Gastrointestinal Illness and Microbial Water Quality in a Large River System Impacted by Combined Sewer Overflows

Dissertation Proposal
Beth M. Haley

ABSTRACT

Fecal contamination of surface waters is a widespread environmental and public health challenge in the United States (US). Each year, an estimated 6.5 million and 50 million cases of gastrointestinal (GI) illness in the US result from contamination of surface drinking water sources and recreational waters, respectively. Drinking water-associated GI illness is associated with heavy rainfall while illness among water recreators is associated with increased concentrations of pathogens and intensity of water contact. Human sources of fecal contamination are more likely to cause GI illness than fecal matter from other animals due to human-specific enteric pathogens, yet it is common to use non-human specific fecal indicator bacteria to identify possible contamination. Combined sewer overflow (CSO) activations are important contributors of untreated sewage to surface waterbodies in over 700 cities in the US. CSO activations increase concentrations of microbial contaminants in receiving waterbodies, but few studies have directly assessed the health risks associated with CSO activations. The proposed research is an integrated approach to investigate the impacts of CSOs on microbial water quality and gastrointestinal illness in a CSO-impaired river system that is used as a drinking water source and for recreation.

The research outlined in this proposal will contribute to our current understanding of the impacts of CSOs on microbial water quality and gastrointestinal illness and generate knowledge useful for improved management of combined sewer systems. This research will 1) develop a model for prediction of fecal bacteria levels throughout a river that serves as both a drinking water source and a recreational resource, 2) address a gap in our understanding of the epidemiology of GI illness and CSO events, and 3) use a human microbial source tracker and fecal indicator bacteria for prediction of GI illness for recreational users in the Merrimack River.

SPECIFIC AIMS

Fecal contamination of surface waters is a major environmental and public health challenge in the United States[1]. Each year, an estimated 6.5 million[2] and 50 million[3] cases of gastrointestinal (GI) illness in the US result from contamination of surface drinking water sources and recreational waters, respectively. Rainfall is an important driver of drinking water-associated GI illness[4–6], and illness among water recreators is associated with both the concentration of pathogens and degree of contact with the water[7]. In combined sewer systems (CSS), stormwater and sewage are collected in the same pipe. CSS are designed to discharge untreated sewage and stormwater directly into surface water bodies when the volume of water in the collection pipe exceeds treatment capacity—called a combined sewer overflow (CSO) event—often during heavy rainfall. Increased concentrations of microbial contaminants in receiving waterbodies following CSO activations are well documented[9–13]. Despite the clear conceptual connection between pathogens in CSO discharge and public health risk, few studies have directly assessed the health impacts or risks associated with CSO activations.

The primary goal of this proposal is to improve our understanding of the water quality and health risks associated with untreated sewage released during CSO activations in a river system used for recreation and as a drinking water source. We propose to use the Merrimack River, located in New Hampshire and northern Massachusetts, as a model CSO-impaired river system[14]. We will leverage an existing dataset of fecal indicator bacteria to develop a spatial statistical model that can be used to predict concentrations of bacteria throughout the river. Associations between CSO events and GI illness in the Merrimack region will be examined using electronic medical records of emergency room visits. We will also conduct a quantitative microbial risk assessment (QMRA) for recreational users of the Merrimack River, leveraging field measurements of a human microbial source tracker and fecal indicator bacteria collected under wet and dry weather conditions. We propose the following specific aims:

Aim 1: Develop a predictive spatial statistical model for *Escherichia coli* in the rivers and streams of the Merrimack watershed.

Objective 1a: Develop a spatial stream network model for *Escherichia coli* (*E. coli*) using historical *E. coli* concentration data, meteorological patterns, point source locations, and land use and land cover characteristics.
Objective 1b: Evaluate model performance and predictive accuracy of model variations that differ by autocovariance structure.
Objective 1c: Apply the best model to generate predications of *E. coli* concentrations at equidistant unsampled locations throughout the river and stream network in the Merrimack watershed to identify impaired river sections, frequency of impairment, and areas with greatest uncertainty.

Aim 2: Evaluate the relationship between CSO activations and acute gastrointestinal illness (AGI) in communities that source their drinking water from the Merrimack River and communities with a different drinking water source.
Hypothesis 2a: Cases of AGI will be associated with upstream CSO activations, after controlling for confounding factors, only among communities that source their drinking water exclusively from the Merrimack River.
Hypothesis 2b: Among communities with a CSO-impacted drinking water source, CSO activations will be a stronger predictor of AGI than heavy ($\geq 90^{th}$ or $\geq 95^{th}$ percentile) or extreme ($\geq 99^{th}$ percentile) precipitation.

Aim 3: Conduct a quantitative microbial risk assessment (QMRA) to estimate risk of gastrointestinal (GI) illness associated with recreational exposure in the Merrimack River.
Objective 3a: Estimate the risk of GI illness for recreational exposures at two recreational areas in the Merrimack River using fecal indicator bacteria and a human microbial source tracker to estimate human pathogen concentrations.
Hypothesis 3a Predicted risk of GI illness associated with recreational exposure is greater under wet weather conditions than under dry weather conditions.

Impact: CSOs are a legacy problem that threaten human and ecosystem health. CSS persist due to the high costs of separating stormwater from sewage pipes and limited investment in water infrastructure updates. CSO activations are driven by precipitation, making them vulnerable to climate change. Research that helps us understand the risks to health may lend further evidence to decisions about permanent elimination of combined sewer systems.

RESEARCH STRATEGY

A. Significance

Fecal contamination in surface waters threatens human and ecosystem health

Fecal contamination of surface waters is a major environmental and public health challenge in the United States (US). A national inventory of US waterways in 2017 identified fecal pollution as a top cause of impairment[1]. Sources of fecal pollution include agricultural[15,16] and urban runoff[16,17], treated wastewater effluent[18,19], combined sewer overflows[9,17,20], failing septic systems[21], and wildlife excrement[22].

Exposure to pathogens in surface waters can occur through consumption of drinking water sourced from a contaminated surface water body or during water recreation. Adverse health outcomes associated with surface waters include gastrointestinal illness, respiratory infections, eye infections, ear/nose/throat infections, and skin irritations[3,7,23–28]. The primary outcome of interest for the proposed work is gastrointestinal (GI) illness given the abundance of epidemiological data on GI illness among water recreators[3,23], the availability of well-established dose-response models for enteric pathogens[29], and the prevalence of GI illness associated with drinking water exposure[30].

Drinking water and gastrointestinal illness

Microbial contamination of drinking water due to poor source water quality, insufficient treatment, and water distribution network failures can cause illness among drinking water consumers[2,31–34]. In the US, an estimated 16.4 million cases of acute gastrointestinal illness (AGI) are attributable to drinking water each year with more than 40% of those cases attributable to source water quality or insufficient treatment[2].

Rainfall is an important driver of drinking water-associated GI illness, especially for surface water sources[30,32,35]. From 1948 to 1994, 68% of all waterborne disease outbreaks in the US were preceded by precipitation events above the 80th percentile[35]. The majority of epidemiological studies investigating the relationship between GI illness and heavy rainfall found a significant but lagged relationship between precipitation events and increased GI illness with a lag period ranging from 1 day to 6 weeks[5,6,36–41]. Important confounders in the relationship between precipitation and GI illness include temperature[37,38,42], antecedent rainfall[5,6,41,43], day of the week[37], season[6,36,37,42], and long term time trends[37,42]. Several studies found that there was an association between rainfall and GI illness only in communities whose drinking water came from surface water sources[39,42,44].

Water recreation and gastrointestinal illness

An estimated 4 billion surface water recreation events—including swimming, paddling, boating, and fishing—occur each year in the US resulting in approximately 50 million cases of GI illness and 90 million illnesses overall[3]. Leading etiological agents for GI illness outbreaks associated with untreated recreational water in the US are enteric pathogens including norovirus, pathogenic *E. coli*, *Cryptosporidium*, *Campylobacter*, *Giardia*, and *Shigella*[28,45].

The risk of illness during recreation in untreated surface waters is associated with both the concentration of pathogens in the water and the degree of contact with the water[7]. Evidence suggests that children experience the largest attributable illness burden, likely due to having comparatively higher water exposure and being more susceptible to illness than adults[25]. A significant increase in relative risk of GI illness is found for swimmers (2.19 [95% CI: 1.82, 2.63]) and those engaging in moderate contact activities (2.69 [95% CI: 1.04, 6.92]; e.g., surfing, rafting, or whitewater canoeing or kayaking)[7]. While some studies find a nonsignificant increase in GI illness with minimal contact activities (e.g., canoeing, kayaking, rowing, or fishing)[7], others have shown a significant increase in risk among even limited-contact recreators compared to non-recreators[46].

Regulatory criteria for recreational water quality are based on fecal indicator bacteria and epidemiological studies

Prospective cohort studies that found exposure-response relationships between fecal indicator bacteria (FIB, e.g. *E. coli* and enterococci) and GI illness among swimmers in wastewater-impacted waterbodies have served as the foundation for US recreational water quality criteria since the 1970s[47–49]. FIBs are useful indicators of fecal contamination because they are ubiquitous in human feces, exist in detectable concentrations in environmental samples, are rarely infectious to people, and

can be enumerated reliably with well-established field and lab protocols[49]. FIB enumeration methods include those in which the bacteria are grown (cultured) and those in which the DNA is extracted, amplified, and quantified using qPCR[49].

The relationship between FIB concentration and GI illness in recreators is often dependent on the source(s) of fecal pollution impacting the location of the study[8,22,50,51]. Studies at recreation locations impacted by a human-associated point source tend to demonstrate a significant relationship between FIBs and GI illness[23,48] while the relationship between FIBs and GI illness is less predictable at locations impacted by non-point sources[23,52–54].

FIBs exist in the GI tracts of all warm-blooded animals, meaning that tests for FIBs cannot discern between human fecal inputs and feces from other animals[49]. Many of the waterborne pathogens that affect humans are only present in human sources[8], so it is difficult to determine human health risk based on FIBs. Host-specific genetic markers, called microbial source trackers (MST), have recently been used to identify the presence of animal-specific fecal sources, including humans[55,56]. MST studies have found that animal sources of fecal contamination contribute greatly to FIB concentration loads[23], but tend to carry limited risk to human health[15,22,57].

Estimation of human health risk with quantitative microbial risk assessment

Quantitative microbial risk assessment (QMRA) is an established scientific framework for estimating risk of illness caused by infectious pathogens that can be used in place of or to complement epidemiological research[29,58]. A group of reference pathogens are commonly used to characterize risk of illness in a QMRA because they are detectable in environmental samples, are responsible for the majority of waterborne illnesses, have reliable dose-response relationships, and are representative of broader pathogen groups (i.e., bacteria, protozoa, and viruses)[8,29,59]. The ingested pathogen dose is estimated based on pathogen concentration in the water sample, rate of water ingestion, and the frequency and duration of exposure[29]. Dose-response relationships estimate the probability of infection, given an ingested dose of each reference pathogen. Finally, risk characterization estimates the cumulative probability of infection and the likelihood of illness given infection[29].

Obstacles to measuring the concentration of reference pathogens in water samples include the variety of pathogens, low concentrations of pathogens in environmental samples, and the cost and complexity of enumeration[60,61]. Although human health risk cannot be estimated from FIB concentration data alone, several methods have been proposed that either allow for source apportionment of FIBs using MSTs to estimate reference pathogen concentrations[8,62] or estimate risk of illness directly from MST data[51,59,63].

Combined sewer overflows are important sources of human fecal contamination

Combined sewer systems (CSS) exist in approximately 700 communities in the US where stormwater and sewage are collected in the same pipe and flow together to a wastewater treatment plant. CSS are designed to discharge untreated sewage and stormwater directly into surface water bodies when the volume of water in the collection pipe exceeds treatment capacity, often during heavy rainfall. CSO events are driven by frequency and intensity of precipitation events, and so tend to vary over space and time[64].

CSO events are significant sources of chemical and microbial pollutants including FIBs[10–13,17]; enteric viruses[65,66] and protozoa[10,65,67]; human MSTs[20,68]; and antibiotic resistance genes[69]. CSOs can lead to one to two order of magnitude increases in the concentrations of FIBs[10,11,13], pathogens[10], and MSTs[20] compared to dry weather concentrations.

Increased precipitation in many areas due to climate change is expected to increase incidence of GI illness[70,71], adding urgency to research that seeks to understand the relationships between rainfall, CSO events, and GI illness.

Health risks associated with CSOs are understudied

Despite the clear conceptual connection between pathogens in CSO discharge and public health risk, few studies have directly assessed the health impacts or risks associated with CSO activations.

Of the three retrospective time series analyses investigating the relationship between CSO or sewage bypass events and GI illness through drinking water exposure, two suggest that sewage discharge events are associated with GI illness in systems where drinking water is sourced from the affected water body[4,72]. Using a similar methodology, Drayna et al[36] found a significant association between extreme rainfall events and ER visits independent of CSO events. Two studies that combined

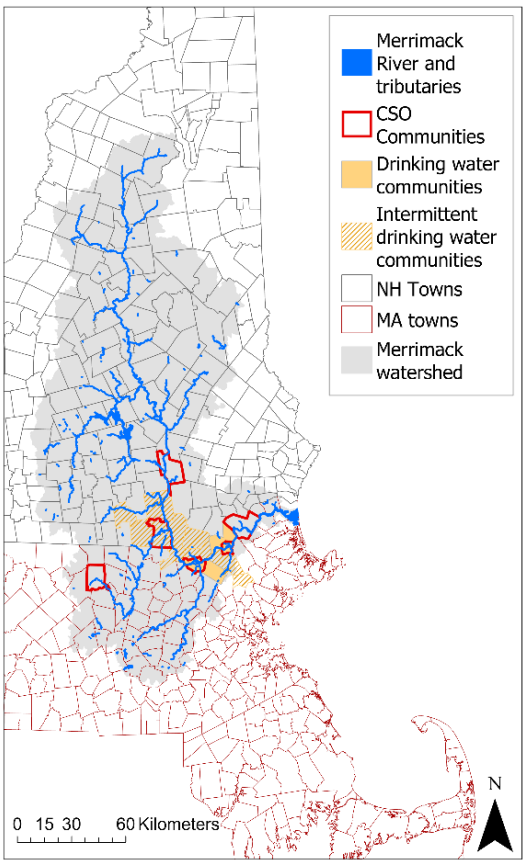
water quality modeling of sewage discharges to surface source waters with QMRA of GI illness due to drinking water contamination found that drinking water treatment is adequate for removal of pathogens to meet the health target only when treatment performance is optimal[73,74]. These studies emphasize the importance of understanding pathogen loading upstream of drinking water intakes and effective treatment of surface source waters.

Recreational exposure to CSO-impaired water for a cohort of triathletes resulted in a 42% attack rate of GI illness compared to 8% among triathletes in non-CSO-impaired water (RR: 5.0 [95% CI: 4.0,6.39])[75]. A QMRA conducted in a CSO-impaired river for several recreational exposure scenarios found that the probability of contracting gastrointestinal illness over the course of a year varied from 0.14 to nearly 0.70 depending on the level of water contact[65]. A case-crossover study of childhood emergency room (ER) visits found that children who lived within 500m of a CSO outfall location had an elevated risk of a GI-related ER visit two days after a CSO event (OR: 1.16 [95% CI: 1.04,1.30]), possibly due to direct contact with CSO discharge through recreational activities or aerosolized pathogens[76].

The Merrimack River: An ideal study system

Located in New Hampshire and northern Massachusetts, the 117-mile Merrimack River is an ideal system for studying the health implications of CSOs (Figure 1). The watershed spans over 5,000 square miles and contains 6 communities with combined sewers (63 CSO outfalls in total)[14]. The river and its tributaries serve as a recreational resource for the 2.6 million people who live in the watershed[14]; including swimming, fishing, kayaking, boating, and rowing[77]. More than 500,000 people also rely on the Merrimack as a primary or intermittent drinking water source[14], including the Massachusetts environmental justice communities of Lowell and Lawrence. CSOs are a concern for many community members and politicians in the watershed[78], but barriers remain to eliminating CSOs from this river system including the cost associated with CSO mitigation, coordination between state regulators, and reliable information on the true impacts of CSOs in the watershed[77].

Figure 1: The Merrimack River Watershed



B. Innovation

The proposed innovative research provides an integrated approach to investigating the dynamics of microbial water quality and gastrointestinal illness related to CSOs in the Merrimack River which has never been done in this waterway. Other innovative aspects of the proposed work include:

- Direct investigation of the relationship between CSO events and GI illness using retrospective time series analysis and QMRA methods for drinking water and recreational water exposures, respectively
- Creation of a novel spatial statistical model for *E. coli* that includes CSO outfall and wastewater treatment plant locations
- A novel application of a proposed source apportionment method of FIBs in environmental samples to estimate pathogen concentrations using FIB concentrations enumerated with droplet digital PCR and a human MST, *Bacteroides* HF183

C. Approach

Aim 1: Develop a predictive spatial statistical model for *Escherichia coli* in the rivers and streams of the Merrimack watershed.

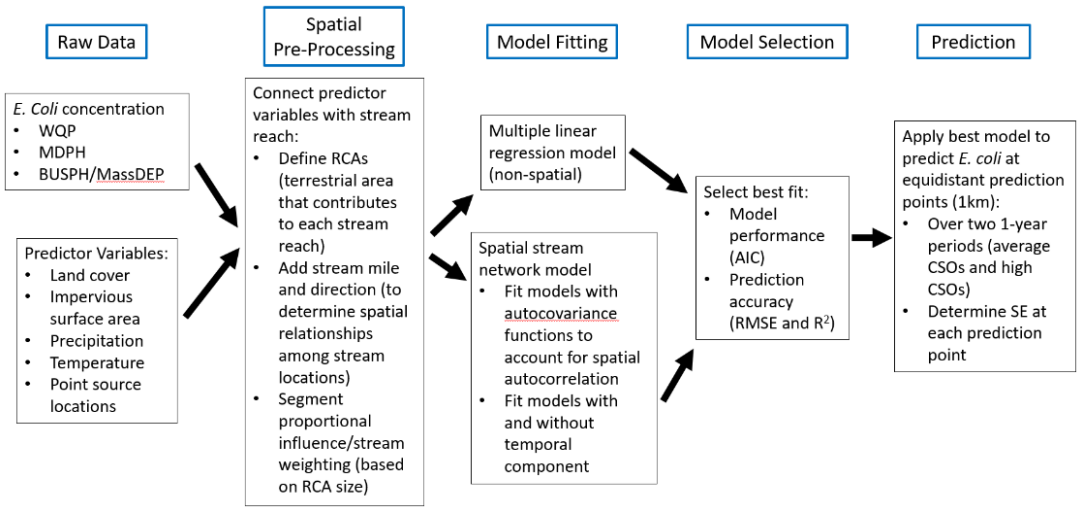
Objective 1a: Develop a spatial stream network model for *Escherichia coli* (*E. coli*) using historical *E. coli* concentration data, meteorological patterns, point source locations, and land use and land cover characteristics.

Objective 1b: Evaluate model performance and predictive accuracy of model variations that differ by autocovariance structure.

Objective 1c: Apply the best model to generate predications of *E. coli* concentrations at equidistant unsampled locations throughout the river and stream network in the Merrimack watershed to identify impaired river sections, frequency of impairment, and areas with greatest uncertainty.

Overview: The objective of this Aim is to develop and evaluate a predictive model for *E. coli* concentrations in the Merrimack River that accounts for spatial autocorrelation and the natural constraints of a river network (i.e., stream channel boundaries and flow direction). Spatial statistical models build on traditional multiple linear regression by including a spatial autocovariance function to account for random effects in the variance due to spatial autocorrelation. Terrestrial spatial statistical models account for spatial autocorrelation according to Euclidean distance, but spatial stream network (SSN) models are designed for stream networks where stream distance and flow direction are also relevant for spatial relationships. SSN models allow for a mixed model approach where the autocovariance structure can be defined in terms of flow-connected and/or Euclidean distance. We will fit SSN models with different autocovariance structures and use the best fitting model to estimate *E. coli* concentrations and uncertainty at unsampled locations. The full model development and applications process is shown in Figure 2. This model can be used to inform future monitoring programs by identifying potential hot spots and areas with high uncertainty.

Figure 2: Summary of SSN model development process



Study Area: The Merrimack River flows from Franklin, NH to the Atlantic Ocean near Newburyport, MA. The watershed covers 5,007 square miles and includes five major tributaries and over 6,000 miles of rivers and streams (Figure 1).

Model Data:

Water Quality Data: Culturable *E. coli* concentration data will be compiled from three sources: the national Water Quality Portal (WQP), Massachusetts Department of Public Health (MDPH) recreational water surveillance program, and a pilot study conducted by BUSPH in partnership with the Massachusetts Department of Environmental Protection (MassDEP) (Table 1). Only concentration data enumerated through an EPA-approved method for *E.coli* quantification[79] will be included in this study.

Table 1: Summary of E. coli samples collected from the Merrimack River watershed

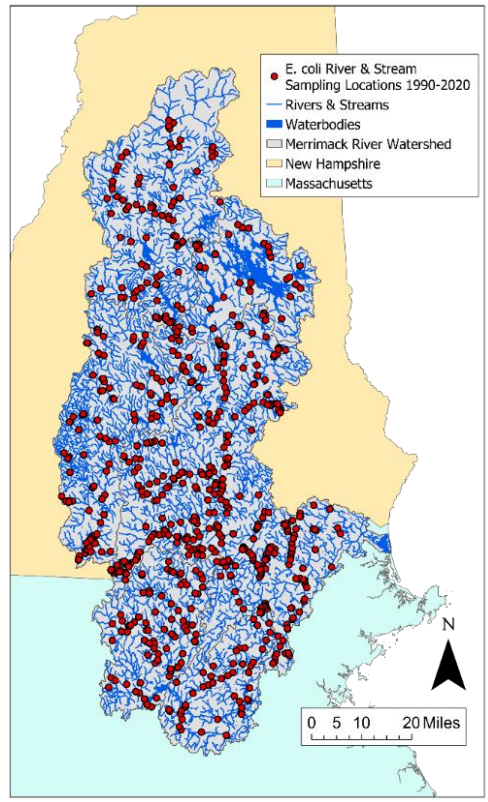
Source	Years	Location(s)	Number of Samples
WQP	1990-2020	782 river and stream locations throughout the watershed	8,151
MDPH	2002-2018	1 location on the Merrimack River in Lowell, MA	202
BUSPH Pilot Study with MassDEP	2019-2021	24 locations on the Merrimack River from Manchester, NH to Newburyport, MA	406 (estimated)
Total	**1990-2021**	**807**	**8,759**

The WQP is a cooperative platform sponsored by the United States Geological Survey, Environmental Protection Agency, and National Water Quality Monitoring Council that provides public access to data collected by over 400 state, federal, tribal, and local agencies and organizations[80]. The WQP includes *E. coli* concentration data for 8,151 samples collected from 782 river and stream locations throughout the Merrimack River watershed between 1990 and 2020 (Figure 3). We will omit lake samples and samples that were not validated or finalized.

Additional *E. coli* concentration data are available from routine recreational beach water quality sampling conducted by MDPH at one freshwater location in the Merrimack River (202 samples collected between 2002 and 2018)[81]. *E.coli* concentration data from a BUSPH pilot water quality monitoring study of 206 samples collected from 22 locations, primarily in the Lower Merrimack River, in 2019 and 2020 and approximately 200 samples collected from the same locations in 2021 will be included. All *E. coli* concentration data will include location and date of collection. The limit of detection (LOD) varies from 1 to 10 (cfu or MPN)/100mL based on the culturable enumeration method and dilution used. We will include non-detects in the dataset at the value of the LOD[49]. We anticipate that *E. coli* concentration data will be right skewed, and so plan to use log-transformed concentration values. We will assess the temporal variation in the frequency and location of water quality sampling, and include adjustments in the model as needed for seasonal or long-term time trends that may be present.

Predictor Variables: Meteorological data, land use/land cover characteristics, river characteristics, and point source data will be included in all models as predictor variables. Shapefiles of the stream network that have been reconditioned to be compatible with SSN analyses are available through the United States Forest Service Rocky Mountain Research Station[82]. The United States Geological Survey (USGS) National Hydrography Dataset (NHD)[83] will be used to delineate watershed boundaries and include waterbodies (e.g., lakes, rivers, and reservoirs) into the SSN model. Land cover and land use data are available through the StreamCat dataset, which includes upstream characteristics for each stream reach in the NHD stream network[84]. Total daily precipitation and daily mean temperature will be included at 4km resolution. These modeled data are part of a gridded dataset developed by the PRISM Climate Group at Oregon State University[85]. Combined sewer overflow (CSO) outfall[14] and wastewater treatment plant (WWTP)[86] locations will be included in the model as potential point sources of *E. coli*.

Figure 3: River and stream E. coli sampling locations in the Merrimack River watershed, 1990-2020



Model Development:

Spatial Pre-Processing: We will use the STARS package[87] in ArcGIS Pro[88] to generate the spatial information required to fit the SSN model including a landscape network, a digital elevation model, prediction points, distance from the stream outlet, segment proportional influence, and reach contributing areas (RCA) for each stream segment. The RCAs will be used to calculate precipitation, temperature, land cover, impervious surface area, and point sources (CSO outfalls and WWTPs) for each stream reach. These spatial data will be exported as .ssn objects and directly imported to R statistical software[89] using the SSN package[90].

Model Components: The SSN package consists of tools needed to fit generalized linear models for stream network data that incorporates spatial autocorrelation. SSN models account for spatial structure in residual errors by including an autocovariance function to describe the spatial relationships among the model elements[91].

Considering the log-transformed *E. coli* concentration data as the response variable, a SSN model can be represented by Equation 1:

$$\ln(\mathrm{EC}(s, t)) = A(s, t)\,\beta + Z(s, t) + \varepsilon \qquad (1)$$

where (EC(s, t)) is a vector of observations according to a spatial coordinate matrix (s) and time vector (t), $A(s, t)$ is a design matrix for fixed effects (predictor variables), β is a vector of coefficients for the fixed effects, $Z(s, t)$ comprises the variance explained by spatiotemporal autocorrelation in the data (random effects defined by a spatial autocovariance structure), and ε is a vector of independent random errors that comprises the variance unexplained by the model.

Multiple Linear Regression Model: We will fit a multiple linear regression (MLR) model as a non-spatial comparison to the SSN models. The MLR model will be fit with the same fixed effects (predictor variables) included in initial SSN models

(i.e., precipitation, temperature, land cover, impervious surface area, and point sources). Unlike the SSN models, the assumption in the MLR model is that the random errors (ε) are independent of spatial relationships, meaning that the MLR model will not account for spatial autocorrelation. Following the notation of Equation 1, the MLR model can be expressed as:

$$\ln(\mathrm{EC}(s, t)) = A(s, t)\,\beta + \varepsilon \qquad (2)$$

Model Fitting: We will consider different definitions of the error function $Z(s, t)$ to compare models with varying spatial autocovariance functions. We will also fit a series of SSN models that to not account for temporal autocorrelation. We will use a mixed model approach that combines flow-connected and Euclidean distance autocovariance functions to account for a variety of spatial relationships including those that occur within the stream (flow-connected) and those that result from the straight-line distances over the surrounding terrestrial area (Euclidean)[92] . For example, a flow-connected autocovariance function may better account for stream discharge whereas a Euclidean autocovariance function may better account for land cover and precipitation parameters[91]. We will utilize tail-up covariance models for flow-connected parameters given the passive, downstream movement of *E. coli* in a stream system[92]. The size of the contributing catchment area (RCA) will be used to determine stream segment weights, which is an important factor in the construction of a tail-up covariance function[92].

Model Selection: We will use Akaike Information Criterion (AIC) to evaluate overall model fit[93] and use leave-one-out cross validation (LOOCV) to compare the predictive performance of all models[91]. Prediction accuracy will be assessed through the root-mean-square error (RMSE) as a measure of predictive capability and the proportion of variance in the observations explained by the fixed and random effects included in each model (R^2)[91].

Prediction:

We will use the model with the best fit to estimate *E. coli* concentrations at unsampled locations in the watershed as designated by a network of 1 km equidistant prediction points. Universal kriging[91] will be used to estimate *E. coli* concentrations at all prediction points over the course of two years of the study period: one year with average occurrence of CSO events and the year with the most CSO events. Candidate years for this analysis will be selected based on the completeness of CSO event data. We will define water quality impairment as *E. coli* concentrations greater than the recommended 2012 Recreational Water Quality Criteria for a single sample maximum (410 cfu/100 mL)[49]. Impaired river sections and frequency of impairment will be assessed over the two prediction years.

Prediction Uncertainty:

A benefit of using a SSN model in contrast to a nonspatial regression model is that spatially unique standard errors will also be calculated for each prediction point in addition to an estimated mean concentration[92]. The standard errors for prediction points throughout the network will provide information on the precision of the predictions at each location; areas with large standard errors may warrant additional targeted sampling in the future, informing future water quality monitoring strategies. Previous studies suggest that standard errors tend to be smaller near sampling locations and increase with distance from a sampling location[91,94,95]. We will calculate standard errors for all prediction points in the Merrimack River watershed and identify areas with the largest uncertainty.

Strengths and Limitations:

Strengths: The SSN approach improves upon non-spatial regression modeling by accounting for spatial autocorrelation among sampling locations and allowing for different kinds of spatial relationships between points (flow-connected and Euclidean distance). Using this method and a diverse set of predictor variables, we are able to leverage a large existing water quality dataset without additional time and financial resources that would be necessary for extensive water quality monitoring in such a large watershed. Location-specific uncertainty provided by the SSN model will be useful in improving future monitoring efforts by identifying areas of greatest uncertainty and potential hot spots. As additional *E. coli* data become available through water quality monitoring programs in the watershed, they can be incorporated into the model, leading to iterative improvement. Through this process, ongoing monitoring efforts and the SSN model become processes working in tandem, each informing the other.

Limitations: A geostatistical model like the SSN model relies on correlations as opposed to mechanistic or causal relationships. As a result, future research questions about specific changes to wastewater infrastructure elements in the

watershed (e.g., separation of combined sewers) cannot be addressed with the model proposed here. However, compared to mechanistic models, geostatistical models are much less complicated to implement due to fewer data and expertise requirements and they tend to still have good estimation accuracy[96,97]; in the two published studies using SSN models to predict fecal indicator organisms, R^2 values are 0.71[94] and nearing 1[95], respectively. We are proposing to use *E. coli* concentration data that were collected over 31 years during which time anthropogenic activities (e.g., changes to wastewater infrastructure, meteorological trends, increased development) may have impacted microbiological inputs to the streams and rivers in the watershed. We will assess temporal patterns in the concentration data and consider fitting separate models for shorter periods of time (e.g., five year increments) or including temporal variables as predictors.

Aim 2: Evaluate the relationship between CSO activations and acute gastrointestinal illness (AGI) in communities that source their drinking water from the Merrimack River and communities with a different drinking water source.

Hypothesis 2a: Cases of AGI will be associated with upstream CSO activations, after controlling for confounding factors, only among communities that source their drinking water exclusively from the Merrimack River.

Hypothesis 2b: Among communities with a CSO-impacted drinking water source, CSO activations will be a stronger predictor of AGI than heavy ($\geq$90th or $\geq$95th percentile) or extreme ($\geq$99th percentile) precipitation.

Overview: The objective of this Aim is to assess whether the association between upstream CSO activations and AGI differs depending on primary community drinking water source (Merrimack River or other drinking water source) over the study period (2014-2019). We will also evaluate the relationship between heavy precipitation and AGI to determine which predictor (CSO activations or precipitation) has a stronger association with AGI in communities with a CSO-impaired drinking water source.

Exposure Groups: Community-level exposure status will be assigned based on drinking water source (Figure 4). The Merrimack River is the primary drinking water source for four Massachusetts communities: Lowell, Lawrence, Methuen, and Tewksbury (Exposed Group). Eight communities adjacent to the Merrimack River in Massachusetts that do not rely on the river as a drinking water source at any point in the year will be selected as the Unexposed Group (Table 2)[98].

Figure 4: Exposure groups for Aim 2 based on drinking water source

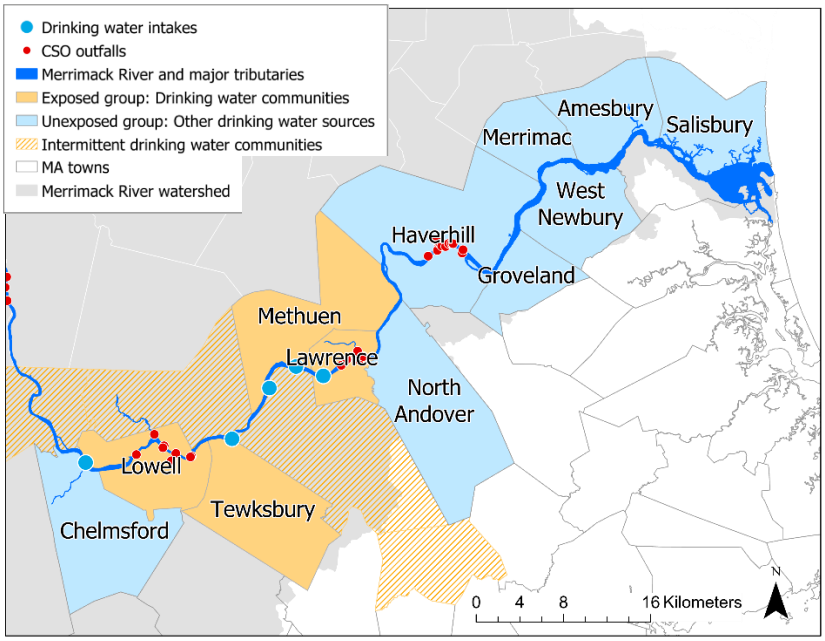


Table 2: Descriptive statistics for Aim 2 exposure groups

Characteristics	Exposed	Unexposed
No. of Cities/towns	4	8
Total population	272,708	171,117
Age Group		
<5 years	17,186	9,519
5-19 years	52,455	30,062
20-64 years	168,880	104,474
≥65 years	34,187	27,062
Population density (people/km^2)[a]	2,646	537
Median household income (USD)[a]	$62,298	$91,934
Average household size[a]	2.80	2.58

[a] *Weighted by population*
All demographic data is from the 2019 American Community Survey 5-year estimates

Outcome – Emergency Room Visits for Gastrointestinal Illness: Data on emergency room (ER) visits for gastrointestinal illness will be used to determine community-level counts of AGI. Individual-level ER visit data to all Massachusetts hospitals for the study period (January 1, 2014 – December 31, 2019) will be accessed via the Emergency Department Database (EDD) from the Center for Health Information and Analysis (CHIA)[99]. Data from two other CHIA databases—

the Hospital Inpatient Discharge Database and Outpatient Observations Database—over the same time period will be used to identify cases when a patient was admitted to the hospital or discharged to outpatient observation, respectively, from the emergency department. Data for each visit will include diagnosis code, town of residence, age, sex, and date of visit. AGI cases for this study will be defined as those diagnosed with specified gastroenteritis (*International Classification of Disease-10*[100] codes A00-A09), unspecified gastroenteritis (K52.9), nausea and vomiting (R11), and diarrhea (R19.7). Town of residence will be used to calculate the daily total number of AGI cases per 100,000 people for each exposure group throughout the study period.

Exposure – CSO Activations and Precipitation:

CSO Activations: Information on CSO activations in the Merrimack River watershed during the study period has been collected from the EPA Region 1 office. Data for each activation include date, location and discharge volume (Table 3). A daily time series of CSO events will be developed for each exposure group for the study period. Both the daily number of upstream CSO events and total upstream CSO discharge volume (millions of gallons per day, MGD) will be summed for each study community, and then totaled for the respective exposure group. This will allow for exposure classification at the group level that still considers the geospatial relationship of each community to CSO outfalls.

Table 3: Summary of CSO events in the Merrimack River watershed, 2014-2019

	2014		2015		2016		2017		2018	
	# of events	Total volume (MG)	# of events	Total volume (MG)	# of events	Total volume (MG)	# of events	Total volume (MG)	# of events	Total volume (MG)
Manchester	394	377.01	283	160.31	151	130.51	176	226.87	177	364.28
Nashua	52	50.83	35	5.52	22	9.79	23	9.61	68	42.51
Fitchburg	127	17.72	112	5.60	111	8.43	159	12.28	182	31.25
Lowell	105	278.47	76	112.87	77	118.42	95	108.04	153	292.36
Lawrence	10	6.15	4	12.66	8	36.00	16	17.77	30	93.08
Haverhill	202	43.21	70	8.08	81	21.43	180	30.77	197	49.52

Precipitation: We will access daily rainfall data from the National Centers for Environmental Information (NCEI)[101]. We will select a representative weather station in the Lower Merrimack region based on proximity to study communities and completeness of available data over the study period. We will create three binary exposure variables for heavy precipitation to identify days with rainfall above a certain percentile cutoff ($\geq$90th, 95th and 99th percentile)[4,35,37] over the course of the 6 years in the study period.

Covariates: We will consider additional time and environmental variables in our analyses as covariates. Time covariates will include day of the week (categorical), Federal holidays (binary), and season (categorical). A spline function will be included to adjust for long-term time trends. Average daily temperature will be included as a covariate in all models, as temperature has been associated with increased rates of AGI in previous studies[38,102]. Daily average temperature will be calculated from the maximum and minimum daily temperature as recorded by the same weather station used to quantify precipitation.

Statistical Analysis: We will conduct a retrospective time-series analysis to investigate the relationship between CSO activations and daily ER visits for AGI in the Merrimack Valley of Massachusetts from 2014 to 2019. We will also assess associations between heavy precipitation and AGI cases. We will examine the correlation between CSO activations and daily precipitation and consider including an interaction term when investigating CSO activations as the primary independent variable.

We will use a Poisson regression with a distributed lag nonlinear regression model[103] (DLNM) to investigate the relationship between CSO activations/ heavy precipitation and AGI within each exposure group (Equation 1). A Poisson regression model is appropriate because the response variable is composed of counts (number of daily AGI cases); we will use a quasi-Poisson distribution if the data are overdispersed. The DLNM framework allows for non-linear relationships between exposure and outcome as well as temporally lagged associations. The variable temporal lag between surface water contamination and emergency room cases of AGI can be attributed to environmental transport, drinking water treatment time and distribution, disease incubation, and care-seeking behavior. We will also evaluate lag periods of 4, 8, 15, and 28 days based on the literature[4,5,37,43]. We will calculate cumulative exposure-response functions by summing the effects across all lags.

We will conduct a sensitivity analysis to evaluate the relationship between AGI and variations of the predictor variables, namely CSO activations (frequency and cumulative CSO volume) and heavy rainfall above a percentile cutoff (greater than 90th, 95th, and 99th percentiles). The cumulative relative risk of AGI following an exposure event (CSO activation or heavy rainfall) and confidence intervals will be used to assess and compare predictors of AGI. We will conduct sub-analyses by stratifying by age group and by season to explore age-related and seasonal patterns within each exposure group. We will also assess the relationship between CSO activations/heavy rainfall and all daily ER visits in each exposure group to establish whether the predictor variables are drivers of ER visits for all complaints[5,37]. All analyses will be conducted in R[89] using the dlnm package[104] for the distributed lag model.

Strengths and Limitations:

Strengths: Our research will build on a 2015 study in Massachusetts that found a 13% increase in the expected rate of emergency room visits due to gastrointestinal illness among all age groups following 99th percentile rainfall in the Lower Merrimack region from 2003 to 2007[4]. The proposed study will address sources of potential exposure misclassification in the 2015 study by incorporating CSO event data (instead of rainfall as a proxy), town drinking water source, and the geospatial relationship of CSO events and drinking water intakes into our analysis. The use of the DLNM framework is also a strength of the proposed study as it simultaneously allows for nonlinear exposure-response relationships and variable temporal lags.

Limitations: As with any ecological epidemiology study, this analysis is limited by a lack of individual-level exposure data; as a result, no conclusions can be made regarding the cause of AGI in this study population. Additionally, exposure classification based on drinking water source does not address possible exposure to waterborne pathogens during recreation or consumption of contaminated water from other sources. Gastrointestinal illness caused by waterborne pathogens is often self-limiting; the use of ER visit data to quantify AGI presents the limitation of only taking into account extreme cases in what might be a more widespread issue. Relying on hospital admissions data of any kind excludes cases among people who lack access to healthcare or are unlikely to seek care. Because the total number of AGI cases is likely to exceed what is recorded in ER data, use of ER data likely biases the results towards the null, indicating that any detected association could be larger if outcome data were improved.

Aim 3: Conduct a quantitative microbial risk assessment (QMRA) to estimate risk of gastrointestinal (GI) illness associated with recreational exposure in the Merrimack River.

Objective 3a: Estimate the risk of GI illness for recreational exposures at two recreational areas in the Merrimack River using fecal indicator bacteria and a human microbial source tracker to estimate human pathogen concentrations.

Hypothesis 3a Predicted risk of GI illness associated with recreational exposure is greater under wet weather conditions than under dry weather conditions.

Overview: The goal of this study is to characterize the risk of developing GI illness due to recreational exposure at two sites in the Merrimack River using a QMRA framework. We will use concentration measurements of two common fecal indicators bacteria (FIB), *Escherichia coli* (*E. coli*) and enterococci, and a human-specific microbial source tracker (MST), *Bacteroides* HF183, to estimate enteric pathogen dose. We will use a probabilistic approach to account for uncertainty in pathogen concentration estimates and water ingestion rates. We will characterize and compare predicted risks under dry and wet weather conditions.

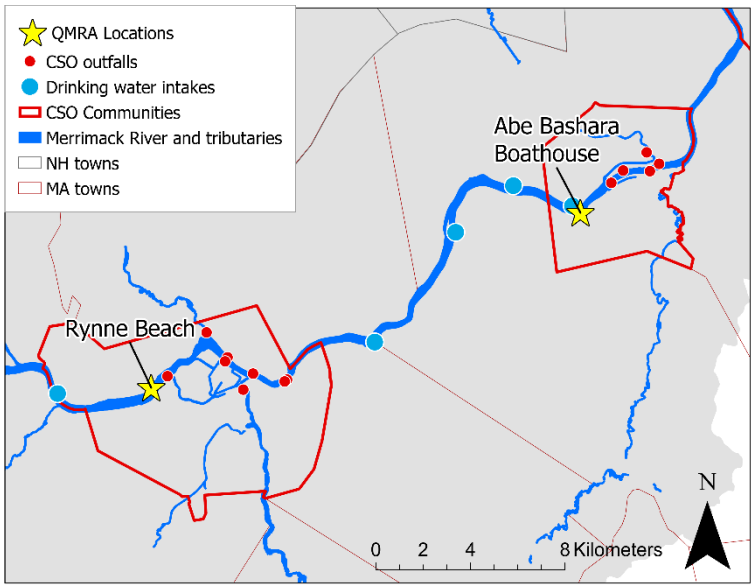
Recreation Locations: We will estimate risk of GI illness at two sites in the Merrimack River with documented recreational use that are located downstream of CSO outfalls: Rynne Beach in Lowell, MA, and the Abe Bashara Boathouse in Lawrence, MA (Figure 5). Rynne Beach is the only public freshwater beach on the Lower Merrimack River[105,106]. Programs at Abe Bashara Boathouse engage thousands of children each year in small watercraft recreation including kayaking, canoeing, rowing, sailing, and stand up paddleboarding[107]. Pathogen concentration estimates will be calculated from water samples collected from these sites in 2020 and 2021.

Data Sources:

Water quality data: Water samples are collected from the two recreation locations included in the Merrimack River during the spring, summer, and fall of 2020 and 2021. Samples from 2020 (July through November) were collected from the center

channel of the river, while samples from 2021 will be collected along the shore of the Merrimack River. Center channel samples may be more representative of limited-contact recreation exposures (e.g., kayaking, boating, and canoeing), while shore samples may be more representative of full-contact recreation exposures (e.g., swimming and wading). Samples are collected in sterile 500mL glass bottles approximately 0.1 meter below the river surface and stored on ice until processed in the laboratory. A field duplicate, rotated among the sampling locations, is collected for each sampling event. A field blank is included at each sampling event. All samples are placed on ice and delivered to the microbiological laboratory at the Massachusetts Department of Environmental Protection (MassDEP) Wall Experiment Station within six hours of sample collection.

Figure 5: Recreation and QMRA locations in Lowell and Lawrence, MA



Two-100mL portions of each sample are filtered for the recovery of *Bacteroides* HF183 human fecal marker and molecular quantification of *E. coli* and enterococci[108]. For each batch of field samples, the following set of laboratory quality control samples are filtered and processed through the entire analytical procedure: method blank (sterile reagent water), lab fortified blank (100mL sterile reagent water and 10mL of wastewater), lab fortified matrix (100mL of one of the field samples with 10mL of wastewater), and a lab duplicate (filtration of an additional 100mL portion of one of the field samples). Wastewater used in the lab fortified blank and lab fortified matrix will be collected from a local wastewater treatment plant. The resulting sample concentrates were subjected to nucleic acid extraction as previously described[108]. Sample extracts are analyzed with droplet digital PCR as previously described[109] for the HF183 human fecal marker, enterococci, and *E. coli*, and reported as genetic copies per 100mL (gc/100mL).

Precipitation: Daily rainfall data will be accessed from the National Centers for Environmental Information (NCEI)[101]. We will select a representative weather station for each sampling location based on proximity to the site and completeness of available data over the study period (2020-2021). Days with total daily rainfall of 0.1 inches or above, as well as the 72 hours after the rain event, will be considered wet weather days.

<u>Exposure Assessment:</u> We will determine the ingested dose of each reference pathogen for three exposure groups: adult swimmers, child swimmers, and adult kayakers. The primary route of exposure that leads to GI illness is incidental ingestion of water during recreation[110]. Incidental ingestion can include ingesting water while swimming, exposure to pathogens via hand-to-mouth behavior, and swallowing inhaled water aerosols[111,112]. We will estimate GI risk for full- and limited-contact recreational activities that occur in the Merrimack River: swimming and kayaking, respectively. Other limited contact recreation that occur on the Merrimack, such as canoeing, fishing, sailing, rowing, and boating, result in similar ingestion rates and so are represented by the kayaking risk estimate[113].

Water ingestion rates: A range of ingestion rates have been published for swimmers and for kayakers [113,114]. For swimming, we will characterize risk based on the 95% confidence intervals of ingestion rates for adults (95% CI: 11-14 mL/h) and children (95% CI: 17-33 mL/h)[114]. For kayaking, age-specific estimates are not available but we will use the ingestion rate estimates for kayakers who capsize for this limited-contact recreation scenario (95% CI: 0-16.5 mL/h)[113].

Exposure duration and frequency: Although activity duration and frequency may vary substantially among individuals and between swimmers and kayakers, duration and frequency of the activity will be standardized at a single, one-hour recreation event to allow for direct comparison of risk between activities.

Pathogen concentration estimates: Reference pathogens in this study are consistent with many QMRA studies[8,58,115] and EPA methods[29]: norovirus, *Cryptosporidium* spp., *Giardia lamblia*, *Campylobacter jejuni*, *E. coli* O157:H7, and *Salmonella*

enterica. We will use measured concentrations of *E. coli* and enterococci to estimate the dose of each reference pathogen (μ^S_{rp}) according to the following equation, adapted from Soller et al[8]:

$$\mu^S_{rp} = \frac{C_{FIB}}{F^S_{FIB} \times 100} \times R^S_{rp} \times p^S_{rp} \times D \times G \quad (3)$$

where

- S is the fecal contamination source;
- C_{FIB} is the concentration of the FIB attributable to the source in the sample (gc/100mL);
- F^S_{FIB} is the concentration of FIB in the source (e.g., sewage; gc/L);
- R^S_{rp} is the concentration of the reference pathogen in the source (pathogens/L);
- p^S_{rp} is the fraction of human-infectious pathogenic strains from source S (assumed to be 1 for human sources);
- D is the duration of the recreation activity (assumed to be 1 hour for this study);
- G is the rate of ingestion for the given recreation scenario (mL/hour).

Ranges for pathogen (R^S_{rp}) and indicator (F^S_{FIB}) input parameters will be determined based on the literature[8,116].

Given the comparatively higher health risk associated with fecal contamination from a human source as opposed to non-human sources[8,22], we will focus on characterizing risk attributable to human sources (i.e., treated wastewater effluent and combined sewer discharge) in the Merrimack River. However, to use Equation 3 to estimate pathogen concentrations from FIB concentration data, we need to know the proportion of FIB concentration that can be attributed to human sources. We will leverage the ratio of *Bacteroides* HF183—a human fecal marker—to FIBs in raw sewage and the ratio of the same parameters in Merrimack River water samples to estimate the proportion of the FIB concentration that can be attributed to human sources[62]. This calculation, adapted from Wang et al[62] can be described as follows:

$$P_{hum,observed} = \hat{R}_M / R_{1_feces,mean(\tau)} \quad (4)$$

where $P_{hum,observed}$ represents the fractional contribution from a human source (e.g., treated wastewater effluent or untreated human sewage) to the overall FIB concentration in the water sample, $\hat{R}_M$ is the observed value of the ratio between the concentrations of the source-specific marker and the FIB, and $R_{hum,mean(\tau)}$ is the mean of the ratio values of the concentrations of the source-specific marker and the FIB from representative samples of the human source at age τ[62]. Equation 1 can then be expressed as:

$$\mu^S_{rp} = \frac{C_{FIB,total} \times P_{hum,observed}}{F^S_{FIB} \times 100} \times R^S_{rp} \times p^S_{rp} \times D \times G \quad (5)$$

where $C_{FIB,total}$ is the total concentration of the FIB in the sample (gc/100mL). The application of this ratio-based source allocation method to environmental samples is limited by variable decay rates between FIBs and the source-specific marker[62]. Molecular enumeration of *E. coli* and enterococci have shown similar rates of decay as *Bacteroides* HF183 in freshwater over 3-4 and 2-3 days, respectively, indicating that using molecular concentrations of FIBs will result in more reliable HF183:FIB ratios in water samples than would be possible with culturable FIB concentrations[62].

Uncertainty Assessment: Many of the input parameters in the risk model have inherent variability and uncertainty. We will conduct probabilistic analysis using Monte Carlo simulation using R software[89] to describe the range of estimated ingested doses of the reference pathogens and the relative contribution to health risk. Consistent with previous work, log-uniform distributions will be used to characterize pathogen (R^S_{rp}) and indicator (F^S_{FIB}) input parameters, while water ingestion rates (G) will be modeled using a lognormal distribution[8]. $P_{hum,observed}$ will be calculated by weather condition (dry and wet weather) for each site as the average of concentration ratios, and it will be modeled in the Monte Carlo analysis using a normal distribution. Each model run will consist of 10,000 iterations for each of the two sites, for each of the three exposure groups, and in wet and dry conditions.

Dose Response Assessment: Dose-response relationships are used in QMRAs to estimate the risk of infection given the dose of the pathogen[29]. Using the dose estimates for each reference pathogen generated from the Monte Carlo analysis, we will estimate the risk of infection (P_{inf}) according to well-established dose-response relationships (Table 4)[8,115].

Risk Characterization: Risk of illness given infection with a reference pathogen (P_{ill_rp}) will be calculated according to morbidity data from the literature[8]. The total probability of illness from a specific source (P^S_{ill}) will be calculated using the probability of illness from each reference pathogen ($P^S_{ill_rp}$):

$$P^S_{ill} = 1 - \Pi_{rp}(1 - P^S_{ill_rp}) \quad (6)$$

Risk will be characterized as total GI illnesses per 1000 recreational users at each site for each activity in wet and dry conditions. Estimated risks for wet days will be compared with estimated risk for dry weather days. Given that past studies have shown increased rates of illness among recreators following wet weather[7,26,75], we expect that estimated risks will be higher on wet weather days.

Table 4: Dose-response relationships for reference pathogens[8,115]

Reference pathogen	Model	Parameter(s)	Parameter Values	Reference
Norovirus	Hypergeometric	alpha	0.04	Teunis et al, 2008a[117]
		beta	0.055	
Cryptosporidium spp.	Exponential	r	0.09	US EPA 2006[118]
Giardia lamblia	Exponential	r	0.0199	Rose and Gerba, 1991[119]
Campylobacter jejuni	Hypergeometric	alpha	0.024	Teunis et al, 2005[120]
		beta	0.011	
E. coli O157:H7	Beta-Poisson	alpha	0.248	Teunis et al, 2008b[121]
		beta	48.8	
Salmonella enterica	Beta-Poisson	alpha	0.3126	Fazil and Haas, 1996[122]
		beta	2884	

Strengths & Limitations

Strengths: A strength of the proposed work is leveraging FIB and a human MST to estimate human health risk; while molecular quantification of FIB and MST requires more laboratory sophistication than enumeration of culturable FIBs, these molecular methods are more reliable and common than direct measurements of pathogens. Another strength of this project is using a probabilistic approach to account for uncertainty and variability in the parameter inputs. Finally, the proposed QMRA will characterize risk at two locations in the Merrimack River with documented recreation use for three exposure groups and under dry and wet weather conditions, broadening the applicability of the results for community stakeholders.

Limitations: A major limitation of this proposed study is accounting for differential decay rates in FIBs and MSTs[62]. In the Merrimack River, it is possible that several sources of human-associated fecal contamination are present in different times and locations, further complicating the ratios between FIBs and MSTs. The Monte Carlo analysis will account for some of the uncertainty this may introduce, but future work that measures non-human MSTs and pathogens directly could be useful in validating the ratio-based approach to source apportionment proposed here. The exposure assessment may not accurately describe the activities at these two locations, but the use of a recreational use survey throughout the Merrimack River watershed could improve the accuracy.

References

1. US EPA. *National Water Quality Inventory: Report to Congress*.; 2017. https://www.epa.gov/sites/production/files/2017-12/documents/305brtc_finalowow_08302017.pdf
2. Messner M, Shaw S, Regli S, Rotert K, Blank V, Soller J. An approach for developing a national estimate of waterborne disease due to drinking water and a national estimate model application. *J Water Health*. 2006;4 Suppl 2:201-240. doi:10.2166/wh.2006.024
3. DeFlorio-Barker S, Wing C, Jones RM, Dorevitch S. Estimate of incidence and cost of recreational waterborne illness on United States surface waters. *Environmental Health*. 2018;17(1):3. doi:10.1186/s12940-017-0347-9
4. 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):873-879. doi:10.1289/ehp.1408971
5. Tornevi A, Axelsson G, Forsberg B. Association between precipitation upstream of a drinking water utility and nurse advice calls relating to acute gastrointestinal illnesses. *PLoS One*. 2013;8(7):e69918. doi:10.1371/journal.pone.0069918
6. De Roos AJ, Kondo MC, Robinson LF, et al. Heavy precipitation, drinking water source, and acute gastrointestinal illness in Philadelphia, 2015-2017. *PLoS One*. 2020;15(2). doi:10.1371/journal.pone.0229258
7. Russo GS, Eftim SE, Goldstone AE, Dufour AP, Nappier SP, Wade TJ. Evaluating health risks associated with exposure to ambient surface waters during recreational activities: A systematic review and meta-analysis. *Water Research*. 2020;176:115729. doi:10.1016/j.watres.2020.115729
8. Soller JA, Schoen ME, Bartrand T, Ravenscroft JE, Ashbolt NJ. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. *Water Research*. 2010;44(16):4674-4691. doi:10.1016/j.watres.2010.06.049
9. Ham Y-S, Kobori H, Takasago M. Effects of combined sewer overflow and stormwater on indicator bacteria concentrations in the Tama River due to the high population density of Tokyo Metropolitan area. *Environ Monit Assess*. 2008;152(1):459. doi:10.1007/s10661-008-0330-6
10. Rechenburg A, Koch Ch, Claßen Th, Kistemann Th. Impact of sewage treatment plants and combined sewer overflow basins on the microbiological quality of surface water. *Water Science and Technology*. 2006;54(3):95-99. doi:10.2166/wst.2006.454
11. Passerat J, Ouattara NK, Mouchel J-M, Vincent Rocher, Servais P. Impact of an intense combined sewer overflow event on the microbiological water quality of the Seine River. *Water Research*. 2011;45(2):893-903. doi:10.1016/j.watres.2010.09.024
12. Madoux-Humery A-S, Dorner S, Sauvé S, et al. Temporal variability of combined sewer overflow contaminants: Evaluation of wastewater micropollutants as tracers of fecal contamination. *Water Research*. 2013;47(13):4370-4382. doi:10.1016/j.watres.2013.04.030
13. Madoux-Humery A-S, Dorner S, Sauvé S, et al. The effects of combined sewer overflow events on riverine sources of drinking water. *Water Research*. 2016;92:218-227. doi:10.1016/j.watres.2015.12.033
14. US EPA. Merrimack River Watershed. Published 2019. Accessed December 1, 2020. https://epa.maps.arcgis.com/apps/MapSeries/index.html?appid=922e1c016c6e42b199f902d1cfb84bbd
15. Soller J, Bartrand T, Ravenscroft J, et al. Estimated human health risks from recreational exposures to stormwater runoff containing animal faecal material. *Environmental Modelling & Software*. 2015;72:21-32. doi:10.1016/j.envsoft.2015.05.018
16. Bushon RN, Brady AMG, Christensen ED, Stelzer EA. Multi-Year Microbial Source Tracking Study Characterizing Fecal Contamination in an Urban Watershed. *Water Environment Research*. 2017;89(2):127-143. doi:https://doi.org/10.2175/106143016X14798353399412
17. Marsalek J, Rochfort Q. Urban wet-weather flows: Sources of fecal contamination impacting on recreational waters and threatening drinking-water sources. *Journal of Toxicology and Environmental Health, Part A*. 2004;67(20-22):1765-1777. doi:10.1080/15287390490492430
18. Okoh AI, Sibanda T, Gusha SS. Inadequately Treated Wastewater as a Source of Human Enteric Viruses in the Environment. *International Journal of Environmental Research and Public Health*. 2010;7(6):2620-2637. doi:10.3390/ijerph7062620

19. Mansfeldt C, Deiner K, Mächler E, et al. Microbial community shifts in streams receiving treated wastewater effluent. *Science of The Total Environment*. 2020;709:135727. doi:10.1016/j.scitotenv.2019.135727
20. Olds HT, Corsi SR, Dila DK, Halmo KM, Bootsma MJ, McLellan SL. High levels of sewage contamination released from urban areas after storm events: A quantitative survey with sewage specific bacterial indicators. *PLoS Med*. 2018;15(7). doi:10.1371/journal.pmed.1002614
21. Sowah RA, Habteselassie MY, Radcliffe DE, Bauske E, Risse M. Isolating the impact of septic systems on fecal pollution in streams of suburban watersheds in Georgia, United States. *Water Research*. 2017;108:330-338. doi:10.1016/j.watres.2016.11.007
22. Schoen ME, Ashbolt NJ. Assessing Pathogen Risk to Swimmers at Non-Sewage Impacted Recreational Beaches. *Environmental Science & Technology*. 2010;44:2286-2291.
23. Fewtrell L, Kay D. Recreational Water and Infection: A Review of Recent Findings. *Curr Environ Health Rep*. 2015;2(1):85-94. doi:10.1007/s40572-014-0036-6
24. Leonard AFC, Singer A, Ukoumunne OC, Gaze WH, Garside R. Is it safe to go back into the water? A systematic review and meta-analysis of the risk of acquiring infections from recreational exposure to seawater. *Int J Epidemiol*. 2018;47(2):572-586. doi:10.1093/ije/dyx281
25. Arnold BF, Wade TJ, Benjamin-Chung J, et al. Acute Gastroenteritis and Recreational Water: Highest Burden Among Young US Children. *Am J Public Health*. 2016;106(9):1690-1697. doi:10.2105/AJPH.2016.303279
26. Arnold BF, Schiff KC, Ercumen A, et al. Acute Illness Among Surfers After Exposure to Seawater in Dry- and Wet-Weather Conditions. *Am J Epidemiol*. 2017;186(7):866-875. doi:10.1093/aje/kwx019
27. Säve-Söderbergh M, Åkesson A, Simonsson M, Toljander J. Endemic gastrointestinal illness and change in raw water source and drinking water production – A population-based prospective study. *Environment International*. 2020;137:105575. doi:10.1016/j.envint.2020.105575
28. Graciaa DS, Cope JR, Roberts VA, et al. Outbreaks Associated with Untreated Recreational Water — United States, 2000–2014. *American Journal of Transplantation*. 2018;18(8):2083-2087. doi:https://doi.org/10.1111/ajt.15002
29. US EPA. *Microbiological Risk Assessment (MRA) Tools, Methods, and Approaches for Water Media*.; 2014.
30. Guzman Herrador BR, de Blasio BF, MacDonald E, 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. doi:10.1186/s12940-015-0014-y
31. Reynolds KA, Mena KD, Gerba CP. Risk of waterborne illness via drinking water in the United States. *Rev Environ Contam Toxicol*. 2008;192:117-158. doi:10.1007/978-0-387-71724-1_4
32. Schuster CJ, Aramini JJ, Ellis AG, et al. Infectious Disease Outbreaks Related to Drinking Water in Canada, 1974–2001. *Can J Public Health*. 2005;96(4):254-258. doi:10.1007/BF03405157
33. Murphy HM, Thomas MK, Medeiros DT, McFADYEN S, Pintar KDM. Estimating the number of cases of acute gastrointestinal illness (AGI) associated with Canadian municipal drinking water systems. *Epidemiol Infect*. 2016;144(7):1371-1385. doi:10.1017/S0950268815002083
34. Moreira NA, Bondelind M. Safe drinking water and waterborne outbreaks. *J Water Health*. 2017;15(1):83-96. doi:10.2166/wh.2016.103
35. 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):1194-1199. doi:10.2105/ajph.91.8.1194
36. Drayna P, McLellan SL, Simpson P, Li S-H, Gorelick MH. Association between rainfall and pediatric emergency department visits for acute gastrointestinal illness. *Environ Health Perspect*. 2010;118(10):1439-1443. doi:10.1289/ehp.0901671
37. Bush KF, O'Neill MS, Li S, et al. Associations between extreme precipitation and gastrointestinal-related hospital admissions in Chennai, India. *Environ Health Perspect*. 2014;122(3):249-254. doi:10.1289/ehp.1306807
38. Lin S, Sun M, Fitzgerald E, Hwang S-A. Did summer weather factors affect gastrointestinal infection hospitalizations in New York State? *Sci Total Environ*. 2016;550:38-44. doi:10.1016/j.scitotenv.2015.12.153
39. 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. doi:10.1371/journal.pone.0173794

40. Chhetri BK, Takaro TK, Balshaw R, 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. doi:10.2166/wh.2017.100
41. Lai H, Hales S, Woodward A, et al. Effects of heavy rainfall on waterborne disease hospitalizations among young children in wet and dry areas of New Zealand. *Environment International*. 2020;145:106136. doi:10.1016/j.envint.2020.106136
42. Uejio CK, Yale SH, Malecki K, Borchardt MA, Anderson HA, Patz JA. Drinking water systems, hydrology, and childhood gastrointestinal illness in Central and Northern Wisconsin. *Am J Public Health*. 2014;104(4):639-646. doi:10.2105/AJPH.2013.301659
43. Tornevi A, Barregård L, Forsberg B. Precipitation and primary health care visits for gastrointestinal illness in Gothenburg, Sweden. *PLoS One*. 2015;10(5):e0128487. doi:10.1371/journal.pone.0128487
44. Tornevi A, Simonsson M, Forsberg B, Säve-Söderbergh M, Toljander J. Efficacy of water treatment processes and endemic gastrointestinal illness – A multi-city study in Sweden. *Water Research*. 2016;102:263-270. doi:10.1016/j.watres.2016.06.018
45. Vanden Esschert KL, Mattioli MC, Hilborn ED, et al. Outbreaks Associated with Untreated Recreational Water - California, Maine, and Minnesota, 2018-2019. *MMWR Morb Mortal Wkly Rep*. 2020;69(25):781-783. doi:10.15585/mmwr.mm6925a3
46. Dorevitch S, Pratap P, Wroblewski M, et al. Health Risks of Limited-Contact Water Recreation. *Environ Health Perspect*. 2012;120(2):192-197. doi:10.1289/ehp.1103934
47. Prüss A. Review of epidemiological studies on health effects from exposure to recreational water. *Int J Epidemiol*. 1998;27(1):1-9. doi:10.1093/ije/27.1.1
48. Wade TJ, Pai N, Eisenberg JNS, Colford JM. Do U.S. Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. *Environ Health Perspect*. 2003;111(8):1102-1109.
49. US EPA. *Recreational Water Quality Criteria*. Office of Water; 2012.
50. Soller JA, Schoen ME, Varghese A, et al. Human health risk implications of multiple sources of faecal indicator bacteria in a recreational waterbody. *Water Research*. 2014;66:254-264. doi:10.1016/j.watres.2014.08.026
51. Zhang Q, Gallard J, Wu B, et al. Synergy between quantitative microbial source tracking (qMST) and quantitative microbial risk assessment (QMRA): A review and prospectus. *Environ Int*. 2019;130:104703. doi:10.1016/j.envint.2019.03.051
52. Arnold BF, Schiff KC, Griffith JF, et al. Swimmer illness associated with marine water exposure and water quality indicators: impact of widely used assumptions. *Epidemiology*. 2013;24(6):845-853. doi:10.1097/01.ede.0000434431.06765.4a
53. Marion JW, Lee J, Lemeshow S, Buckley TJ. Association of gastrointestinal illness and recreational water exposure at an inland U.S. beach. *Water Res*. 2010;44(16):4796-4804. doi:10.1016/j.watres.2010.07.065
54. Abdelzaher AM, Wright ME, Ortega C, et al. Daily measures of microbes and human health at a non-point source marine beach. *J Water Health*. 2011;9(3):443-457. doi:10.2166/wh.2011.146
55. Holcomb DA, Stewart JR. Microbial Indicators of Fecal Pollution: Recent Progress and Challenges in Assessing Water Quality. *Curr Environ Health Rep*. 2020;7(3):311-324. doi:10.1007/s40572-020-00278-1
56. Boehm AB, Van De Werfhorst LC, Griffith JF, et al. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. *Water Research*. 2013;47(18):6812-6828. doi:10.1016/j.watres.2012.12.046
57. Dufour A, Wade TJ, Kay D. Epidemiological studies on swimmer health effects associated with potential exposure to zoonotic pathogens in bathing beach water – a review. In: Dufour AP, Bartram J, Bos R, Gannon V, eds. *Animal Waste, Water Quality and Human Health*. Emerging issues in water and infectious disease series. Published on behalf of the World Health Organization by IWA Publishing; 2012.
58. Soller JA, Eftim S, Wade TJ, et al. Use of quantitative microbial risk assessment to improve interpretation of a recreational water epidemiological study. *Microbial Risk Analysis*. 2016;1:2-11. doi:10.1016/j.mran.2015.04.001
59. Boehm AB, Graham KE, Jennings WC. Can We Swim Yet? Systematic Review, Meta-Analysis, and Risk Assessment of Aging Sewage in Surface Waters. *Environ Sci Technol*. 2018;52(17):9634-9645. doi:10.1021/acs.est.8b01948

60. Bofill-Mas S, Rusiñol M. Recent trends on methods for the concentration of viruses from water samples. *Current Opinion in Environmental Science & Health*. 2020;16:7-13. doi:10.1016/j.coesh.2020.01.006
61. Haramoto E, Kitajima M, Hata A, et al. A review on recent progress in the detection methods and prevalence of human enteric viruses in water. *Water Research*. 2018;135:168-186. doi:10.1016/j.watres.2018.02.004
62. Wang D, Farnleitner AH, Field KG, Green HC, Shanks OC, Boehm AB. Enterococcus and Escherichia coli fecal source apportionment with microbial source tracking genetic markers – Is it feasible? *Water Research*. 2013;47(18):6849-6861. doi:10.1016/j.watres.2013.02.058
63. Ahmed W, Hamilton KA, Lobos A, et al. Quantitative microbial risk assessment of microbial source tracking markers in recreational water contaminated with fresh untreated and secondary treated sewage. *Environ Int*. 2018;117:243-249. doi:10.1016/j.envint.2018.05.012
64. US EPA O. Combined Sewer Overflow Frequent Questions. US EPA. Published October 14, 2015. Accessed December 1, 2020. https://www.epa.gov/npdes/combined-sewer-overflow-frequent-questions
65. Donovan E, Unice K, Roberts JD, Harris M, Finley B. Risk of gastrointestinal disease associated with exposure to pathogens in the water of the Lower Passaic River. *Appl Environ Microbiol*. 2008;74(4):994-1003. doi:10.1128/AEM.00601-07
66. Hata A, Katayama H, Kojima K, et al. Effects of rainfall events on the occurrence and detection efficiency of viruses in river water impacted by combined sewer overflows. *Science of The Total Environment*. 2014;468-469:757-763. doi:10.1016/j.scitotenv.2013.08.093
67. Arnone RD, Walling JP, Arnone (corresponding RD, Walling JP. Evaluating Cryptosporidium and Giardia concentrations in combined sewer overflow. *Journal of Water and Health*. Published online 2006:157-165.
68. McGinnis S, Spencer S, Firnstahl A, et al. Human Bacteroides and total coliforms as indicators of recent combined sewer overflows and rain events in urban creeks. *Science of The Total Environment*. 2018;630:967-976. doi:10.1016/j.scitotenv.2018.02.108
69. Eramo A, Delos Reyes H, Fahrenfeld NL. Partitioning of Antibiotic Resistance Genes and Fecal Indicators Varies Intra and Inter-Storm during Combined Sewer Overflows. *Front Microbiol*. 2017;8. doi:10.3389/fmicb.2017.02024
70. Kuhn KG, Nygård KM, Guzman-Herrador B, et al. Campylobacter infections expected to increase due to climate change in Northern Europe. *Sci Rep*. 2020;10(1):13874. doi:10.1038/s41598-020-70593-y
71. Chhetri BK, Galanis E, Sobie S, et al. Projected local rain events due to climate change and the impacts on waterborne diseases in Vancouver, British Columbia, Canada. *Environ Health*. 2019;18(1):116. doi:10.1186/s12940-019-0550-y
72. Redman RL, Nenn CA, Eastwood D, Gorelick MH. Pediatric emergency department visits for diarrheal illness increased after release of undertreated sewage. *Pediatrics*. 2007;120(6):e1472-1475. doi:10.1542/peds.2007-0283
73. Sokolova E, Petterson SR, Dienus O, Nyström F, Lindgren P-E, Pettersson TJR. Microbial risk assessment of drinking water based on hydrodynamic modelling of pathogen concentrations in source water. *Science of The Total Environment*. 2015;526:177-186. doi:10.1016/j.scitotenv.2015.04.040
74. Taghipour M, Shakibaeinia A, Sylvestre É, Tolouei S, Dorner S. Microbial risk associated with CSOs upstream of drinking water sources in a transboundary river using hydrodynamic and water quality modeling. *Science of The Total Environment*. 2019;683:547-558. doi:10.1016/j.scitotenv.2019.05.130
75. Harder-Lauridsen NM, Kuhn KG, Erichsen AC, Mølbak K, Ethelberg S. Gastrointestinal Illness among Triathletes Swimming in Non-Polluted versus Polluted Seawater Affected by Heavy Rainfall, Denmark, 2010-2011. *PLOS ONE*. 2013;8(11):e78371. doi:10.1371/journal.pone.0078371
76. Brokamp C, Beck AF, Muglia L, Ryan P. Combined sewer overflow events and childhood emergency department visits: A case-crossover study. *Sci Total Environ*. 2017;607-608:1180-1187. doi:10.1016/j.scitotenv.2017.07.104
77. Westphal K, D'Orlando A. *Technical Memorandum: Integrated Decision Framework for the Merrimack River Basin*. Brown and Caldwell; 2020.
78. Combined Sewer Overflows. Merrimack River Watershed Council. Accessed November 23, 2020. http://merrimack.org/education/cso/
79. US EPA O. Approved CWA Microbiological Methods for Ambient Water. US EPA. Published August 17, 2015. Accessed April 1, 2021. https://www.epa.gov/cwa-methods/approved-cwa-microbiological-methods-ambient-water

80. National Water Quality Monitoring Council. Water Quality Data Home. Accessed March 16, 2021. https://www.waterqualitydata.us/
81. Massachusetts Environmental Public Health Tracking. Recreational Water. Accessed March 17, 2021. https://matracking.ehs.state.ma.us/Environmental-Data/recreational-water/
82. Nagel, D D, Wollrab S, Parkes-Payne S, Peterson E, Isaak D, Ver Hoef J. National Stream Internet hydrography datasets for spatial-stream-network (SSN) analysis. Published online 2017. Accessed March 18, 2021. https://www.fs.fed.us/rm/boise/AWAE/projects/NationalStreamInternet/NSI_network.html
83. USGS. National Hydrography Dataset. Published 2021. Accessed March 18, 2021. https://www.usgs.gov/core-science-systems/ngp/national-hydrography
84. US EPA. StreamCat Dataset. Published September 21, 2015. Accessed March 31, 2021. https://www.epa.gov/national-aquatic-resource-surveys/streamcat-dataset-0
85. PRISM Climate Group. PRISM Climate Data. Accessed March 18, 2021. https://prism.oregonstate.edu/
86. US EPA O. Facility Registry Service (FRS). US EPA. Published April 23, 2018. Accessed March 18, 2021. https://www.epa.gov/frs
87. Peterson EE, Hoef JMV. STARS: An ArcGIS Toolset Used to Calculate the Spatial Information Needed to Fit Spatial Statistical Models to Stream Network Data. *J Stat Softw*. 2014;56(2):1-17.
88. Esri Inc. *ArcGIS Pro*.; 2019.
89. R Core Team. *R: The R Project for Statistical Computing*.; 2021. Accessed March 18, 2021. https://www.r-project.org/
90. Hoef JV, Peterson E, Clifford D, Shah R. SSN: An R Package for Spatial Statistical Modeling on Stream Networks. *Journal of Statistical Software*. 2014;56(1):1-45. doi:10.18637/jss.v056.i03
91. Isaak DJ, Peterson EE, Hoef JMV, et al. Applications of spatial statistical network models to stream data. *WIREs Water*. 2014;1(3):277-294. doi:https://doi.org/10.1002/wat2.1023
92. Peterson EE, Ver Hoef JM. A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*. 2010;91(3):644-651. doi:10.1890/08-1668.1
93. Hoeting JA, Davis RA, Merton AA, Thompson SE. Model Selection For Geostatistical Models. *Ecological Applications*. 2006;16(1):87-98. doi:https://doi.org/10.1890/04-0576
94. Holcomb DA, Messier KP, Serre ML, Rowny JG, Stewart JR. Geostatistical Prediction of Microbial Water Quality Throughout a Stream Network Using Meteorology, Land Cover, and Spatiotemporal Autocorrelation. *Environ Sci Technol*. 2018;52(14):7775-7784. doi:10.1021/acs.est.8b01178
95. Neill AJ, Tetzlaff D, Strachan NJC, et al. Using spatial-stream-network models and long-term data to understand and predict dynamics of faecal contamination in a mixed land-use catchment. *Science of The Total Environment*. 2018;612:840-852. doi:10.1016/j.scitotenv.2017.08.151
96. Jat P, Serre ML. A novel geostatistical approach combining Euclidean and gradual-flow covariance models to estimate fecal coliform along the Haw and Deep rivers in North Carolina. *Stoch Environ Res Risk Assess*. 2018;32(9):2537-2549. doi:10.1007/s00477-018-1512-6
97. Mainali J, Chang H, Chun Y. A review of spatial statistical approaches to modeling water quality. *Progress in Physical Geography: Earth and Environment*. 2019;43(6):801-826. doi:10.1177/0309133319852003
98. US Census. American Community Survey 5-Year Data (2009-2019). The United States Census Bureau. Accessed April 7, 2021. https://www.census.gov/data/developers/data-sets/acs-5year.html
99. CHIA. Case Mix Data. Published 2021. Accessed March 20, 2021. https://www.chiamass.gov/case-mix-data/
100. World Health Organization. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Fifth Edition*. World Health Organization; 2015. https://apps.who.int/iris/handle/10665/246208
101. NOAA. National Centers for Environmental Information (NCEI). Published 2021. Accessed March 20, 2021. https://www.ncei.noaa.gov/
102. Morral-Puigmal C, Martínez-Solanas È, Villanueva CM, Basagaña X. Weather and gastrointestinal disease in Spain: A retrospective time series regression study. *Environ Int*. 2018;121(Pt 1):649-657. doi:10.1016/j.envint.2018.10.003
103. Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Statistics in Medicine*. 2010;29(21):2224-2234. doi:https://doi.org/10.1002/sim.3940

104. Gasparrini A. Distributed Lag Linear and Non-Linear Models in R: The Package dlnm. *J Stat Softw*. 2011;43(8):1-20.
105. City of Lowell Recreation Office. Michael Rynne Bath House - Rynne Beach. Published 2021. Accessed April 7, 2021. https://lowellrec.com/info/facilities/area_info.aspx?FacilityID=14742&AreaID=14957
106. Commonwealth of Massachusetts. Lowell Heritage State Park. Accessed April 7, 2021. https://www.mass.gov/locations/lowell-heritage-state-park
107. GLCBP. *Greater Lawrence Community Boating Program 2018 Annual Report*. GLCBP
108. Hill VR, Narayanan J, Gallen RR, Ferdinand KL, Cromeans T, Vinjé J. Development of a Nucleic Acid Extraction Procedure for Simultaneous Recovery of DNA and RNA from Diverse Microbes in Water. *Pathogens*. 2015;4(2):335-354. doi:10.3390/pathogens4020335
109. Cao Y, Raith MR, Griffith JF. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. *Water Res*. 2015;70:337-349. doi:10.1016/j.watres.2014.12.008
110. World Health Organization, ed. *Guidelines for Safe Recreational Water Environments*. World Health Organization; 2003.
111. US EPA. *Update for Chapter 3 of the Exposure Factors Handbook: Ingestion of Water and Other Select Liquids*. National Center for Environmental Assessment, Office of Research and Development; 2019:157.
112. Sunger N, Haas CN. Quantitative Microbial Risk Assessment for Recreational Exposure to Water Bodies in Philadelphia. *Water Environment Research*. 2015;87(3):211-222. doi:https://doi.org/10.2175/106143015X14212658613073
113. Dorevitch S, Panthi S, Huang Y, et al. Water ingestion during water recreation. *Water Research*. 2011;45(5):2020-2028. doi:10.1016/j.watres.2010.12.006
114. Dufour AP, Behymer TD, Cantú R, Magnuson M, Wymer LJ. Ingestion of swimming pool water by recreational swimmers. *Journal of Water and Health*. 2017;15(3):429-437. doi:10.2166/wh.2017.255
115. Soller JA, Schoen ME, Varghese A, et al. Human health risk implications of multiple sources of faecal indicator bacteria in a recreational waterbody. *Water Research*. 2014;66:254-264. doi:10.1016/j.watres.2014.08.026
116. Schoen ME, Soller JA, Ashbolt NJ. Evaluating the importance of faecal sources in human-impacted waters. *Water Research*. 2011;45(8):2670-2680. doi:10.1016/j.watres.2011.02.025
117. Teunis PFM, Moe CL, Liu P, et al. Norwalk virus: How infectious is it? *Journal of Medical Virology*. 2008;80(8):1468-1476. doi:https://doi.org/10.1002/jmv.21237
118. US EPA. *National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rule (LT2ESWTR)*.; 2006.
119. Rose JB, Gerba CP. Use of Risk Assessment for Development of Microbial Standards. *Water Science and Technology*. 1991;24(2):29-34. doi:10.2166/wst.1991.0025
120. Teunis P, Brandhof WVD, Nauta M, Wagenaar J, Kerkhof HVD, Pelt WV. A reconsideration of the Campylobacter dose–response relation. *Epidemiology & Infection*. 2005;133(4):583-592. doi:10.1017/S0950268805003912
121. Teunis PFM, Ogden ID, Strachan NJC. Hierarchical dose response of E. coli O157:H7 from human outbreaks incorporating heterogeneity in exposure. *Epidemiology & Infection*. 2008;136(6):761-770. doi:10.1017/S0950268807008771
122. Fazil A, Haas C (Eds.). *A Quantitative Risk Assessment Model for Salmonella*. AWWA; 1996.