects of flooding on specific infections could lead to considerable variation in the overall burden of disease and strain on the healthcare system (Table 1.2). Legionnaires' disease and *Pseudomonas* infections, in particular, almost always require hospitalization and have comparatively high mortality rates whereas people infected with *Giardia* often do not seek care [1]. Previous research on the associations between environmental drivers and cases has largely varied by pathogen, which suggests that the effects of flooding on waterborne infectious diseases are not likely to be uniform (Table 1.3). This underscores the need for pathogen-specific analyses to elucidate how the effects of flooding may change depending on location, drinking water or sanitation system, and severity of flood event.

Table 1. 1 Description of basic biological characteristics of waterborne pathogens

Pathogen	Type	Size (μm)	Infection mechanism	Strains
Campylobacter **Phylum: Protobacteria**	Intracellular, Gram-neg., obligate microaerophilic bacteria; motile, spiral [16]	0.5 - 5	Bacteria bypass mucus layer and enter extracellular matrix via M cells; invade intestinal epithelial cells through endocytosis; replicate in vacuole in host cells; cause inflammatory response [30]	Two strains responsible for most human illness, *C. jejuni* and *C. coli*, multiple serotypes [16]
E. coli **Phylum: Protobacteria**	Extracellular[a], Gram-neg., facultative anaerobic bacteria; motile and non-motile strains [23]	1 - 2	Unique mechanism for each strain but generally bacteria adhere to epithelial cells in small intestine or colon, secrete toxins, and cause inflammation and cell death; adhesion can include forming biofilms or creating pedestal-like lesions [21, 31]	Six strains that cause illness in humans, each with multiple serogroups or serotypes defined by O and H antigens [21]
Salmonella **Phylum: Protobacteria**	Intracellular, Gram-neg., facultative anaerobic bacteria [23]; motile, straight rods	2 - 5	Bacterial proteins alter cytoskeleton of intestinal epithelial to allow entry; use immune response to out-compete commensal bacteria; alter host cell membrane to create membrane-bound organelles that infect other cells [22]	>2,600 serovars that vary by O and H antigens; have different reservoirs, virulence, and geographic locations [15, 16, 32]
Shigella **Phylum: Protobacteria**	Intracellular, Gram-neg. bacteria; non-spore-forming, non-motile, rod-like [23]	0.4 – 0.6	Bacteria are internalized via endosomes by epithelial cells in colon; replicate in cytoplasm; release endotoxins to cause inflammatory response [23]	Four main strains, each with multiple serotypes [33]
Cryptosporidium **Phylum: Apicomplexan**	Intracellular, non-flagellated, parasite	4.2 – 5.4	Fully sporulated oocysts release four sporozoites in the GIT and enter epithelial cells; alter host cell membrane to create a parasitophorous vacuole that protects parasite from immune response;	*C. parvum* and *C. hominis* cause most disease in humans, 3 other species also common [34]

			undergoes asexual and sexual reproduction before releasing new oocyst[25]	
Giardia **Phylum: Metamonada**	Extracellular, anaerobic, flagellated parasite	10 – 14	Cysts release two trophozoites in the distal intestine; adhere to intestinal wall using ventral adhesive disks; obtain nutrients from host while undergoing binary fission; new trophozoites released and transform back into cysts [35]	*G. duodenalis* and *G. lamblia* infect humans, 7 genotypes within them [14]
Legionella **Phylum: Proteobacteria**	Intracellular, Gram-neg. bacteria; *L. pneumophilia* is aerobic, non-spore-forming, flagellated	2 – 20	Bacteria are phagocytosed by alveolar macrophages; alter host membrane to avoid phagolysom. fusion; create vacuole that supports bacterial replication [36]	*L. pneumophilia* causes Legionnaires' disease; 24 species can infect humans
Non-tuberculous *Mycobacteria* **Phylum: Actinobacteria**	Intracellular, Gram-pos., facultative, aerobic bacteria; non-motile, non-spore-forming, irregular rods	1 – 10	Bacteria are phagocytosed by alveolar macrophages; lipids on bacterial cell wall modulate phagosome maturation so avoid degradation; bacteria replicate in their protected vacuoles [37, 38]	>170 species, most human respiratory infections caused by *M. avium* and *M. abscessus*; two main phenotypic groups are slow-growing (SGM) and rapidly growing (RGM) [39]
Pseudomonas **Phylum: Proteobacteria**	Extracellular, Gram-neg., aerobic bacteria; motile, flagellated	1.5 – 3	Bacteria adhere to damaged epithelial cells using flagella and pili; secrete virulence factors that damage host lung and alter immune response; secrete extracellular matrix for biofilm formation; biofilm grows and protects bacteria from immune cells and antibiotics [40, 41]	>140 species, 25 of which infect humans; most *Pseudomonas* pneumonia cases caused by *P. aeruginosa* [42]

[a]One strain that causes illness in humans, EIEC, is intracellular [21].

1.2.1 Enteric bacteria

Campylobacter, Escherichia coli, Salmonella, and *Shigella* primarily cause gastrointestinal illness characterized by diarrhea, vomiting, nausea, and high fever [33]. *Shigella* is unique among them because humans are its only reservoir and person-to-person transmission is common, especially in crowded settings and vulnerable populations [43, 44]. The other pathogens, however, are well-suited for waterborne transmission as they have adaptable genomes that enable them to flourish in multiple reservoirs and persist in the natural environment [22, 31, 45]. *Salmonella* and *E. coli* strains, in particular, exchange genes that improve their ability to withstand harsh environments. Within the *E. coli* species, there is a small core genome and a much larger, flexible pangenome that includes genes that confer the ability to obtain nutrients from a range of sources and endure temperature fluctuations [31, 45, 46]. Similarly, the more than 2,600 *Salmonella* serotypes exchange plasmids that make them robust to environmental conditions, in addition to enhancing their virulence [22].

Table 1. 2 Description of clinical characteristics of waterborne pathogens

Pathogen	Est. annual cases [1]	Est. % hospitalized [1]	Est. % waterborne [1]	Incubation period (range) [16]	Serious sequalae
Campylobacter	1,540,000 (597,000-3,250,000)	19.5	13	2–3 days (1-10)	Guillen-Barre syndrome [16]
E. coli	283, 200 (93,000-681,000)	30.5	5	0.5–4 days (0.5 – 10)	Hemolytic uremic syndrome [31]
Salmonella	1,350,000 (733,000-2,450,000)	28.4	6	0.5–2 days	
Shigella	449,000 (97,800-1,350,000)	24.4	4	1–3 days (0.5 – 4)	
Cryptosporidium	823,000 (243,000-2,160,000)	19.2	43	7 days (2-12)	Chronic infections; stunted growth [20]
Giardia	1,070,000 (727,000-1,560,000)	7.9	44	7 days (1-14)	Recurring infections [6]
Legionella	11,400 (8,920-13,600)	98. 1	97	5-6 days (2-10)	
Non-tuberculous *Mycobacteria*	97,000 (75,700-122,000)	78.4	72	7-10 days	
Pseudomonas	31,700 (19,300-46,000)	97.2	51	1-3 days	Chronic infections in people with Cystic Fibrosis [47]

Salmonella can persist in river water for a month and has the highest survival rate in aquatic environments compared to the other bacteria [28, 48], which are more sensitive to external conditions [31, 49]. *E. coli* and *Campylobacter* are frequently detected in environmental water and can survive under favorable conditions [5, 49], but their concentrations are thought to be due more to constant input into environmental waters rather than long-term persistence [50]. Storms have been associated with rapid increases in *E. coli* concentrations and the small size of the bacteria makes it unlikely to sediment during high flow events [51]. All four pathogens persist the longest in agricultural slurry, the mixture composed of animal waste, soil, and water that is stored on farms [9, 50, 52]; this may be a contamination risk when floods occur on

agricultural land, particularly because the pathogens also persist in soil. Survival in soil is a complex function of temperature, nutrient availability, and the microbial community, and thus varies by location, but *Salmonella* typically survives for 1 to 3 months [9, 48] and *E. coli* for several weeks [53, 54].

Previous research on the association between meteorological variables and enteric bacteria suggest that the effect of flooding on disease may depend on the type of flood and pathogen. The effect of seasonal rainfall on cases is inconsistent [55-57], but increased cases and outbreaks have been associated with extreme rain (Table 3). Salmonellosis cases have risen after extreme rain events [32, 58] and the effect has been even stronger when the extreme events follow very wet or very dry periods [59]. *E. coli* cases have also been associated with extreme rain [60], including extreme rain with antecedent dry periods [61], as has Campylobacteriosis, though the effect may be restricted to certain geographic locations [62-65]. Examination of the effect of different flood types, across multiple regions, would help determine whether flooding has a similar, nonlinear effect on cases of enteric waterborne disease.

.

Table 1. 3 Average effect of meteorological variables on waterborne infectious diseases[1]

Pathogen	Seasonality	Rainfall	Temperature	Extreme wet conditions	River conditions
Campylobacter	Summer	Predominantly positive	Predominantly positive	Predominantly positive	-
E. coli	Summer	Mixed	Positive	Positive	-
Salmonella	Summer	Mixed	Positive	Positive	-
Shigella	None	-	Positive	Positive[2]	-
Cryptosporidium	Mixed	Predominantly positive	Positive	Positive	Mixed
Giardia	Late summer	Positive[3]	Positive	-	-
Legionella	Late summer/ early fall	Positive	Mixed		Positive
Non-tuberculous *Mycobacteria*	-	-	-	-	-
Pseudomonas	Weak summer[4]	-	-	-	-

[1]Associations were considered positive if 90% or more of the studies found a positive association and predominately positive if between 75% and 90% of the studies found a positive association.
[2]Based on a single outbreak [44].
[3]Based on a single study [66].
[4]Based on a single study [67].

1.2.2 Parasites

The symptoms of *Cryptosporidium* and *Giardia* infections are similar to the enteric bacteria, but the parasites are biologically distinct. Unlike the highly flexible bacteria, the parasitic pathogens have reduced and conserved genomes [6, 68], as well as fewer strains that are pathogenic in humans (Table 1.1). The differences between the primary Cryptosporidiosis species, however, are remarkably informative; *C. parvum* and *C. hominis* have separate main reservoirs, so the type of infection provides insight into the contamination sources and transmission routes. *C. parvum* infect a broad range of animals and cases suggest contact with animal waste, whereas *C. hominis* only infect humans and cases indicate exposure to human sewage [34, 69]. As a result, species-specific risk varies by location with a higher risk of *C. parvum* infections in rural areas and *C. hominis* in cities [20]. The simple parasitic genomes are

also ideally suited for waterborne transmission as they can persist in a range of environmental conditions.

Cryptosporidium oocysts and *Giardia* cysts do not independently replicate in the environment but can persist for months in water and soil [4]. Oocysts are small and remain suspended in flowing water for long periods, though even when they deposit out into sediment they remain viable [4]. They can also persist in water distribution systems, often as members of biofilm communities [70]. For both parasites, survival on surfaces is prolonged when organic matter is present and thus soil provides an excellent niche for long-term persistence [71]; *Cryptosporidium* can survive for up to 18 months in this environment [72] and *Giardia* for almost a year, though they typically cannot withstand freeze-thaw cycles during the winter [73]. Oocysts generally stay in the top 2-cm of soil, which suggests that flooding and overland flow easily remobilize them [74]. This is supported by research finding that planting vegetation between crops is associated with a reduction in oocysts in drinking and irrigation water because it thwarts the movement of contaminated floodwater across land [75].

Rain is central to the movement of cysts and oocysts in the environment [76], but its effect on cases of disease varies by region [77] and is not consistently positive [78]. The effect of hydrometeorological variables on cases is likely complex, as increased incidence has often been associated with rain [66, 79, 80] but occasionally with dry periods as well [81-83]. Similarly, Cryptosporidiosis seasonality exhibits several patterns and differs across regions [84-86]. Some countries report peaks in the spring and fall [87], whereas others just in spring or fall [84]. Insights derived from the strain-specific studies may help explain the various seasonal patterns. Cases of *C. parvum*, the species common in rural areas, are more prevalent in the spring and may be associated with the calving season [87]; a high proportion of calves are colonized with *C.*

parvum and runoff contaminated with their waste may introduce high loads of parasite into waterways [69]. *C. hominis* cases, conversely, are more prevalent in the fall [87] and may reflect transmission due to contact with sewage, potentially after floods [60], or foodborne exposure [88].

1.2.3 Biofilm-forming bacteria

Legionella, *Pseudomonas*, and Nontuberculous *mycobacterium* (NTM) are transmitted via the inhalation or aspiration of contaminated water and cause pneumonia, though they can also lead to disseminated infections [19, 89]. They are natural inhabitants of the environment and, unlike the other waterborne pathogens, replicate in water, soil, and water distribution systems [17]. *Legionella* and NTM parasitize amoebae [36, 90], which protect them from external conditions. *Pseudomonas* are metabolically versatile, allowing them to survive in harsh environments [91]. This durability underlies the ability to colonize almost all components of drinking water infrastructure [17]; pipe material and condition, temperature, and nutrients all influence growth, but the biofilms persist in most conditions [90]. These biofilm-forming bacteria are ubiquitous in drinking water [92], including public and private sources [93-95], but their presence is not necessarily a sign of contamination. They are abundant in most water sources and immunocompetent people are thought to be unaffected by exposure to low concentrations [96]. The burden of disease is unknown, however, as mild infections or cases of community-acquired pneumonia (CAP) are not currently detected by healthcare systems.

Biofilm-forming bacteria are opportunistic pathogens and infections identified by epidemiological surveillance are typically those that occur among immunocompromised or elderly people [41]. *Pseudomonas* and NTM infections, in particular, are associated with nosocomial transmission and reasonably studied as hospital-acquired infections (HAI), given that

Pseudomonas can be recovered from 50-60% of hospitalized patients [17] and is the second-most common cause of ventilator acquired pneumonia [41]. Most NTM and Legionnaires' disease cases, however, are community-acquired [92, 97] though it is difficult to estimate their true incidence; infections among immunocompetent people may go undetected because they do not seek care, in the event of mild illness, or are diagnosed with community-acquired pneumonia (CAP) that is not attributed to a specific pathogen [98]. CAP infections are only regularly identified for cystic fibrosis patients, for whom infections are severe and typically acquired in the community [41, 99]. Legionnaires' disease surveillance has improved over the last 30 years and is increasingly examined as a waterborne disease with environmental drivers, but the effect of meteorological variables on NTM and *Pseudomonas* infections has not been widely studied.

Legionnaires' disease has a consistent late summer or early fall seasonality in the US and has been associated with rainfall [100-102] and river height [103]. Associations between hydrometeorological variables and NTM or *Pseudomonas* infections have not been explored, but there is limited evidence that HAI from these pathogens do not exhibit seasonal patterns. Specifically, no seasonality of NTM cases in three states [99] or of *Pseudomonas* infections treated at naval hospitals in the US [104] was observed. An analysis of NTM in cystic fibrosis patients, however, found that cases peaked in the fall, which indicates that there may be seasonality to community-acquired cases [67].

1.3 Floods

Flooding occurs when weather events deliver more water to a drainage basin than can be absorbed or stored [105]; however, this basic process belies the complex mechanisms that determine where and when flood events arise [105, 106]. While most regions are vulnerable to flooding, seasonality, severity, and dominant flood types vary throughout the US [2, 107]. This

variability is a function of the large-scale climatic processes that determine seasonal meteorological conditions and weather events, as well as regionally-specific hydroclimatology, topology, and soil geomorphology [3, 105]. Flood severity, in particular, depends on the interaction between the type of weather event and surface characteristics; large thunderstorms, for example, can have no effect in rural areas with large drainage basins but lead to devastating flash floods in cities that lack permeable surfaces. Similarly, an unseasonably warm period can cause rapid snowmelt and heavy runoff in regions with extensive snow packs whereas areas without snow are unaffected. Both the hydroclimatology and land surface characteristics of a region regulate the effect of weather events and the types of floods that they can generate [3].

Hydroclimatology describes the movement of water between and within the atmosphere and land surface, synthesizing interactions among rainfall, water vapor, streamflow, and soil moisture [108, 109]. This framework helps explain why the effect of weather events depends on underlying meteorological and land surface conditions. In urban areas, convective storms are the primary driver of flooding so rainfall intensity or cumulative rainfall are typically the best flood indicators [110, 111]. Heavy rainfall does not always lead to flooding, however, because an unsaturated drainage basin may be able to absorb the influx; when soil is fully saturated, however, even light rain will generate surface runoff [105]. Premature snowmelt often leads to flooding because it occurs in the early spring when soil is at peak saturation after months of minimal evapotranspiration [105]. While rainfall is the critical environmental driver for some flood types, soil moisture and snowmelt regulate flooding for most of the US [3]. A mechanistic understanding of flood drivers enables identification of the variables, for example snowmelt in the Upper Midwest or duration of cyclonic storms along the East Coast, that best define flooding for different regions which is crucial because flooding cannot be measured by a single metric.

Hydroclimatology also determines the dominant flood type in a region and can help explain variability in the associations between flooding and specific waterborne infections, as different floods can lead to contamination from distinct sources. River floods typically occur in areas dominated by snowmelt or rain-on-snow hydroclimatology [106] and affect transmission by mobilizing pathogens that persist in floodplains or churning up sediment that harbors environmental pathogens [112]. River floods also create pools of standing water that may increase the likelihood of groundwater contamination [113, 114]; this is particularly relevant to the transmission of waterborne disease as groundwater is thought to be protected from contamination and thus is untreated [112]. Many areas of the US are vulnerable to flash floods but they are dominant in regions that experience thunderstorms and have small catchment areas [106]. While a single intense thunderstorm can cause flooding, flash floods usually occur when multiple thunderstorms sequentially affect the same area in a process called training thunderstorms [106]. Flash floods are more common in cities but arise in rural areas that have experienced drought, which reduces soil permeability, or have steep terrain with thin soil layers [106]. In urban areas, flash floods are often associated with sewage contamination from combined sewer overflow (CSO) systems and inundated waste treatment plants [113].

Coastal floods are primarily driven by storm surges associated with convective storms, though unusually high tides and large waves can also lead to flooding. Tropical cyclones are typically the most destructive, as associated storm surges can reach up to a mile inland [115]. These floods affect regions that encompass multiple land surface types, including rural and urban locations, and as a result can lead to contamination with agricultural and industrial runoff in addition to sewage [116, 117]. Lake floods are less common than other flood types and are generated by various weather events depending on the size, type, and location of the lake [106].

They are similar to coastal floods in that they are often due to storm surges or seiches, which are pressure difference-driven waves [106], and cause contamination from numerous sources [118, 119].

The dominant flood type in a region provides insight into the expected seasonality of flooding. Seasonality is often weak in places that experience multiple types of flooding, which can occur throughout the year, whereas areas with a single dominant flood type have a clear seasonality [2, 120]. For example, river floods generated by snowmelt or the combination of rain-on-snow generally typically manifest during the spring and exhibit a strong seasonal signal [2]. Summer or early fall flood seasonality is more common in areas that experience tropical cyclones, though this signal is less stable as cyclonic storms can make landfall between May and November along the East Coast [120]. Regions characterized by winter or summer seasonality are not always geographically isolated; in North Carolina, for example, peak flood season in the western mountainous part of the state is between December and April whereas in the central Piedmont and coastal areas flooding is during late summer or early fall [121]. Dominant flood types can vary over relatively short geographical distances and identifying the appropriate flood indicator variables is a challenge when examining the effects of flooding on waterborne disease.

The Eastern US experiences the greatest variety of flood types and includes few areas with pronounced flood seasonality. Appalachia exhibits the weakest seasonality as its topography facilitates flash floods during heavy rain and river floods associated with snowmelt, but it also covers a region that is vulnerable to cyclonic storms [2]. Seasonality is also weak along the Eastern seaboard and the Gulf Coast, where flooding occurs in conjunction with winter extratropical cyclones, particularly in the northeast, and summer convective storms [107, 120]. Among southeastern states five distinct hydroclimatological regions have been identified with

flood peaks occurring during winter, spring, mid-summer, fall or none at all [122]. Seasonality is more pronounced in areas where meteorological conditions enable co-occurring flood types; Florida has a summer flood season associated with the combination of tropical cyclones and summer thunderstorms, and in New England flooding peaks in early spring when cold extratropical cyclones and snowmelt lead to extensive river flooding [107, 120].

The Midwest is roughly divided between two dominant flood generating processes and flood seasons. The Northern Great Plains experience a consistent snowmelt and rain-on-snow driven spring flood season whereas in the south flooding is driven by mesoscale convective systems during the summer [123]. These storm systems bring intense rainfall to the entire Midwest, but do not necessarily lead to flooding in the north where land surface conditions allow for greater absorption. Areas of the West are similarly divided between two flood seasons; however, these distinct hydroclimatological regions are often geographically adjacent. In the Southwest, the North American Monsoon brings heavy rainfall to the region when warm, moist air from the Pacific is conveyed northward to Arizona and New Mexico during the summer [105]. This localized system leads to destructive flash floods due to the region's relatively impermeable soil and topography [124]. The mountainous region in central Arizona exhibit the opposite seasonality, however, as it is dominated by river floods associated with snowmelt [2]. Along the coast, floods are highly seasonal and peak during the winter in conjunction with cold extratropical cyclones and atmospheric rivers, which are bands of concentrated water vapor that carry moist air from the tropics to the midlatitudes [125]. The inland Cascade and Sierra Nevada mountain ranges experience snowmelt-driven flooding in addition to rainfall associated with storms and atmospheric rivers; this combination leads to longer, less pronounced flood seasons that can extend into spring or early summer [2]. Accounting for the diversity of flood generating

processes and types is necessary to analyze thoroughly the effects of flooding on waterborne disease in the US.

1.4 Contamination

Floods mobilize pathogens in the environment and inundate sanitation infrastructure, which can lead to the contamination of drinking water sources and transmission of waterborne disease [113, 126, 127]. Churning floodwater increases pathogen loads by resuspending pathogens in benthic sediment and dislodging biofilms in rivers or streams [128, 129]. Flooding also transports pathogens over land, especially when soil is already saturated, and deposits them into waterbodies or sewers. The influx into sewers can cause sewage overflows and reduce the efficacy of wastewater treatment, which results in heavily contaminated water entering the environment [129, 130]. This is of particular concern during snowmelt-driven flooding because there are fewer restrictions on raw sewage discharges during the winter [131]. In 1993, the largest waterborne disease outbreak recorded in the US occurred after a combination of snowmelt and rain led to river flooding; the contamination affected the primary drinking water source for Milwaukee, Wisconsin and over 400,000 people were infected with *Cryptosporidium* [132]. Cyclonic storms and extreme rainfall have also been associated with elevated cases of waterborne infections due to contamination from agricultural runoff, sewage, and damaged infrastructure [60, 133-135].

Most of the waterborne pathogens that cause enteric disease in humans also infect livestock and are frequently detected in soil or water samples on farms [13, 72, 136, 137]. *Campylobacter* and *Salmonella* strains common in pigs and poultry are prevalent in agricultural watersheds, which suggests that animal waste enters and is transported throughout the environment [137-139]. Newborn calves are often infected with parasites [140, 141] and the

seasonality of *Cryptosporidium* in both river samples and human cases has been attributed to the spring calving season [69, 81]. Large livestock operations, like concentrated animal feeding operations (CAFOs), typically use liquid waste management practices to store animal waste in pits [142]. Fecal matter is mixed with water in wastewater lagoons to create slurry, which is sprayed onto fields planted with cover crops [28, 143]. During extreme storms these lagoons have overflowed, ruptured, and become inundated with floodwater, leading to widespread contamination in fields and neighboring waterways [144]. In addition to transporting slurry and pathogens in soil to drinking water sources, runoff can also contaminate irrigation water that is used on crops [27, 145-147]. *E. coli*, *Salmonella*, and *Campylobacter* persist on produce and outbreaks have been attributed to contaminated irrigation water [136, 145], in some cases driven by extreme rainfall [148]. Most enteric bacterial infections are estimated to be foodborne, but flooding on farmland may play an important role contaminating crops [121].

In urban areas, inadequate sanitation infrastructure is more likely to cause contamination than agricultural runoff. Wastewater is highly pathogenic [149-152] and sediment in sewers act as a sink for pathogens while pipes often harbor biofilm communities containing *Legionella* or *Pseudomonas* [153, 154]. Under normal conditions, sewers deliver wastewater to treatment plants that provide at least secondary treatment before discharge to receiving waters (i.e. bays, lakes, or rivers) [155]. When they are overwhelmed with floodwater, however, untreated wastewater can spill or be discharged directly into the environment [156]. Old sanitation systems are especially vulnerable as they use combined sewer systems that collect sanitary wastewater and water from storm drains into a single pipe [157]. Above capacity, these systems experience combined sewer overflows (CSOs) and discharge untreated wastewater into nearby waterbodies [127]; CSOs are subject to regulations under the Clean Water Act but in practice the water often

exceeds permittable contamination levels [157]. Even without CSOs, excess inflow to treatment plants can result in reduced treatment efficacy or sewage bypasses whereupon wastewater only undergoes primary treatment before discharge [158]. This is a particularly hazardous process when flooding occurs after dry periods, which can result in low-flow conditions that increase the concentration of pathogens in waterbodies and in sewage sediment [153]. Discharging this extremely pathogenic water in a CSO or sewage bypass without treatment can lead to the contamination of drinking water [127, 134]. CSOs have been associated with elevated *Cryptosporidium*, *Giardia*, *Campylobacter*, and *E. coli* concentrations in recreational and drinking water sources [134, 151, 159], as well as with increased cases of gastroenteritis [160, 161].

Wastewater treatment plants (WWTPs) are typically built near coasts, riverbanks, or lakefronts to minimize the cost of conveying treated wastewater to receiving waters, but these locations make them vulnerable to inundation [162]. For coastal WWTPs, this vulnerability is compounded by sea level rise and the potential influx of marine water, which corrodes pipelines and pumps [163]. In the Eastern US, WWTPs are also susceptible to severe damage during cyclonic storms that can disrupt wastewater treatment; after Hurricane Harvey in 2017, for example, 40 WWTPs in Houston were inoperable for weeks [162]. This vulnerability reflects a fundamental challenge to sanitation infrastructure in the US. Redundancy is not built into the system so when a plant is damaged, there is no mechanism to back it up with another water treatment facility [162]. Further, much of wastewater treatment infrastructure is aging; old pipes are prone to infiltration from groundwater or floodwater, thereby increasing the likelihood of CSO events [156]. Even under normal conditions, wastewater treatment is not completely effective and low pathogen concentrations are often detected in lakes and rivers used for

municipal water [152, 164]; after CSOs, however, pathogen concentrations substantially increase [127].

Cases and outbreaks of waterborne disease have been attributed to direct contact with contaminated floodwater [126, 135, 165] and exposure via recreational activities [166]. Cryptosporidiosis, Giardiasis, and *Campylobacter* infections have been associated with exposure to urban floodwater, especially among children, and the clean-up periods immediately after floods have been identified as high-risk for transmission [126, 165]. Exposure associated with recreation poses an even greater risk and is estimated to cause millions of waterborne disease infections annually in the US [167, 168]. Transmission has been associated with swimming, fishing, and boating in lakes and rivers contaminated with pathogenic runoff [169-172], particularly after heavy rainfall [118, 171, 173, 174].

Most cases, however, are due to contaminated drinking water sources. Pathogens are frequently detected in waterbodies used for drinking water, but effective treatment generally removes them before distribution [10]. Flooding can compromise this process, though, by simultaneously increasing the pathogen concentration of source water and disrupting water treatment. Outbreaks of *Giardia*, *Cryptosporidium*, *Campylobacter*, and *Salmonella* have been attributed to drinking water sources in agricultural watersheds after spring runoff and summer storms [175-177]. Slurry and newborn calves have been specifically identified as contamination sources for drinking water-related transmission of *Cryptosporidium* and *Giardia* [24]. Cities are less susceptible to agricultural contamination but their extensive water distribution systems facilitate the growth and spread of biofilm-forming bacteria [178]. Indeed, Legionnaires' disease, NTM, and *Pseudomonas* infections have been associated with domestic water supply, as they

flourish in residential plumbing systems [93]; however, the effects of flooding or meteorological extremes have not been thoroughly examined.

1.5 Drinking water infrastructure

Drinking water infrastructure is highly localized in the US and varies by distribution system, source water, and water treatment method [179, 180]. Most people rely on drinking water from public water systems that control distribution from the entry point (the water treatment plant or the source water, if it is untreated) to the service connection, at which point the property owner becomes responsible for further distribution. The majority are community water systems (CWSs) that supply water to the same population year-round for domestic, agricultural, and commercial uses [181]. CWSs are subject to federal and state regulations, which are designed to ensure safe drinking water quality, though compliance and enforcement vary by system [182]. Private wells, however, are unregulated and individual owners are responsible for the maintenance of safe drinking water standards [183]. The exact number of private wells is unknown but an estimated 14% to 17% of the US population rely on drinking water from private sources, particularly in rural areas [184-186]. In addition to the effects of regulations and treatment methods, drinking water quality also depends on source water; private wells rely on groundwater whereas CWSs use both groundwater and surface water [183]. These factors influence the likelihood of contamination and may mediate the effects of flooding on waterborne disease transmission [112, 179, 187].

CWSs that use groundwater serve different communities and are vulnerable to distinct types of contamination compared to those that use surface water [187]. Groundwater systems are more likely to serve small, rural populations and were largely untreated until EPA enacted the Groundwater Rule in 2006 [188]. Groundwater is stored below the earth's surface, often in

aquifers, and for many years treatment was considered unnecessary [10]. Evidence that groundwater could become contaminated with fecal waste eventually led to the adoption of the Groundwater Rule (GWR), which mandates sanitary surveys to identify deficiencies in groundwater systems and treatment requirements if fecal contamination is identified during Total Coliform monitoring [188]. Compliant CWSs are required to undergo occasional monitoring whereas noncompliant systems must take corrective action to address the source of contamination or to provide enhanced water treatment. Unlike groundwater sources, the vulnerability of surface water to contamination has long been established and treatment standards were set in the 1974 Safe Drinking Water Act (SDWA), the first piece of federal drinking water regulation [10]. The Surface Water Treatment Rule (SWTR) is more stringent than the GWR as it requires filtration and disinfection, including residual disinfectants to ensure continuous treatment, for most drinking water systems [189]. The SWTR and GWR set maximum contaminant level goals (MCLGs) for *Giardia*, *Legionella*, *E. coli*, and viruses at zero because any exposure to these pathogens is considered unsafe [190]. In such cases where the goal is to completely eliminate contamination, the drinking water regulations are defined according to a treatment technique, e.g. a 3-log (99.9%) or 4-log (99.99%) removal, instead of a maximum containment level (MCL) [190].

While regulations have substantially improved drinking water quality since the 1970s, contamination is still prevalent throughout the US [179, 191]. Between 1982 and 2015, an estimated 9 to 45 million people relied on CWSs that violated water quality standards in a given year, an approximation that does not include contamination in private wells [11]. Treatment failures often occur in old or poorly maintained systems, even in the absence of storms or floods [175, 192]. Substandard water quality and disease outbreaks have been attributed to cracks in

water containment facilities, allowing for the infiltration of animal waste, and with disinfectant or filtration failures [193]. Flooding exacerbates these existing deficiencies and causes high-flow conditions, which can inhibit water treatment and increase turbidity in drinking water. Groundwater is particularly vulnerable to contamination because multi-barrier water treatment is not required under the GWR. The rule's objective was to enforce treatment requirements in CWSs at greatest risk for contamination while allowing those with high quality source water flexibility in selecting treatment methods [194]. In practice, however, the rule does not provide a framework for identifying these vulnerable CWSs, so systems with low quality water are not necessarily targeted and can operate using insufficient treatment techniques [11]. Even though groundwater is typically better protected from contamination than surface water, disinfection alone has still been found to be inadequate during high turbidity periods compared to the combination of disinfection and filtration [195].

Advanced water treatment is especially important for water sources vulnerable to contamination with *Cryptosporidium* oocysts and *Giardia* cysts [24, 196]. Due to their size, the parasites are able to penetrate conventional sand filters, and *Cryptosporidium* oocysts are resistant to most disinfectants; *Giardia*, however, is sensitive to common disinfectants and thus removed more readily during chemical water treatment [24]. The majority of *Cryptosporidium* outbreaks reported to CDC occur in systems that ostensibly treat drinking water, underscoring the need for advanced treatment, such as ozone or UV radiation, in more CWSs [187]. This is an especially serious problem in rural CWSs, which report more water quality violations than urban or suburban communities [11], likely due to compounding contamination risk factors that disproportionately affect rural areas. Rural CWSs are often in regions that experience river floods, which can generate standing floodwater that contaminates groundwater sources with

agricultural waste. They are also usually smaller and have fewer resources to implement or enforce SDWA regulations [197, 198].

The highest risk, though, is associated with consuming drinking water from private wells [199]. They are also more common in rural areas and are vulnerable to agricultural runoff, leaking septic tanks, and floodwaters [184, 200]. Private wells are not typically monitored for contamination. One survey in Georgia found that 41% of private wells were contaminated with fecal coliform and that 95% of people in rural Georgia relied on wells [200]. Consuming drinking water from private wells has been associated with increased risk for Salmonellosis, Campylobacteriosis, *E. coli* infection, and norovirus [172, 199, 201, 202]. Urban drinking water systems are better protected from contamination because they generally rely on surface water sources and undergo multi-barrier water treatment, per the SWTR [203]. They are susceptible, however, to the biofilm-forming pathogens that are natural inhabitants of water and flourish in extensive distribution systems [17]. The bacteria are ubiquitous in drinking water and often accumulate in domestic (property owner) systems, rather than CWSs, because premise plumbing has low residual disinfectant and intermittent stagnation that facilitate biofilm-formation [17, 92]. None of the waterborne pathogens are restricted to groundwater or surface water sources, but the overall associations among water sources, contamination routes, and distribution systems provides insight into how the effects of flooding on disease transmission vary by location.

1.6 Dissertation overview

The objective of this work is to quantify the effects of flooding on waterborne infectious diseases in the US and to determine whether they vary by location or drinking water source. The burden of waterborne disease is likely to increase in the future due to more severe seasonal and extreme floods, a problem compounded by aging water and sanitation infrastructure. Our ability

to mitigate the effects of flooding is contingent upon a thorough understanding of flood-disease dynamics. To that end, we seek to address two critical gaps in the flood-disease literature by 1) examination of the effects of flooding on specific pathogens, rather than syndromic conditions like acute gastroenteritis, and 2) use of multiple flood-indicator variables to account for the variety of flood-types common in the US. This is also the first study to examine the effects of tropical cyclones on specific waterborne diseases over multiple storm seasons. Flood-disease dynamics defy simple characterization because they are influenced by factors, namely hydrometeorological conditions and drinking water infrastructure, that vary across space.

In **Chapter 1** above, we propose a framework for analyzing these complex associations, which we apply to the following dissertation work. In **Chapter 2** we use the Healthcare Cost and Utilization Project (HCUP) dataset to examine the effects of seasonal hydrometeorological variables on hospitalizations for 13 different waterborne pathogens. Using Multimodel Inference (MMI) we assess the relative importance of flood-indicator variables and select the most highly-weighted factors to include in further statistical analysis. We then use a generalized linear mixed-model (GLMM) to determine the effects of seasonal hydrometeorology, location (rural/urban), and drinking water source on hospitalizations for the pathogens separately and grouped by pathogen-type (i.e. parasites, enteric bacteria, biofilm-forming bacteria). In **Chapter 3** we use weekly case data for six waterborne diseases from the National Notifiable Disease Surveillance System (NNDSS) to evaluate the effects of named cyclonic storms. We use a conditional quasi-Poisson model to compare cases in weeks with and without storms, and repeat the analysis using different storm exposure definitions based on cumulative rainfall, windspeed, and distance from the storm track. We also combine wind and rainfall exposure to create storm-type categories, e.g. “"high rain-high wind" or "low rain-low wind", to determine whether different types of storms

had a distinct effect waterborne disease. In **Chapter 4** we focus on hospitalizations for Legionnaires' disease, which is increasing in incidence in the US and may be an important, underestimated cause of community-acquired pneumonia (CAP) [98, 204, 205]. We first use the MMI and GLMM approach outlined in chapter 2 to study the effects of seasonal meteorology; next, we examine the effects of extreme floods using two approaches. We use a non-parametric approach to compare hospitalizations in months with extreme hydrometeorology to bootstrapped monthly averages. We also compare Legionnaires' disease hospitalizations in months with and without named tropical cyclones using a conditional logistic model. In **Chapter 5** we present further discussion and conclusions.

Chapter 2

Hydrometeorology and geography affect hospitalizations for waterborne infectious disease in the United States: A retrospective analysis

Victoria D. Lynch[1] and Jeffrey Shaman[1]

Affiliations:
[1]Department of Environmental Health Sciences, Columbia Mailman School of Public Health, Columbia University, New York, NY, United States of America
*Corresponding author: Victoria Lynch, vdl2103@cumc.columbia.edu, 722 W. 168th St. New York, NY 10032

Abstract

Meteorology, hydroclimatology, and drinking water infrastructure influence the transmission of waterborne infectious diseases in the United States, but their roles are not well-understood and may vary by pathogen type or geographic region. These pathogens cause severe intestinal, respiratory, or systemic infections in vulnerable people and pathogens that form biofilms may be an important driver of community-acquired pneumonia. Identifying the mechanisms that underlie contamination events and disease transmission is particularly important given that climate change may lead to more extreme floods, droughts, and seasonal precipitation. We examined the effect of meteorological variables, drinking water source, geographic region, and location (rural/urban) on hospitalizations for 12 waterborne bacterial, parasitic, and viral infections in the United States. Twelve years of hospitalization data from 516 hospitals in 25 states were used to assess seasonality and long-term trends in hospitalizations; we found that hospitalizations for bacterial and parasitic pathogen groups peaked between July and September and that Legionnaires' disease peaked between August and October. We used a multimodel inference approach to identify the most highly-weighted explanatory variables and included these in a generalized linear mixed model (GLMM) framework. There was a 16% (95% CI: 8% - 24%) decrease in hospitalizations for the bacterial pathogen group in urban compared to rural areas; for *Campylobacter*, specifically, there was a 31% (95% CI: 9% - 53%) decrease in urban areas, a 27% (95% CI: 6% - 48%) decrease associated with drinking water from surface water sources, and an 11% (95% CI: 4% - 17%) increase with a 1-standard deviation (SD) increase in runoff. Parasitic hospitalizations increased 9% (95% CI: 4% - 15%) with a 1-SD increase in precipitation, predominantly driven by *Cryptosporidium* hospitalizations, and were greater in areas that relied on groundwater rather than surface water as a drinking water source.

Legionnaires' disease increased 124% (95% CI: 90% - 157%) with a 1-SD increase in soil moisture. Associations between hospitalization rates and meteorological conditions, location, and drinking water source varied among the specific pathogens; the pathogen-group level analyses masked several of these findings and were largely uninformative. Precipitation, runoff, and rural locations were positively associated with some bacterial and parasitic infections; many of these pathogens regularly colonize livestock, and these findings suggest that agricultural areas may be particularly vulnerable to contamination events and disease transmission. Conversely, hospitalizations for biofilm-forming pathogens were associated with soil moisture and hospitalization rates were higher in urban areas. For these pathogens, prolonged wet conditions, and locations with extensive water distribution systems may drive transmission.

2.1 Introduction

Waterborne infectious disease is a persistent problem in the United States, where an estimated 7,150,000 cases occur annually despite safe drinking water regulations and sanitation infrastructure [1]. Waterborne pathogens transmitted via contaminated environmental or drinking water can cause severe respiratory, gastrointestinal, and systemic disease [179]. Drinking water and wastewater treatment substantially reduces the burden of disease but these systems are still vulnerable to contamination, a problem that will likely intensify in conjunction with aging infrastructure [11].

Waterborne pathogens include bacteria that naturally inhabit water (e.g. Nontuberculous mycobacteria, *Pseudomonas)* and thrive in biofilm-forming communities in pipes and water storage facilities [90, 206, 207]; *Legionella*, in particular, is associated with outbreaks linked to plumbing systems [208]. The biofilm-forming bacteria are opportunistic pathogens that when aspirated cause pneumonia among immunocompromised, elderly, or hospitalized people [41, 96, 206]. These pathogens are ubiquitous in drinking water [47, 94] and household plumbing [17], however, and may also cause a considerable proportion of community-acquired pneumonia [97].

Waterborne pathogens that cause gastrointestinal disease (e.g. *Cryptosporidium*, *E. coli*) are introduced into the environment through human or animal waste, and their seasonality suggests that meteorological factors influence the contamination events necessary for transmission [209, 210]. Intense precipitation, flooding, and drought affect the concentration and dispersal of these pathogens that, while not natural inhabitants of water, persist in the environment for months [49, 63, 211, 212]. Floods mobilize pathogens in sediment, soil, and water and overwhelm sanitation infrastructure so that sewage circumvents treatment [161]. Flooding after prolonged dry periods is of particular concern. Low-flow conditions during

droughts can increase pathogen concentration in water distribution systems; this pathogenic water is then flushed out with rapid flood-driven inflow [61, 82]. Most previous research has focused on non-specific gastrointestinal infectious disease and has found positive associations with flooding [213], precipitation [214], dry periods [82, 215], and temperature [55, 216], which affects the survival of some pathogens in the environment. Some pathogen- and location-specific studies have found more inconsistent associations, however, which indicate that the effect of meteorological variability is not uniform across regions or pathogens [77, 217].

The contamination events necessary for transmission are governed by dynamic interactions among hydroclimatology, land use, and water infrastructure. Meteorological conditions that lead to contamination in one setting may have no effect in regions with different hydroclimatology [3, 218, 219]; for example, Cryptosporidiosis has been found to increase with precipitation and temperature in tropical and temperate climates [76], but exhibits no seasonality or association with meteorological variables in arid regions [77]. Environment-disease dynamics can vary even within a small geographic area; precipitation has been positively associated with *Campylobacter* and *Salmonella* bacterial infections in low-lying coastal areas, but not inland regions, within a single state in the US [58, 59, 64].

Drinking water sources from both groundwater [112, 220] and surface water [187] are susceptible to contamination but meteorological drivers, exposure routes, and pathogens may vary by location [130, 179, 220], particularly between urban and rural areas [221]. In cities or places experiencing drought, precipitation on impermeable surfaces can lead to flash floods that cause sewage by-passes or combined sewer overflows (CSOs) [127]; this wastewater is highly pathogenic [131, 153] and can contaminate surface water sources used for drinking water [126, 222]. In agricultural regions with large drainage basins, however, exposure is often driven by

snowmelt [131], which generates standing water in fields and runoff polluted with animal waste [112]. Floodwater carrying pathogens from soil, including biofilm-forming bacteria, and from livestock fecal matter can contaminate drinking water from groundwater sources through direct contact or infiltration [112, 199, 223].

In this study, we examined the effect of meteorological variables on hospitalizations for waterborne infectious diseases and whether these associations were influenced by drinking water source, location (rural/urban), and region. We assessed these associations for bacterial, protozoal, viral, and biofilm-forming pathogen groups, and for each pathogen independently, to determine whether environment-disease dynamics were consistent among pathogens with similar biology. Previous research has examined the effect of precipitation or temperature on cases, but these studies have focused on nonspecific diarrheal illness, narrow geographic regions, or on outbreaks. Waterborne infectious diseases will become a more pressing public health challenge as climate change leads to more severe floods and droughts [224]. A thorough understanding of contamination mechanisms is necessary to identify communities at risk for waterborne illness and to develop effective interventions.

2.2 Methods

2.2.1 Data

Hospitalization data

The Centers for Disease Control and Prevention (CDC) has identified 12 waterborne pathogens that are endemic to the US and can cause severe illness [1]. In this analysis, we used the National Inpatient Sample (NIS) from the Healthcare Cost and Utilization Project (HCUP) to identify hospitalizations for the bacterial, protozoal, and viral pathogens that cause enteric or respiratory disease; we excluded the pathogens that predominantly cause ear and wound infections. We also

included hospitalizations for unspecified intestinal amoebic and protozoal infections which, while not included in CDC surveillance data, are waterborne pathogens known to cause severe disease [225]. We identified infections by ICD-9 code and found the monthly hospitalization count for each of the 12 pathogens at every hospital between 2000 and 2011. We restricted our analysis to hospitals that contributed at least 4 years of data to the NIS dataset, provided monthly counts of hospitalizations, and reported their exact geographic location. We excluded one hospital in Denver, Colorado, that specialized in respiratory infections, including those caused by the biofilm-forming pathogens included in this analysis, and treated patients from across the US. Since the NIS data are de-identified and in a publicly available dataset, Columbia University's Human Research Protection Office does not consider this to be research with human subjects and thus does not require an IRB review.

In the primary analysis, we also restricted hospitals to those that had at least 10 hospitalizations for waterborne infections during the study period. Hospitals with low case counts may be in areas without endemic waterborne pathogens, and the few cases that are identified and treated may be due to travel, foodborne outbreaks, or other sources that do not reflect local waterborne contamination. As a sensitivity analysis for this exclusion criteria, we repeated the analysis using several case count thresholds. We created subsets of our hospitalization data containing hospitals with at least 1, 5, 15, and 20 cases of waterborne disease during the study period; all of the analyses were repeated with these case threshold datasets.

Hospitals were categorized by location (rural/urban) and size (number of hospital beds) according to the definitions used by HCUP. Prior to 2004, urban hospitals were those within Metropolitan Statistical Areas (MSAs), as defined by the US Census Bureau based on 1990 Census data, and rural hospitals were those outside MSAs. From 2004 to 2011, Core Based

Statistical Areas (CBSAs) defined by the US Census Bureau were used to determine location; hospitals within 'Metropolitan' or 'Division' CBSAs were considered urban while those in 'Micropolitan' or 'Rural' CBSAs were rural. Hospital size was determined by the number of hospital beds given the hospital's geographic region, location, and teaching status (teaching hospital or non-teaching hospital).

Meteorological data

Precipitation, soil moisture, surface runoff, and temperature data were obtained from the NASA/ NOAA North American Land Data Assimilation System 2 (NLDAS-2) forcing dataset and were aggregated from hourly temporal resolution to mean monthly values for each hospital location.

Drinking water data

Drinking water data were extracted from the Safe Drinking Water Information System (SDWIS) for the community water system (CWS) that served each hospital. The correlation among drinking water source, primary water source, ownership of water system, and implementation status of source water protection measures was assessed to inform variable selection for the statistical model.

2.2.2 Trend and seasonality analysis by geographic region and pathogen group

We categorized the hospitalizations for waterborne illnesses into "bacterial", "parasitic", "biofilm-forming", and "viral" pathogen groups. The biofilm-forming pathogens are bacterial, but distinct from the other bacterial pathogens in that they are natural inhabitants of environmental water. We also assigned the hospitals to geographic regions according to United States Geological Survey (USGS) categories, with slight modifications to prevent single states from being the only representative in a region. The pathogen group and geographic region

variables were used to assess how seasonality and trends in hospitalizations varied throughout the US during the study period.

For each hospital, we calculated monthly hospitalizations per 10,000 discharges for each waterborne pathogen. We averaged monthly hospitalizations by pathogen group, geographic region, and month to determine if there was a distinct seasonality to hospitalizations. We also averaged monthly hospitalizations by only geographic region to examine the differences in trends between 2000 and 2011 across the US. We repeated the seasonality and trend analyses with the pathogen-specific hospitalizations to evaluate the consistency within the pathogen groups.

2.2.3 Statistical analysis

We modeled the association between waterborne disease hospitalizations and meteorological variables, drinking water source, and location using a negative binomial generalized linear mixed model (GLMM) framework to account for overdispersion in the hospitalization data. Drinking water source was included as a binary variable (groundwater or surface water), and location variables included terms for geographic region and hospital location (rural or urban). The models included a term for year and annual sine and cosine terms to adjust for secular and seasonal trends, respectively, and a random intercept for each hospital. Hospital- and year-specific total annual discharges were used as an offset to obtain the rate of hospitalizations; we present all results from the statistical analysis as percent changes in hospitalization rates. We modeled hospitalization rates for each pathogen separately and as pathogen-type groups.

Multimodel inference was used to compare models with all combinations of the standardized meteorological variables, drinking water source, geographic region, and hospital

location and to determine the importance weight of each explanatory variable. The candidate models varied only in these variables, but otherwise had the same structure. We used log likelihood and the number of parameters to calculate the Akaike weight for each model and ranked them by weight. The top models were the smallest number of models whose weights summed to 0.90 or greater, and the best-fitting model was the one with the highest weight.

Among the top models, variable weight importance for the meteorological, drinking water, and location variables was determined; these importance weights were used to calculate the weighted average effect estimates. Cross-validation was performed by removing 20% of the data and conducting multimodel inference on the remainder; this process was iterated 1,000 times to evaluate the consistency of the weights and effect estimates, and to compare these results to the top full models. These analyses were repeated for each case-count threshold.

2.2.4 Sensitivity analyses

The NIS includes the location of the reporting hospital, but not case residential locations. To address the possibility of misclassification bias, given that flood data are associated with the location of a hospital, we matched the hospitals to Hospital Service Areas (HSA) provided by the Dartmouth Atlas of Healthcare [226]. The HSA is the catchment area for each hospital and includes the zip codes where most Medicare patients receive care from a given hospital. We repeated the analyses using flood data associated with the HSA catchment area, instead of the hospital location, as a sensitivity analysis to assess the consistency of our findings.

2.3 Results

2.3.1 Waterborne disease hospitalizations

There were 57,335 hospitalizations for waterborne disease between 2000 and 2011 from 516 hospitals in the United States (Figure 2.1). The biofilm-forming pathogens comprised nearly

81% of all hospitalizations for waterborne illnesses, with 66% of hospitalizations due to respiratory *Pseudomonas* infections alone (Table 2.1). Apart from the *Pseudomonas*-related hospitalizations, the most common causes for hospitalization were Nontuberculous mycobacteria (NTM) infection (9.6%), *Salmonella* infections (8.0%), and Legionnaires' disease (4.1%) (Table 2.1).

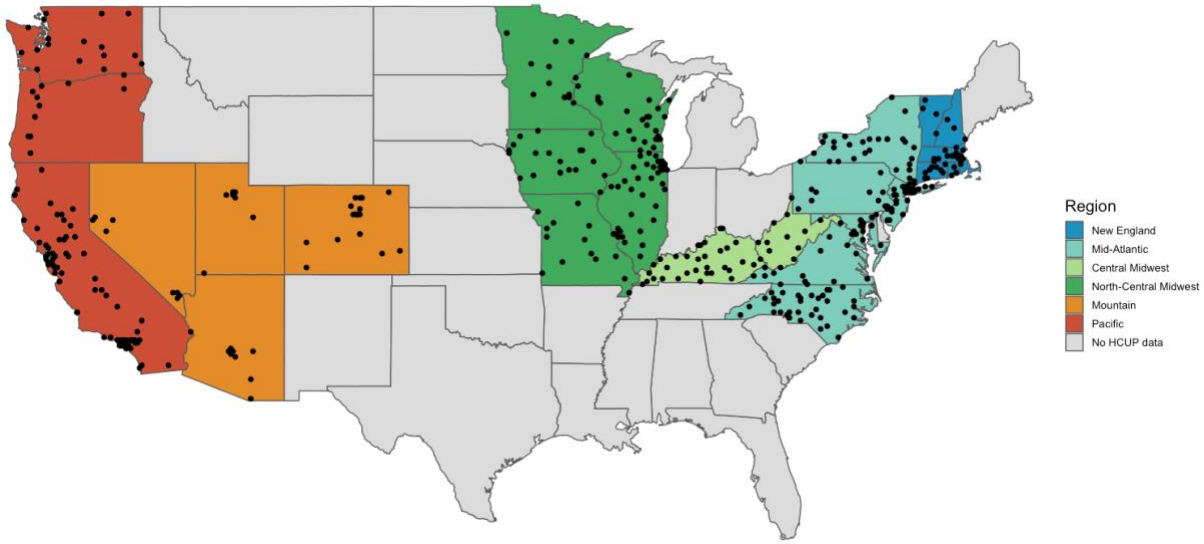

Figure 2. 1: Hospitals included in the analysis.
516 hospitals (black circles) in the HCUP dataset met the inclusion criteria and had a minimum of 10 total hospitalizations for at least one of the waterborne infectious diseases. The dark gray states are those that did not provide monthly geolocated data or did not report to HCUP.

Table 2. 1 Total count of hospitalizations for pathogen groups and specific pathogens between 2000 and 2011

Pathogen group	Specific pathogens	Number of cases	Percent of all hospitalizations
Bacteria		9, 259	16.2
	Salmonella	4,587	8.0
	Shigella	1,024	1.8
	E. coli	1,451	2.5
	Campylobacter	2,197	3.8
Parasite		1,580	2.8
	Giardia	654	1.1
	Cryptosporidium	661	1.2
	Protozoa (multiple species	79	0.1
	Amoeba (multiple species)	186	0.3
Biofilm-forming bacteria		46,221	80.6
	Legionella	2,327	4.1
	Respiratory *pseudomonas*	37,681	65.7
	Intestinal *pseudomonas*	717	1.3
Virus			0.5
	Norovirus	275	0.5

Hospitalizations for intestinal and biofilm-forming bacterial pathogens were significantly higher in areas that used groundwater as a drinking water source instead of surface water (Table 2.2); parasitic hospitalizations were slightly elevated as well, but the difference was insignificant ($p = 0.97$). However, the pathogen groups did not accurately reflect the pathogen-specific differences in hospitalizations by drinking water source (Table 1.S.1). Cryptosporidiosis hospitalizations were almost three times greater in groundwater areas compared to surface water while Giardiasis hospitalizations were slightly higher in latter (Table 1.S.1). Among the intestinal bacteria, Campylobacteriosis and *E. coli* hospitalizations were much higher in groundwater while Salmonellosis and Shigellosis were evenly split between drinking water categories (Table 1.S.1).

Hospitalizations for intestinal and biofilm-forming bacteria were also significantly higher in areas with privately owned CWSs, and in state-owned systems just for biofilm hospitalizations

(Table 2.S.2). Bacterial and parasitic hospitalizations did not substantially vary by primary water source (i.e. purchased groundwater, groundwater under influence of surface water, etc.) or by whether or not source water protection had been implemented (Table 2.S.2). Among biofilm-forming pathogens, however, hospitalizations were much higher in areas that relied on purchased groundwater and that had not implemented source water protection measures (Table 2.S.2).

Table 2. 2 Average monthly hospitalizations per 10,000 discharges by hospital location and type for the pathogen groups

Hospital characteristics[a]	Bacteria Mean ± SD	P	Biofilm-forming Mean ± SD	P	Parasite Mean ± SD	P	Virus Mean ± SD	P
Number of hospitals	302		516		89		13	
Hospital bed size								
Small	0.35 (1.02)		2.6 (14.81)		0.26 (1.02)		0.32 (0.58)	
Medium	0.29 (0.94)		0.89 (2.04)		0.17 (0.61)		0.77 (5.88)	
Large	0.23 (0.55)	<0.001	0.73 (1.45)	<0.001	0.1 (0.27)	<0.001	0.14 (0.64)	0.24
Hospital location								
Rural	0.48 (1.42)		1.33 (3.49)		0.35 (1.23)		1.06 (8.39)	
Urban	0.22 (0.50)	<0.001	1.13 (8.01)	<0.001	0.11 (0.38)	0.41	0.24 (0.89)	0.001
Region								
New England	0.21 (0.51)	0.26	0.70 (1.34)	<0.001	0.08 (0.17)	0.90	-	
Mid-Atlantic	0.29 (0.86)	0.37	1.16 (5.38)	<0.001	0.16 (0.61)	0.96	0.50 (4.66)	0.52
Central Midwest	0.25 (0.70)	-	1.41 (9.81)	-	0.12 (0.34)	-	0.21 (0.84)	-
North-Central Midwest	0.35 (1.02)	<0.001	2.6 (14.81)	<0.001	0.26 (1.02)	0.05	0.32 (0.58)	0.24
Mountain	0.29 (0.94)	<0.001	0.89 (2.04)	<0.001	0.17 (0.61)	0.001	0.77 (5.88)	0.24
Pacific	0.23 (0.55)	-	0.73 (1.45)	-	0.1 (0.27)	-	0.14 (0.64)	-
Water source								
Groundwater	0.30 (0.88)		1.39 (10.16)		0.17 (0.70)		0.55 (5.62)	
Surface water	0.24 (0.68)	0.13	1.06 (4.47)	<0.001	0.11 (0.31)	0.97	0.28 (0.99)	0.26

[a]Differences between or among hospital types were assessed for each pathogen group using Kruskal-Wallis test (for multiple groups) and Mann-Whitney U test (two groups) for non-parametric continuous data.

Hospitalizations for all of the pathogen groups were greater in small and rural hospitals, especially for the parasitic infections (Table 2.2). The pathogen-specific analysis demonstrated, however, that Legionnaires' disease hospitalizations were higher in urban areas unlike the other

pathogens in its group (Table 2.S.1). There were starker differences among group hospitalizations by geographic region; they were highest in the North-Central Midwest and Central Midwest regions for parasitic infections and in the Mountain and Pacific regions for biofilm-related infections. Hospitalizations for intestinal bacterial infections were relatively consistent across the geographic regions, though slightly higher in the North-Central Midwest (Table 2.2). An estimated 0.4% of all Norovirus cases lead to hospitalizations [1] and there were few in the dataset. Among the hospitals that reported cases, Norovirus hospitalizations were greater in medium-sized, rural hospitals and in the Pacific states (Table 2.2). These findings were not skewed by the specific HCUP hospitals included in the analysis; the number, size, and geographic breakdown of the hospitals was relatively consistent across pathogen group, though hospitals contributing to the parasitic pathogen group were disproportionately located in the North-Central Midwest and less likely to be located in the Pacific compared to the other geographic regions (Table 2.S.3). Most of the hospitals in the analysis were large facilities and located in urban areas with the exception of the hospitals restricted by *Pseudomonas* case thresholds; among these hospitals, which had at least 10 *Pseudomonas* infections, 30.4% were in rural areas and 56% were small- or medium-sized (Table 2.S.4).

2.3.2 Seasonality and time series trends in hospitalizations

The seasonality of waterborne disease hospitalizations varied by pathogen group and region in the United States (Figure 2.2). The bacterial pathogens exhibited the most consistent seasonality with hospitalizations peaking between July and September in all geographic regions (Figure 2.2a). During peak months, the average hospitalization rate for bacterial infections was greatest in the Central and North-Central Midwest compared to the other regions; this difference was not evident throughout the rest of the year, when hospitalizations for bacterial infections

were comparable among the geographic regions. These findings were consistent across the specific bacterial pathogens, though Campylobacter hospitalizations peaked earlier in the year in all geographic regions (Figure 2.S.1).

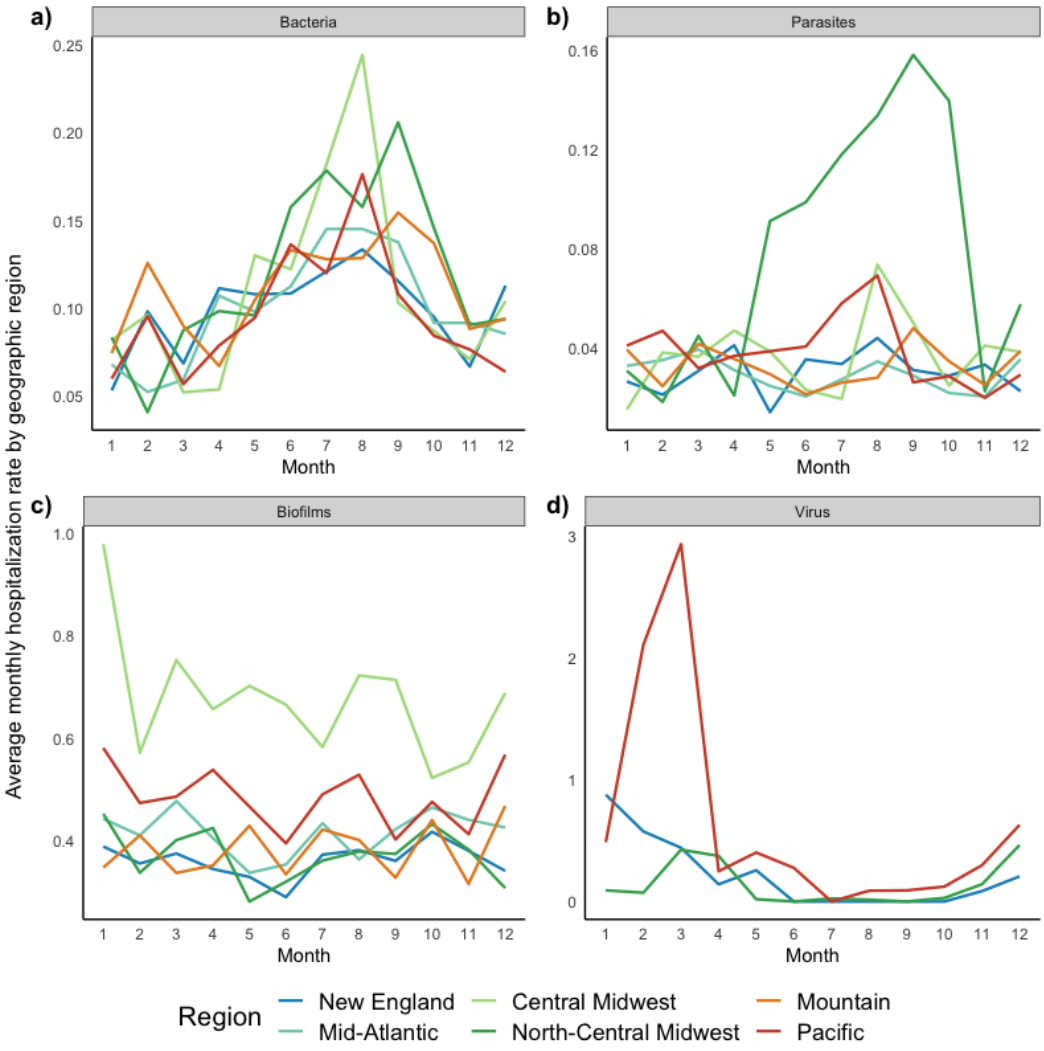

Figure 2. 2: Seasonality of hospitalizations by pathogen group and geographic region. Average monthly hospitalizations per 10,000 discharges for a) bacterial infections peaked between July and September for all regions; b) parasitic hospitalizations exhibited a seasonality similar to bacterial infections but only in the Midwest and Pacific regions. There was no clear seasonality to hospitalizations for c) biofilm-related infections and d) Norovirus hospitalizations peaked during winter months, though data were limited to only 13 hospitals in New England, North-Central Midwest, and Pacific states.

Seasonality for the parasitic pathogen-group was driven by hospitalizations for *Cryptosporidium* infections in the North-Central Midwest, which exhibited a sharp increase in May and then consistently increased until September (Figure 2.S.2); in the other geographic regions, peaks occurred throughout the year. There was no seasonality to hospitalizations for any of the other parasitic pathogens (Figure 2.S.2). Amoebic and protozoal hospitalizations were higher in the Mountain and Pacific regions throughout the year, but there were too few cases to assess seasonality.

There was no clear seasonality to hospitalizations for infections caused by biofilm-forming pathogens (Figure 2.2c), though this group-level analysis obscured the seasonality of specific pathogens. Hospitalizations for Legionnaires' disease peaked between August and October in all regions except the Pacific states, and for intestinal *Pseudomonas* infections in the late fall and winter (Figure 2.S.3). Finally, the only hospitals that met the 10-case threshold for Norovirus were in New England, the North-Central Midwest, and the Pacific, and hospitalizations peaked between January and March in all of those regions (Figure 2.2d).

Between 2000 and 2011 there was no significant change in monthly hospitalizations for any of the waterborne pathogen groups (Table 2.S.5). This was consistent across the specific pathogens except for Legionnaires' disease and Nontuberculous *mycobacterium*, which increased in New England, Mid-Atlantic, and the North-Central Midwest (Figure 2.S.4). In the latter half of the time series, biofilm-related hospitalizations increased in areas served by surface water and decreased in areas that used groundwater for drinking water (Figure 2.S.5); this trend was not evident among the bacterial or parasitic hospitalizations.

2.3.3 Statistical analysis

The most highly weighted meteorological variables identified by multimodel inference varied both among and within the pathogen groups, though drinking water source and hospital location were at least moderately weighted for most of the pathogens (Figure 2.3). The biofilm-forming group was the most consistent, with soil moisture and drinking water source highly weighted for the overall group and for all of the specific pathogens, other than intestinal *Pseudomonas* hospitalizations. Region was highly weighted only for respiratory *Pseudomonas* while hospital location (rural/urban) was moderately weighted for all of the other biofilm-forming pathogens. Multimodel inference for the bacterial pathogen group also moderately weighted drinking water source and soil moisture, though the latter was due to *Salmonella* hospitalizations (Figure 2.3). The pathogen-specific models were not well-aligned with the overall model; region was highly weighted only for *Salmonella*, while hospital location was highly or moderately weighted for *Campylobacter*, *E. coli*, and *Shigella*. Runoff was highly weighted in the *Campylobacter* model and precipitation was moderately weighted for *E. coli* and *Shigella* (Figure 2.3). Water source, hospital location, precipitation, and runoff were moderately weighted in the parasitic pathogen groups and more highly weighted for *Cryptosporidium* on its own (Figure 2.3). There were not enough amoeba and protozoal cases to assess the effects of region, hospital location, or drinking water source. Finally, in the Norovirus model none of the explanatory variables had high importance weight.

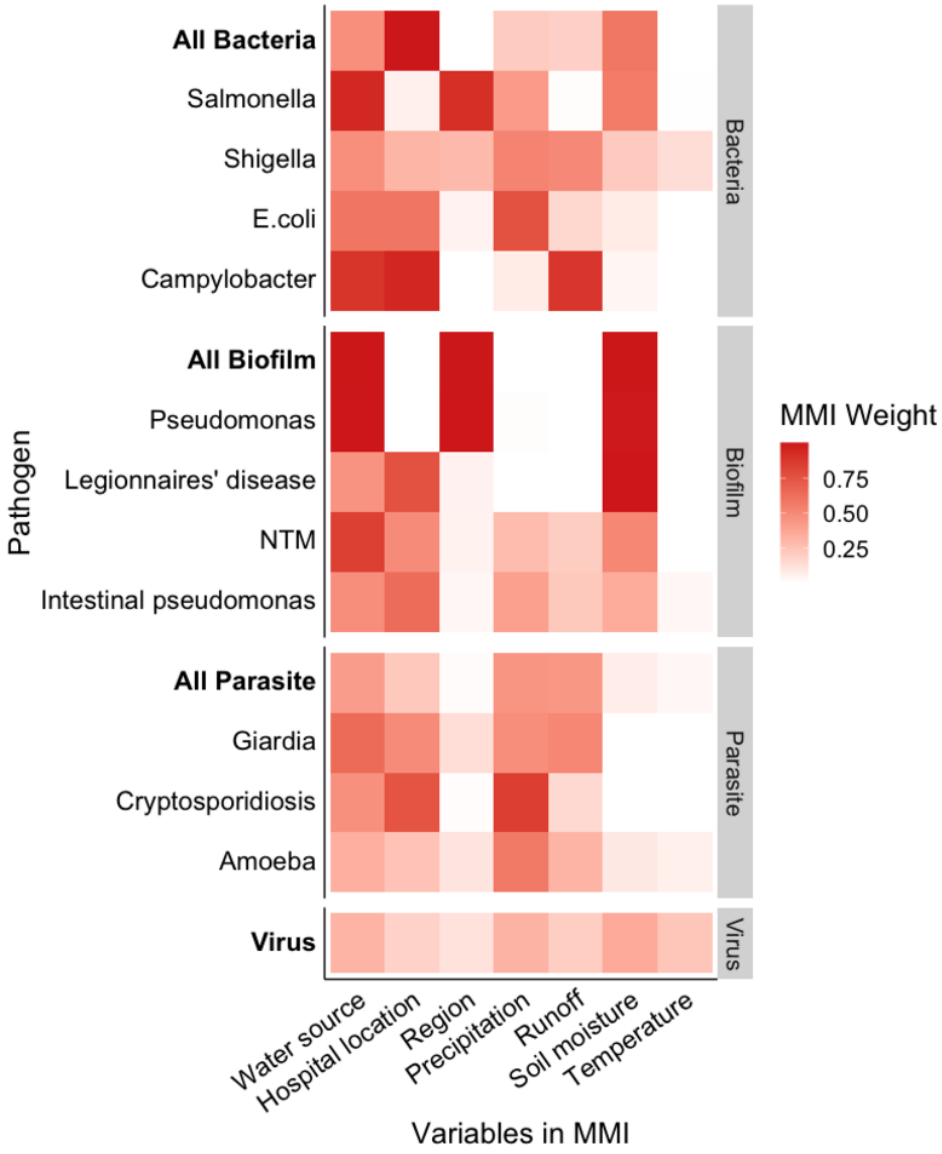

Figure 2. 3: Importance weights identified by multimodel inference.
Consistency between the group-level and pathogen-specific weights varied. Drinking water source and hospital location (rural/urban) were highly or moderately weighted in most of the models, but the importance weights for the meteorological variables were inconsistent between the group-level and pathogen-specific models. Runoff was highly weighted for *Campylobacter* while precipitation was moderately weighted for the other intestinal bacteria. Soil moisture was highly weighted for most of the biofilm-forming pathogens. Among the parasitic pathogens, only precipitation in the *Cryptosporidium* model was highly weighted.

There was a 16% (95% CI: 8% - 25%) decrease in hospitalization rates for the bacterial pathogen group in urban compared to rural locations (Figure 2.4), which was largely driven by a 31% (95% CI: 9% - 53%) decrease in *Campylobacter* hospitalizations in urban areas (Table 2.3). *Campylobacter* hospitalization rates also increased 11% (95% CI: 4% - 17%) in association with a 1-standard deviation (SD) increase in runoff and decreased 27% (95% CI: 6% - 48%) in areas that used drinking water from surface water instead of groundwater sources (Table 2.3). *E. coli* hospitalization rates increased in rural areas but decreased 14% (95% CI: -29 % - 1%) with a 1-SD increase in precipitation, though these effects were marginally significant.

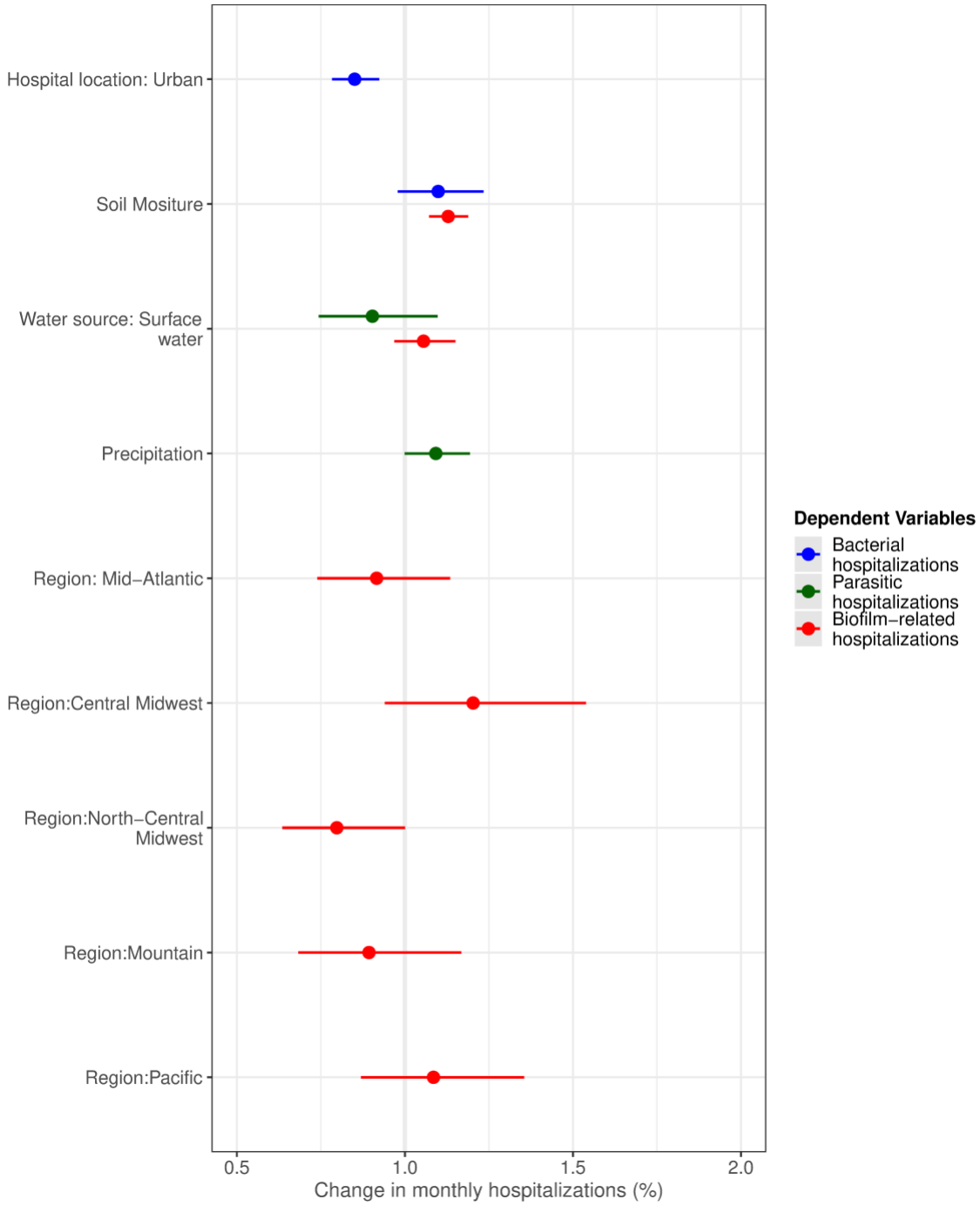

Figure 2. 4: Effect estimates from top model for each pathogen group.
There was a 16% decrease in the hospitalization rate for bacterial infections (blue) in urban compared to rural areas and soil moisture was included in the top model, but the positive association was marginally insignificant. Biofilm-related hospitalization rates (red) increased 12% with a 1-standard deviation increase in soil moisture and were greater in areas that used drinking water from surface water sources (Table 2.2), though this association was marginally insignificant in the model. A 1-SD increase in precipitation was associated with a 9% increase in hospitalization rates for parasitic infections (green).

Table 2. 3 Associations between hospitalization rates and meteorology, drinking water source, and location

Pathogen group	Precipitation	Soil moisture	Runoff	Temperature	Surface water	Urban
Bacteria	-	0.09 (-0.02, 0.21)		-	-	**-0.16 (-0.24, -0.08)**
Salmonella	-	0.09 (-0.09, 0.28)	-		0.01 (-0.12, 0.15)	-
Shigella	0.04 (-0.13, 0.21)	-	-		-0.14 (-0.65, 0.36)	-
E. coli	-0.14 (-0.29, 0.01)	-	-		-	-0.15 (-0.49, 0.18)
Campylobacter	-	-	**0.11 (0.04, 0.17)**		**-0.27 (-0.48, -0.06)**	**-0.31 (-0.53, -0.09)**
Biofilm	-	**0.12 (0.07, 0.17)**		-	0.05 (-0.03, 0.14)	-
Legionnaires' disease	-	**1.24 (0.9, 1.57)**	-		-	0.16 (-0.04, 0.37)
NTM	-	-0.09 (-0.25, 0.07)	-		-0.17 (-0.38, 0.03)	-
Respiratory *pseudomonas*	-	**0.09 (0.04, 0.15)**	-		0.08 (-0.01, 0.17)	-
Intestinal *pseudomonas*	0.16 (-0.11, 0.42)	-	-		-	**-0.62 (-1.21, -0.02)**
Parasite	**0.09 (0, 0.18)**	-		-	-0.1 (-0.3, 0.09)	-
Cryptosporidium	**0.22 (0.01, 0.44)**	-			-	-0.33 (-0.83, 0.17)
Giardia	-0.01 (-0.2, 0.19)	-	-		-	0.08 (-0.44, 0.59)
Protozoa	-	-	-0.47 (-1.92, 0.98)		-	-
Amoeba	0.24 (-0.02, 0.51)	-	-		-	-
Virus	-	-		-15.7 (-37.2, 5.71)	-	-

Hospitalization rates for biofilm-related infections increased 12% (95% CI: 7% - 17%) with a 1-SD increase in soil moisture (Figure 2.4), but the group-level findings obscured pathogen-specific associations. A 1-SD increase in soil moisture was associated with a 124% (95% CI: 90% - 157%) increase in Legionnaires' disease and a 9% (95% CI: 4% - 15%) increase in respiratory *Pseudomonas* hospitalizations (Table 2.3). Drinking water from surface water sources was also associated with an 8% (95% CI: -1% - 17%) increase in respiratory *Pseudomonas*, though the effect was marginally significant. Intestinal *Pseudomonas*

hospitalization rates, meanwhile, decreased 62% (95% CI: 2% - 121%) in urban areas, unlike the other biofilm-forming pathogens that were higher in urban locations, and were positively, though not significantly, associated with precipitation.

Cryptosporidium and amoebic hospitalization rates exhibited a similar relationship with precipitation, though location and drinking water variables were not included when modeling amoebas because data were too sparse. A 1-SD increase in precipitation was associated with a 22% (95% CI: 1% - 44%) increase in Cryptosporidium and a 24% (95% CI: -2% - 51%) increase in amoebic infections, though these effects was marginally insignificant (Table 2.3). Norovirus did not have a significant relationship with any of the meteorological variables in either the best model or the average of the top models. The importance weights, best model, and effect estimates were consistent across the hospitalization thresholds (Figure 2.S.6) and hospital service areas.

2.4 Discussion

Hospitalization rates for waterborne infectious diseases were associated with meteorological conditions, location, and drinking water source throughout the United States; however, the strength and direction of the relationships varied among pathogens. Rurality, runoff, and precipitation were associated with some bacterial and parasitic infections that are also common among livestock; conversely, soil moisture had an effect on hospitalization rates for biofilm-related infections and for Legionnaires' disease, hospitalizations were higher in urban areas. In general, the pathogen groups obfuscated important pathogen-specific associations and were ineffective at identifying trends.

Pathogen-specific water quality monitoring is onerous and expensive [227], and as a result infrequently conducted; these results suggest, however, that it may be necessary to

establish accurate associations between meteorological variables and waterborne disease. The need for pathogen-specific analyses was further underscored by the variability in seasonal peaks and general seasonal patterns, especially among pathogens within the same group. Seasonal variation can be a powerful tool for disentangling relationships between meteorological variables and infectious diseases because deviations from seasonal patterns can provide insight into the drivers of transmission [228, 229]. Given long time-series, even small changes in the seasonal environmental variables can indicate important, and potentially obscured, factors related to infectious disease dynamics. Extreme departures from seasonal norms, like rainfall during cyclonic storms, are also informative but their relative infrequency is limiting. Most waterborne diseases are considered highly seasonal, but we found considerable variability by geographic region and pathogen. Salmonellosis hospitalization rates peaked sharply in August in the Central Midwest but not in neighboring regions (North-Central Midwest, Mid-Atlantic) with similar meteorological seasonality. Conversely, in the North-Central Midwest *E. coli* hospitalizations peaked during the same time of year while Salmonellosis did not. More geographically and temporally resolved epidemiological data would allow a broader examination of why different regions exhibit distinct seasonality.

Bacterial and parasitic hospitalization rates were higher in small, rural hospitals and in Midwestern regions. Much of the Midwest experiences a wet spring season where the combination of snowmelt and intense precipitation can lead to flooding and heavy runoff [3]. Rural communities typically use drinking water from private wells, which are vulnerable to inundation during floods, or groundwater sources, which are often undertreated relative to surface drinking water [112, 187]. This is of particular concern in agricultural regions; both

increased pathogen concentrations in water and illnesses have been associated with wet conditions near farms [59, 81].

The ability to persist in the environment or evade water treatment measures varies by pathogen and may help explain why the effect of meteorological conditions is not uniform. Only *Campylobacter* hospitalizations were significantly associated with environmental or drinking water variables. *Campylobacter* can enter a dormant state in the environment, persisting for weeks in water or sewage, but do not replicate outside of animal hosts [5, 147]; this suggests that contaminated water may be as important a driver of transmission as foodborne exposure. *Campylobacter* hospitalizations were positively associated with runoff, drinking water from groundwater sources, and rurality, results that are consistent with previous research identifying associations with precipitation, rural coastal areas, and untreated well water [5, 64].

Among the parasitic pathogens, we found that *Cryptosporidium* hospitalizations increased with average monthly precipitation. *Giardia* is a cyst-forming parasite but unlike *Cryptosporidium* (which forms oocysts), hospitalizations were not associated with environmental variables and demonstrated no discernable seasonality. The difference between these pathogens underscores the roles of pathogen biology and water treatment in transmission dynamics. While both pathogens colonize livestock and have been positively associated with wet conditions [66], they differ in their persistence in the environment and response to water treatment [211]. *Giardia* has been associated with high flowrates, indicating that runoff and flood conditions dilute its concentration and flush it out of the environment [130, 230]; *Giardia* is also easily removed from water, so treatment is highly effective [130]. *Cryptosporidium*, however, persists in water, potentially as part of biofilm communities, and is highly resistant to chlorination [211].

Waterborne diseases have also been associated with drought conditions when pathogens are concentrated in diminished waterbodies; we found some evidence for this in hospitalizations for *E. coli*, which increased in months with lower precipitation, though this association was marginally significant. *E. coli* cases have been found to increase during dry periods and, in particular, during intense precipitation with antecedent dry periods [82].

Biofilm-forming bacteria may be an important source of community-acquired pneumonia (CAP) but their transmission dynamics outside hospital environments have not been thoroughly examined. We found hospitalizations for biofilm-related infections were positively associated with soil moisture, which integrates rainfall and snowmelt and reflects more extreme hydrological conditions including floods and droughts [231, 232]. Prolonged wet periods and overland flow likely mobilize these pathogens that naturally inhabit soil. The group-level association was driven by Legionnaires' disease and respiratory *Pseudomonas*-related hospitalizations, though the importance of environmental drivers on transmission differed between them. The effect of soil moisture on Legionnaires' disease was 10-times stronger compared to the group and while Legionnaires' hospitalizations demonstrated consistent seasonality across geographic regions, there was no seasonality to *Pseudomonas* hospitalizations. This suggests that respiratory *Pseudomonas* is less tightly coupled to environmental variability, though this finding may be due to the inability to distinguish community-acquired and nosocomial infections.

Legionnaires' disease hospitalization rates were higher in urban areas and in places that used drinking water from surface water sources; these associations were not statistically significant in the model framework but provide important guidance for future research with more temporally or geographically resolved data. Cities have complex distribution systems and a large

number of premise plumbing systems that provide locations (e.g. pipes, holding tanks) for biofilm formation [17]. Rural drinking water sources are still vulnerable to contamination but non-centralized systems, and private wells in particular, offer fewer opportunities for biofilms to form or grow. Intestinal *Pseudomonas* hospitalizations, however, were substantially higher in rural areas and associated with precipitation at marginally significant levels; these associations closely mirror those of *Cryptosporidium*, and suggest that both infections share similar transmission mechanisms. The similarity between intestinal *Pseudomonas* and *Cryptosporidium*, in addition to the overall inconsistency between pathogen group-level and pathogen-specific findings, demonstrate the complexity of factors that influence waterborne transmission and indicate that they may not be adequately captured by broad categorization.

Our findings are constrained by several limitations. The monthly resolution of the hospitalization data prevented examination of the effect of rapid changes in meteorological conditions, which may increase contamination by concentrating and then flushing pathogens [61, 82]. Data geolocation also introduces the potential for misclassification bias, given that meteorological data were associated with hospital locations, which may not reflect conditions at patients' work and home. We aimed to address these limitations by repeating the study using hospital catchment areas as a sensitivity analysis, which was consistent with the primary findings. The analysis also does not include data from the Southeast because these states did not report monthly data to HCUP; this is a major limitation as many Southeastern states include agricultural regions and experience substantial flooding associated with cyclonic storms.

The severity of floods and droughts are likely to change in conjunction with atmospheric warming; identifying the effect of environmental factors on waterborne infectious diseases is necessary to prepare for these events. Future research should aim to develop a comprehensive

mechanistic model of contamination events by incorporating water quality data from environmental and drinking water sources. Detailed microbiological data would enable an exploration of the interactions of waterborne pathogens in water with multiple contaminants. In lab studies, biofilm formation was enhanced in water with biofilm-forming bacteria (*Legionella* and nontuberculous *Mycobacterium*) and amoebas [225]; there were too few amoebic hospitalizations in this dataset to assess their relationship with biofilm-related infections, but in the future associations between microbiological contamination and infections should be examined. As most cases of waterborne disease are not hospitalized, future work should also expand to include all reportable cases; this is particularly important for understanding the burden of community-acquired pneumonia due to biofilm-forming pathogens. Some previous studies have found associations between waterborne diseases and extreme climatic events, including floods and droughts. These potential nonlinear effects are not captured in this analysis, and future work should examine cases and outbreaks due to extreme events.

2.5 Supplementary Materials

Table 2. S. 1 Average hospitalizations per 10,000 annual discharges by hospital location and type

Pathogen	Salm.	Shigella	E. coli	Campy.	Crypto.	Giardia	Protozoa	Amoeba	Legionnaires' disease	NTM	Pseudomonas	Int. pseudo	Norovirus
No. cases	4,587	1,024	1,451	2,197	661	654	79	186	2,327	5,496	37,681	717	275
No. hospitals	173	26	40	71	20	27	2	11	75	147	496	9	13
Hospital Location(%)													
Rural	0.306	0.069	0.125	0.144	0.338	0.092	0	0.003	0.053	0.075	1.716	0.049	1.058
Urban	0.153	0.038	0.055	0.085	0.068	0.062	0.007	0.015	0.07	0.511	1.495	0.029	0.24
Hospital Bedsize (%)													
Small	0.184	0.055	0.088	0.153	0.215	0.111	0.004	0.006	0.076	1.105	2.583	0.055	0.323
Medium	0.19	0.053	0.067	0.102	0.102	0.08	0.005	0.014	0.057	0.114	1.288	0.027	0.766
Large	0.178	0.037	0.064	0.079	0.059	0.05	0.007	0.015	0.062	0.135	1.176	0.031	0.139
Region (%)													
New England	0.174	0.023	0.062	0.124	0.047	0.057	0.006	0.006	0.097	0.144	1.188	0.031	0.216
Mid-Atlantic	0.22	0.033	0.036	0.078	0.051	0.056	0.004	0.011	0.095	0.143	1.547	0.021	-
Central Midwest	0.184	0.048	0.105	0.079	0.074	0.086	0	0.006	0.062	0.06	2.269	0.085	-
North-Central Midwest	0.159	0.04	0.112	0.132	0.203	0.07	0.003	0.007	0.049	0.089	1.184	0.035	0.138
Mountain	0.173	0.071	0.085	0.072	0.031	0.059	0.016	0.023	0.049	3.03	1.156	0.032	-
Pacific	0.153	0.061	0.061	0.097	0.041	0.074	0.011	0.033	0.031	0.113	1.834	0.032	0.7
Water Source (%)													
Groundwater	0.186	0.05	0.086	0.114	0.146	0.055	0.005	0.012	0.055	0.756	1.393	0.039	0.553
Surface water	0.181	0.045	0.059	0.087	0.052	0.068	0.007	0.015	0.071	0.127	1.675	0.033	0.281

Table 2. S. 2 Average monthly hospitalizations per 10,000 annual discharges by drinking water variables for the pathogen groups

CWS characteristics[a]	**Bacteria**	**Biofilm-forming pathogens**	**Parasite**	**Virus**
Water Source				
Groundwater	0.299*	0.977*	0.168	0.553
Surface water	0.244	1.066	0.110	0.281
Ownership of Water System				
Federal	0.207*	0.582*	0.129	-
Local	0.264	0.961	0.135	0.423
Private/public	0.161	0.393	0.044	-
Native American	0.273	0.673	-	-
Private	0.321	1.831	0.086	0.225
State	0.202	1.536	0.035	-
Primary Water Source				
Groundwater infl. by surface water	0.392	1.001*	0.179	-
Purchased groundwater infl. by surface water	0.248	0.702*	-	-
Groundwater	0.302	0.945	0.168	0.609
Purchased groundwater	0.158	1.842	-	0.233
Surface water	0.240	0.986	0.101	0.411
Purchased surface water	0.237	1.182	0.118	0.105
Water Source Protection				
Water source protection not implemented	0.254	1.017*	0.115	0.205
Water source protection implemented	0.211	0.696	0.076	-
Not reported	0.289	1.171	0.158	0.503

[a]Differences between or among hospital types were assessed for each pathogen group using Kruskal-Wallis test (for multiple groups) and Mann-Whitney U test (two groups) for non-parametric continuous data.

*Indicates significant differences between or among CWS characteristics ($p < 0.05$).

Table 2. S. 3 Description of the hospitals included in the analysis by pathogen group using HCUP variables and drinking water source data

Hospital characteristics[a]	**Bacteria**	**Parasites**	**Biofilm-forming bacteria**	**Virus**	**Overall**
No. of hospitals	302	89	516	13	524
Hosp. Location (%)					
Rural	19.2	6.2	31.7	20	24.9
Urban	80.8	93.8	68.3	80	75.1
Hospital Bedsize (%)					
Small	13.9	8.8	25.2	1.7	19.5
Medium	30.2	26.9	30.6	41.7	30.2
Large	55.9	64.3	44.2	56.7	50.3
Region (%)					
New England	10	10.2	7.9	28.3	9.1
Mid-Atlantic	32.6	34.3	30.1	0	30.9
Central Midwest	7.5	5.5	12.6	0	10
North-Central Midwest	19	23.8	19.4	28.3	19.8
Mountain	10.6	13.1	8.9	0	9.8
Pacific	20.3	13.1	21.1	43.3	20.3
Water Source (%)					
Groundwater	33.2	38.2	34.4	46.2	36.3
Surface water	66.8	61.8	65.6	53.8	63.7
Annual Discharge					
Mean (SD)	16,100 (11,100)	24,100 (13,500)	11,300 (10,600)	17,800 (10,700)	14,200 (11,800)
Median (Min, Max)	14,000 (1,040, 65,800)	23,100 (1,500, 65,800)	7,670 (76, 65,800)	17,100 (503, 42,600)	10,900 (76, 65,800)

[a]Differences between or among hospital types were assessed for each pathogen group using Kruskal-Wallis test (for multiple groups) and Mann-Whitney U test (two groups) for non-parametric continuous data. There were significant differences by hospital characteristics ($p < 0.05$) for all of the categories except water source. Differences among the pathogen groups, however, were mostly insignificant.

Table 2. S. 4 Description of the hospitals by specific pathogen using HCUP variables and drinking water source data

Pathogen	Salm.	Shigella	E. coli	Campy.	Crypto.	Giardia	Protozoa	Amoeba	Legionnaires' disease	NTM	Pseudomonas	Int. pseudo	Norovirus
No. cases	4,587	1,024	1,451	2,197	661	654	79	186	2,327	5,496	37,681	717	275
No. hospitals	173	26	40	71	20	27	2	11	75	147	496	9	13
Hospital Location (%)													
Rural	13.6	5	14.4	4.9	9.8	0	0	2.9	6.4	5.6	30.4	14.6	20
Urban	86.4	95	85.6	85.1	90.2	100	100	97.1	93.6	94.4	69.6	85.4	80
Hospital Bedsize (%)													
Small	9.1	9.3	6.7	6.1	10.8	4.4	0	0	9.8	9.3	24.5	0	1.7
Medium	23.2	30	17.4	25	15.7	19	0	18.8	24.9	26.5	31	39	41.7
Large	67.6	60.7	75.9	68.9	73.5	76.6	100	81.2	65.4	64.1	44.6	61	56.7
Region (%)													
New England	9	9.3	17.4	12.8	5.9	10.2	64.3	0	14	10.7	8.2	14.6	28.3
Mid-Atlantic	41.3	28.6	24.6	29.4	37.3	40.9	0	31.9	52.2	39	30.3	31.7	0
Central Midwest	6	3.6	6.7	1.2	4.9	6.6	0	0	5.9	4.1	12.7	9.8	0
North-Central Midwest	14.8	9.3	23.1	19.2	38.2	19.7	0	7.2	10.9	11.4	19.4	12.2	28.3
Mountain	9.6	22.9	16.9	12.5	4.9	13.9	35.7	11.6	12.3	14.1	8.1	22	0
Pacific	19.2	26.4	11.3	25	8.8	8.8	0	49.3	4.7	20.7	21.3	9.8	43.3
Water Source (%)													
Groundwater	32	30.8	27.5	29.6	47.6	25	0	100	35.5	27.2	37.5	22.2	46.2
Surface water	68	69.2	72.5	70.4	52.4	75	100	0	64.5	72.8	62.5	77.8	53.8

Primary Water Source													
GW	30.9	30.8	27.5	29.6	47.6	25		100	35.5	27.2	35.9	22.2	38.5
GU	1.7	0	2.5	2.8	4.8	0		0	0	1.3	1.8	11.1	0
GWP	36.6	34.6	42.5	36.6	33.3	50		0	36.8	39.1	33.9	55.6	30.8
SW	29.1	34.6	27.5	31	14.3		0	27.6	32.5	26.3	0	23.1	
SWP	0.6	0	0	0	0		0	0	0	0.4	11.1	0	<0.001
Annual Discharge													
Mean (SD)	20,500 (11,800)	25,100 (15,000)	24,700 (15,200)	23,200 (13,000)	28,100 (18,000)	37,400 (4,950)	25,700 (12,800)	23,300 (13,100)	21,600 (11,600)	11,600 (10,700)	18,900 (15,300)	17,800 (10,700)	
Median (Min, Max)	19,200 (2,130, 65,800)	24,600 (1,570, 64,700)	22,100 (4,320, 65,800)	21,700 (2,170, 64,700)	25,000 (3,610, 65,800)	36,300 (28,600, 43,500)	23,900 (6,170, 65,800)	20,800 (1,760, 65,800)	20,100 (76, 65,800)	8,080 (128, 65,800)	15,500 (4,500, 65,800)	17,100 (503, 42,600)	

Table 2. S. 5 Assessment of time series trends by pathogen group using Mann-Kendall test

Pathogen group	Mann-Kendall p-value	Mann-Kendall slope	Mann-Kendall indicator
Bacteria	0.338	7.01E-06	0.429
Biofilm	0.613	6.14E-04	0.744
Parasite	0.685	6.76E-06	1
Virus	<0.001	8.89E-04	2.34

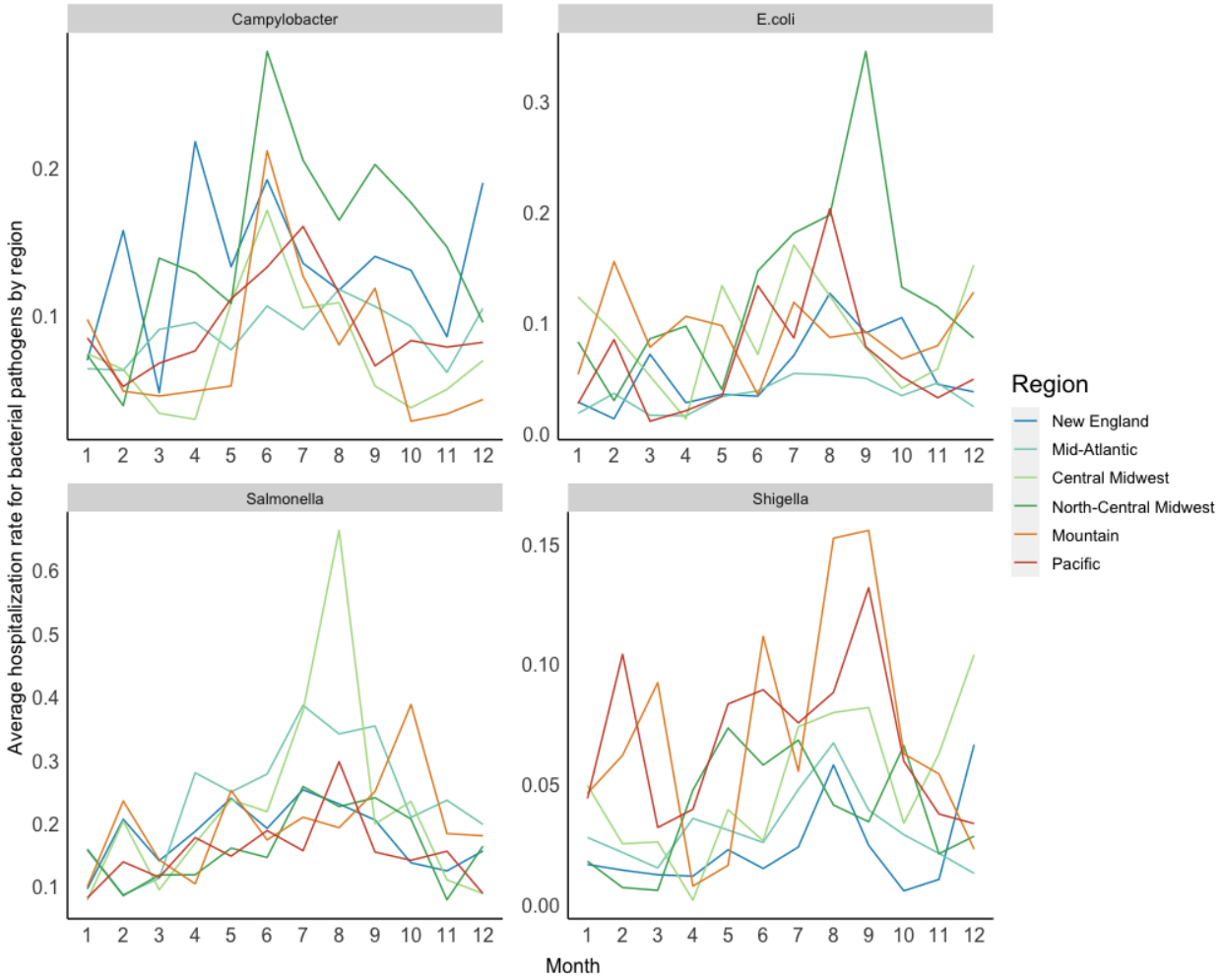

Figure 2.S. 1: Seasonality of bacterial hospitalizations by geographic region.
Average monthly hospitalizations per 10,000 discharges for all bacterial pathogens exhibited clear seasonality, with most peaking in the late summer or early fall. *Campylobacter* hospitalizations peaked earlier in the year compared to the other bacterial pathogens.

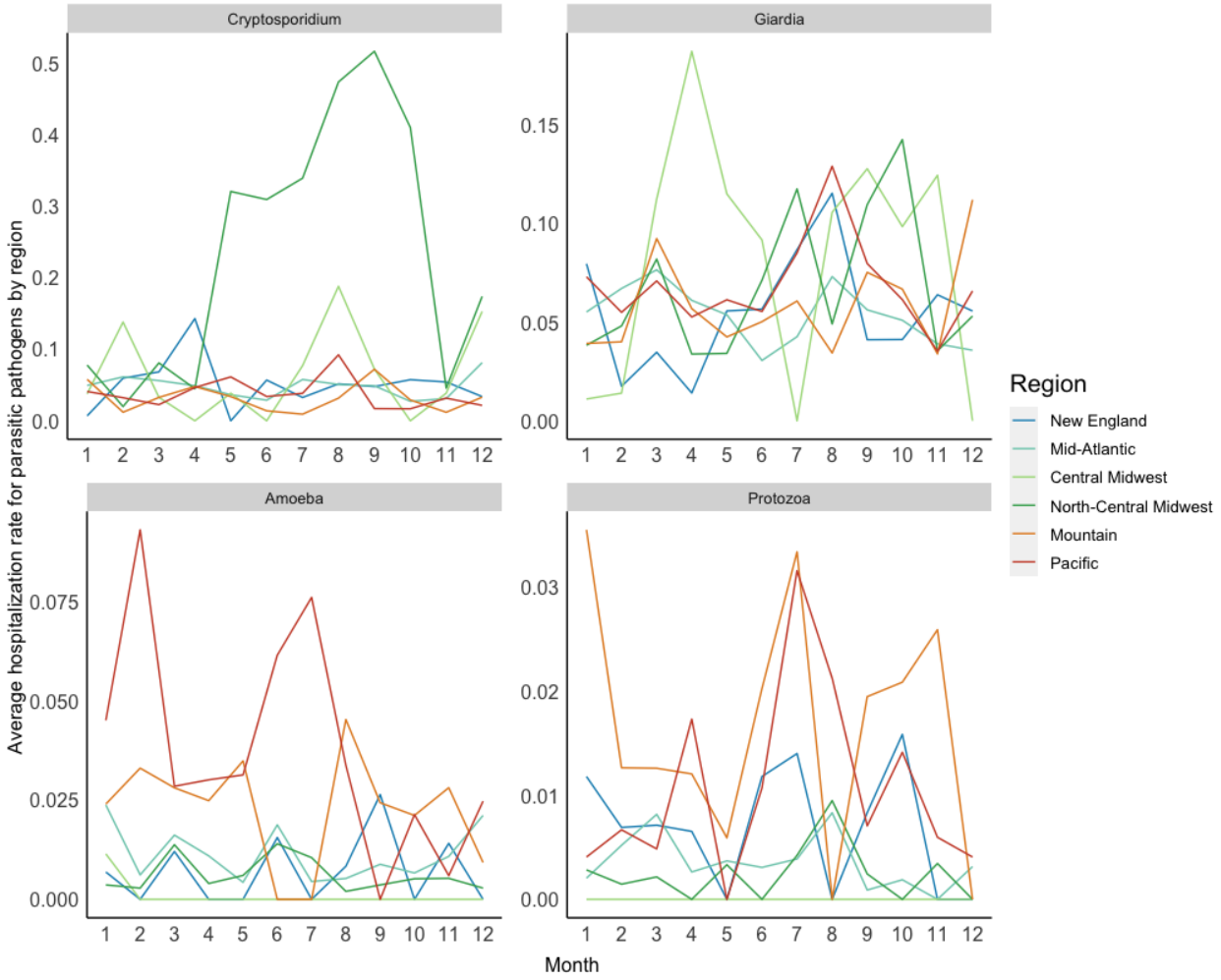

Figure 2.S. 2: Seasonality of parasitic hospitalizations by geographic region.
Average monthly *Cryptosporidium* hospitalizations per 10,000 discharges in the Midwestern regions were the only parasitic hospitalizations to show strong seasonality. *Giardia* hospitalizations in the Pacific region also demonstrated a seasonal peak in August, but there were few hospitalizations in that area.

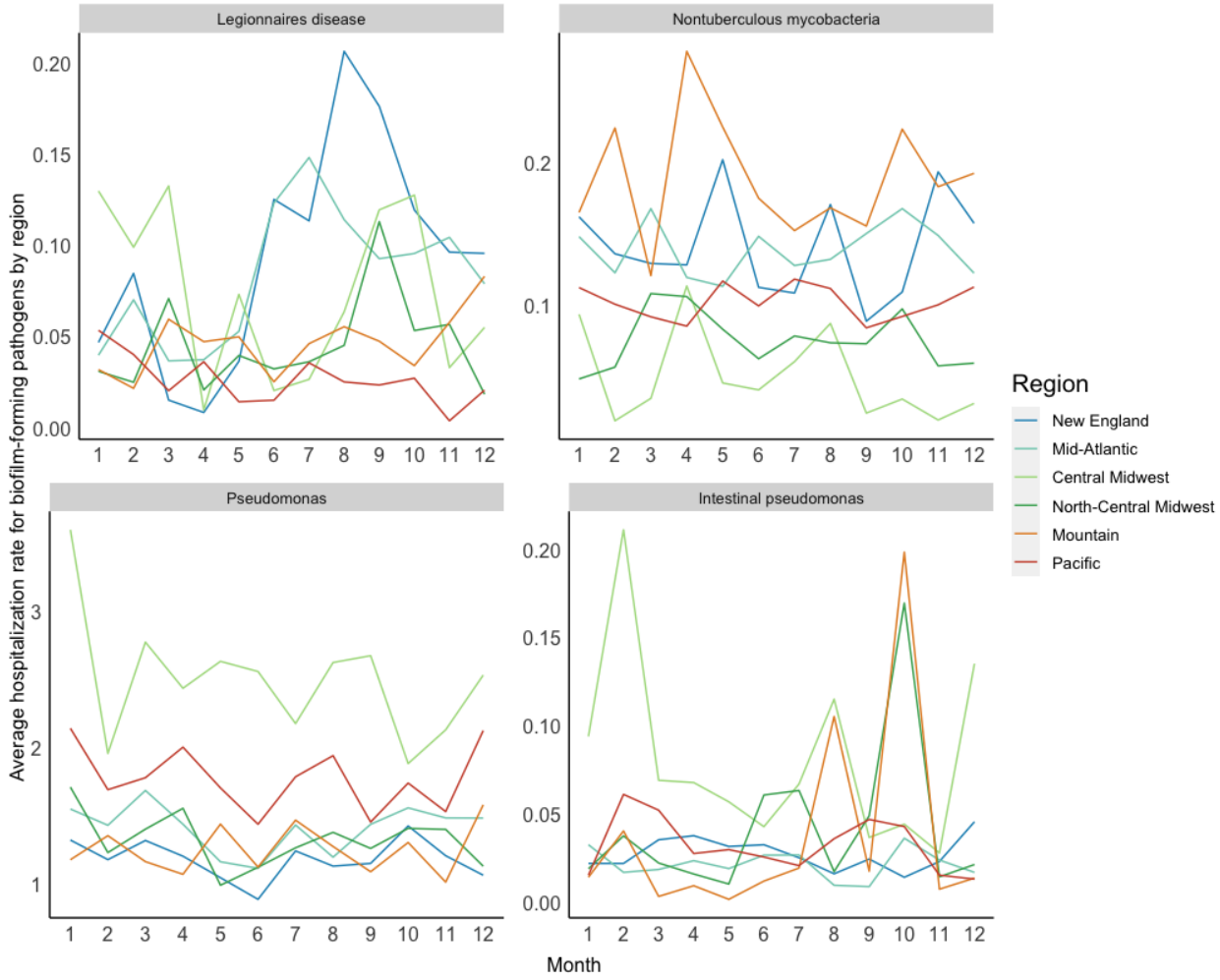

Figure 2.S. 3: Seasonality of biofilm-related hospitalizations by geographic region. Average monthly Legionnaires' disease hospitalizations per 10,000 discharges peaked in New England and Mid-Atlantic states earlier in the year (July – August) compared to Midwestern and Mountain states (September – October). The other respiratory biofilm-forming pathogens exhibited no discernible seasonality. Intestinal *pseudomonas* hospitalizations peaked between October and February in some regions but there were few hospitals in the intestinal *Pseudomonas*-specific dataset.

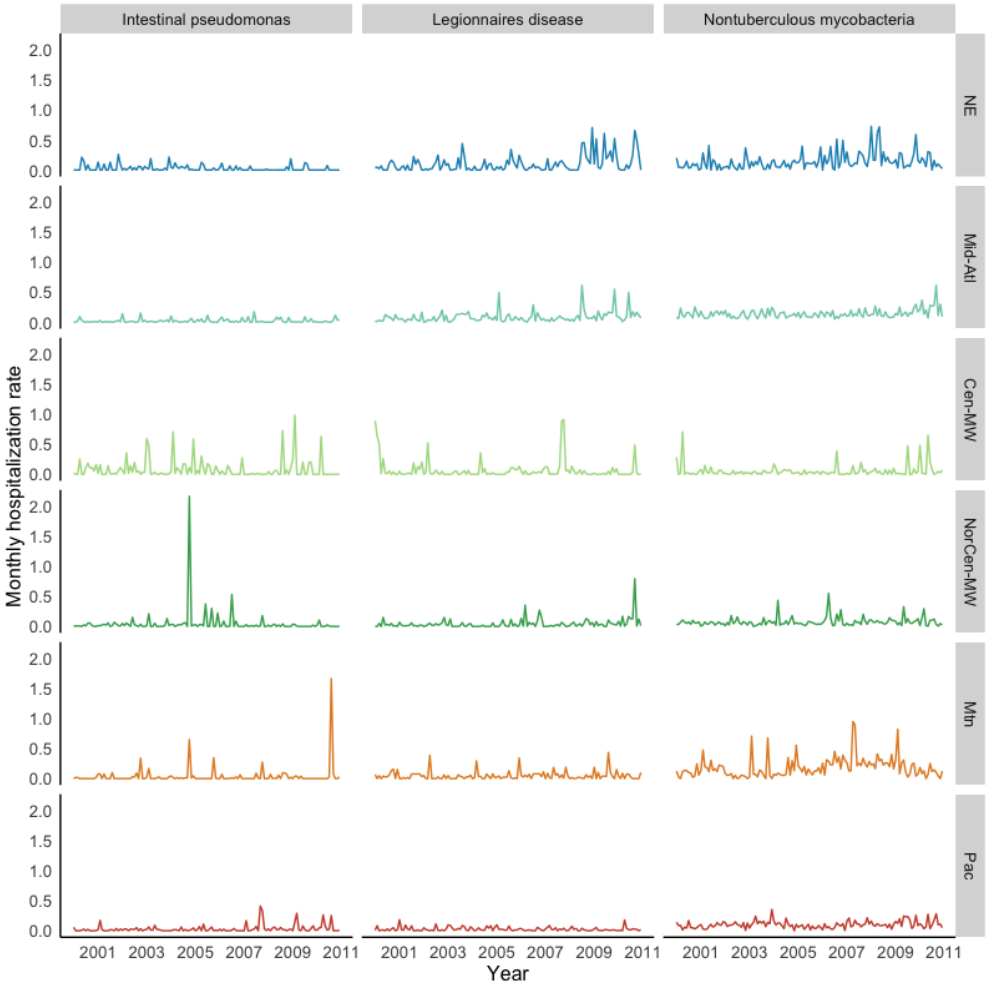

Figure 2.S. 4: Time series for biofilm-related hospitalizations per 10,000 discharges averaged by geographic regions.
Hospitalizations for Legionnaires' disease and NTM increased between 2000 and 2011 in New England, Mid-Atlantic, and Midwestern hospitals.

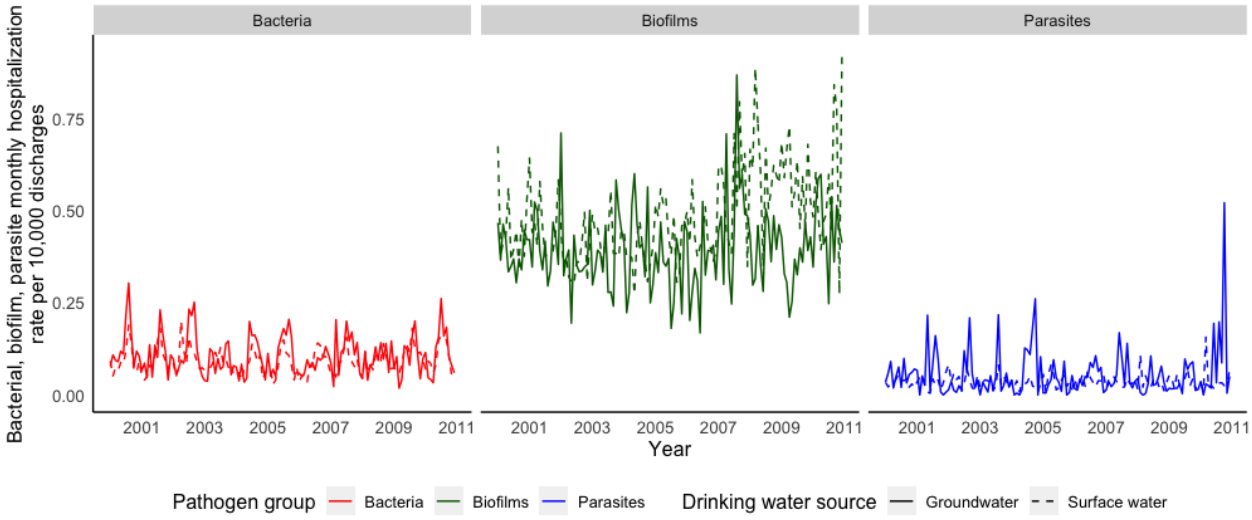

Figure 2.S. 5: Time series for pathogen-group hospitalizations by drinking water source. After 2006, biofilm-related hospitalizations increased in areas served by surface water and decreased in areas that used groundwater for drinking water.

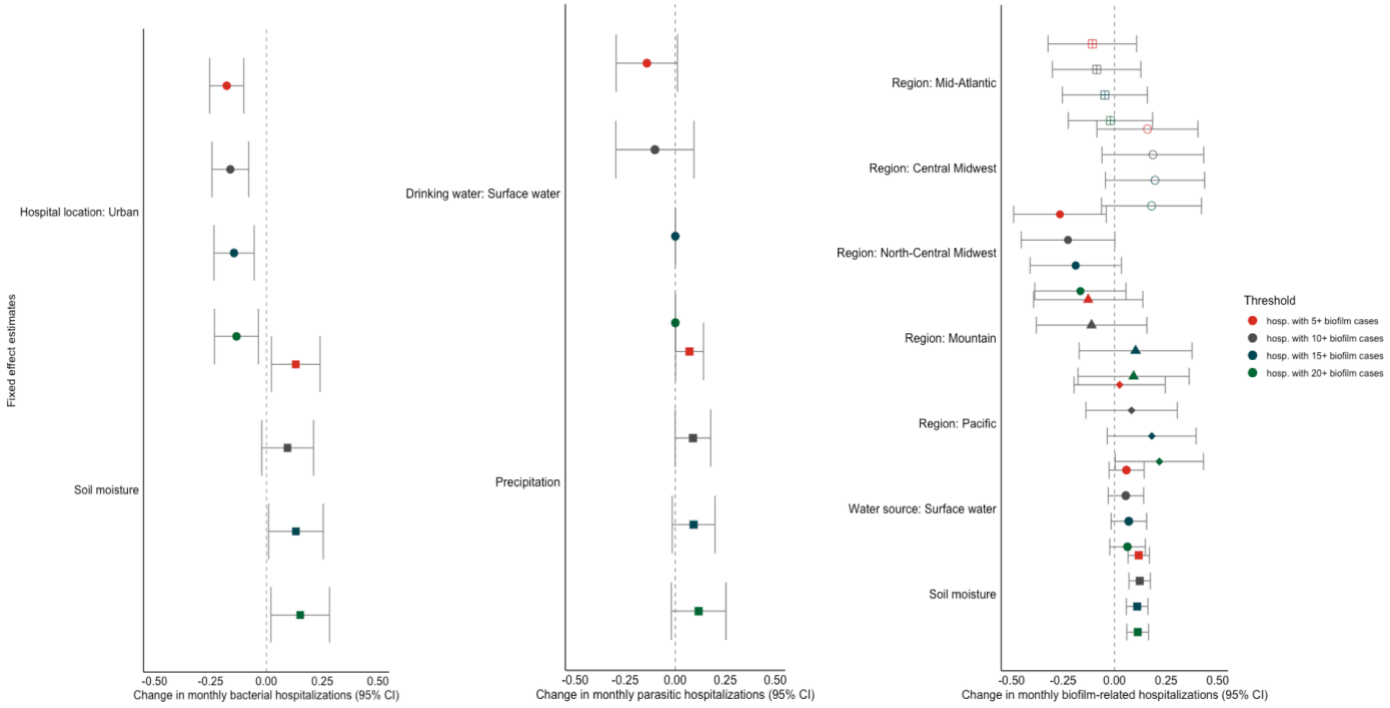

Figure 2.S. 6: Best model effect estimates for each pathogen group across different case-count thresholds.

As a sensitivity analysis, the data were restricted use 5-, 10-, 15-, and 20-case thresholds as cutoffs for inclusion in the hospitalization dataset. The effect estimates were consistent across the case-count thresholds.

Chapter 3

Waterborne infectious disease and tropical cyclone exposure in the Eastern United States

Victoria D. Lynch[1] and Jeffrey Shaman[1]

Affiliations:
[1]Department of Environmental Health Sciences, Columbia Mailman School of Public Health, Columbia University, New York, NY, United States of America
*Corresponding author: Victoria Lynch, vdl2103@cumc.columbia.edu, 722 W. 168th St. New York, NY 10032

Abstract

Tropical cyclones cause destructive flooding in the Eastern United States that can lead to a range of adverse health outcomes. Storm-driven flooding is known to contaminate environmental, recreational, and drinking water sources, but its effect on specific waterborne infections has not been examined over time. In this analysis, we used 23 years of tropical cyclone exposure and weekly state-level case data to assess the effects of storms on six waterborne diseases in a conditional quasi-Poisson model. Storm exposure was defined separately for windspeed, rainfall, and proximity to the storm track; in a secondary analysis, we combined wind and rain exposure to determine the effect of storm types, e.g. high wind-high rain storms, on cases. We found that exposure to storm-related rainfall was associated with increases in Shiga-toxin producing *E. coli* infections (48% increase, 95% confidence interval [CI]: 27-69%) and Legionnaires' disease (42%, 95% CI: 22-62%) one- and two-weeks post-storm, respectively. Cryptosporidiosis cases increased sharply during the storm exposure week (52%, 95% CI: 42-62%) and remained elevated but declined over ensuing weeks. High rain-high wind storms had no effect on cases except for Cryptosporidiosis, which increased sharply three-weeks post-storm, while high rain-low wind storms were associated with delayed increases in Legionnaires' disease cases. Tropical cyclones are a risk to public health that will likely become more serious with increasing storm severity and aging water infrastructure systems. These findings suggest that storm preparedness efforts should focus on identifying and addressing sources of contamination, and on protecting source waters.

3.1 Introduction

Tropical cyclones are a seasonal occurrence in the Eastern United States where they cause widespread destruction and endanger public health [233-235]. Among many storm-related hazards, extreme flooding is a concern because it can lead to the contamination of recreational, irrigation, and drinking waters [112, 145, 236], and may facilitate the transmission of waterborne infectious diseases [187]. Elevated cases and outbreaks [237] have been attributed to individual storms but the effect of tropical cyclones on specific waterborne infections has not been evaluated over multiple storm seasons. Understanding the transmission of waterborne pathogens is a pressing public health challenge, as the burden of disease will likely increase in conjunction with an aging population and deteriorating drinking and wastewater treatment systems [11].

Bacterial, parasitic, and viral pathogens cause an estimated 7,150,000 cases of waterborne disease annually in the US [1]. Infections are typically mild but can lead to life-threatening enteric or respiratory illness for immunocompromised, young, or elderly people [14, 15, 206]. Cyclonic storms may be an important driver of transmission because floodwater mobilizes pathogens in the environment and inundates water system infrastructure, which causes further contamination through ineffective treatment or sewage overflows [160, 238]. High pathogen loads are frequently detected in floodwater [8, 239], as well as environmental [174] and drinking water sources [240, 241], following cyclonic storms. Floods may also affect foodborne transmission by contaminating irrigation water used on crops [148]; several pathogens of concern are predominantly foodborne, but flood-driven contamination may still influence their transmission.

Contamination does not necessarily lead to transmission, however, and while extreme events have been associated with gastrointestinal illness [133] or specific outbreaks [176, 242],

some storms have been found to have no effect on cases [243]. These inconsistent associations reflect the importance of pathogen-specific factors, particularly pathogen biology and primary reservoirs, in determining the effects of storms on transmission. Pathogens that form oocysts or are members of biofilm communities persist in environmental waters for weeks, which may increase the likelihood of transmission [244, 245]; other pathogens that do not persist in the environment may be flushed from waterways by flooding [24]. Pathogen biology also affects the efficacy of water treatment; *Cryptosporidium* and *Legionella*, in particular, are resistant to common decontamination methods [211, 246] whereas *Giardia* is readily removed from water [35]. Cyclonic storms may also lead to different types of contamination depending on the land-use and drinking water or sanitation infrastructure of affected regions. Cattle and poultry are the primary reservoirs for several gastrointestinal pathogens, and flooding near livestock production can contaminate drinking water sources with animal waste [81]. This is of particular concern in rural agricultural regions where many people rely on private wells that are untreated and vulnerable to inundation [179]. Storms in densely populated areas, meanwhile, often lead to floodwater contaminated with human sewage [247]. Urban flooding can also damage water treatment or distribution systems that serve entire cities, leading to large outbreaks [248].

The effect of cyclonic storms on waterborne disease may also depend on the storm characteristics that determine the extent of flooding and destruction. Storms are generally defined by windspeed and rainfall, factors that are often weakly correlated upon landfall [249] and lead to different conditions in affected areas. Slow-moving storms tend to cause greater accumulation of rain and more severe flooding whereas tropical cyclones with high windspeeds may cause more damage but bring less rain [233, 250]. Storm type may also dictate disaster management decisions and individual-level response to storm events. Evacuation orders, and the

ability to comply with them, may depend on storm type and could affect the number of people exposed to contaminated water. Storm severity also influences healthcare-seeking behavior and healthcare infrastructure. Storm-related disruptions may dissuade people with mild or moderate conditions form seeking care [251] whereas catastrophic storms may prevent people with urgent needs from accessing healthcare systems [252].

In this study, we examined the effects of tropical cyclones on waterborne infectious diseases and whether these associations varied by pathogen or type of storm exposure. Previous research has largely focused on specific storms and outbreaks, or on non-specific gastrointestinal illness; associations over multiple storm seasons, however, have not been examined. Storm severity is projected to increase with atmospheric warming, so developing a thorough understanding of storm effects on waterborne diseases could aid climate adaptation and public health policies.

3.2 Methods

3.2.1 Data

Case data

We used surveillance data from the National Notifiable Diseases Surveillance System (NNDSS) to identify weekly cases of Cryptosporidiosis, Legionnaires' disease, *E. coli* infections, Giardiasis, Salmonellosis, and Shigellosis from each state in the US between 1996 and 2018. The six infections included in this study are caused by bacterial and parasitic pathogens and can lead to severe gastrointestinal or respiratory illness. Of the six *E. coli* strains, only Shiga-toxin producing *E. coli* (STEC) infections are tracked in the NNDSS. The data consist of laboratory-confirmed cases from hospitalizations, emergency department visits, and primary care visits that are reported to local health departments and compiled by state health departments to submit to

Centers for Disease Control and Prevention (CDC), which manages the NNDSS. We restricted our analyses to the 30 states, and Washington, D.C., that experienced at least one tropical cyclone during the study period and to the months of the Atlantic storm season (June – November). We also used US Census data to determine county and state populations throughout the study period.

Storm data

Storm track, windspeed, and rainfall data for tropical cyclones that made landfall in the US between 1996 and 2018 were obtained from the hurricaneexposure (version 0.1.1) and hurricaneexposuredata (0.1.0) R packages. For each county, the day with the shortest distance between the geographic county center and the storm track was defined as the primary exposure day. Storm track and surface windspeed data were from the National Hurricane Center's HURDAT-2 dataset and included maximum and sustained windspeeds on the primary exposure day. Rainfall data were from the NASA/NOAA North American Land Data Assimilation System 2 (NLDAS-2) and were included in the dataset as total daily rainfall in each exposed county from five days before to three days after the primary exposure day. Correlations among distance, wind, and rainfall variables, including total and daily maximum rainfall, were assessed to inform the selection of exposure variables used in the analysis.

3.2.2 Storm exposure definition

Informed by the correlation analysis of storm variables, we defined county-level exposure to storms according to total rainfall, sustained gust windspeed, and distance from the storm track. In the primary analysis, we defined exposure separately for each variable and repeated the analyses using several exposure thresholds. Counties were considered exposed if they experienced 50-mm, 75-mm, or 100-mm of total rainfall associated with the storm or were

within 500-km, 250-km, or 150-km of the storm track. For windspeed, counties were considered exposed to tropical storms if they experienced sustained gusts of at least 34 knots (gale-force wind on the Beaufort scale) and to hurricanes if sustained gusts exceeded 64 knots. Correlations among the exposure thresholds were assessed. To determine the state-level exposure, we calculated the percent of the state population in exposed counties during storm weeks and classified the state as exposed if 75%, 50%, 25%, 5% or any of the population was exposed; we repeated the analysis for each of these population thresholds.

In the secondary analysis, we combined storm exposure variables to describe categories of cyclonic storms. We categorized storms as: 1) "high rain-high wind" if total rainfall was greater than 100-mm and windspeed exceeded 64 knots, 2) "high rain-low wind" if total rainfall was greater than 100-mm and windspeed was between 34 and 64 knots, or 3) "low rain-low wind" if total rainfall was less than 100-mm and windspeeds were between 34 and 64 knots; a low rain-high wind category was not included as no storms met the definition. Counties were considered exposed to a specific storm type if they met both the rainfall and windspeed criteria. Hurricane-force winds are rare and usually affect a small proportion of a state's population (S2 Table); therefore, population-exposure thresholds were defined by rainfall exposure as in the primary analysis. A state was considered exposed to a given storm type if: 1) it met the rainfall-based population exposure threshold, and 2) any of the counties were exposed to the given storm type.

3.2.3 Statistical analysis

We modeled the association between exposure to tropical cyclones and cases of waterborne diseases using a conditional quasi-Poisson model, which accounted for overdispersion in the case data (S1 Model) [253]. The number of cases in weeks with and

without storms was compared across matched strata based on state and week-of-the-year. This structure addresses potential confounding due to variation among states, i.e. different state policies regarding storm preparedness or case reporting, and controls for seasonality. Cyclonic storm occurrence was modeled as a binary exposure variable and lagged from 0 to 3 weeks to account for the incubation periods of the pathogens and the potential for delays in seeking healthcare after destructive storms. The model included a flexibly adjusted term for year to control for long-term trends that could affect storm exposure or cases of waterborne infectious diseases. Annual state population was used as an offset to obtain the rate of cases and we modeled cases for each pathogen separately. We present all results as percent changes in weekly case rates. The analysis was repeated for all exposure definitions and population exposure thresholds. 95% CIs were adjusted for multiple comparisons using the Bonferroni-Holmes method. Finally, this method was repeated with counties stratified by drinking water source or location (rural/urban); this supplementary analysis is described in S1 Appendix.

3.3 Results

The number of cases reported to NNDSS varied by pathogen with enteric bacteria comprising the majority of infections (Table 3.1). Most peaked in the late summer or early fall but the amplitude of seasonality differed among pathogens and by geographic region (Figure 3.1).

Table 3. 1 Description of pathogens included in the analysis

Pathogen	Cases in NNDSS (% of total)	Pathogen type	Incubation period (range) [16]	Estimated % waterborne[1]	Years in NNDSS
Legionella	129,559 (3.4)	Biofilm-forming bacteria	5-6 days (2 – 10)	97	1996 – 2018
Cryptosporidium	283,030 (7.4)	Parasite	7 days (2 – 12)	43	1998 - 2018
Giardia	534,911 (14)	Parasite	7 days (1 – 14)	44	2002 - 2018
E. coli	239,354 (6.3)	Bacteria	0.5 – 4 days (0.5 – 10)	5	1996 - 2018
Salmonella	1,845,428 (48.5)	Bacteria	0.5 – 2 days (0.5 – 16)	6	1998 - 2018
Shigella	775,563 (20.4)	Bacteria	1 – 3 days (0.5 – 7)	4	1998 - 2018

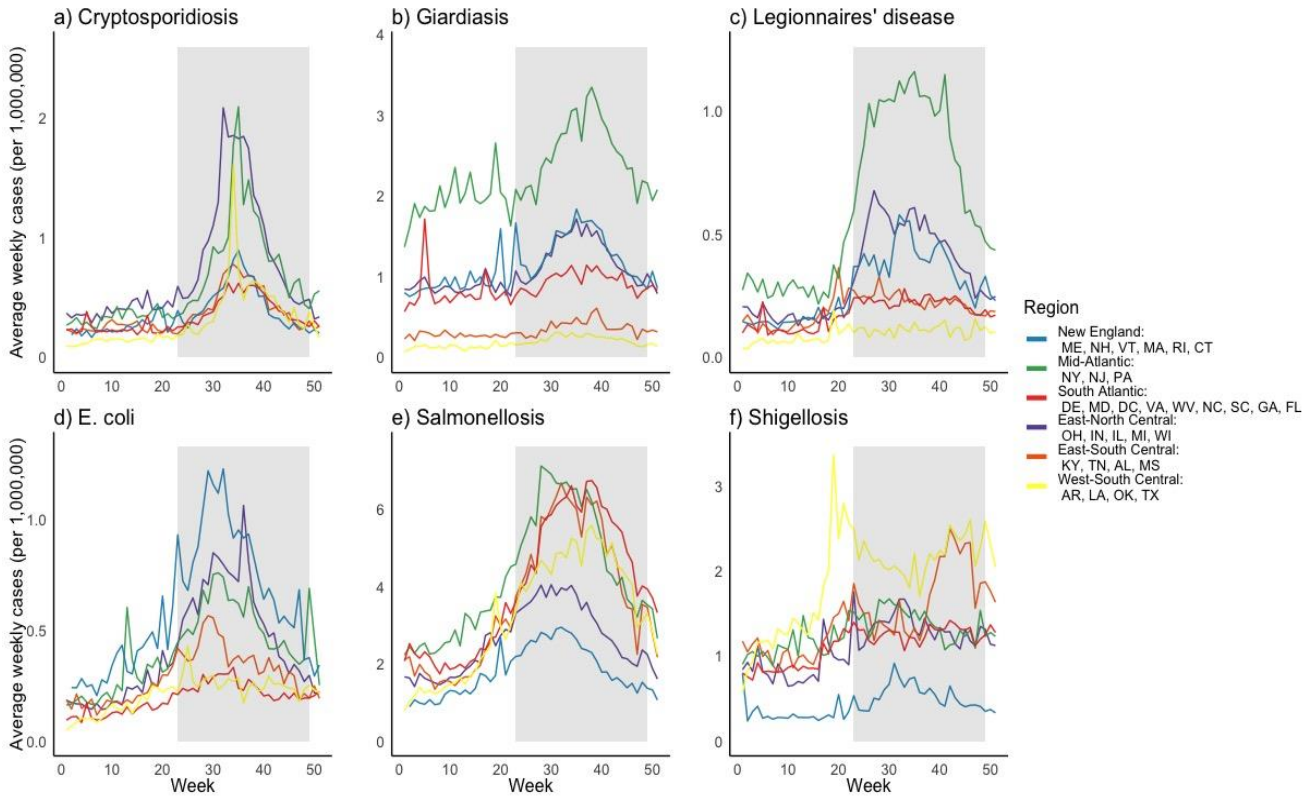

Figure 3. 1:Average weekly cases by geographic region and infection.
Average weekly cases per 1,000,000 people by geographic region (colors) and infectious disease (a-f) reported to NNDSS between 1996 and 2018; not all infections were reported for the entire study period (Table 3.1). The shaded region represents the weeks encompassed in the Atlantic storm season (June 1 to November 30). The geographic regions reflect the reporting areas used by CDC for infectious disease surveillance.

Cryptosporidiosis exhibited the strongest and most consistent seasonality as cases peaked in September in all geographic regions. In the Northeast and Upper Midwest, Legionnaires' disease cases were consistently elevated between June and October but did not exhibit a singular peak. In most states, Legionnaires' disease and the parasitic infections displayed only a moderate increase during summer months (Figure 3.S.1). Enteric bacterial infections were more common across all states and Salmonellosis showed a strong summer seasonality in most states (Figure 3.S.2). Between 1996 and 2018, Legionnaires' disease and Cryptosporidiosis cases increased, and Giardiasis decreased, in all geographic regions; the other pathogens were relatively consistent over time (Figure 3.S.3).

Wind, rainfall, and distance variables were not highly correlated, though different measures of the same variable, i.e. maximum rainfall and total rainfall, were correlated (Figure 3.S.4). Among the storm variable thresholds used to determine county-level exposure, hurricane- and gale-force wind exposure were not highly correlated (r = 0.21) while ≥ 50-mm, ≥ 75-mm, and ≥ 100-mm rainfall exposure thresholds were (r = 0.50 to 0.72) (Figure 3.S.5). There were 134 cyclonic storms during the study period using the most inclusive storm exposure threshold (gale-force wind) (Table 3.2). These storms affected 2,363 counties in 30 states and Washington, D.C. over 177 weeks. Counties with the greatest number of wind exposure storm weeks were concentrated along the coast, particularly in North and South Carolina (Figure 3.2). Rainfall exposure was most common in South Florida but was overall more widespread and uniform than the wind and distance metrics (Figure 3.2). There was no long-term trend in the number of cyclonic storms during the study period (Figure 3.S.6).

Table 3. 2 Description of cyclonic storm exposure definitions

Storm exposure variable	Exposure definition	Number of storms	Number of affected counties
Total rainfall	50-mm	98	2,165
	75-mm	96	2,041
	100-mm	87	1,732
Sustained gusts	Gale-force winds (kts)	134	1,025
	Hurricane-force winds (kts)	31	136
Distance from storm track	500-km	134	2,363
	250-km	134	2,179
	150-km	117	2,072

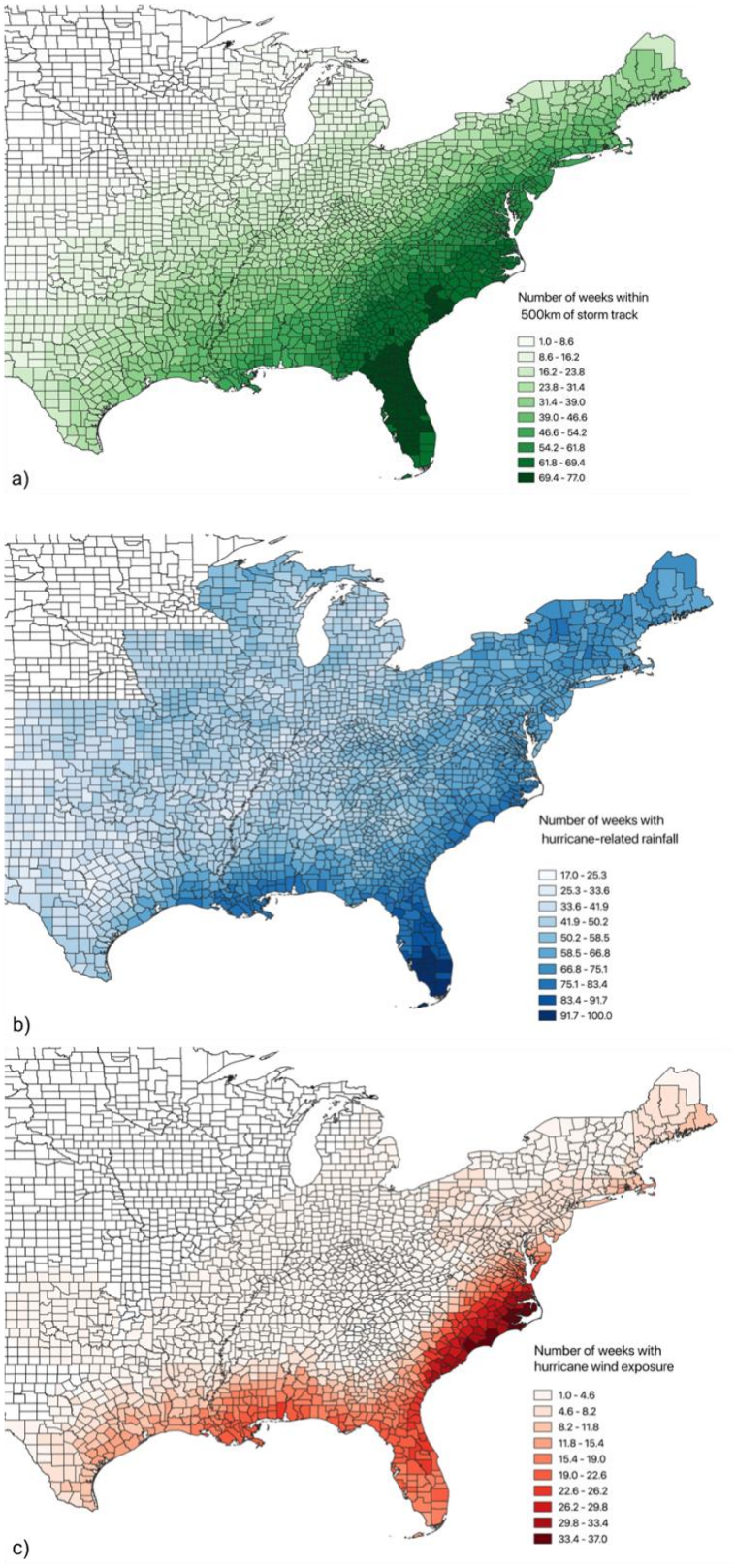

Figure 3. 2: Number of weeks of storm exposure per county by storm hazard.
Number of weeks of storm exposure per county when exposure is defined by a) distance to storm track (within 500-km of track), b) cumulative rainfall (minimum 75-mm), and c) sustained gusts above gale-force wind (minimum 34 knot

Cryptosporidiosis cases significantly increased during storm weeks at low population exposure thresholds; storms that brought ≥ 75-mm of rainfall were associated with a 40% (95% CI: 31% - 50%) increase in weekly cases when any of the state's population was exposed and a 52% (95% CI: 42% - 62%) increase when at least 5% was exposed (Figure 3.3). Similar associations persisted across lagged exposures, though the effects were weaker, ranging from 8% (95% CI: -2%, 19%) to 19% (95% CI: 8% - 31%) increases in the post-storm weeks (Table 3.S.1). Legionnaires' disease cases were also significantly associated with storm exposure but the effect was strongest 2 and 3 weeks after a storm and at higher population exposure thresholds (Figure 3.3). When 75% of the state population was exposed to a storm, cases increased by 31% (95% CI: 1% - 52%), 42% (95% CI: 22% - 62%), and 39% (95% CI: 26% - 54%) in lagged weeks 1-3 (Table 3.S.1). *E. coli* cases exhibited a clearer peak and decline associated with lagged storm events. After an initial decrease during the storm week, cases increased 48% (95% CI: 27% - 69%) and 33% (95% CI: 11% - 56%) in the first two weeks post-storm when 75% of the state population was exposed (Figure 3.3). Salmonellosis and Giardiasis were not significantly associated with storm exposure and Shigellosis cases slightly decreased during storm weeks (Figure 3.3).

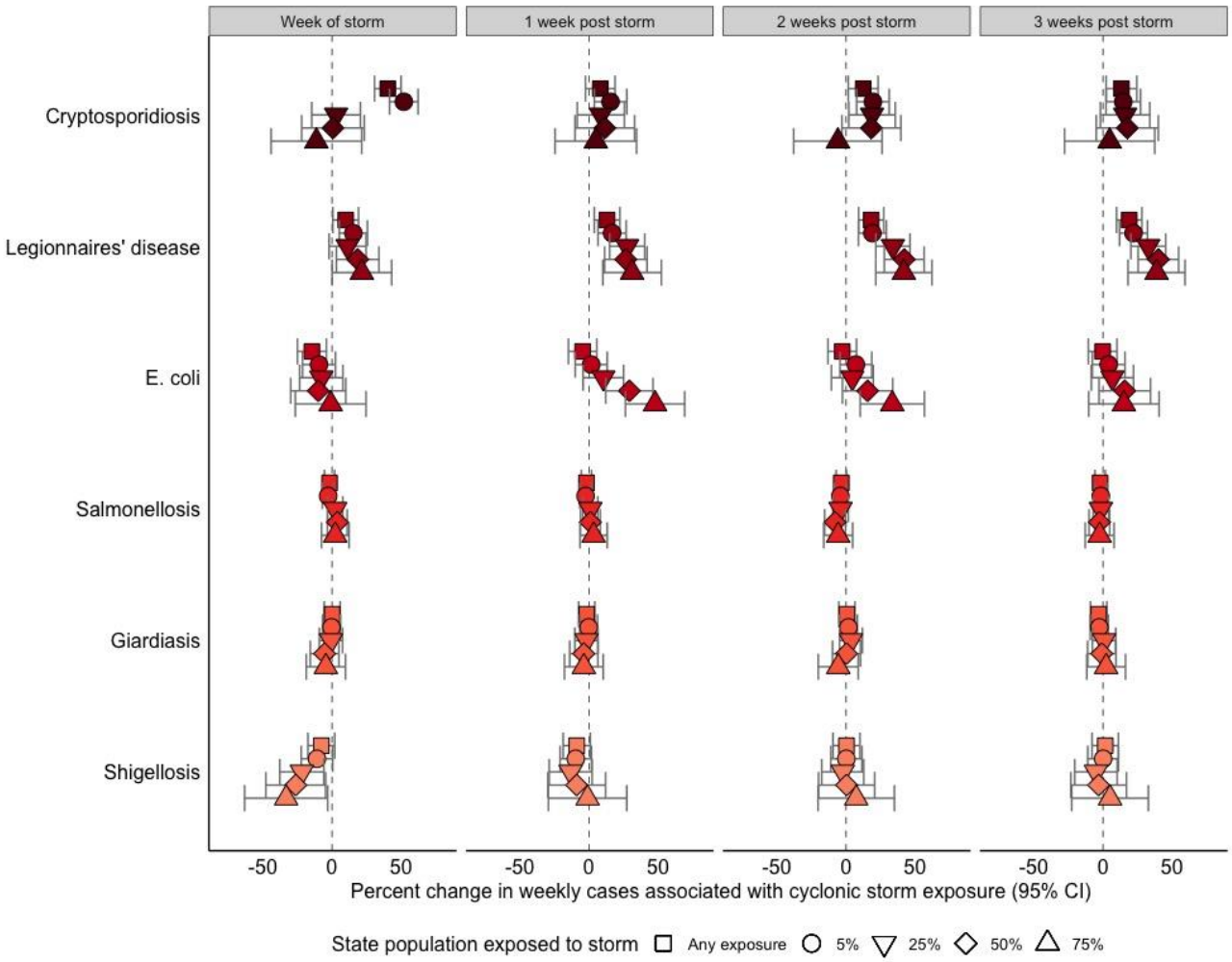

Figure 3. 3: Change in case rates associated with exposure to 75-mm of storm-related rainfall. Percent change in weekly case rates associated with exposure to ≥ 75-mm of storm-related rainfall. The estimates and Bonferroni-corrected 95% confidence intervals are reported for each infectious disease (shade) and population-exposure threshold (shape); estimates are reported for week of the storm and 1 to 3 weeks post-storm.

The associations between storm-related rainfall and Cryptosporidiosis, Legionnaires' disease, and *E. coli* case rates were consistent across different exposure definitions (Figure 3.4). Storms with more (≥100-mm) and less (≥50-mm) rainfall were associated with significant initial increases in Cryptosporidiosis case rates that attenuated over weeks 1 to 3. The strength of the association between Legionnaires' disease case rates and storm exposure increased in conjunction with population exposure threshold and amount of rainfall (Figure 3.4). Similarly,

the lagged increase in *E. coli* case rates was more pronounced in storms with ≥ 100-mm rainfall. The associations between case rates and storm exposure were similar when exposure was defined by distance from the storm track instead of rainfall (Figure 3.S.7). Stratifying exposure by drinking water source or rural/urban location also yielded similar results; the lagged effect on *E. coli* and Legionnaires' disease case rates was slightly more pronounced when restricted to rural or groundwater-reliant counties, but associations were otherwise consistent (S1 Appendix).

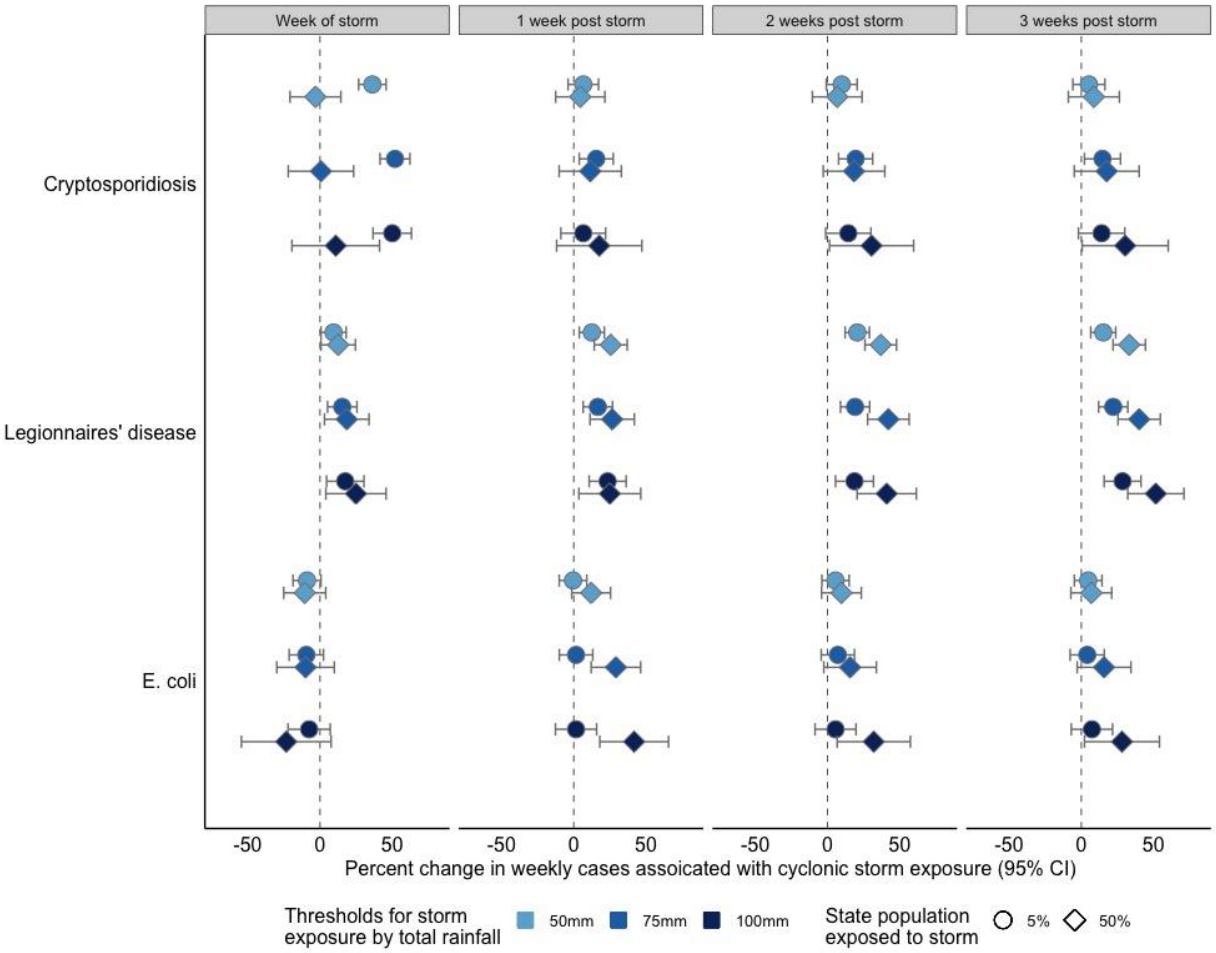

Figure 3. 4: Change in case rates associated with exposure to storm-related rainfall. Exposure is defined by three cumulative rainfall thresholds: 1) ≥ 50-mm (light blue), 2) ≥ 75-mm (medium blue), or ≥ 100-mm (dark blue) and for two population-exposure thresholds (shape). The estimates and Bonferroni-corrected 95% confidence intervals are reported for Cryptosporidiosis, Legionnaires' disease, and *E. coli* infections for the week of the storm and 1 to 3 weeks post-storm.

Storm exposure defined by hurricane-force winds was associated with increased Cryptosporidiosis case rates 2 and 3 weeks after storms but otherwise had no effect on cases (Figure 3.S.8). Conversely, gale-force wind exposure was associated with decreased Cryptosporidiosis and Giardiasis case rates during the storm week and had no effect in the lagged weeks post-storm (Figure 3.S.8).

Combining wind and rainfall exposure in storm-type categories supported the findings of the wind exposure analysis. High rain-high wind, high rain-low wind, and low rain-low wind storms were all associated with decreased Giardiasis case rates during the storm week before returning to baseline 1 week post storm (Figure 3.5). Consistent with the rainfall analysis, high rain-low wind storms were positively associated with Cryptosporidiosis case rates up to 2 weeks post-storm but, unlike rainfall alone, case rates also increased 3 weeks-post high rain-high wind and low rain-low wind storms: a 58% (95% CI: 30% - 78%) increase in Cryptosporidiosis case rates when at least 5% of the population was exposed to high wind-high rain storms and a 17% (95% CI: 2% - 35%) increase after low rain-low wind storms (Figure 3.5). Finally, there was no effect of high rain-high wind storms on Legionnaires' disease case rates and the effect of high rain-low wind storms and low rain-low wind storms was only apparent 3 weeks after storms (Figure 3.5).

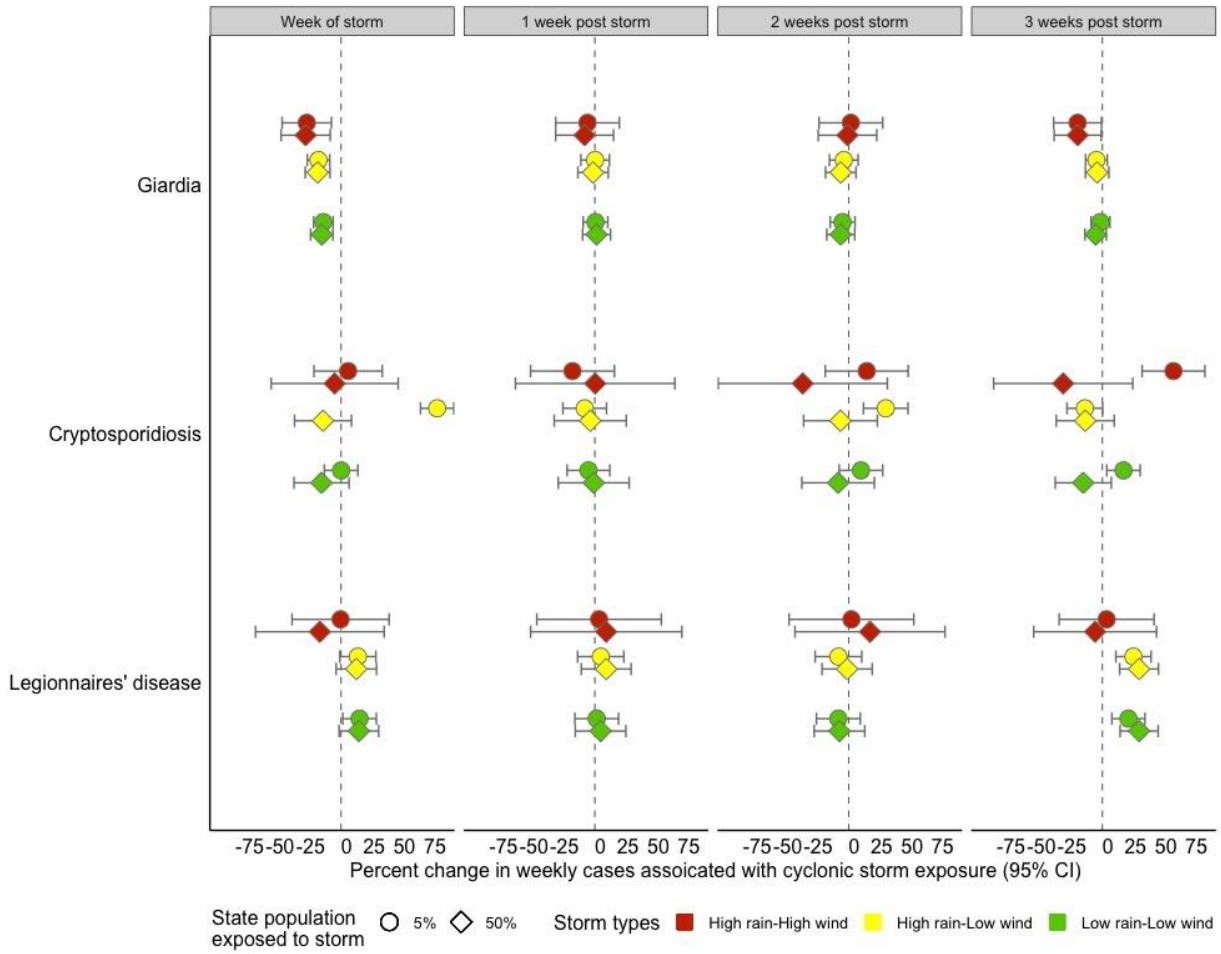

Figure 3. 5: Change in weekly case rates associated with exposure to storm types. Exposure is defined by three categories according to rainfall and wind thresholds: 1) High rain-High wind storms (red), 2) High rain-Low wind storms (yellow), and 3) Low rain-Low wind storms (green). The estimates and Bonferroni-corrected 95% confidence intervals are reported for Giardiasis, Cryptosporidiosis, and Legionnaires' disease at two population-exposure thresholds (shape), and for the week of the storm and 1 to 3 weeks post-storm. The population-exposure thresholds refer to the percent of the state population exposed to storm-related rainfall only; no hurricane-force winds affect >25% of the state population.

3.4 Discussion

Tropical cyclones were associated with waterborne diseases in the United States, though the effect magnitude varied by exposure. The associations also differed among the specific pathogens; Legionnaires' disease, *E. coli*, and Cryptosporidiosis increased with rainfall whereas Salmonellosis, Shigellosis, and Giardiasis were unaffected, or decreased, during storm weeks.

These divergent associations likely reflect factors that mediate the relationship between storms and disease, including pathogen biology, transmission routes, and severity of infection.

Legionnaires' disease and *E. coli* case rates consistently increased with rainfall and population exposure thresholds, though the timing of the effects differed between these infections. *E. coli* case rates peaked 1 week after storms and returned to baseline by week 3 whereas Legionnaires' disease case rates were highest 3 weeks post storm. These findings support microbiological studies that have analyzed bacterial counts in streams [254] and water systems [255] after specific hurricanes; elevated *E. coli* loads were reported 12 to 24 hours after a storm started whereas *Legionella* increased 4 to 5 days later [254]. *Legionella* are natural inhabitants of aquatic environments and replicate in water, typically in biofilm communities that colonize household plumbing and water infrastructure systems [17, 206]. The *Legionella* load, therefore, can increase over time whereas for other bacterial pathogens, which do not replicate in the environment, the bacterial load typically peaks after the initial contamination event and dissipates over time [49].

Cryptosporidiosis case rates also increased with storm-related rainfall but only at low population thresholds, which may be due to the limited overlap between areas with high disease burden and vulnerability to tropical cyclones. Cryptosporidiosis infections were most common in the North-Central Midwest, a region that infrequently experiences tropical storms or hurricanes severe enough to affect more than 25% of the population. The substantial increase in case rates concurrent with storm week may be driven by several widespread outbreaks attributed to specific storm events that damaged water treatment facilities [256]. *Cryptosporidium* is resistant to standard chemical disinfectants and is small enough to pass through sand filtration systems

common in water treatment plants [24]; as a result, when the parasite contaminates water distribution systems that serve large populations, it can lead to massive outbreaks [237].

County-level exposure to heavy rainfall and cyclonic windspeed was often uncorrelated, which is characteristic of tropical cyclones [249], and the effect of extreme wind on cases differed from that of rainfall for several pathogens. Gale-force wind was associated with a lagged increase in Legionnaires' disease, but the effect on *E. coli* and Cryptosporidiosis was insignificant; hurricane-force wind was only associated with increased Cryptosporidiosis case rates 3 weeks post storm. Such attenuated effects may reflect the intricate, and possibly opposing, factors that influence transmission. High windspeeds are typically associated with destructive storms that can damage sanitation infrastructure, increasing the probability of transmission [8], but also lead to displacement [257, 258] that reduces the likelihood of people coming into contact with contaminated water. Extreme storms can also disrupt healthcare systems or alter healthcare seeking behavior, which can lead to a reduction in detection or reporting of cases [259].

For areas that experienced both rainfall and cyclonic wind, we combined exposures into storm-type categories; the results underscored the importance of pathogen-specific analyses when evaluating the effect of tropical cyclones and the limitations inherent in studying events that rarely occur. The high rain-high wind category encompassed the most devasting storms that occurred during the study period (i.e. Hurricane Katrina, Hurricane Ivan) but represented a small fraction of all storms. These events were associated with a significant decrease in Giardiasis but had no effect on Legionnaires' disease. *Giardia* transmission often occurs in recreational waters, i.e. swimming pools, rivers, etc., and may be thwarted during storm weeks when people are less likely to engage in those activities. The burden of Legionnaires' disease, meanwhile, was highest

in regions that infrequently experience hurricane-force winds. High rain-high wind storms were associated with a substantial increase in Cryptosporidiosis case rates 3 weeks post-storm, though the effect may have been driven by a two-month span in 2008 when Texas experienced two hurricanes and a tropical cyclone in succession, and reported extremely high case counts for an extended period.

Unlike the other infections, Salmonellosis was unaffected by cyclonic storms at all population thresholds. *Salmonella* transmission is predominantly foodborne and outbreaks attributed to contaminated food are common, particularly during the summer [58]. The high frequency of outbreaks makes it difficult to detect elevated case-counts associated with storms, given that comparison weeks may coincide with foodborne outbreaks. Storm-related rainfall was associated with a slight decrease in Shigellosis at high population thresholds during storm weeks. These infections are typically mild and the negative association may reflect a reduction in seeking healthcare for minor illness after the disruption of storm events.

With the exception of Shigellosis, cases exhibited a summer seasonality that coincided with the cyclonic storm season in the US. The inconsistent associations between storms and specific pathogens, however, demonstrates that the effects were not simply driven by overlapping seasonal patterns. Salmonellosis and *E. coli* cases peaked during the same weeks in most regions but storm-related rainfall had no effect on the former and a strong positive effect on the latter. This study demonstrated the need for more pathogen-specific analyses that combine microbiological water quality data from multiple sources with epidemiological data; this approach could help explain why pathogens with similar biology or transmission routes exhibited divergent associations with cyclonic storms.

This analysis is constrained by several limitations, the most prominent of which is the spatial mismatch between the case and storm data. Aggregating from county- to state-level storm exposure introduced the possibility of misclassification bias given that state-level exposure may be inconsistent with the conditions experienced by cases. We aimed to address this limitation by repeating the analysis using several population thresholds to define exposure and assessing the consistency of the associations. This type of nondifferential misclassification would also be biased toward the null and underestimate the associations [260]. Given the spatial resolution, we could only perform a rough estimate of the effect of storms stratified by drinking water source or rural/urban location using county-level averages; highly resolved water source and location data could provide insight into the mechanisms underlying the associations between storms and some waterborne diseases. Identifying the drivers of transmission, and opportunities for intervention, is crucial as the US has aging sanitation infrastructure [11] and an aging population that is more susceptible to serious waterborne disease [206].

3.5 Supplementary Materials

Table 3.S. 1 Effect estimates with rainfall exposure

	Population expo. (%)	Week of storm	1 week post storm	2 weeks post storm	3 weeks post storm
Cryptosporidiosis	Any	0.4 (0.31, 0.5)	0.08 (-0.02, 0.19)	0.12 (0.02, 0.23)	0.13 (0.02, 0.24)
	5	0.52 (0.42, 0.62)	0.16 (0.04, 0.27)	0.19 (0.08, 0.31)	0.15 (0.02, 0.27)
	25	0.03 (-0.14, 0.2)	0.09 (-0.08, 0.25)	0.19 (0.02, 0.35)	0.16 (-0.01, 0.33)
	50	0.01 (-0.22, 0.23)	0.11 (-0.1, 0.33)	0.18 (-0.03, 0.39)	0.18 (-0.04, 0.4)
	75	-0.11 (-0.43, 0.21)	0.05 (-0.24, 0.34)	-0.06 (-0.37, 0.25)	0.05 (-0.27, 0.37)
Legionnaires' disease	Any	0.1 (0.01, 0.19)	0.13 (0.04, 0.22)	0.18 (0.09, 0.27)	0.19 (0.1, 0.28)
	5	0.15 (0.05, 0.25)	0.17 (0.07, 0.27)	0.19 (0.09, 0.29)	0.22 (0.12, 0.32)
	25	0.11 (-0.02, 0.24)	0.28 (0.15, 0.4)	0.34 (0.22, 0.46)	0.33 (0.2, 0.45)
	50	0.19 (0.03, 0.34)	0.27 (0.12, 0.42)	0.42 (0.28, 0.56)	0.4 (0.26, 0.54)
	75	0.22 (0, 0.43)	0.31 (0.1, 0.52)	0.42 (0.22, 0.62)	0.39 (0.19, 0.59)
E. coli	Any	-0.15 (-0.25, -0.04)	-0.05 (-0.15, 0.05)	-0.03 (-0.13, 0.07)	0 (-0.1, 0.1)
	5	-0.09 (-0.21, 0.02)	0.02 (-0.1, 0.13)	0.07 (-0.04, 0.18)	0.04 (-0.07, 0.16)
	25	-0.08 (-0.23, 0.07)	0.1 (-0.04, 0.25)	0.05 (-0.1, 0.19)	0.07 (-0.08, 0.22)
	50	-0.1 (-0.3, 0.1)	0.29 (0.12, 0.46)	0.16 (-0.02, 0.33)	0.16 (-0.03, 0.34)
	75	-0.01 (-0.26, 0.24)	0.48 (0.27, 0.69)	0.33 (0.11, 0.56)	0.15 (-0.1, 0.4)
Salmonellosis	Any	-0.02 (-0.05, 0.02)	-0.02 (-0.05, 0.02)	-0.03 (-0.07, 0)	-0.02 (-0.06,0.02)
	5	-0.03 (-0.07, 0.01)	-0.03 (-0.07, 0.02)	-0.04 (-0.08, 0)	-0.02 (-0.06,0.03)
	25	0.02 (-0.03, 0.08)	0.01 (-0.04, 0.06)	-0.04 (-0.09, 0.02)	-0.02 (-0.07,0.04)
	50	0.04 (-0.03, 0.11)	0.01 (-0.06, 0.08)	-0.08 (-0.15, -0.01)	-0.03 (-0.1, 0.05)
	75	0.02 (-0.07, 0.12)	0.03 (-0.06, 0.13)	-0.06 (-0.16, 0.05)	-0.02 (-0.13,0.08)
Giardiasis	Any	-0.02 (-0.05, 0.02)	-0.02 (-0.05, 0.02)	-0.03 (-0.07, 0)	-0.02 (-0.06,0.02)
	5	-0.03 (-0.07, 0.01)	-0.03 (-0.07, 0.02)	-0.04 (-0.08, 0)	-0.02 (-0.06,0.03)
	25	0.02 (-0.03, 0.08)	0.01 (-0.04, 0.06)	-0.04 (-0.09, 0.02)	-0.02 (-0.07,0.04)
	50	0.04 (-0.03, 0.11)	0.01 (-0.06, 0.08)	-0.08 (-0.15, -0.01)	-0.03 (-0.1, 0.05)
	75	0.02 (-0.07, 0.12)	0.03 (-0.06, 0.13)	-0.06 (-0.16, 0.05)	-0.02 (-0.13,0.08)
Shigellosis	Any	-0.02 (-0.05, 0.02)	-0.02 (-0.05, 0.02)	-0.03 (-0.07, 0)	-0.02 (-0.06,0.02)
	5	-0.03 (-0.07, 0.01)	-0.03 (-0.07, 0.02)	-0.04 (-0.08, 0)	-0.02 (-0.06,0.03)
	25	0.02 (-0.03, 0.08)	0.01 (-0.04, 0.06)	-0.04 (-0.09, 0.02)	-0.02 (-0.07,0.04)
	50	0.04 (-0.03, 0.11)	0.01 (-0.06, 0.08)	-0.08 (-0.15, -0.01)	-0.03 (-0.1, 0.05)
	75	0.02 (-0.07, 0.12)	0.03 (-0.06, 0.13)	-0.06 (-0.16, 0.05)	-0.02 (-0.13,0.08)

Effect of exposure to cyclonic storms on cases when tropical cyclone exposure is defined as a minimum of 75-mm cumulative rainfall attributed to the storm.

Table 3.S. 2 Total number of storm weeks by population-and exposure-threshold

Storm exposure variable	Any pop. exposed	5% pop. exposed	25% pop. exposed	50% pop. exposed	75% pop. exposed
50-mm rainfall	4,453	3,734	2,502	1,751	1,100
75-mm rainfall	3,199	2,487	1,481	902	476
100-mm rainfall	2,210	1,580	865	547	213
Gale-force wind	1,437	1,070	586	284	163
Hurricane-force wind	215	110	6	0	0
500-km distance	9,108	8,461	7,167	6,107	5,123
250-km distance	5,219	4,584	3,400	2,577	1,737
150-km distance	3,758	3,151	2,105	1,304	769

Note: There are 20,442 weeks included in the analysis given 27 weeks in the storm season, 30 states and Washington, D.C. in the affected region, and 23 years of data; West Virginia did not report to the NNDSS in 1996 and 1997.

S1 Model. Description of the conditional quasi-Poisson model

In the conditional quasi-Poisson framework, the effect of storm exposure is first determined within matched strata created by matching on state and week. For week *i* and state *s* the within-strata model is:

$$\log\left(\mathrm{E}[Y_{is}]\right) = \alpha_0 + \alpha_{is} + \sum_{l=0}^{4} \beta_l \,\mathrm{storm}_{lis} + ns(\mathrm{year}_s) + \log\left(\mathrm{population}_{is}\right)$$

where α_{is} is the stratum-specific intercept, β_l the lag-specific coefficient (log rate ratio) for storm exposure, storm_{lis} the binary storm exposure variable, $ns(\mathrm{year}_s)$ the spline term for year with two degrees of freedom, and $\mathrm{population}_{is}$ the population for each state and year. Instead of estimating model parameters, this approach conditions them out by conditioning on the sum of cases for each week *i* in a multinomial model [253] such that:

$$(Y_{is} | \textstyle\sum_s Y_{is}) \sim \text{Multinomial}\ (\pi_s)$$

where:

$$\pi_s = \frac{e^{\boldsymbol{\beta^T x_s}}}{\sum_{j \in i} e^{\boldsymbol{\beta^T x_j}}}$$

The $\boldsymbol{\beta^T x_s}$ and $\boldsymbol{\beta^T x_j}$ terms describe row vectors of coefficients, $\boldsymbol{\beta}$, and variables, $\boldsymbol{x}$, from the quasi-Poisson model where *j* is the subset of *s* that includes the observations for each week *i*.

S1 Appendix. Supplemental methods and results

Supplementary Methods

Data sources

The storm and case data were the same as described in the main text. Drinking water data were extracted from the Safe Drinking Water Information System (SDWIS), which reports the drinking water source (groundwater or surface water) and population served for each community water system (CWS) in the US. The primary drinking water source for the county was determined by aggregating the CWSs within the county and calculating the proportion of the population served by groundwater and surface water; counties where ≥50% of the population was served by groundwater sources were categorized as groundwater counties and ≥50% served by surface water as surface water counties. Rurality categories were extracted from the US Department of Agriculture (USDA) Rural-Urban Continuum Codes [261], which categorizes metro areas by their population size and nonmetro areas by degree of urbanization and proximity to metro areas. Each county in the study area was classified as urban if it was in a metro category and rural if in a nonmetro category.

Storm exposure definition and statistical approach

The process for defining state-level exposure was the same as in the main analysis except the counties were stratified by drinking water source and rural or urban location. The population thresholds were any exposure, 5%, 25%, 50%, and 75% but only applied to counties that met the drinking water or location criteria. For example, a state was considered exposed to a storm if 25% of the population who lives in rural counties was exposed. We used the conditional quasi-Poisson statistical framework outlined in the main text and S1 Model.

Supplementary Results

The effect of storm exposure on cases when exposure was restricted to groundwater counties (A1 Table) or rural counties (A2 Table) was essentially the same as the main analysis. When storm exposure was defined by ≥75-mm rainfall, Cryptosporidiosis cases increased during the storm week at low population thresholds but had no significant effect in lagged weeks. Legionnaires' disease cases increased with lagged weeks and higher population exposure thresholds, and the effects were similar to those in the main analysis. At the 50% and 75% population exposure thresholds, the increase in *E. coli* cases 2 to 3 weeks post-storm was more pronounced in groundwater-reliant and rural areas. When exposure was restricted to surface water or rural counties, few states were considered exposed and the effect of storms on most cases was insignificant.

A1 Table. Effect of exposure to cyclonic storms on cases when exposure is restricted to counties with groundwater drinking water sources. Storm exposure is defined by ≥75-mm rainfall.

	Population expo. (%)	Week of storm	1 week post storm	2 weeks post storm	3 weeks post storm
Cryptosporidiosis	Any	0.42 (0.32, 0.51)	0.08 (-0.02, 0.19)	0.12 (0.02, 0.23)	0.13 (0.02, 0.24)
	5	0.52 (0.41, 0.62)	0.13 (0.01, 0.25)	0.2 (0.08, 0.31)	0.15 (0.03, 0.27)
	25	0.03 (-0.15, 0.2)	0.08 (-0.09, 0.25)	0.2 (0.03, 0.37)	0.16 (-0.01, 0.34)
	50	0.02 (-0.21, 0.26)	0.08 (-0.14, 0.31)	0.15 (-0.07, 0.38)	0.2 (-0.03, 0.43)
	75	-0.05 (-0.41, 0.3)	0.07 (-0.26, 0.39)	-0.01 (-0.34, 0.33)	0.12 (-0.23, 0.46)
Legionnaires' disease	Any	0.1 (0.01, 0.19)	0.14 (0.05, 0.23)	0.18 (0.09, 0.27)	0.19 (0.1, 0.28)
	5	0.13 (0.03, 0.24)	0.17 (0.07, 0.27)	0.18 (0.08, 0.28)	0.22 (0.12, 0.32)
	25	0.11 (-0.03, 0.24)	0.27 (0.15, 0.4)	0.35 (0.22, 0.47)	0.35 (0.22, 0.47)
	50	0.17 (0.01, 0.33)	0.28 (0.12, 0.44)	0.38 (0.23, 0.54)	0.41 (0.26, 0.56)
	75	0.23 (0, 0.46)	0.33 (0.1, 0.55)	0.38 (0.16, 0.61)	0.34 (0.12, 0.57)
E. coli	Any	-0.15 (-0.25, -0.04)	-0.05 (-0.15, 0.05)	-0.02 (-0.12, 0.08)	-0.01 (-0.11, 0.1)
	5	-0.1 (-0.22, 0.01)	0.02 (-0.1, 0.13)	0.06 (-0.05, 0.18)	0.05 (-0.07, 0.16)
	25	-0.08 (-0.23, 0.08)	0.12 (-0.03, 0.26)	0.04 (-0.11, 0.2)	0.09 (-0.06, 0.24)
	50	-0.1 (-0.31, 0.1)	0.28 (0.1, 0.45)	0.17 (-0.01, 0.36)	0.2 (0.01, 0.39)
	75	-0.03 (-0.3, 0.24)	0.5 (0.28, 0.72)	0.39 (0.15, 0.63)	0.27 (0.01, 0.53)
Salmonellosis	Any	-0.02 (-0.06, 0.02)	-0.02 (-0.06, 0.02)	-0.04 (-0.07, 0)	-0.02 (-0.06,0.01)
	5	-0.02 (-0.07, 0.02)	-0.03 (-0.07, 0.02)	-0.05 (-0.09, 0)	-0.02 (-0.06,0.03)
	25	0.03 (-0.03, 0.08)	0.02 (-0.03, 0.07)	-0.03 (-0.09, 0.02)	-0.01 (-0.06,0.05)
	50	0.04 (-0.03, 0.12)	0.02 (-0.05, 0.09)	-0.07 (-0.15, 0)	-0.03 (-0.11,0.05)
	75	0.05 (-0.06, 0.15)	0.06 (-0.04, 0.17)	-0.04 (-0.16, 0.07)	-0.02 (-0.13, 0.1)
Giardiasis	Any	0 (-0.06, 0.05)	-0.01 (-0.07, 0.04)	0 (-0.05, 0.06)	-0.03 (-0.09,0.03)
	5	-0.01 (-0.07, 0.06)	-0.01 (-0.07, 0.06)	0.02 (-0.05, 0.08)	-0.03 (-0.09,0.04)
	25	-0.02 (-0.1, 0.07)	-0.02 (-0.1, 0.07)	0.04 (-0.05, 0.12)	0.01 (-0.08, 0.09)
	50	-0.05 (-0.16, 0.06)	-0.04 (-0.15, 0.06)	0.01 (-0.1, 0.11)	0 (-0.1, 0.1)
	75	-0.02 (-0.17, 0.14)	-0.06 (-0.21, 0.1)	-0.03 (-0.19, 0.12)	0 (-0.15, 0.16)
Shigellosis	Any	-0.08 (-0.18, 0.02)	-0.09 (-0.18, 0.01)	0.01 (-0.09, 0.1)	0.02 (-0.08, 0.11)
	5	-0.11 (-0.22, 0.01)	-0.08 (-0.2, 0.03)	0.01 (-0.1, 0.12)	0.01 (-0.1, 0.12)
	25	-0.19 (-0.35, -0.03)	-0.11 (-0.26, 0.05)	-0.01 (-0.16, 0.14)	-0.03 (-0.19,0.12)
	50	-0.26 (-0.48, -0.04)	-0.07 (-0.28, 0.14)	0.02 (-0.18, 0.23)	-0.01 (-0.21,0.19)
	75	-0.35 (-0.67, -0.02)	-0.04 (-0.35, 0.27)	0.11 (-0.18, 0.41)	0.09 (-0.2, 0.39)

A2 Table. Effect of exposure to cyclonic storms on cases when exposure is restricted to rural counties. Storm exposure is defined by ≥75-mm rainfall.

	Population exposed (%)	Week of storm	1 week post storm	2 weeks post storm	3 weeks post storm
Cryptosporidiosis	Any	0.44 (0.34, 0.54)	0.1 (-0.01, 0.21)	0.16 (0.05, 0.26)	0.16 (0.05, 0.28)
	5	0.56 (0.46, 0.66)	0.17 (0.05, 0.29)	0.24 (0.12, 0.35)	0.18 (0.05, 0.3)
	25	0.03 (-0.15, 0.21)	0.09 (-0.09, 0.26)	0.22 (0.05, 0.39)	0.17 (-0.01, 0.35)
	50	0.06 (-0.18, 0.29)	0.16 (-0.06, 0.39)	0.24 (0.01, 0.46)	0.23 (0, 0.47)
	75	-0.42 (-0.89, 0.05)	-0.31 (-0.73, 0.11)	-0.41 (-0.88, 0.05)	-0.22 (-0.67, 0.24)
Legionnaires' disease	Any	0.1 (0.01, 0.2)	0.14 (0.05, 0.23)	0.17 (0.08, 0.26)	0.2 (0.1, 0.29)
	5	0.15 (0.04, 0.25)	0.18 (0.07, 0.28)	0.19 (0.09, 0.29)	0.23 (0.13, 0.34)
	25	0.13 (0, 0.26)	0.3 (0.18, 0.43)	0.37 (0.24, 0.49)	0.34 (0.22, 0.47)
	50	0.21 (0.05, 0.37)	0.27 (0.11, 0.44)	0.42 (0.27, 0.57)	0.4 (0.25, 0.56)
	75	0.28 (0.05, 0.51)	0.3 (0.07, 0.54)	0.41 (0.18, 0.63)	0.45 (0.23, 0.67)
E. coli	Any	-0.13 (-0.24, -0.03)	-0.03 (-0.14, 0.07)	0.01 (-0.09, 0.11)	0.02 (-0.09, 0.12)
	5	-0.1 (-0.22, 0.02)	0.02 (-0.1, 0.14)	0.07 (-0.04, 0.19)	0.07 (-0.05, 0.19)
	25	-0.08 (-0.24, 0.08)	0.13 (-0.02, 0.28)	0.06 (-0.09, 0.22)	0.11 (-0.05, 0.26)
	50	-0.11 (-0.32, 0.1)	0.33 (0.15, 0.51)	0.22 (0.03, 0.41)	0.11 (-0.1, 0.31)
	75	0 (-0.3, 0.3)	0.51 (0.26, 0.77)	0.49 (0.22, 0.75)	0.23 (-0.07, 0.53)
Salmonellosis	Any	-0.02 (-0.05, 0.02)	-0.02 (-0.05, 0.02)	-0.03 (-0.07, 0)	-0.02 (-0.06,0.01)
	5	-0.02 (-0.06, 0.03)	-0.03 (-0.07, 0.01)	-0.06 (-0.11, -0.02)	-0.02 (-0.06,0.03)
	25	0.03 (-0.02, 0.09)	0.01 (-0.05, 0.07)	-0.03 (-0.09, 0.02)	0 (-0.06, 0.06)
	50	0.03 (-0.05, 0.1)	-0.02 (-0.1, 0.06)	-0.09 (-0.17, 0)	-0.02 (-0.1, 0.06)
	75	0.05 (-0.08, 0.17)	-0.02 (-0.15, 0.1)	-0.1 (-0.24, 0.03)	-0.01 (-0.14,0.12)

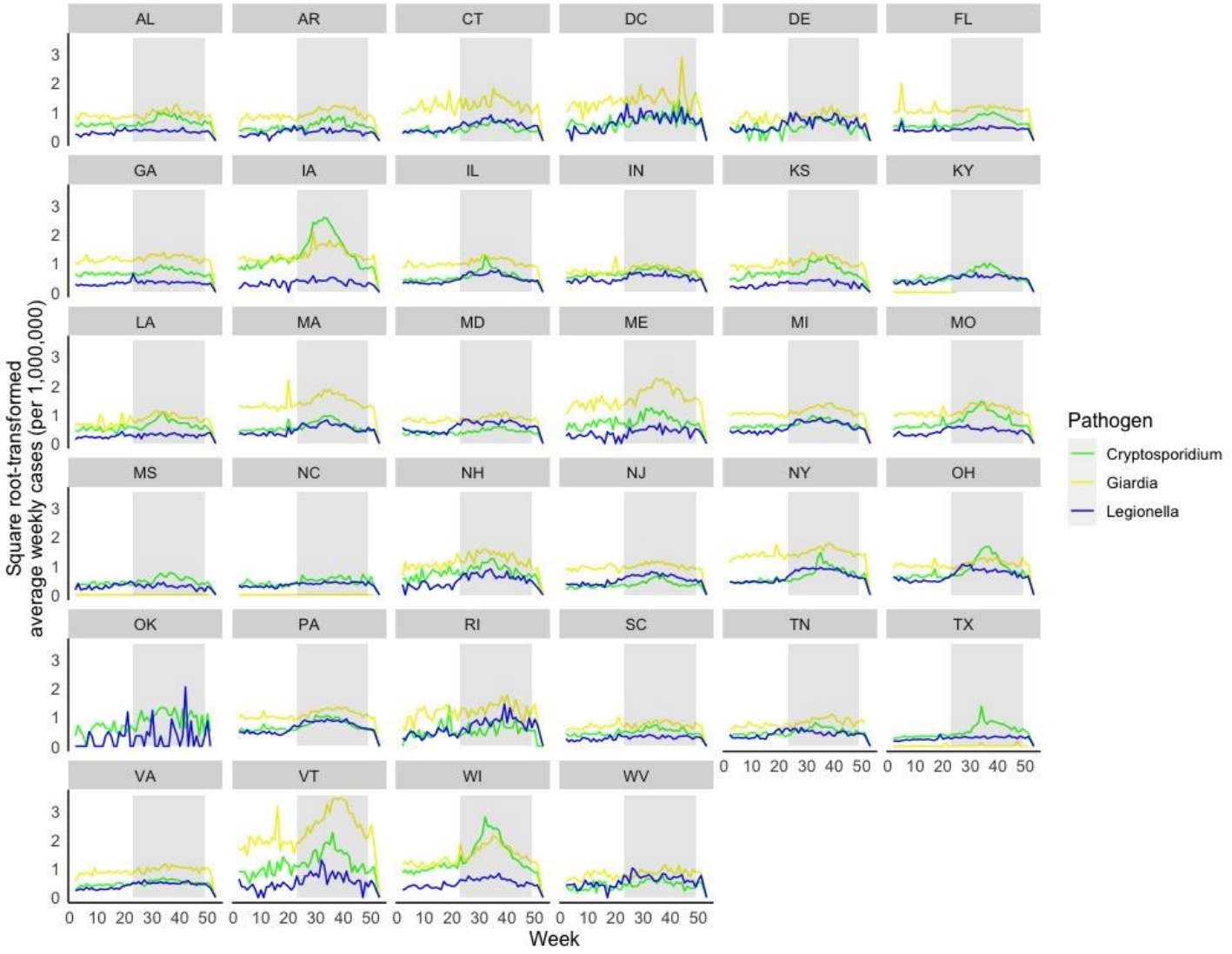

Figure 3.S. 1: Average weekly cases of Cryptosporidiosis, Giardiasis, and Legionnaires' disease by state.

Average weekly cases (square-root transformed) per 1,000,000 people by state for Cryptosporidiosis (green), Giardiasis (yellow), and Legionnaires' disease (blue) reported to NNDSS between 1996 and 2018; not all infections were reported for the entire study period (Table 3.1). The shaded region represents the weeks encompassed in the Atlantic storm season (June 1 to November 30).

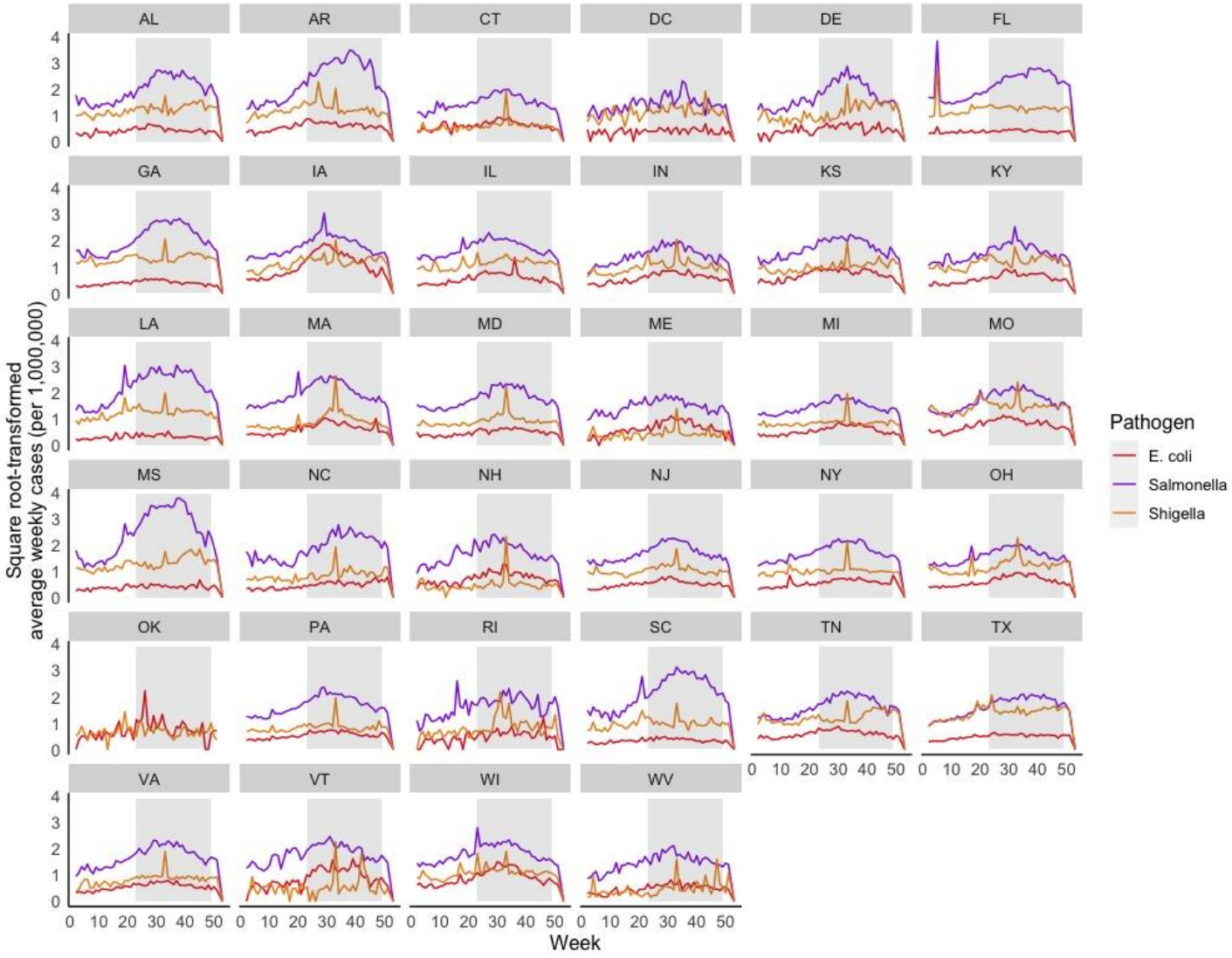

Figure 3.S. 2: Average weekly cases for E. coli, Salmonellosis, and Shigellosis by state. Average weekly cases (square-root transformed) per 1,000,000 people by state for *E. coli* infections (red), Salmonellosis (purple), and Shigellosis (orange) reported to NNDSS between 1996 and 2018; not all infections were reported for the entire study period (Table 3.1). The shaded region represents the weeks encompassed in the Atlantic storm season (June 1 to November 30).

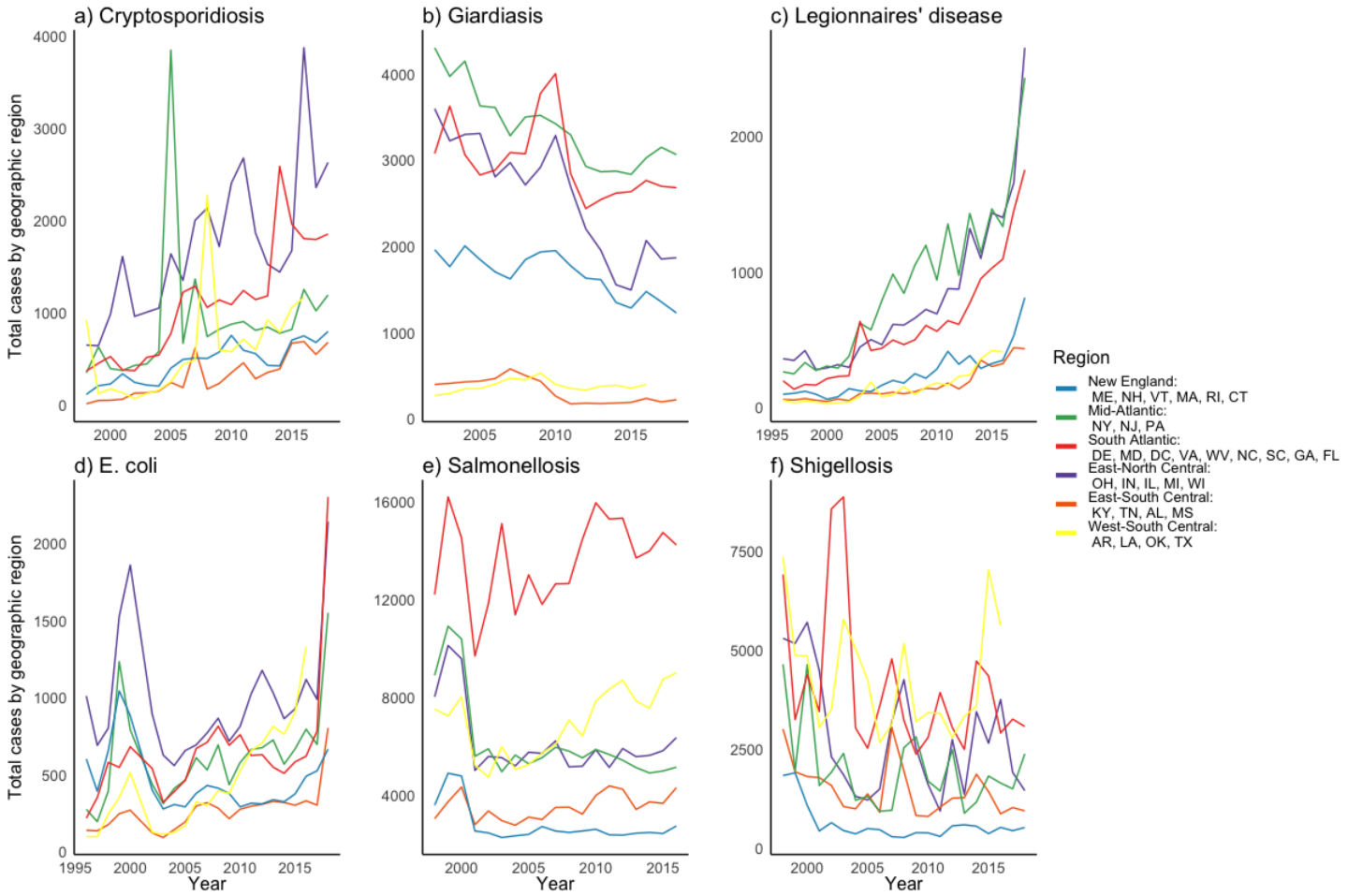

Figure 3.S. 3: Total weekly cases by geographic region and infection.
Total weekly cases by geographic region (colors) and infectious disease (a-f) reported to NNDSS between 1996 and 2018; not all infections were reported for the entire study (Table 3.1).

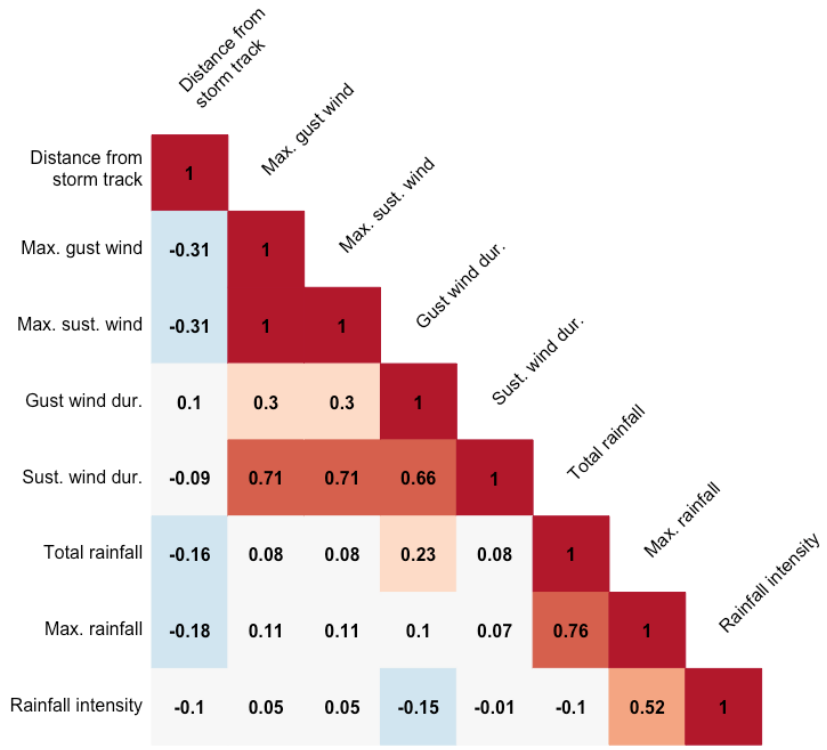

Figure 3.S. 4: Correlation matrix for storm characteristics.

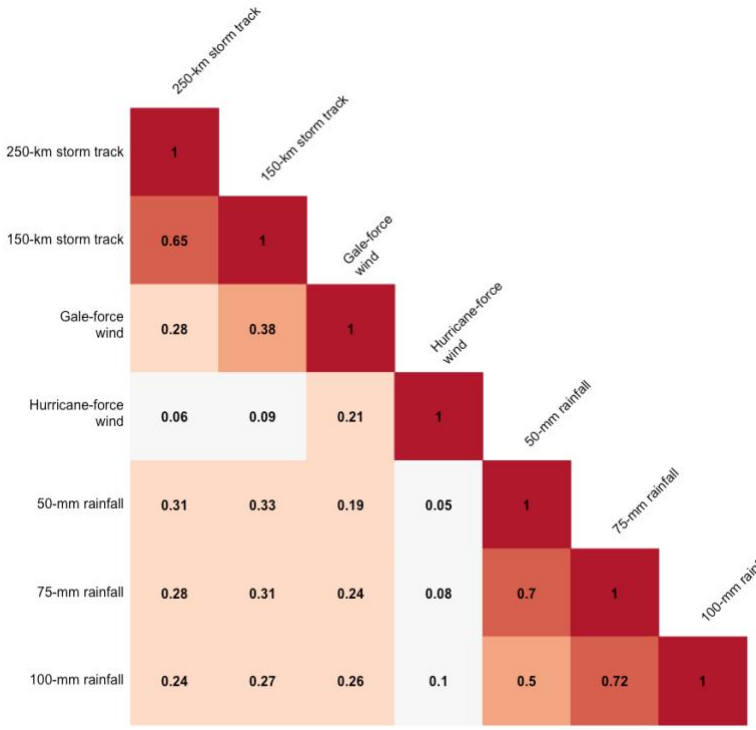

Figure 3.S. 5: Correlation matrix for storm exposure thresholds.

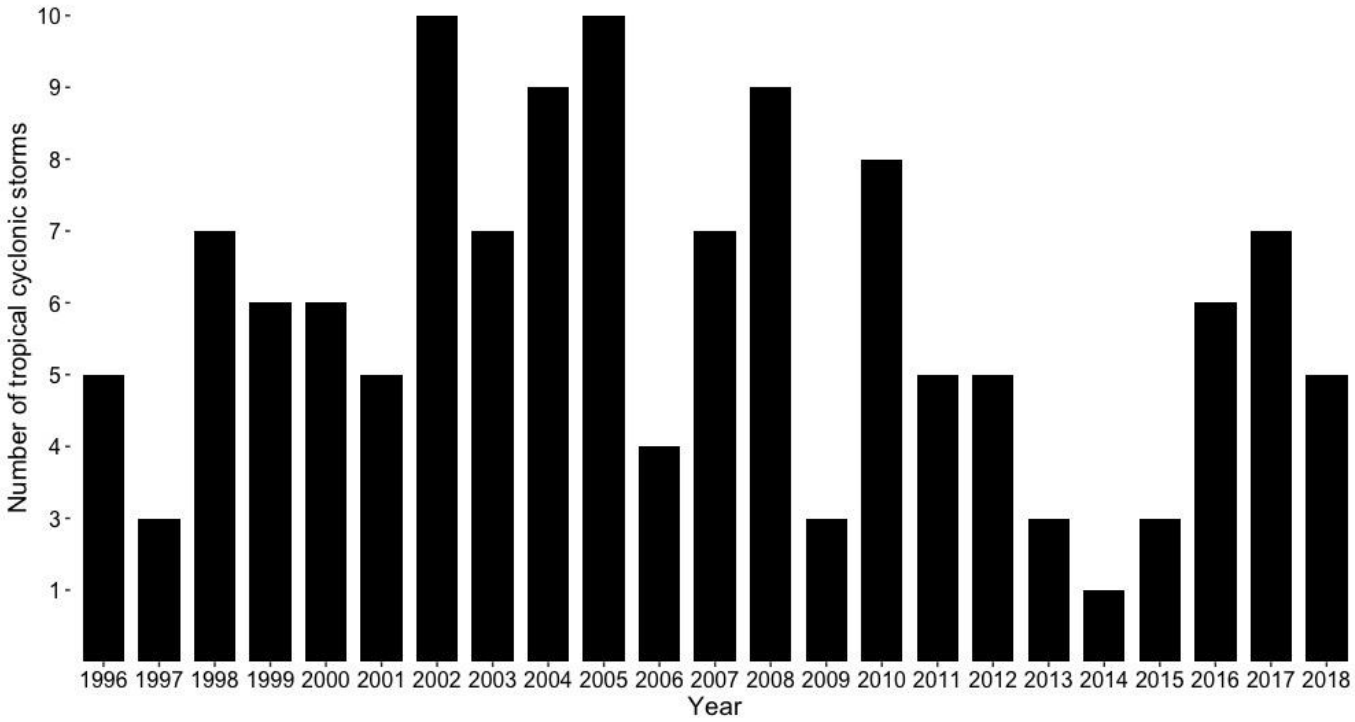

Figure 3.S. 6: Number of tropical cyclones that made landfall in the US between 1996 and 2018.

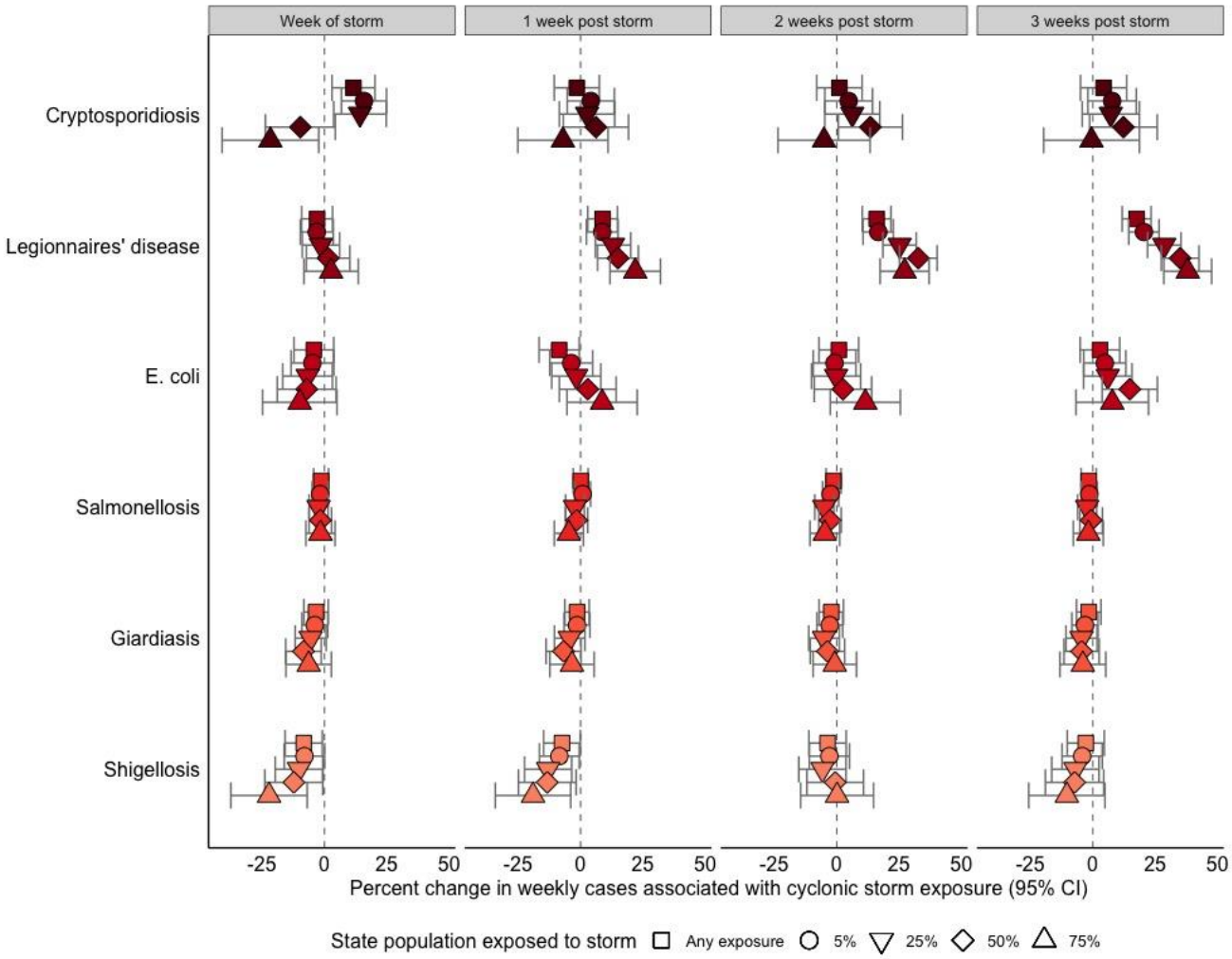

Figure 3.S. 7: Change in case rates associated with distance from storm track.
Average percent change in weekly case rates associated with exposure to tropical cyclones where exposure is defined as being within 250-km of the storm track. The estimates and Bonferroni-corrected 95% confidence intervals are reported for each infectious disease (shade) and population-exposure threshold (shape); estimates are reported for week of the storm and 1 to 3 weeks post-storm.

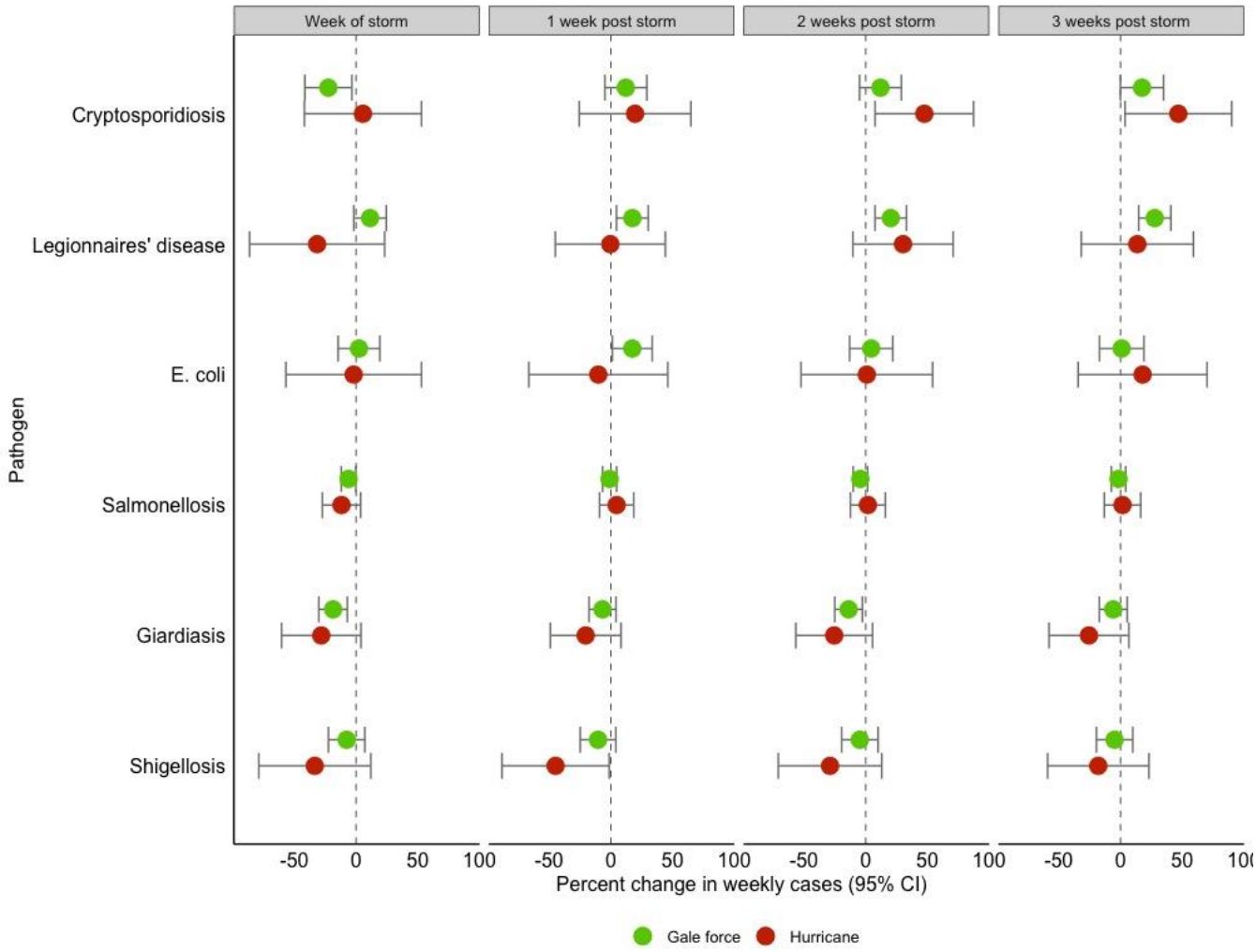

Figure 3.S. 8: Change in case rates associated with windspeed.
Average percent change in weekly case rates associated with exposure to tropical cyclones where exposure is defined by sustained gust gale-force windspeed (≥ 34 knots, green) or hurricane-force windspeed (≥ 64 knots, red). The estimates and Bonferroni-corrected 95% confidence intervals are reported for each infectious disease 5% state population exposed threshold; estimates are reported for week of the storm and 1 to 3 weeks post-storm.

Chapter 4

The effect of seasonal and extreme floods on hospitalizations for Legionnaires' disease in the United States, 2000 – 2011

Victoria D. Lynch[1] and Jeffrey Shaman[1]

Affiliations:
[1]Department of Environmental Health Sciences, Columbia Mailman School of Public Health, Columbia University, New York, NY, United States of America
*Corresponding author: Victoria Lynch, vdl2103@cumc.columbia.edu, 722 W. 168th St. New York, NY 10032

Published:
Lynch VD, Shaman J. The effect of seasonal and extreme floods on hospitalizations for Legionnaires' disease in the United States, 2000 – 2011. BMC Infect Dis. 2022; 22(1):550.

Abstract

An increasing severity of extreme storms and more intense seasonal flooding are projected consequences of climate change in the United States. In addition to the immediate destruction caused by storm surges and catastrophic flooding, these events may also increase the risk of infectious disease transmission. We aimed to determine the association between extreme and seasonal floods and hospitalizations for Legionnaires' disease in 25 US states during 2000-2011. We used a nonparametric bootstrap approach to examine the association between Legionnaires' disease hospitalizations and extreme floods, defined by multiple hydrometeorological variables. We also assessed the effect of extreme flooding associated with named cyclonic storms on hospitalizations in a generalized linear mixed model (GLMM) framework. To quantify the effect of seasonal floods, we used multi-model inference to identify the most highly weighted flood-indicator variables and evaluated their effects on hospitalizations in a GLMM. We found a 32% increase in monthly hospitalizations at sites that experienced cyclonic storms, compared to sites in months without storms. Hospitalizations in months with extreme precipitation were in the 89th percentile of the bootstrapped distribution of monthly hospitalizations. Soil moisture and precipitation were the most highly weighted variables identified by multi-model inference and were included in the final model. A 1-standard deviation (SD) increase in average monthly soil moisture was associated with a 49% increase in hospitalizations; in the same model, a 1-SD increase in precipitation was associated with a 26% increase in hospitalizations. This analysis is the first to examine the effects of flooding on hospitalizations for Legionnaires' disease in the United States using a range of flood-indicator variables and flood definitions. We found evidence that extreme and seasonal flooding is associated with increased hospitalizations; further research

is required to mechanistically establish whether floodwaters contaminated with *Legionella* bacteria drive transmission.

4.1 Background

Legionnaires' disease is among the most severe and costly waterborne illnesses in the United States, where it is responsible for an estimated 15% of all deaths related to waterborne infectious disease [1] and between 3% and 9% of all cases of community-acquired pneumonia [97, 205]. Legionnaires' disease was so named in 1977, when a cooling tower contaminated with the bacteria was found to be the cause of a pneumonia outbreak among guests at a hotel [262], and its incidence has substantially increased since 2000 [263, 264]. Outbreaks of Legionnaires' disease have decreased over the last 40 years, however, with only 4% of reported cases since 2000 linked to a cluster [206]. Over 80% of cases are sporadic [89], and the source of infection is never identified for the majority of these cases. Legionnaires' disease cases typically peak in late summer or early fall, and this consistent seasonality suggests that environmental factors affect transmission [265, 266] and may help explain the origin of these sporadic infections.

Environmental conditions affect the proliferation of *Legionella* bacteria in lakes, streams, and estuaries [267, 268], and the contamination events that may lead to disease transmission [113, 214]. The bacteria are abundant in aqueous environments [269] and survive by parasitizing amoebae, including many that persist in environmental biofilms [270, 271]. The bacteria optimally grow in wet, warm conditions (between 25°C and 42°C) and flourish in sessile biofilm communities [245, 272] in the natural and built environment [273, 274]. Environmental events that mobilize biofilms may be an important driver of infection by increasing the bacterial load in plumbed water [275-277], water used for industrial processes [278, 279], and surface water where direct exposure can occur [280]. Susceptible individuals can become infected by inhaling aerosolized bacteria from these contaminated water sources.

Previous studies have found positive associations between cases of Legionnaires' disease and rainfall [100-102, 281], relative humidity [266, 282], and streamflow [103], and inconsistent associations with proximity to rivers or river height [103, 267]. Temperature has been positively associated with cases in several studies, though its effect is often attenuated when adjusting for other seasonal factors [101, 266]. While many of these hydrometeorological variables are associated with flood events, the relationship between Legionnaires' disease and flooding has not been formally evaluated. Flooding is known to mobilize bacteria-rich biofilms in water bodies [239, 259], which may lead to increased bacterial colonization of the built environment. Churning flood waters may also lead to the direct aerosolization of bacteria and increased risk of exposure for individuals close to flood waters.

Flooding during extreme storms may be of particular concern because high winds and storm surges can damage or overwhelm the water treatment infrastructure necessary to address contamination events [8, 127]. The effect of extreme floods on waterborne infectious diseases has not been systematically examined in the US; rather, it has only been assessed after specific storm events (e.g. Superstorm Sandy [283, 284], Hurricane Katrina [117]). Increased incidence of intestinal illness has been reported after major storms, however, and post-storm microbiological analyses have found high concentrations of pathogenic bacteria in floodwater [8, 285].

Floods can be measured with a range of hydrometeorological variables and those that best describe extreme or seasonal events often vary by region to reflect local hydroclimatology, geography, and the built environment [3, 107]. These factors determine the conditions under which a flood occurs and help explain, for example, how a single heavy precipitation event can lead to a devastating flash flood in an urban area with a small watershed, whereas the same

amount of precipitation has no effect in a rural area with a large drainage basin [2]. Precipitation has traditionally been the primary variable used to determine flood magnitude; however, recent research has demonstrated that soil moisture, snowmelt, and precipitation excess might better characterize flooding in many regions [232]. Most studies that have examined the association between floods and health outcomes have used a single hydrologic indicator [64, 65] or observed storms records [237]. Given that floods cannot be defined by the same set of hydrometeorological variables across all locations, this approach does not allow for the identification of all major flood types in the US (i.e. river, coastal, and flash floods as well as flooding after cyclonic storms).

Understanding the association between Legionnaires' disease infections and flood events is particularly important given that the severity of flooding is predicted to increase in conjunction with rising temperatures [123, 286]. The severity and timing of river floods is projected to increase due to earlier snowmelt and more intense precipitation [287, 288]. The number of major, billion-dollar floods has increased by 5% each year in the US since 1980 [289]; this is a trend that is likely to continue under future global warming, as more severe cyclonic storms and coastal flood events are projected to occur in the coming decades [224].

In this study, we used nonparametric and generalized linear mixed models to determine the effect of extreme and seasonal floods on hospitalizations for Legionnaires' disease across the US. Previous research has examined the association between single hydrometeorological variables and cases, but a thorough examination of the effect of flooding on Legionnaires' disease has not been conducted. Earlier studies have also been limited to small geographic regions, primarily in the northeastern US, whereas this study includes hospitalizations from 25 states throughout the US. Using this national dataset, we have quantified the effects of extreme

and seasonal floods, measured using multiple flood-indicator variables, on hospitalizations for Legionnaires' disease across the US.

4.2 Methods

4.2.1 Data

Hospitalization data

Legionnaires' disease infections occur primarily among older or immunocompromised individuals, and an estimated 97% of identified cases are hospitalized [1, 272]. We used the National Inpatient Sample (NIS) from the Healthcare Cost and Utilization Project (HCUP) to identify Legionnaires' disease hospitalizations between 2000 and 2011 throughout the US. The NIS is the largest publicly available all-payer inpatient database in the US; it captures 20% of hospitalizations per year and is designed to be representative of all hospitalizations nationwide. We identified infections by ICD-9 code (482.84) and found the monthly Legionnaires' disease hospitalization count for each hospital. We restricted our analysis to hospitals that contributed at least 4 years of data to the NIS dataset, provided monthly counts of hospitalizations, and reported their geographic location.

Hospitals that reported no Legionnaires' disease cases were excluded from the analysis because the absence of cases could indicate that *Legionella* were not present in environmental or household water sources in that region, or because a hospital was not testing for Legionnaires' disease among patients with pneumonia. Many hospitals reported only one case during the study period; as a sensitivity analysis, we repeated the analyses using several case count thresholds to further restrict the included hospitals. We created subsets of our hospitalization data containing hospitals with at least 1, 5, 10, 15, and 20 Legionnaires' disease cases during the study period; all of the analyses were repeated with these case count threshold datasets.

The NIS includes the location of the reporting hospital, but not the cases' residential locations. To address the possibility of misclassification bias, given that the flood data are associated with the location of the hospital, we matched the hospitals to Hospital Service Areas (HSA) provided by the Dartmouth Atlas of Healthcare [226]. The HSA is the catchment area for each hospital and includes the zip codes where most Medicare patients receive care from the given hospital. We repeated the analyses using flood data associated with catchment area, instead of the hospital location, as a sensitivity analysis to assess the consistency of our findings.

Flood data

Flooding can be characterized by several hydrometeorological variables, and we used multiple flood-indicator variables to account for the range of flood-types found across the study sites (e.g. river floods, coastal floods, flash floods), and to distinguish between extreme and seasonal events. Precipitation, soil moisture, and surface runoff data were obtained from the NASA/ NOAA North American Land Data Assimilation System 2 (NLDAS-2) forcing dataset and were aggregated from an hourly temporal resolution to mean monthly values for each hospital location [290]. We used the United States Geological Survey (USGS) National Water Information System to find the stream gages closest to each hospital, for those that had a stream gage in the same zip code, and obtained daily median and maximum stream discharge measurements, which were aggregated to monthly means [291].

Data on flooding associated with tropical cyclones were obtained from the NOAA Storm Event Database, which tracks the location, type, and severity of named storms in the Atlantic Storm Basin [292, 293]. For each named storm that occurred during the study period, we extracted county-level data on: 1) storm-related precipitation, 2) reported flooding, and 3)

distance from the storm track. Exposure to each of these extreme flood-related indicators was assessed for each hospital and month in the study period.

4.2.2 Statistical analysis

Extreme floods associated with cyclonic storms

Two methods for identifying extreme floods were used to account for the range of flood types that occur in the US. In the first approach, we defined extreme floods as those associated with named cyclonic storms, and restricted the dataset to the hospitals that experienced these storms and to the months of the Atlantic Basin storm season (June – November).

We modeled the association between Legionnaires' disease hospitalizations and extreme storm-related floods using a negative binomial generalized linear mixed model (GLMM) framework to account for the over-dispersed hospitalization data. The counties with HCUP-contributing hospitals were categorized as exposed or unexposed to storms for each month during the storm season between 2000 and 2011. A county was considered exposed if it was within 150km of the storm track and unexposed if it was outside of that range. In addition to the binary exposure variable, we assessed storm-related precipitation and proximity to the storm track as continuous variables and as categorical variables grouped by quartile.

The model included a binary location variable to assess differences between rural and urban hospitals and hospital-specific monthly discharges as an offset to obtain the rate of Legionnaires' disease hospitalizations. We also included hospital-specific random intercepts nested within state-specific random intercepts to account for underlying differences in hospitalization policies (e.g. testing, reporting, and admitting practices) as well as state-level responses to extreme events (Model S1, Additional file 1). The storm-related variables were modeled separately and jointly, and model fit was assessed using the Akaike Information

Criterion (AIC). To assess the consistency of our findings, this analysis was repeated for each Legionnaires' disease case threshold to determine whether a storm in the preceding month was associated with hospitalizations.

Extreme floods associated with anomalous hydrometeorology

In the second analysis, we classified months with anomalously high precipitation, soil moisture, surface runoff, or streamflow discharge as those with extreme flooding. For each hospital, we found the months with mean hydrometeorological variables above the 95th percentile and averaged the number of Legionnaires' disease hospitalizations in this "extreme group". We compared the hospitalizations in the extreme group to a bootstrapped distribution of monthly Legionnaires' disease hospitalizations.

The bootstrap generated a sampling distribution by randomly selecting 5% of months in the time series, with replacement, averaging the number of hospitalizations in those months, and then repeating the process 10,000 times. To control for seasonality, the sample was selected from the same range of months as those included in the extreme group for each hospital (i.e. if the extreme group for a given hospital did not include hospitalizations for November, then other November months in that hospital's time series were not selected during the bootstrapping process). The probability of the Legionnaires' disease hospitalizations in the extreme group was determined by comparison to the empirical cumulative distribution generated by the bootstrap. The bootstrap process was repeated for the all of the hydrometeorological flood indicators and for each case threshold.

Seasonal floods

In a third analysis, we used a multimodel inference approach to determine the effect of seasonal flood indicators on Legionnaires' disease hospitalizations for the whole time series, not

restricted to months with extreme floods or during the Atlantic Basin hurricane season. Multimodel inference was conducted on candidate models that varied only in the explanatory hydrometeorological variables, but that otherwise had the same structure. All combinations of standardized precipitation, soil moisture, surface runoff, and observed flood count were included in the candidate models; temperature was also included, given that the growth of *Legionella* has been associated with temperature seasonality. The models also included terms to control for seasonal and secular trends and a random intercept for each hospital (Model S2, Additional file 1). The streamflow variables were excluded from this analysis due to missing data for hospitals that were not near USGS stream gages.

We used the log likelihood and number of parameters to calculate the Akaike weight for each model. The models were ranked by weight, and the top models, the smallest number of models whose weights added to 0.90, were selected as the best-fitting models. Among the top models, variable weight importance for the hydrometeorological and temperature variables was determined. Cross-validation was performed by removing 20% of the data and conducting multimodel inference on the remainder; this process was iterated 1,000 times to evaluate the consistency of the weights and effect estimates, and to compare them to the top full models. These analyses were repeated for each Legionnaires' disease case threshold.

4.3 Results

There were 1,376 Legionnaires' disease hospitalizations between 2000 and 2011 at the 75 hospitals that met our inclusion criteria for the primary analysis (Figure 4.1a). Most of these hospitals were large facilities (65.4%) and located in urban areas in the Northeast (66.2%) or Midwest (16.8%). The number, size, and geographic breakdown of the hospitals was relatively consistent across years in the study period, with the exception of 2008 when there were no rural

hospitals in the dataset (Table 4.1). The rural/urban location and hospital bed-size variables were not included in the 2011 HCUP dataset, but the mean annual discharge and geographic region breakdown for this year are consistent with previous years.

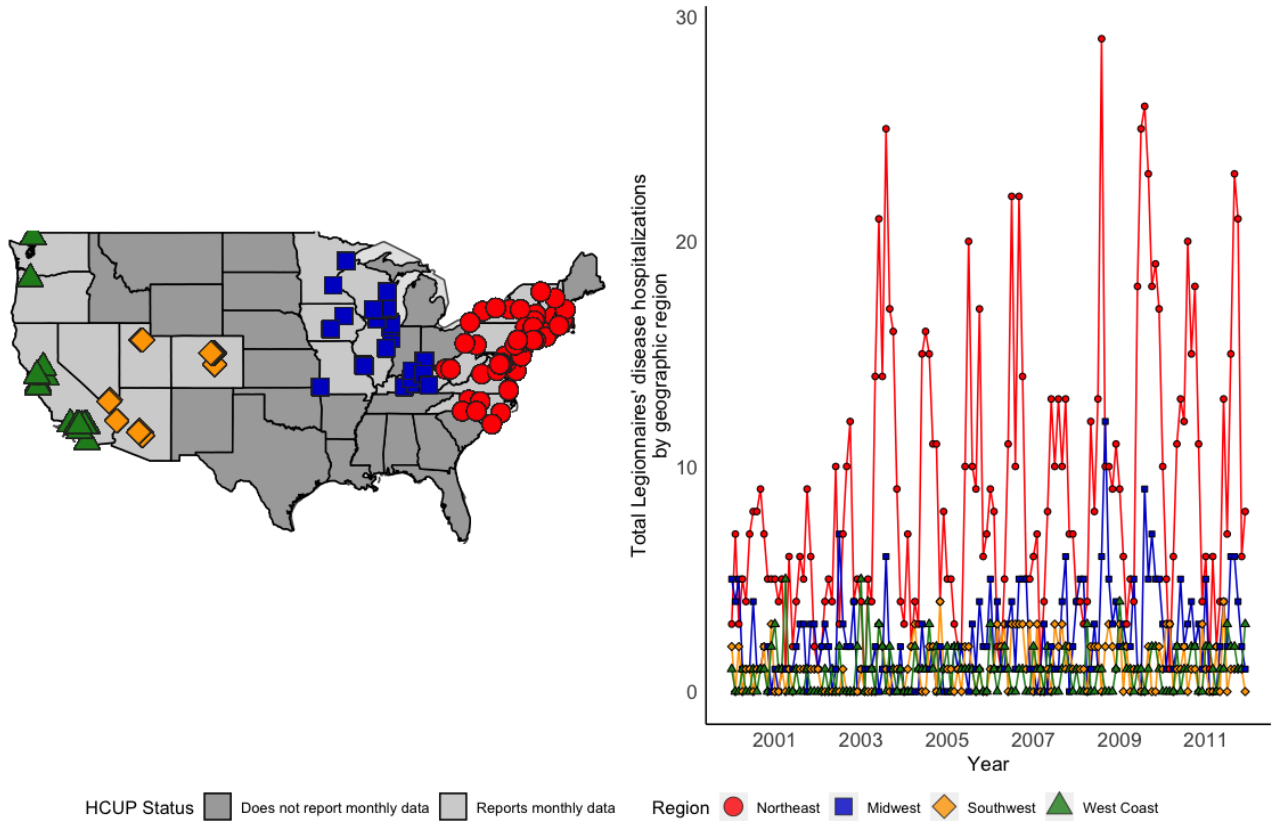

Figure 4. 1: Time series of Legionnaires' disease between 2000 and 2011.
a) The 75 hospitals in the HCUP dataset with a minimum of 10 total Legionnaires' disease cases; dark gray states are those that do not participate in HCUP or do not provide monthly data. b) Total Legionnaires' disease hospitalizations among the included hospitals between 2000 and 2011 by geographic region.

Table 4. 1 Description of hospitals from the HCUP dataset included in the primary analysis, 2000 – 2011

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Overall[1]
Number of Hospitals	36	24	36	34	29	30	37	29	38	32	33	31	
Number of Cases	92	55	83	133	90	88	139	100	149	182	141	124	1376
Hospital Location (%)													
Rural	11.1	8.3	11.1	5.9	3.4	3.3	8.1	3.4	0	6.2	9.1	-	6.4
Urban	88.9	91.7	88.9	94.1	96.6	96.7	91.9	96.6	100	93.8	90.9	-	93.6
Hospital Bedsize (%)													
Small	8.3	8.3	16.7	11.8	17.2	10	5.4	3.4	5.3	12.5	9.1	-	9.8
Medium	33.3	29.2	22.2	35.3	13.8	20	21.6	24.1	34.2	18.8	18.2	-	24.9
Large	58.4	62.5	61.1	52.9	69	70	73	72.5	60.5	68.7	72.7	-	65.4
Geographic Region (%)													
Northeast	63.9	70.8	61.1	73.5	72.4	66.7	62.2	69	57.9	71.9	63.6	64.5	66.2
Midwest	25	16.7	25	8.8	6.9	13.3	16.2	13.8	21.1	18.8	15.2	19.4	16.8
Southwest	8.3	8.3	8.3	11.8	17.2	13.3	16.2	13.8	18.4	6.2	12.2	9.7	12.3
West Coast	2.8	4.2	5.6	5.9	3.5	6.7	5.4	3.4	2.6	3.1	9	6.5	4.7
Mean Annual Discharge (SD)	20,600 (12,600)	22,600 (12,800)	19,700 (9,620)	21,400 (14,300)	24,900 (13,900)	23,700 (11,000)	24,000 (15,800)	27,600 (13,500)	25,200 (11,200)	25,500 (14,400)	21,900 (13,200)	24,900 (11,200)	23,300 (13,100)
Number of Hospitals	58.4	62.5	61.1	52.9	69	70	73	72.5	60.5	68.7	72.7	-	65.4

[1]75 hospitals were included in the primary analysis, each of which contributed at least 4 years of data; the number per year refers to the number, out of the 75, that contribute in that given year.

Seasonality and secular trends in hospitalizations varied by geographic region (Figure 4.1). In the Northeast and Midwest, hospitalizations increased between July and October, peaking in August, and also increased over time (Figure 4.1b). In the Southwest, hospitalizations exhibited an attenuated seasonality, with increased hospitalizations between March and October, and fluctuated over time. There was no clear seasonal or secular trend in hospitalizations among the hospitals in western states.

The hospital characteristics varied considerably across the datasets with different case count thresholds used in the secondary analysis. The subset of hospitals with at least one Legionnaires' disease case included 378 hospitals, with many located on the West Coast (17.6%) and in rural areas (23.1%) (Table S1, Additional file 1). At higher case count thresholds, the included hospitals on average had larger bed capacity and were concentrated in urban areas in the Northeast; among the 15-case and 20-case threshold hospitals, none were from rural areas or located on the West Coast (Table S1, Additional file 1). Seasonal and secular trends were consistent across the different case count thresholds.

Fifteen named hurricanes or tropical cyclones affected counties with hospitals included in the dataset (Table 4.S.2, Additional file 1). Among the hospitals that experienced these storms, there was a significant increase in Legionnaires' disease hospitalizations during months with a storm compared to months during the Atlantic storm season when a storm did not occur (Figure 4.2). There was a 32% increase in monthly Legionnaires' disease hospitalizations among hospitals that experienced a cyclonic storm compared to those that did not. This association was consistent across the case count thresholds, though it was insignificant in the 1-case and 5-case subsets and stronger in the 15-case and 20-case subsets, where there was a 46% and 54%

increase in hospitalizations, respectively, in months with a cyclonic storm compared to those without storms (Figure 4.2).

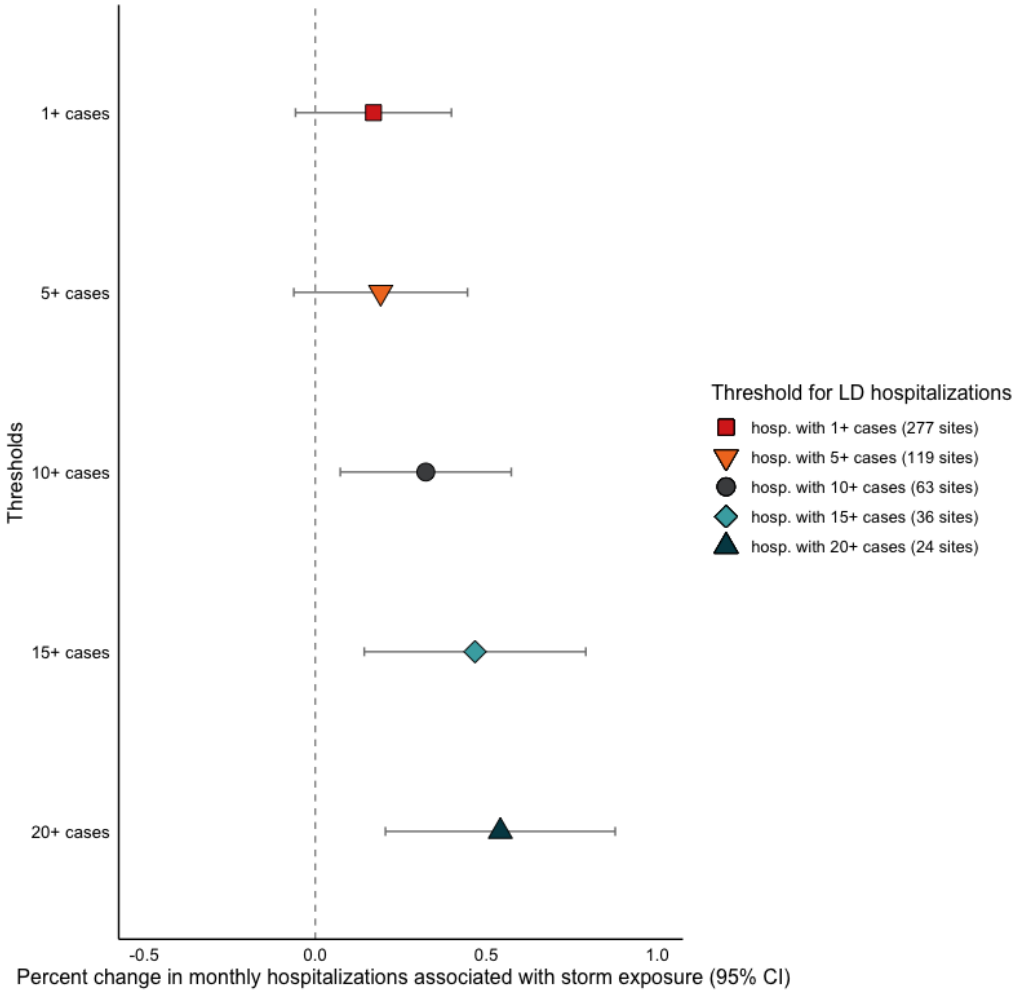

Figure 4. 2: Legionnaires' disease hospitalizations and cyclonic storm exposure.
Change in monthly Legionnaires' disease hospitalizations among hospitals that experienced a cyclonic storm in the same month compared to hospitals that did not experience a storm; analysis was restricted hospitals in regions that experience cyclonic storms from the Atlantic storm basin and to the months of the Atlantic storm season (June – November). Symbols represent the effect estimates from models using the different Legionnaires' disease case count thresholds.

The intensity of storm-related precipitation and proximity to the storm track were not significantly associated with Legionnaires' disease hospitalizations (Figure 4.S.1, Additional file

1). Hospitals that experienced the most intense storm-related precipitation (quartile 4 of maximum rainfall) had an increase in hospitalizations compared to hospitals that did not experience storm-related precipitation, but the difference was insignificant (Figure 4.S.1a, Additional file 1). Among the hospitals in the 15-case and 20-case subsets, however, this association was significant; hospitalizations increased by 81% and 90%, respectively, with moderate storm-related precipitation (quartile 3 of maximum rainfall) (Figure 4.S.1a, Additional file 1). Proximity to the storm track was not associated with hospitalizations for any of the case count thresholds (Figure 4.S.1b, Additional file 1). The sensitivity analysis with storm data aggregated to each hospital's catchment area yielded results consistent with the primary analysis. There was a 50% increase in monthly Legionnaires' disease hospitalizations among hospitals in HSAs that experienced a cyclonic storm compared to those that did not (Figure 4.S.2a, Additional file 1), and no significant association with precipitation intensity (Figure 4.S.2b, Additional file 1) or proximity to the storm track (Figure 4.S.2c, Additional file 1).

The average number of Legionnaires' disease hospitalizations in months with extreme precipitation was in the 89th percentile of the bootstrapped distribution (Figure 4.3), which was substantially higher than the average number of hospitalizations for all other causes in the same months. The strength of this association increased among the 15-case and 20-case threshold subsets to the 92nd and 94th percentiles, respectively (Table 4.2). Across all case-count thresholds, Legionnaires' disease hospitalizations in months with extreme runoff, soil moisture, or temperature did not significantly vary from the bootstrapped averages (Table 4.2). These findings are supported by the sensitivity analysis using meteorological data aggregated to the hospitals' catchment areas; the average number of hospitalizations in months with extreme

precipitation was in the 84th percentile and increased to the 91st and 93rd percentiles among the higher case thresholds (Table 4.S.3, Additional file 1).

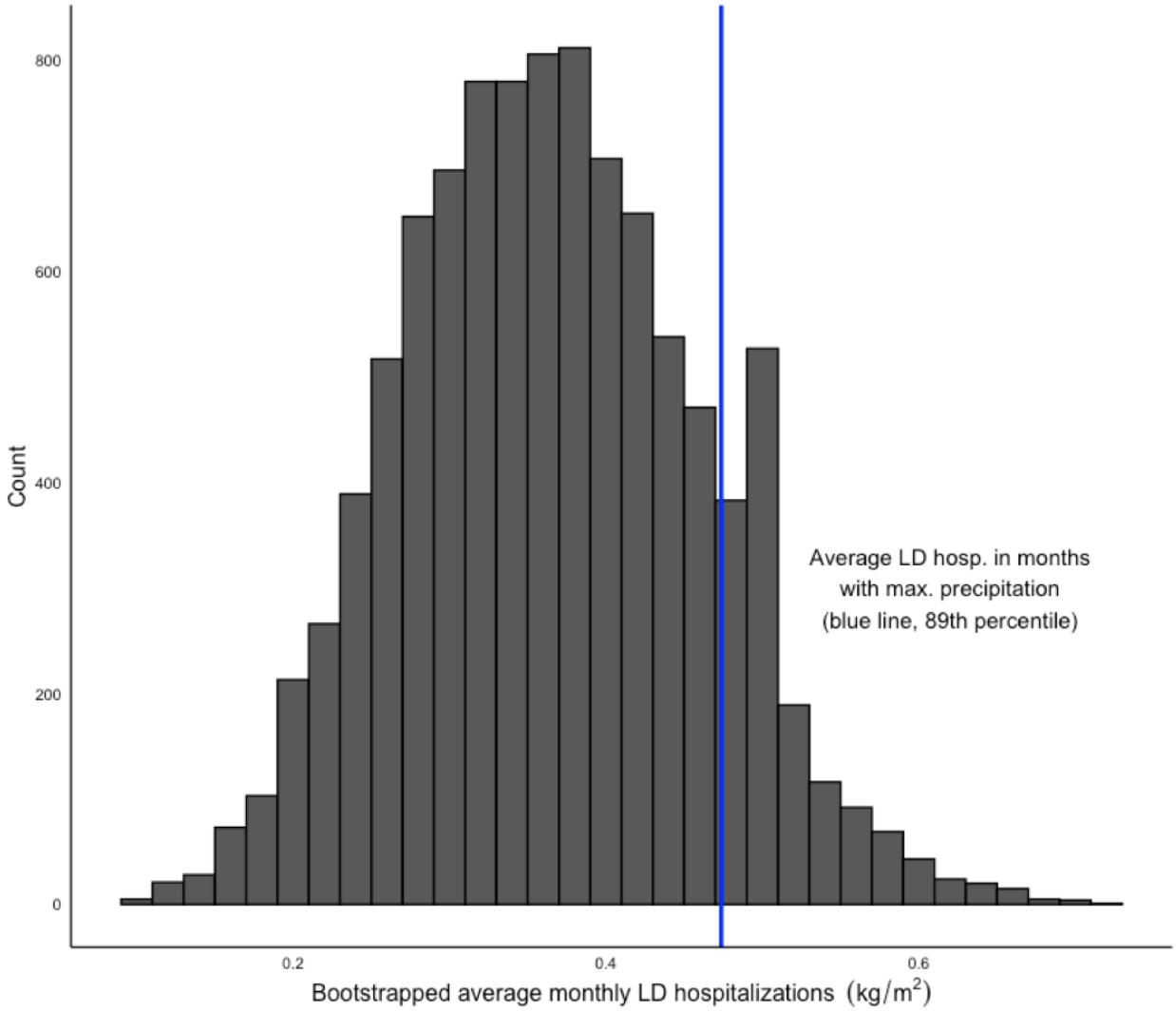

Figure 4. 3: Anomalous precipitation associated with Legionnaires' disease hospitalizations. Among the 75 hospitals with at least 10 cases, the average number of Legionnaires' disease hospitalizations in months with extreme precipitation is in the 89th percentile of the bootstrapped distribution of monthly Legionnaires' disease hospitalizations.

Table 4. 2 Associations between anomalous meteorology and hospitalizations
Percentile of average monthly hospitalizations in months with extreme meteorological conditions compared to bootstrapped distribution of average monthly hospitalizations

Hospitalization Threshold	Precipitation	Runoff	Soil Moisture	Temperature
1+ Case	0.87	0.57	0.38	0.52
5+ Cases	0.80	0.60	0.29	0.51
10+ Cases	0.89	0.58	0.50	0.46
15+ Cases	0.92	0.76	0.60	0.47
20+ Cases	0.94	0.70	0.60	0.56

The hydrometeorological flood-indicator variables exhibited seasonal patterns that varied by geographic region (Figure 4.S.3, Additional file 1). Precipitation typically peaked between June and September in the Northeast, Midwest, and Southwest, whereas along the West Coast it was driest during the summer and peaked in December or January (Figure 4.S.3a, Additional file 1). Soil moisture seasonality was consistent across the US, with maxima occurring in January or February and minima mid-summer, but the range varied by region (Figure 4.S.3b, Additional file 1). In the Northeast and Midwest, monthly soil moisture was relatively stable, whereas in the Southwest and on the West Coast there was a steep decline in soil moisture during the summer. Surface runoff exhibited the most distinct seasonality by region; on the West Coast it peaked during the winter, coinciding with the precipitation peaks, whereas in the Northeast surface runoff peaked in late spring, prior to the precipitation peak. Many areas in the Midwest experienced two peaks, one in the early spring and one in the later summer (Figure 4.S.3c, Additional file 1).

Soil moisture and precipitation were the most highly weighted variables identified by the importance weighting and multimodel inference (Figure 4.4). Both variables were positively associated with a significant increase in monthly Legionnaires' disease hospitalizations and were included in all of the top models (Table 4.3). A 1-standard deviation increase in average soil

moisture was associated with a 49% increase in hospitalizations in the most highly weighted model (Table 4.3). In the same model, a 1-standard deviation increase in average precipitation was associated with a 26% increase in hospitalizations. Temperature and the other hydrometeorological variables were not significantly associated with Legionnaires' disease hospitalizations in any of the top models. The importance weights, top models, and effect estimates were consistent across all hospitalization thresholds (Table 4.S.4, Additional file 1) and in the cross-validation sensitivity analysis. Similarly, multimodel inference using flood-indicator data aggregated to the hospitals' catchment areas identified the same top models and comparable effect estimates. In the most highly weighted model, a 1-standard deviation increase in soil moisture and precipitation at the catchment level was associated with a 53% and 26% increase in hospitalizations, respectively (Table 4.S.5, Additional file 1).

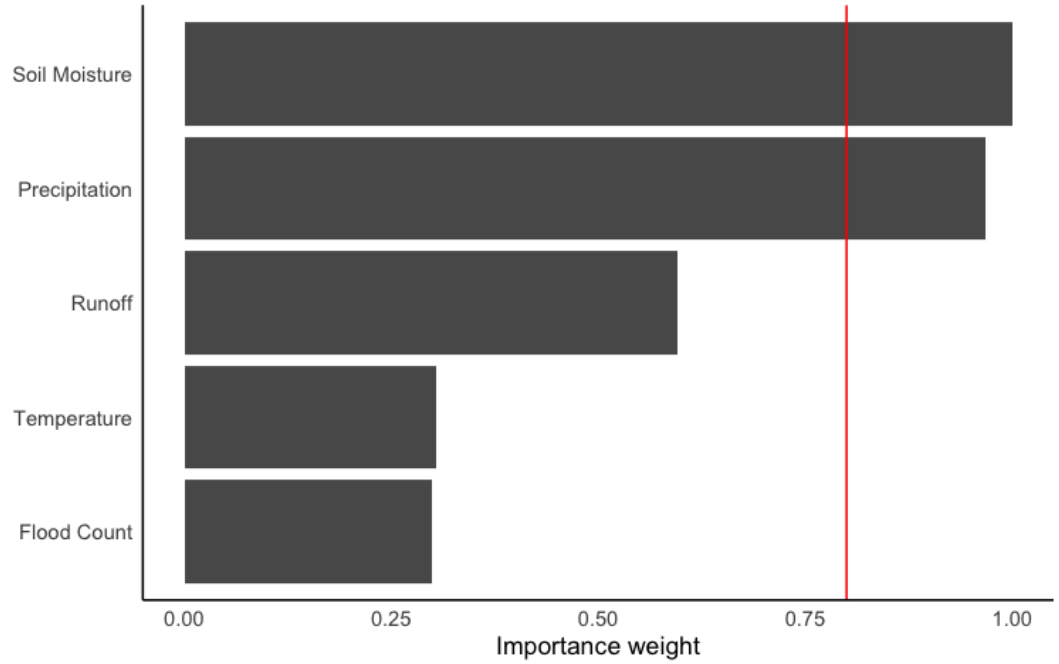

Figure 4. 4: Multimodel inference importance weights.
Soil moisture and precipitation were the most highly weighted flood-indicator variables assessed in the multimodel inference analysis; these variables were highly weighted in 98% and 96% of the candidate models, respectively. The red line indicates where variables are highly weighted in at least 80% of the candidate models; variables that exceed this importance threshold are included in the final model.

Table 4. 3 Association between hospitalizations and meteorological variables in the most highly weighted models

Model	Precipitation	Soil moisture	Temperature	Runoff	Flood count	Model weight
1	0.26 (0.14, 0.38)	0.49 (0.24, 0.74)		-0.08 (-0.16, 0.00)		0.261
2	0.28 (0.16, 0.40)	0.48 (0.23, 0.73)		-0.07 (-0.13, -0.01)	-0.03 (-0.11, 0.05)	0.176
3	0.19 (0.01, 0.29)	0.44 (0.20, 0.68)				0.133
4	0.23 (0.13, 0.33)	0.42 (0.17, 0.67)			-0.04 (-0.10, 0.02)	0.125
5	0.26 (0.14, 0.38)	0.47 (0.20, 0.74)	-1.28 (-6.49, 3.93)	-0.08 (-0.16, 0.00)		0.104
6	0.28 (0.16, 0.40)	0.50 (0.23, 0.77)	-1.27 (-6.54, 4.00)	-0.07 (-0.13, -0.01)	-0.04 (-0.10, 0.02)	0.070
7	0.19 (0.01, 0.29)	0.42 (0.17, 0.67)	-1.31 (-6.74, 4.12)			0.053

Note: Effect estimates are the change in monthly hospitalizations associated with a 1-standard deviation increase in the meteorological variables; values in parentheses indicate the 95% confidence interval.

4.4 Discussion

The incidence of sporadic Legionnaires' disease has increased for over 20 years, but the association between cases and flooding as a potential driver of disease has not been thoroughly examined in the US. Previous studies have found positive associations among cases, rainfall, and relative humidity, but most have focused on specific cities or small geographic regions [266, 267, 294]. Many have also relied on weather data from a single source [101, 281] or a single source per state [100], which may obscure local variation in meteorological conditions. In this study, the association between flooding, measured by several hydrometeorological variables, and Legionnaires' disease hospitalizations was analyzed across 75 hospitals in 25 states in the US, a geographic scope that encompasses a range of climatological regimes and demographics.

This work suggests that flooding, which can lead to the contamination of household and recreational water sources [8, 285], may be associated with hospitalizations for Legionnaires' disease. While previous work has implied that rainfall influences the spread of disease via

contamination [129, 241], none have focused on identifying or quantifying flood events as the driver of transmission. To address this gap, we used several methods to characterize extreme and seasonal floods; we found that hospitalizations increased during months with flooding due to extreme storms and were positively associated with monthly precipitation and soil moisture, which are common flood-indicator variables.

The seasonality and intensity of flooding varies considerably throughout the US, and these events cannot be measured with a single flood-indicator variable. We used two definitions of extreme events to account for the variety of flood types that occur, including those associated with hurricanes or tropical storms as well as those unrelated to cyclonic storms (e.g. due to intense precipitation and snowmelt). In the first approach, hospitalizations increased 32% in months with named storms during the Atlantic storm basin season among hospitals in the mid-Atlantic and Northeast. The second approach reinforced this finding and determined that hospitalizations throughout the US increased in months with anomalously high precipitation, not just those affected by cyclonic storms. The extreme event analysis supported a 2005 study that found legionellosis was positively associated with high atmospheric pressure more 10 days before occurrence and low atmospheric pressure within 5 days of occurrence, consistent with the transition that occurs when a storm front moves through an area [266].

In addition to extreme floods, many parts of the US experience seasonal flooding, including floods associated with snowmelt, frequent thunderstorms, and flash floods after droughts. We used multiple flood-indicator variables to characterize these seasonal floods and found that monthly soil moisture and precipitation are associated with increased Legionnaires' disease hospitalizations. The association between rainfall and cases is well-supported, but this is the first to assess soil moisture, which functions as an integrator of rainfall and is an important

flood indicator. *Legionella* bacteria thrive in extremely warm environments but in our analysis, extreme or seasonal temperature was not significantly associated with hospitalizations. This effect of temperature on Legionnaires' disease is inconsistent with previous studies; temperature lagged from 1 to 9 weeks was predictive of cases in some studies [102, 281] and associated with a decrease in disease rates in others. Given laboratory studies demonstrating that *Legionella* bacteria preferentially grow at high temperatures, it is likely that environmental temperature influences transmission. Our findings suggest, however, that extreme or seasonal flood events are more strongly associated with increased hospitalizations whereas temperature alone is not.

While the effect of flooding on Legionnaires' disease has not been examined in the US, our findings are supported by earlier research on the relationship between flood-indicator variables and outbreaks of other waterborne diseases [64, 65, 239]. An analysis of 42 years of outbreaks in the US found that 51% were preceded by extreme rainfall and that 60% were attributed to drinking water contamination [214]. This study did not examine the mechanisms by which flooding affects Legionnaires' disease hospitalizations, but previous research has identified *Legionella* in environments that are vulnerable to flooding. *Legionella* have been detected in surface runoff [129, 241, 280], which can directly contaminate drinking water sources or overwhelm water treatment systems during floods. The bacteria have also been found in wastewater and sewage treatment plants [135, 150], which are prone to overflows and contamination events associated with floods [162]. Our findings indicate an association between flooding and Legionnaires' disease, and future research should focus on examining the mechanisms by which flooding could lead to contamination and drive transmission.

Sources of contamination typical in Legionnaires' disease outbreaks, namely cooling towers, plumbing systems, and recreational or decorative pools [208], are often not the source for

sporadic cases [281]. Transmission of sporadic, community-acquired cases may instead be driven by household water and environmental exposures. Previous studies have attributed up to 40% of sporadic cases to potable water [295], and an elevated risk of infection has been associated with water from private wells [296] and from surface water (compared to potable water from groundwater sources) [241]. Water quality data from a range of sources could help determine the primary modes of exposure to *Legionella* for sporadic cases not associated with point source contamination. Detailed exposure analyses would also lead to an improved understanding of how infection occurs; while Legionnaires' disease transmission is thought to occur primarily via the inhalation of aerosolized bacteria, some studies suggest that infection also occurs via aspiration [246, 297]. Contaminated drinking water may be a crucial source of exposure if infection occurs via aspiration, as aerosolization by a household item (e.g. a showerhead, faucet, or hose) would not be required for transmission.

Our findings are constrained by a number of limitations related to the availability and resolution of the hospitalization data. The analysis does not include any data from the Southeastern US because these states either do not contribute to the HCUP dataset or do not provide monthly data; this is a major limitation, as states in this region are most prone to cyclonic storms. However, regions with the highest incidence of Legionnaires' disease were included in the analysis, and states that did not contribute to the HCUP dataset generally had lower incidence compared to the national average [204]. A recent analysis of Legionnaires' disease epidemiological trends in the United States between 1992 and 2018 found that age-standardized average incidence was higher in the Northeast and Midwest compared to the South and West, and highest in New England and the Mid-Atlantic states [204]; these geographic differences in incidence were more pronounced later in the time series (after 2002), which

overlaps with most of the study period in this analysis. Future studies should examine the associations among hydrometeorological conditions and Legionnaires' disease throughout the US, particularly in the Southeast, but the regions included in the study capture the states with the highest Legionnaires' disease burden.

Despite rising incidence, hospitalizations for Legionnaires' disease are relatively uncommon and as such our study relies on a small number of cases. During the study period, the total number of annual cases, not just hospitalizations, in the US reported to the Centers for Disease Control and Prevention (CDC) ranged from 969 to 3,676 [298]. To address this limitation, we repeated the analysis using several case-count thresholds in order to examine the consistency of our findings when different hospitals were included in the dataset. The stability of the associations, even when hospitals with a single case were included in the dataset, indicate that the findings are robust.

The National Inpatient Sample only provides monthly hospitalization data, which prohibits a more temporally resolved analysis, and the geographic location of the hospital, not the residential locations of the cases. The absence of more temporally or geographically resolved data introduces the possibility of misclassification bias, given that the flood data associated with the hospital's zip code may not accurately reflect the conditions at the cases' residential zip codes. We aimed to address these limitations by including a large number of hospitals in the study from rural, urban, and suburban areas and evaluating the consistency of our findings across different study sites. Our findings are also consistent with small-scale studies that used daily case data [266] or had residential location data [103].

4.5 Conclusion

Both seasonal and extreme flooding is projected to increase in conjunction with warming atmospheric temperatures, and our ability to mitigate the effect of these floods is contingent upon a thorough understanding of flood-disease dynamics and how they geographically vary. Our findings suggest that the increase of Legionnaires' disease across the US may be explained by flooding and that mitigating the effects of these events in the future is key to reducing the spread of disease. These results also suggest that current flood or contamination control measures are insufficient with respect to *Legionella* and may indicate that more rigorous water and wastewater treatment policies are required. The findings may also be of use to clinicians treating patients with respiratory symptoms in the wake of extreme events or during seasonal flood periods. While awareness of and testing for legionnaires' disease has increased, it remains substantially underdiagnosed and underreported among younger and immune-competent individuals. Future analysis should incorporate detailed water quality data from natural and built environments to better understand the routes of exposure, and how hydrological events affect transmission.

4.6 Supplementary Materials

Table 4. S. 1 Description of HCUP hospitals grouped by Legionnaires' disease case count thresholds

LD Case Threshold	1+ Case	5+ Cases	10+ Cases	15+ Cases	20+ Cases
Number of Hospitals	378	151	75	36	25
Number of LD Cases	2,361	1,885	1,376	980	715
Hospital Location (%)					
Rural	23.1	8.3	6.4	0	0
Urban	76.9	91.7	93.6	100	100
Hospital Bedsize (%)					
Small	23	13.2	9.8	11.4	11.8
Medium	27.3	24.9	24.9	25.7	19.1
Large	49.6	61.9	65.4	62.9	69.1
Geographic Region (%)					
Northeast	48	60.2	66.2	72.9	76.5
Midwest	25.4	18.2	16.8	2.9	16.2
Southwest	9	9.7	12.3	14.3	7.4
West Coast	17.6	11.9	4.7	0	0
Mean Annual Discharge (SD)	13,700 (11,200)	20,400 (12,000)	23,300 (13,100)	26,100 (13,700)	29,900 (14,000)

Table 4. S. 2 Cyclonic storms that affected counties with HCUP hospitals between 2000 and 2011

Storm	Year	Number of affected counties with HCUP hospitals
Allison	2001	46
Charley	2004	58
Dennis	2005	16
Frances	2004	32
Gustav	2008	31
Ike	2008	32
Irene	2011	41
Isabel	2003	22
Isidore	2002	20
Ivan	2004	37
Jeanne	2004	48
Katrina	2005	6
Lili	2002	2
Noel	2007	2
Rita	2005	4

Table 4. S. 3 Bootstrapped associations by HSA
Percentile of average monthly hospitalizations in months with extreme meteorological conditions averaged across Hospital Service Areas (HSAs) compared to bootstrapped distribution of average monthly hospitalizations

Hospitalization Threshold	Precipitation	Runoff	Soil Moisture	Temperature
1+ Cases	0.75	0.49	0.35	0.43
5+ Cases	0.81	0.57	0.37	0.45
10+ Cases	0.84	0.44	0.50	0.41
15+ Cases	0.91	0.66	0.52	0.41
20+ Cases	0.93	0.76	0.79	0.58

Table 4. S. 4 Association between Legionnaires' disease hospitalizations and meteorological variables in the most highly weighted model for each hospitalization threshold

Hospitalization Threshold	Precipitation	Soil moisture	Temperature	Runoff	Model weight
1+ Case	0.15 (0.09, 0.21)	0.20 (0.10, 0.29)		-0.07 (-0.13, -0.01)	0.21
5+ Cases	0.19 (0.07, 0.31)	0.55 (0.43, 0.67)		-0.06 (-0.14, 0.02)	0.29
10+ Cases	0.26 (0.14, 0.38)	0.49 (0.24, 0.74)		-0.08 (-0.16, 0.00)	0.26
15+ Cases	0.37 (0.21, 0.53)	0.61 (0.20, 1.02)		-0.10 (-0.20, 0.00)	0.26
20+ Cases	0.36 (0.18, 0.54)	0.92 (0.35, 1.49)	-11.64 (-23.97, 0.69)	-0.10 (-0.22, 0.02)	0.24

Note: Effect estimates are the change in monthly hospitalizations associated with a 1-standard deviation increase in the meteorological variables; values in parentheses indicate the 95% confidence interval.

Table 4. S. 5 Association between Legionnaires' disease hospitalizations and meteorological variables averaged across Hospital Service Areas (HSAs) in the most highly weighted models

Model	Precipitation	Soil moisture	Temperature	Runoff	Flood count	Model weight
1	0.26 (0.12, 0.40)	0.53 (0.22, 0.85)		-0.08 (-0.16, -0.00)		0.294
2	0.28 (0.14, 0.42)	0.52 (0.21, 0.83)		-0.08 (-0.16, -0.00)	-0.03 (-0.09, 0.03)	0.179
3	0.19 (0.07, 0.31)	0.48 (0.17, 0.79)				0.125
4	0.22 (0.10, 0.34)	0.48 (0.17, 0.79)			-0.04 (-0.10, 0.02)	0.118
5	0.25 (0.11, 0.39)	0.55 (0.22, 0.88)	1.25 (-5.57, 8.07)	-0.08 (-0.16, -0.00)		0.084
6	0.27 (0.13, 0.41)	0.54 (0.21, 0.87)	1.31 (-5.57, 8.19)	-0.08 (-0.16, -0.00)	-0.03 (-0.11, 0.05)	0.079
7	0.18 (0.06, 0.30)	0.50 (0.19, 0.81)	1.38 (-5.62, 8.38)			0.051

Note: Effect estimates are the change in monthly hospitalizations associated with a 1-standard deviation increase in the meteorological variables averaged across HSAs; values in parentheses indicate the 95% confidence interval.

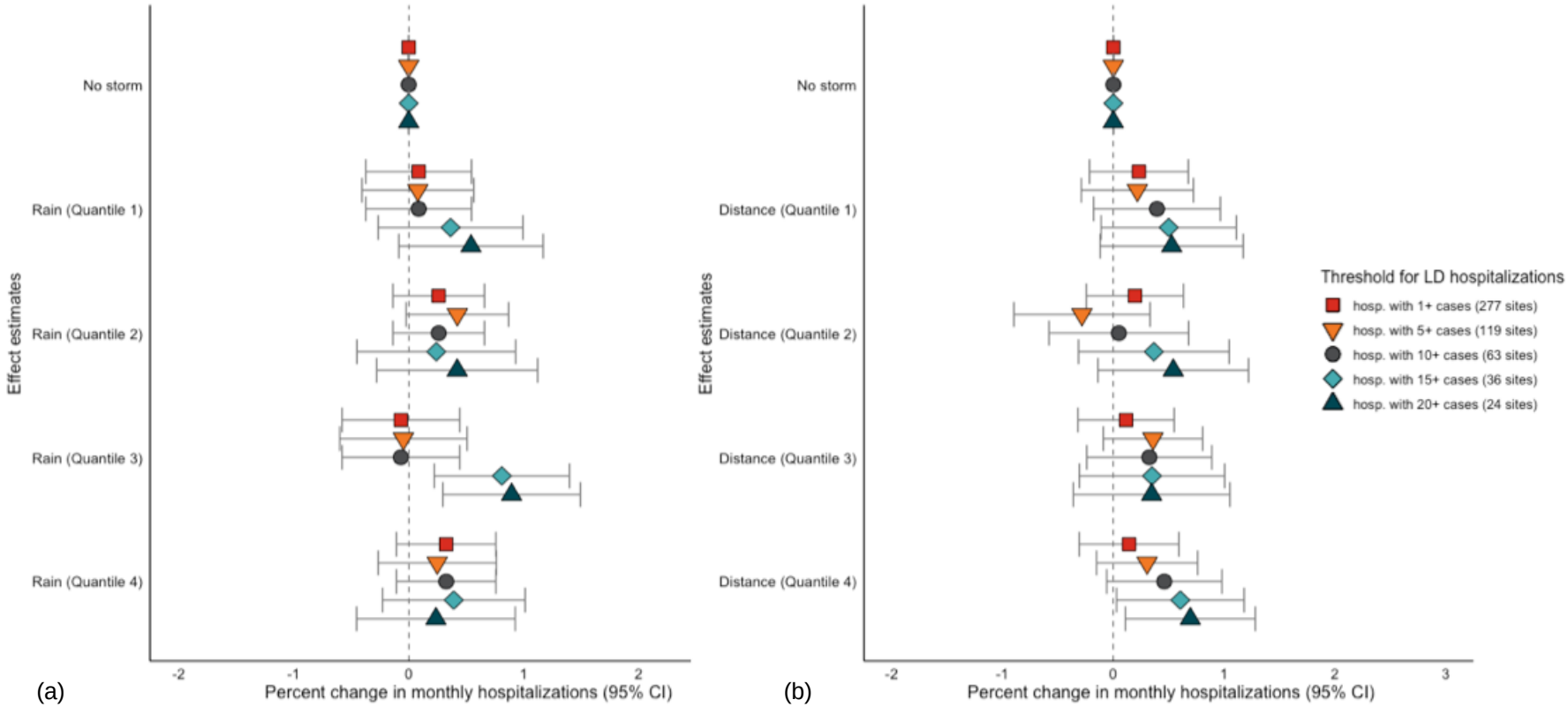

Figure 4. S. 1: Effect of exposure to cyclonic storms by case thresholds.
a) Precipitation associated with cyclonic storms and b) proximity to the storm track were not associated with a consistent significant change in monthly Legionnaires' disease hospitalizations among hospitals that experienced the storms, compared to hospitals that were unexposed to the storms. Moderately intense precipitation (quartile 3) was associated with a significant increase in hospitalizations among the hospitals with a minimum of 15 and 20 total cases, but this association was insignificant at different precipitation levels and case thresholds. The analysis was restricted hospitals in regions that experience cyclonic storms from the Atlantic storm basin and to the months of the Atlantic storm season (June – November).

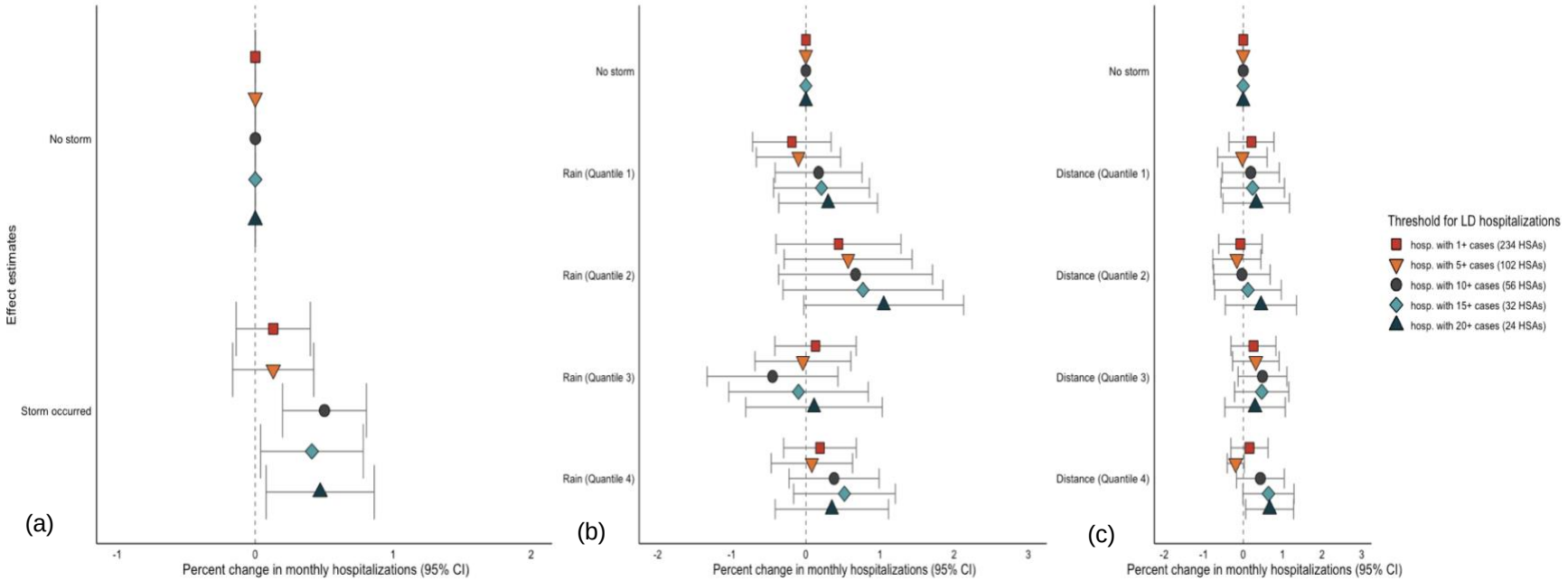

Figure 4. S. 2: Exposure to cyclonic storm at HSA level of analysis.
The association between exposure to cyclonic storms at the Hospital Service Area (HSA) level of analysis and monthly Legionnaires' disease hospitalizations did not substantially differ from the associations identified using the county-level storm data. a) Among hospitals in the 10-, 15-, and 20-case thresholds, hospitals in HSAs exposed to storms had a significant increase in hospitalizations compared to those in HSAs unexposed to storms. b) Cyclonic-storm related precipitation and c) proximity to storm tracks at the HSA level were not associated with significant changes in monthly hospitalizations; these findings are consistent with the analyses using county-level storm data (Fig.2, Supp. Fig

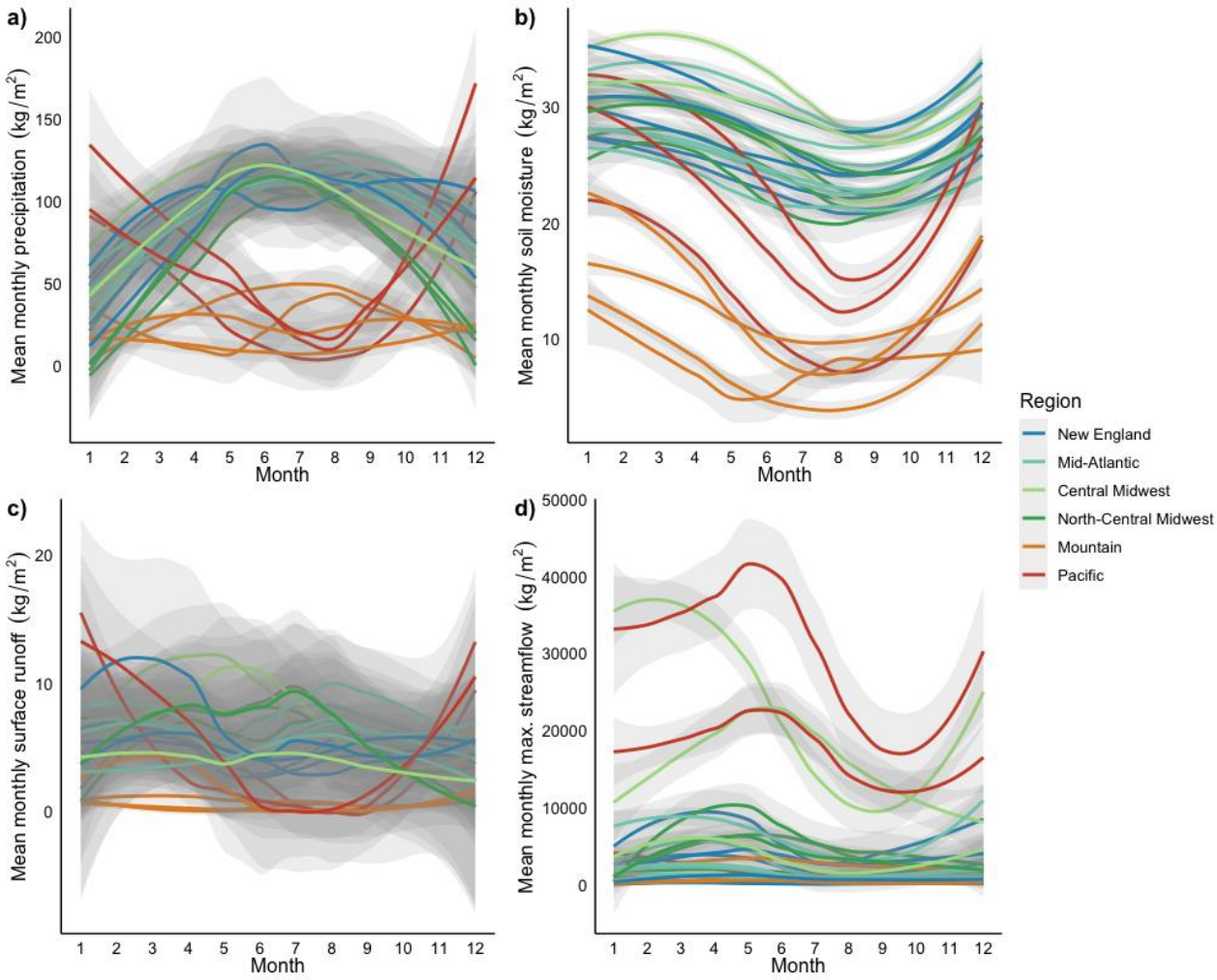

Figure 4. S. 3: Seasonality of hydrometeorological variables by region.
Monthly hydrometeorological flood-indicator variables averaged across the 75 hospitals in the primary analysis between 2000 and 2011, grouped by state (lines) and geographic region (color). The seasonality of a) precipitation and c) runoff differs in the Northeast and Midwest compared to the Southwest and West Coast, with peaks typically occurring in opposite months of the year. The seasonal pattern of b) soil moisture and d) streamflow is more consistent across the US, but the magnitude of the seasonal variation differs by region.

Supplementary Model Description 1.

Observed number of hospitalizations in hospital *i*, Y_i, is assumed to be distributed as a negative binomial variable:

$$Y_i \sim NB\ (Y_i \mid \mu_i, \theta)$$

where μ_i is the mean and θ is the shape parameter.

The model structure is:

$$\log(\mu_i) = X_i\beta + Z_i b + \log(P_i)$$

Where X_i are the variables of interest, which are storm occurrence and location, β is the vector of fixed effects for X_i, b is the vector of random effects for sample variables Z_i, and P_i is the offset, which is the total monthly hospitalizations for hospital *i*. The vector of K random effects, b, is assumed to follow a normal distribution: $b \sim N_k\ (0, \Psi)$ where Ψ is a positive definite variance-covariance matrix that determines the random effects. The statistical analysis was performed in R.

Supplementary Model Description 2.

Observed number of hospitalizations in hospital *i* at month *t*, Y_{it}, is assumed to be distributed as a negative binomial variable:

$$Y_{it} \sim NB\ (Y_{it} \mid \mu_{it}, \theta)$$

where μ_{it} is the mean, θ is the shape parameter and $t = 1,\ldots, n$. Here n (144) is the number of months in the study period.

The model structure is:

$$\log(\mu_{it}) = X_{it}\beta + Z_{it}b + \log(P_{it})$$

Where X_{it} are the variables of interest, which are the monthly standardized average hydrometeorological and temperature variables, monthly sine and cosine terms to account for seasonality ($\sin(2\pi \times \frac{\text{month}_t}{12})$ and $\cos(2\pi \times \frac{\text{month}_t}{12})$), and a term for year to account for long-term trends. β is the vector of fixed effects for X_{it}, b is the vector of random effects for sample variables Z_{it}, and P_{it} is the offset, which is the total monthly hospitalizations for hospital *i* in time *t*. The vector of K random effects, b, is assumed to follow a normal distribution: $b \sim N_k\ (0, \Psi)$ where Ψ is a positive definite variance-covariance matrix that determines the random effects. The statistical analysis was performed in R.

Chapter 5

The objective of this work was to quantify the effect of flooding on waterborne infectious diseases across common flood types in the US. We expanded on previous research by examining associations between specific pathogens and multiple flood-indicator meteorological variables, including those related to cyclonic storms, over many years. We incorporated data on hydroclimatology, location (rural/urban), and drinking water source to account for factors that may influence flood-disease dynamics. By including this wide range of variables, we sought to provide insight into the sources of contamination and transmission routes that govern waterborne disease in the US. We applied this framework with particular attention to Legionnaires' disease, a respiratory infection with high mortality that is rapidly increasing in incidence and may be an underestimated cause of community acquired pneumonia. In this final chapter, we first summarize our findings and consider them in the context of flood-disease dynamic literature (5.1). We then discuss the necessary next steps and future directions for waterborne disease research in the US (5.2) and end with a brief conclusion (5.3).

5.1 Discussion of the results in the context of flood-disease dynamic research

5.1.1. Summary of findings

We proposed a framework for considering pathogen biology, hydroclimatology, and infrastructure when studying the effect of flooding on waterborne diseases (Chapter 1), and applied this approach to our analysis of seasonal flooding on pathogen groups (Chapter 2). Hospitalization rates for all pathogen groups were higher in rural locations and, with the exception of biofilm-forming pathogens, in areas that relied on groundwater for drinking water sources. Waterborne disease is considered seasonal in temperate and subtropical areas [55, 77, 217], but we found substantial variability in the strength of seasonality by region and pathogen.

This underscores the need for more regionally-specific analyses in the US; previous work on waterborne infections has largely been conducted in a few states along the East Coast [58, 59, 64, 161], though we found hospitalization rates to be higher in the Midwest and Southwest. The group-level analysis was limited as associations with meteorological and location variables were driven by specific pathogens and did not accurately reflect all pathogens in the group, which demonstrates the importance of pathogen-specific analyses. To our knowledge, this was the first study to examine the effect of environmental variables on *Pseudomonas* and NTM infections and the first to include soil moisture for any of the pathogens. We found that Legionnaires' disease and *Pseudomonas* infections were positively associated with soil moisture.

Cyclonic storms often cause severe floods that are more destructive than those driven by seasonal hydrometeorology [115, 250, 299]. We evaluated the effect of tropical storm or hurricane exposure on cases of six waterborne infections using several exposure definitions and thresholds (Chapter 3). Storm hazards are often uncorrelated [249], i.e. areas that experience hurricane-force winds do not necessarily receive extreme rainfall, and we found that storm-related rainfall and wind had different effects on cases. Rainfall had a strong effect on Cryptosporidiosis, Legionnaires' disease, and Shiga toxin-producing *E. coli* infections while hurricane-force wind was only associated with a delayed increase in Cryptosporidiosis. We combined storm hazards to create storm type categories and examined their effect on cases; we found that high wind-high rain storms only had an effect on Cryptosporidiosis. This is a notable contribution to the field as the association between cyclonic storms and specific infections has not been studied across multiple storms seasons.

Seasonal meteorological conditions can lead to extreme flooding unrelated to tropical cyclones, especially in areas prone to spring snowmelt and intense rainfall [3, 107]. We assessed

the effect of extreme rainfall, soil moisture, runoff, and temperature on hospitalizations for Legionnaires' disease (Chapter 4). Hospitalizations increased significantly in months with extreme average rainfall, which indicates different flood-disease dynamics than those that determine the effect of seasonal flooding on infections. In the multi-pathogen analysis (Chapter 2), we found that soil moisture was positively associated with monthly Legionnaires' disease hospitalizations, a finding that we confirmed in this study. Under extreme meteorological conditions, however, rainfall was associated with Legionnaires' disease while soil moisture was not; this suggests that contamination sources or transmission routes may differ between seasonal and extreme floods.

5.1.2. Rurality and vulnerability to waterborne disease

Flooding and seasonal meteorology are important drivers of transmission in countries with endemic waterborne disease [213, 300]. Outbreaks of cholera, typhoid, and general diarrheal illness have been associated with monsoon seasons and extreme floods in places that lack adequate sanitation infrastructure [219, 301, 302]. These infections were also common in the US prior to the introduction of water treatment and organized waste management [303, 304] in the early 20th century. Since then, federal regulation under SDWA, discussed in 1.5, and advanced treatment methods have further improved drinking water quality [10]. National water quality and regulatory compliance assessments do not accurately represent the conditions for many communities in the US, however, and mask important regional variability in drinking water safety [179]. Clean drinking water is standard in many parts of the US but millions of people, predominantly in rural areas, rely on inadequate drinking water from noncompliant CWSs and private wells [11, 305, 306]. Rural CWSs have the highest rates of SDWA violations and report the worst drinking water quality [11, 307], problems that will likely worsen in the

future in conjunction with aging infrastructure. In these regions, high rates of diarrheal illness linked to environmental variability may still be an urgent, if unacknowledged, public health problem.

Rural CWSs are particularly vulnerable to contamination due to overlapping water source, land-use, and flooding risk factors [112, 179, 199]. People in rural areas generally receive drinking water from small CWSs that rely on groundwater sources or from private wells [185], both of which have been associated with increased contamination after flood events [112, 308, 309]. As described in Chapter 1, rural flooding is especially harmful because of animal waste in the environment and the insufficiency, or absence, of water treatment. Private wells present the greatest risk as they are entirely unregulated, but groundwater systems also suffer from inadequate treatment and poor enforcement of SDWA rules. It is a challenge for small CWSs to comply with monitoring and treatment standards mandated by the GWR; these systems typically have limited technical resources and small budgets, which make them dependent on securing federal grants for water treatment projects [306]. Given these factors, it is unsurprising that waterborne disease outbreaks disproportionately occur in CWSs that rely on groundwater [310]. Most are attributed to system failures in water or wastewater treatment [187, 311], but the consistent seasonality in drinking and recreational water outbreaks suggests that environmental drivers still influence transmission [167, 168, 187].

Rural counties experience health disparities with respect to many common chronic diseases and have a higher rate of adjusted-mortality compared to non-rural counties [312]. Our findings suggest that geographic disparities also exist for waterborne disease hospitalizations, which is consistent with the growing body of research on rural health [313, 314]. We found that hospitalization rates were higher in rural hospitals for all of the diseases except amebic

infections. Cryptosporidiosis, Shigellosis, Campylobacteriosis, and *E. coli* hospitalization rates were also higher in areas that relied on groundwater, which is the dominant source for drinking water in rural CWSs. Many rural regions are vulnerable to multiple flood types including snowmelt-driven river floods, which generate persistent standing floodwater that can contaminate groundwater [3]. Compounding this risk are demographic factors; rural counties comprise 85% of the "older-age" counties in the US [314] and age is the greatest risk factor for serious waterborne infections. Rural counties generally have fewer physicians per capita and weaker healthcare infrastructure, which means that cases are probably under detected and underreported [315]. Specific rural areas are extraordinarily vulnerable to waterborne infection and are likely hotspots for transmission. In some parts of rural Alabama, for example, straight piping sewage is common practice and leads to the discharge of raw waste within several yards of residences [316]. Surrounding soil and water is often highly pathogenic and direct contamination of drinking water is frequent. Given these overlapping risk factors, focusing on flood-disease dynamics in rural areas should be a priority for public health research.

5.1.3. Growing importance of opportunistic pathogens

Diarrheal illness has traditionally been the focus of waterborne disease research but the emergence of Legionnaires' disease has demonstrated the need for a more expansive framework that incorporates respiratory infections caused by waterborne pathogens. Flood-related polymicrobial pneumonia has been observed after extreme floods in Southeast Asia [317] and tropical cyclones have been associated with increased hospitalization rates for pneumonia in the US [234], but few pathogen-specific analyses have been conducted and none over multiple years. Identifying flood-disease dynamics for respiratory infections is critical given the severity of these infections and their disproportionate burden among vulnerable people [17]. In the US, more

people are living longer with immunocompromising conditions and the overall population is aging [17, 318], which indicates that the number of people susceptible to these opportunistic infections will increase in the future. To address this gap, we studied the effect of multiple storm exposures on Legionnaires' disease cases and found a substantial increase associated with storm-related rainfall. This is an important first step and invites further research into possible routes of exposure in the aftermath of storms. Transmission due to contaminated plumbing has been well-established but a critical, and largely unexamined, question remains as to the role of environmental exposure; churning floodwater, for example, may aerosolize bacteria in waterways and cause direct transmission in the environment. While Legionnaires' disease is the most prominent, other biofilm-forming pathogens may be significant contributors to waterborne respiratory disease.

Respiratory *pseudomonas* and NTM infections are opportunistic pathogens that share biological and epidemiological characteristics with Legionnaires' disease but their association with environmental variables has not been examined in the US. While primarily considered nosocomial pathogens, they are also an established cause of community acquired pneumonia (CAP) in immunocompromised people [67]. The bacteria are ubiquitous in treated and environmental water, so determining their association with flooding could provide insight into periods of high transmission risk for vulnerable groups. As opportunistic pathogens, they may also be indirectly related to flooding as sequelae of initial post-flood health effects. Flooding has been associated with elevated rates of asthma and COPD [234, 319], which are conditions that may increase susceptibility to these pathogens. We found that respiratory *Pseudomonas* hospitalizations had similar, though attenuated, associations with meteorological variables as Legionnaires' disease. Further work should focus on better-estimating the role of *Pseudomonas*

as a causative agent for CAP and specifically studying these cases, as non-nosocomial infections may be more sensitive to environmental variables.

5.1.4 Extreme floods and modeling rare events

The number of major, billion-dollar floods has increased by 5% each year in the US since 1980 [289, 320], a trend that is likely to continue under future climate scenarios [320]. Warmer atmospheric temperatures will affect all of the flood generating processes discussed in Chapter 1; earlier snowmelt and intense rainfall will make both river floods and flash floods more severe, and increased cyclonic storm severity will worsen coastal flooding [321-323]. These events are generally more destructive and generate more expansive floodwater compared to normal flooding. As a result, extreme events may facilitate transmission through different pathways; direct contact with contaminated water may be a more prominent transmission route after cyclonic storms, for example, than it is during a standard flood season. Our intention with including several flood-indicator variables and using multiple definitions for extreme events, as described in Chapter 1, was to capture the range of flood types that occur in the US. We found some support for this approach in our analysis of Legionnaires' disease in Chapter 4. Hospitalizations were most strongly associated with soil moisture in the analysis of seasonal flood-indicator variables, but anomalously high monthly rainfall had a stronger effect in the nonparametric approach. Further research should explicitly study transmission routes during extreme and seasonal flooding so as to determine the most effective disaster management responses and to identify points for intervention.

Among flood anomalies, some events are extraordinarily destructive and likely have an outsized influence on waterborne disease transmission. Quantifying their effect is a fundamental challenge in flood epidemiology, however, as they are rare events. It is difficult to evaluate this

subsection of extreme events in most statistical frameworks because their small sample size generates large amounts of uncertainty [324]. In our analysis of storm types, we found that high rain-high wind storms had no effect on cases except for Cryptosporidiosis with a three-week lag. This finding illustrates the limitations inherent in studying extremes; the effect may reflect a consistent, delayed association between cases and high rain-high wind storms or it may be driven by an unusual occurrence of successive hurricanes and tropical storms in Texas (described in Chapter 4). The absence of an effect at any other lag or with other infections is also difficult to interpret because it may be a result of the small number of high rain-high wind storms. Conversely, it could a genuine effect driven by the catastrophic nature of these events. People evacuated prior to storm landfall avoid floodwater or contaminated drinking water and some of those exposed may be displaced in the aftermath, so surveillance systems close to the flood zone would not capture case reports [258]. Healthcare systems are also disrupted, which inhibits case detection, reporting, and epidemiological surveillance [325, 326].

5.2 Next steps for waterborne disease research in the US

Some of the current challenges and limitations to waterborne disease research in the US were outlined in 5.1. Here, we discuss the effect of these limitations and outline the necessary steps to address gaps in our understanding of waterborne disease transmission. We also identify important areas of research that could inform public health and flood management policies to reduce the burden of disease.

5.2.1 Geographic and temporal resolution of data

Geographically and temporally resolved epidemiological data are crucial for conducting informative analyses of flood-disease dynamics. These data enable more rigorous assessments of exposure routes by linking cases to the CWS that provides household water, determining their

proximity to environmental or recreational water sources, and specifically characterizing storm exposure. State and local health departments conduct surveillance of water- and foodborne infections [58, 64, 222, 327] but the data are largely inaccessible to researchers outside these organizations and CDC. States report deidentified data on single cases, i.e. non-outbreak cases, of notifiable diseases to NNDSS, which publicly releases these data at state resolution (data used in Chapter 3) [298, 328]. Through the FoodNet program, health departments and CDC partner to perform exemplary epidemiological surveillance for several pathogens that are primarily foodborne in 10 participating states (e.g. *Salmonella*, STEC) [329]. The data are mainly used to track outbreaks, however, and access to them is generally limited to CDC and its partner health departments. As a result, the majority of studies that use these resolved, individual-level data are conducted in several states along the East Coast [32, 205, 264].

Flood-disease dynamic studies typically lack water quality data, which is necessary for establishing mechanistic associations between floods and cases. Many of the pathogens included in this analysis are not subject to mandatory sampling or compliance standards, so data collection is limited to specific research projects that focus on a single waterbody over a short period of time [130, 211]. SDWA requires states to monitor drinking water sources for regulated contaminants (e.g. *E. coli*), but the availability of that data varies within and among states [194]. The data are also reported at monthly intervals, which is not useful for assessing the short-term effect of flooding on water quality. USGS programs conduct continuous monitoring or estimation of environmental water quality that generate datasets with excellent spatial and temporal resolution, but do not measure pathogen concentrations [330]. In lieu of adequate time-series data, many studies rely on SDWA violations to characterize drinking water quality [11, 305, 331]. CWSs are required to report violations, which SDWIS compiles and makes public, but

in practice there is variability in compliance monitoring and reporting [311] (Chapter 1). The lack of time-series water quality data, especially data with broad geographic coverage, is arguably the greatest challenge for flood-disease dynamic research. The cost associated with regular, pathogen-specific monitoring is likely prohibitive particularly given the chronic underfunding of many CWSs.

Our work was constrained by the use of monthly hospitalization data (Chapters 2 and 4) and the absence of residential location for hospitalized and case patients (Chapters 2-4). We tested the sensitivity of our findings with exposure thresholds, and supplemental analyses using meteorological data associated with hospital catchment area, but the potential for spatial and temporal mismatch persisted. It is unlikely that the data resolution yielded overestimated effects, as the misclassification bias would be toward the null [260], but it did prevent modeling nonlinear, spatial, or lagged effects (except for the weekly storm data in Chapter 3) [228]. Hospitalization data only capture the most severe cases, which are a small fraction of overall waterborne infections and disproportionately occur among older people. Hospitals also vary in their policies with respect to diagnostic testing and admissions, which could affect the ICD-9 codes reported and admission or discharge dates. To address this limitation, we included random effect terms for hospitals throughout our statistical analyses; this approach controls for hospital-specific differences in diagnostic testing, admission policies, and reporting of ICD-9 codes.

Improving the quality and availability of epidemiological surveillance and drinking water quality data is central for addressing the most urgent public health challenges related to waterborne disease.

5.2.2. Health disparities related to waterborne disease

One of the most harmful consequences of data limitations is the inability to rigorously examine health disparities related to waterborne infections. For decades public health institutions have been ostensibly committed to eliminating health disparities but they have persisted, and in some cases worsened, across most health outcomes [332, 333]. Few studies in the US have examined waterborne infections by sociodemographic or socioeconomic variables and none across broad geographic ranges. Case rates for waterborne infections have been associated with county-level socioeconomic and sociodemographic variables, but the underlying mechanisms have not been evaluated. Counties with populations that are primarily low-income, racial or ethnic minorities, or foreign-born report higher rates of waterborne infections [32, 334-337]. These studies do not consider systemic factors that could influence risk, however, and instead attribute differences to individual-level behavioral factors (e.g. food preparation, dietary habits). The FoodNet surveillance system discussed in 5.2.1 collects individual-level data on occupation, household water source, and exposure to recreational water sources; expanding the program to include all waterborne pathogens would be transformative for flood-disease dynamic and health disparities research. Rather than solely documenting differences by sociodemographic variables, it would provide insight into the source of health disparities. With geographically resolved epidemiological data, location-specific factors related to transmission, including proximity to WWTPs and land-use categories, could be assessed in conjunction with sociodemographic data.

While informative pathogen-specific analyses are limited, disparities related to drinking water infrastructure have been more thoroughly examined. Higher rates of SDWA violations have been documented in CWSs that serve communities with lower average socioeconomic status (SES), lower insurance rates, and higher proportions of non-Hispanic Black and Hispanic

residents [305, 331, 338, 339]. This has been attributed to people with lower SES living in areas with older infrastructure and relying on CWSs that use groundwater [305]; as discussed in Chapter 1, groundwater CWSs are vulnerable to contamination because water is often undertreated. Consistent with this analysis, small (≤ 10,000 people served) and rural CWSs were also at greater risk for SDWA violations [305, 340]. The data used in these studies are not specific to bacterial contamination, however, and associations may differ by violation type. For example, increased SDWA violations have been associated with counties in the Southwestern US and on tribal land, but these are also regions with high levels of arsenic in soil that may cause frequent violations unassociated with biological contamination [340, 341]. Future research should assess CWS-level SES variables and violations stratified by type.

SDWA violations identify CWSs at risk for contamination, but they do not capture the groups most vulnerable to waterborne disease. An estimated 1.1 million people in 471,000 households do not have access to piped water and areas where this is common may represent hotspots for transmission [340, 342]. Unlike the CWS disparities associated with rurality, lack of piped water is more common in densely populated urban areas. People in these households are more likely to be from racial or ethnic minority groups and to be renters [342]. Numerous health effects including respiratory infections, diarrheal illness, and malnutrition are associated with the absence of household water, though these studies are typically conducted in low- and middle-income countries [343]. Implementing interventions in areas of extreme health disparity and risk, including the rural households that lack sewage infrastructure discussed in 5.1.2, should be a public health priority for local and state governments.

The clearest and most consistent identified health disparities relevant to flood-disease dynamics are driven by cyclonic storm exposure. Vulnerability to flooding mirrors other

indicators of social inequity; people who are racial or ethnic minorities, recent immigrants, or have physical disabilities are disproportionately affected by extreme events [344, 345]. Low income families are more likely to have their homes destroyed during natural disasters [346] and federally-subsidized housing is overly represented among properties in floodplains [347]. There is spatial heterogeneity in vulnerability to natural disasters across the US that has changed over time in conjunction with shifting demographics [348]. Vulnerability has increased in communities with aging populations, especially in small towns in the Upper Midwest, compared to earlier decades. It has also increased in predominantly Hispanic and low-income communities along the Mexico border, where risk is associated with language barriers, and in the deep south [348, 349].

Health disparities related to storm exposure are a result of overlapping risk factors; people who are most vulnerable to exposure are also those who have fewer resources to evacuate and recover from them [347]. Susceptibility to waterborne disease may be associated with both of these factors, but assessing these interactions requires comprehensive epidemiological surveillance to establish baseline risk as well as post-storm epidemiological studies.

5.2.3 Preparing for the effect of climate change on burden of disease

Climate change will likely have a profound influence on flood-disease dynamics as its effects will alter the severity of seasonal and extreme flooding. The ability to anticipate the consequences for specific health outcomes is hampered, however, but our limited understanding of current dynamics [350, 351]. The effect of warmer atmospheric temperatures on river floods, flash floods, and cyclonic storms is established (and discussed in 1.3); extreme rainfall and flood severity, in some regions, have already increased in the US [352-354]. There is considerable uncertainty, however, surrounding the effects of climate change on disease dynamics;

temperature, rainfall, and humidity changes may influence the ability of pathogens to persist in the environment [65, 237] and the seasonality of infections in animals [355]. This may be particularly important for biofilm-forming bacteria because they actively grow, rather than merely persist, in the environment and are sensitive to meteorological conditions. The effect of extreme storms and floods on most health outcomes is also uncertain; recent work has focused on characterizing a range of health effects [234], but the field is developing and constrained by data availability.

The effects of cyclonic storms are often hyperlocal and epidemiological analyses tailored to specific locations may best describe flood-associated risks. As a result, however, it is difficult to establish a comprehensive understanding of flood-disease dynamics that has adequate external validity [356]. Post-storm analyses vary by data collection, methodological approaches, and outcomes, which generates distinct post-storm surveillance data. The work presented in Chapter 3 aims to address this gap in flood epidemiology by applying consistent exposure and outcome definitions over 23 years of extreme storms. Future research should adapt this approach with more geographically resolved data to better characterize exposure and to incorporate community-level information that could provide insight into possible drivers of health disparities. The most glaring source of uncertainty concerns the effects of climate change, and flooding specifically, on sociodemographic trends [357]. The effects of climate change may lead to temporary [358, 359] and permanent displacement [360], which affects disease transmission, drinking water infrastructure, and healthcare systems in addition to innumerable social effects.

5.3 Conclusion

Decades of public policy designed to improve drinking water quality and sanitation infrastructure led to dramatic reductions in the burden of waterborne disease throughout the 20th

century. In some regions and communities, however, contamination persists and may lead to disease transmission that is not captured by epidemiological surveillance or healthcare systems. Vulnerability to inadequate infrastructure and waterborne disease is concentrated among groups within the population that often experience multiple sources of health and social inequity. The effects of climate change will likely exacerbate these disparities. Determining how climate-driven processes influence health, and identifying who is most at risk, is therefore central to environmental justice work. The goal of this dissertation work was to provide some insight into flood-disease dynamics and to develop a framework for approaching waterborne disease research. The ability to examine sociodemographic and socioeconomic risk factors will require a more nuanced understanding of these dynamics, which in turn depends on improved infectious disease and water quality surveillance. In the face of aging water infrastructure and more extreme floods, this should be a priority for public health researchers and policymakers.

References

1. Collier SA, Deng L, Adam EA, Benedict KM, Beshearse EM, Blackstock AJ, et al. Estimate of Burden and Direct Healthcare Cost of Infectious Waterborne Disease in the United States. Emerg Infect Dis. 2021;27(1):140-9.

2. Villarini G. On the seasonality of flooding across the continental United States. Advances in Water Resources. 2016;87:80-91.

3. Berghuijs WR, Woods RA, Hutton CJ, Sivapalan M. Dominant flood generating mechanisms across the United States. Geophysical Research Letters. 2016;43(9):4382-90.

4. King BJ, Monis PT. Critical processes affecting Cryptosporidium oocyst survival in the environment. Parasitology. 2007;134(Pt 3):309-23.

5. Stelzer W, Jacob J, Schulze E. Environmental aspects of Campylobacter infections. Zentralblatt für Mikrobiologie. 1991;146(1):3-15.

6. Plutzer J, Ongerth J, Karanis P. Giardia taxonomy, phylogeny and epidemiology: Facts and open questions. Int J Hyg Environ Health. 2010;213(5):321-33.

7. Robertson LJ. Giardia and Cryptosporidium infections in sheep and goats: a review of the potential for transmission to humans via environmental contamination. Epidemiol Infect. 2009;137(7):913-21.

8. Amaral-Zettler LA, Rocca JD, Lamontagne MG, Dennett MR, Gast RJ. Changes in microbial community structure in the wake of Hurricanes Katrina and Rita. Environ Sci Technol. 2008;42(24):9072-8.

9. Nicholson FA, Groves SJ, Chambers BJ. Pathogen survival during livestock manure storage and following land application. Bioresour Technol. 2005;96(2):135-43.

10. Weinmeyer R, Norling, A, Kawarsk, M, Higgins, H. The Safe Drinking Water Act of 1974 and its role in providing access to safe drinking water in the United States. AMA J Ethics. 2017;19(10):1018-26.

11. Allaire M, Wu H, Lall U. National trends in drinking water quality violations. Proc Natl Acad Sci U S A. 2018;115(9):2078-83.

12. Hutchison ML, Walters LD, Avery SM, Synge BA, Moore A. Levels of zoonotic agents in British livestock manures. Lett Appl Microbiol. 2004;39(2):207-14.

13. Rodriguez A, Pangloli P, Richards HA, Mount JR, Draughon FA. Prevalence of Salmonella in diverse environmental farm samples. J Food Prot. 2006;69(11):2576-80.

14. Fletcher SM, Stark D, Harkness J, Ellis J. Enteric protozoa in the developed world: a public health perspective. Clin Microbiol Rev. 2012;25(3):420-49.

15. Crump JA, Sjolund-Karlsson M, Gordon MA, Parry CM. Epidemiology, Clinical Presentation, Laboratory Diagnosis, Antimicrobial Resistance, and Antimicrobial Management of Invasive Salmonella Infections. Clin Microbiol Rev. 2015;28(4):901-37.

16. DWK A. Food and waterborne illnesses. Encyclopedia of Microbiology. 2009:365-81.

17. Falkinham JO, 3rd, Hilborn ED, Arduino MJ, Pruden A, Edwards MA. Epidemiology and Ecology of Opportunistic Premise Plumbing Pathogens: Legionella pneumophila, Mycobacterium avium, and Pseudomonas aeruginosa. Environ Health Perspect. 2015;123(8):749-58.

18. Rose JB. Environmental ecology of Cryptosporidium and public health implications. Annu Rev Public Health. 1997;18:135-61.

19. Reynolds D, Kollef M. The Epidemiology and Pathogenesis and Treatment of Pseudomonas aeruginosa Infections: An Update. Drugs. 2021;81(18):2117-31.

20. Lake IR, Nichols G, Bentham G, Harrison FC, Hunter PR, Kovats SR. Cryptosporidiosis decline after regulation, England and Wales, 1989-2005. Emerg Infect Dis. 2007;13(4):623-5.

21. Kaper JB, Nataro JP, Mobley HL. Pathogenic Escherichia coli. Nat Rev Microbiol. 2004;2(2):123-40.

22. Aljahdali NH, Sanad YM, Han J, Foley SL. Current knowledge and perspectives of potential impacts of Salmonella enterica on the profile of the gut microbiota. BMC Microbiol. 2020;20(1):353.

23. Cabral JP. Water microbiology. Bacterial pathogens and water. Int J Environ Res Public Health. 2010;7(10):3657-703.

24. Karanis P, Kourenti C, Smith H. Waterborne transmission of protozoan parasites: a worldwide review of outbreaks and lessons learnt. J Water Health. 2007;5(1):1-38.

25. Tzipori S, Ward H. Cryptosporidiosis: biology, pathogenesis and disease. Microbes Infect. 2002;4(10):1047-58.

26. Strachan NJ, Rotariu O, Smith-Palmer A, Cowden J, Sheppard SK, O'Brien SJ, et al. Identifying the seasonal origins of human campylobacteriosis. Epidemiol Infect. 2013;141(6):1267-75.

27. Li B, Vellidis G, Liu H, Jay-Russell M, Zhao S, Hu Z, et al. Diversity and antimicrobial resistance of Salmonella enterica isolates from surface water in Southeastern United States. Appl Environ Microbiol. 2014;80(20):6355-65.

28. Jacobsen CS, Bech TB. Soil survival of Salmonella and transfer to freshwater and fresh produce. Food Research International. 2012;45(2):557-66.

29. Painter JA, Hoekstra RM, Ayers T, Tauxe RV, Braden CR, Angulo FJ, et al. Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998-2008. Emerg Infect Dis. 2013;19(3):407-15.

30. van Putten JPM, van Alphen LB, Wösten MMSM, de Zoete MR. Molecular Mechanisms of Campylobacter Infection. In: Sasakawa C, editor. Molecular Mechanisms of Bacterial Infection via the Gut. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009. p. 197-229.

31. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. Recent advances in understanding enteric pathogenic Escherichia coli. Clin Microbiol Rev. 2013;26(4):822-80.

32. Morgado ME, Jiang C, Zambrana J, Upperman CR, Mitchell C, Boyle M, et al. Climate change, extreme events, and increased risk of salmonellosis: foodborne diseases active surveillance network (FoodNet), 2004-2014. Environ Health. 2021;20(1):105.

33. Heymann DL. Control of communicable diseases manual: An official report of the American Public Health Association. Washington, D.C. : American Public Health Association; 2015.

34. Xiao L. Molecular epidemiology of cryptosporidiosis: an update. Exp Parasitol. 2010;124(1):80-9.

35. Adam RD. Biology of Giardia lamblia. Clin Microbiol Rev. 2001;14(3):447-75.

36. Newton Hayley J, Ang Desmond KY, van Driel Ian R, Hartland Elizabeth L. Molecular Pathogenesis of Infections Caused by Legionella pneumophila. Clinical Microbiology Reviews. 2010;23(2):274-98.

37. Shamaei M, Mirsaeidi M. Nontuberculous Mycobacteria, Macrophages, and Host Innate Immune Response. Infect Immun. 2021;89(8):e0081220.

38. McMurray DN. Mycobacteria and Nocardia. In: th, Baron S, editors. Medical Microbiology. Galveston (TX)1996.

39. Turenne CY. Nontuberculous mycobacteria: Insights on taxonomy and evolution. Infect Genet Evol. 2019;72:159-68.

40. Curran CS, Bolig T, Torabi-Parizi P. Mechanisms and Targeted Therapies for Pseudomonas aeruginosa Lung Infection. American journal of respiratory and critical care medicine. 2018;197(6):708-27.

41. Fujitani S, Sun HY, Yu VL, Weingarten JA. Pneumonia due to Pseudomonas aeruginosa: part I: epidemiology, clinical diagnosis, and source. Chest. 2011;139(4):909-19.

42. Iglewski BH. Pseudomonas. In: th, Baron S, editors. Medical Microbiology. Galveston (TX)1996.

43. Adams DA, Thomas KR, Jajosky RA, Foster L, Baroi G, Sharp P, et al. Summary of notifiable infectious diseases and conditions -- United States, 2015. MMWR Morbidity and mortality weekly report. 2017;65(53).

44. Hines JZ, Jagger MA, Jeanne TL, West N, Winquist A, Robinson BF, et al. Heavy precipitation as a risk factor for shigellosis among homeless persons during an outbreak — Oregon, 2015–2016. Journal of Infection. 2018;76(3):280-5.

45. van Elsas JD, Semenov AV, Costa R, Trevors JT. Survival of Escherichia coli in the environment: fundamental and public health aspects. ISME J. 2011;5(2):173-83.

46. Sidhu JP, Ahmed W, Hodgers L, Toze S. Occurrence of virulence genes associated with Diarrheagenic pathotypes in Escherichia coli isolates from surface water. Appl Environ Microbiol. 2013;79(1):328-35.

47. Mena KD, Gerba CP. Risk assessment of Pseudomonas aeruginosa in water. Rev Environ Contam Toxicol. 2009;201:71-115.

48. Chao WL, Ding RJ, Chen RS. Survival of pathogenic bacteria in environmental microcosms. Zhonghua Min Guo Wei Sheng Wu Ji Mian Yi Xue Za Zhi. 1987;20(4):339-48.

49. Winfield MD, Groisman EA. Role of nonhost environments in the lifestyles of Salmonella and Escherichia coli. Appl Environ Microbiol. 2003;69(7):3687-94.

50. Jones K. Campylobacters in water, sewage and the environment. Symp Ser Soc Appl Microbiol. 2001(30):68S-79S.

51. McKergow LA, Davies-Colley RJ. Stormflow dynamics and loads ofEscherichia coliin a large mixed land use catchment. Hydrological Processes. 2009:n/a-n/a.

52. Burrows MR, Rankin JD. A further examination of the survival of pathogenic bacteria in cattle slurry. Br Vet J. 1970;126(8):xxxii+.

53. Alonso JL, Alonso MA, Usera MA, Echeita A. The occurrence of Salmonella serotypes in marine recreational waters of Valencia, Spain. Microbiologia. 1992;8(1):44-8.

54. Temple KL, Camper AK, McFeters GA. Survival of two enterobacteria in feces buried in soil under field conditions. Appl Environ Microbiol. 1980;40(4):794-7.

55. Philipsborn R, Ahmed SM, Brosi BJ, Levy K. Climatic Drivers of Diarrheagenic Escherichia coli Incidence: A Systematic Review and Meta-analysis. J Infect Dis. 2016;214(1):6-15.

56. Bi P, Cameron AS, Zhang Y, Parton KA. Weather and notified Campylobacter infections in temperate and sub-tropical regions of Australia: an ecological study. J Infect. 2008;57(4):317-23.

57. Kuhn KG, Nygard KM, Guzman-Herrador B, Sunde LS, Rimhanen-Finne R, Tronnberg L, et al. Campylobacter infections expected to increase due to climate change in Northern Europe. Sci Rep. 2020;10(1):13874.

58. Jiang C, Shaw KS, Upperman CR, Blythe D, Mitchell C, Murtugudde R, et al. Climate change, extreme events and increased risk of salmonellosis in Maryland, USA: Evidence for coastal vulnerability. Environ Int. 2015;83:58-62.

59. Lee D, Chang HH, Sarnat SE, Levy K. Precipitation and Salmonellosis Incidence in Georgia, USA: Interactions between Extreme Rainfall Events and Antecedent Rainfall Conditions. Environ Health Perspect. 2019;127(9):97005.

60. 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. Int J Environ Health Res. 2006;16(3):167-80.

61. Effler E, Isaäcson M, Arntzen L, Heenan R, Canter P, Barrett T, et al. Factors contributing to the emergence of Escherichia coli O157 in Africa. Emerg Infect Dis. 2001;7(5):812-9.

62. Kuhn KG, Nielsen EM, Molbak K, Ethelberg S. Epidemiology of campylobacteriosis in Denmark 2000-2015. Zoonoses Public Health. 2018;65(1):59-66.

63. Auld H, MacIver D, Klaassen J. Heavy rainfall and waterborne disease outbreaks: the Walkerton example. J Toxicol Environ Health A. 2004;67(20-22):1879-87.

64. Soneja S, Jiang C, Romeo Upperman C, Murtugudde R, C SM, Blythe D, et al. Extreme precipitation events and increased risk of campylobacteriosis in Maryland, U.S.A. Environ Res. 2016;149:216-21.

65. Guzman Herrador BR, de Blasio BF, MacDonald E, Nichols G, Sudre B, Vold L, et al. Analytical studies assessing the association between extreme precipitation or temperature and drinking water-related waterborne infections: a review. Environ Health. 2015;14:29.

66. Lal A, Baker MG, Hales S, French NP. Potential effects of global environmental changes on cryptosporidiosis and giardiasis transmission. Trends Parasitol. 2013;29(2):83-90.

67. Psoter KJ, De Roos AJ, Wakefield J, Mayer J, Rosenfeld M. Season is associated with Pseudomonas aeruginosa acquisition in young children with cystic fibrosis. Clin Microbiol Infect. 2013;19(11):E483-9.

68. Tandel J, English ED, Sateriale A, Gullicksrud JA, Beiting DP, Sullivan MC, et al. Life cycle progression and sexual development of the apicomplexan parasite Cryptosporidium parvum. Nat Microbiol. 2019;4(12):2226-36.

69. McLauchlin J, Amar C, Pedraza-Diaz S, Nichols GL. Molecular epidemiological analysis of Cryptosporidium spp. in the United Kingdom: results of genotyping Cryptosporidium spp. in 1,705 fecal samples from humans and 105 fecal samples from livestock animals. J Clin Microbiol. 2000;38(11):3984-90.

70. Percival SL, Walker, J.T., & Hunter, P.R. Microbiological Aspects of Biofilms and Drinking Water 1st ed. ed: CRC Press; 2000.

71. Alum A, Absar IM, Asaad H, Rubino JR, Ijaz MK. Impact of environmental conditions on the survival of cryptosporidium and giardia on environmental surfaces. Interdiscip Perspect Infect Dis. 2014;2014:210385.

72. Graczyk TK, Evans BM, Shiff CJ, Karreman HJ, Patz JA. Environmental and geographical factors contributing to watershed contamination with Cryptosporidium parvum oocysts. Environ Res. 2000;82(3):263-71.

73. Robertson L, Gjerde B. Effects of the Norwegian winter environment on Giardia cysts and Cryptosporidium oocysts. Microbial ecology. 2004;47(4):359-65.

74. Mawdsley JL, Brooks AE, Merry RJ. Movement of the protozoan pathogen Cryptosporidium parvum through three contrasting soil types. Biology and Fertility of Soils. 1996;21(1):30-6.

75. Trask JR, Kalita PK, Kuhlenschmidt MS, Smith RD, Funk TL. Overland and near-surface transport of Cryptosporidium parvum from vegetated and nonvegetated surfaces. Journal of Environmental Quality. 2004;33(3):984-93.

76. Ikiroma IA, Pollock KG. Influence of weather and climate on cryptosporidiosis-A review. Zoonoses Public Health. 2021;68(4):285-98.

77. Jagai JS, Castronovo DA, Monchak J, Naumova EN. Seasonality of cryptosporidiosis: A meta-analysis approach. Environ Res. 2009;109(4):465-78.

78. Hu W, Mengersen K, Fu SY, Tong S. The use of ZIP and CART to model cryptosporidiosis in relation to climatic variables. Int J Biometeorol. 2010;54(4):433-40.

79. Britton E, Hales S, Venugopal K, Baker MG. The impact of climate variability and change on cryptosporidiosis and giardiasis rates in New Zealand. J Water Health. 2010;8(3):561-71.

80. 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. J Water Health. 2005;3(2):185-96.

81. Lake IR, Bentham G, Kovats RS, Nichols GL. Effects of weather and river flow on cryptosporidiosis. J Water Health. 2005;3(4):469-74.

82. Nichols G, Lane C, Asgari N, Verlander NQ, Charlett A. Rainfall and outbreaks of drinking water related disease and in England and Wales. J Water Health. 2009;7(1):1-8.

83. Chhetri BK, Takaro TK, Balshaw R, Otterstatter M, Mak S, Lem M, et al. Associations between extreme precipitation and acute gastro-intestinal illness due to cryptosporidiosis and giardiasis in an urban Canadian drinking water system (1997-2009). J Water Health. 2017;15(6):898-907.

84. Semenza JC, Nichols G. Cryptosporidiosis surveillance and water-borne outbreaks in Europe. Euro Surveill. 2007;12(5):E13-4.

85. 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 Rep. 2000;115(5):436-47.

86. Lal A, Hales S, French N, Baker MG. Seasonality in human zoonotic enteric diseases: a systematic review. PLoS One. 2012;7(4):e31883.

87. Lake IR, Pearce J, Savill M. The seasonality of human cryptosporidiosis in New Zealand. Epidemiol Infect. 2008;136(10):1383-7.

88. Lake IR, Harrison FC, Chalmers RM, Bentham G, Nichols G, Hunter PR, et al. Case-control study of environmental and social factors influencing cryptosporidiosis. Eur J Epidemiol. 2007;22(11):805-11.

89. Marston BJ, Lipman HB, Breiman RF. Surveillance for Legionnaires' Disease: Risk Factors for Morbidity and Mortality. Archives of Internal Medicine. 1994;154(21):2417-22.

90. Falkinham JO, 3rd, Norton CD, LeChevallier MW. Factors influencing numbers of Mycobacterium avium, Mycobacterium intracellulare, and other Mycobacteria in drinking water distribution systems. Appl Environ Microbiol. 2001;67(3):1225-31.

91. Gellatly SL, Hancock REW. Pseudomonas aeruginosa : new insights into pathogenesis and host defenses. Pathogens and Disease. 2013;67(3):159-73.

92. Donohue MJ, Wymer L. Increasing Prevalence Rate of Nontuberculous Mycobacteria Infections in Five States, 2008-2013. Ann Am Thorac Soc. 2016;13(12):2143-50.

93. Falkinham JO, 3rd, Iseman MD, de Haas P, van Soolingen D. Mycobacterium avium in a shower linked to pulmonary disease. J Water Health. 2008;6(2):209-13.

94. Hilborn ED, Covert TC, Yakrus MA, Harris SI, Donnelly SF, Rice EW, et al. Persistence of nontuberculous mycobacteria in a drinking water system after addition of filtration treatment. Appl Environ Microbiol. 2006;72(9):5864-9.

95. Perkins SD, Mayfield J, Fraser V, Angenent LT. Potentially pathogenic bacteria in shower water and air of a stem cell transplant unit. Appl Environ Microbiol. 2009;75(16):5363-72.

96. Li T, Abebe LS, Cronk R, Bartram J. A systematic review of waterborne infections from nontuberculous mycobacteria in health care facility water systems. Int J Hyg Environ Health. 2017;220(3):611-20.

97. Stout JE, Yu VL. Legionellosis. New England Journal of Medicine. 1997;337(10):682-7.

98. Viasus D, Di Yacovo S, Garcia-Vidal C, Verdaguer R, Manresa F, Dorca J, et al. Community-acquired Legionella pneumophila pneumonia: a single-center experience with 214 hospitalized sporadic cases over 15 years. Medicine (Baltimore). 2013;92(1):51-60.

99. Donohue MJ. Increasing nontuberculous mycobacteria reporting rates and species diversity identified in clinical laboratory reports. BMC Infectious Diseases. 2018;18(1):163.

100. Hicks LA, Rose CE, Jr., Fields BS, Drees ML, Engel JP, Jenkins PR, et al. Increased rainfall is associated with increased risk for legionellosis. Epidemiol Infect. 2007;135(5):811-7.

101. Karagiannis I, Brandsema P, M VDS. Warm, wet weather associated with increased Legionnaires' disease incidence in The Netherlands. Epidemiol Infect. 2009;137(2):181-7.

102. Halsby KD, Joseph CA, Lee JV, Wilkinson P. The relationship between meteorological variables and sporadic cases of Legionnaires' disease in residents of England and Wales. Epidemiol Infect. 2014;142(11):2352-9.

103. Cassell K, Gacek P, Warren JL, Raymond PA, Cartter M, Weinberger DM. Association Between Sporadic Legionellosis and River Systems in Connecticut. J Infect Dis. 2018;217(2):179-87.

104. Gierhart S, Chukwuma, U. Annual Surveillance Summary: Pseudomonas aeruginosa Infections in the Military Health System (MHS), 2015. Navy and Marine Corps Public Health Center, Department EC; 2017.

105. Hirschboeck KK. Climate and floods. US Geological Survey Water - Supply Paper. 1991;2375:67-88.

106. Doswell CA. Hydrology, Floods and Droughts | Flooding. Encyclopedia of Atmospheric Sciences2015. p. 201-8.

107. Villarini G, Smith JA. Flood peak distributions for the eastern United States. Water Resources Research. 2010;46(6).

108. Teale N. Hydroclimatology and Climate Variability. In: Houser P, editor. Hydroclimatology and Hydrometeorology. The International Encyclopedia of Geography 2020.

109. Mather JR. A History of Hydroclimatology. Physical Geography. 1991;12(3):260-73.

110. Bayliss AC, Jones RC. Peaks-over-threshold flood database: Institute of Hydrology; 1993.

111. Robson A, Reed D. Flood estimation handbook: statistical procedures for flood frequency estimation: Institute of Hydrology; 1999.

112. Andrade L, O'Dwyer J, O'Neill E, Hynds P. Surface water flooding, groundwater contamination, and enteric disease in developed countries: A scoping review of connections and consequences. Environ Pollut. 2018;236:540-9.

113. Rose JB, Daeschner S, Easterling DR, Curriero FC, Lele S, Patz JA. Climate and waterborne disease outbreaks. Journal AWWA. 2000;92(9):77-87.

114. Semenza JC, Menne B. Climate change and infectious diseases in Europe. Lancet Infect Dis. 2009;9(6):365-75.

115. Monmonier M. Cartographies of danger: Mapping hazards in America. Chicago: University of Chicago Press; 1997.

116. Parker JK, McIntyre D, Noble RT. Characterizing fecal contamination in stormwater runoff in coastal North Carolina, USA. Water Res. 2010;44(14):4186-94.

117. Presley SM, Rainwater TR, Austin GP, Platt SG, Zak JC, Cobb GP, et al. Assessment of pathogens and toxicants in New Orleans, LA following Hurricane Katrina. Environ Sci Technol. 2006;40(2):468-74.

118. Bush KF, Fossani CL, Li S, Mukherjee B, Gronlund CJ, O'Neill MS. Extreme precipitation and beach closures in the great lakes region: evaluating risk among the elderly. Int J Environ Res Public Health. 2014;11(2):2014-32.

119. Patz JA, Vavrus SJ, Uejio CK, McLellan SL. Climate change and waterborne disease risk in the Great Lakes region of the U.S. Am J Prev Med. 2008;35(5):451-8.

120. Smith J, Villarini G, Baeck M. Mixture Distributions and the Hydroclimatology of Extreme Rainfall and Flooding in the Eastern United States. Journal of Hydrometeorology. 2011;12.

121. Lecce SA. Seasonality of Flooding in North Carolina. Southeastern Geographer. 2000;40(2):168-75.

122. Gamble DW. The Relationship between Drainage Basin Area and Annual Peak-Flood Seasonality in the Southeastern United States. Southeastern Geographer. 1997;37(1):61-75.

123. Peterson TC, Heim RR, Hirsch R, Kaiser DP, Brooks H, Diffenbaugh NS, et al. Monitoring and Understanding Changes in Heat Waves, Cold Waves, Floods, and Droughts in the United States: State of Knowledge. Bulletin of the American Meteorological Society. 2013;94(6):821-34.

124. Adams DK, Comrie AC. The north American monsoon. Bulletin of the American Meteorological Society. 1997;78(10):2197-214.

125. Dettinger MD. Climate change, atmospheric rivers, and floods in California - a multimodel analysis of storm frequency and magnitude changes. Journal of the American Water Resources Association. 2011;47(3):514-23.

126. de Man H, van den Berg HH, Leenen EJ, Schijven JF, Schets FM, van der Vliet JC, et al. Quantitative assessment of infection risk from exposure to waterborne pathogens in urban floodwater. Water Res. 2014;48:90-9.

127. Gibson CJ, Stadterman KL, States S, Sykora J. Combined sewer overflows: A source of Cryptosporidium and Giardia? Water Science and Technology. 1998;38(12):67-72.

128. Jamieson RC, Joy DM, Lee H, Kostaschuk R, Gordon RJ. Resuspension of sediment-associated Escherichia coli in a natural stream. J Environ Qual. 2005;34(2):581-9.

129. Schalk JA, Docters van Leeuwen AE, Lodder WJ, de Man H, Euser S, den Boer JW, et al. Isolation of Legionella pneumophila from pluvial floods by amoebal coculture. Appl Environ Microbiol. 2012;78(12):4519-21.

130. Tolouei S, Burnet JB, Autixier L, Taghipour M, Bonsteel J, Duy SV, et al. Temporal variability of parasites, bacterial indicators, and wastewater micropollutants in a water resource recovery facility under various weather conditions. Water Res. 2019;148:446-58.

131. Madoux-Humery AS, Dorner S, Sauve S, Aboulfadl K, Galarneau M, Servais P, et al. Temporal variability of combined sewer overflow contaminants: evaluation of wastewater micropollutants as tracers of fecal contamination. Water Res. 2013;47(13):4370-82.

132. 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.

133. Wade TJ, Lin CJ, Jagai JS, Hilborn ED. Flooding and emergency room visits for gastrointestinal illness in Massachusetts: a case-crossover study. PLoS One. 2014;9(10):e110474.

134. Jalliffier-Verne I, Leconte R, Huaringa-Alvarez U, Madoux-Humery AS, Galarneau M, Servais P, et al. Impacts of global change on the concentrations and dilution of combined sewer overflows in a drinking water source. Sci Total Environ. 2015;508:462-76.

135. Loenenbach AD, Beulens C, Euser SM, van Leuken JPG, Bom B, van der Hoek W, et al. Two Community Clusters of Legionnaires' Disease Directly Linked to a Biologic Wastewater Treatment Plant, the Netherlands. Emerg Infect Dis. 2018;24(10):1914-8.

136. Cooley M, Carychao D, Crawford-Miksza L, Jay MT, Myers C, Rose C, et al. Incidence and tracking of Escherichia coli O157:H7 in a major produce production region in California. PLoS One. 2007;2(11):e1159.

137. Vereen E, Jr., Lowrance RR, Jenkins MB, Adams P, Rajeev S, Lipp EK. Landscape and seasonal factors influence Salmonella and Campylobacter prevalence in a rural mixed use watershed. Water Res. 2013;47(16):6075-85.

138. Haley BJ, Cole DJ, Lipp EK. Distribution, diversity, and seasonality of waterborne salmonellae in a rural watershed. Appl Environ Microbiol. 2009;75(5):1248-55.

139. Rodriguez S, Araujo R. Occurrence of thermotolerant Campylobacter species in surface waters of a Mediterranean area and in its prevailing pollution sources. J Appl Microbiol. 2010;109(3):1027-34.

140. Miller WA, Lewis DJ, Lennox M, Pereira MG, Tate KW, Conrad PA, et al. Climate and on-farm risk factors associated with Giardia duodenalis cysts in storm runoff from California coastal dairies. Appl Environ Microbiol. 2007;73(21):6972-9.

141. Miller WA, Lewis DJ, Pereira MD, Lennox M, Conrad PA, Tate KW, et al. Farm factors associated with reducing Cryptosporidium loading in storm runoff from dairies. J Environ Qual. 2008;37(5):1875-82.

142. Wing S, Freedman S, Band L. The potential impact of flooding on confined animal feeding operations in eastern North Carolina. Environ Health Perspect. 2002;110(4):387-91.

143. Erickson MC, Habteselassie MY, Liao J, Webb CC, Mantripragada V, Davey LE, et al. Examination of factors for use as potential predictors of human enteric pathogen survival in soil. J Appl Microbiol. 2014;116(2):335-49.

144. Mallin MA. Impacts of Industrial Animal Production on Rivers and Estuaries. American Scientist. 2000;88:26.

145. Weller D, Brassill N, Rock C, Ivanek R, Mudrak E, Roof S, et al. Complex Interactions Between Weather, and Microbial and Physicochemical Water Quality Impact the Likelihood of Detecting Foodborne Pathogens in Agricultural Water. Front Microbiol. 2020;11:134.

146. Luo Z, Gu G, Ginn A, Giurcanu MC, Adams P, Vellidis G, et al. Distribution and Characterization of Salmonella enterica Isolates from Irrigation Ponds in the Southeastern United States. Appl Environ Microbiol. 2015;81(13):4376-87.

147. Rosef O, Rettedal G, Lageide L. Thermophilic campylobacters in surface water: a potential risk of campylobacteriosis. Int J Environ Health Res. 2001;11(4):321-7.

148. Holvoet K, Sampers I, Seynnaeve M, Uyttendaele M. Relationships among hygiene indicators and enteric pathogens in irrigation water, soil and lettuce and the impact of climatic conditions on contamination in the lettuce primary production. International Journal of Food Microbiology. 2014;171:21-31.

149. Oda T KM, Uga S. Detection of Giardia cysts in sewage and estimations of giardiasis prevalence among inhabitants in Hyogo Prefecture, Japan. Tropical Medicine and Health. 2005;33(1):1-5.

150. Caicedo C, Beutel S, Scheper T, Rosenwinkel KH, Nogueira R. Occurrence of Legionella in wastewater treatment plants linked to wastewater characteristics. Environmental Science and Pollution Research. 2016;23(16):16873-81.

151. Rechenburg A, Kistemann T. Sewage effluent as a source of Campylobacter sp. in a surface water catchment. Int J Environ Health Res. 2009;19(4):239-49.

152. Lemarchand K, Lebaron P. Occurrence of Salmonella spp and Cryptosporidium spp in a French coastal watershed: relationship with fecal indicators. FEMS Microbiol Lett. 2003;218(1):203-9.

153. Madoux-Humery AS, Dorner SM, Sauve S, Aboulfadl K, Galarneau M, Servais P, et al. Temporal analysis of E. coli, TSS and wastewater micropollutant loads from combined sewer overflows: implications for management. Environ Sci Process Impacts. 2015;17(5):965-74.

154. Passer JK, Danila RN, Laine ES, Como-Sabetti KJ, Tang W, Searle KM. The association between sporadic Legionnaires' disease and weather and environmental factors, Minnesota, 2011-2018. Epidemiol Infect. 2020;148:e156.

155. EPA U. Primer for municipal wastewater treatment systems. OoW Management. 2004.

156. Su X, Liu T, Beheshti M, Prigiobbe V. Relationship between infiltration, sewer rehabilitation, and groundwater flooding in coastal urban areas. Environmental Science and Pollution Research. 2020;27(13):14288-98.

157. Tibbetts J. Combined sewer systems: down, dirty, and out of date. Environ Health Perspect. 2005;113(7):A464-7.

158. Brière FG. Drinking-water distribution, sewage, and rainfall collection: Presses inter Polytechnique; 2014.

159. Donovan EP, Staskal DF, Unice KM, Roberts JD, Haws LC, Finley BL, et al. Risk of gastrointestinal disease associated with exposure to pathogens in the sediments of the Lower Passaic River. Appl Environ Microbiol. 2008;74(4):1004-18.

160. Jagai JS, DeFlorio-Barker S, Lin CJ, Hilborn ED, Wade TJ. Sanitary Sewer Overflows and Emergency Room Visits for Gastrointestinal Illness: Analysis of Massachusetts Data, 2006-2007. Environ Health Perspect. 2017;125(11):117007.

161. Jagai JS, Li Q, Wang S, Messier KP, Wade TJ, Hilborn ED. Extreme Precipitation and Emergency Room Visits for Gastrointestinal Illness in Areas with and without Combined Sewer Systems: An Analysis of Massachusetts Data, 2003-2007. Environ Health Perspect. 2015;123(9):873-9.

162. Hummel M, Hummel M, Berry M, Berry M, Stacey M, Stacey M. Sea Level Rise Impacts on Wastewater Treatment Systems Along the U.S. Coasts. Earth's Future. 2018;6(4).

163. Friedrich E, Kretzinger D. Vulnerability of wastewater infrastructure of coastal cities to sea level rise: A South African case study. Water SA. 2011;38:755-64.

164. Edge TA, Khan IU, Bouchard R, Guo J, Hill S, Locas A, et al. Occurrence of waterborne pathogens and Escherichia coli at offshore drinking water intakes in lake Ontario. Appl Environ Microbiol. 2013;79(19):5799-813.

165. Fewtrell L, Kay D, Watkins J, Davies C, Francis C. The microbiology of urban UK floodwaters and a quantitative microbial risk assessment of flooding and gastrointestinal illness. Journal of Flood Risk Management. 2011;4(2):77-87.

166. DeFlorio-Barker S, Wing C, Jones RM, Dorevitch S. Estimate of incidence and cost of recreational waterborne illness on United States surface waters. Environ Health. 2018;17(1):3.

167. Graciaa DS, Cope JR, Roberts VA, Cikesh BL, Kahler AM, Vigar M, et al. Outbreaks Associated with Untreated Recreational Water - United States, 2000-2014. MMWR Morb Mortal Wkly Rep. 2018;67(25):701-6.

168. Hlavsa MC, Cikesh BL, Roberts VA, Kahler AM, Vigar M, Hilborn ED, et al. Outbreaks Associated with Treated Recreational Water - United States, 2000-2014. MMWR Morb Mortal Wkly Rep. 2018;67(19):547-51.

169. Denno DM, Keene WE, Hutter CM, Koepsell JK, Patnode M, Flodin-Hursh D, et al. Tri-county comprehensive assessment of risk factors for sporadic reportable bacterial enteric infection in children. J Infect Dis. 2009;199(4):467-76.

170. Eze JI, Scott EM, Pollock KG, Stidson R, Miller CA, Lee D. The association of weather and bathing water quality on the incidence of gastrointestinal illness in the west of Scotland. Epidemiol Infect. 2014;142(6):1289-99.

171. Till D, McBride G, Ball A, Taylor K, Pyle E. Large-scale freshwater microbiological study: rationale, results and risks. J Water Health. 2008;6(4):443-60.

172. Ravel A, Pintar K, Nesbitt A, Pollari F. Non food-related risk factors of campylobacteriosis in Canada: a matched case-control study. BMC Public Health. 2016;16(1):1016.

173. Martinez-Urtaza J, Saco M, de Novoa J, Perez-Pineiro P, Peiteado J, Lozano-Leon A, et al. Influence of environmental factors and human activity on the presence of Salmonella serovars in a marine environment. Appl Environ Microbiol. 2004;70(4):2089-97.

174. Cho S, Hiott LM, Barrett JB, McMillan EA, House SL, Humayoun SB, et al. Prevalence and characterization of Escherichia coli isolated from the Upper Oconee Watershed in Northeast Georgia. PLoS One. 2018;13(5):e0197005.

175. Schuster CJ, Ellis AG, Robertson WJ, Charron DF, Aramini JJ, Marshall BJ, et al. Infectious disease outbreaks related to drinking water in Canada, 1974-2001. Can J Public Health. 2005;96(4):254-8.

176. Hrudey SE, Payment P, Huck PM, Gillham RW, Hrudey EJ. A fatal waterborne disease epidemic in Walkerton, Ontario: comparison with other waterborne outbreaks in the developed world. Water Sci Technol. 2003;47(3):7-14.

177. Clark CG, Price L, Ahmed R, Woodward DL, Melito PL, Rodgers FG, et al. Characterization of waterborne outbreak-associated Campylobacter jejuni, Walkerton, Ontario. Emerg Infect Dis. 2003;9(10):1232-41.

178. Le Dantec C, Duguet JP, Montiel A, Dumoutier N, Dubrou S, Vincent V. Occurrence of mycobacteria in water treatment lines and in water distribution systems. Appl Environ Microbiol. 2002;68(11):5318-25.

179. Reynolds KA, Mena KD, Gerba CP. Risk of waterborne illness via drinking water in the United States. Rev Environ Contam Toxicol. 2008;192:117-58.

180. Rizak S, Hrudey SE. Drinking-water safety: challenges for community-managed systems. J Water Health. 2008;6 Suppl 1:33-41.

181. (ASCE) ASCE. Infrastructure report card: Drinking Water. 2017.

182. Tiemann M. Safe Drinking Water Act: A Summary of the Act and Its Major Requirements. 2014.

183. EPA. Private Drinking Water Wells: Environmental Protection Agency 2022 [Available from: https://www.epa.gov/privatewells.

184. Murray A, Hall A, Weaver J, Kremer F. Methods for Estimating Locations of Housing Units Served by Private Domestic Wells in the United States Applied to 2010. JAWRA Journal of the American Water Resources Association. 2021;57(5):828-43.

185. Maupin MA, Kenny JF, Hutson SS, Lovelace JK, Barber NL, Linsey KS. Estimated use of water in the United States in 2010. Report. Reston, VA; 2014. Report No.: 1405.

186. Johnson TD, Belitz K, Lombard M. Estimating domestic well locations and populations served in the contiguous U.S. for years 2000 and 2010. Science of the Total Environment. 2019;687:1261-73.

187. Craun GF, Brunkard JM, Yoder JS, Roberts VA, Carpenter J, Wade T, et al. Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. Clin Microbiol Rev. 2010;23(3):507-28.

188. EPA. Ground Water Rule: Environmental Protection Agency; 2022 [Available from: https://www.epa.gov/dwreginfo/ground-water-rule.

189. EPA. Surface Water Treatment Rules Environmental Protection Agency2022 [Available from: https://www.epa.gov/dwreginfo/surface-water-treatment-rules.

190. National primary drinking water regulations: Long Term 1 Enhanced Surface Water Treatment Rule. Final rule. Fed Regist. 2002;67(9):1811-44.

191. Moreira NA, Bondelind M. Safe drinking water and waterborne outbreaks. J Water Health. 2017;15(1):83-96.

192. Pike WA. Modeling Drinking Water Quality Violations with Bayesian Networks. JAWRA Journal of the American Water Resources Association. 2004;40(6):1563-78.

193. Ailes E, Budge P, Shankar M, Collier S, Brinton W, Cronquist A, et al. Economic and health impacts associated with a Salmonella Typhimurium drinking water outbreak-Alamosa, CO, 2008. PLoS One. 2013;8(3):e57439.

194. EPA US. The Data Management and Quality Assurance/Quality Control Process for the Third Six-Year Review Information Collection Rule Dataset. EPA-810-R-16-015 2016 [Available from: https://www.epa.gov/sites/production/files/2016-12/documents/810r16015_0.pdf.

195. LeChevallier MW, Evans TM, Seidler RJ. Effect of turbidity on chlorination efficiency and bacterial persistence in drinking water. Appl Environ Microbiol. 1981;42(1):159-67.

196. Yoder JS, Hlavsa MC, Craun GF, Hill V, Roberts V, Yu PA, et al. Surveillance for waterborne disease and outbreaks associated with recreational water use and other aquatic facility-associated health events--United States, 2005-2006. MMWR Surveill Summ. 2008;57(9):1-29.

197. Barron G, Buchanan S, Hase D, Mainzer H, Ransom MM, Sarisky J. New approaches to safe drinking water. J Law Med Ethics. 2002;30(3 Suppl):105-8.

198. Laufenberg SD. The struggle of cities to implement the Safe Drinking Water Act in the context of intergovernmental relations. Drake J Agric L. 1998;3:495.

199. Uhlmann S, Galanis E, Takaro T, Mak S, Gustafson L, Embree G, et al. Where's the pump? Associating sporadic enteric disease with drinking water using a geographic information system, in British Columbia, Canada, 1996-2005. J Water Health. 2009;7(4):692-8.

200. Clarkson LS, Tobin-D'Angelo M, Shuler C, Hanna S, Benson J, Voetsch AC. Sporadic Salmonella enterica serotype Javiana infections in Georgia and Tennessee: a hypothesis-generating study. Epidemiol Infect. 2010;138(3):340-6.

201. Olsen SJ, Miller G, Breuer T, Kennedy M, Higgins C, Walford J, et al. A waterborne outbreak of Escherichia coli O157:H7 infections and hemolytic uremic syndrome: implications for rural water systems. Emerg Infect Dis. 2002;8(4):370-5.

202. Gelting R, Sarisky J, Selman C, Otto C, Higgins C, Bohan PO, et al. Use of a systems-based approach to an environmental health assessment for a waterborne disease outbreak investigation at a snowmobile lodge in Wyoming. Int J Hyg Environ Health. 2005;208(1-2):67-73.

203. LeChevallier MW, Au K-K. Water Treatment and Pathogen Control: Process Efficiency in Achieving Safe Drinking-water: IWA Publishing; 2013. Available from: https://doi.org/10.2166/9781780405858.

204. Barskey AE, Derado G, Edens C. Rising Incidence of Legionnaires' Disease and Associated Epidemiologic Patterns, United States, 1992-2018. Emerg Infect Dis. 2022;28(3):527-38.

205. Marston BJ, Plouffe JF, File TM, Jr, Hackman BA, Salstrom S-J, Lipman HB, et al. Incidence of Community-Acquired Pneumonia Requiring Hospitalization: Results of a

Population-Based Active Surveillance Study in Ohio. Archives of Internal Medicine. 1997;157(15):1709-18.

206. Cunha BA, Burillo A, Bouza E. Legionnaires' disease. Lancet. 2016;387(10016):376-85.

207. Masak J, Cejkova A, Schreiberova O, Rezanka T. Pseudomonas biofilms: possibilities of their control. FEMS Microbiol Ecol. 2014;89(1):1-14.

208. van Heijnsbergen E, Schalk JA, Euser SM, Brandsema PS, den Boer JW, de Roda Husman AM. Confirmed and Potential Sources of Legionella Reviewed. Environ Sci Technol. 2015;49(8):4797-815.

209. Hellberg RS, Chu E. Effects of climate change on the persistence and dispersal of foodborne bacterial pathogens in the outdoor environment: A review. Crit Rev Microbiol. 2016;42(4):548-72.

210. Eisenberg JN, Desai MA, Levy K, Bates SJ, Liang S, Naumoff K, et al. Environmental determinants of infectious disease: a framework for tracking causal links and guiding public health research. Environ Health Perspect. 2007;115(8):1216-23.

211. Kitajima M, Haramoto E, Iker BC, Gerba CP. Occurrence of Cryptosporidium, Giardia, and Cyclospora in influent and effluent water at wastewater treatment plants in Arizona. Sci Total Environ. 2014;484:129-36.

212. Robertson LJ, Campbell AT, Smith HV. Survival of Cryptosporidium parvum oocysts under various environmental pressures. Appl Environ Microbiol. 1992;58(11):3494-500.

213. Levy K, Woster AP, Goldstein RS, Carlton EJ. Untangling the Impacts of Climate Change on Waterborne Diseases: a Systematic Review of Relationships between Diarrheal Diseases and Temperature, Rainfall, Flooding, and Drought. Environ Sci Technol. 2016;50(10):4905-22.

214. Curriero FC, Patz JA, Rose JB, Lele S. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994. Am J Public Health. 2001;91(8):1194-9.

215. 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. Am J Epidemiol. 2014;179(3):344-52.

216. Carlton EJ, Woster AP, DeWitt P, Goldstein RS, Levy K. A systematic review and meta-analysis of ambient temperature and diarrhoeal diseases. Int J Epidemiol. 2016;45(1):117-30.

217. Kovats RS, Edwards SJ, Charron D, Cowden J, D'Souza RM, Ebi KL, et al. Climate variability and campylobacter infection: an international study. Int J Biometeorol. 2005;49(4):207-14.

218. Galway LP, Allen DM, Parkes MW, Takaro TK. Seasonal variation of acute gastro-intestinal illness by hydroclimatic regime and drinking water source: a retrospective population-based study. J Water Health. 2014;12(1):122-35.

219. Bertuzzo E, Mari L, Righetto L, Gatto M, Casagrandi R, Rodriguez-Iturbe I, et al. Hydroclimatology of dual-peak annual cholera incidence: Insights from a spatially explicit model. Geophysical Research Letters. 2012;39(5):n/a-n/a.

220. Fong TT, Mansfield LS, Wilson DL, Schwab DJ, Molloy SL, Rose JB. Massive microbiological groundwater contamination associated with a waterborne outbreak in Lake Erie, South Bass Island, Ohio. Environ Health Perspect. 2007;115(6):856-64.

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

222. Gleason JA, Fagliano JA. Effect of drinking water source on associations between gastrointestinal illness and heavy rainfall in New Jersey. PLoS One. 2017;12(3):e0173794.

223. Jean JS, Guo HR, Chen SH, Liu CC, Chang WT, Yang YJ, et al. The association between rainfall rate and occurrence of an enterovirus epidemic due to a contaminated well. J Appl Microbiol. 2006;101(6):1224-31.

224. Melillo JM, Richmond, T.C., Yohe, G.W. 2014: Highlights of Climate Change Impacts in the United States: The Third National Climate Assessment. US Global Change Research Program. 2014;Retrieved from: https://nca2014.globalchange.gov/downloads/low/NCA3_Highlights_LowRes.pdf.

225. Marciano-Cabral F, Jamerson M, Kaneshiro ES. Free-living amoebae, Legionella and Mycobacterium in tap water supplied by a municipal drinking water utility in the USA. J Water Health. 2010;8(1):71-82.

226. TDI. Research Methods. The Dartmouth Atlast of Health Care The Dartmouth Institute for Health Policy and Clinical Practice. 2020; https://www.dartmouthatlas.org/downloads/methods/research_methods.pdf [accessed 15 November 2020].

227. Gomez A. Rural Water Infrastructure: Federal Agencies Provide Funding but Could Increase Coordination to Help Communities. United States Government Accountability Office; 2015.

228. Metcalf CJE, Walter KS, Wesolowski A, Buckee CO, Shevliakova E, Tatem AJ, et al. Identifying climate drivers of infectious disease dynamics: recent advances and challenges ahead. Proc Biol Sci. 2017;284(1860).

229. Pitzer VE, Viboud C, Alonso WJ, Wilcox T, Metcalf CJ, Steiner CA, et al. Environmental Drivers of the Spatiotemporal Dynamics of Respiratory Syncytial Virus in the United States. PLOS Pathogens. 2015;11(1):e1004591.

230. Atherholt TB, LeChevallier MW, Norton WD, Rosen JS. Effect of rainfall on giardia and crypto. American Water Works Association. 1998;90(9):66-80.

231. Lakshmi V, Piechota T, Narayan U, Tang C. Soil moisture as an indicator of weather extremes. Geophysical Research Letters. 2004;31(11):n/a-n/a.

232. Parinussa RM, Lakshmi V, Johnson FM, Sharma A. A new framework for monitoring flood inundation using readily available satellite data. Geophysical Research Letters. 2016;43(6):2599-605.

233. Rappaport EN. Fatalities in the United States from Atlantic Tropical Cyclones: New Data and Interpretation. Bulletin of the American Meteorological Society. 2014;95(3):341-6.

234. Parks RM, Anderson GB, Nethery RC, Navas-Acien A, Dominici F, Kioumourtzoglou MA. Tropical cyclone exposure is associated with increased hospitalization rates in older adults. Nat Commun. 2021;12(1):1545.

235. Erickson TB, Brooks J, Nilles EJ, Pham PN, Vinck P. Environmental health effects attributed to toxic and infectious agents following hurricanes, cyclones, flash floods and major hydrometeorological events. J Toxicol Environ Health B Crit Rev. 2019;22(5-6):157-71.

236. Young I, Smith BA, Fazil A. A systematic review and meta-analysis of the effects of extreme weather events and other weather-related variables on Cryptosporidium and Giardia in fresh surface waters. J Water Health. 2015;13(1):1-17.

237. Cann KF, Thomas DR, Salmon RL, Wyn-Jones AP, Kay D. Extreme water-related weather events and waterborne disease. Epidemiol Infect. 2013;141(4):671-86.

238. Sauer EP, Vandewalle JL, Bootsma MJ, McLellan SL. Detection of the human specific Bacteroides genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. Water Res. 2011;45(14):4081-91.

239. Wade TJ, Sandhu SK, Levy D, Lee S, LeChevallier MW, Katz L, et al. Did a severe flood in the Midwest cause an increase in the incidence of gastrointestinal symptoms? Am J Epidemiol. 2004;159(4):398-405.

240. Beaudeau P, Schwartz J, Levin R. Drinking water quality and hospital admissions of elderly people for gastrointestinal illness in Eastern Massachusetts, 1998-2008. Water Res. 2014;52:188-98.

241. Den Boer JW, Coutinho RA, Yzerman EP, van der Sande MA. Use of surface water in drinking water production associated with municipal Legionnaires' disease incidence. J Epidemiol Community Health. 2008;62(4):e1.

242. Rangel JM, Sparling PH, Crowe C, Griffin PM, Swerdlow DL. Epidemiology of Escherichia coli O157:H7 outbreaks, United States, 1982-2002. Emerg Infect Dis. 2005;11(4):603-9.

243. Bloom MS, Palumbo J, Saiyed N, Lauper U, Lin S. Food and Waterborne Disease in the Greater New York City Area Following Hurricane Sandy in 2012. Disaster Med Public Health Prep. 2016;10(3):503-11.

244. Lefebvre M, Razakandrainibe R, Villena I, Favennec L, Costa D. Cryptosporidium-Biofilm Interactions: a Review. Appl Environ Microbiol. 2021;87(3).

245. Berendt RF. Survival of Legionella pneumophila in Aerosols: Effect of Relative Humidity. The Journal of Infectious Diseases. 1980;141(5):689-.

246. Breiman RF, Butler JC. Legionnaires' disease: clinical, epidemiological, and public health perspectives. Semin Respir Infect. 1998;13(2):84-9.

247. Sidhu JP, Hodgers L, Ahmed W, Chong MN, Toze S. Prevalence of human pathogens and indicators in stormwater runoff in Brisbane, Australia. Water Res. 2012;46(20):6652-60.

248. Collinet-Adler S, Ward HD. Cryptosporidiosis: environmental, therapeutic, and preventive challenges. Eur J Clin Microbiol Infect Dis. 2010;29(8):927-35.

249. Anderson GB, Ferreri J, Al-Hamdan M, Crosson W, Schumacher A, Guikema S, et al. Assessing United States County-Level Exposure for Research on Tropical Cyclones and Human Health. Environ Health Perspect. 2020;128(10):107009.

250. Kruk MC, Gibney EJ, Levinson DH, Squires M. A Climatology of Inland Winds from Tropical Cyclones for the Eastern United States. Journal of Applied Meteorology and Climatology. 2010;49(7):1538-47.

251. Radcliff TA, Chu K, Der-Martirosian C, Dobalian A. A Model for Measuring Ambulatory Access to Care Recovery after Disasters. J Am Board Fam Med. 2018;31(2):252-9.

252. Noe RS, Schnall AH, Wolkin AF, Podgornik MN, Wood AD, Spears J, et al. Disaster-related injuries and illnesses treated by American Red Cross disaster health services during Hurricanes Gustav and Ike. South Med J. 2013;106(1):102-8.

253. Armstrong BG, Gasparrini A, Tobias A. Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis. BMC Medical Research Methodology. 2014;14(1):122.

254. Ulrich N, Rosenberger A, Brislawn C, Wright J, Kessler C, Toole D, et al. Restructuring of the Aquatic Bacterial Community by Hydric Dynamics Associated with Superstorm Sandy. Appl Environ Microbiol. 2016;82(12):3525-36.

255. Brigmon RL, Turick CE, Knox AS, Burckhalter CE. The Impact of Storms on Legionella pneumophila in Cooling Tower Water, Implications for Human Health. Front Microbiol. 2020;11:543589.

256. Health TDo. Epidemiology in Texas 1998 Annual Report. Texas Department of Health; 1999.

257. Saulnier DD, Brolin Ribacke K, von Schreeb J. No Calm After the Storm: A Systematic Review of Human Health Following Flood and Storm Disasters. Prehosp Disaster Med. 2017;32(5):568-79.

258. Shukla MA, Woc-Colburn L, Weatherhead JE. Infectious Diseases in the Aftermath of Hurricanes in the United States. Current Tropical Medicine Reports. 2018;5(4):217-23.

259. Lane K, Charles-Guzman K, Wheeler K, Abid Z, Graber N, Matte T. Health effects of coastal storms and flooding in urban areas: a review and vulnerability assessment. J Environ Public Health. 2013;2013:913064.

260. Carroll RJ, Ruppert, D., Stefanski, L.A., & Crainiceanu, C.M. . Measurement Error in Nonlinear Models: A Modern Perspective. Second Edition (2nd ed.) ed: Chapman and Hall/CRC.; 2006.

261. USDA. Rural Urban Continuum Codes (RUCC). In: Service USDoAER, editor. 2013.

262. Fraser DW, Tsai TR, Orenstein W, Parkin WE, Beecham HJ, Sharrar RG, et al. Legionnaires' disease: description of an epidemic of pneumonia. N Engl J Med. 1977;297(22):1189-97.

263. CDC. Legionnaires' Disease Surveillance Summary Report, United States 2016 - 2017. https://wwwcdcgov/legionella/health-depts/surv-reporting/2016-17-surv-report-508pdf. 2020.

264. Neil K, Berkelman R. Increasing incidence of legionellosis in the United States, 1990-2005: changing epidemiologic trends. Clin Infect Dis. 2008;47(5):591-9.

265. Phin N, Parry-Ford F, Harrison T, Stagg HR, Zhang N, Kumar K, et al. Epidemiology and clinical management of Legionnaires' disease. Lancet Infect Dis. 2014;14(10):1011-21.

266. Fisman DN, Lim S, Wellenius GA, Johnson C, Britz P, Gaskins M, et al. It's not the heat, it's the humidity: wet weather increases legionellosis risk in the greater Philadelphia metropolitan area. J Infect Dis. 2005;192(12):2066-73.

267. Ng V, Tang P, Jamieson F, Drews SJ, Brown S, Low DE, et al. Going with the flow: legionellosis risk in Toronto, Canada is strongly associated with local watershed hydrology. Ecohealth. 2008;5(4):482-90.

268. States SJ, Conley LF, Kuchta JM, Oleck BM, Lipovich MJ, Wolford RS, et al. Survival and multiplication of Legionella pneumophila in municipal drinking water systems. Appl Environ Microbiol. 1987;53(5):979-86.

269. Fliermans CB, Cherry WB, Orrison LH, Smith SJ, Tison DL, Pope DH. Ecological distribution of Legionella pneumophila. Applied and Environmental Microbiology. 1981;41(1):9-16.

270. Declerck P. Biofilms: the environmental playground of Legionella pneumophila. Environ Microbiol. 2010;12(3):557-66.

271. Fields BS, Benson RF, Besser RE. Legionella and Legionnaires' disease: 25 years of investigation. Clin Microbiol Rev. 2002;15(3):506-26.

272. Edelstein PH, Cianciotto NP. Legionella Species and Legionnaires' Disease. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. The Prokaryotes: A Handbook on the Biology of Bacteria Volume 6: Proteobacteria: Gamma Subclass. New York, NY: Springer New York; 2006. p. 988-1033.

273. Declerck P, Behets J, van Hoef V, Ollevier F. Detection of Legionella spp. and some of their amoeba hosts in floating biofilms from anthropogenic and natural aquatic environments. Water Res. 2007;41(14):3159-67.

274. Rogers J, Dowsett AB, Dennis PJ, Lee JV, Keevil CW. Influence of temperature and plumbing material selection on biofilm formation and growth of Legionella pneumophila in a model potable water system containing complex microbial flora. Appl Environ Microbiol. 1994;60(5):1585-92.

275. Stout JE, Yu VL, Yee YC, Vaccarello S, Diven W, Lee TC. Legionella pneumophila in residential water supplies: environmental surveillance with clinical assessment for Legionnaires' disease. Epidemiol Infect. 1992;109(1):49-57.

276. Best M, Yu VL, Stout J, Goetz A, Muder RR, Taylor F. Legionellaceae in the hospital water-supply. Epidemiological link with disease and evaluation of a method for control of nosocomial legionnaires' disease and Pittsburgh pneumonia. Lancet. 1983;2(8345):307-10.

277. Garrison LE, Kunz JM, Cooley LA, Moore MR, Lucas C, Schrag S, et al. Vital Signs: Deficiencies in Environmental Control Identified in Outbreaks of Legionnaires' Disease - North America, 2000-2014. MMWR Morb Mortal Wkly Rep. 2016;65(22):576-84.

278. Walczak M, Kletkiewicz H, Burkowska A. Occurrence of Legionella pneumophila in lakes serving as a cooling system of a power plant. Environ Sci Process Impacts. 2013;15(12):2273-8.

279. Morton S, Bartlett CL, Bibby LF, Hutchinson DN, Dyer JV, Dennis PJ. Outbreak of legionnaires' disease from a cooling water system in a power station. Br J Ind Med. 1986;43(9):630-5.

280. Sakamoto R, Ohno A, Nakahara T, Satomura K, Iwanaga S, Kouyama Y, et al. Legionella pneumophila in rainwater on roads. Emerg Infect Dis. 2009;15(8):1295-7.

281. Brandsema PS, Euser SM, Karagiannis I, Den Boer JW, Van Der Hoek W. Summer increase of Legionnaires' disease 2010 in The Netherlands associated with weather conditions and implications for source finding. Epidemiol Infect. 2014;142(11):2360-71.

282. Ricketts KD, Charlett A, Gelb D, Lane C, Lee JV, Joseph CA. Weather patterns and Legionnaires' disease: a meteorological study. Epidemiol Infect. 2009;137(7):1003-12.

283. Greene SK, Wilson EL, Konty KJ, Fine AD. Assessment of reportable disease incidence after Hurricane Sandy, New York City, 2012. Disaster Med Public Health Prep. 2013;7(5):513-21.

284. Ridpath AD, Bregman B, Jones L, Reddy V, Waechter H, Balter S. Challenges to implementing communicable disease surveillance in New York City evacuation shelters after Hurricane Sandy, November 2012. Public Health Rep. 2015;130(1):48-53.

285. Sinigalliano CD, Gidley ML, Shibata T, Whitman D, Dixon TH, Laws E, et al. Impacts of Hurricanes Katrina and Rita on the microbial landscape of the New Orleans area. Proc Natl Acad Sci U S A. 2007;104(21):9029-34.

286. Arnell NW, Gosling SN. The impacts of climate change on river flood risk at the global scale. Climatic Change. 2014;134(3):387-401.

287. Hirsch RM, Ryberg KR. Has the magnitude of floods across the USA changed with global CO2 levels? Hydrological Sciences Journal. 2012;57(1):1-9.

288. Georgakakos A, Fleming P, Dettinger M, Peters-Lidard C, Richmond T, Reckhow K, et al. Ch. 3: Water Resources. Climate Change Impacts in the United States: The Third National Climate Assessment, J. M. Melillo, Terse (T.C.) Richmond, and G.W. Yohe, Eds. U.S. Global Change Research Program. 2014:69-112.

289. Smith AB, Katz RW. US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases. Natural Hazards. 2013;67(2):387-410.

290. Mocko DNGH. NLDAS Primary Forcing Data L4 Monthly 0.125 x 0.125 degree V002. 2012;USA, Goddard Earth Sciences Data and Information Services Center (GES DISC).

291. De Cicco LA, Hirsch, R.M., Lorenz, D., Watkins, W.D. dataRetrieval: R packages for discovering and retrieving water data available from Federal hydrologic web services. 2021.

292. Anderson B YM, Ferreri J, Crosson W, Al-Hamdan M, Schumacher A and Eddelbuettel D. hurricaneexposure: Explore and Map County-Level Hurricane Exposure in the United States. R package version 0.1.1, <URL: http://CRAN.R-project.org/package=hurri. 2020.

293. Anderson B SA, Crosson W, Al-Hamdan M, Yan M, Ferreri J, Chen Z, Quiring S and Guikema S. . hurricaneexposuredata: Data Characterizing Exposure to Hurricanes in United States Counties. R package version 0.1.0. <URL: https://github.com/geanders/hurricaneexposuredata>. 2020.

294. Garcia-Vidal C, Labori M, Viasus D, Simonetti A, Garcia-Somoza D, Dorca J, et al. Rainfall is a risk factor for sporadic cases of Legionella pneumophila pneumonia. PLoS One. 2013;8(4):e61036.

295. Stout JE, Yu VL, Muraca P, Joly J, Troup N, Tompkins LS. Potable water as a cause of sporadic cases of community-acquired legionnaires' disease. N Engl J Med. 1992;326(3):151-5.

296. Straus WL, Plouffe JF, File TM, Jr., Lipman HB, Hackman BH, Salstrom SJ, et al. Risk factors for domestic acquisition of legionnaires disease. Ohio legionnaires Disease Group. Arch Intern Med. 1996;156(15):1685-92.

297. Yu VL. Could aspiration be the major mode of transmission for Legionella? Am J Med. 1993;95(1):13-5.

298. Division of Health Informatics and Surveillance. National Notifiable Diseases Surveillance System, Weekly Tables of Infectious Disease Data. Centers for Disease Control and Prevention. Atlanta, GA. 2011.

299. Czajkowski J, Villarini G, Montgomery M, Michel-Kerjan E, Goska R. Assessing Current and Future Freshwater Flood Risk from North Atlantic Tropical Cyclones via Insurance Claims. Sci Rep. 2017;7:41609.

300. Patz JA, Epstein PR, Burke TA, Balbus JM. Global climate change and emerging infectious diseases. JAMA. 1996;275(3):217-23.

301. Pascual M, Rodo X, Ellner SP, Colwell R, Bouma MJ. Cholera dynamics and El Nino-Southern Oscillation. Science. 2000;289(5485):1766-9.

302. Kunii O, Nakamura S, Abdur RM, Wakai S. The impact on health and risk factors of the diarrhoea epidemics in the 1998 Bangladesh floods. Public health. 2002;116 2:68-74.

303. Armstrong GL, Conn LA, Pinner RW. Trends in infectious disease mortality in the United States during the 20th century. JAMA. 1999;281(1):61-6.

304. CDC. Achievements in Public Health, 1900-1999. 1999.

305. McDonald YJ, Jones NE. Drinking Water Violations and Environmental Justice in the United States, 2011–2015. American Journal of Public Health. 2018;108(10):1401-7.

306. Switzer D, Teodoro MP, Karasik S. The Human Capital Resource Challenge Recognizing and Overcoming Small Utility Workforce Obstacles. Journal (American Water Works Association). 2016;108(8):E416-E24.

307. Strosnider H, Kennedy C, Monti M, Yip F. Rural and Urban Differences in Air Quality, 2008-2012, and Community Drinking Water Quality, 2010-2015 - United States. MMWR Surveill Summ. 2017;66(13):1-10.

308. Dai D, Rhoads WJ, Katner A, Strom L, Edwards MA, Pruden A, et al. Molecular survey of Legionella and Naegleria fowleri in private well water and premise plumbing following the 2016 Louisiana flood. Environmental Science: Water Research & Technology. 2019;5(8):1464-77.

309. Pieper KJ, Jones CN, Rhoads WJ, Rome M, Gholson DM, Katner A, et al. Microbial Contamination of Drinking Water Supplied by Private Wells after Hurricane Harvey. Environmental Science & Technology. 2021;55(12):8382-92.

310. Jin Y, Flury M. Fate and Transport of Viruses in Porous Media. In: Sparks DL, editor. Advances in Agronomy. 77: Academic Press; 2002. p. 39-102.

311. Benedict KM, Reses H, Vigar M, Roth DM, Roberts VA, Mattioli M, et al. Surveillance for Waterborne Disease Outbreaks Associated with Drinking Water - United States, 2013-2014. MMWR Morb Mortal Wkly Rep. 2017;66(44):1216-21.

312. Moy E, Garcia MC, Bastian B, Rossen LM, Ingram DD, Faul M, et al. Leading Causes of Death in Nonmetropolitan and Metropolitan Areas- United States, 1999-2014. MMWR Surveill Summ. 2017;66(1):1-8.

313. Miller CE, Vasan RS. The southern rural health and mortality penalty: A review of regional health inequities in the United States. Social Science & Medicine. 2021;268:113443.

314. Cromartie JB, editor Rural America At A Glance 2018 Edition2017.

315. Meit M, Knudson A, Gilbert T, Yu AT-C, Tanenbaum E, Ormson E, et al. The 2014 update of the rural-urban chartbook. Bethesda, MD: Rural Health Reform Policy Research Center. 2014;10.

316. McKenna ML, McAtee S, Bryan PE, Jeun R, Ward T, Kraus J, et al. Human Intestinal Parasite Burden and Poor Sanitation in Rural Alabama. Am J Trop Med Hyg. 2017;97(5):1623-8.

317. Ivers LC, Ryan ET. Infectious diseases of severe weather-related and flood-related natural disasters. Curr Opin Infect Dis. 2006;19(5):408-14.

318. Vespa J M, L, Armstrong, DM. Demographic Turning Points for the United States: Population Projections for 2020 to 2060. United States Census; 2020.

319. Peirce AM, Espira LM, Larson PS. Climate Change Related Catastrophic Rainfall Events and Non-Communicable Respiratory Disease: A Systematic Review of the Literature. Climate. 2022;10(7):101.

320. (NCEI) NNCfEI. U.S. Billion-Dollar Weather and Climate Disasters 2022 [Available from: www.ncei.noaa.gov/access/billions/.

321. Woodruff JD, Irish JL, Camargo SJ. Coastal flooding by tropical cyclones and sea-level rise. Nature. 2013;504(7478):44-52.

322. Prein AF, Rasmussen RM, Ikeda K, Liu C, Clark MP, Holland GJ. The future intensification of hourly precipitation extremes. Nature Climate Change. 2016;7(1):48-52.

323. Kirshen P, Watson C, Douglas E, Gontz A, Lee J, Tian Y. Coastal flooding in the Northeastern United States due to climate change. Mitigation and Adaptation Strategies for Global Change. 2007;13(5-6):437-51.

324. Malik N, Ozturk U. Rare events in complex systems: Understanding and prediction. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2020;30(9):090401.

325. Axelrod C, Killam PP, Gaston MH, Stinson N. Primary health care and the Midwest flood disaster. Public Health Rep. 1994;109(5):601-5.

326. Curry MD, Larsen PG, Mansfield CJ, Leonardo KD. Impacts of a flood disaster on an ambulatory pediatric clinic population. Clin Pediatr (Phila). 2001;40(10):571-4.

327. Birkhead G, Vogt RL. Epidemiologic surveillance for endemic Giardia lamblia infection in Vermont. The roles of waterborne and person-to-person transmission. Am J Epidemiol. 1989;129(4):762-8.

328. NNDSS. What is case surveillance? https://www.cdc.gov/nndss/about/index.html: Centers for Disease Control, National Notifiable Disease Surveillance System 2022 [

329. Allos BM, Moore MR, Griffin PM, Tauxe RV. Surveillance for Sporadic Foodborne Disease in the 21st Century: The FoodNet Perspective. Clinical Infectious Diseases. 2004;38(Supplement_3):S115-S20.

330. Dashboard NW. National Water Dashboard. USGS2022 [Available from: https://dashboard.waterdata.usgs.gov/app/nwd/?aoi=default.

331. Switzer D, Teodoro MP. The Color of Drinking Water: Class, Race, Ethnicity, and Safe Drinking Water Act Compliance. Journal AWWA. 2017;109(9):40-5.

332. AHRQ. National Healthcare Quality and Disparities Report. Rockville, MD: Agency for Healthcare Research and Quality; 2021.

333. Fiscella K, Sanders MR. Racial and Ethnic Disparities in the Quality of Health Care. Annu Rev Public Health. 2016;37:375-94.

334. Shaw KS, Cruz-Cano R, Jiang C, Malayil L, Blythe D, Ryan P, et al. Presence of animal feeding operations and community socioeconomic factors impact salmonellosis incidence rates: An ecological analysis using data from the Foodborne Diseases Active Surveillance Network (FoodNet), 2004–2010. Environmental research. 2016;150:166-72.

335. Hadler JL, Clogher P, Libby T, Wilson E, Oosmanally N, Ryan P, et al. Relationship Between Census Tract–Level Poverty and Domestically Acquired Salmonella Incidence: Analysis of Foodborne Diseases Active Surveillance Network Data, 2010–2016. The Journal of Infectious Diseases. 2020;222(8):1405-12.

336. Quinlan JJ. Foodborne illness incidence rates and food safety risks for populations of low socioeconomic status and minority race/ethnicity: a review of the literature. International journal of environmental research and public health. 2013;10(8):3634-52.

337. Becker DJ, Oloya J, Ezeamama AE. Household socioeconomic and demographic correlates of Cryptosporidium seropositivity in the United States. PLoS neglected tropical diseases. 2015;9(9):e0004080.

338. Balazs C, Morello-Frosch R, Hubbard A, Ray I. Social disparities in nitrate-contaminated drinking water in California's San Joaquin Valley. Environmental Health Perspectives. 2011;119(9):1272-8.

339. Balazs CL, Morello-Frosch R, Hubbard AE, Ray I. Environmental justice implications of arsenic contamination: a cross-sectional, cluster-design examining exposure and compliance in community drinking water systems. Environmental health: a global access science source [electronic resource]. 2012;11(1):84-.

340. VanDerslice J. Drinking water infrastructure and environmental disparities: evidence and methodological considerations. American journal of public health. 2011;101(S1):S109-S14.

341. Nigra Anne E, Chen Q, Chillrud Steven N, Wang L, Harvey D, Mailloux B, et al. Inequalities in Public Water Arsenic Concentrations in Counties and Community Water Systems across the United States, 2006–2011. Environmental Health Perspectives.128(12):127001.

342. Meehan K, Jurjevich Jason R, Chun Nicholas MJW, Sherrill J. Geographies of insecure water access and the housing–water nexus in US cities. Proceedings of the National Academy of Sciences. 2020;117(46):28700-7.

343. Prüss-Ustün A, Wolf J, Bartram J, Clasen T, Cumming O, Freeman MC, et al. Burden of disease from inadequate water, sanitation and hygiene for selected adverse health outcomes: an updated analysis with a focus on low-and middle-income countries. International journal of hygiene and environmental health. 2019;222(5):765-77.

344. Zahran S, Brody SD, Peacock WG, Vedlitz A, Grover H. Social vulnerability and the natural and built environment: a model of flood casualties in Texas. Disasters. 2008;32(4):537-60.

345. Lindell M PC. Assessing community impacts of natural disasters. Natural Hazards. 2003;4(4):176-85.

346. Fothergill A PL. Poverty and disasters in the United States: A review of recent sociological findings. Natural Hazards. 2004;32:89-110.

347. Yager J, Rosoff, S. Housing in U.S. Floodplains. NYU Furman Center; 2017.

348. Cutter SL, Finch C. Temporal and spatial changes in social vulnerability to natural hazards. Proc Natl Acad Sci U S A. 2008;105(7):2301-6.

349. Gabe T, Falk, G, McCart, M. Hurricane Katrina: Social-demographic characteristics of impacted areas. The LIbrary of Congress; 2005.

350. Charron D, Thomas M, Waltner-Toews D, Aramini J, Edge T, Kent R, et al. Vulnerability of waterborne diseases to climate change in Canada: a review. J Toxicol Environ Health A. 2004;67(20-22):1667-77.

351. Bell J, Herring S, Jantarasami L, Adrianopoli C, Benedict K, Conlon K, et al. Ch. 4: Impacts of Extreme Events on Human Health. 2016. p. 99-128.

352. Archfield SA, Hirsch RM, Viglione A, Blöschl G. Fragmented patterns of flood change across the United States. Geophysical Research Letters. 2016;43(19):10,232-10,9.

353. Wuebbles DJ, Fahey DW, Hibbard KA. Climate science special report: fourth national climate assessment, volume I. 2017.

354. Wehner MF, Arnold JR, Knutson T, Kunkel KE, LeGrande AN. Droughts, floods, and wildfires. Climate science special report: fourth national climate assessment. 2017;1(GSFC-E-DAA-TN49033).

355. Liu C, Hofstra N, Franz E. Impacts of climate change on the microbial safety of pre-harvest leafy green vegetables as indicated by Escherichia coli O157 and Salmonella spp. Int J Food Microbiol. 2013;163(2-3):119-28.

356. Anderson GB, Schumacher A, Done JM, Hurrell JW. Projecting the Impacts of a Changing Climate: Tropical Cyclones and Flooding. Curr Environ Health Rep. 2022;9(2):244-62.

357. Vollset SE, Goren E, Yuan C-W, Cao J, Smith AE, Hsiao T, et al. Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. The Lancet. 2020;396(10258):1285-306.

358. Centers for Disease C, Prevention. Infectious disease and dermatologic conditions in evacuees and rescue workers after Hurricane Katrina--multiple states, August-September, 2005. MMWR Morb Mortal Wkly Rep. 2005;54(38):961-4.

359. Du W, FitzGerald GJ, Clark M, Hou XY. Health impacts of floods. Prehosp Disaster Med. 2010;25(3):265-72.

360. Lê F, Tracy M, Norris FH, Galea S. Displacement, county social cohesion, and depression after a large-scale traumatic event. Social psychiatry and psychiatric epidemiology. 2013;48(11):1729-41.