Identifying public water facilities with low spatial variability of disinfection by-products for epidemiological investigations

A F Hinckley, A M Bachand, J R Nuckols, J S Reif

Background and Aims: Epidemiological studies of disinfection by-products (DBPs) and reproductive outcomes have been hampered by misclassification of exposure. In most epidemiological studies conducted to date, all persons living within the boundaries of a water distribution system have been assigned a common exposure value based on facility-wide averages of trihalomethane (THM) concentrations. Since THMs do not develop uniformly throughout a distribution system, assignment of facility-wide averages may be inappropriate. One approach to mitigate this potential for misclassification is to select communities for epidemiological investigations that are served by distribution systems with consistently low spatial variability of THMs.

Methods and Results: A feasibility study was conducted to develop methods for community selection using the Information Collection Rule (ICR) database, assembled by the US Environmental Protection Agency. The ICR database contains quarterly DBP concentrations collected between 1997 and 1998 from the distribution systems of 198 public water facilities with minimum service populations of 100 000 persons. Facilities with low spatial variation of THMs were identified using two methods; 33 facilities were found with low spatial variability based on one or both methods. Because brominated THMs may be important predictors of risk for adverse reproductive outcomes, sites were categorised into three exposure profiles according to proportion of brominated THM species and average THM concentration. The correlation between THMs and haloacetic acids (HAAs) in these facilities was evaluated to see whether selection by total trihalomethanes (TTHMs) corresponds to low spatial variability for HAAs. TTHMs were only moderately correlated with HAAs (r = 0.623).

Conclusions: Results provide a simple method for a priori selection of sites with low spatial variability from state or national public water facility datasets as a means to reduce exposure misclassification in epidemiological studies of DBPs.
Main messages

- Epidemiological studies of disinfection by-products (DBPs) and reproductive outcomes have been hampered by misclassification of exposure, in part due to spatial variability of these contaminants within water distribution systems.

- One approach to mitigate exposure misclassification is to select communities for epidemiological investigations that are served by distribution systems with consistently low spatial variability of DBPs.

- We investigated two approaches to identify low spatial variability: the first based on a statistical analysis of variability (ANOVA) and the second based on concentration cut-points for total trihalomethanes (TTHMs) derived from prior epidemiological studies of birth outcomes. Both methods appear to be useful for community selection.

Policy implications

- Application of these techniques to identify communities with low spatial variability will reduce exposure misclassification and make the results of studies of DBPs and reproductive outcomes more useful for regulatory purposes and policy formulation.

- Performing epidemiological investigations of DBPs and adverse reproductive outcomes in communities served by water facilities with characteristically low spatial variability of THM contaminants will minimise the bias due to spatial variability and will help to elucidate the health effects these contaminants.

METHODS

Under the guidelines of the ICR, water samples were collected from the distribution systems of 198 water facilities. Samples were collected at the following locations: (1) two locations of the “average” residence time of water in the distribution system (AVG1 and AVG2); (2) the point farthest along in the distribution system or at a dead-end representing the maximum residence time of water in the distribution system (MAX); and (3) at a location with a known retention time within the distribution system (termed the distribution system equivalent, DSE). Samples for each quarterly “sampling event” were collected during a single 24 hour window.

We limited our study to facilities in the ICR database that had collected data for total THMs (TTHMs), the sum of individual THM species (chloroform (CHCl3), bromodichloromethane (CHBrCl2), dibromochloromethane (CHBr2Cl), and bromoform (CHBr3)), over four consecutive quarters (all seasons) and at four different sampling points in the distribution system from November 1997 to October 1998. Eighty four of the 198 facilities in the ICR database met this requirement. We extracted data for total (TTHMs) and individual THM species, and the sum of five haloacetic acids (HAA5) (monobromoacetic, monochloroacetic, dibromoacetic, dichloroacetic, and trichloroacetic acids) for each of these facilities.

We selected facilities exhibiting low spatial variability in each of the four quarters using two methods: (1) a method based on two way analysis of variance (ANOVA); and (2) a method based on concentration cut-points for TTHM derived from prior epidemiological studies of birth outcomes. We compared these methods with respect to factors that limited their validity and number of sites found eligible.

Two way ANOVA

We used a two way ANOVA procedure to compare TTHM concentrations between the four ICR sampling points in the presence of an extraneous variable, quarter, which reflected season. This allowed us to evaluate spatial variability between sampling points independently from temporal variability between quarters. Facilities with low intra-system spatial variability were defined as those where spatial variability did not depend on season according to Tukey’s test for non-additivity, and where there was no evidence of a statistically significant difference (p ≥ 0.05) between concentrations of TTHM by sampling point.

Cut-points

Epidemiological studies investigating the effects of TTHMs have suggested increased risks of low birth weight, small for gestational age, spontaneous abortion, and selected birth defects at specific estimated exposure levels of TTHMs. In general, positive associations have been described at TTHM levels greater than 80 µg/l, whereas TTHM levels below 40 µg/l do not appear not to be associated with adverse effects. Facilities exhibiting consistently low intra-system spatial variability through four quarters were defined as those where, in a given quarter, all four sample values fell within one of three exposure groups (low, <40 µg/l; medium, 40–80 µg/l; or high, >80 µg/l).

Validation

To ensure that sites selected as having low intra-system spatial variability would also be selected in subsequent years, we performed validation analyses on results from both ANOVA and cut-point methods by substituting TTHM concentrations from the fall 1997 and summer 1998 seasons with fall measurements from 1998 and summer measurements from 1997. The additional data were available for 52 (62%) of the 84 sites that had data for six quarters. The validation was conducted by calculating the sensitivity and specificity of our spatial variability classification methods. We cross-tabulated the number of sites identified by classification as low or high spatial variability for each method according to the original and repeat analyses. The results from the original analyses were used as the reference or “gold standard” data.

Concentration profiles and brominated THMs

We classified those facilities in the USEPA ICR database selected as having consistently low spatial variability by either the ANOVA or cut-point method into one of three
exposure profiles: (1) high TTHM (≥80 μg/l), high proportion of brominated THMs; (2) high TTHM (≥80 μg/l), low proportion of brominated THMs; and (3) low overall TTHMs (<40 μg/l). A facility had to meet one of three criteria to be classified as having predominantly brominated THMs: (i) the mean concentration of brominated THM species for all sampling points exceeded 50 μg/l across all four quarters; (ii) the proportion of brominated species exceeded 50% at all sampling points during all four quarters; or (iii) the average bromine incorporation factor (BIF) across quarters was greater than or equal to 1.5. The BIF describes the molar contribution of all brominated species and is equal to:

$$0 \times [\text{CHCl}_3] + 1 \times [\text{CHBrCl}_2] + 2 \times [\text{CHBr}_2\text{Cl}] + 3 \times [\text{CHBr}_3]$$

$$[\text{CHCl}_3] + [\text{CHBrCl}_2] + [\text{CHBr}_2\text{Cl}] + [\text{CHBr}_3]$$

where in the numerator, the stoichiometric number of bromine moles in each THM species is multiplied by their respective concentration and the denominator is the concentration of TTHMs. Our cut-point of 1.5 is the midpoint of the possible range in BIFs, 0 to 3.

**Haloacetic acids**

The HAAs are a class of non-volatile by-products with demonstrated reproductive effects in animals at higher doses; however, with few exceptions, measurements of haloacetic acids have not been incorporated into epidemiological studies to date. In order to determine whether TTHM concentrations could predict HAA concentrations at a water facility, we evaluated the relation between concentrations of HAAs and TTHMs for the four sampling points using Spearman correlation coefficients for all sites identified by our methods as having consistently low spatial variability. We performed the same analyses to determine whether brominated THM concentrations could be used to predict brominated HAA (monobromoacetic and dibromoacetic acid) concentrations.

**RESULTS**

Descriptive statistics for TTHMs for the 84 US sites with complete reporting for four consecutive quarters are shown in table 1. Stratification by quarter revealed the expected seasonal fluctuation in TTHM concentrations with highest levels from July to September and lowest levels from January to March. The means at the four sampling points in the distribution system in each quarter suggest that TTHM concentrations increased with water residence time in the distribution system, with the lowest values seen at the earliest (DSE) and the highest values seen at the latest (MAX) sample points.

<table>
<thead>
<tr>
<th>ICR time period</th>
<th>Sampling point</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quarter 1 (Oct-Dec 97)</td>
<td>DSE</td>
<td>39.20</td>
<td>22.33</td>
<td>33.30</td>
<td>ND–108.9</td>
</tr>
<tr>
<td>AVG1</td>
<td>41.24</td>
<td>23.44</td>
<td>35.95</td>
<td>ND–130.8</td>
<td></td>
</tr>
<tr>
<td>AVG2</td>
<td>42.41</td>
<td>23.15</td>
<td>38.10</td>
<td>ND–104.4</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>48.76</td>
<td>25.35</td>
<td>42.35</td>
<td>ND–129.0</td>
<td></td>
</tr>
<tr>
<td>Quarter 2 (Jan–Mar 98)</td>
<td>DSE</td>
<td>30.91</td>
<td>16.59</td>
<td>28.35</td>
<td>3.8–74.8</td>
</tr>
<tr>
<td>AVG1</td>
<td>32.82</td>
<td>18.76</td>
<td>28.30</td>
<td>5.9–96.1</td>
<td></td>
</tr>
<tr>
<td>AVG2</td>
<td>32.91</td>
<td>16.26</td>
<td>31.30</td>
<td>ND–80.5</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>38.71</td>
<td>21.09</td>
<td>34.20</td>
<td>3.6–108.5</td>
<td></td>
</tr>
<tr>
<td>Quarter 3 (Apr–Jun 98)</td>
<td>DSE</td>
<td>45.17</td>
<td>27.37</td>
<td>39.15</td>
<td>1.1–136.0</td>
</tr>
<tr>
<td>AVG1</td>
<td>47.36</td>
<td>26.59</td>
<td>39.80</td>
<td>7.2–133.1</td>
<td></td>
</tr>
<tr>
<td>AVG2</td>
<td>48.10</td>
<td>28.34</td>
<td>42.05</td>
<td>3.2–136.0</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>54.68</td>
<td>28.68</td>
<td>48.60</td>
<td>4.3–136.4</td>
<td></td>
</tr>
<tr>
<td>Quarter 4 (Jul–Sept 98)</td>
<td>DSE</td>
<td>56.19</td>
<td>35.31</td>
<td>47.35</td>
<td>ND–193.0</td>
</tr>
<tr>
<td>AVG1</td>
<td>58.15</td>
<td>33.65</td>
<td>50.40</td>
<td>1.2–188.0</td>
<td></td>
</tr>
<tr>
<td>AVG2</td>
<td>61.12</td>
<td>38.31</td>
<td>53.00</td>
<td>ND–222.0</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>67.75</td>
<td>35.99</td>
<td>58.90</td>
<td>ND–204.7</td>
<td></td>
</tr>
</tbody>
</table>

**Validation**

Twelve of 17 facilities characterised as having consistently low spatial variability according to the ANOVA method had six quarters of data. When TTHM concentrations from the summer and fall quarters were replaced, 10 (sensitivity = (10/12) = 83%) were again found to exhibit low spatial variability using ANOVA. Of the sites originally excluded by the ANOVA method, 40 had data for six quarters. After replacing quarters, 28 (specificity = (28/40) = 70%) of these were found to have high spatial variability and 12 were considered to have low spatial variability.

We also reevaluated facilities selected by the cut-point method. The facility with high spatial variability and 12 were considered to have low spatial variability using the ANOVA method. Twelve of the sites originally selected by the cut-point method had six quarters of data in the ICR database. In this case, 10 of the 12 facilities (sensitivity = (10/12) = 83%) retained the classification of having consistently low spatial variability. Of the two facilities that were excluded, one had consistent low spatial variability of TTHM concentrations for three quarters, but a single abnormally high value for the MAX sample site led to high spatial variability in the last quarter. The other site was excluded for having a range (26–55 μg/l) including the lower cut-point of TTHM concentrations across sample locations for the summer quarter. All sites (n = 40) originally excluded due to spatial variability by
the cut-point method were again excluded (specificity = (40/40) = 100%) in the validation analyses.

Concentration profiles and brominated THMs
Table 2 describes the facilities with low spatial variability by exposure profile. Of the 17 water facilities selected by the ANOVA method, one had high and predominantly brominated TTHMs, three facilities had high TTHM levels with a low proportion of brominated compounds, and six had low overall TTHM levels. Seven facilities selected by the ANOVA had medium TTHM concentrations (40 μg/l ≤ mean TTHM < 80 μg/l) and were not classifiable into any profile.

Among the 20 facilities that met the criteria for low spatial variability using the cut-point method, there were no water facilities with high TTHM concentrations that were predominantly brominated (profile 1). Three facilities met the definition of low proportions of brominated THMs with high mean TTHM concentrations (profile 2). Thirteen facilities with low TTHM exposure fit into profile 3 with a mean concentration of 21.2 μg/l TTHMs; five of these were classified as predominantly brominated sites. Four facilities selected by the cut-point method had medium mean TTHM concentrations (40 μg/l ≤ mean TTHM < 80 μg/l) and were not classifiable into any profile.

In table 3, the mean concentration of the three brominated species, proportion of brominated compounds, BIF, and exposure profile are presented for the nine sites (one site was chosen by both methods) characterised as having low spatial variability and a predominantly brominated THM mixture. The exposure profile for these nine sites was obtained from their categorisation in table 2. The summary statistics suggested that the most useful predictor of a highly brominated system was the proportion of brominated compounds; the other criteria appeared to be highly variable.

HAAs
A moderate correlation (r = 0.667, p < 0.001) was found between TTHM values and HAA values for sampling events from facilities characterised as having consistently low spatial variability. The correlation was similar between mean concentrations of the brominated THM species and mean concentrations of the brominated HAA species (r = 0.642, p < 0.001).

DISCUSSION
We investigated two approaches to identify low spatial variability of DBPs in water distribution systems: the first based on a statistical ANOVA, and the second based on concentration cut-points for TTHMs derived from prior epidemiological studies of birth outcomes. Both methods appear to be useful for community selection. Identification of communities with low spatial variability will reduce exposure misclassification and make the results of studies of DBPs and reproductive outcomes more useful for regulatory purposes and policy formation.

In most epidemiological studies conducted to date, exposure assessment has been based on water facility data without accounting for spatial variability within the distribution system. A few investigators have attempted to account for intra-system spatial variability. Klotz and Pyrch and Dodds et al used residential tap water sampling to validate exposure for each subject. However, performing comprehensive water sampling for multiple DBP species is time and cost intensive and is difficult to accomplish in large studies. In addition, misclassification can occur when the samples are taken after the birth has occurred or a participant has moved. Gallagher et al incorporated a hydraulic model of the distribution system into a geographic information system, and assigned exposures to individual census block groups. While this method shows promise for prediction of DBP concentrations throughout the system, extensive validation of the model for each class of by-product is necessary. Recently, Waller et al used a weighting procedure to reduce the influence on risk estimates of individuals with less accurate exposure values. Weightings were based on degree of spatial variability in individual distribution systems and were assigned to each subject by residential location. This method may be appropriate for adjustment of biased exposure estimates when systems with high spatial variability must be utilised; however, as recently reported in a paper by Wright and Bateson, the applicability of this method may depend on degree of spatial variability and magnitude of “true” risk estimated for a given birth outcome.

We evaluated two methods for selection of populations served by water distribution systems with minimal spatial variability in the concentrations of TTHMs. Using the two way ANOVA method, we compared TTHM levels between four sampling points in the distribution system controlling
for the influence of season and identified 17 sites with low spatial variability in four consecutive seasons. A limitation of this method was that an arbitrary alpha level for site inclusion was chosen ($\alpha = 0.05$). Further, the power of the F-test was potentially differential for each facility, depending on the relative difference between sampling point means. Lastly, applying an inherently statistical approach does not incorporate biological plausibility. To illustrate, consider a hypothetical scenario in which the four TTHM levels measured during one quarter in the distribution system of facility A are 150, 170, 180, and 200 (variance = 433). In the distribution system of facility B, levels 70, 70, 90, and 90 were measured (variance = 133). Although the variance in facility B is much lower, facility A is a better choice for a high exposure profile number for nine predominantly brominated sites selected using ANOVA and cut-point methods

### Table 3

<table>
<thead>
<tr>
<th>Method</th>
<th>Facility</th>
<th>Mean TTHM (µg/l)</th>
<th>Mean brominated THMs (µg/l)</th>
<th>% brominated species</th>
<th>BIF</th>
<th>Exposure profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>126</td>
<td>118.4</td>
<td>51.3</td>
<td>54.0</td>
<td>1.54</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>56†</td>
<td>43.7</td>
<td>23.0</td>
<td>52.9</td>
<td>1.50</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>163</td>
<td>32.2</td>
<td>19.0</td>
<td>60.1</td>
<td>1.51</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>413</td>
<td>19.4</td>
<td>9.6</td>
<td>51.3</td>
<td>1.33</td>
<td>3</td>
</tr>
<tr>
<td>Cut-points</td>
<td>132†</td>
<td>51.8</td>
<td>36.6</td>
<td>62.7</td>
<td>1.50</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>535</td>
<td>24.1</td>
<td>18.9</td>
<td>78.7</td>
<td>1.77</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>415</td>
<td>19.9</td>
<td>9.8</td>
<td>51.0</td>
<td>1.34</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>355</td>
<td>19.4</td>
<td>10.0</td>
<td>52.5</td>
<td>1.36</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>413</td>
<td>19.4</td>
<td>9.6</td>
<td>51.3</td>
<td>1.33</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>416</td>
<td>19.0</td>
<td>9.2</td>
<td>50.0</td>
<td>1.32</td>
<td>3</td>
</tr>
<tr>
<td>Summary‡</td>
<td>Mean</td>
<td>34.9</td>
<td>17.8</td>
<td>55.7</td>
<td>1.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>32.2</td>
<td>13.6</td>
<td>9.1</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>20.2</td>
<td>10.0</td>
<td>52.5</td>
<td>1.36</td>
<td></td>
</tr>
</tbody>
</table>

†Site had medium average TTHM concentrations (40 µg/l < mean TTHM < 80 µg/l) and was not classifiable into any profile.
‡Site 413 was counted only once in the summary statistics.

*Bromodichloromethane, dibromochloromethane, bromoform.

The cut-point method may be more helpful in identifying sites with consistent low spatial variability from available state or national databases.

The cut-point method can be used to select water facilities having low spatial variability for a variety of study objectives. Of the 20 facilities identified by this method, 12 were classified into the same exposure category in each of the four consecutive quarters and would be appropriate for inter-community studies in which disease rates in communities with higher TTHM were compared with those in communities with lower TTHM levels. The exposure group classification (high, medium, low) of the eight other facilities with consistently low spatial variability changed over time. These facilities may be most suitable for studies employing an in-community approach where the effects of an exceedance above a given threshold concentration are evaluated. This approach is well suited to the study of birth defects since the critical period of development is likely to be known and one can investigate the effects of high exposures during that window of gestation within a single community.

Most previous epidemiological studies of reproductive outcomes have relied on TTHMs as the relevant exposure metric, while recent studies suggest that the composition of the mixture and the concentration of specific DBPs, especially brominated DBPs, may be of critical importance.13–15 18 22 Approximately 60% of sites selected by ANOVA (10/17) and 80% of sites selected by the cut-point (16/20) method fell into one of the exposure profiles that was selected for future studies. The sites selected fell mainly into profiles of high TTHM concentrations with minimal proportions of brominated species (profile 2) and low overall TTHM concentrations (profile 3). Only one site, selected by ANOVA, had a high proportion of brominated THMs and TTHM levels that met the criteria for profile 1. The paucity of sites identified for this exposure profile is likely due to the relative rarity of the description, rather than an inherent limitation of the classification method. The analysis was designed to provide a comparison between tools available to identify a site with predominantly brominated DBPs. The summary statistics suggested that the most useful predictor of a highly brominated facility was the proportion of brominated compounds since the other criteria appeared to be highly variable.

Toxicological data1 suggest that HAAs may be important contributors to risk for adverse reproductive outcomes, but with two exceptions,18 21 exposure to HAAs has not been...
incorporated into epidemiological studies. This is partly due to the deficiency of HAA data in the USA prior to 2002. Using data from the ICR for sites identified as having low spatial variability of TTHM concentrations, we found a moderate correlation with TTHM and HAA5. These inter-species correlations were generally consistent with the relation between TTHMs and the sum of five HAAs (r = 0.815) described by Villanueva and colleagues. In a recent study, King et al. found correlation coefficients of 0.74 and 0.52 between TTHM and HAA values for tap water samples taken in two regions of Canada. These findings indicate that the concentrations of TTHMs are, at best, moderately associated with HAAs and suggest that researchers should not assume that selection of sites with low spatial variability in TTHMs will produce corresponding low spatial variability exposures to HAAs.

Improvements in exposure assessment for DBPs have been difficult to implement. New methods of exposure assessment such as distribution system modelling, and use of biomarkers in blood, urine or exhaled breath have yet to be validated or applied in large epidemiological studies. Recent studies in the USA have relied on quarterly sampling data collected at multiple sampling sites to meet regulatory requirements. Future studies will likely continue to use these data, due to their relative efficiency. The methods in this paper permit selection of sites with limited spatial variability from state or national DBP datasets to be used for future prospective and retrospective epidemiology studies. Ultimately, simple techniques such as these may improve both the quality of data and our understanding of the true risks associated with exposure to DBPs.

ACKNOWLEDGEMENTS
Funding for this project was provided by the US Environmental Protection Agency through grant OD-5375-NAEX. The authors thank Dr Fred Hauchman and Kenneth Eliston of US EPA for their helpful suggestions during the design of this study. We thank Thomas Keele for helpful advice regarding the ANOVA analyses.

Authors’ affiliations
A F Hinckley, A M Bachand, J R Nuckols, J S Reif, Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, Colorado, USA

Competing interests: none declared

REFERENCES
Identifying public water facilities with low spatial variability of disinfection by-products for epidemiological investigations

A F Hinckley, A M Bachand, J R Nuckols and J S Reif

Occup Environ Med 2005 62: 494-499
doi: 10.1136/oem.2004.017798