Exposure to Disinfection By-products, Fetal Growth, and Prematurity

A Systematic Review and Meta-analysis

James Grellier,1 James Bennett,1 Evridiki Patelarou,2 Rachel B. Smith,1 Mireille B. Toledano,1 Lesley Rushton,1 David J. Briggs,1 and Mark J. Nieuwenhuijsen1

Background: Exposure to total trihalomethanes in drinking water has been associated with several adverse birth outcomes relating to fetal growth and prematurity.

Methods: We carried out a systematic review and meta-analysis of epidemiologic studies featuring original peer-reviewed data on the association of total trihalomethane exposure and health outcomes related to fetal growth and prematurity.

Results: A comprehensive literature search yielded 37 studies, 15 of which were selected for the extraction of relative risks relating adverse birth outcomes to trihalomethane exposure. Sufficient data were available for meta-analyses to be carried out for 4 adverse birth outcomes: low birth weight (LBW), term low birth weight (term LBW), preterm delivery, and small for gestational age (SGA) (including intra uterine growth retardation). We found little or no evidence for associations between trihalomethane exposure and LBW (odds ratio per 10 μg total trihalomethane/L = 1.00 [95% confidence interval = 0.97–1.03]), term LBW (1.03 [0.93–1.15]), or preterm delivery (0.99 [0.98–1.00]), but some evidence for SGA (1.01 [1.00–1.02]).

Conclusions: There was little or no evidence for associations between total trihalomethane concentration and adverse birth outcomes to trihalomethane exposure. However, the possible exception of SGA. We discuss these findings and the uncertainties—relating particularly to exposure—that may have affected them.

(Epidemiology 2010;21: 300–313)

Supplies of drinking water were first disinfected using chlorine at the start of the 20th century, primarily as a means of reducing mortality and morbidity associated with waterborne infectious disease. Chlorination was widespread in cities across the developed world by the 1920s, and the method remains a relatively inexpensive and effective means of disinfecting drinking water.

Chlorine reacts with organic compounds such as fulvic and humic acids in the source water to produce disinfection by-products. First identified in disinfected drinking water in the 1970s, trihalomethanes are generally the most abundant of the disinfection by-products, but many other chemicals may also be present. Over 600 disinfection by-products have been reported, and their presence and relative concentration vary seasonally and geographically, due to differences in the chemical character and physical properties of the source water and in the treatment and distribution systems.

PREGNANCY OUTCOMES ASSOCIATED WITH DISINFECTION BY-PRODUCTS

Over the last 2 decades, human studies have assessed the association of disinfection by-products with various outcomes related to fetal growth and prematurity. These outcomes have included low birth weight (LBW), term LBW, very LBW, small for gestational age (SGA), intra uterine growth retardation (IUGR), preterm delivery (PTD), and very PTD, and fetal death (miscarriage, spontaneous abortion, and stillbirth). Six systematic reviews of the epidemiologic evidence for reproductive and developmental effects of exposure to disinfection by-products have been published: 2 narrative reviews, 2 comprehensive weight-of-evidence reviews, and 2 meta-analyses of chlorination and birth defects.

The results from studies of fetal growth and prematurity are mixed, varying in both direction and magnitude of effect. Existing reviews present a useful synthesis and critique of the available literature, but they have not attempted to produce summary measures of effect. The weight of evidence is suggestive of small, positive associations between trihalomethane con-

Submitted 24 March 2009; accepted 2 October 2009.

From the 1Department of Epidemiology and Public Health, Imperial College London, London, United Kingdom; 2Department of Social Medicine, University of Crete, Heraklion, Greece; and 3Center for Research in Environmental Epidemiology (CREAL), IMIM, CIBERESP, Barcelona, Spain.

Supported by the Integrated Assessment of Health Risks of Environmental Stressors in Europe (INTARESE) project, which is funded under the European Commission under the Sixth Framework Programme (2002–2006) grant GOCE-CT-2005–018385.
centrations in drinking water and some adverse birth outcomes related to fetal growth restriction (term LBW, SGA, IUGR), although evidence is not conclusive.

**OBJECTIVES**

Our objectives were to systematically review existing epidemiologic evidence and to carry out a meta-analysis of these data, to produce best-estimate exposure-response slopes of total trihalomethane exposure and adverse birth outcomes relating to fetal growth and prematurity. Ultimately, the objective is that these quantitative results would be suitable for the estimation of burden of disease using routine monitoring data on drinking water quality.

**METHODS**

**Search Methods**

We carried out a systematic review of the existing literature on trihalomethanes and adverse birth outcomes related to fetal growth and prematurity, using the following review question: “Given existing epidemiologic evidence, what is the exposure-response relationship between exposure of pregnant women to trihalomethanes in drinking water and the risk of various adverse birth outcomes related to fetal growth and prematurity?” We drew up a review protocol for the meta-analysis in advance, broadly following guidelines laid out in Egger et al. We conducted and reported on our search methods and results following standards outlined in the QUOROM statement and the MOOSE Guidelines.

We carried out a systematic, comprehensive bibliographic search using the US National Library of Medicine Medline database for the years 1980–2007, using the PubMed interface. Full details of the search are provided in the Appendix (http://links.lww.com/EDE/A379). We checked the list of studies identified thus far for completeness against studies referenced in existing reviews.

We defined a priori eligibility criteria for studies. We retained only studies that were reported in peer-reviewed journals or published by a reputable independent body such as the World Health Organization (WHO) or the US Environmental Protection Administration (EPA). Studies were included if they were published in English, were epidemiologic studies, used maternal residence for exposure estimation, and presented odds ratio (OR), relative risk (RR), or other comparable measure of effect for at least one adverse birth outcome associated with exposure to disinfection by-products. Studies not meeting these criteria were excluded and the specific reasons for their exclusion were noted.

The list of included studies was narrowed down further on the basis of the exposure assessment methods used: only those that characterized disinfection by-product exposure using at least 3 exposure categories were included. We excluded studies that dichotomized exposure, primarily because such a measure offers only a crude index of exposure; early epidemiologic studies that classified exposure according to water treatment methods have been criticized for their failure to capture a more detailed picture of exposure to disinfection by-products. Second, we considered it impractical to combine relative risks from studies with binary exposure characterization and continuous/categorical exposure. Third, because trihalomethane concentrations from routine monitoring data on drinking water are generally available an estimate of a continuous odds ratio slope was considered to provide health impact assessments with the most useful information. Fourth, in developed countries, the reporting of drinking water treatment type is generally not mandatory, whereas reporting of trihalomethane concentrations is a legal obligation. Lastly, mixing of drinking water that has undergone different treatments is common practice in many countries.

The following data were extracted systematically from each included study by 2 researchers using a standard data collection form: study design, exposure characterization, definitions of exposure categories, and measures of effect and confidence intervals for each exposure category (Table 1). We checked the 2 datasets against one another and addressed any inconsistencies. The final set of studies was reviewed qualitatively to assess between-study heterogeneities.

**Statistical Methods**

In each of the studies reviewed, exposure had been presented in terms of concentration of total trihalomethanes or, in one case, trichloromethane (chloroform), using one of 2 measures: either parts per million (ppm) or micrograms per liter (μg/L). We considered concentrations in ppm as equivalent to μg/L, because at the low concentrations present in drinking water these are virtually equivalent. To include the one study reporting only chloroform concentrations, we multiplied reported exposure categories by a factor of 1.33, on the assumption that chloroform might make up 75% of the total trihalomethane mixture and that concentrations of chloroform and total trihalomethanes in drinking water are highly correlated.

The majority of studies presented their results as ORs with 95% confidence intervals (CIs), although some used hazard ratio (HR) or relative risk, or risk ratio (both RR). For the purposes of this analysis, these measures were assumed to be equivalent to odds ratios. One study presented results at nested levels of confidence other than 95%. In this instance, we calculated the standard error on the OR from the 99% CIs provided using the formula:

\[
\text{standard error} = \frac{\ln(\text{upper 99% CI)} - \ln(\text{lower 99% CI})}{(2 \times 2.575)}
\]

Upper and lower 95% confidence intervals were then calculated as follows:

- upper 95% CI = exp(ln[OR]) + (1.96 × standard error)
- lower 95% CI = exp(ln[OR]) − (1.96 × standard error)
| Author(s) and Location of Study | Study Period | Study Design | Live Birth Outcomes Investigated | Live Birth Outcomes Relevant to Meta-analysis | Outcome Measure Used | Data for TTHM and/or Other Chlorination By-products | Exposure | Location | Timing |
|---------------------------------|--------------|--------------|----------------------------------|---------------------------------------------|----------------------|-----------------------------------------------|----------|----------|-------|
| Bove et al, 21 New Jersey | 1985-1988 | Retrospective (cross-sectional) cohort | BW | PTD (n = 7,167) | SGA | Teal LBW (not reported in detail) | Monthly estimates of TTHM from state department monitoring data. | Mother’s town of residence at the time of birth was assumed to be the town of residence for the entire pregnancy. | Monitoring data assigned to each gestational month and exposure averaged over entire pregnancy. |
| Gallagher et al, 13 Colorado | 1990-1993 | Retrospective cohort | LBW (n = 72) | Term LBW (n = 259) | PTD (n = 68) | Total n = 1,244 births for which THM data available | Routine monitoring TTHM values from 1 yr prior to study. | Individual exposures assigned according to maternal residence within a municipality by way of hydraulic modeling of TTHM values in the distribution network. | Exposure score estimated for third trimester. |
| Dodds et al, 12 Nova Scotia, Canada | 1988-1995 | Retrospective cohort | SGA (n = 4,675) | LBW (n = 239) | SGA | Term LBW | Relative risk | Routine monitoring TTHM values and linear regression modeling used to estimate TTHM for those time periods missing sampling data. | Mother’s residence at time of delivery was linked to the geographic area served by each water company, providing individual measures of exposure. | TTHM exposure calculated with a regression model for the third trimester. |
| Wright et al, 24 Massachusetts | 1990 | Retrospective cohort | Term LBW (n = 1,325) | SGA (n = 5,310) | PTD (n = 3,173) | Total n = 56,513 singleton infants | Maternal THM exposure for the 3rd trimester estimated from the quarterly average TTHM concentration. | Maternal town of residence. | Exposures estimated for all 3 trimesters and a total pregnancy average. Trimester specific and pregnancy average exposures were assigned based on the month of birth. |
| Wright et al, 35 Massachusetts | 1995-1998 | Retrospective cohort | BW | SGA | PTD | Total n = 196,000 residents with singleton infants | Maternal zip code at birth and infant month of birth used to assign third-trimester specific exposure data. | Maternal residence (zip code at birth) (extrapolated from estimated exposure at month of birth) A also averaged over whole pregnancy. | Exposure estimated for third trimester. |
| Hinckley et al, 22 Arizona | 1998-2002 | Retrospective cohort | IUGR (n = 43,461) | Term LBW (n = 1,010) | PTD (n = 4,080) | Total n = 48,119 live births and fetal deaths | Monthly TTHM exposures estimated from quarterly and monthly samples. Spline regression techniques used to estimate exposure for specific periods. | Maternal residence (zip code at birth) (assumed to be the same as third trimester). | Exposure estimated for third trimester (extrapolated from estimated exposure at month of birth) and other specific time windows (for IUGR and LBW) (only preterm and very preterm exposure evaluated for specific time windows). |
| Porter et al, 11 Maryland | 1998-2002 | Retrospective cohort | IUGR | Total n = 15,315 births | OR | Monthly TTHM concentrations for 4 sampling points in study area | Maternal residence (zip code at birth) | Exposure estimated for both entire pregnancy and each trimester. | (Continued) |
| Author(s) and Location of Study | Study Period | Study Design | Live Birth Outcomes Investigated | Live Birth Outcomes Relevant to Meta-analysis | Outcome Measure Used | Data for TTHM and/or Other Chlorination By-products | Exposure | Location | Timing |
|--------------------------------|--------------|-------------|----------------------------------|-----------------------------------------------|----------------------|-------------------------------------------------|----------|----------|--------|
| Toledano et al, United Kingdom | 1992–1998    | Retrospective cohort | BW LBW Very LBW SB | LBW Very LBW | OR | Weighted average of modeled quarterly TTHM estimates for last 93 d before birth. | Exposure modelled for each water zone using maternal residence (by postcode) at time of birth. |                |          |        |
| Yang et al, Taiwan            | 2000–2002    | Retrospective cohort | Term LBW (n = 2766) PTD (n = 2818) SGA (n = 8938) Selton births | Term LBW SGA PTD | OR | Exposure established as an average of TTHM monitoring data over a 2-year period for each municipality. | Municipality of residence at birth (assuming continual residence at that location throughout pregnancy). Exposure estimated for pregnancy average. |                |          |        |
| Hoffman et al, USA            | 2000–2004    | Community-based prospective cohort study | SGA (n = 113) Total n = 1958 live births (restricted to those born at 37 wk or later) | SGA | Risk ratio | Dedicated sampling at representative locations in the distribution system (for TTHM, all 4 individual THM species, 9 HAAs, total organic halide (TOH)). | Two exposure metrics used: (1) estimated residential concentration and (2) estimated personal DBP exposure. Individual trimesters. |                |          |        |
| Hoffman et al, USA            | 2000–2004    | Community-based prospective cohort study | PTD (n = 185) Total n = 2039 births | PTD | Risk ratio | Two exposure metrics used: (1) estimated residential concentration and (2) personal DBP exposure, estimated as uptake through showering and bathing for TTHMs and by intake through ingestion for HAAs. | Reported on second trimester. |                |          |        |
| Kramer et al, Iowa            | 1989–1990    | Population-based case-control | IUGR (n = 187) LBW (n = 159) PTD (n = 342) | IUGR LBW PTD | OR | THM levels from 1987 water survey. | Assigned to maternal residence in a given municipality at birth. Exposure estimated over entire pregnancy. |                |          |        |
| Savitz et al, North Carolina  | 1988–1989, 1988–1991 | Population-based case-control | PTD (n = 586) LBW (n = 464) MC (n = 418) | LBW MC PTD | OR | Quarterly average THM values recorded by appropriate water supplier. | Mother’s addresses used to assign them to one of 5 public water supplies. Dates of pregnancy used to assign reported nearest average quarterly THM value to each mother. Assignment time varied between outcomes: MC (4th week), PTD/LBW (28th week). |                |          |        |
| Lewis et al, Massachusetts     | 1999–2001    | Population-based case-control | Term LBW total n = 36,529 births | Term LBW | OR | Weekly TTHM monitoring data from 4 sampling sites. | Each birth assigned to based on maternal residence at birth (from birth certificate). TTHM exposure estimate calculated for each gestational period. |                |          |        |
| Lewis et al, Massachusetts     | 1999–2001    | Population-based case-control | PTD (n = 2813) Total n = 37,498 singleton births | PTD | HR | Weekly TTHM monitoring data from 4 sampling sites. | Each birth assigned to based on maternal residence at birth (from birth certificate). TTHM exposure estimate calculated for each gestational period. (results reported for lot, 2nd trimester, and 4-wk risk sets, and pregnancy average). |                |          |        |

TTHM indicates total trihalomethane; THM, trihalomethane; MC, miscarriage.
The majority of the studies reported measures of effect adjusted for confounders. Adjustment had been carried out for a range of covariates that varied across the studies, but the majority adjusted for the same important factors (maternal age, parity, smoking, and social deprivation). Several studies did not provide unadjusted results and so adjusted results were used in the meta-analysis. We used unadjusted results in the one case where only unadjusted results were reported. Details regarding individual studies are presented in Table 2.

We sought to minimize between-study heterogeneity relating to exposure assessment methods. Thus, only those studies characterizing exposure based on maternal residence were included in this analysis. We subsequently grouped studies according to the exposure agent measured, the type of measure used, and the timing of exposure that was assumed. The timing of exposure in each study was either categorized by trimester or summarized for the whole pregnancy. The number of exposure categories used in studies varied from 3 to 6. Given the variation in the exposure assessment among studies, we carried out a 2-stage subset analysis to investigate differences in exposure agent and exposure timing for each health outcome. The analysis was divided on the basis of including the study that used chloroform as the exposure agent.

For each of these 2 subsets, analysis was further divided according to exposure timing. The first subset included studies that reported measures of effect associated solely with exposure in the third trimester, because most fetal growth occurs in this period; the second included only those reporting on entire pregnancy exposure; for completeness—and because exposure in different periods are likely correlated—the third subset included all studies regardless of exposure timing (where both third trimester and entire pregnancy exposure were reported in the same study, measures of effect for third trimester were used). We carried out meta-analyses only for those subsets including at least 4 studies.

It was not practical to quantitatively explore other heterogeneities among the studies for a number of reasons: studies were relatively similar in overall design; differences among studies were not consistently presented; and, where meta-analysis might have been stratified on the basis of between-study variability (overall study design, variables adjusted for, geographical location of study, etc.), the number of studies in such subgroups was too few for the application of meta-analytical methods.

Techniques for pooling correlated estimates to compute regression slopes across exposure categories in individual studies have been described previously. All studies provided measures of effect for several exposure categories, although cut-off points of these categories differed among studies (Table 2). Meta-analysis was carried out with the R software package using scripts adapted from those developed by Key et al.

For each study, we fitted a weighted least-squares regression of ln(OR) against exposure, the weight being inversely proportional to the variance on ln(OR) at each exposure category midpoint. If there was no upper limit to the topmost exposure category, a midpoint was derived using half the width of the preceding category.

The use of various dose-response models to obtain study-specific slope estimates has been explored previously, and slope estimates (and standard errors) were observed to be higher when using dose, as compared with ln(dose). Using Bayes information criterion to assess the fit of each dose-response model, it has been demonstrated that neither dose nor ln(dose) in a linear model was more advantageous. Therefore, we carried out a regression of ln(OR) against exposure. In the regression, we assumed that exposure to zero disinfection by-products from water was unlikely, because exposure to volatile disinfection by-products such as trihalomethanes can occur in the domestic environment through several routes (ingestion, inhalation, dermal absorption) and through a variety of pathways (drinking, eating, cooking, washing); therefore the intercepts of the regression slopes were not constrained to go through the origin. The reference categories used in each study differed (Table 2), which further supported this decision.

In addition to the qualitative investigation of heterogeneity between the studies, as described above, Cochran Q-statistic was used to test for between-study heterogeneity. Regression was carried out using both fixed effects and random effects models, and the results compared. The overall choice of a random effects model was informed by the findings of these analyses. Regression slopes of exposure-response derived from individual studies were plotted, together with the summary slopes produced from the meta-analysis (Fig. 1) and forest plots (Fig. 2).

To investigate the role of publication bias and other biases in the meta-analysis, we produced funnel plots (eFigure 2, http://links.lww.com/EDE/A379) for visual inspection of the symmetry of the data, as well as carrying out the Egger regression test.

We investigated the relative influence of individual studies on summary measures of effect using a leave-one-out sensitivity analysis for every subset analysis. Differences between the magnitude and direction of summary measures of effect for each study left out were investigated.

We calculated the risk of each pregnancy outcome for third trimester exposure to total trihalomethane at levels currently prescribed as guidelines in the United States and the European Union (80 μg/L and 100 μg/L, respectively).

RESULTS

Results of Search, Data Extraction, and Study Evaluation

Figure 3 shows the numbers of studies identified and selected/excluded in each phase of the search. No additional studies were identified by means of searching in databases other
| Study | Definition of Outcome | Exposure Agent | Measure of Effect Used in Meta-analysis | Exposure Timing Used in Meta-analysis | Exposure Categories | Measures of Effect (95% Confidence Intervals) |
|-------|----------------------|----------------|---------------------------------------|--------------------------------------|---------------------|---------------------------------------------|
| Kramer et al. | <2500 g | Chloroform OR (adjusted) | Entire pregnancy Nondetect | 1.00 | 1.0 (0.7–1.6) | — |
| Savitz et al. | <2500 g | TTHM OR (adjusted) | Third trimester 40.8–63.3 ppb | 1.00 | 1.1 (0.7–1.6) | 1.3 (0.8–2.2) |
| Gallagher et al. | ≤5 pounds, 8 ounces (≤2495 g) | TTHM OR (adjusted) | Third trimester | 1.00 | 1.5 (1.0–2.3) | 1.3 (0.8–2.1) |
| Dodds et al. | <2500 g | TTHM Relative risk (adjusted) | Third trimester | 1.00 | 1.0 (0.6–1.1) | 0.8 (0.3–1.7) |
| Toledano et al. | <2500 g | TTHM OR (adjusted) | Third trimester | 1.00 | 1.07 (0.8–1.4) | 1.11 (0.9–1.3) |
| Gallagher et al. | ≥37 wk of gestation and ≤5 pounds, 8 ounces (≤2495 g) | TTHM OR (adjusted) | Third trimester | 1.00 | 1.05 (0.8–1.3) | 1.2 (0.9–1.7) |
| Hanckey et al. | ≤37 completed weeks of gestation and weighing ≤500 g | TTHM OR (adjusted) | Third trimester | 1.00 | 1.05 (0.8–1.3) | 1.2 (0.9–1.7) |
| Wright et al. | <2500 g among term births | TTHM OR (adjusted) | Third trimester, entire pregnancy average | 1.00 | 1.08 (0.88–1.3) | 1.09 (0.8–1.3) |
| Lewis et al. | Birth weight <2500 g and born after 36 wk of gestation | TTHM OR (adjusted) | Third trimester, entire pregnancy average | 1.00 | 0.97 (0.9–1.1) | 0.95 (0.8–1.1) |
| Yang et al. | 437 gestational weeks and <2500 g | TTHM OR (adjusted) | Entire pregnancy average | 1.00 | 0.99 (0.9–1.0) | 1.00 (0.9–1.0) |
| Kramer et al. | <37 wk of gestation | Chloroform OR (adjusted) | Entire pregnancy Nondetect | 1.00 | 1.0 (0.8–1.3) | 1.0 (0.8–1.3) |
| Savitz et al. | <37 wk completed gestation | TTHM OR (adjusted) | Third trimester 40.8–63.3 ppb | 1.00 | 1.1 (0.8–1.4) | 1.1 (0.8–1.4) |
| Gallagher et al. | <37 wk completed gestation | TTHM OR (adjusted) | Third trimester | 1.00 | 1.05 (0.8–1.3) | 1.05 (0.8–1.3) |
| Dodds et al. | <37 wk of gestation | TTHM Relative risk (adjusted) | Third trimester | 1.00 | 1.05 (0.8–1.3) | 1.05 (0.8–1.3) |
| Wright et al. | <37 gestational weeks | TTHM OR (adjusted) | Third trimester, entire pregnancy average | 1.00 | 1.05 (0.8–1.3) | 1.05 (0.8–1.3) |
| Wright et al. | <37 gestational weeks | TTHM OR (adjusted) | Third trimester | 1.00 | 1.05 (0.8–1.3) | 1.05 (0.8–1.3) |

(Continued)
| Study                      | Definition of Outcome                                                                 | Exposure Agent | Measure of Effect Used in Meta-analysis | Exposure Timing Used in Meta-analysis | Measures of Effect (95% Confidence Intervals) |
|---------------------------|-------------------------------------------------------------------------------------|----------------|----------------------------------------|--------------------------------------|-----------------------------------------------|
| Lewis et al34             | Analyses were restricted to infants between 32 and 45 gestational weeks with a birth weight 500–5000 g | TTHM           | HR (adjusted)                          | Entire pregnancy average              | <40 40-60 μg/L ≥60 μg/L                     |
|                           |                                                                                      |                |                                        |                                      | 0.92 (0.82–1.02) 0.85 (0.74–0.97)           |
| Yang et al26              | <37 gestational weeks                                                                | TTHM           | OR (adjusted)                          | Entire pregnancy average              | <4.93 μg/L≥40 4.93–11.11 μg/L            |
|                           |                                                                                      |                |                                        |                                      | 1.03 (0.94–1.13) 1.08 (0.98–1.18)           |
| Hoffman et al33           | Delivery before 37 wk gestation                                                      | TTHM           | Risk ratio (adjusted)                   | Second trimester (assumed as entire pregnancy average in meta-analysis) | 2.2–4.6 μg/L≥40 33.1–55.0 μg/L |
|                           |                                                                                      |                |                                        |                                      | 0.89 (0.50–1.30) 0.90 (0.60–1.40)           |
| Kramer et al36d           | <5th percentile of weight for gestational age (excluding unrealistic gestational ages of ≤22 wk or ≥46 wk) | Chloroform     | OR (adjusted)                          | Entire pregnancy                      | 1.0 ≥10 μg/L                              |
|                           |                                                                                      |                |                                        |                                      | 1.3 (0.9–1.8) 1.8 (1.1–2.9)               |
| Bove et al21              | Live births at or below their race-, sex-, and gestational week-specific 5th percentile weight | TTHM           | OR (unadjusted/ adjusted)              | Entire pregnancy average              | <20 ppb 60–80 ppb ≥100 ppb               |
|                           |                                                                                      |                |                                        |                                      | 0.98 (0.84–1.15) 1.11 (1.01–1.22)           |
| Dodds et al39             | Lowest 10th of weight distribution among Canadian live births for each week of gestation for each sex | TTHM           | Relative risk (adjusted)               | Third trimester                      | 0–49 μg/L 50–74 μg/L 100 or more          |
|                           |                                                                                      |                |                                        |                                      | 1.00 (0.97–1.11) 1.01 (0.92–1.11)           |
|                           |                                                                                      |                |                                        |                                      | 1.08 (0.99–1.18)                             |
| Wright et al34            | Lowest 10th percentile of birth weight for each gestational week stratified by infant race, sex, and maternal race (only term births) | TTHM           | OR (adjusted)                          | Third trimester, entire pregnancy average | 0–60 μg/L ≥60–80 μg/L 80 μg/L           |
|                           |                                                                                      |                |                                        |                                      | 0.98 (0.89–1.09) 1.03 (0.94–1.14)           |
|                           |                                                                                      |                |                                        |                                      | 1.00 (0.92–1.09) 1.14 (1.02–1.26)           |
| Wright et al35            | Lowest decile of birth weight for each gestational week stratified by infant race and maternal race (only term births from 37 to 45 gestational weeks) | TTHM           | OR (adjusted)                          | Third trimester                      | 0–33 μg/L ≥33–74 μg/L ≥74–163 μg/L        |
|                           |                                                                                      |                |                                        |                                      | 1.06 (1.02–1.10) 1.13 (1.07–1.20)           |
| Hinckley et al224d        | Birth weight < lowest 10th percentile of birth weights by race, ethnicity, and gestational age (term or preterm) | TTHM           | OR (adjusted)                          | Third trimester                      | <40 μg/L 40–53 μg/L ≥53 μg/L            |
|                           |                                                                                      |                |                                        |                                      | 0.98 (0.90–1.07) 1.09 (1.00–1.18)           |

(Continued)
### TABLE 2. (Continued)

| Study                  | Definition of Outcome                                                                 | Exposure Agent | Measure of Effect Used in Meta-analysis | Exposure Timing Used in Meta-analysis | Measures of Effect (95% Confidence Intervals) | Exposure Categories |
|------------------------|---------------------------------------------------------------------------------------|----------------|----------------------------------------|---------------------------------------|-----------------------------------------------|--------------------|
| Porter et al<sup>31-d</sup> | Birth weight <10th percentile for gestational age (adjusted for sex and race)           | TTHM           | OR (adjusted)                          | Third trimester, entire pregnancy average | 29.48 µg/L < 30.00 µg/L                         | 64.48 µg/L —       |
|                        |                                                                                       |                |                                        |                                       | 37.38 µg/L ≤ 43.67 µg/L                      |                    |
|                        |                                                                                       |                |                                        |                                       | 49.13 µg/L ≤ 59.00 µg/L                      |                    |
| Yang et al<sup>26</sup>  | Birth weight ≤10th percentile for each gestational week stratified by infant sex      | TTHM           | OR (adjusted)                          | Entire pregnancy average               | 4.93<sup>b</sup> – 13.11<sup>c</sup> µg/L     |                    |
|                        |                                                                                       |                |                                        |                                       | 0.96<sup>d</sup> – 1.00<sup>d</sup> µg/L      |                    |
| Hoffman et al<sup>29</sup> | Birth weight <10th percentile for gestational age at birth, sex, maternal race and parity | TTHM           | Risk ratio (adjusted)                  | Third trimester                        | 2.2<sup>e</sup> – 6.6<sup>e</sup> µg/L        |                    |
|                        |                                                                                       |                |                                        |                                       | 1.0<sup>e</sup>  – 1.0<sup>e</sup> µg/L       |                    |

Where studies presented both adjusted and crude ORs, only the adjusted are given here as these were the ones used (unless otherwise stated). Only third trimester and entire pregnancy average exposure timings were used (unless otherwise stated). Some studies provided results for additional exposure timings that were not investigated in this meta-analysis.

*Where the exposure category was defined in Kramer et al<sup>16</sup> as below the limit of detection, it was assumed to be <1 µg/L.

+aReference category.

+bExposure range at low DBP site: not continuous with upper exposure categories but used as referent category.

+cStudies reporting on “IUGR” using definitions considered to be equivalent to SGA.

+dAdjusted results given when difference between adjusted & unadjusted was >15% of unadjusted value.

+Quoted 95% CIs recalculated from data published at different levels of confidence.

---

**FIGURE 1.** Plots of individual study slopes (solid colored lines) and the random-effects regression slope (dashed blue line) (both per 10 µg/L TTHM) estimated from these three trimester exposure to TTHM only, for (A) LBW, (B) term LBW, (C) PTD, and (D) SGA. Crosses indicate midpoints of exposure categories versus OR in that category.
A

B

C

D

Grellier et al
Epidemiology

proportionally to their weight in the meta-analysis regression; and (D) SGA. Study OR slopes are plotted with squares sized

term exposure to TTHM for (A) LBW, (B) term LBW, (C) PTD,

FIGURE 2. Forest plots of OR slopes per 10 µg TTHM/L for third trimester exposure to TTHM for (A) LBW, (B) term LBW, (C) PTD, and (D) SGA. Study OR slopes are plotted with squares sized proportionally to their weight in the meta-analysis regression; than Medline. Manual searching of bibliographies provided additional studies that met broad eligibility criteria: all but one were later excluded on the basis of more detailed criteria. A QUOROM diagram demonstrates the search method and the reasoning behind the exclusion of studies (Fig. 3). Further data were provided for the study by Porter et al11 to give exposure category quintiles for the analyses of interest that had not been presented in the published paper. Ultimately, fifteen studies were deemed suitable for inclusion in the meta-analysis. Characteristics of the studies included in the analysis are given in Table 1. The meta-analysis included 2 population-based case-control studies,16,17 2 cross-sectional studies,21,24 1 cohort study,13 2 retrospective cohort studies,12,22 2 prospective pregnancy studies,29,33 and 5 studies for which the design type was not explicitly named.18,26,31,34,35 For the purposes of this review, the studies were defined as population case-control studies, retrospective pregnancy cohort studies, or prospective pregnancy cohort studies (Table 1). The qualitative review of between-study heterogeneities found that the studies differed in their geographical location, their quoted measure of effect, adjustment for confounders, exposure characterization and categorization, and the definitions of health outcomes.

Eleven studies were conducted in the United States, 1 in the United Kingdom, 1 in Canada, and 1 in Taiwan. Three studies used data from Massachusetts, but these could be combined because the time periods did not overlap. Two studies looked at the same populations in the United States, but reported on different outcomes.29,33 The majority of studies reported their results as odds ratios; 1 study reported relative risk,12 2 reported risk ratios,29,33 and 1 reported a hazard ratio (HR).34 (Table 2). Eight studies provided only adjusted measures of effect; 5 provided crude and adjusted results, and 1 study provided crude figures where the difference between crude and adjusted was less than 15%.34 Apart from this last exception, adjusted measures of effect were used in the meta-analysis. Adjustment for confounding in all studies had been done using logistic regression analysis, except one study that had used a Poisson regression model.12 The covariates adjusted for in each study are shown in eFigure 2 (http://links.lww.com/EDE/A379).

The search retrieved studies in which exposure characterization differed, particularly in terms of exposure assessment. Studies not characterizing exposure with quantitative DBP concentration measurements were excluded. Exposure assessment methods used in the studies are given in Table 1. The types of measure included concentrations, either from sampling or monitoring data. Only one of the studies did not use total trihalomethane as an exposure agent, but instead
used trichloromethane (chloroform). Total trihalomethane concentration was by far the most common exposure agent across the studies. Many studies characterized exposure simply by taking the concentrations for the area (eg, water company, municipality, etc.) encompassing the maternal place of residence at birth. One study used hydraulic modeling to assign specific exposures to mothers, while most studies made use of routine monitoring data. Two provided measures of effect both for residential trihalomethane concentration derived from sampling, and for personal exposure calculated using published algorithms.

There were some disparities in the definitions of adverse birth outcomes among studies (Table 2). LBW was universally defined as birthweight <2500 g (or imperial equivalent). Term LBW was also universally defined as <2500 g for term births (themselves defined as ≥37 weeks of gestation). PTD was generally defined as a birth of <37 weeks of gestation, although one study used a definition that incorporated limits on gestational age and birth weight. The definitions of SGA (including IUGR) varied the most, with differences in the age-weight distributions and cut-off points, and whether only term births were included. Definitions of SGA also varied in terms of the population weight percentile cut-off points, and whether only term births were included. Differences among studies contributed to our decision to employ a random effects model in the meta-analysis.

Results of Meta-analysis

Figure 2 shows the study-specific exposure-response slopes and the pooled slope for each of the outcomes investigated. Results of the Q-test suggested that there was no heterogeneity among studies. The Q-test, however, has limited ability to detect heterogeneity when numbers of studies are small. Differences in results with fixed effects and random effects were scarcely distinguishable. In the light of these findings, and given the results of the qualitative review of between-study heterogeneities, we applied the more conservative approach of using the random effects model. The results of the random-effects meta-analysis are summarized in Table 3. These are given as odds ratio slopes (OR per 10 g trihalomethane/L) with 95% confidence intervals; Cochran Q-statistics are also provided for each subgroup analysis. Overall, we found little or no evidence for associations between trihalomethane concentration and the pregnancy outcomes examined.

Forest plots for the various pregnancy outcomes are given in Figure 2, assuming only total trihalomethane as a measure of exposure for the third trimester. We considered the distribution of studies in funnel plots (total trihalomethane only and third trimester exposure) to indicate that further investigation of bias would be justified, particularly in the case
of PT (although the low number of studies made their interpretation difficult). The results of weighted and unweighted Egger’s regression tests (eTable 1, http://links.lww.com/EDE/A379) provided no evidence for publication bias (or similar biases) in any of the subset analyses.

The leave-one-out sensitivity analysis results were tabulated, and differences between the results of each iteration and the original full subset analysis were calculated. Full results of the sensitivity analysis are presented in eTable 2 (http://links.lww.com/EDE/A379). Some very small changes of magnitude and changes of direction of effect were noted. Nevertheless, in none of the subset analyses did omitting an individual study change the summary measure of effect by more than 2%, with most differences being several orders of magnitude less. The direction of effect was altered only for analyses looking at LBW. This finding can be attributed to the summary OR slope being extremely close to 1.00. Removing the only study using chloroform as an exposure index instead of total trihalomethane16 had an effect on the direction of only one analysis (LBW, third trimester)—again the summary OR slope was very close to 1.00.

DISCUSSION

We used quantitative meta-analysis techniques to investigate associations between exposure to total trihalomethane in drinking water and indicators of fetal growth and prematurity. Meta-analytic techniques can increase the statistical power to detect small excess risks. Nonetheless, we found little or no evidence for associations with most indicators of fetal growth and prematurity, with the exception of SGA.

These results are broadly in line with narrative reviews carried out previously, which have found evidence for an association of disinfection by-product exposure with SGA but not with LBW or PTD.42,43 In contrast to previous qualitative results of these reviews, this meta-analysis did not find a positive association with term LBW.

We carried out subset analyses to investigate the effects of exposure timing and the inclusion of a study using chloroform as the exposure agent; small positive effects for SGA were reported only for analyses that included total trihalomethane as the exposure agent and third-trimester exposure or any exposure timing. We consider SGA to be the best characterized of these fetal growth outcomes because it takes gestational age of the fetus into account. As such, with SGA we expect to have more power to detect small risks relating to retarded fetal growth.

The Cochran test for homogeneity indicated a lack of heterogeneity among the studies. This was in contrast to the findings of our qualitative review of the studies, which showed study differences in the characteristics of the study populations, in the degree to which confounding was controlled, and in definitions of health outcomes. In addition, because total trihalomethane acts as a surrogate for exposure to an unknown putative agent, the actual concentrations of this agent (or agents) might differ among the studies. The outcome for which the meta-regression graphs display the least between-study heterogeneity (in terms of gradient) is that of SGA (Fig. 1D), where all but one of the studies indicate a positive slope. Because of these qualitative findings, and the fact that the Q-test is known to have a low power when the number of studies is small,58 we considered a random effects model to be most appropriate for the regression of the study-specific slopes.59 Other tests of heterogeneity, such as the I 2-test, were not employed as these are similarly limited when studies are few.60

The OR slopes should be viewed in the context of levels of total trihalomethane typically present in drinking water, and where potentially large populations are exposed.

### TABLE 3. Summary Table of Results of Meta-analyses for all Health Outcomes, Including Results of Subset Analyses for Exposure Agent and Exposure Timing

| Exposure Agent                | Exposure Timing | Health Outcome | No. Studies Included | OR Slope per 10 μg/L | 95% CI         | Q-Statistic |
|-------------------------------|-----------------|----------------|----------------------|----------------------|----------------|-------------|
| Only TTHM                     | Third trimester | LBW            | 4                    | 0.9999 (0.9735–1.0270)| 2.244          |             |
|                               |                 | Term LBW       | 4                    | 1.0337 (0.9272–1.1525)| 3.987          |             |
|                               |                 | PTD            | 6                    | 0.9896 (0.9781–1.0013)| 1.840          |             |
|                               |                 | SGA            | 6                    | 1.0100 (1.0006–1.0194)| 3.569          |             |
|                               | Any exposure timing | LBW       | 5                    | 1.0013 (0.9748–1.0286)| 2.495          |             |
|                               |                 | Term LBW       | 5                    | 1.0228 (0.9456–1.1063)| 4.008          |             |
|                               |                 | PTD            | 8                    | 0.9894 (0.9777–1.0007)| 4.124          |             |
|                               |                 | SGA            | 8                    | 1.0096 (1.0009–1.0184)| 4.641          |             |
| TTHM and chloroform           | Entire pregnancy| SGA            | 4                    | 1.0105 (0.9712–1.0514)| 4.659          |             |
|                               | Any exposure timing | LBW       | 5                    | 1.0001 (0.9737–1.0272)| 2.495          |             |
|                               |                 | PTD            | 9                    | 0.9894 (0.9777–1.0007)| 4.125          |             |
|                               | Entire pregnancy| PTD            | 4                    | 0.9696 (0.9139–1.0286)| 1.441          |             |
We applied our summary estimates of effect to United States and European guidelines (80 μg/L and 100 μg/L, respectively). As an example, we found that the risks of SGA for third trimester exposure to total trihalomethane at these levels were OR = 1.08 (95% CI = 1.01–1.17) and 1.10 (1.01–1.21), respectively. Results for the other 3 outcomes are provided in eTable 3 (http://links.lww.com/EDE/A379).

We carried out this meta-analysis under the assumption that the log-odds of the response variables varied linearly against concentration of total trihalomethane; this was in the absence of data to support other exposure-response relationships. This is a limitation of our analysis, and should be taken into account when using the slope estimates, particularly when extrapolating to high concentrations of total trihalomethane. Were it possible to pool all original data from these studies, specific exposure cut-offs might be examined, thereby facilitating investigation of exposure-response slopes.

The few number of studies included in some meta-analysis subsets limited the degree to which we could investigate differences in exposure assessment. Although some studies reported various exposure timings, these have not been extensively explored in the available literature; the majority of studies looked only at the third trimester, which is regarded as the most critical exposure period for these outcomes. For SGA, slightly stronger evidence was found for an association with exposure in the third trimester, which might be expected given that weight gain occurs mainly in the third trimester. Few studies reported exposure specifically to chloroform, limiting the analysis of different exposure agents. However, total trihalomethane and chloroform both presumably serve merely as indicators for the unknown putative agent.

In the leave-one-out sensitivity analysis, individual studies had little effect on the magnitude of the OR slopes, although direction of the effect was altered in some instances. The large study by Dodds et al exerted considerable influence on the summary measure. Inspection of the meta-analysis regression slopes (Fig. 1C) showed that a study with very narrow exposure categories tended to produce slopes with tight confidence intervals, which thus increased their weighting in the meta-analysis. Results changed very little in the leave-one-out sensitivity analysis for any of the SGA subgroup analyses, further supporting evidence of an association for this outcome.

Interpretation of the funnel plots was hampered by the small number of studies. Although the results of Egger’s regression test (both weighted and unweighted) demonstrated that there was no notable publication bias in results of any subset analysis, the robustness of this test was limited.

Although definitions of LBW, term LBW, and PTD were consistent across all studies, definitions for SGA (sometimes called IUGR, in spite of differences between the 2 outcomes) differed in the weight percentile cut-off points and the degree to which reference curves had been adjusted for various factors (Table 2).

It was not possible to explore the effects of varying the exposure categories because the studies did not present the distribution of their exposure data in sufficient detail. The selection of exposure category midpoints may have introduced bias into the model for the uppermost exposure categories which, if open-ended, were set using the midpoint from the preceding category. Use of the exposure-response slope in the assessment of population health risks should take this into account.

No toxicologic data were incorporated into the analysis. An investigation has been published previously on the use of Bayesian methods for the combination of epidemiologic and toxicologic studies. Trihalomethane exposure and LBW were used for illustration; combining study-specific dose-response slope estimates. Results were found to be contingent on robust data and consistent definitions for health outcomes in humans and in animals. Furthermore, epidemiologic studies commonly use total trihalomethane concentration in water as a proxy for exposure, rather than a measure of ingested dose. In addition, in normalizing the epidemiologic studies to toxicologic ones, the assumption is made that epidemiologic studies have reported on trihalomethanes as the putative agent and that all exposure is through ingestion. The validity of these assumptions may be questioned.

We expected Berkson error associated with aggregate total trihalomethane data to dominate over random error for residential exposure estimates in the individual studies, and hence in the summary estimate. Berkson error may have reduced the power of the studies, but the risk estimates were probably not attenuated as they might have been if random error were dominant. Mobility of women during their pregnancies, and other factors such as changing residence, between areas with different exposure, may have led to exposure misclassification and attenuation of the summary measures of effect.

Elevated risks of restricted fetal growth have been associated with exposure to total trihalomethane of those mothers and infants carrying a genetic polymorphism for CYP2E1, the enzyme primarily involved in the metabolism of low doses of chloroform. If these data are corroborated, people carrying the CYP2E1 variant could have considerably greater risk of SGA than what we report here for pooled populations.

Studies generally used indirect estimates of exposure based on monitoring data linked to maternal residence at birth. As such, exposure data were aggregated in both space and time, due to marked variations in trihalomethane concentration occurring from home to home and throughout each pregnancy. Hundreds of disinfection by-products might be present in any one drinking water sample. Only studies using area-level concentration of total trihalomethane (and, in one
instance, chloroform) in drinking water were combined in this meta-analysis. Some studies estimated exposure through different routes or pathways, but we included only those based on maternal residence. Area-level total trihalomethane data represent the most practicable means of categorizing exposure in large studies; the costs of accurately estimating intake in large populations are prohibitively high. As long as the putative agent in the DBP mixture remains unknown, the results of this meta-analysis may be useful in health impact assessment or other estimations of burden of disease attributable to disinfection by-products, where routine total trihalomethane monitoring data are available. It would be worthwhile to examine the potential effects of individual disinfection by-products, if such data become available.

Large, well-designed epidemiologic studies are needed that take into account relevant confounders and characterization of disinfection by-product exposure, and with carefully defined health outcomes. In the absence of such studies, meta-analysis provides the best possible estimate measure for use in risk assessment and public health policy.

ACKNOWLEDGMENTS

We thank Chad Porter for providing additional information on exposure category cut-off points for his group’s study.

REFERENCES

1. Wigle DT. Safe drinking water: a public health challenge. Chronic Dis Can. 1998;19:103–107.
2. Galal-Gorchev H. Chlorine in water disinfection. Pure Appl Chem. 1996;68:1731–1735.
3. Cutler D, Miller G. The role of public health improvements in health advances: the twentieth-century United States. Demography. 2005;42:1–2
4. Bogart TA, Lichtenberg JJ, Kroner RC. The occurrence of organohalides in chlorinated drinking water. J Am Water Works Assoc. 1974;66:703–706.
5. Rook J. Chlorination reactions of fulvic acids in natural waters. Environ Sci Technol. 1977;11:478–482.
6. Nieuwenhuijzen MJ, Toledano MB, Eaton NE, Fawell J, Elliott P. Chlorination disinfection by-products in water and their association with adverse reproductive outcomes: a review. Occup Environ Med. 2000;57:73–85.
7. World Health Organization (WHO). Guidelines for Drinking Water—First Addendum to the Third Edition, Vol. 1, Recommendations. 3rd ed. Geneva: World Health Organization; 2006.
8. Richardson SD, Plewa MJ, Wagner ED, Schoeny R, Demarini DM. Occurrence, genotoxicity, and carcinoogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. Mutat Res. 2007;636:178–242.
9. Richardson SD. Drinking water disinfection by-products. In: Meyers RA, ed. The Encyclopedia of Environmental Analysis & Remediation, Vol. 3. New York: John Wiley & Sons; 1998:1898–1921.
10. International Programme on Chemical Safety (ICPS). Disinfectant and Disinfectant By-Products. Environmental Health Criteria 216. Geneva: United Nations Environment Programme (UNEP), International Labour Organization (ILO), World Health Organization (WHO); 2000.
11. Nieuwenhuijzen MJ, Toledano MB, Elliott P. Uptake of chlorination disinfection by-products; a review and a discussion of its implications for exposure assessment in epidemiological studies. J Expo Anal Environ Epidemiol. 2000;10:586–599.
12. Dodds L, King W, Woodcott C, Pole J. Trihalomethanes in public water supplies and adverse birth outcomes. Epidemiology. 1999;10:233–237.
13. Gallagher MD, Nuckols JR, Stallones L, Savitz DA. Exposure to trihalomethanes and adverse pregnancy outcomes. Epidemiology. 1998;9:484–489.
14. Jaakkola JJ, Magnus P, Skrondal A, Hwang BF, Becher G, Dybing E. Foetal growth and duration of gestation relative to water chlorination. Occup Environ Med. 2001;58:437–442.
15. Kanitz S, Franco Y, Patrone V, et al. Association between drinking water disinfection and somatic parameters at birth. Environ Health Perspect. 1996;104:516–520.
16. Kramer MD, Lynch CF, Isaacson P, Hanson JW. The association of waterborne chloroform with intrauterine growth retardation. Epidemiology. 1992;3:401–413.
17. Shute DA, Andrews KO, Pastore LM. Drinking water and pregnancy outcome in central North Carolina: source, amount, and trihalomethane levels. Environ Health Perspect. 1995;103(6):592–6.
18. Toledano MB, Nieuwenhuijzen MJ, Best N, et al. Relation of trihalomethane concentrations in public water supplies to stillbirth and birth weight in three water regions in England. Environ Health Perspect. 2005;113:225–232.
19. Tuthill RW, Giusti RA, Moore GS, Calabrese EJ. Health effects among newborns after prenatal exposure to ClO2-disinfected drinking water. Environ Health Perspect. 1982:46:39–45.
20. Yang CY. Drinking water chlorination and birth outcomes in Taiwan. Toxicology. 2004;198:249–254.
21. Bove FJ, Fulcomer MC, Klotz JB, Esmart J, Dufficy EM, Savrin JE. Public drinking water contamination and birth outcomes. Am J Epidemiol. 1995;141:850–862.
22. Hinckley AF, Bachand AM, Reif JS. Late pregnancy exposures to disinfection by-products and growth-related birth outcomes. Environ Health Perspect. 2005;113:1808–1813.
23. Lewis C, Suffet IH, Ritz B. Estimated effects of disinfection by-products on birth weight in a population served by a single water utility. Am J Epidemiol. 2006;163:38–47.
24. Wright JM, Schwartz J, Dockery DW. Effect of trihalomethane exposure on fetal development. Occup Environ Med. 2003;60:173–180.
25. Yang CY, Cheng BH, Tsai SS, Wu TN, Lin MC, Lin KC. Association between chlorination of drinking water and adverse pregnancy outcome in Taiwan. Environ Health Perspect. 2000;108:765–78.
26. Yang CY, Xiao ZP, Ho SC, Wu TN, Tsai SS. Association between trihalomethane concentrations in drinking water and adverse pregnancy outcome in Taiwan. Environ Res. 2007;104:390–395.
27. Källen BA, Robert E. Drinking water chlorination and delivery outcome—a registry-based study in Sweden. Reprod Toxicol. 2000;14:303–309.
28. Boorman GA. Drinking water disinfection by-products: review and approach to toxicity evaluation. Environ Health Perspect. 1999;107(Suppl 1):207–217.
29. Hoffman CS. Drinking water disinfection by-product exposure and fetal growth. Epidemiology. 2008;19:729.
30. Infante-Rivard C. Drinking water contaminants, gene polymorphisms, and pregnancy outcome in central North Carolina: source, amount, and trihalomethane levels. Environ Health Perspect. 1992;3:407–413.
31. Gallagher MD, Nuckols JR, Stallones L, Savitz DA. Exposure to trihalomethanes and adverse pregnancy outcomes. Epidemiology. 1998;9:484–489.
32. Jaakkola JJ, Magnus P, Skrondal A, Hwang BF, Becher G, Dybing E. Foetal growth and duration of gestation relative to water chlorination. Occup Environ Med. 2001;58:437–442.
33. Kanitz S, Franco Y, Patrone V, et al. Association between drinking water disinfection and somatic parameters at birth. Environ Health Perspect. 1996;104:516–520.
34. Kramer MD, Lynch CF, Isaacson P, Hanson JW. The association of waterborne chloroform with intrauterine growth retardation. Epidemiology. 1992;3:401–413.
35. Wright JM, Schwartz J, Dockery DW. The effect of disinfection by-products and mutagenic activity on birth weight and gestational duration. Environ Health Perspect. 2004;112:920–925.
36. Swan SH, Waller K, Hopkins B, et al. A prospective study of spontaneous abortion: relation to amount and source of drinking water consumed in early pregnancy. Epidemiology. 1998;9:126–133.
37. Waller K, Swan SH, DeLorenze G, Hopkins B. Trihalomethanes in drinking water and spontaneous abortion. Epidemiology. 1998;9:134–140.
38. Waller K, Swan SH, Windham GC, Fenster L. Influence of exposure assessment methods on risk estimates in an epidemiologic study of total trihalomethane exposure and spontaneous abortion. J Expo Anal Environ Epidemiol. 2001;11:522–531.

39. Dodds L, King W, Allen AC, Armson BA, Fell DB, Nimrod C. Trihalomethanes in public water supplies and risk of stillbirth. Epidemiology. 2004;15:179–186.

40. King WD, Dodds L, Allen AC. Relation between stillbirth and specific chlorination by-products in public water supplies. Environ Health Perspect. 2000;108:883–886.

41. Reif JS, Hatch MC, Bracken M, Holmes LB, Schwetz BA, Singer PC. Reproductive and developmental effects of disinfection by-products in drinking water. Environ Health Perspect. 1996;104:1056–1061.

42. Graves CG, Matanoski GM, Tardiff RG. 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.

43. Tardiff RG, Carson ML, Ginevan ME. Updated weight of evidence for an association between adverse reproductive and developmental effects and exposure to disinfection by-products. Regul Toxicol Pharmacol. 2006;45:185–205.

44. Hwang BF, Jaakkola JJ. Water chlorination and birth defects: a systematic review and meta-analysis. Arch Environ Health. 2003;58:83–91.

45. Nieuwenhuijsen MJ, Martinez D, Grellier J, et al. Exposure to disinfection by-products and congenital malformations—a review and meta-analysis. Environ Health Perspect. 2009;117:1486–1493.

46. Egger M, Davey SG, Altman DG. Systematic Reviews in Health Care: Meta-Analysis in Context. London, UK: BMJ Publishing Group; 2001.

47. Huedo-Medina TB, Sanchez-Meca J, Marin-Martinez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I² index? Psychol Methods. 2006;11:193–206.