Refined assessment of associations between drinking water residence time and emergency department visits for gastrointestinal illness in Metro Atlanta, Georgia

Karen Levy, Mitchel Klein, Stefanie Ebelt Sarnat, Samina Panwhar, Alexandra Huttinger, Paige Tolbert and Christine Moe

ABSTRACT

Recent outbreak investigations suggest that a substantial proportion of waterborne disease outbreaks are attributable to water distribution system issues. In this analysis, we examine the relationship between modeled water residence time (WRT), a proxy for probability of microorganism intrusion into the distribution system, and emergency department visits for gastrointestinal (GI) illness for two water utilities in Metro Atlanta, USA during 1993–2004. We also examine the association between proximity to the nearest distribution system node, based on patients' residential address, and GI illness using logistic regression models. Comparing long ($\geq$90th percentile) with intermediate WRTs (11th to 89th percentile), we observed a modestly increased risk for GI illness for Utility 1 (OR = 1.07, 95% CI: 1.02–1.13), which had substantially higher average WRT than Utility 2, for which we found no increased risk (OR = 0.98, 95% CI: 0.94–1.02). Examining finer, 12-hour increments of WRT, we found that exposures >48 h were associated with increased risk of GI illness, and exposures of >96 h had the strongest associations, although none of these associations was statistically significant. Our results suggest that utilities might consider reducing WRTs to <2–3 days or adding booster disinfection in areas with longer WRT, to minimize risk of GI illness from water consumption.

Key words | diarrhea, emergency department, gastrointestinal illness, water distribution system, water quality

Karen Levy (corresponding author)
Mitchel Klein
Stefanie Ebelt Sarnat
Paige Tolbert
Department of Environmental Health,
Rollins School of Public Health, Emory University,
1518 Clifton Rd NE,
Atlanta, GA 30322, USA
E-mail: *karen.levy@emory.edu*

Karen Levy
Samina Panwhar
Alexandra Huttinger
Christine Moe
Center for Global Safe Water,
Sanitation and Hygiene at Emory University,
1518 Clifton Rd NE,
Atlanta, GA 30322, USA

Samina Panwhar
Alexandra Huttinger
Christine Moe
Hubert Department of Global Health,
Rollins School of Public Health, Emory University,
1518 Clifton Rd NE,
Atlanta, GA 30322, USA

Samina Panwhar
Present address: Oregon Health Authority,
500 Summer St. NE, E52,
Salem, OR 97301, USA

INTRODUCTION

The aging water delivery infrastructure is becoming a growing issue in the United States, with deferred maintenance creating a system that is 'deeply stressed,' 'over-worked and under-budgeted', according to the US Environmental Protection Agency (USEPA 2011a). There are an estimated 240,000 water main breaks per year in the USA, and the number of breaks increases substantially near the end of the system's service life (USEPA 2011b). Pipe breaks and/or low and negative pressure events in the drinking water distribution system can result in intrusion of pathogenic microorganisms if an external source of contamination is present (Besner *et al.* 2011).

Several outbreaks of waterborne disease have been attributed to public drinking water systems, including the largest reported waterborne disease outbreak ever documented, in Milwaukee, Wisconsin, in 1993, with over 400,000 people affected (MacKenzie *et al.* 1995). Between 1971 and 2006, 282 outbreaks of acute gastrointestinal (GI) illness in community water systems were reported to the US Waterborne Disease and Outbreak Surveillance System (WBDOSS).

Contamination within the distribution system and in-premise plumbing accounted for 79 (9.9%) and 65 (8.1%), respectively, of the non-legionellosis waterborne disease outbreaks where a deficiency in the water system could be identified. While there has been a decrease in the number of reported outbreaks over time, in this period there was no change in the annual proportion of outbreaks associated with distribution system deficiencies (Craun *et al.* 2010). Between 2009 and 2010, distribution system deficiencies accounted for 15.2% (5/33) of identified outbreak deficiencies in drinking water-associated disease outbreaks (CDC 2013).

However, outbreaks of waterborne disease represent only the disease events that impact a large enough number of people within a relatively short period of time to be recognized by public health agencies. Linking illness to drinking water is inherently difficult through outbreak investigation methods, because most persons have daily exposure to tap water (Tostmann *et al.* 2012; CDC 2013). Epidemiological studies can be used to evaluate the endemic burden of waterborne disease in the population, and to examine risks associated with distribution system deficiencies and intermittent contamination. A number of observational studies have found associations between GI illness and drinking tap water compromised by low pressure events (Hunter *et al.* 2005; UK), main breaks or maintenance work (Nygard *et al.* 2007; Norway), declines of residual chlorine concentrations (Egorov *et al.* 2002; Russia), increased water age (Tinker *et al.* 2009; USA), increased water turbidity (Egorov *et al.* 2003; Russia; Tinker *et al.* 2010; USA), water system outages (CDC 2011; USA), and virus detection in non-disinfected groundwater (Borchardt *et al.* 2012; USA). On the other hand, Malm *et al.* (2013; Sweden) found no evidence of increased complaints of GI illness associated with disturbances at the water works or in the distribution network compared with control periods without disturbances in the same geographical area, in a study using Swedish national Health Call Center data. A recent review of the impact of distribution system deficiencies on endemic GI illness found that tap water consumption in malfunctioning distribution networks and system deficiencies, such as water outages, were both associated with GI disease (Ercumen *et al.* 2014). Risk assessment models also suggest that distribution systems can be a source of GI illness (Teunis *et al.* 2010; USA; Lambertini *et al.* 2012; USA).

Intervention studies comparing homes with water treatment devices versus those without have had mixed results, with some reporting increased risk of GI illness attributable to drinking tap water (Payment *et al.* 1991, 1997; Canada), some reporting increased risk only in sensitive subpopulations (Colford *et al.* 2009; USA), and others reporting no difference (Hellard *et al.* 2001; Australia; Colford *et al.* 2005; USA). In their meta-analysis, Ercumen *et al.* (2014) found elevated risk of GI illness for consumers of tap water versus point-of-use treated water in unblinded studies, but no association for studies that blinded participants to their point-of-use water treatment status.

Epidemiological studies of distribution system contributions to GI illness are challenging to design and implement because of the difficulty in estimating exposure to waterborne pathogens, detecting the relevant health outcomes, and controlling for the effect of confounding factors. Our group has taken advantage of an extensive dataset of emergency department (ED) visits and detailed information on water distribution systems in Metro Atlanta, GA to address this challenge. We previously conducted an analysis of the associations between ED visits for acute GI illness and water residence time (WRT) in the distribution system, as estimated using hydraulic models developed by two water utilities. We consider WRT as a proxy for microorganism intrusion into the distribution system, because when it takes longer for water to reach the consumer there is a higher probability of an intrusion event occurring. As hypothesized, we observed a modest increased risk for GI illness, approximately 5–7%, among people living in ZIP codes served by water with a long average WRT (top 10%) compared with intermediate residence times (11th to 89th percentile), after controlling for potential confounding factors such as patient age and markers of socioeconomic status (Tinker *et al.* 2009).

In this previous analysis, we based the WRT exposure assignment on a ZIP code average level. However, important variation in WRT likely exists within ZIP codes, as the lengths of water pipes between the fixed-location water treatment plants and homes vary, depending on the size of the ZIP code and the location of homes. Because we are ultimately interested in the water that reaches the end user, we hypothesized that a more spatially refined characterization of WRT reduces exposure misclassification and thus may reduce any bias to the null of our previous epidemiologic findings. Therefore, in the analysis presented here, we refined our exposure metric by estimating patient-level

WRT based on patients' residential address and proximity to the nearest distribution system node in a calibrated hydraulic model of the distribution system.

METHODS

Study site

The greater Atlanta metropolitan area is served by six major water utilities. The water treatment technology, age of the infrastructure, size of the service area, demographics of the population served, and various other factors differ among these utilities. In this way, Atlanta provides a good example of the range of conditions that exist in drinking water distribution systems in major US cities. This study focuses on two utilities in the greater Atlanta area. Utility 1 and Utility 2 each operate two treatment plants that supply water to their respective distribution pipeline networks. Utility 1 serves 680,000 customers over a 650 square mile (1,680 km^2) area, and Utility 2 serves 1.2 million customers over a 348 square mile (900 km^2) area. The demographics of the populations served differ substantially: Utility 1 serves a 65% Caucasian population with an average Census 2000 block-level median household income of $60,569 (s.d. $20,017), whereas Utility 2 serves an 85% African American population with an average Census 2000 block-level median household income of $28,894 (s.d. $16,674).

Emergency department data

Our ED database comprised over 10 million ED visit records from hospitals in the 20-county Atlanta metropolitan area during 1993–2004, in which 41 of 42 acute care hospitals provided data for all or part of the study period. Relevant data elements for the current analysis included: patient medical record number, unique visit number, date of admission, primary and secondary International Classification of Diseases, Ninth Revision (ICD-9) diagnostic codes, patient age, date of birth, gender, race, ZIP code of residence, residential street address, and method of payment for the visit (e.g. Medicaid).

The current analysis was based on a subset of our overall database, and a subset of those records analyzed by Tinker *et al.* (2009). In this analysis, records were selected for inclusion for patients who visited any of the participating hospital EDs during the study period, had full residential address data available, and resided in the service area of the selected water companies at the time of the ED visit. Our initial selection resulted in 1,772,787 records from 15 hospitals for which full residential address information was available, and which were made by patients with residential ZIP codes in either the Utility 1 (884,643 visits) or Utility 2 (888,144 visits) service areas, as defined by Tinker *et al.* (2009). This dataset differs from that used in the Tinker *et al.* (2009) analysis, which included 2,092,735 records from 27 hospitals, because that analysis was not restricted by a requirement for full residential address information. A flow chart of the data processing steps we followed is shown in Figure 1, and a comparison of the data used in the present analysis to the Tinker *et al.* (2009) analysis is presented in Table 1.

Our GI illness outcome encompassed ED visits for which the primary, or any available secondary ICD-9, code had one of the following diagnoses: infectious GI illness (001–004, 005.0, 005.4, 005.89, 005.89, 005.9, 006–007, 008.0, 008.42–008.44, 008.47, 008.49, 008.5, 008.6,
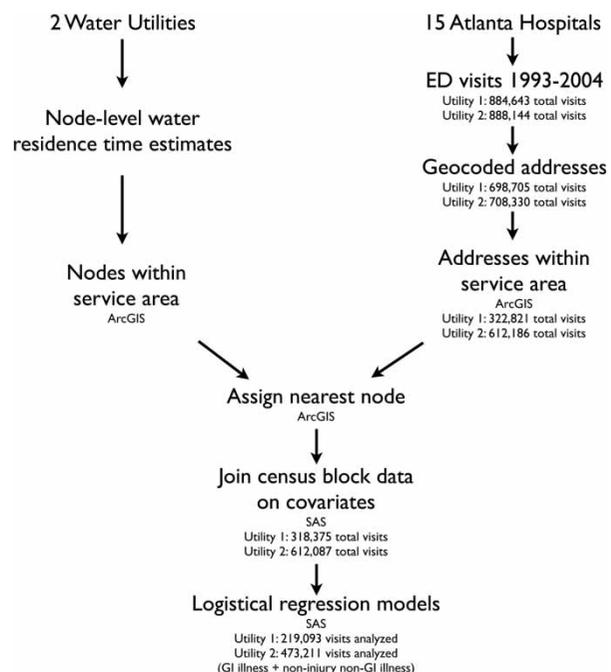
Figure 1 | Data sources and data processing steps involved in the analysis of the association between emergency department (ED) visits and WRT for two water utilities in Metro Atlanta, GA, USA.

Table 1 | Comparison of approaches and data used in the present analysis versus that used by Tinker et al. (2009)

	Tinker et al. (2009)		Present analysis	
Exposure assessment approach	WRT aggregated by ZIP code		WRT assigned by matching geocoded address to nearest node in the distribution system	
Differences in calculation of control variables	• Median household income and percentage minority aggregated by ZIP code • Distance from hospital calculated to zip code centroid		• Median household income and percentage minority by census block • Distance from hospital calculated to geocoded residence	
# of hospitals included	28		15	
	Utility 1	Utility 2	Utility 1	Utility 2
Study period	1996–2003*	1993–2004	1999–2003*	1993–2004
# of ED visits analyzed**	721,982	1,205,816	219,093	473,211
# of GI illness ED visits	63,652	101,285	16,694	35,600
% of ED visits for GI illness	8.8%	8.4%	7.6%	7.5%
Mean WRT (Range)	32.8 h (4.5–88.4)	23.3 h (4.7–144.1)	55.9 h (0.5–336.0)	16.1 h (0.2–335.9)
Mean WRT, short category (≤10th percentile)	10.1 h (s.d. 0.9)	5.9 h (s.d. 0.7)	9.8 h (s.d. 2.9)	4.1 h (s.d. 1.1)
Mean WRT, medium category (11th–89th percentile)	33.4 h (s.d. 9.7)	18.5 h (s.d. 7.8)	44.1 h (s.d. 18.1)	14.0 h (s.d. 5.2)
Mean WRT, long category (≥90th percentile)	74.4 h (s.d. 11.8)	60.4 h (s.d. 32.6)	197.7 h (s.d. 80.5)	45.1 h (s.d. 28.7)

WRT, water residence time; ED, emergency department; GI, gastrointestinal.
*The period of study changed between the two analyses because a different hydraulic model was used to derive WRT estimates in Utility 1 for 1996–1998 vs. 1999–2003 and in the present analysis there was not sufficient sample size to run a separate analysis for the time period 1996–1998.
**Number of ED visits analyzed includes all visits for GI illness (cases) + all non-injury non-GI illness visits (comparison group).

008.8, 009), non-infectious GI illness (558.9), and nausea and vomiting plausibly related to GI illness (787.01–787.03, 787.91). Non-infectious GI illness was included in the case definition because previous research has shown that many infectious cases of GI illness are misclassified into this diagnostic category (Lew et al. 1991; Gangarosa et al. 1992; Schwartz et al. 1997). The control group included non-injury ED visits without GI illness. The spatial distribution of ED usage for injuries might be different from the spatial distribution of the source population, so we excluded ED visits for injuries in the comparison group in an effort to better track the spatial distribution of the source population (i.e. those that would go to the ED for a GI illness if they had a GI illness). Repeat visits by a patient within a single day were counted as a single visit.

Ethics approval

This study protocol was approved by the Social, Humanist, and Behavioral Committee for the Protection of Human Subjects, by the Institutional Review Board at Emory University, Atlanta, GA, USA (Protocol #IRB00045761).

Water residence time data

WRT was estimated using extended period hydraulic models provided by the two water utilities. These models estimate the water travel times to all pipe intersections ('nodes'), which represents a flow-weighted average time for all travel paths between the treatment plant and the node within the distribution system pipe network. Further details of water age calculations are provided in Tinker et al. (2009). Tinker et al. (2009) utilized a usage-weighted average of the node WRT by ZIP code. Here we used individual node-level WRT estimates. WRT was estimated for Utility 1 for 1999–2003 and Utility 2 for 1993–2004.

Spatial data processing

All ED visit records were geocoded using patient residential address information to enable assignment of WRT at the

closest node in the water distribution system. Geocoding was conducted using a standard geocoding algorithm in Streetmap for ArcGIS version 9.2 (ESRI, Redlands, CA). Records with invalid address information and observations with a geocoding confidence score of <50% were removed (Figure 1).

We joined node-level WRT data, geocoded ED visit data, and census block-level median household income and percentage minority data from the 2000 US Census (United States Census Bureau 2000) in ArcMap version 9.3 (ESRI, Redlands, CA), separately for each of the two service areas. The joining function assigned the nearest water distribution system node to each ED patient residence and calculated the Euclidean distance for each pairing. At this step, we excluded ED visits with residences that were located outside of the service area, as defined by geographic boundary data obtained from each utility (Figure 1). This addressed the issue that in addition to serving individual household customers directly, both utilities sell their water to several wholesale city customers that are located within the service areas, but whose water nodes are not included in the hydraulic models.

Statistical analysis

We used a logistic regression model to estimate the odds of ED patients having GI illness as a function of WRT at the node nearest their residence, comparing ED patients with GI illness (GI = 1) to non-injury ED patients without GI illness (GI = 0). The basic logistic model had the following form:

$$\mathrm{LogitP(GI = 1)} = \alpha + \sum_i \beta_i (\mathrm{WRT}_i) + \mathrm{covariates} \tag{1}$$

where WRT_i represents the ith category of WRT.

Separate models were run for Utility 1 and Utility 2. First, WRT was modeled in three categories, with *a priori* cut points: ≤10th percentile, 10th–90th percentile, ≥90th percentile. The middle category (10th–90th percentile) was the reference category, following Tinker *et al.* (2009). We also modeled WRT for Utility 2 using the cut points for Utility 1 as a comparison, in order to compare the same absolute levels for WRT decile categorization. Second, for ease of interpretation for water systems engineers, WRT was modeled using absolute

number of hours, in nine categories of 12-hour increments, with <12 h as the reference category.

Control variables (covariates) included: Medicaid status (patient level indicator – yes, no, or missing), median household income (census block level – continuous), percentage minority (census block level – continuous), age (patient level indicator with four levels: 0–5, 6–18, 19–64, ≥65), Euclidean distance between the hospital where the ED visit occurred and the geocoded patient residence (patient level indicator with three levels: <20th, 20th–80th, >80th percentile), hospital indicator variables denoting the hospital in which the ED visit occurred, year indicator variables, day-of-week indicator variables, season indicator variable, and product terms between patient age and hospital, patient age and distance to hospital, and patient age and Medicaid status. Control variables were chosen on the basis of being potential risk factors for GI illness that also vary spatially, and thus could also be associated with WRT.

For both utilities, we also used a logistic regression model to compare the odds of ED patients having GI illness as a function of straight-line Euclidian distance between x-y coordinates from households to nodes, controlling for the same set of variables as described above. We compared distances of 30–150 m and > 150 m with a reference category of <30 m.

RESULTS

Summary statistics

Frequencies of visits for GI illness among non-injury ED visits during the study periods are shown in Table 1. The geocoding procedure resulted in 698,705 visits with valid geocodes for Utility 1 and 708,330 visits with valid geocodes for Utility 2. This represents an overall success rate of 80.6%. Total numbers of visits and nodes included in the analysis are presented in Table 1. Geocoded data were available for only approximately half of the hospitals and total visits included in the analysis by Tinker *et al.* (2009). Utility 1 had a significantly higher average WRT than Utility 2 (*p* < 0.0001), with a mean of 55.9 h (range: 0.51–336 h) versus 16.1 h (range: 0.24–336 h), as expected given the difference in the size of their service areas.

Epidemiologic analysis

Figure 2 shows the results of the logistic regression models, comparing the exposure assessment approach used here to the results we previously reported in Tinker *et al.* (2009). Using geocoded addresses, we found similar results for Utility 1 to those found by Tinker *et al.* (2009), with a slightly more elevated association between long WRT ($\geq$103.6 h) and GI illness (OR = 1.07, 95% CI: 1.02–1.13), and no protective effect of short WRT ($\leq$13.8 h) and GI illness (OR = 1.00, 95% CI: 0.95–1.06). For Utility 2, we did not observe an increased risk of longer WRT ($\geq$29.0 h) and GI illness in this analysis (OR = 0.98, 95% CI: 0.94–1.02), nor a protective effect of short WRT ($\leq$5.8 h) and GI illness (OR = 1.00, 95% CI: 0.96–1.04).

Because the cut point for long WRT was substantially greater for Utility 1 than for Utility 2, and because we observed stronger associations with GI ED visits for Utility 1, we also considered the decile cut points for Utility 1 in the analysis for Utility 2. This allowed for a direct comparison of the two utilities relative to the same WRT categorization. This analysis showed elevated risk for longer WRT, but with a very large confidence interval for the long WRT category risk ratio.

Figure 3 shows the results of the logistic regression models evaluating finer increments of WRT. While confidence intervals overlap among estimated effects, the strongest effects were observed at the longest WRT category (>96 h) for both utilities, and with a suggestion of an upward trend with step-wise increases in WRT for Utility 2.

We also evaluated the odds of increased risk of GI illness ED visits with Euclidian distance between geocoded patient residences and the location of the nearest distribution system node (Table 2). For Utility 1, we observed an increased odds of GI illness for patients living >150 m from the nearest node, compared with a reference category of <30 m (OR = 1.09, 95% CI: 1.03–1.15).

DISCUSSION

Ultimately, exposure to microbial contamination within the distribution system is dependent upon the occurrence of low-pressure conditions, the presence of a source of
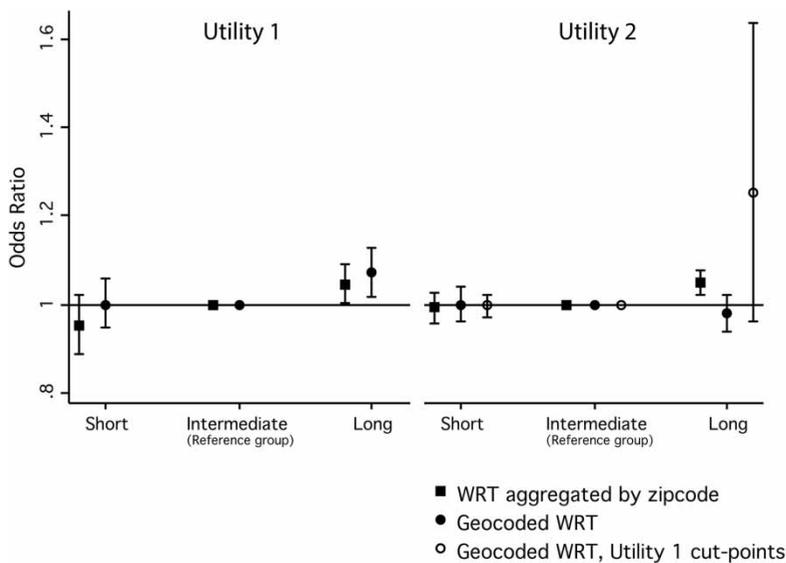

Figure 2 | Comparison of results of logistic regression models using different approaches to exposure assessment. Odds ratio (OR) estimates and 95% confidence intervals are shown for Utility 1 and Utility 2 for the association between water residence time (WRT) and visits to the ED for GI illness. WRT was classified as 'Short' ($\leq$10th percentile), 'Intermediate' (11th–89th percentile), or 'Long' ($\geq$90th percentile). Squares indicate OR using WRT aggregated by ZIP code, as reported by Tinker *et al.* (2009). Circles indicate OR using geocoded WRT, with address of patient matched to WRT of nearest node. Closed circles represent OR using the intermediate WRT category of each utility as its own reference. Open circles represent OR for Utility 2 relative to the WRT category percentile cut-off values of Utility 1. Models controlled for Medicaid status, median household income at the census block level, percentage minority at the census block level, patient age, Euclidean distance between the hospital where the ED visit occurred and the geocoded patient residence, indicator variables for hospital, year, day-of-week, season, and product terms for age*hospital, age*distance to hospital, age*Medicaid status.
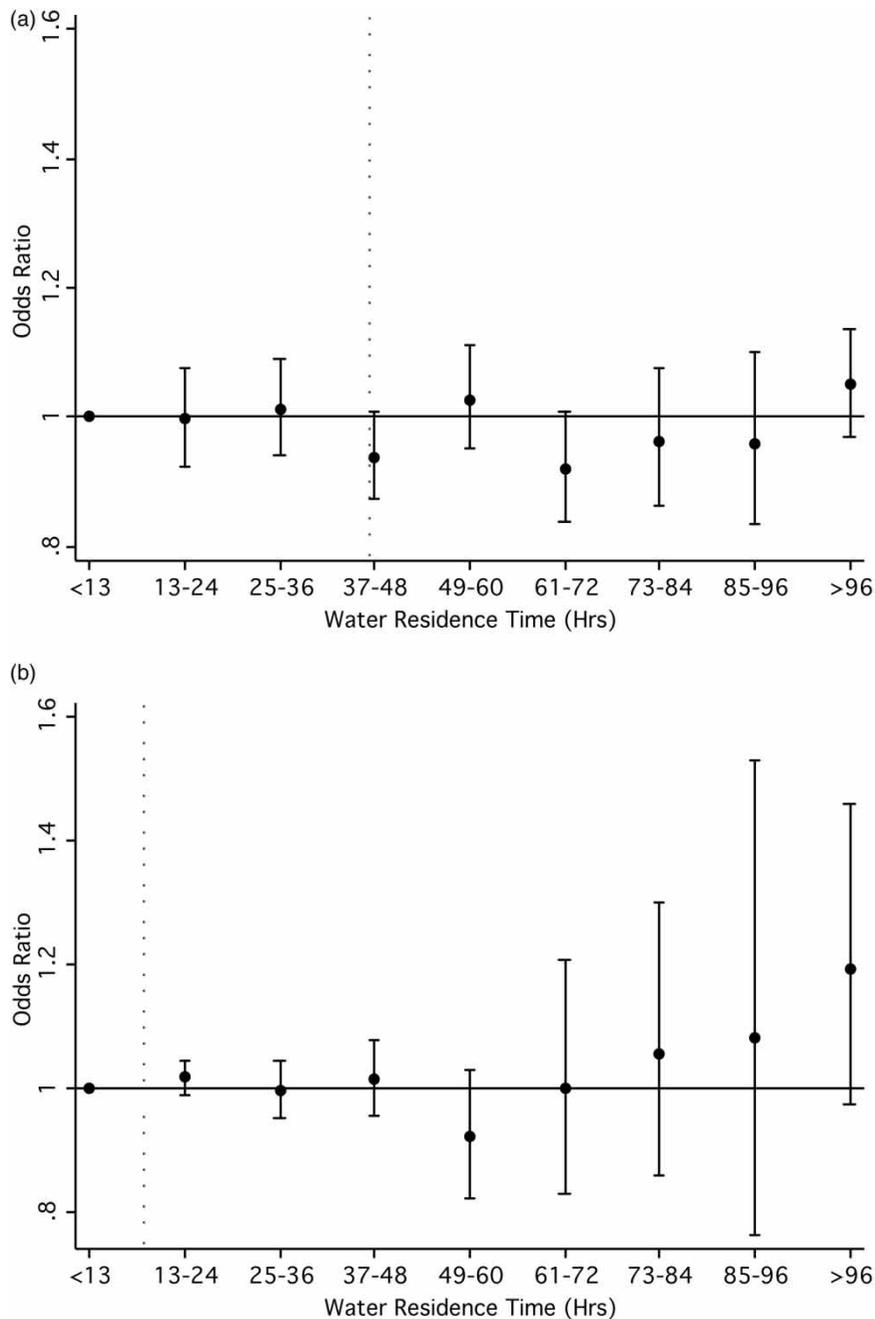
Figure 3 │ Results of logistic regression models examining the relationship between WRT (12-h increments) vs. ED visits for GI illness. Odds ratio (OR) estimates and 95% confidence intervals are shown for (a) Utility 1 and (b) Utility 2, using geocoded WRT estimates, with address of patient matched to WRT of nearest node. Reference condition for each estimate is <12 h. Mean WRT for each utility is shown with a dotted line. Models controlled for Medicaid status, median household income at the census block level, percentage minority at the census block level, patient age, Euclidean distance between the hospital where the ED visit occurred and the geocoded patient residence, indicator variables for hospital, year, day-of-week, and product terms for age*hospital, age*distance to hospital, age*Medicaid status.

contamination, and the existence of a pathway for entry of external contaminants, which together can create the necessary conditions for intrusion events. The concentrations of microbes and duration of intrusion events are key factors

influencing exposure (Besner *et al.* 2011). We use WRT here as a proxy for a large number of processes within the drinking water distribution system that determine whether a person opens their tap at the time when a contaminant is passing

Table 2 | Results of logistic regression model estimating the odds of ED visits for GI illness as a function of Euclidian distance between households and the nearest node in the drinking water distribution system

	Distance to nearest node	
	30–150 m vs. < 30 m	> 150 m vs. < 30 m
Utility 1	OR = 1.02	OR = 1.09
n = 219,093	(0.98–1.06)	(1.03–1.15)
Utility 2	OR = 0.98	OR = 0.96
n = 473,211	(0.95–1.01)	(0.92–1.00)

Models controlled for Medicaid status, median household income at the census block level, percentage minority at the census block level, patient age, Euclidean distance between the hospital where the ED visit occurred and the geocoded patient residence, indicator variables for hospital, year, day-of-week, and product terms for age*hospital, age*distance to hospital, age*Medicaid status.

through and drinks that water. Our assumption is that longer WRT increases the probability of a number of intrusion events occurring.

We conducted this analysis to explore how a more geographically resolved exposure assignment might affect the associations between WRT and ED visits for GI illness observed by Tinker *et al.* (2009), who assigned aggregated WRT by ZIP code. We performed this analysis for two drinking water utilities in the greater Atlanta metropolitan area, with different demographic and distribution system characteristics. In particular, Utility 1 had a substantially higher average WRT compared with Utility 2.

Comparing long WRT (≥90th percentile) with intermediate residence times (11th to 89th percentile), we observed a modestly increased risk for GI illness for Utility 1 (OR = 1.07, 95% CI: 1.02–1.13) but not for Utility 2 (OR = 0.98, 95% CI: 0.94–1.02), whereas Tinker *et al.* (2009) observed a modestly increased risk for GI illness for both utilities (Utility 1: OR = 1.07, 95% CI: 1.03–1.10; Utility 2: OR = 1.05, 95% CI: 1.02–1.08). When we compared the same categories of WRT for both utilities (i.e. based on quintiles of WRT at Utility 1), we observed an increased odds of illness but with a large amount of uncertainty in the estimates (OR = 1.25, 95% CI: 0.96–1.64), because of the few WRT observations that occurred in the long WRT category at Utility 2 in this analysis. Fewer than 1% of the Utility 2 observations occurred in the highest 20th percentile of Utility 1 WRTs.

It is important to note that Tinker *et al.* (2009) made an *a priori* decision to use the intermediate WRT category as the reference group and to explore the possibility of increased

risk related to short WRTs due to reduced contact time with disinfectant in the distribution system, in addition to exploring the possibility of increased risk related to longer WRTs due to increased probability of contamination within the distribution system. In order to maintain consistency between the analyses, we also used the intermediate WRT category as the reference group.

The analysis of 12-hour increments of WRT in relation to ED visits for GI illness provided additional insights that are more interpretable to water treatment plant operators. Examining these finer increments of WRT, we found that only exposures >48 h were associated with increased risk of GI illness, and exposures of >96 h had the strongest associations (Figure 3), although none of these relationships was statistically significant. While there was very little data for Utility 2 in the longest WRT categories, in both utilities we observed an upward trend in risk at the longest absolute WRTs. Again, these relationships were not statistically significant.

By comparison, Payment *et al.* (1991, 1997) reported WRT ranging from 0.3 h to 34 h in Montreal, in studies that found no correlation between WRT and highly credible GI illness. Hellard *et al.* (2001) reported average WRTs of 24–36 h in Melbourne, where no difference was found between groups with functional versus sham household water treatment devices. Thus, the distribution systems observed in these studies may not have reached values of WRT to cause elevated risk of GI illness. However, because WRT is a proxy for likelihood of pathogen intrusion, and each distribution system is unique, we would not necessarily expect the same absolute values across systems (e.g. an older system may have more intrusions within the same WRT relative to a newer system).

Exposure assessment

This analysis provided the opportunity to compare different approaches to WRT exposure assignment. The node-level WRT used here provided a more refined exposure assignment than WRT aggregated by ZIP code, and provided the opportunity for an individual-level, rather than ecological analysis. Aggregation of WRT estimates across nodes within a ZIP code smooths out heterogeneous exposures across an area. Thus, a small number of ZIP codes can have a large influence on the results, and might misrepresent the overall exposure for a particular ZIP code.

However, it is important to note that there were differences in the two datasets analyzed (Table 1), and our refined exposure assignment approach presented several limitations. Assigning node-level WRT to ED patients reduced the sample size available for the analysis as it required limiting the dataset to ED visits with patient residential address information; this reduced the number of hospitals and our total sample size by about half compared with the Tinker *et al.* (2009) analysis. More specific exposure assignment also opens up the possibility of errors related to address and node assignment. Assigning WRT to a specific node in the distribution system might assume too much specificity; i.e. the nearest node by Euclidian distance might not in fact be the node from which the household receives its water, due to complexities in the distribution system. In addition, individuals may consume water outside of the home, but still in the same general vicinity, for example at a local school or childcare center. More aggregated exposure assessment averages out these types of errors.

Our exposure assessment was also limited to describing WRT at distribution system nodes, rather than the tap, at the point of use. The ideal measure of WRT would describe time for water to reach the user at their residence. The availability of geocoded data for visitors to the ED, as well as the locations of nodes, allowed us to examine Euclidian distance between residence and node in the distribution system. We found that for Utility 1, ED visitors living >150 m from the nearest node in the distribution system had 9% increased odds (95% CI: 3–15%) of having GI illness compared with those living <30 m from the nearest node (Table 2). This analysis must be interpreted with caution, because Euclidian distance to a distribution system node does not necessarily represent whether a household receives water from that distribution system node. However, distance between households and drinking water delivery nodes might be an area to explore in future investigations.

It is also important to note that we had no data on water consumption patterns, so our analyses relied on the assumption that patients were consuming water directly from the distribution system, and we also had no information on quality of premise plumbing. Consumption of purchased water or point-of-use treatment of distribution system water would therefore affect the results we observed. Additionally, the assignment of WRT at each node is subject to error due to factors such as misspecifications of the hydraulic model,

mixing of water of different ages, and assignment of only one WRT measurement per node per year.

CONCLUSIONS

Our results suggest that drinking water utilities might consider reducing absolute WRTs to <2–3 days or consider adding booster disinfection in areas of the distribution system with longer WRT in order to minimize the risk of GI illness from consumption of municipal water. Continued investment is needed to maintain the current quality of drinking water infrastructure.

ACKNOWLEDGEMENTS

REFERENCES

Besner, M., Prévost, M. & Regli, S. 2011 Assessing the public health risk of microbial intrusion events in distribution systems: conceptual model, available data, and challenges. *Water Res.* **45**, 961–979.

Borchardt, M. A., Spencer, S. K., Kieke, B. A., Lambertini, E. & Loge, F. J. 2012 Viruses in nondisinfected drinking water from municipal wells and community incidence of acute gastrointestinal illness. *Environ. Health Perspect.* **120** (9), 1272–1279.

CDC 2011 Community health impact of extended loss of water service: Alabama, January 2010. *MMWR Morb. Mortal. Wkly Rep.* **60** (6), 161–166.

CDC 2013 Surveillance for waterborne disease outbreaks associated with drinking water and other nonrecreational water – United States, 2009–2010. *MMWR Morb. Mortal. Wkly Rep.* **62** (35), 714–720.

Colford Jr, J. M., Wade, T. J., Sandhu, S. K., Wright, C. C., Lee, S., Shaw, S., Fox, K., Burns, S., Benker, A., Brookhart, M. A., van der Laan, M. & Levy, D. A. 2005 A randomized, controlled trial of in-home drinking water intervention to reduce gastrointestinal illness. *Am. J. Epidemiol.* **161** (5), 472–482.

Colford Jr, J. M., Hilton, J. F., Wright, C. C., Arnold, B. F., Saha, S., Wade, T. J., Scott, J. & Eisenberg, J. N. 2009 The Sonoma water evaluation trial: a randomized drinking water intervention trial to reduce gastrointestinal illness in older adults. *Am. J. Public Health* **99** (11), 1988–1995.

Craun, G. F., Brunkard, J. M., Yoder, J. S., Roberts, V. A., Carpenter, J., Wade, T., Calderon, R. L., Roberts, J. M., Beach, M. J. & Roy, S. L. 2010 Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. *Clin. Microbiol. Rev.* **23** (3), 507–528.

Egorov, A., Ford, T., Tereschenko, A., Drizhd, N., Segedevich, I. & Fourman, V. 2002 Deterioration of drinking water quality in the distribution system and gastrointestinal morbidity in a Russian city. *Int. J. Environ. Health Res.* **12** (3), 221–233.

Egorov, A. I., Naumova, E. N., Tereschenko, A. A., Kislitsin, V. A. & Ford, T. E. 2003 Daily variations in effluent water turbidity and diarrhoeal illness in a Russian city. *Int. J. Environ. Health Res.* **13** (1), 81–94.

Ercumen, A., Gruber, J. S. & Colford, J. M. 2014 Water distribution system deficiencies and gastrointestinal illness: a systematic review and meta-analysis. *Environ. Health Perspect.* **122** (7), 651–660.

Gangarosa, R. E., Glass, R. I., Lew, J. F. & Boring, J. R. 1992 Hospitalizations involving gastroenteritis in the United States, 1985: the special burden of the disease among the elderly. *Am. J. Epidemiol.* **135**, 281–290.

Hellard, M. E., Sinclair, M. I., Forbes, A. B. & Fairley, C. K. 2001 A randomized, blinded, controlled trial investigating the gastrointestinal health effects of drinking water quality. *Environ. Health Perspect.* **109**, 773–778.

Hunter, P. R., Chalmers, R. M., Hughes, S. & Syed, Q. 2005 Self-reported diarrhea in a control group: a strong association with reporting of low-pressure events in tap water. *Clin. Infect. Dis.* **40** (4), e32–e34.

Lambertini, E., Borchardt, M. A., Kieke Jr, B. A., Spencer, S. K. & Loge, F. J. 2012 Risk of viral acute gastrointestinal illness from nondisinfected drinking water distribution systems. *Environ. Sci. Technol.* **46** (17), 9299–9307.

Lew, J. F., Glass, R. I., Gangarosa, R. E., Cohen, I. P., Bern, C. & Moe, C. L. 1991 Diarrheal deaths in the United States, 1979 through 1987. A special problem for the elderly. *JAMA* **265**, 3280–3284.

MacKenzie, W. R., Schell, W. L., Blair, K. A., Addiss, D. G., Peterson, D. E., Hoxie, N. J., Kazmierczak, J. J. & Davis, J. P. 1995 Massive outbreak of waterborne cryptosporidium infection in Milwaukee, Wisconsin: recurrence of illness and risk of secondary transmission. *Clin. Infect. Dis.* **21** (1), 57–62.

Malm, A., Axelsson, G., Barregard, L., Ljungqvist, J., Forsberg, B., Bergstedt, O. & Pettersson, T. J. 2013 The association of drinking water treatment and distribution network disturbances with Health Call Centre contacts for gastrointestinal illness symptoms. *Water Res.* **47** (13), 4474–4484.

Nygard, K., Wahl, E., Krogh, T., Tveit, O. A., Bohleng, E., Tverdal, A. & Aavitsland, P. 2007 Breaks and maintenance work in the water distribution systems and gastrointestinal illness: a cohort study. *Int. J. Epidemiol.* **36** (4), 873–880.

Payment, P., Richardson, L., Siemiatycki, J., Dewar, R., Edwardes, M. & Franco, E. 1991 Randomized trial to evaluate the risk of gastrointestinal disease due to consumption of drinking water meeting microbiological standards. *Am. J. Public Health* **81**, 703–708.

Payment, P., Siemiatycki, J., Richardson, L., Renaud, G., Franco, E. & Prevost, M. 1997 A prospective epidemiological study of gastrointestinal health effects due to the consumption of drinking water. *Intern. J. Environ. Health Res.* **7**, 5–31.

Schwartz, J., Levin, R. & Hodge, K. 1997 Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* **8**, 615–620.

Teunis, P. F., Xu, M., Fleming, K. K., Yang, J., Moe, C. L. & Lechevallier, M. W. 2010 Enteric virus infection risk from intrusion of sewage into a drinking water distribution network. *Environ. Sci. Technol.* **44** (22), 8561–8566.

Tinker, S. C., Moe, C. L., Klein, M., Flanders, W. D., Uber, J., Amirtharajah, A., Singer, P. & Tolbert, P. E. 2009 Drinking water residence time in distribution networks and emergency department visits for gastrointestinal illness in Metro Atlanta, Georgia. *J. Water Health* **7** (2), 332–343.

Tinker, S. C., Moe, C. L., Klein, M., Flanders, W. D., Uber, J., Amirtharajah, A., Singer, P. & Tolbert, P. E. 2010 Drinking water turbidity and emergency department visits for gastrointestinal illness in Atlanta, 1993–2004. *J. Expo. Sci. Environ. Epidemiol.* **20** (1), 19–28.

Tostmann, A., Bousema, T. & Oliver, I. 2012 Investigation of outbreaks complicated by universal exposure. *Emerg. Infect. Dis.* **18** (11), 1717–1722.

United States Census Bureau 2000 General Housing Characteristics: 2000. In: *American FactFinder*. United States Census Bureau, Washington, DC.

USEPA 2011a *Statement of Lisa P. Jackson Administrator, U.S. Environmental Protection Agency October 15, 2010. Quoted in: Aging Water Infrastructure Research: Science and Engineering for a Sustainable Future*. Report# EPA/600/F-11/010, Office of Research and Development NRMRL.

USEPA 2011b *Aging Water Infrastructure Research: Science and Engineering for a Sustainable Future*. Report# EPA/600/F-11/010, Office of Research and Development NRMRL.