Journal of Exposure Analysis and Environmental Epidemiology (2005) 15, 212–216

www.nature.com/jea

A sensitivity analysis of bias in relative risk estimates due to disinfection by-product exposure misclassification

J. MICHAEL WRIGHT[a] AND THOMAS F. BATESON[b]

[a]*Office of Research and Development, US EPA, National Center for Environmental Assessment, 26 W. Martin Luther King Drive, Cincinnati, Ohio 45268, USA*
[b]*Apex Epidemiology, Washington DC, USA*

We conducted a sensitivity analysis of relative risk estimates using local area mean disinfection by-product exposures. We used Monte Carlo simulations to generate data representing 100 towns, each with 100 births ($n = 10{,}000$). Each town was assigned a mean total trihalomethane (TTHM) exposure value (mean = 45, SD = 28) based on a variable number of sampling locations (range 2–10). True maternal TTHM exposure was randomly assigned from a lognormal distribution using that town's true mean value. We compared the effect of a 20 μg/l increase in TTHM exposure on the risk of small-for-gestational age infancy using the true maternal exposure compared to various weighting measures of the town mean exposures. The exposure metrics included: (1) unweighted town mean, (2) town mean weighted by the inverse variance of the town mean, (3) town mean weighted by the inverse standard deviation of the town mean, (4) town mean weighted by 1−(standard deviation of sites per town/mean across all towns), and (5) a randomly selected value from one of the sites within the town of residence. To estimate the magnitude of misclassification bias from using the town mean concentrations, we compared the true exposure odds ratios (1.00, 1.20, 1.50, and 2.00) to the mean exposure odds ratios from the five exposure scenarios. Misclassification bias from the use of unweighted town mean exposures ranged from 19 to 39%, increasing in proportion to the size of the true effect estimates. Weighted town mean TTHM exposures were less biased than the unweighted estimates of maternal exposure, with bias ranging from 0 to 23%. The weighted town mean analyses showed that attenuation of the true effect of DBP exposure was diminished when town mean concentrations with large variability were downweighted. We observed a trade-off between bias and precision in the weighted exposure analyses, with the least biased effects estimates having the widest confidence intervals. Effect attenuation due to intrasystem variability was most evident in absolute and relative terms for larger odds ratios.
Journal of Exposure Analysis and Environmental Epidemiology (2005) **15,** 212–216. doi:10.1038/sj.jea.7500389
Published online 30 June 2004

Keywords: *exposure assessment, misclassification bias, measurement error, disinfection by-products, trihalomethanes, fetal development.*

Introduction

There is a growing body of epidemiologic evidence suggesting a relationship between disinfection by-products (DBPs) and indices of fetal development, such as low birth weight and small-for-gestational age infancy (Nieuwenhuijsen et al., 2000; Graves et al., 2001; Bove et al., 2002). Indirect exposure assessment is a major limitation of previous studies, since maternal exposures are often estimated from aggregate (i.e., town-level) monitoring data. These data may not be reflective of the exposure experience of the study population due to interindividual differences in water usage and unmeasured temporal and spatial variability in DBP formation. The impact of these sources of variability on previous epidemiologic findings is unclear, but can result in bias due to exposure misclassification in studies using town-level data to assign individual-level exposures.

Intrasystem spatial variation in DBP concentrations results from formation and degradation processes occurring over time (Stevens et al., 1989) and space (Chen and Weisel, 1998). Spatial variability in trihalomethane formation has been reported with distribution system levels up to two (Rodriquez and Serodes, 2001) and three (Sohn et al., 2001) times higher than finished water leaving the treatment plant. This is in contrast to nonvolatile compounds such as haloacetic acids (HAAs), where maximum concentrations have been found at sampling locations closest to the point of disinfection (Chen and Weisel, 1998). Reductions in HAAs and other nonvolatile DBPs within the distribution system can result from abiotic and biotic degradation (Hozalski et al., 2001; Krasner et al., 1989; Singer, 1994). These intrasystem differences highlight the limitation of using town average concentrations as a surrogate of maternal exposure to DBPs.

Several reproductive epidemiologic studies of DBPs have used town average concentrations to estimate maternal exposure (Kramer et al., 1992; Bove et al., 1995; Savitz et al., 1995; Gallagher et al., 1998; Waller et al., 1998; Dodds

1. Address all correspondence to: Dr. J. Michael Wright, Office of Research and Development, US EPA, National Center for Environmental Assessment, 26 W. Martin Luther King Drive, Cincinnati, OH 45268, USA. Tel.: +1-513-569-7922. Fax: +1-513-569-7916.
E-mail wright.michael@epa.gov
Received 13 August 2003; accepted 20 April 2004; published online 30 June 2004

et al., 1999; Klotz and Pyrch, 1999; Wright et al., 2003, 2004). Gallagher et al. (1998) used geographic information system mapping and water hydraulic models to limit the impact of spatial variability on exposure misclassification. After restricting the study population to residents of census blocks best represented by trihalomethane monitoring locations, total trihalomethane (TTHM) exposures greater than 60 μg/l (compared to less than or equal to 20 μg/l) were associated with low birth weight among term infants (odds ratio (OR) = 5.9, 95% CI 2.0–19.0) and all infants (OR = 2.1, 95% CI 1.0–4.8). Klotz and Pyrch (1999) addressed spatial variability in a study of neural tube defects and trihalomethanes by collecting residential tap water samples a year after the critical period of exposure. A 60% (OR = 1.6, 95% CI 0.9–2.7) increased risk of neural tube defects for utility average TTHM concentrations greater than or equal to 40 μg/l (compared to less than 5 μg/l) was observed in the overall population, while subjects with residential tap water samples had a two-fold increased risk (OR = 2.0, 95% CI 0.9–4.9). Residential tap water samples collected a year after pregnancy likely reduced the impact of spatial variability, but may have also introduced exposure misclassification if the samples were not reflective of maternal exposures a year earlier. Using a similar sampling approach in a case–control study of stillbirths, King et al. (2004) reported a small average absolute difference (9.8 μg/l) in TTHM concentrations from samples collected a year after pregnancy with distribution system measurements collected during pregnancy ($r = 0.87$).

The aforementioned studies have attempted to limit the impact of temporal and spatial variability in DBP formation through improved exposure assessment, but very few have quantified the amount of misclassification that occurs from using town average exposures. In a re-analysis of a prospective cohort study of 4212 women (Waller et al., 1998), Waller et al. (2001) compared relative risk estimates of spontaneous abortions for a variety of DBP exposure metrics. Unweighted utility average exposure estimates were compared to weighted utility averages and population subsets restricted by level of intradistribution system variability. Overall, they reported consistent reductions in effect estimates due to misclassification of exposures based on unweighted utility averages. We conducted a sensitivity analysis to further examine the adequacy of using arithmetic town mean concentrations to estimate individual exposure. Monte Carlo simulations were used to quantify the amount of bias due to nondifferential exposure misclassification from the use of weighted and unweighted exposure metrics.

Methods

Statistical Analysis Systems (SAS) software version 8.2 was used to simulate data representing disinfection by-product concentrations in 100 public drinking water systems (SAS, 2000). Each town was assigned a base TTHM exposure value (μ_t) from a lognormal distribution with an arithmetic mean of 45 μg/l and standard deviation of 28 μg/l, since arithmetic distributions are commonly reported in the DBP literature.

$$\log(\mu_t) \sim N(3.643, 0.572W)$$

where

$$W \sim N(0,1) \tag{1}$$

A random number of samples were simulated for each town TTHM level independent of the town base exposure value. The sampling distribution from a previous study was used to assign the number of sampling locations per town (Wright et al., 2004). In all, 26 of the towns had two sampling sites, 47 of the towns had four sites, five towns had five sites, 16 towns had eight sites, and two towns each had six, eight and 10 sites, respectively. Based on the assigned number of sampling sites per town, we independently sampled TTHM exposure values (x_{ti}) using the same town base distribution each time — allowing each town to have a different standard deviation.

$$\log(x_{ti}) \sim N\left(\left(\log(\mu_t) - \frac{(0.5+0.1X)^2}{2}\right), (0.5+0.1X)Y\right)$$

where

$$\begin{aligned} X &\sim N(0,1) \\ Y &\sim N(0,1) \end{aligned} \tag{2}$$

The sampled exposure values were used to calculate the average town exposure concentrations. Each of the 100 simulated mothers per town was independently assigned a true exposure value (x_m) randomly chosen from the TTHM distribution for her town.

$$\log(x_m) \sim N\left(\left(\log(\mu_t) - \frac{(0.5+0.1Z)^2}{2}\right), (0.5+0.1Z)A\right)$$

where

$$\begin{aligned} Z &\sim N(0,1) \\ A &\sim N(0,1) \end{aligned} \tag{3}$$

The probability of each mother giving birth to an infant classified as small for gestational age was 5% in the absence of exposure. We simulated four possible odds ratios (ORs = 1.00, 1.20, 1.50, 2.00) per 20 μg/l TTHM.

Logistic regression analyses were conducted using Proc Logistic in SAS. We compared the true exposure odds ratios to various exposure scenarios including weighted and unweighted town mean exposures. The exposure weights were functions of the standard deviation of exposure across the samples for each town and included the inverse of the arithmetic standard deviation of the town mean, the inverse of the arithmetic variance of the town mean, and a weighting scheme used by Waller et al. (2001) [1−(arithmetic standard deviation of sites per town/mean across all towns)]. We

assessed the effect of gross misclassification by evaluating the odds ratios associated with a randomly chosen TTHM value within a mother's town of residence. We simulated 1000 iterations of each exposure scenario and calculated the mean of the logistic regression coefficients, since the distribution of log odds ratios is approximately normally distributed as opposed to the highly skewed distribution of untransformed odds ratios. We present the exponentiated mean log odds ratios with 95% confidence intervals. We estimated the magnitude of bias by the absolute difference of the mean log odds ratio for each exposure weighting schemes compared to the log odds ratio for the true maternal exposure to standardize the comparisons of bias across different effect sizes. Pearson's correlation coefficients were calculated to determine the relationship between true maternal exposure and town mean concentrations.

Results

Table 1 shows the distribution of town and individual mean DBP exposures for a simulated population of 10,000. The mean of the town base TTHM exposure was 45.1 μg/l with a standard deviation of 27.9 μg/l. The mean number of DBP sampling sites per town was 4.4 with a standard deviation of 2.1. The mean of the town average TTHM based on selected samples was 45.1 μg/l with a standard deviation of 32.2 μg/l. The mean of the true maternal exposures was 45.1 μg/l with a standard deviation of 41.1 μg/l. The correlation between the true maternal exposure and the town mean concentration was 0.62.

The simulated mean odds ratios based on the actual assigned maternal exposures were equal to the true odds ratios across the different exposure metrics (Table 2). The simulations of the null (OR = 1.00) yielded nearly unbiased effect estimates with ORs that ranged from 0.98 to 1.00. Compared to true maternal exposures measured without error, the random site exposure was the most biased exposure metric (48–82%) increasing in proportion to the size of the effect estimates (Table 3). The bias from the unweighted town average exposure also increased in proportion to size of the effect estimates. For a true relative risk of 1.20, the odds ratio based on an unweighted town mean was attenuated by 19% (OR = 1.16). For true relative risks of 1.50 and 2.00, effect estimate attenuation for unweighted town average exposures was 26 and 39%, respectively.

Overall, the odds ratios for the weighted town average metrics were less biased but had wider confidence intervals compared to the unweighted averages (Table 2). For a true relative risk of 1.20, the inverse standard deviation weight was the most biased weighted exposure metric (OR = 1.16). The town mean exposure weighted by one minus the ratio of the standard deviation of the town mean divided by the overall mean across towns was unbiased for a relative risk of 1.20. Negligible bias was detected for the inverse of the variance of the town mean (OR = 1.50), while 12% bias

Table 1. Total trihalomethane exposure characteristics for a simulated population of 10,000 mothers and their town of residence.

TTHM concentration	Arithmetic mean	Arithmetic standard deviation	Arithmetic 5th %	Arithmetic 95th %	Geometric mean	Geometric standard deviation
Town base	45.1	27.9	15.0	98.0	38.2	1.8
Town mean[a]	45.1	32.2	12.8	105.3	36.8	1.9
Maternal exposure	45.1	41.1	9.5	118.3	33.6	2.2

[a]Based on arithmetic mean of 4.4 sampling sites per town, with a standard deviation of 2.1.

Table 2. Effect of simulated total trihalomethane exposure on the risk of small for gestational age[a] infancy in the absence and presence of nondifferential exposure misclassification.

Exposure classification	OR = 1.00[b]	OR = 1.20[b]	OR = 1.50[b]	OR = 2.00[b]
Mother's true exposure	1.00 (0.96, 1.04)	1.20 (1.17, 1.23)	1.50 (1.46, 1.54)	2.00 (1.93, 2.07)
Random site in town	1.00 (0.96, 1.04)	1.10 (1.06, 1.13)	1.20 (1.16, 1.24)	1.30 (1.25, 1.36)
Town mean — unweighted	1.00 (0.94, 1.06)	1.16 (1.09, 1.24)	1.35 (1.23, 1.48)	1.53 (1.36, 1.71)
Town mean — weight 1[c]	1.00 (0.90, 1.11)	1.20 (1.10, 1.31)	1.46 (1.30, 1.64)	1.73 (1.52, 1.97)
Town mean — weight 2[d]	0.99 (0.80, 1.23)	1.16 (1.09, 1.24)	1.43 (1.22, 1.68)	1.70 (1.39, 2.08)
Town mean — weight 3[e]	0.98 (0.49, 1.94)	1.19 (0.70, 2.02)	1.50 (0.88, 2.57)	1.86 (0.66, 5.24)

[a]Probability of small-for-gestational age infancy in unexposed population = 0.05.
[b]Mean regression coefficients exponentiated to the odds ratio scale.
[c]Weight 1: 1−town standard deviation/overall mean across towns; if weight < 0 then weight = 0.
[d]Weight 2: 1/town standard deviation.
[e]Weight 3: 1/town variance.

Table 3. The percent bias[a] in the means odds ratios for small for gestational age[b] infancy due to nondifferential misclassification of total trihalomethane exposures.

Exposure classification	OR = 1.00 (%)	OR = 1.20 (%)	OR = 1.50 (%)	OR = 2.00 (%)
Mother's true exposure	0	0	0	0
Random site in town	0	−48	−55	−82
Town mean — unweighted	0	−19	−26	−39
Town mean — weight 1[c]	0	0	−7	−21
Town mean — weight 2[d]	−1	−19	−12	−23
Town mean — weight 3[e]	−2	−5	0	−10

[a]Percent bias is calculated from the difference between the true and observed odd ratios on the log scale.
[b]Probability of small-for-gestational age infancy in unexposed population = 0.05.
[c]Weight 1: 1−town standard deviation/overall mean across towns; if weight < 0 then weight = 0.
[d]Weight 2: 1/town standard deviation.
[e]Weight 3: 1/town variance.

(OR = 1.43) was found for the inverse standard deviation metric. For a true odds ratio of 2.00, the inverse variance weight was the least biased weighted exposure estimate (OR = 1.86). Across the four selected relative risk values, the inverse variance weights resulted in the least amount of bias (0–10%). Overall, the inverse standard deviation weights were the most biased weighted exposure metric (12–23%).

Discussion

We used Monte Carlo simulations to examine the degree of nondifferential misclassification bias resulting from the use of town average TTHM concentrations to estimate individual-level exposures. The simulations based on the true maternal exposures were equal to the true odds ratios (1.00, 1.20, 1.50, 2.00) indicating that the simulations were unbiased. For a null effect of DBP exposure (OR = 1.00), we observed negligible bias (0–2%) across the different exposure metrics. The unweighted sample mean estimates of personal exposure produced results that were attenuated reflecting the effect of nondifferential measurement error in assigning exposure to individuals based on the average of several samples taken in the area. The amount of effect attenuation due to nondifferential misclassification increased with the size of the effect estimates. Since town mean exposures with greater variability across samples were less likely to accurately estimate a particular mother's exposure, we constructed exposure weights that were functions of the standard deviations of TTHM exposures across the samples for each town. Overall, the weighted town mean analyses produced less misclassification bias at the cost of greater variability in the effect estimates compared to the unweighted results.

We have shown that weighted analyses using the town mean DBP exposure as a proxy for maternal exposure produces less nondifferential misclassification and less attenuation of the true effect of exposure. It is unclear from these results as to which of the weighting schemes will prove most generally useful, since the least biased results were not consistent across the range of relative risks that were examined. The inverse variance weights were less biased but not as efficient as the 1−(standard deviation of sites per town/overall mean across towns) weight used by Waller et al. (2001). Compared to groundwater users, they reported a larger relative risk for weighted (OR = 1.5 *versus* 1.3 for unweighted) TTHM exposures greater than 80 μg/l. We found bias similar in magnitude for an odds ratio of 1.50 (weighted OR = 1.45 *versus* unweighted OR = 1.36), despite differences in exposure scaling. These findings suggest that constraints on the size of the weights are useful, but should be further explored with respect to the individual weighting scenarios. Our results were based on moderate to high spatial variability in DBP exposure data. Town-level surrogate measures of DBP exposure derived from water systems with less spatial variability would be expected to have less measurement error than we observed.

We examined only one source of measurement error that can influence the relationship between true exposure and surrogate measures of DBPs. Other potential sources of measurement error include temporal differences in DBP formation and interindividual variability in water usage. Our simulations were based on a correlation of 0.62 between town mean and true maternal TTHM exposure. Although this relationship is typically unknown in observational epidemiologic studies, it is comparable to that observed in previous DBP studies (Whitaker et al., 2003b; King et al., 2004). Whitaker et al. (2003b) and King et al. (2004) used individual level water usage data to estimate DBP exposure during pregnancy and reported exposure misclassification rates in excess of 40% for town average trihalomethane classification. The impact of this exposure misclassification on previously reported association is unknown, but additional research is needed to better understand the relative contribution of various sources of measurement error in epidemiologic studies.

Acknowledgements

We thank Dr. Kevin Brand for his review of the manuscript and providing constructive comments and suggestions.

References

Bove F.J., Fulcomer M.C., and Klotz J.B., et al. Public drinking water contamination and outcomes. *Am J Epidemiol* 1995: 141: 850–862.

Bove F., Shim Y., and Zeitz P. Drinking water contaminants and adverse pregnancy outcomes: a review. *Environ Health Perspect* 2002: 110(Suppl 1): 61–74.

Chen W.J., and Weisel C.P. Halogenated DBP concentrations in a distribution system. *J Am Water Works Assoc* 1998: 90(4): 151–163.

Dodds L., King W., and Woolcott C., et al. Trihalomethanes in public water supplies and adverse birth outcomes. *Epidemiology* 1999: 10(3): 233–237.

Gallagher M.D., Nuckols J.R., and Stallones L., et al. Exposure to trihalomethanes and adverse pregnancy outcomes. *Epidemiology* 1998: 9(5): 484–489.

Graves C.G., Matanoski G.M., and Tardiff R.G. Weight of evidence for an association between adverse reproductive and developmental effects and exposure to disinfection by products: a critical review. *Regul Toxicol Pharmacol* 2001: 34: 103–124.

Hozalski R.M., Zhang L., and Arnold W.A. Reduction of haloacetic acids by FeO: implications for treatment and fate. *Environ Sci Technol* 2001: 35(11): 2258–2263.

King W.D., Dodds L., Armson B.A., Allen A.C., Fell D.B., and Nimrod C Exposure assessment in epidemiologic studies of adverse pregnancy outcomes and disinfection byproducts. *J Expos Anal Environ Epidemiol* 2004: 14(6): 466–472.

Klotz J.B., and Pyrch L.A. Neural tube defects and drinking water disinfection by-products. *Epidemiology* 1999: 10(4): 383–390.

Kramer M., Lynch C., Isacson P., and Hanson J.W. The association of waterborne chloroform with intrauterine growth retardation. *Epidemiology* 1992: 3(5): 407–413.

Krasner S.W., McGuire M.J., Jacangelo J.G., Patania N.L., Reagan K.M., and Aieta E.M. The occurrence of disinfection by-products in US drinking water. *J Am Water Works Assoc* 1989: 81: 41–53.

Nieuwenhuijsen M.J., Toledano M.R., and Eaton N.E., et al. Chlorination disinfection by-products in water and their association with adverse reproductive outcomes: a review. *Occup Environ Med* 2000: 57: 73–85.

Rodriquez M.J., and Serodes J.-B. Spatial and temporal evolution of trihalomethanes in three water distribution systems. *Water Res* 2001: 35(6): 1572–1586.

SAS. Version 8.2. SAS Institute, Inc., Cary, NC, 2000.

Savitz D., Andrews K., and Pastore L. Drinking water and pregnancy outcomes in central North Carolina: source, amount, and trihalomethane levels. *Environ Health Perspect* 1995: 103(6): 592–596.

Singer P. Control of disinfection by-products in drinking water. *J Environ Eng* 1994: 120(4): 727–744.

Sohn J., Gatel D., and Amy G. Monitoring and modeling of disinfection by-products (DBPs). *Environ Monit Assess* 2001: 70(1–2): 211–222.

Stevens A.A., Moore L.A., and Miltner R.J. Formation and control of non-trihalomethane disinfection by-products. *J Am Water Works Assoc* 1989: 81(8): 54–60.

Waller K., Swan S.H., and DeLorenze G., et al. Trihalomethanes in drinking water and spontaneous abortion. *Epidemiology* 1998: 9(2): 134–140.

Waller K., Swan S.H., Windham G.C., and Fenster L. Influence of exposure assessment methods on risk estimates in an epidemiologic study of total trihalomethane exposure and spontaneous abortion. *J Expos Anal Environ Epidemiol* 2001: 11(6): 522–531.

Whitaker H., Nieuwenhuijsen M.J., Best N., Fawell J., Gowers A., and Elliott P. Description of trihalomethane levels in three UK water suppliers. *J Expos Anal Environ Epidemiol* 2003a: 13: 17–23.

Whitaker H.J., Nieuwenhuijsen M.J., and Best N.G. The relationship between water concentrations and individual uptake of chloroform: a simulation study. *Environ Health Perspect* 2003b: 111(5): 688–694.

Wright J.M., Schwartz J., and Dockery D.W. Effect of trihalomethane exposure on fetal development. *Occup Environ Med* 2003: 60(3): 173–180.

Wright J.M., Schwartz J., and Dockery D.W. The effect of disinfection by-products and mutagenic activity on birth weight and gestation duration. *Environ Health Perspect* 2004: 112: 920–925.