# Drinking water residence time in distribution networks and emergency department visits for gastrointestinal illness in Metro Atlanta, Georgia

Sarah C. Tinker, Christine L. Moe, Mitchel Klein, W. Dana Flanders, Jim Uber, Appiah Amirtharajah, Philip Singer and Paige E. Tolbert

## ABSTRACT

We examined whether the average water residence time, the time it takes water to travel from the treatment plant to the user, for a zip code was related to the proportion of emergency department (ED) visits for gastrointestinal (GI) illness among residents of that zip code. Individual-level ED data were collected from all hospitals located in the five-county metro Atlanta area from 1993 to 2004. Two of the largest water utilities in the area, together serving 1.7 million people, were considered. People served by these utilities had almost 3 million total ED visits, 164,937 of them for GI illness. The relationship between water residence time and risk for GI illness was assessed using logistic regression, controlling for potential confounding factors, including patient age and markers of socioeconomic status (SES). We observed a modestly increased risk for GI illness for residents of zip codes with the longest water residence times compared with intermediate residence times (odds ratio (OR) for Utility 1 = 1.07, 95% confidence interval (CI) = 1.03, 1.10; OR for Utility 2 = 1.05, 95% CI = 1.02, 1.08). The results suggest that drinking water contamination in the distribution system may contribute to the burden of endemic GI illness.

**Key words** | drinking water, epidemiology, gastrointestinal illness

**Sarah C. Tinker** (corresponding author)
**Mitchel Klein**
**Paige E. Tolbert**
Department of Environmental and Occupational
Health,
Emory University, 1518 Clifton Road NE,
Atlanta, GA 30322,
USA
Tel.: +1 404-727-3697
Fax: +1 404-727-8744
E-mail: sceaser@sph.emory.edu

**Christine L. Moe**
Hubert Department of Global Health,
Emory University, 1518 Clifton Road NE,
Atlanta, GA 30322,
USA

**W. Dana Flanders**
Department of Epidemiology,
Emory University, 1518 Clifton Road NE,
Atlanta, GA 30322,
USA

**Jim Uber**
Department of Civil and Environmental
Engineering,
University of Cincinnati,
765 Baldwin Hall, Cincinnati, OH 45221,
USA

**†Appiah Amirtharajah**
Georgia Institute of Technology,
School of Civil and Environmental Engineernig,
Atlanta, Georgia 30332-0512,
USA

**Philip Singer**
Department of Environmental Sciences and
Engineering,
University of North Carolina,
144 Rosenau Hall, Chapel Hill, NC 27599,
USA

## INTRODUCTION

There is increasing recognition of the role that ageing pipe networks used to distribute treated drinking water play in infectious disease outbreaks in the United States. Between 1971 and 1998, drinking water distribution system contamination caused 113 reported outbreaks in the US, resulting in 21,000 cases of illness, 498 hospitalizations and 13 deaths (Craun & Calderon 2001). From 1999 through 2004, half (11/22) of all reported outbreaks in

community water systems were due to distribution system deficiencies, resulting in 202 additional cases of water-borne disease (Lee *et al.* 2002; Blackburn *et al.* 2004; Liang *et al.* 2006; Nygard *et al.* 2007). While most reported distribution system outbreaks were caused by significant incidents (i.e. main breaks) or continuing contamination, low-level or transient contamination is also likely to occur in distribution systems, and the potential for endemic disease transmission through this route is substantial. Randomized-controlled trial results from Canada have suggested a role for the distribution system in contributing to endemic drinking water-related gastrointestinal (GI) illness (Payment *et al.* 1997). These results have been supported by observational studies in Russia and Sweden that found an association between distance from the treatment plant and incidence of GI illness (Egorov *et al.* 2002; Nygard *et al.* 2004) and also studies in the United Kingdom and Sweden in which an association was observed between pressure-loss events in the distribution system and incidence of GI illness (Nygard *et al.* 2004, 2007; Hunter *et al.* 2005). The US Environmental Protection Agency is currently revising one of the primary regulations governing drinking water distribution system water quality in the US, the Total Coliform Rule (US Environmental Protection Agency 1989). Research on the contribution of distribution system water quality to endemic GI illness in the US can better inform this process.

Measuring the frequency and magnitude of contamination events in a distribution system is not currently feasible, and therefore indicators of the potential for contamination may be a useful alternative. Residence time, or water age, is the amount of time water spends in the distribution system between the treatment plant and the consumer. The longer the water is in the distribution system, the more likely it is to encounter contamination. Residence time is a function of flow rate, distance from the treatment plant, storage, system demand and distribution system architecture, among other factors. It is estimated using hydraulic models that incorporate these factors to simulate the flow of water through the distribution system pipes. We present the results of a study that used residence time estimates in a novel way to assess the association between distribution system contamination

and endemic GI illness among the population served by two large drinking water utilities in the metro Atlanta, Georgia, area.

## METHODS

### Emergency department data

The emergency department (ED) database included information on visits from all of the hospitals operating within the five-county Atlanta area and from hospitals located outside the study area that contributed a substantial number of visits by five-county residents (28 hospitals). The information provided by the hospitals included medical record number, date of admission, International Classification of Diseases, 9th Revision (ICD-9) diagnosis codes, zip code of residence, payment method, and age or date of birth. Data from one hospital had to be excluded from the analysis due to missing covariate information regarding method of payment. Data from two hospitals operated by the same healthcare system could not be separated and were considered together in the analytical model.

We defined cases of GI illness using the primary and all available secondary ICD-9 diagnostic codes. This case definition included the following diagnoses: infectious GI illness (001–004, 005.0, 005.4, 005.89, 005.9, 006–007, 008.0, 008.42–008.44, 008.47, 008.49, 008.5, 008.6, 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 (Gangarosa *et al.* 1992; Hoxie *et al.* 1997; Schwartz *et al.* 1997).

### Distribution system residence time data

Residence time was estimated using hydraulic models from two of the six major utilities serving Atlanta and its suburbs, referred to in this paper as Utility 1 and Utility 2. Utility 1 and Utility 2 each operate two treatment plants that supply water to their respective distribution pipeline networks.

These 'extended period' hydraulic models simulate the flow patterns of water throughout the distribution system over a typical diurnal water usage cycle, taking into account pipe layout and characteristics, system storage, customer water usage and utility operating rules. The models had been developed previously by Utility 1 and Utility 2, and their consultants, to assist with design and planning of water system improvements and extensions.

The models estimate the water travel times to all pipe intersections, called 'nodes'. A water distribution pipe network typically has redundant travel paths between a treatment plant source and any node, and so the calculated travel time at a node is the flow-weighted average time for all travel paths; hydraulic engineers refer to this travel time as the 'water age', denoted $a_{ik}$, where $i$ is the node index and $k$ the simulation time step. To account for the time variation of water age at any one location due to diurnal variations in demand, we calculated the average age of water used at node $i$ over one diurnal water usage cycle, $a_i = (1/Q_i)\sum_{k=1}^{n} q_{ik} a_{ik}$, where $q_{ik}$ is the water usage rate at node $i$ during time $k$, $Q_i = \sum_{k=1}^{n} q_{ik}$ is the total water usage at node $i$ during one diurnal water usage cycle, and $n$ is the number of time steps in one cycle. When calculating this average water age, initial values were assumed for each node and the simulations were conducted for a time period of 15 days to allow the value to stabilize. The average water age over the final diurnal cycle (day 15) was taken as the water age for that node.

For this analysis, we aggregated the water age values, $a_i$, over all nodes in each zip code in the service area, separately for each year from 1996 through 2003 for Utility 1 and from 1993 to 2004 for Utility 2. Since these values were intended to represent the quality of water used within a zip code, the aggregated value was defined as the usage-weighted average of the node water age

$$r_j = \frac{1}{\sum_{i \in S_j} Q_i} \sum_{i \in S_j} Q_i a_i,$$

where $S_j$ is the set of nodes within zip code $j$ and $r_j$ are the zip code 'residence times' used in the analysis of GI illness data.

Only zip codes served by the utility that had their centroid located within the boundary of the five-county metro Atlanta area were included in the analysis (Utility 1: 19 zip codes; Utility 2: 37 zip codes). One zip code served by Utility 1 and two zip codes served by Utility 2 were excluded from the analysis due to missing covariate census data (described in more detail below). Two hydraulic models from Utility 1 were utilized. One described the distribution system from 1996 through 1998 and the other from 1999 through 2003. Only one hydraulic model from Utility 2 was utilized, covering 1993 to 2004.

## Other covariate data

We obtained data regarding zip code level demographic characteristics, including median income and percentage minority, from the 2000 US Census (United States Census Bureau 2000).

## Analytic methods

All analyses were performed using SAS statistical software (SAS 2002–2003). We developed unconditional logistic regression models a priori to consider the association between residence time and ED visits for GI illness. Residence time was considered as a three-level categorical variable, with cut points at the 10th and 90th percentile, weighted by population (res_time10, res_time90). Residence time estimates from each of the hydraulic models were considered separately (cut-points: 8.34 hours and 36.37 hours, respectively, for Utility 1, hydraulic model 1; 11.18 hours and 51.78 hours, respectively, for Utility 1, hydraulic model 2; 8.62 and 40.31 hours, respectively, for Utility 2). The models also included the following covariates to control for potential confounding: a four-level categorical age variable (age 0 to 5, 6 to 18, 19 to 64, and 65 + years), indicator variables for year, season and hospital; a three-level categorical variable for the distance from zip code centroid to hospital (dis_h$_1$, dis_h$_2$; cut-points at the 20th and 80th percentiles); continuous variables for zip code median income and zip code percentage minority, and a dichotomous variable for Medicaid payment status. The models took the following

conceptual form:

$$\text{Logit}[\text{P}(Y_{z,a,y,s,h,m} = 1)]$$

$$= \alpha + \beta_1(\text{res\_time10}) + \beta_2(\text{res\_time90})$$

$$+ \sum_i \chi_i(\text{age}_i) + \sum_j \delta_j(\text{year}_j)$$

$$+ \sum_k \varepsilon_k(\text{season}_k + \sum_l \phi_l(\text{hospital}_l)$$

$$+ \sum_n \gamma_n(\text{dis\_h}_n) + \tau(\text{median income})$$

$$+ \varphi(\text{percent minority}) + \kappa(\text{Medicaid})$$

$$+ \sum_p \sum_q \lambda_{pq}(\text{dis\_h}_p{}^*\text{age}_q) + \sum_r \sum_s \pi_{rs}$$

$$(\text{dis\_h}_r{}^*\text{hospital}_s) + \sum_t \theta_t(\text{Medicaid}^*\text{age}_t)$$

This model estimates the probability that a person visiting an ED for a non-injury cause had a GI illness, given the estimated drinking water residence time of their zip code of residence, while controlling for potentially confounding factors.

Models were run separately for each utility and, for Utility 1, for the periods 1996 through 1998 and 1999 through 2003, corresponding to the time period covered by each of the hydraulic models used to estimate water residence time. We calculated a weighted average of the risk ratio parameter estimates for Utility 1 based on the inverse of the variance. The analyses were stratified in this manner because the water residence time estimates derived from the different hydraulic models were not directly comparable. The earlier hydraulic model from Utility 1 included details of only the larger pipes, while the later hydraulic model incorporated details on all of the pipes. These differences in the extent of skeletonization of the network led the water residence time estimates derived from the earlier model to be systematically lower than those derived from the later model.

In addition to demographic and temporal covariates, we also considered the hospital at which the ED visit occurred and the distance from the hospital to the zip code centroid as covariates in the model. The proportion of non-injury ED visits for GI illness varied by hospital. This variation was probably greater than that explained by differences in drinking water exposure between the hospital catchment areas, which led to concern about the potential for confounding by other factors associated with ED visits for GI illness. We considered the distance from zip code centroid to hospital as a potential confounding factor because results from preliminary descriptive analyses suggested that the rate of ED visits for GI illness in the population of a given zip code decreased as the distance between that zip code and the hospital increased. This association varied based on age group and hospital, so interaction terms for these variables were also included in the model.

We considered models stratified by year, season and age category, and we used likelihood ratio tests to assess effect measure modification.

## RESULTS

Twenty-seven hospitals provided data on 2,092,735 non-injury ED visits in the two utilities' service areas; 164,937 (7.9%) of these visits were for GI illness. There were no reported drinking waterborne disease outbreaks in metro Atlanta during the study period. Of the ED visits for GI illness, 29% were among children age 5 years or younger, while only 17% of non-GI illness visits were among people in this age group (Table 1). The highest proportion of ED visits for GI illness occurred in winter months, while the highest proportion of all other non-injury visits occurred during autumn months. Twenty-two per cent of GI illness patients and 18% of non-injury, non-GI illness patients paid with Medicaid.

The average residence time derived from the first hydraulic model from Utility 1, covering 1996 through 1998, was 24.7 hours (standard deviation (SD) = 16.1 hours) and from the second hydraulic model from Utility 1, covering 1999 through 2003, was 32.8 hours (SD = 18.2 hours). The maximum residence time was 88.4 hours and the minimum was 4.5 hours. The average water residence time for the Utility 2 service area was 23.3 hours (SD = 23.5 hours). The maximum and minimum water residence times for Utility 2 were 144.1 hours and 4.7 hours, respectively. (The mean water residence times for each exposure category are described in Table 2.) As expected,

**Table 1** │ Descriptive statistics for emergency department visits for gastrointestinal (GI) illness and non GI-illness among residents of zip codes in the service areas of two drinking water utilities, metro Atlanta, 1993–2004

| | Visits for GI illness (% of all visits for GI illness) | | Visits for non-GI illness[*] (% of all visits for non-GI illness) | |
| --- | --- | --- | --- | --- |
| | Utility 1 | Utility 2 | Utility 1 | Utility 2 |
| **Year** | | | | |
| 1993 | N/A[†] | 2,024 (2.0) | N/A | 27,615 (2.3) |
| 1994 | N/A | 3,829 (3.8) | N/A | 54,882 (4.6) |
| 1995 | N/A | 7,788 (7.7) | N/A | 91,045 (7.6) |
| 1996 | 2,345 (3.7) | 7,233 (7.1) | 31,263 (4.3) | 87,322 (7.2) |
| 1997 | 5,662 (8.9) | 8,648 (8.5) | 65,734 (9.1) | 106,689 (8.8) |
| 1998 | 7,758 (12.2) | 8,988 (8.9) | 83,326 (11.5) | 107,292 (8.9) |
| 1999 | 8,225 (12.9) | 8,653 (8.5) | 92,173 (12.8) | 103,782 (8.6) |
| 2000 | 9,254 (14.5) | 9,724 (9.6) | 107,900 (14.9) | 122,123 (10.1) |
| 2001 | 9,859 (15.5) | 10,106 (10.0) | 109,337 (15.1) | 116,647 (9.7) |
| 2002 | 9,368 (14.7) | 10,075 (9.9) | 110,678 (15.3) | 118,227 (9.8) |
| 2003 | 11,181 (17.6) | 11,491 (11.3) | 121,571 (16.8) | 126,618 (10.5) |
| 2004 | N/A | 12,726 (12.6) | N/A | 143,574 (11.9) |
| **Age category** | | | | |
| 0–5 Years | 19,336 (30.4) | 27,892 (27.5) | 123,953 (17.2) | 199,863 (16.6) |
| 6–18 Years | 7,512 (11.8) | 10,419 (10.3) | 78,351 (10.9) | 119,724 (9.9) |
| 19–64 Years | 32,095 (50.4) | 53,209 (52.5) | 428,639 (59.4) | 710,450 (58.9) |
| 65 + Years | 4,709 (7.4) | 9,765 (9.6) | 91,039 (12.6) | 175,779 (14.6) |
| **Season** | | | | |
| Winter | 18,503 (29.1) | 29,248 (28.9) | 184,011 (25.5) | 316,545 (26.3) |
| Spring | 16,682 (26.2) | 26,638 (26.3) | 176,898 (24.5) | 299,000 (24.8) |
| Summer | 13,315 (20.9) | 22,453 (22.2) | 173,021 (24.0) | 287,752 (23.9) |
| Autumn | 15,152 (23.8) | 22,946 (22.7) | 188,052 (26.0) | 302,519 (25.1) |
| **Medicaid payment status** | | | | |
| Paid with Medicaid | 13,612 (21.4) | 22,415 (22.1) | 120,614 (16.7) | 219,040 (18.2) |
| Did not pay with Medicaid | 50,040 (78.6) | 78,870 (77.9) | 601,368 (83.3) | 986,776 (81.8) |
| Total | 63,652 | 101,285 | 721,982 | 1,205,816 |

[*]Does not include emergency department visits for injuries.
[†]Water residence time estimates were not available for Utility 1 for 1993–1995 and 2004.

**Table 2** │ Descriptive statistics[*] of drinking water residence time exposure categories[†], metro Atlanta

| | Short water residence time category, mean (standard deviation) | Intermediate water residence time category, mean (standard deviation) | Long water residence time category, mean (standard deviation) |
| --- | --- | --- | --- |
| Utility 1, hydraulic model 1 (1996–1998) | 6.83 (1.36) | 21.96 (9.54) | 47.40 (11.76) |
| Utility 1, hydraulic model 2 (1999–2003) | 10.12 (0.89) | 33.43 (9.73) | 74.41 (12.80) |
| Utility 2 (1993–2004) | 5.85 (0.65) | 18.45 (7.79) | 60.39 (32.58) |

[*]In hours. Summarized over years and zip codes.
[†]Short water residence time defined as ≤ 10th percentile of all water residence time estimates, stratified by hydraulic model; intermediate = 11th to 89th percentile; long = ≥ 90th percentile.

those zip codes closest to the treatment plants tended to have the shortest residence time (Figure 1(a) and (b)).

The results of our analyses suggested that there was no difference in the rate of ED visits for GI illness among people living in zip codes with short residence time compared with that for people living in zip codes with intermediate residence times (odds ratio (OR) for Utility 1 = 1.00, 95% confidence interval (CI) = 0.96, 1.03; OR for Utility 2 = 0.99, 95% CI = 0.96, 1.03) (Figure 2). However, we observed a modest increased risk for ED visits for GI illness among people living in zip codes with the longest residence time compared with people living in zip codes with intermediate residence times (OR for Utility 1 = 1.07, 95% CI = 1.03, 1.10; OR for Utility 2 = 1.05, 95% CI = 1.02, 1.08).

The strength of the association between water residence time and risk for ED visits for GI illness varied by year for both utilities, although no overall pattern was discernable in the year-specific results (Figure 3). For Utility 2, the results stratified by season and age suggested stronger associations between long water residence times and GI illness during non-summer months (Figure 4) and among children 18 years and younger (Figure 5).

## DISCUSSION

The results of the study suggest that people served by drinking water that has spent the greatest amount of time in the distribution system may have a slightly increased risk of ED visits for GI illness compared with people who receive water that spends less time in the distribution system. These results support the recent focus on the distribution system as a source of waterborne illness and the current deliberations about the revisions to the Total Coliform Rule in the US.

Studies of distribution system water quality have demonstrated that, as distance from the plant increases, the level of bacterial contamination can increase (Payment *et al.* 1988). There are many routes through which transient distribution system contamination may occur. Low-pressure events, typically caused by main breaks, sudden changes in demand, uncontrolled pump starting or stopping, opening and closing of fire hydrants, power failures, sudden air
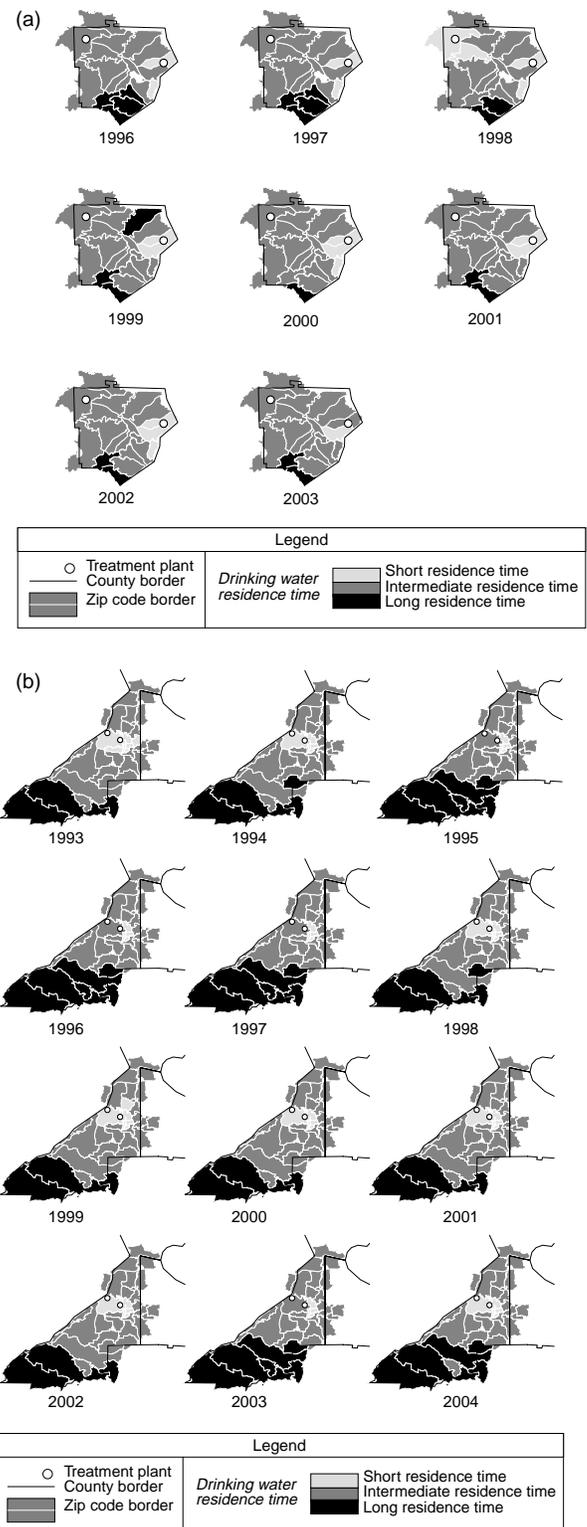


**Figure 1** | Classification of mean water residence time by zip code by year (a) Utility 1; (b) Utility 2.
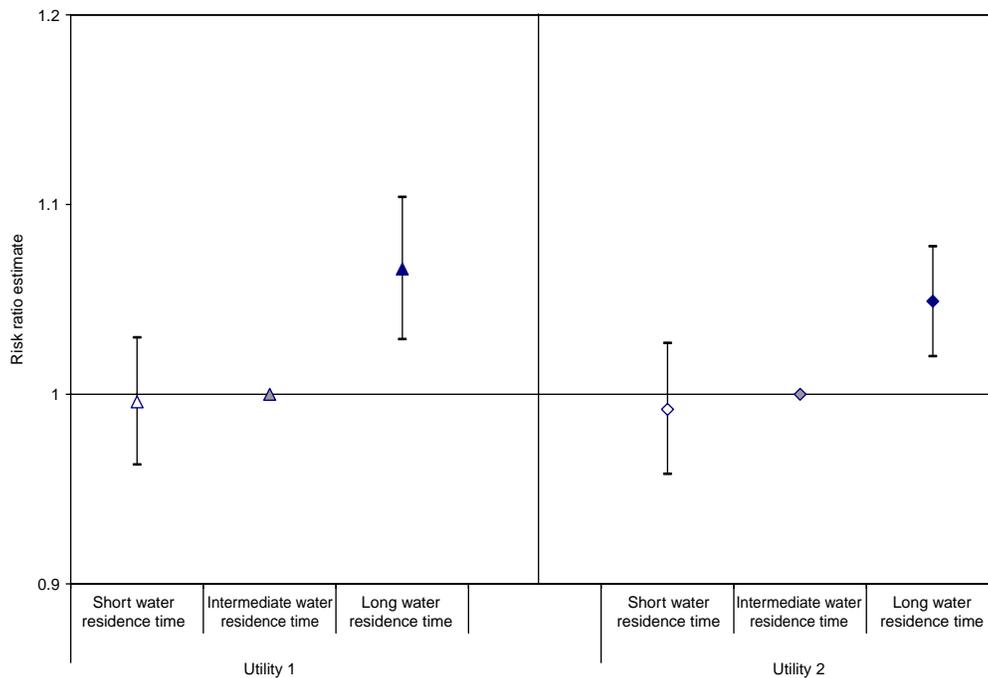
**Figure 2** │ Risk ratio estimates for emergency department visits for gastrointestinal illness among people living in zip codes with short (≤10th percentile) drinking water residence times and long (≥90th percentile) drinking water residence times compared with intermediate (11th–89th percentile) drinking water residence times, metro Atlanta, stratified by utility.

valve closures and flushing operations can decrease the normally positive pressure maintained in the pipes and increase the chance that non-potable water is drawn in from outside the pipes through cracks in the pipe wall or at pipe joints (LeChevallier *et al.* 2003). Leaks are common in distribution systems throughout the world, and account for 10% or more of the piped water that is lost (American Water Works Association Research Foundation 1992; Kirmeyer *et al.* 2001; LeChevallier *et al.* 2003). Microbiological studies have demonstrated that there can be high levels of faecal contamination and human viruses in the soil and water adjacent to distribution system pipes (Karim *et al.* 2003; LeChevallier *et al.* 2003). Although engineering standards call for a minimum separation of 0.5 to 3 m between drinking water pipes and sewer lines, microbes can migrate several metres in short periods of time under certain conditions (Abu-Ashour *et al.* 1994; LeChevallier *et al.* 2003). This combination of negative pressure and proximity to contamination can result in pathogenic organisms entering the distribution system.

Distribution system contamination can also occur through cross-connections with non-potable water sources,

during main breaks and repair activities, and biofilm disruption. Biofilms form when corrosion in the distribution system pipes produces tubercles, thereby increasing the surface area of the pipe and providing niches where bacteria and other organisms are protected from disinfection (LeChevallier *et al.* 1996). In addition, biofilms accelerate further pipe corrosion, consume disinfectant residual and promote the exchange of resistance or virulence factors between the mixed microbial population in close proximity with each other (Ford 1999). Biofilms can release pathogenic organisms into drinking water when flow disruptions occur, such as during routine flushing (Trussell 1999).

While regulations in the US require a disinfectant residual to be maintained in water throughout the distribution system, such a residual is not always present in all parts of the distribution system. Even if a residual is present, it may not be present in sufficient concentration to inactivate a large influx of pathogens. This may be especially true at the far ends of the distribution system where the residual may be low. However, we observed only a small decrease in the average chlorine residual in the zip codes in the 90th percentile of water residence time compared with
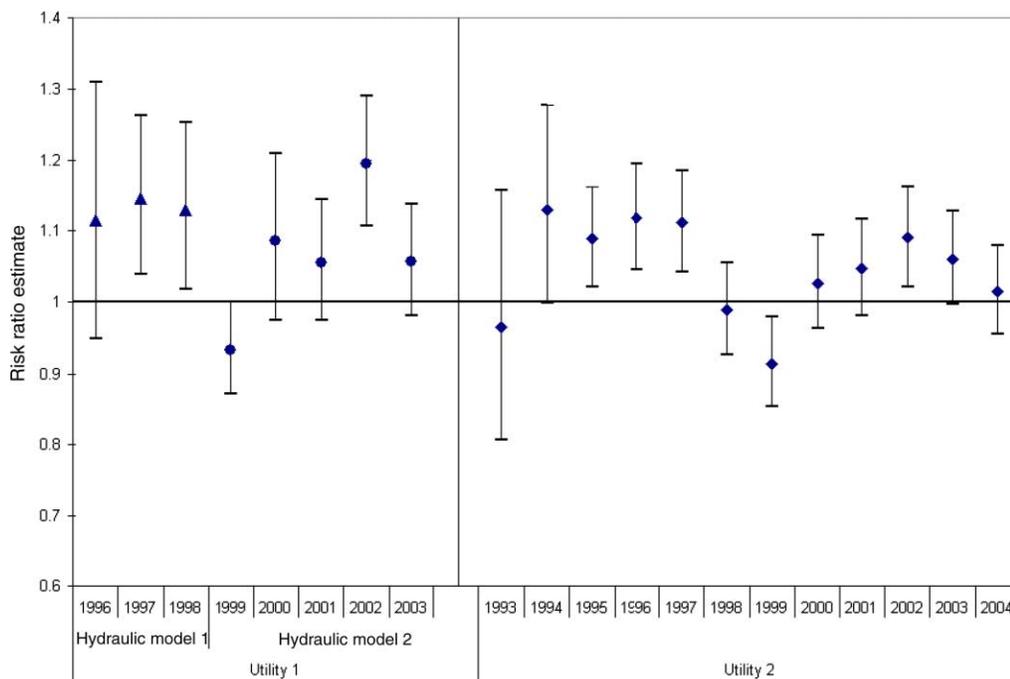
**Figure 3** | Risk ratio estimates for emergency department visits for gastrointestinal illness among people living in zip codes with long (≥90th percentile) drinking water residence times compared with intermediate (11th–89th percentile) drinking water residence times, metro Atlanta, stratified by year.

the average chlorine residual in the zip codes with shorter residence times (results not shown). Payment (1999) has argued that typical chlorine residuals are insufficient to inactivate most microorganisms in the distribution system. In addition, some waterborne pathogens are quite resistant to disinfection.

The transient nature of low pressure events results in the total volume of contaminated water to be small, and therefore less likely to be detected during routine sampling (Karim *et al.* 2003). In a study in England, Hunter *et al.* (2005) found an odds ratio of 12.5 for the association between a self-reported experience of a water pressure loss in the preceding two weeks and the occurrence of a diarrhoeal episode (Hunter *et al.* 2005). Norwegian investigators recently reported results of a cohort study in which documented main breaks or other pressure loss events were associated with an increased risk of GI illness in the following week (Nygard *et al.* 2007).

Few studies have examined the association between deteriorating water quality in the distribution system and GI illness. Secondary data analysis from a randomized trial found no correlation between estimated drinking water residence time and incidence of GI illness (Payment *et al.* 1997). Correlations take into account only linear relationships, however, and our results suggest an increased risk only for those served by water travelling for the longest amount of time. A study in Russia found that, as the distance from the plant increased, the chlorine residual decreased, and that this decreased chlorine residual was associated with higher rates of GI illness (Egorov *et al.* 2002). Secondary analyses of data from a different randomized-controlled trial also suggested higher rates of GI illness as distance from the plant increased (Payment *et al.* 1993).

We chose the intermediate residence category as the referent based on the a priori hypothesis that short residence time has the potential for association with both decreased and increased risk of GI illness and long residence time has the potential to increase risk of GI illness, as well as the stability of the intermediate residence time category (50,936 ED visits for GI illness). The water at the beginning of the distribution system generally has the lowest water residence time and has less opportunity to encounter contamination within the distribution system;
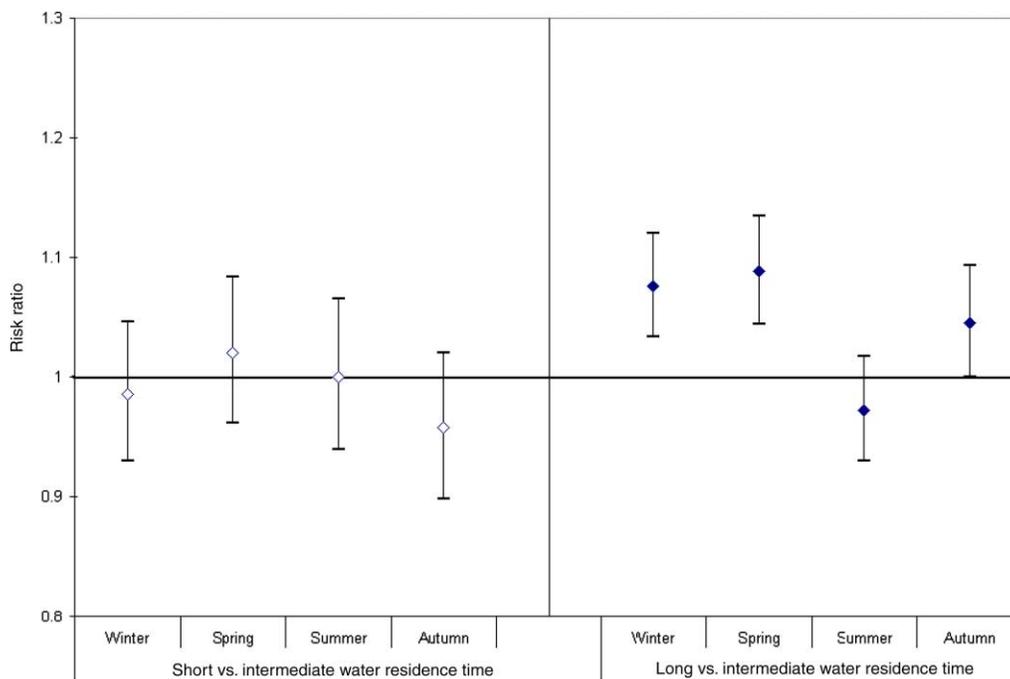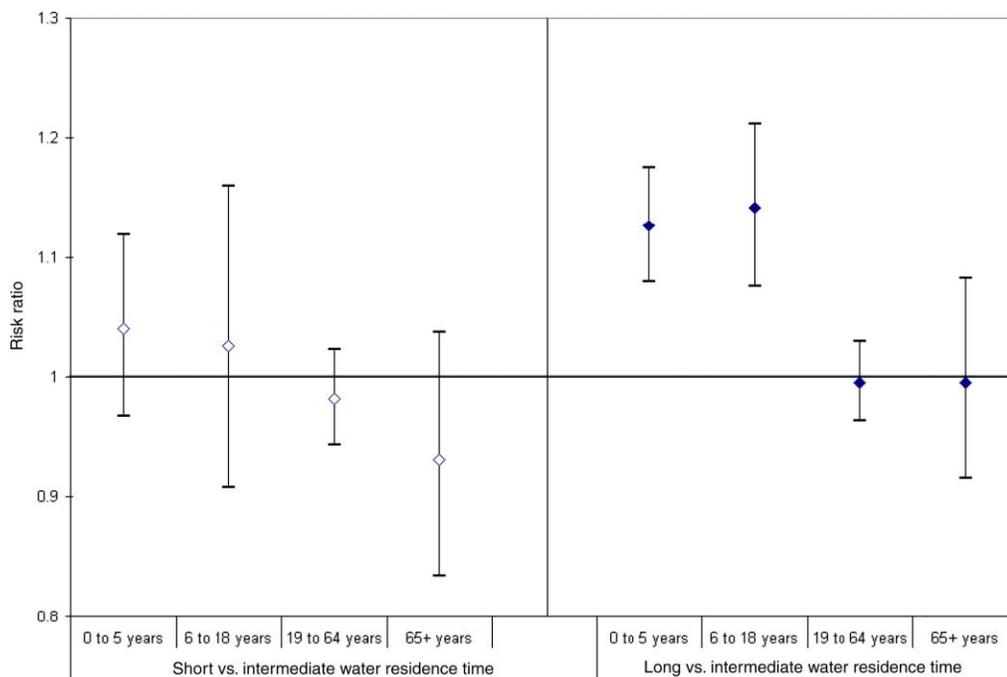
**Figure 4** | Risk ratio estimates for emergency department visits for gastrointestinal illness among people living in zip codes with short (≤10th percentile) drinking water residence times and long (≥90th percentile) drinking water residence times compared with intermediate (11th–89th percentile) drinking water residence times, metro Atlanta, stratified by season, Utility 2 only.

therefore people receiving this water may be at lower risk of GI illness. Alternatively, the water at the beginning of the distribution system also has a shorter disinfection contact time with the disinfectant residual in the distribution system, and if contamination from the raw water source was not eliminated at the plant, this water may pose a greater health risk than water that has been exposed to the disinfectant for a longer amount of time. A long water residence time is associated with a greater probability of encountering distribution system contamination.

In sensitivity analyses, we considered alternative coding schemes for the water residence time exposure variable to assess whether our choice of cut-points for the short and long water residence time categories, at the 10th and 90th percentiles of all water residence times, respectively, were appropriate. Water residence time was considered as: 1) a continuous, cubic variable; 2) a ten-level categorical variable based on deciles of the distribution of water residence times; and 3) a three-level categorical variable with alternative cut-points. The results of these analyses suggest that the association between water residence time and GI illness was confined to the top decile of the distribution of water residence times, consistent with our a priori coding scheme.

A number of limitations of our analysis must be considered when interpreting the results. Water residence time is a marker for the quality of water consumed. People will consume water outside of their home and people often drink bottled water or water that has received point-of-use treatment, such as in-home filtration. These sources of misclassification were probably independent of water residence time. Therefore, any associated bias would be expected to bias the risk ratio estimates toward the null. Another source of exposure misclassification comes from the water residence time estimates themselves. Although they offer a major improvement over previous studies, they were estimated from hydraulic models intended to represent typical water use and operational patterns. We summarized multiple water residence time estimates for the nodes in each zip code and assigned this single value to all of the individuals in that zip code. In reality, residence time among different residents of a single zip code may vary appreciably. We assigned only a single average water residence time to each zip code for a given year, yet we

**Figure 5** | Risk ratio estimates for emergency department visits for gastrointestinal illness among people living in zip codes with short (≤10th percentile) drinking water residence times and long (≥90th percentile) drinking water residence times compared with intermediate (11th–89th percentile) drinking water residence times, metro Atlanta, stratified by age group, Utility 2 only.

know that actual residence times vary temporally, on seasonal and hourly time scales.

Our hospital dataset captured only those cases of GI illness severe enough to warrant a visit to an ED, a small proportion of all cases of GI illness. However, these more severe cases may be those of greatest interest in terms of the public health impact of waterborne disease contracted through drinking water. The large size of our ED dataset provided sufficient power to detect modest associations, which we expected based on the results observed in the majority of previously published studies on the association between drinking water distribution system exposures and endemic GI illness (Payment *et al.* 1997; Egorov *et al.* 2002; Nygard *et al.* 2004, 2007), and the potential for bias to the null discussed above. While the inclusion of only two utilities may limit the generalizability of the results, the consistency of the results observed for two independently operated utilities lends strength to the validity of our observations.

Although we attempted to control for factors that may have confounded the association between water residence time and ED visits for GI illness, some confounding

probably remained. Because residence time was assigned based on zip code, it is important to account for any zip code characteristics that might have been associated with GI illness. A challenging factor to control for in this analysis was socioeconomic status (SES). We attempted to account for SES using two group-level variables, zip code median income and zip code percentage minority, and one individual-level variable, whether or not the patient paid for the ED visit using Medicaid. There are likely to be other spatially varying and individual factors that we were not able to consider that may have influenced our results, including other risk factors for GI illness, such as day care attendance, recreational water exposure or food consumption habits.

There is also the potential for confounding if an outbreak of GI illness due to a cause other than drinking water led to a cluster of cases in parts of the study area. A recreational waterborne outbreak of *Escherichia coli* occurred at a water park located in the service area of Utility 1 in 1998. However, there were only 26 confirmed cases and the park was located in a zip code classified in the referent category, and therefore any clustering of cases in

this area would be expected to bias results negatively (Gilbert & Blake 1998). In sensitivity analyses we excluded the outbreak time period and our conclusions were unchanged.

## CONCLUSIONS

The results of this study suggest that there is a slightly increased risk of GI illness associated with drinking water that spends the longest amount of time in the distribution system. These results also highlight the useful metrics that can be estimated using hydraulic models. Given the growing body of evidence that links drinking water distribution systems and endemic GI illness, hydraulic models should continue to be used to inform health studies and identify areas of vulnerability, with the ultimate goal of improving distribution system integrity, decreasing the water age to the extent possible and delivering safe water to consumers.

## ACKNOWLEDGEMENTS

## REFERENCES

Abu-Ashour, J., Joy, D. M., Lee, H., Whiteley, H. R. & Zelin, S. 1994 Transport of microorganisms through soil. *Water Air Soil Pollut.* **75**, 141–158.

American Water Works Association Research Foundation 1992 *Water Industry Database: Utility Profiles*. American Water Works Association, Denver, Colorado.

Blackburn, B. G., Craun, G. F., Yoder, J. S., Hill, V., Calderon, R. L., Chen, N., Lee, S. H., Levy, D. A. & Beach, M. J. 2004 Surveillance for waterborne-disease outbreaks associated with drinking water: United States, 2001–2002. *MMWR Surveill. Summ.* **53**(8), 23–45.

Craun, G. F. & Calderon, R. L. 2001 Waterborne disease outbreaks caused by distribution system deficiencies. *J. Am. Water Works Assoc.* **93**(9), 64–75.

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.

Ford, T. E. 1999 Microbiological safety of drinking water: United States and global perspectives. *Environ. Health Perspect.* **107**(Suppl 1), 191–206.

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**(3), 281–290.

Gilbert, L. & Blake, P. 1998 Outbreak of *Escherichia coli* O157:H7 infections associated with a water park. *Georgia Epidemiol. Rep.* **14**(7), 1–2.

Hoxie, N. J., Davis, J. P., Vergeront, J. M., Nashold, R. D. & Blair, K. A. 1997 Cryptosporidiosis-associated mortality following a massive waterborne outbreak in Milwaukee, Wisconsin. *Am. J. Public Health* **87**(12), 2032–2035.

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.

Karim, M. R., Abbaszadegan, M. & LeChevallier, M. W. 2003 Potential for pathogen intrusion during pressure transients. *J. Am. Water Works Assoc.* **95**(5), 134–146.

Kirmeyer, G. J., Friedman, M., Martel, K., Howie, D., LeChevallier, M. W., Abbaszadegan, M., Karim, M. R., Funk, J. E. & Harbour, J. 2001 *Pathogen Intrusion Into the Distribution System*. American Water Works Association Research Foundation and American Water Works Association, Denver, Colorado.

LeChevallier, M. W., Welch, N. J. & Smith, D. B. 1996 Full-scale studies of factors related to coliform regrowth in drinking water. *Appl. Environ. Microbiol.* **62**(7), 2201–2211.

LeChevallier, M. W., Gullick, R. W., Karim, M. R., Friedman, M. & Funk, J. E. 2003 The potential for health risks from intrusion of contaminants into the distribution system from pressure transients. *J. Water Health* **1**(1), 3–14.

Lee, S. H., Levy, D. A., Craun, G. F., Beach, M. J. & Calderon, R. L. 2002 Surveillance for waterborne-disease outbreaks: United States, 1999–2000. *MMWR Surveill. Summ.* **51**(8), 1–47.

Liang, J. L., Dziuban, E. J., Craun, G. F., Hill, V., Moore, M. R., Gelting, R. J., Calderon, R. L., Beach, M. J. & Roy, S. L. 2006 Surveillance for waterborne disease and outbreaks associated with drinking water and water not intended for drinking: United States, 2003–2004. *MMWR Surveill. Summ.* **55**(12), 31–65.

Nygard, K., Andersson, Y., Rottingen, J. A., Svensson, A., Lindback, J., Kistemann, T. & Giesecke, J. 2004 Association between

environmental risk factors and *Campylobacter* infections in Sweden. *Epidemiol. Infect.* **132**(2), 317–325.

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. 1999 Poor efficacy of residual chlorine disinfectant in drinking water to inactivate waterborne pathogens in distribution systems. *Can. J. Microbiol.* **45**(8), 709–715.

Payment, P., Gamache, F. & Paquette, G. 1988 Microbiological and virological analysis of water from two water filtration plants and their distribution systems. *Can. J. Microbiol.* **34**(12), 1304–1309.

Payment, P., Franco, E. & Siemiatycki, J. 1993 Absence of relationship between health effects due to tap water consumption and drinking water quality parameters. *Water Sci. Technol.* **27**(3–4), 137–143.

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

SAS 2002–2003 *SAS 9.1.* SAS Institute, Inc., Cary, North Carolina.

Schwartz, J., Levin, R. & Hodge, K. 1997 Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* **8**(6), 615–620.

Trussell, R. 1999 Safeguarding distribution system integrity. *J. Am. Water Works Assoc.* **91**, 46–54.

United States Census Bureau 2000 *General Housing Characteristics: 2000. Census Summary File 1 (SF 1) 100-Percent Data.* American FactFinder, United States Census Bureau, Washington, DC.

United States Environmental Protection Agency 1989 Total coliforms (including fecal coliforms and *E. coli*). Final rule. *Fed. Regist.* **54**(124), 27544–27568.