



36 **Abstract**

37 *Background*

38 Combined sewer overflow (CSO) events release untreated wastewater into surface waterbodies  
39 during heavy precipitation or snowmelt. Combined sewer systems serve approximately 40  
40 million people in the United States, primarily in urban and suburban municipalities in the  
41 Midwest and Northeast. Predicted increases in heavy precipitation events driven by climate  
42 change underscore the importance of quantifying potential health risks associated with CSO  
43 events.

44 *Objectives*

45 The aims of this study were: to 1) estimate the association between CSO events (2014-2019) and  
46 emergency department (ED) visits for acute gastrointestinal illness (AGI) among Massachusetts  
47 municipalities that border a CSO-impacted river, and 2) determine if associations differ by  
48 municipal drinking water source.

49 *Methods*

50 A case time series design was used to estimate the association between daily cumulative  
51 upstream CSO discharge in the prior four days and ED visits for AGI, adjusting for temporal  
52 trends, temperature, and precipitation. Associations between CSO events and AGI were also  
53 compared by municipal drinking water source (CSO-impacted river vs. other sources).

54 *Results*

55 Extreme upstream CSO discharge events (>95th percentile by cumulative volume) were  
56 associated with a cumulative risk ratio (CRR) of AGI of 1.22 (95% CI: 1.05, 1.42) over the next  
57 four days for all municipalities, and the association was robust after adjusting for precipitation  
58 (1.17 [0.98, 1.39]). In municipalities with CSO-impacted drinking water sources, the adjusted

59 association was somewhat less pronounced following 95th percentile CSO events (1.05 [0.82,  
60 1.33]). The adjusted CRR of AGI was 1.62 in all municipalities following 99th percentile CSO  
61 events (95% CI: 1.04, 2.51) and not meaningfully different across strata defined by drinking  
62 water source.

### 63 *Discussion*

64 In municipalities bordering a CSO-impacted river in Massachusetts, extreme CSO events are  
65 associated with higher risk of AGI within four days. The largest CSO events are associated with  
66 increased risk of AGI regardless of drinking water source.

67

68 **Introduction**

69 In more than 700 municipalities in the United States, combined sewer systems (CSS)  
70 collect residential and non-residential sewage, industrial waste, and stormwater that flow  
71 together to a wastewater treatment facility. CSS discharge untreated combined wastewater into  
72 rivers and lakes when the volume of water in the common collection pipe exceeds treatment  
73 capacity, potentially posing risks to the health of communities and ecosystems downstream of  
74 discharge points. Discharge events—called combined sewer overflow (CSO) events—are driven  
75 by heavy precipitation and/or runoff from snowmelt.<sup>1,2</sup> Observed and predicted increases in the  
76 intensity and frequency of precipitation events due to climate change in the Northeast and  
77 Midwest where most CSS are located underscore the importance of understanding and  
78 addressing the health risks associated with CSO events.<sup>3–5</sup>

79 CSO events introduce viral, bacterial, and protozoan pathogens to surface waterbodies, as  
80 well as chemical and physical contaminants.<sup>1,2,6–14</sup> Compared to dry weather conditions, CSO  
81 events can lead to one to two orders of magnitude increases in the concentration of waterborne  
82 pathogens and fecal indicator organisms in receiving waterbodies.<sup>2,6,15</sup> Exposure to sewage-  
83 associated pathogens can lead to gastrointestinal, respiratory, and soft tissue infections among  
84 recreational and drinking water users.<sup>16–21</sup> The association between fecal contamination of  
85 surface waters and gastrointestinal illness is well-documented in settings with modern  
86 wastewater infrastructure,<sup>22–25</sup> as is the relationship between precipitation, which is the primary  
87 driver of CSO events, and gastrointestinal illness.<sup>26–28</sup> However, relatively little research has  
88 been conducted to determine the direct association between CSO events and health outcomes.

89 The few studies that have investigated the relationship between CSO events and health  
90 outcomes suggest a link between discharge events and acute gastrointestinal illness (AGI).

91 Studies of recreational use of CSO-impacted waterbodies suggest that the risk of AGI among  
92 swimmers can increase up to five times after a CSO event compared to when no CSO event  
93 occurred,<sup>29</sup> and that the likelihood that recreators will develop AGI can vary widely by activity  
94 and level of water contact.<sup>9</sup> Increases in pediatric cases of AGI following large volume partially  
95 treated sewage discharge events in a drinking water source<sup>30</sup> and increased risk of AGI following  
96 extreme precipitation events in a region with a CSO-impaired drinking water source<sup>31</sup> suggest  
97 that an underexplored pathway of exposure may be contaminated drinking water. Importantly,  
98 neither of these studies used CSO event data in their analyses, but instead used partially treated  
99 sewage discharge event or precipitation data to characterize exposure, respectively.<sup>30,31</sup> Research  
100 has demonstrated elevated rates of AGI following large volume CSO events among residents of  
101 Atlanta<sup>32</sup> and increased odds of AGI among children living near a CSO outfall after any CSO  
102 event.<sup>33</sup> In all studies, the association between discharge events and reported AGI has a lag of  
103 two to eight days, which could be due to pathogen travel time in the waterbody, timing of  
104 recreational activity, drinking water treatment and distribution, pathogen infection and  
105 incubation, and/or decision to seek care.<sup>30-33</sup>

106         There remain key gaps in our understanding of the association between CSO events and  
107 health, including the robustness of the association, relevant exposure pathways, and the role of  
108 precipitation events. Accordingly, we sought to assess the association between CSO events and  
109 emergency department (ED) visits for AGI in the Massachusetts municipalities that border the  
110 Merrimack River, a drinking water and recreational resource for surrounding communities. We  
111 evaluate whether the association between CSO events and AGI differs by municipal drinking  
112 water source and after adjusting for precipitation events. We hypothesize that 1) upstream CSO  
113 events increase risk of AGI in downstream municipalities after adjusting for precipitation; and 2)

114 the risk of AGI in municipalities that exclusively source their drinking water from the Merrimack  
115 River is more strongly associated with upstream CSO discharge than in municipalities with other  
116 drinking water sources. To our knowledge, this is the first study that directly investigates the role  
117 of drinking water source in the association between CSO event discharge volume and AGI.

## 118 **Methods**

### 119 *Study Site*

120 The Merrimack River flows 115 miles through New Hampshire (NH) and northern  
121 Massachusetts (MA) before flowing into the Atlantic Ocean.<sup>34</sup> At 5,010 square miles, the  
122 Merrimack watershed is the fourth largest in New England and home to over 2.6 million  
123 people.<sup>34</sup> Over 500,000 people rely on the Merrimack as a drinking water source, and it is a  
124 regional recreational resource for many more.<sup>35</sup>

125 The main stem of the Merrimack River flows through five of the six CSS municipalities  
126 in the watershed (Figure 1). The sixth CSS community in the watershed is located 51 miles  
127 upstream from the confluence with the main stem of the Merrimack River along a tributary and  
128 not included in this study due to its distance from the main stem of the river. Four municipalities  
129 in Massachusetts exclusively source their drinking water from the river, while three  
130 intermittently depend on the Merrimack River as a drinking water source. All 17 Massachusetts  
131 municipalities that border the Merrimack River are included in this study.

### 132 *Exposure Data and Classification*

133 US Environmental Protection Agency (EPA) Region 1 and the City of Manchester, NH  
134 provided daily CSO event data for the study period. Locations of CSO outfalls are documented  
135 and publicly available via an online dashboard.<sup>35</sup> CSO event data were aggregated by  
136 municipality and date, yielding a time series of total daily CSO discharge volume per CSS

137 municipality. Cumulative daily upstream CSO discharge volume was calculated for each study  
138 municipality based on which CSS municipalities are located upstream. Figure S1 shows the 17  
139 study municipalities and their spatial relationships to CSS.

140 Total daily precipitation and mean daily temperature were estimated using PRISM  
141 (Parameter-elevation Relationships on Independent Slopes Model) Climate Data.<sup>36</sup> The centroid  
142 for each municipality in the study was used to identify the corresponding 4 km<sup>2</sup> PRISM grid cell  
143 and generate a daily time series of total precipitation and average temperature for each  
144 municipality. Cross-correlations were performed to determine how closely and at what point in  
145 the time series precipitation and CSO discharge are most correlated.

146 Study days were categorized in one of the following ways: 1) extreme CSO/precipitation  
147 events ( $\geq 95^{\text{th}}$  percentile events on the basis of cumulative daily upstream CSO discharge or total  
148 daily precipitation, respectively), 2) all other CSO/precipitation events ( $<95^{\text{th}}$  percentile events),  
149 or 3) no CSO/precipitation events. Days with no CSO/precipitation events were used as the  
150 referent category. By characterizing CSO discharge volume and precipitation categorically, risk  
151 of AGI following the largest CSO/precipitation events are compared to risk of AGI following  
152 days without any events.

### 153 *Outcome Data*

154 Statewide emergency department (ED) visit data for Massachusetts were obtained from  
155 the Center for Health Information and Analysis (CHIA), a state-level agency that collects and  
156 normalizes patient-level and financial data from all hospitals across Massachusetts. The  
157 Massachusetts Acute Hospital Case Mix Database contains data on ED visits as well as hospital  
158 inpatient discharges and outpatient observation stays for patients who originated in the ED.<sup>37</sup>  
159 Data available for each visit include date, municipality of residence, patient age and sex, and

160 primary and secondary diagnoses as defined by the International Classification of Diseases (ICD)  
161 9<sup>th</sup> or 10<sup>th</sup> revision. Claims were de-identified and visits are assumed for analysis to represent  
162 separate individuals. Claims data were available from January 1, 2014 through September 30,  
163 2019.

164 All individuals with AGI who entered a Massachusetts hospital through the ED during  
165 the study period, and whose permanent or temporary municipality of residence is in the study  
166 area are included. Cases of AGI were defined as any visit that included one or more of the  
167 following as a primary or secondary diagnosis: specified and unspecified infectious intestinal  
168 diseases (ICD-9: 001-009 or ICD-10: A00-A09), other and unspecified noninfectious  
169 gastroenteritis (ICD-9: 558.9 or ICD-10: K52.9), vomiting (ICD-9: 787.0, 787.01, 787.03 or  
170 ICD-10: R11.1, R11.10-R11.12, R11.2), and diarrhea (ICD-9: 787.91 or ICD-10: R19.7). Daily  
171 AGI cases were aggregated to the municipal level based on residence for each study municipality  
172 over the study period. This study was designated as exempt from review by the Boston  
173 University Medical Campus Institutional Review Board (IRB number: H-42193).

#### 174 *Statistical Analysis*

175 A case time series design was used to assess associations between CSO/precipitation  
176 events and daily municipality-level cases of AGI. The case time series design is a self-matched  
177 method which allows for assessment of short-term risks of intermittent exposure events by  
178 comparing outcomes that occur at different times within individuals or small areas, such as  
179 municipalities.<sup>38,39</sup> This approach allows for investigation of potentially lagged relationships and  
180 provides control for time-invariant characteristics by design, similar to other self-matched  
181 designs.<sup>38,39</sup>

182 The case time series model (Model 1) is specified as:



183 
$$g[E(y_{it})] = \xi_{i(k)} + f(x_{it}, l) + ns(temp, f) + I(federal\ holiday) \quad (1)$$

184 where  $y_{it}$  is the number of AGI cases in municipality  $i$  at time  $t$ . The baseline risks for different  
185 risk sets (e.g., temporal strata;  $k=1, 2, \dots, k_i$ ) are represented by  $\xi_{i(k)}$ . The temporal strata used to  
186 define within-municipality comparison included year, month, and day of week. The function  
187  $f(x_{it}, l)$  specifies the association with the exposure in municipality  $i$  at time  $t$  along the lag  
188 dimension  $l$ . A uniform weighted lag structure was used that assumes that the probability of  
189 illness is constant across the lag period. The term  $ns(temp, f)$  describes a natural spline of  
190 average daily temperature with  $f$  degrees of freedom (three degrees of freedom was assumed).  
191 The temporal strata in the case time series model control for temporal trends across years,  
192 months, and days of the week, but  $I(federal\ holiday)$  is included as an indicator for public  
193 federal holidays because patterns of ED use may differ on holidays.<sup>40,41</sup> Results generated by the  
194 case time series model include cumulative exposure-response relationships reported as  
195 cumulative risk ratio (CRR) as well as exposure-lag-response associations reported as daily risk  
196 ratio (DRR) over the lag period. CRR represents an adjusted estimate comparing the number of  
197 cases that occur over the lag period after an extreme event to the number of cases over the same  
198 lag period after days that are temporally matched by day of the week, month, and year when no  
199 event occurred.

200 Model 1 was used to estimate the association between extreme CSO events and AGI, and  
201 separately to estimate the association between extreme precipitation events and AGI (2 distinct  
202 models). Model 1 was also applied to the same two exposure metrics (CSO and precipitation  
203 events) stratified by municipality drinking water source: 1) exclusive use of the Merrimack River  
204 as a drinking water source, or 2) exclusive use of other drinking water sources (4 additional  
205 model iterations).



229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251

## Results

### *Descriptive Statistics*

Over the course of the study period, CSS municipalities released a total of 3,012.6 million gallons (MG) of CSO discharge into the Merrimack River (Table S1). CSO events occurred in at least one CSS in the watershed on 447 days (21%) of the study period (January 1, 2014 – September 31, 2019). Both the frequency and size of CSO events varied among municipalities, with Manchester and Lowell contributing approximately 83% of total CSO discharge volume to the system. Cumulative daily upstream CSO discharge volume was highly variable during the study period, ranging from less than 0.1 MG to 189.0 MG with the 95<sup>th</sup> percentile value at 27.3 MG and the 99<sup>th</sup> percentile at 78.7 MG (Figure S2).

Precipitation events occurred in at least one of the municipalities in the study area on 972 days (46%) of the study period (Table S1). Precipitation events ranged from less than 0.1 inches of daily accumulation to 4.67 inches daily, with a 95<sup>th</sup> percentile value of 1.24 inches and a 99<sup>th</sup> percentile value of 1.97 inches. A positive correlation was observed between daily municipality-level precipitation and upstream cumulative CSO discharge on the day of the events ( $r=0.53$ ) (Figure S3).

Massachusetts EDs documented 100,206 cases of AGI among residents of the study municipalities over the study period (Table 1). AGI patients were majority female (59.9%), and approximately 30% of patients were in age groups most susceptible to AGI (adults over 65 years old and children under 5 years old).<sup>20</sup> The annual number of AGI ED visits increased by year over the study period, with 16,079 cases occurring in 2014 and 18,141 cases in 2018. Figure S4 shows the per capita rate of AGI among all study municipalities over the study period compared to the overall rate of AGI in the state of Massachusetts per the CHIA dataset.<sup>37</sup>

252 *Model Results*

253 All municipalities

254 In models adjusted for temporal trends and temperature but not adjusted for precipitation,  
255 the cumulative risk ratio (CRR) of AGI was 22% higher (CRR: 1.22 [95% CI: 1.05, 1.42]) in the  
256 four days following 95<sup>th</sup> and 63% higher (1.63 [1.07, 2.47]) in the four days following 99<sup>th</sup>  
257 percentile CSO events compared to days with no CSO events when considering all  
258 municipalities (Figure 2a). A marginal association that was not statistically significant was  
259 observed between 90<sup>th</sup> percentile CSO events and AGI (1.07 [0.96, 1.19]). No increase in risk  
260 was observed following CSO events below the 90<sup>th</sup> percentile. These associations remain largely  
261 unchanged among all municipalities after adjusting for precipitation, with a 17% increase in  
262 CRR of AGI in the 4 days after 95<sup>th</sup> percentile CSO events (1.17 [0.98, 1.39]) and a 62%  
263 increase after 99<sup>th</sup> percentile CSO events (1.62 [1.04, 2.51]) (Figure 2b). Extreme precipitation is  
264 also associated with a 13% increase in CRR of AGI in the 4 days following 95<sup>th</sup> percentile  
265 precipitation events (1.13 [1.00, 1.27]), and a marginal increase following 99<sup>th</sup> percentile  
266 precipitation events (1.07 [0.82, 1.41]) (Figure 2c). Results of all models, including those that  
267 assessed the association between extreme events and AGI among municipalities that use any  
268 water from the Merrimack and those that used a seven-day lag period, are included in Tables S2  
269 and S3.

270 Stratification by drinking water source

271 We did not observe a difference in the unadjusted CRR of AGI in municipalities with  
272 different drinking water sources following 95<sup>th</sup> percentile CSO events after stratification by  
273 drinking water source (Merrimack River: 1.19 [0.97, 1.46]; Other sources: 1.19 [0.93, 1.54]).  
274 After adjusting for precipitation, CRR of AGI was only marginally elevated following 95<sup>th</sup>

275 percentile CSO events among municipalities that exclusively get their drinking water from the  
276 river (1.05 [0.82, 1.33]), while the CRR of AGI among municipalities that do not use the  
277 Merrimack River as a drinking water source was slightly higher (1.27 [0.94, 1.70]) (Figure 2b).  
278 CRR of AGI was elevated regardless of drinking water source after 99<sup>th</sup> percentile CSO events  
279 compared to days with no CSO events (Merrimack River: 1.57 [0.86, 2.88]; Other sources: 1.84  
280 [0.88, 3.85]) (Figure 2b).

### 281 Daily relative risk

282 Comparison of daily risk ratios (DRR) of AGI in municipalities with different drinking  
283 water sources over the four-day lag period shows that the magnitude of the strongest observed  
284 increase in DRR was approximately the same between drinking water groups (Merrimack River:  
285 1.36 [1.05, 1.76]; Other sources: 1.45 [1.07, 1.96]), but occurred on different days within that  
286 period (Figure 3). The strongest association between extreme CSO events and AGI was observed  
287 two days after an event among municipalities that use the Merrimack as their exclusive drinking  
288 water source, and one day after an event among municipalities with other drinking water sources.  
289 A similar temporal pattern was observed in the four days following 95<sup>th</sup> percentile CSO events,  
290 though the magnitude of the associations were attenuated (Figure S5).

### 291 Sensitivity analyses

292 Results from the sensitivity analysis assessing the relationship between CSO events from  
293 individual CSS municipalities and cases of AGI in downstream municipalities are shown in  
294 Table S4. Elevated CRR of AGI is observed among downstream municipalities 4-days after CSO  
295 events in Manchester (1.09 [0.93, 1.29]) and Lowell (1.16 [0.95, 1.42]), though neither of these  
296 associations are statistically significant.

## 297 **Discussion**

298 In this analysis of the association between CSO events and ED visits for AGI, we  
299 observed higher risk of AGI in the 4 days following extreme CSO events compared to days with  
300 no CSO events. Results were robust to adjustment for potential confounding by precipitation.  
301 The magnitude of CSO discharge volume was an important factor in the association between  
302 CSO events and AGI, with the largest associations observed following  $\geq 95^{\text{th}}$  and  $\geq 99^{\text{th}}$  percentile  
303 CSO events. Stratification by drinking water source (Merrimack River or other source) suggests  
304 that the association between  $\geq 95^{\text{th}}$  percentile CSO events and AGI after adjusting for extreme  
305 precipitation events was slightly more pronounced in municipalities that do not rely on the  
306 Merrimack River as a drinking water source, though CRR of AGI was elevated after the largest  
307 CSO events ( $>99^{\text{th}}$  percentile) regardless of drinking water source. The temporal pattern of daily  
308 risk in the 4 days following  $99^{\text{th}}$  percentile CSO events differed between municipalities that do  
309 and do not use the Merrimack River as a drinking water source, which may suggest different  
310 exposure pathways.

311 The magnitude of the associations observed in this analysis are similar to or larger than  
312 those found in previous studies, though methodological differences and study populations make  
313 direct comparisons difficult. Using a similar self-matched study design, Brokamp et al. found a  
314 16% higher odds of AGI after any CSO event among children living within 500m of a CSO  
315 outfall (OR: 1.16 [1.04, 1.30]).<sup>33</sup> This study differed in notable ways from our study, including in  
316 the approach to exposure classification, study population, and the inclusion of CSO volume and  
317 precipitation, but the magnitude of the association was similar to the CRR we observed  
318 following  $95^{\text{th}}$  percentile CSO events. On the other hand, Miller et al. did not use a self-matched  
319 study design and found a smaller but more precise association between large CSO events and  
320 AGI after adjusting for precipitation among Atlanta residents of all ages (rate ratio: 1.09 [1.03,

321 1.14]).<sup>32</sup> The definition of a large CSO event in Miller et al. was  $\geq 56.6$  MG of discharge, which  
322 falls in the range between 95<sup>th</sup> (27.3 MG) and 99<sup>th</sup> percentile (78.7 MG) CSO discharge events in  
323 our study. In an analysis in the Merrimack Valley region of Massachusetts that used extreme  
324 precipitation as a proxy for CSO events, Jagai et al. found a 13% increase in cumulative risk of  
325 AGI following extreme precipitation events (equivalent to 1.97 in., CRR: 1.13 [1.00, 1.28]).<sup>31</sup>  
326 This result is remarkably similar to the association observed in this analysis between 95<sup>th</sup>  
327 percentile precipitation events (equivalent to 1.24 in.) and cumulative risk of AGI over four days  
328 observed in our study. Overall, we find that the strength of the associations observed in this study  
329 between 95<sup>th</sup> percentile CSO or precipitation events and AGI are fairly consistent with previous  
330 findings, but that the CRR of AGI following 99<sup>th</sup> percentile CSO events is notably higher than  
331 what has been reported in similar studies.

332 Our findings suggest that a nonlinear dose-response relationship exists between upstream  
333 CSO discharge volume and AGI, a finding consistent with previous studies.<sup>31,32</sup> Miller et al.  
334 similarly found that only large CSO events ( $\geq 75^{\text{th}}$  percentile CSO events by volume, equivalent  
335 to 56.5 MG) were significantly associated with AGI in the week following an event, but models  
336 estimating the association between any CSO event and AGI resulted in no association.<sup>32</sup> Jagai et  
337 al. found that only the most extreme precipitation events (99<sup>th</sup> percentile, equivalent to 1.97 in)  
338 were significantly associated with increased cumulative risk of AGI; lower magnitude, not  
339 statistically significant associations were found for lower percentile precipitation events.<sup>31</sup> These  
340 results are particularly relevant as the frequency and intensity of extreme precipitation events  
341 (and extreme CSO events as a result) are expected to increase due to climate change in the  
342 Northeast and Midwest where most CSS in the US are located.<sup>3</sup>

343           The results of this analysis suggest that there is an increase in risk of AGI following 99<sup>th</sup>  
344 percentile CSO events in municipalities that source their drinking water from the Merrimack  
345 River that could be related to contaminated drinking water, though these results should be  
346 interpreted with caution. The later peak in DRR of AGI following extreme CSO events in  
347 municipalities that get their drinking water from the Merrimack compared to those that do not  
348 could also suggest a longer lag due to additional time for source water intake, treatment, and  
349 distribution. However, extreme precipitation events were associated with AGI most prominently  
350 in municipalities that use the Merrimack River as their exclusive or intermittent drinking water  
351 source (Table S2C), which could instead suggest contamination of drinking water due to  
352 pathogen infiltration or cross contamination of leaky sewage and water distribution pipes.<sup>43,44</sup>  
353 The large confidence intervals and lack of statistical significance among any of the stratified  
354 drinking water groups in this study, possibly due to smaller case counts, further limits the  
355 conclusions that can be made about the potential for drinking water contamination.

356           The evidence in the literature is inconclusive on the potential for drinking water  
357 contamination after CSO events. While it is conceptually possible for treated drinking water to  
358 be contaminated by CSO discharge,<sup>45,46</sup> studies that incorporated drinking water source  
359 contamination following discharge events either did not include CSO event data in the  
360 exposure<sup>30,31</sup> or found that controlling for CSO events as a binary covariate did not alter the  
361 association between precipitation and AGI.<sup>42</sup> We are also unaware of an accounting of how  
362 many municipalities have a drinking water source that is impacted by CSO discharge, so it is  
363 unclear how widespread this issue might be, though the results of this analysis suggest that  
364 drinking water contamination may not be the only important exposure pathway in a CSO-  
365 impaired river system. Understanding the critical exposure pathway(s) following discharge



366 events is an important direction for future work that has direct implications for infrastructure  
367 investment decisions and policies designed to protect public health.

368         The findings of this study suggest that CSO events and AGI were more strongly  
369 associated across the four days following an event compared to a longer lag period of seven  
370 days. These findings are consistent with previous studies that found significant associations  
371 between sewage discharge events and AGI over lag periods from 2-8 days.<sup>30-33</sup> A four-day lag  
372 period is also consistent with the incubation times of some of the pathogens found in waterways  
373 contaminated with sewage.<sup>1,9,47</sup> Some AGI-causing viral and bacterial pathogens have incubation  
374 periods that are consistent with the short lag identified in this study including noroviruses (12-48  
375 hours), rotavirus (1-3 days), *Shigella* (12 hours-7 days), and enterotoxigenic *E. coli* (12-72  
376 hours).<sup>47</sup> Viral waterborne pathogens can be infectious at low doses, persistent in environmental  
377 media, and less effectively removed by conventional drinking water treatment compared to  
378 bacteria.<sup>48,49</sup> Another consideration with a short lag time is time between CSO event and  
379 exposure, which could differ based on exposure pathway. It is possible that people exposed  
380 primarily through recreational activity or aerosolized pathogens could ingest contaminated water  
381 within hours after a CSO event, though our analysis was not spatially refined enough to assess  
382 the possibility of exposure through the latter pathway.

383         This study has a number of potential limitations. First, the analysis is limited by a lack of  
384 individual-level information on exposure, including rates of consumption or quality of drinking  
385 water, amount or timing of contact with recreational waters, and/or recreational surface water  
386 quality experienced by individuals within the study population. The lack of individual exposure  
387 data limits our ability to gain insights about specific exposure pathways. Moreover, lack of  
388 information about local water contamination and potential local responses to CSO events (e.g.,

389 adjustments to drinking water treatment processes), may have led to substantial exposure  
390 misclassification and likely biased results towards the null hypothesis of no association. Second,  
391 the AGI case data were deidentified, precluding identification of patients that may have been  
392 treated for AGI more than once. Our inability to account for the correlation among repeated  
393 episodes within an individual may have resulted in confidence intervals that are perhaps too  
394 narrow. Third, people seeking care for AGI in the ED likely represent the most severe cases of  
395 gastrointestinal illness, given that 10% or fewer of all cases of AGI come to medical attention in  
396 acute care settings.<sup>50</sup> Thus, it is unclear whether these results are generalizable to milder episodes  
397 of AGI or to other communities across the country.

398         On the other hand, this study has some notable strengths. First, the case time series design  
399 minimizes potential confounding by differences between municipalities and effectively controls  
400 for temporal trends. Second, by using daily measures of CSO discharge volume instead of  
401 approaches taken in other studies to measure CSO exposure, such as precipitation as a proxy for  
402 CSO events,<sup>31</sup> modeling CSO events as binary exposures,<sup>33</sup> or aggregating to a weekly time  
403 step,<sup>32</sup> this study more directly evaluates the relationship between CSO discharge and health. Our  
404 use of daily CSO and precipitation data allowed for a refined temporal analysis in the days  
405 following an exposure event, and CSO volume data provided the opportunity to assess potential  
406 dose-response relationships between CSO discharge and AGI. We also considered the spatial  
407 relationship between CSS municipalities and downstream municipalities in our approach to  
408 assigning exposure. While there are limitations to characterization of exposure in this study by  
409 considering upstream discharge, this approach is representative of how CSO discharge moves  
410 through the area and accumulates in downstream stretches of the river. The PRISM dataset  
411 provided spatial resolution of precipitation estimates that may account for heterogeneity in

412 rainfall that is not captured by a regional weather station, the latter of which has been the source  
413 for precipitation data in other studies.<sup>31,32</sup> Finally, reliance on the Merrimack River as a drinking  
414 water source downstream of CSO outfalls provided an opportunity to compare associations  
415 between municipalities that do and do not get their drinking water from this CSO-impacted  
416 source. To our knowledge, only three studies report the association between sewage-related  
417 discharge events and health outcomes in regions with modern wastewater infrastructure where  
418 drinking water intakes exist downstream of outfalls.<sup>30,31,42</sup> In those studies, the exposure was  
419 either extreme rainfall as a proxy for CSO events,<sup>31</sup> undertreated sewage discharge events,<sup>30</sup> or  
420 precipitation with CSO discharge included as a covariate.<sup>42</sup>

421 Our findings suggest that extreme CSO events increase the risk of ED visits for AGI in  
422 downstream municipalities in the four days following an event, and that a non-linear dose-  
423 response relationship exists between CSO discharge and ED visits for AGI. The largest CSO  
424 events are associated with AGI across drinking water sources and concurrent precipitation levels,  
425 although we observed differences in the timing and magnitude of associations based on drinking  
426 water source. Although our study is ecological in design, our findings indicate that CSS may  
427 negatively impact public health. As precipitation events intensify due to climate change, large-  
428 volume CSO events may become more common and the risk of AGI in downstream  
429 municipalities may increase.

430

431 **Acknowledgements:** The authors acknowledge staff at US EPA Region 1 and the City of  
432 Manchester for their help locating CSO discharge data. We are also grateful to Kevin Brander of  
433 the Massachusetts Department of Environmental Protection for sharing his expertise and  
434 knowledge of the study system and to Aine Studdert-Kennedy for her assistance with data entry.

435

436 **Funding:** This project was supported by an Early Stage Urban Research Award from the Boston  
437 University Initiative on Cities. BMH was partially supported by National Institute of

438 Environmental Health Sciences (NIEHS) grant T32ES014562 and a National Science  
439 Foundation Research Traineeship (NRT) grant to Boston University (DGE 1735087).  
440

441

## References

- 442 1. US EPA. *Report to Congress on Impacts and Control of Combined Sewer Overflows and*  
443 *Sanitary Sewer Overflows.*; 2004. Accessed May 10, 2023.  
444 [https://www.epa.gov/sites/default/files/2015-10/documents/csosortc2004\\_full.pdf](https://www.epa.gov/sites/default/files/2015-10/documents/csosortc2004_full.pdf)
- 445 2. Madoux-Humery AS, Dorner S, Sauvé S, et al. The effects of combined sewer overflow  
446 events on riverine sources of drinking water. *Water Research.* 2016;92:218-227.  
447 doi:10.1016/j.watres.2015.12.033
- 448 3. USGCRP. *Fourth National Climate Assessment.* U.S. Global Change Research Program,  
449 Washington, DC; 2018:1-470. Accessed May 12, 2023.  
450 <https://nca2018.globalchange.gov/chapter/3>
- 451 4. US EPA. Combined Sewer Overflow Frequent Questions. US EPA. Published October 14,  
452 2015. Accessed December 1, 2020. [https://www.epa.gov/npdcs/combined-sewer-overflow-](https://www.epa.gov/npdcs/combined-sewer-overflow-frequent-questions)  
453 [frequent-questions](https://www.epa.gov/npdcs/combined-sewer-overflow-frequent-questions)
- 454 5. Harp RD, Horton DE. Observed Changes in Daily Precipitation Intensity in the United  
455 States. *Geophysical Research Letters.* 2022;49(19):e2022GL099955.  
456 doi:10.1029/2022GL099955
- 457 6. Rechenburg A, Koch Ch, Claßen Th, Kistemann Th. Impact of sewage treatment plants and  
458 combined sewer overflow basins on the microbiological quality of surface water. *Water*  
459 *Science and Technology.* 2006;54(3):95-99. doi:10.2166/wst.2006.454
- 460 7. Passerat J, Ouattara NK, Mouchel JM, Vincent Rocher, Servais P. Impact of an intense  
461 combined sewer overflow event on the microbiological water quality of the Seine River.  
462 *Water Research.* 2011;45(2):893-903. doi:10.1016/j.watres.2010.09.024
- 463 8. Madoux-Humery AS, Dorner S, Sauvé S, et al. Temporal variability of combined sewer  
464 overflow contaminants: Evaluation of wastewater micropollutants as tracers of fecal  
465 contamination. *Water Research.* 2013;47(13):4370-4382. doi:10.1016/j.watres.2013.04.030
- 466 9. Donovan E, Unice K, Roberts JD, Harris M, Finley B. Risk of gastrointestinal disease  
467 associated with exposure to pathogens in the water of the Lower Passaic River. *Appl Environ*  
468 *Microbiol.* 2008;74(4):994-1003. doi:10.1128/AEM.00601-07
- 469 10. Hata A, Katayama H, Kojima K, et al. Effects of rainfall events on the occurrence and  
470 detection efficiency of viruses in river water impacted by combined sewer overflows.  
471 *Science of The Total Environment.* 2014;468-469:757-763.  
472 doi:10.1016/j.scitotenv.2013.08.093
- 473 11. Arnone RD, Walling JP. Evaluating Cryptosporidium and Giardia concentrations in  
474 combined sewer overflow. *Journal of Water and Health.* Published online 2006:157-165.  
475 doi:10.2166/wh.2006.064

- 476 12. Nickel JP, Fuchs S. Micropollutant emissions from combined sewer overflows. *Water*  
477 *Science and Technology*. 2019;80(11):2179-2190. doi:10.2166/wst.2020.035
- 478 13. Di Nunno F, Granata F, Parrino F, Gargano R, de Marinis G. Microplastics in Combined  
479 Sewer Overflows: An Experimental Study. *Journal of Marine Science and Engineering*.  
480 2021;9(12):1415. doi:10.3390/jmse9121415
- 481 14. Treilles R, Gasperi J, Saad M, et al. Macro and Microplastics in Stormwater and Combined  
482 Sewer Overflows in Paris Megacity. In: Cocca M, Di Pace E, Errico ME, et al., eds.  
483 *Proceedings of the 2nd International Conference on Microplastic Pollution in the*  
484 *Mediterranean Sea*. Springer Water. Springer International Publishing; 2020:145-151.  
485 doi:10.1007/978-3-030-45909-3\_23
- 486 15. Olds HT, Corsi SR, Dila DK, Halmo KM, Bootsma MJ, McLellan SL. High levels of sewage  
487 contamination released from urban areas after storm events: A quantitative survey with  
488 sewage specific bacterial indicators. *PLoS Med*. 2018;15(7).  
489 doi:10.1371/journal.pmed.1002614
- 490 16. Fewtrell L, Kay D. Recreational Water and Infection: A Review of Recent Findings. *Curr*  
491 *Environ Health Rep*. 2015;2(1):85-94. doi:10.1007/s40572-014-0036-6
- 492 17. Arnold BF, Wade TJ, Benjamin-Chung J, et al. Acute Gastroenteritis and Recreational  
493 Water: Highest Burden Among Young US Children. *Am J Public Health*. 2016;106(9):1690-  
494 1697. doi:10.2105/AJPH.2016.303279
- 495 18. DeFlorio-Barker S, Wing C, Jones RM, Dorevitch S. Estimate of incidence and cost of  
496 recreational waterborne illness on United States surface waters. *Environmental Health*.  
497 2018;17(1):3. doi:10.1186/s12940-017-0347-9
- 498 19. Russo GS, Eftim SE, Goldstone AE, Dufour AP, Nappier SP, Wade TJ. Evaluating health  
499 risks associated with exposure to ambient surface waters during recreational activities: A  
500 systematic review and meta-analysis. *Water Research*. 2020;176:115729.  
501 doi:10.1016/j.watres.2020.115729
- 502 20. Reynolds KA, Mena KD, Gerba CP. Risk of waterborne illness via drinking water in the  
503 United States. *Rev Environ Contam Toxicol*. 2008;192:117-158. doi:10.1007/978-0-387-  
504 71724-1\_4
- 505 21. Messner M, Shaw S, Regli S, Rotert K, Blank V, Soller J. An approach for developing a  
506 national estimate of waterborne disease due to drinking water and a national estimate model  
507 application. *J Water Health*. 2006;4 Suppl 2:201-240. doi:10.2166/wh.2006.024
- 508 22. Collier SA, Deng L, Adam EA, et al. Estimate of Burden and Direct Healthcare Cost of  
509 Infectious Waterborne Disease in the United States. *Emerg Infect Dis*. 2021;27(1):140-149.  
510 doi:10.3201/eid2701.190676

- 511 23. Craun GF, Brunkard JM, Yoder JS, et al. Causes of Outbreaks Associated with Drinking  
512 Water in the United States from 1971 to 2006. *Clin Microbiol Rev.* 2010;23(3):507-528.  
513 doi:10.1128/CMR.00077-09
- 514 24. Moreira NA, Bondelind M. Safe drinking water and waterborne outbreaks. *J Water Health.*  
515 2017;15(1):83-96. doi:10.2166/wh.2016.103
- 516 25. Schuster CJ, Aramini JJ, Ellis AG, et al. Infectious Disease Outbreaks Related to Drinking  
517 Water in Canada, 1974–2001. *Can J Public Health.* 2005;96(4):254-258.  
518 doi:10.1007/BF03405157
- 519 26. Curriero FC, Patz JA, Rose JB, Lele S. The association between extreme precipitation and  
520 waterborne disease outbreaks in the United States, 1948-1994. *Am J Public Health.*  
521 2001;91(8):1194-1199. doi:10.2105/ajph.91.8.1194
- 522 27. Guzman Herrador BR, de Blasio BF, MacDonald E, et al. Analytical studies assessing the  
523 association between extreme precipitation or temperature and drinking water-related  
524 waterborne infections: a review. *Environ Health.* 2015;14:29. doi:10.1186/s12940-015-0014-  
525 y
- 526 28. Exum N, Betanzo E, Schwab K, Chen T, Guikema S, Harvey D. Extreme Precipitation,  
527 Public Health Emergencies, and Safe Drinking Water in the USA. *Current Environmental*  
528 *Health Reports.* 2018;5. doi:10.1007/s40572-018-0200-5
- 529 29. Harder-Lauridsen NM, Kuhn KG, Erichsen AC, Mølbak K, Ethelberg S. Gastrointestinal  
530 Illness among Triathletes Swimming in Non-Polluted versus Polluted Seawater Affected by  
531 Heavy Rainfall, Denmark, 2010-2011. *PLOS ONE.* 2013;8(11):e78371.  
532 doi:10.1371/journal.pone.0078371
- 533 30. Redman RL, Nenn CA, Eastwood D, Gorelick MH. Pediatric emergency department visits  
534 for diarrheal illness increased after release of undertreated sewage. *Pediatrics.*  
535 2007;120(6):e1472-1475. doi:10.1542/peds.2007-0283
- 536 31. Jagai JS, Li Q, Wang S, Messier KP, Wade TJ, Hilborn ED. Extreme Precipitation and  
537 Emergency Room Visits for Gastrointestinal Illness in Areas with and without Combined  
538 Sewer Systems: An Analysis of Massachusetts Data, 2003-2007. *Environ Health Perspect.*  
539 2015;123(9):873-879. doi:10.1289/ehp.1408971
- 540 32. Miller AG, Ebel S, Levy K. Combined Sewer Overflows and Gastrointestinal Illness in  
541 Atlanta, 2002–2013: Evaluating the Impact of Infrastructure Improvements. *Environmental*  
542 *Health Perspectives.* 2022;130(5):057009. doi:10.1289/EHP10399
- 543 33. Brokamp C, Beck AF, Muglia L, Ryan P. Combined sewer overflow events and childhood  
544 emergency department visits: A case-crossover study. *Sci Total Environ.* 2017;607-  
545 608:1180-1187. doi:10.1016/j.scitotenv.2017.07.104
- 546 34. US EPA. About the Merrimack. Published August 15, 2016. Accessed May 17, 2023.  
547 <https://www.epa.gov/merrimackriver/about-merrimack>

- 548 35. US EPA. Merrimack River Watershed. Published 2019. Accessed December 1, 2020.  
549 <https://epa.maps.arcgis.com/apps/MapSeries/index.html?appid=922e1c016c6e42b199f902d1>  
550 cfb84bbd
- 551 36. PRISM Climate Group. PRISM Climate Data. Accessed March 18, 2021.  
552 <https://prism.oregonstate.edu/>
- 553 37. CHIA. Case Mix Data. Published 2021. Accessed March 20, 2021.  
554 <https://www.chiamass.gov/case-mix-data/>
- 555 38. Gasparrini A. The Case Time Series Design. *Epidemiology*. 2021;32(6):829-837.  
556 doi:10.1097/EDE.0000000000001410
- 557 39. Gasparrini A. A tutorial on the case time series design for small-area analysis. *BMC Med Res*  
558 *Methodol*. 2022;22(1):129. doi:10.1186/s12874-022-01612-x
- 559 40. Zheng W, Muscatello DJ, Chan AC. Deck the halls with rows of trolleys ... emergency  
560 departments are busiest over the Christmas holiday period. *Medical Journal of Australia*.  
561 2007;187(11-12):630-633. doi:10.5694/j.1326-5377.2007.tb01451.x
- 562 41. Salazar A, Corbella X, Sánchez JL, Argimón JM, Escarrabill J. How to manage the ED crisis  
563 when hospital and/or ED capacity is reaching its limits. Report about the implementation of  
564 particular interventions during the Christmas crisis. *European Journal of Emergency*  
565 *Medicine*. 2002;9(1):79.
- 566 42. Drayna P, McLellan SL, Simpson P, Li SH, Gorelick MH. Association between rainfall and  
567 pediatric emergency department visits for acute gastrointestinal illness. *Environ Health*  
568 *Perspect*. 2010;118(10):1439-1443. doi:10.1289/ehp.0901671
- 569 43. Ercumen A, Gruber JS, Colford JM. Water distribution system deficiencies and  
570 gastrointestinal illness: a systematic review and meta-analysis. *Environ Health Perspect*.  
571 2014;122(7):651-660. doi:10.1289/ehp.1306912
- 572 44. Renwick DV, Heinrich A, Weisman R, Arvanaghi H, Rotert K. Potential Public Health  
573 Impacts of Deteriorating Distribution System Infrastructure. *J Am Water Works Assoc*.  
574 2019;111(2):42-53. doi:10.1002/awwa.1235
- 575 45. Sokolova E, Petterson SR, Dienus O, Nyström F, Lindgren PE, Pettersson TJR. Microbial  
576 risk assessment of drinking water based on hydrodynamic modelling of pathogen  
577 concentrations in source water. *Science of The Total Environment*. 2015;526:177-186.  
578 doi:10.1016/j.scitotenv.2015.04.040
- 579 46. Taghipour M, Shakibaeinia A, Sylvestre É, Tolouei S, Dorner S. Microbial risk associated  
580 with CSOs upstream of drinking water sources in a transboundary river using hydrodynamic  
581 and water quality modeling. *Science of The Total Environment*. 2019;683:547-558.  
582 doi:10.1016/j.scitotenv.2019.05.130



- 583 47. CDC. *Waterborne Disease Outbreak Investigation Toolkit.*; 2023.  
584 <https://www.cdc.gov/healthywater/emergency/pdf/waterborne-disease-outbreak-toolkit-h.pdf>
- 585 48. Gibson KE. Viral pathogens in water: occurrence, public health impact, and available control  
586 strategies. *Current Opinion in Virology*. 2014;4:50-57. doi:10.1016/j.coviro.2013.12.005
- 587 49. Betancourt WQ, Rose JB. Drinking water treatment processes for removal of  
588 *Cryptosporidium* and *Giardia*. *Vet Parasitol*. 2004;126(1-2):219-234.  
589 doi:10.1016/j.vetpar.2004.09.002
- 590 50. Schmidt MA, Groom HC, Rawlings AM, et al. Incidence, Etiology, and Healthcare  
591 Utilization for Acute Gastroenteritis in the Community, United States. *Emerging Infectious*  
592 *Diseases*. 2022;28(11). doi:10.3201/eid2811.220247
- 593

594

## Tables

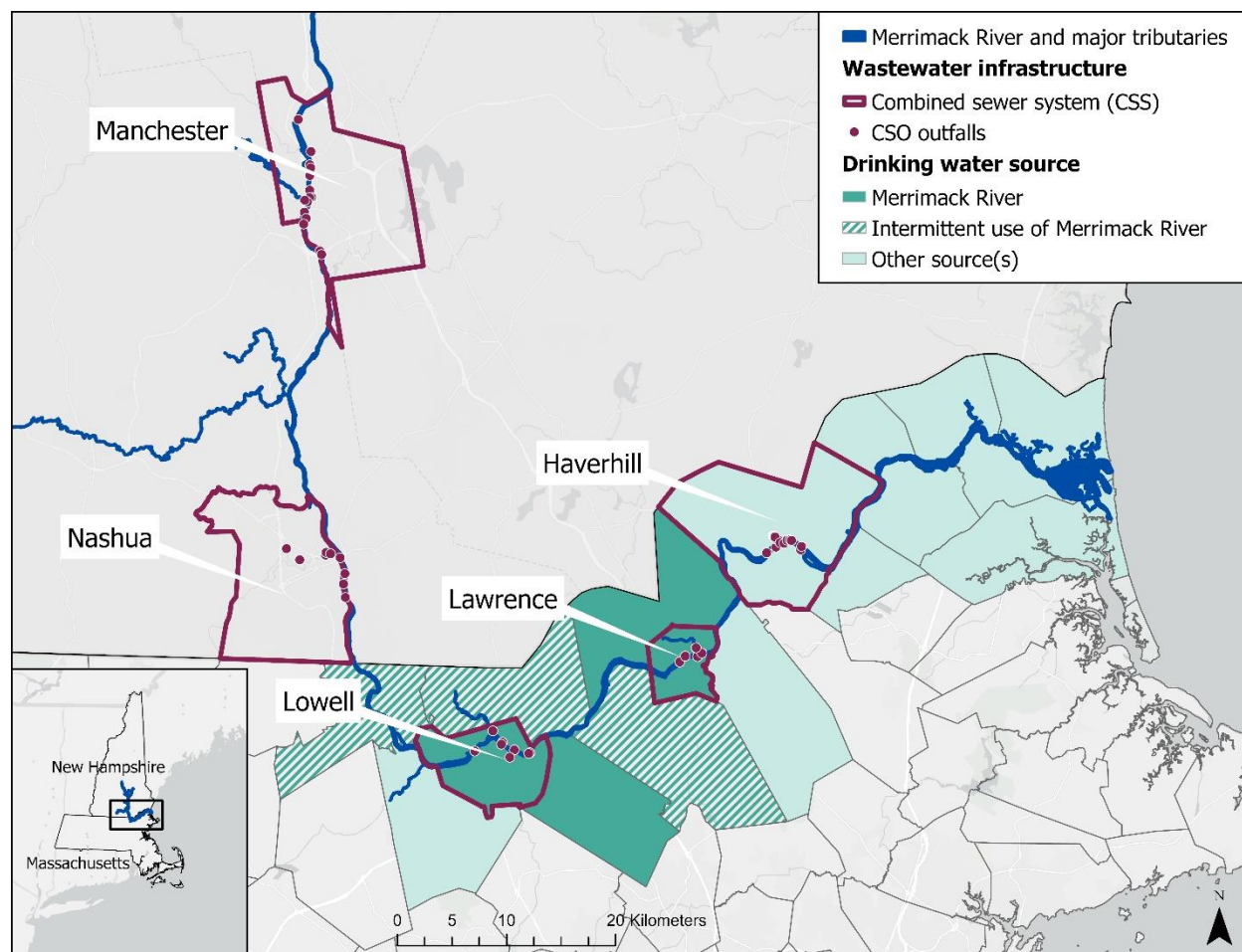
595 **Table 1: Descriptive statistics for cases of acute gastrointestinal illness in 17 Massachusetts**  
 596 **municipalities bordering the Merrimack River, January 1, 2014 – September 30, 2019.**

Characteristic	n (%)
Total number of AGI cases	100,206 (100%)
Cases of AGI by year	
2014	16,079 (16.0%)
2015	16,881 (16.8%)
2016	16,903 (16.9%)
2017	17,727 (17.7%)
2018	18,141 (18.1%)
2019 (though September 30)	14,475 (14.4%)
Average number of cases of AGI per year	17,427
Cases of AGI by sex	
Female	59,989 (59.9%)
Cases of AGI by age group	
< 5 years old	14,845 (14.8%)
5-19 years old	15,094 (15.1%)
20-64 years old	53,999 (53.9%)
65+ years old	16,268 (16.2%)
Cases of AGI by municipal drinking water source	
Merrimack River exclusively (4 municipalities)	64,278 (64.2%)
Merrimack River intermittently (3 municipalities)	7,649 (7.6%)
Other sources (10 municipalities)	28,279 (28.2%)

597

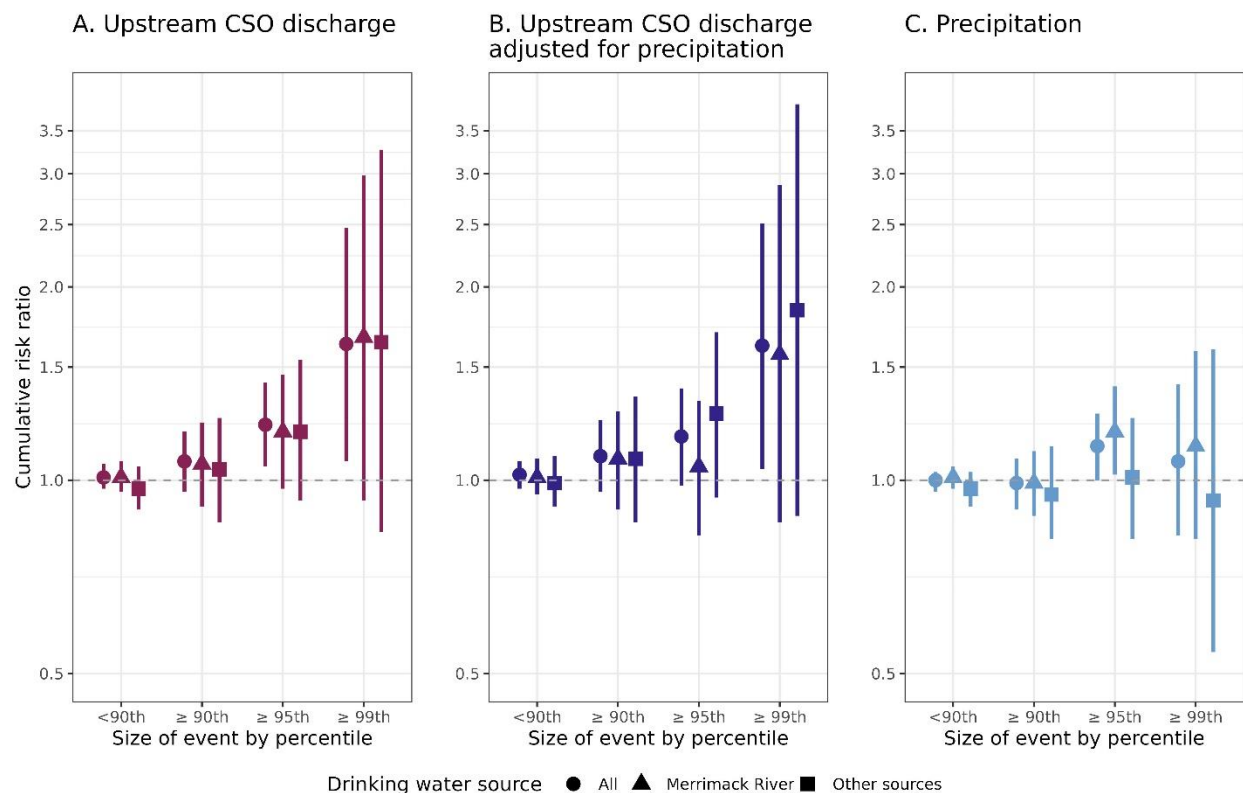
598  
599

## Figures

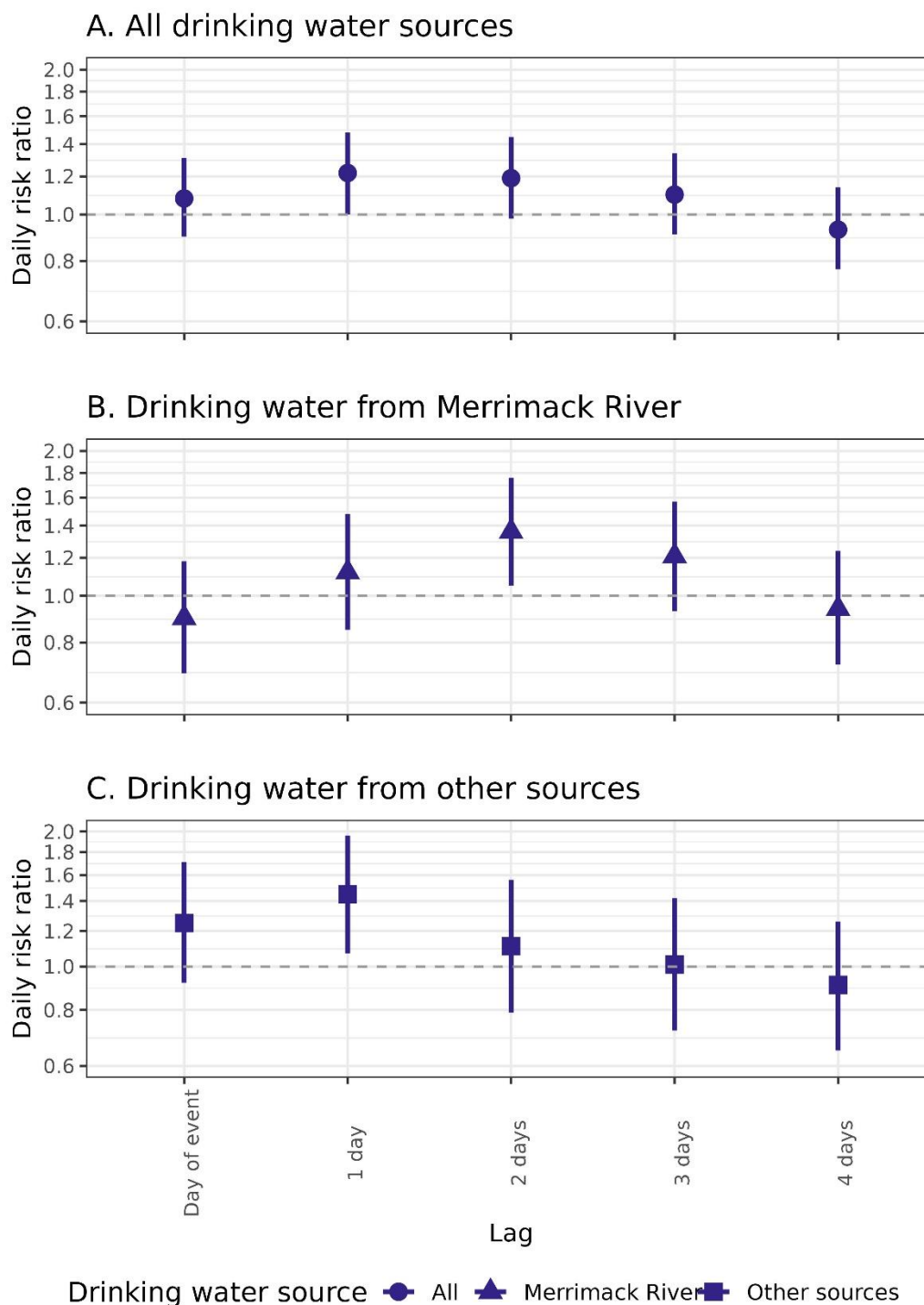


600  
601  
602  
603  
604

**Figure 1: Map of the study area in the lower reaches of the Merrimack River.**



605  
606 **Figure 2: Cumulative risk ratio (CRR) of acute gastrointestinal illness (AGI) over the four**  
607 **days following A) combined sewer overflow (CSO) events, B) CSO events adjusted for**  
608 **precipitation, and C) precipitation events.** Model results are shown for all municipalities and  
609 by drinking water source (Merrimack River or other sources), as well as events of different sizes:  
610 <90<sup>th</sup> percentile, ≥90<sup>th</sup> percentile, ≥95<sup>th</sup> percentile, and ≥99<sup>th</sup> percentile events. Municipalities  
611 that intermittently use the Merrimack River as a drinking water source were included in the  
612 analysis for all municipalities but not in the drinking water stratification analysis.  
613  
614  
615



616  
 617 **Figure 3: Daily risk ratio (DRR) of acute gastrointestinal illness (AGI) over a 4-day lag**  
 618 **period following of 99<sup>th</sup> percentile combined sewer overflow (CSO) events and adjusted for**  
 619 **precipitation in A) municipalities with all drinking water sources, B) municipalities that**  
 620 **exclusively get their drinking water from the Merrimack River, and C) municipalities that**  
 621 **exclusively get their drinking water from sources other than the Merrimack. Municipalities that**  
 622 **intermittently use the Merrimack River as a drinking water source were included in the analysis**  
 623 **for all municipalities but not in the drinking water stratification analysis.**