# Exposure to Disinfection By-products, Fetal Growth, and Prematurity: A Systematic Review and Meta-analysis

Grellier, James[a]; Bennett, James[a]; Patelarou, Evridiki[b]; Smith, Rachel B.[a]; Toledano, Mireille B.[a]; Rushton, Lesley[a]; Briggs, David J.[a]; Nieuwenhuijsen, Mark J.[a,c]

[a]Department of Epidemiology and Public Health, Imperial College London, London, United Kingdom; [b]Department of Social Medicine, University of Crete, Heraklion, Greece; and [c]Center for Research in Environmental Epidemiology (CREAL), IMIM, CIBERESP, Barcelona, Spain.

## ABSTRACT

**OBJECTIVES**: We aimed to provide quantitative estimates of exposure-response relationships between total trihalomethanes in drinking water and several adverse birth outcomes relating to fetal growth and prematurity, suitable for use in assessments of attributable burden of disease.

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

**RESULTS**: A comprehensive literature search yielded 37 studies for consideration, 15 of which were selected for the extraction of relative risks relating adverse birth outcomes to TTHM exposure. Sufficient data were available for meta-analyses to be carried out for four adverse birth outcomes: low birth weight (LBW), term low birth weight (TLBW), preterm delivery (PTD) and small for gestational age (SGA) (including intra uterine growth retardation (IUGR)). We found little or no evidence for associations between third trimester TTHM exposure and LBW (OR per 10µg TTHM/L = 0.9999 95% CI 0.9735, 1.0270), TLBW (OR per 10µg TTHM/L = 1.0337 95% CI 0.9272, 1.1525) or PTD (OR per 10µg TTHM/L = 0.9896 95% CI 0.9781, 1.0013). We found evidence for an association with SGA (OR per 10µg TTHM/L = 1.0100 95% CI 1.0006, 1.0194).

**CONCLUSIONS**: We found little or no evidence for associations between TTHM concentration and most adverse birth outcomes relating to fetal growth and prematurity. We did find evidence for an association between TTHM concentration and SGA. We

discuss these findings and the uncertainties—relating particularly to exposure—which may have affected them.

# BACKGROUND

## Drinking water chlorination and disinfection by-products

Supplies of drinking water were first disinfected using chlorine at the start of the 20th century,1 primarily as a means of reducing mortality and morbidity associated with waterborne disease.2,3 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 (DBPs). First identified in disinfected drinking water in the 1970s,[4,5] trihalomethanes are generally the most abundant of the DBPs, but many other chemicals may also be present.[6,7] Over 600 DBPs have been reported,[8,9] 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.[10,11]

## Health outcomes associated with disinfection by-products

The health outcomes investigated in human studies related to fetal growth and prematurity over the following two decades have included LBW,[12-20] TLBW,[13,21-26] very LBW,[12,18,21,27] SGA,[13,15,18-20,24,28,29] IUGR,[16,22,27,30,31] preterm delivery (PTD[12-17,20,22,24-26,32-35] and very PTD[22]), fetal death (miscarriage,[17] spontaneous abortion,[36-38] and stillbirth[12,18,39,40]).

Six systematic reviews of the epidemiological evidence for reproductive and developmental effects of exposure to DBPs have been published to date, including two narrative reviews,[6,41] two comprehensive weight of evidence reviews,[42,43] and two meta-analyses looking at chlorination and birth defects.[44,45]

The results presented by individual studies on adverse birth outcomes relating to fetal growth and prematurity are mixed, with study results varying in direction and magnitude of effect. Existing reviews present a useful synthesis and critique of the available literature but have not attempted to arrive at quantitative summary measures of effect for

any outcomes related to fetal growth. These reviews concur that the weight of evidence is suggestive of small, positive associations between THM concentrations in drinking water and some adverse birth outcomes related to fetal growth restriction (TLBW, SGA, IUGR), although evidence is not conclusive.

### Objectives

The objectives of this meta-analysis were to systematically review existing epidemiological evidence and to carry out a meta-analysis of these data, to produce best-estimate exposure-response slopes of TTHM exposure and adverse birth outcomes relating to fetal growth and prematurity suitable for application in the estimation of burden of disease using routine drinking water quality monitoring data.

## METHODS

### Search methods

We carried out a systematic review of the existing literature on THMs and adverse birth outcomes related to fetal growth and prematurity, using the following review question: "Given existing epidemiological evidence, what is the exposure-response relationship between exposure of pregnant women to THMs 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. (2001)[46] and carried out and reported on both search methods and results following standards outlined in the QUOROM statement[47] and the MOOSE Guidelines.[48]

We carried out a systematic, comprehensive bibliographic search using the US National Library of Medicine (USNLM) Medline database for the years 1980-2007, using the PubMed interface. Full details of the search are provided in Annex 1. We checked the list of studies identified thus far for completeness against studies referenced in existing reviews.[6,10,41-43]

We defined *a priori* eligibility criteria to restrict the studies included. We only retained studies if they were peer-reviewed journal articles either (a) already published or (b) available on-line but awaiting formal publication, or were studies published by a highly reputable independent body such as WHO or USEPA. Studies were only included if they were published in English, were epidemiological studies, used maternal residence for

exposure estimation, and presented odds ratio (OR) or relative risk (or other comparable measure of effect) of at least one adverse birth outcome associated with exposure to DBPs. Studies not meeting these criteria were excluded and the specific reasons for their exclusion were noted. Studies meeting the criteria were shortlisted for inclusion in the meta-analysis. The list was narrowed down on the basis of the exposure assessment methods used: only those which characterised DBP exposure using $\geq 3$ exposure categories were included. Studies using binary characterisation of exposure were excluded primarily because such a measure offers only a crude index of exposure and early epidemiological studies classifying exposure according to water treatment methods have been criticised for their failure to capture a more detailed picture of exposure to DBPs.[11] Secondly, it was considered impracticable to combine relative risks from studies with binary exposure characterisation and continuous/categorical exposure. Thirdly, TTHM concentrations from routine drinking water monitoring data are generally available, hence an estimate of a continuous odds ratio slope was considered to provide health impact assessments with the most useful information. Fourthly, in developed countries, the reporting of drinking water treatment type is generally not mandatory, whereas reporting of TTHM 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 two researchers using a pre-designed standard data collection form: study design, exposure characterisation, definitions of exposure categories, and measures of effect and confidence intervals for each exposure category (Table 1). The two datasets were checked against one another and any inconsistencies addressed. 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 TTHM or, in one case, trichloromethane (chloroform), using one of two measures: either parts per million (ppm) or micrograms per litre (µg/L). We considered concentrations given in ppm as equivalent to µg/L, since at the low concentrations present in drinking water these are virtually equivalent. In order to include the one study reporting only chloroform concentrations as an exposure measure,[16] we multiplied reported exposure

categories by a factor of 1.33, on the assumption that chloroform might make up 75% of the TTHM mixture and that concentrations of chloroform and TTHM in drinking water are highly correlated.[49]

The majority of studies presented their results as ORs with 95% confidence intervals (CIs), although some expressed their results using other measures of effect (e.g. hazard ratio, HR, or relative risk, RR, or risk ratio); 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%.[21] In this instance, the standard error on the OR was calculated from the 99% CIs provided using the formula:

$$\text{standard error} = (\ln(\text{upper 99\% CI}) - \ln(\text{lower 99\% CI}))/(2*2.575)$$

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

$$\text{upper 95\% CI} = \exp(\ln(\text{OR}) + (1.96*\text{standard error})$$

$$\text{lower 95\% CI} = \exp(\ln(\text{OR}) - (1.96*\text{standard error})$$

The majority of the studies reported measures of effect adjusted for confounders. Adjustment had been carried out for a range of covariates which varied across the studies, but in the majority of cases the same important factors had been adjusted for (maternal age, parity, smoking, social deprivation). Several studies did not provide unadjusted results so adjusted results were used in the meta-analysis. In the one case where only unadjusted results were reported, these were used. The measures of effect for each health outcome, per exposure category, together with 95% CIs were summarised in tables (Tables 2a-d).

Only those studies characterising exposure using maternal residence were included in this analysis as we considered it desirable to minimise between-study heterogeneities relating to exposure assessment methods. Studies were subsequently grouped 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 categorised either by trimester or given for the whole pregnancy. The number of exposure categories used in studies varied from three to six. Given the variation in the exposure assessment between studies, we decided to carry out a two-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.[16] For each of these two 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, since this 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 $\geq 4$ studies. It was not practicable to quantitatively explore other heterogeneities between the studies for a number of reasons: studies were relatively similar in overall design; differences between studies were not presented sufficiently consistently; 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.), we considered that the number of studies included in such subgroups was prohibitively low for the application of meta-analytical methods.

Techniques for pooling correlated estimates to compute regression slopes across different exposure categories in individual studies have been described previously.[50] All included studies provided measures of effect for several exposure categories. Cut-off points of these categories differed between studies (Tables 2a-d). Meta-analysis was carried out with the R software package[51] using scripts adapted from those developed by Key et al. (2006).[52] 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. In cases where there was no upper limit to the topmost exposure category, a midpoint was derived using the half the width of the preceding category.

The use of different dose-response models to obtain study-specific slope estimates has been explored previously,[53] and slope estimates (and standard errors) were observed to be higher when dose was used, as compared to 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 than the other. Therefore we carried out a regression of ln(OR) against exposure. In the regression, we

assumed that exposure to zero DBPs from water was unlikely, since exposure to volatile DBPs such as THMs can occur in the domestic environment through several routes (ingestion, inhalation, dermal absorption) and through a variety of pathways (drinking, eating, cooking, washing) and therefore the intercepts of the regression slopes were not constrained to go through the origin. The reference categories used in each study differed (see Tables 2a-d), which further supported this decision.

In addition to the qualitative investigation of heterogeneity between the studies, as described above, Cochran's 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 (Figures 1a-d) and forest plots (Figures 2a-d).

In order to investigate the role of publication, and other, biases in the meta-analysis, we produced funnel plots (Annex 2) for visual inspection of the symmetry of the data, as well as carrying out the Egger regression test.[54]

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 health outcome for third trimester exposure to TTHM only at levels currently prescribed as guidelines in the US and the European Union (80µg/L and 100µg/L respectively).[55,56]

# 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 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 was drawn up demonstrating the search method and the

reasoning behind the exclusion of studies (Figure 3). Further data were provided for the study by Porter et al. (2005),[31] 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 two population-based case-control studies,[16,17] two cross-sectional studies,[21,24] one cohort study,[13] two retrospective cohort studies,[12,22] two prospective pregnancy studies[29,33] and five 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 (see 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 characterisation and categorisation, and the definitions of health outcomes.

Eleven studies were conducted in the USA, one in the UK, one in Canada and one in Taiwan. Three studies used data from Massachusetts, but these could be combined since the time periods of each study did not overlap. Two studies looked at the same populations in the USA, but reported on different outcomes.[29,33]

The majority of studies reported their results as odds ratios (OR); one study reported relative risk (RR),[12] two reported risk ratios[29,33] and reported a hazard ratio (HR)[34] (Tables 2a-d). Eight studies provided only adjusted measures of effect; five provided crude and adjusted results, and one study provided crude figures where the difference between crude and adjusted was less than 15%.[21] Apart from this 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 Annex 3.

The search retrieved studies in which exposure characterisation differed, particularly in terms of the exposure assessment used. Studies not characterising exposure with quantitative DBP concentration measurements were excluded. Exposure assessment methods used in the studies are given in Table 1. The types of measure used included concentrations of these agents, either from sampling or monitoring data. Only one of the included studies did not use TTHM as an exposure agent, but instead used trichloromethane (chloroform).[16] TTHM concentration was by far the most common

exposure agent across the studies. Many studies characterised exposure simply by taking the concentrations for the area (e.g. water company, municipality etc.) encompassing the maternal place of residence at birth. One study used hydraulic modelling to assign specific exposures to mothers,[13] while most studies made use of routine monitoring data. Two provided measures of effect both for residential TTHM concentration derived from sampling, and for personal exposure calculated using published algorithms.[29,33]

There were some disparities in the definitions of adverse birth outcomes between studies (Tables 2a-d). LBW was universally defined as birthweight <2500g (or imperial equivalent). Term LBW was also universally defined as <2500g (or imperial equivalent) 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.[34] The definitions of SGA (including IUGR) varied the most, with differences in the age-weight distributions and cut-off points, and whether or not only term births were included. Definitions of SGA also varied in terms of the population weight percentile cut-off points.

The presence of differences between studies contributed to our decision to employ a random effects model in the meta-analysis.

**Results of meta-analysis**

Figures 2a-d show 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 between the studies. The Q-test has, however, been shown to be limited in usefulness in detecting heterogeneity when numbers of studies are small.[57] Differences in results produced 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 summarised in Table 3. These are given as odds ratio slopes (OR per 10µg TTHM/L) with 95% confidence intervals; Cochran's Q-statistics are also provided for each subgroup analysis. Overall, we found little or no evidence for associations between TTHM concentration and LBW, TLBW or PTD. We did, however, find some evidence for an association between TTHM concentration and SGA.

Forest plots for LBW, TLBW, PTD and SGA respectively are given in Figures 2a-d, assuming only TTHM as a measure of exposure, for the third trimester. We considered the distribution of studies in funnel plots for LBW, TLBW, PTD and SGA (TTHM only and third trimester exposure) (Annex 2) to indicate that further investigation of bias would be justified, particularly in the case of PTD, although the low number of studies made their interpretation difficult. The results of weighted and unweighted Egger's regression tests (Annex 4) 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 Annex 5. 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 TTHM[16] had an effect only on the direction of one analysis (LBW, third trimester) – again the summary OR slope was very close to 1.00.

## DISCUSSION

In this study we have brought together the existing body of epidemiological evidence using quantitative meta-analysis techniques to investigate potential associations between exposure to TTHM in drinking water and indicators of fetal growth and prematurity. Grouping together generally small studies—the risk estimates of which may have been attenuated by non-differential misclassification bias—using meta-analytical techniques increased our statistical power to detect small excess risks of adverse birth outcomes related to exposure to TTHM in drinking water. The summary measures produced demonstrated that there is little or no evidence for association in the case of most indicators of fetal growth and prematurity, with the exception of SGA, for which we found some evidence of an association.

The results of this meta-analysis are broadly in line with the narrative reviews that have been carried out previously, which have found evidence for an association between DBP exposure and SGA, and no association for LBW or PTD.[42,43] In contrast to the qualitative results of these reviews, this meta-analysis did not find evidence of a positive association with TLBW.

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 TTHM as the exposure agent and the third trimester exposure or any exposure timing. We consider that SGA is the best characterized of those fetal growth outcomes investigated because it takes gestational age of the fetus into account. As such, with SGA we expect to have a higher power to detect small risks relating to retarded fetal growth.

The results of Cochran's test for homogeneity implied no significant degree of heterogeneity between the studies. This was in contrast to the findings of our qualitative review of the studies, which showed that the studies differed in the characteristics of the study populations, in the degree to which confounding was controlled, and in differences in definitions of health outcomes. In addition, because TTHM acts as a surrogate for exposure to an unknown putative agent, the actual concentrations of this agent (or agents) might differ between the studies. The outcome for which the meta-regression graphs display the least between-study heterogeneity in terms of gradient is that of SGA (Figure 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 included 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 it has been demonstrated that this is similarly limited when study numbers are low.[60]

The OR slopes that we have reported should be viewed in the context of levels of TTHM typically present in drinking water, and where potentially large populations are exposed. We applied our summary estimates of effect to US and European guidelines (80µg/L and 100µg/L, respectively). As an example, we found that the risks of SGA for third trimester exposure to TTHM at these levels were found to be OR = 1.08 (95% CI 1.01, 1.17) and

OR = 1.10 (95% CI 1.01, 1.21), respectively. Results for the other three outcomes are provided in Annex 6.

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

The low number of studies included in some meta-analysis subsets limited the degree to which we could investigate differences in exposure assessment between studies. Although some studies reported on different exposure timings, these have not been extensively explored in the available literature; the majority of studies only looked at the third trimester, as this is regarded as the most critical exposure period for these outcomes. For SGA, slightly stronger evidence was found for an association in the third trimester of exposure, which might be expected given that cell growth—and therefore weight gain—occurs mainly in the third trimester.[43] Few studies reported solely on exposure to chloroform, limiting the scope for an analysis of different exposure agents. Importantly, while reviewing the included studies it became apparent that this meta-analysis should be carried out with the assumption that TTHM and chloroform both probably serve only as indicators for the unknown putative agent.

While the leave-one-out sensitivity analysis illustrated that omitting individual studies had little effect on the magnitude of the OR slopes in any of the analyses, direction of the effect was altered in some instances. The study by Dodds et al. (1999) was a large study and its inclusion exerted considerable influence on the summary measure.[35] Inspection of the meta-analysis regression slopes (Figure 1c) showed that a study with very narrow exposure categories[22] tended to produce slopes with tight confidence intervals, which thus increased their weighting in the meta-analysis. We observed that the results changed very little independent of which study was removed 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 sparseness of included 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 by the low numbers of studies.

Whereas definitions of LBW, TLBW and PTD were consistent across all studies, definitions for SGA and IUGR (which was grouped together with SGA) differed in terms of the weight percentile cut-off points and the degree to which reference curves had been adjusted for various factors (Table 2b). We did not investigate IUGR separately from SGA since the definitions used for both terms were actually describing SGA; the results of our study should be interpreted in terms of SGA. In practice, IUGR is a clinical diagnosis made during pregnancy, whereas SGA status is ascribed to infants at birth on the basis of birth weight and gestational age compared on standardised curves, a distinction that was not made in studies ostensibly reporting on IUGR.

It was not possible to explore the effects of using other methods of setting midpoints for each exposure category 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 only be done taking this into account.

No toxicological data were incorporated into the analysis. An investigation into the use of Bayesian methods for the combination of epidemiological and toxicological studies using THM exposure and LBW for illustration which combined study-specific dose-response slope estimates has been published previously.[53] While these methods showed potential, meaningful results were found to be contingent on robust data and the existence of consistent definitions for health outcomes in humans and in animals. Furthermore, epidemiological studies commonly use TTHM concentration in water as a proxy for exposure, rather than a measure of ingested dose. In addition, in normalising the epidemiological studies to toxicological ones, the assumption is made that epidemiological studies have reported on THMs as the putative agent and that all exposure is through ingestion, the validity of which may be questioned.

Berkson error associated with aggregate TTHM data was expected 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 moving house, between areas with different exposure may have led to exposure misclassification and attenuation of the summary measures of effect.

Greatly elevated risks of restricted fetal growth have been associated with exposure to TTHM of those mothers and infants carrying a genetic polymorphism for CYP2E1, the enzyme primarily involved in the metabolism of low doses of chloroform.[30] While it is unlikely that only one gene was responsible for this environment-gene interaction, if these data are corroborated with further research, for those members of the population carrying the CYP2E1 variant the excess risk of SGA could be considerably greater than the figure we report here, which is based on the studies' pooled populations.

The studies combined in this meta-analysis generally used indirect estimates of exposure based on monitoring data linked to individual mothers by maternal residence at birth. As such, exposure data were considerably aggregated in both spatial and temporal dimensions, marked variations in THM concentration occurring from home to home, and throughout each pregnancy. Many hundreds of DBPs might be present in any one drinking water sample. Only studies using area level concentration of TTHM in drinking water (and in one instance, chloroform) assigned to maternal residence were combined in this meta-analysis. Some of the included studies estimated exposure through different routes or pathways but we chose to include those which characterised THM exposure using maternal residence. Area level TTHM data represent the most practicable means of categorising exposure in large studies, since the data requirements and costs of accurately estimating uptake profiles in a large population are prohibitively high. As long as the putative agent in the DBP mixture remains unknown, the results of this meta-analysis may be considered useful in health impact assessment or other means of estimating burden of disease attributable to DBPs, where routine TTHM monitoring data might be readily used. In future work (including systematic reviews and meta-analyses, should sufficient epidemiological studies be available) it would be worthwhile to examine the potential effects of individual disinfection by-products.

Recommendations stemming from this study generally echo those of previous reviews. There is a need to carry out large, well-designed epidemiological studies which take into account relevant confounders and characterisation of exposure, and take care to use meaningful health outcomes that are properly defined.[11,42,43] In the absence of such studies, the use of meta-analysis was justified as a means to produce a best estimate measure for use in health impact/risk assessment.

## CONCLUSIONS

In this study we have brought together the existing body of epidemiological evidence, and generally found little or no evidence for associations between TTHM concentration and most adverse birth outcomes relating to fetal growth and prematurity. We did find evidence for an association between TTHM concentration and SGA. Although the magnitudes of the summary risk slopes reported for SGA are small, the potential for exposure across very large populations is high, particularly given widespread exposure to DBPs through showering, bathing, and cooking etc., as well through consumption of drinking water. The burden of SGA attributable to DBP exposure may, therefore, not be inconsiderable in areas with sustained, high levels of DBPs. We investigated and discussed uncertainties relating particularly to exposure, and their effects our summary estimates. These factors should be considered by health impact/risk assessors making use of our summary estimates, and represent areas for further work.

## REFERENCES

1. Wigle DT. Safe drinking water: a public health challenge. *Chronic Dis Can.* 1998;19(3):103-7.

2. Galal-Gorchev H. Chlorine in water disinfection. *Pure and Applied Chemistry.* 1996;68(9):1731-1735.

3. Cutler D, Miller G. The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States. *Demography.* 2005;42:1-22.

4. Bellar 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. Nieuwenhuijsen MJ, Toledano MB, Eaton NE, Fawell J, Elliott P. Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: a review. *Occup Environ Med.* 2000;57(2):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 carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. *Mutat Res.* 2007;636(1-3):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;1898-1421.

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. Nieuwenhuijsen 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(6 Pt 1):586-99.

12. Dodds L, King W, Woolcott C, Pole J. Trihalomethanes in public water supplies and adverse birth outcomes. *Epidemiology.* 1999;10(3):233-7.

13. Gallagher MD, Nuckols JR, Stallones L, Savitz DA. Exposure to trihalomethanes and adverse pregnancy outcomes. *Epidemiology.* 1998;9(5):484-9.

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(7):437-42.

15. Kanitz S, Franco Y, Patrone V, Caltabellotta M, Raffo E, Riggi C, Timitilli D, Ravera G. Association between drinking water disinfection and somatic parameters at birth. *Environ Health Perspect.* 1996;104(5):516-20.

16. Kramer MD, Lynch CF, Isacson P, Hanson JW. The association of waterborne chloroform with intrauterine growth retardation. *Epidemiology.* 1992;3(5):407-13.

17. Savitz DA, Andrews KW, 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, Nieuwenhuijsen MJ, Best N, Whitaker H, Hambly P, de Hoogh C, Fawell J, Jarup L, Elliott P. Relation of trihalomethane concentrations in public water supplies to stillbirth and birth weight in three water regions in England. *Environ Health Perspect.* 2005;113(2):225-32.

19. Tuthill RW, Giusti, R.A., Moore, G.S., Calabrese, E.J. . 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 adverse birth outcomes in Taiwan. *Toxicology.* 2004;198(1-3):249-54.

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(9):850-62.

22. Hinckley AF, Bachand AM, Reif JS. Late pregnancy exposures to disinfection by-products and growth-related birth outcomes. *Environ Health Perspect.* 2005;113(12):1808-13.

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(1):38-47.

24. Wright JM, Schwartz J, Dockery DW. Effect of trihalomethane exposure on fetal development. *Occup Environ Med.* 2003;60(3):173-80.

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(8):765-8.

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.

27. Källen BA, Robert E. Drinking water chlorination and delivery outcome-a registry-based study in Sweden. *Reprod Toxicol.* 2000;14(4):303-9.

28. Boorman GA. Drinking water disinfection byproducts: review and approach to toxicity evaluation. *Environ Health Perspect.* 1999;107 Suppl 1:207-17.

29. Hoffman CS. Drinking Water Disinfection By-Product Exposure and Fetal Growth. *Epidemiology.* 2008;19(5):729.

30. Infante-Rivard C. Drinking water contaminants, gene polymorphisms, and fetal growth. *Environ Health Perspect.* 2004;112(11):1213-6.

31. Porter CK, Putnam SD, Hunting KL, Riddle MR. The effect of trihalomethane and haloacetic acid exposure on fetal growth in a Maryland county. *Am J Epidemiol.* 2005;162(4):334-44.

32. Aggazzotti G, Righi E, Fantuzzi G, Biasotti B, Ravera G, Kanitz S, Barbone F, Sansebastiano G, Battaglia MA, Leoni V, Fabiani L, Triassi M, Sciacca S. Chlorination by-products (CBPs) in drinking water and adverse pregnancy outcomes in Italy. *J Water Health.* 2004;2(4):233-47.

33. Hoffman CS. Drinking Water Disinfection By-Product Exposure and Duration of Gestation. *Epidemiology.* 2008;19(5):738.

34. Lewis C, Suffet IH, Hoggatt K, Ritz B. Estimated effects of disinfection by-products on preterm birth in a population served by a single water utility. *Environ Health Perspect.* 2007;115(2):290-5.

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(8):920-5.

36. Swan SH, Waller K, Hopkins B, Windham G, Fenster L, Schaefer C, Neutra RR. A prospective study of spontaneous abortion: relation to amount and source of drinking water consumed in early pregnancy. *Epidemiology.* 1998;9(2):126-33.

37. Waller K, Swan SH, DeLorenze G, Hopkins B. Trihalomethanes in drinking water and spontaneous abortion. *Epidemiology.* 1998;9(2):134-40.

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(6):522-31.

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(2):179-86.

40. King WD, Dodds L, Allen AC. Relation between stillbirth and specific chlorination by-products in public water supplies. *Environ Health Perspect.* 2000;108(9):883-6.

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(10):1056-61.

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(2):103-24.

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(2):185-205.

44. Hwang BF, Jaakkola JJ. Water chlorination and birth defects: a systematic review and meta-analysis. *Arch Environ Health.* 2003;58(2):83-91.

45. Nieuwenhuijsen MJ, Martinez D, Grellier J, Bennett JB, Best NG, Iszatt N, Vrijheid M, MB T. Exposure to disinfection by-products and congenital malformations – a review and meta-analysis. *Environ Health Perspect.* 2009 (In press).

46. Egger M, Davey Smith, G., Altman, D.G. Systematic Reviews in Health Care: Meta-Analysis in Context. London, UK.: BMJ Publishing Group, 2001.

47. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet.* 1999;354(9193):1896-900.

48. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of Observational Studies in Epidemiology A Proposal for Reporting. Vol. 283 Am Med Assoc, 2000;2008-2012.

49. Whitaker H, Nieuwenhuijsen MJ, Best N, Fawell J, Gowers A, Elliot P. Description of trihalomethane levels in three UK water suppliers. *J Expo Anal Environ Epidemiol.* 2003;13(1):17-23.

50. Greenland S, Longnecker M. Methods for Trend Estimation from Summarized Dose-Response Data, with Applications to Meta-Analysis. *Am J Epidemiol.* 1992;135(11):1301-9.

51. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2008.

52. Key J, Hodgson S, Omar RZ, Jensen TK, Thompson SG, Boobis AR, Davies DS, Elliott P. Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues. *Cancer Causes Control.* 2006;17(6):759-70.

53. Peters JL, Rushton, L., Sutton, A.J., Jones, D.R., Abrams, K.R., Mugglestone, M.A. Bayesian methods for the cross-design synthesis of epidemiological and toxicological evidence. *Applied Statistics.* 2005;54(1):159–172.

54. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7109):629-634.

55. EC Directive: 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. European Union, 1998.

56. USEPA. National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts, Final Rule, 40 CGR part 9. 141 and 142;. 1998.

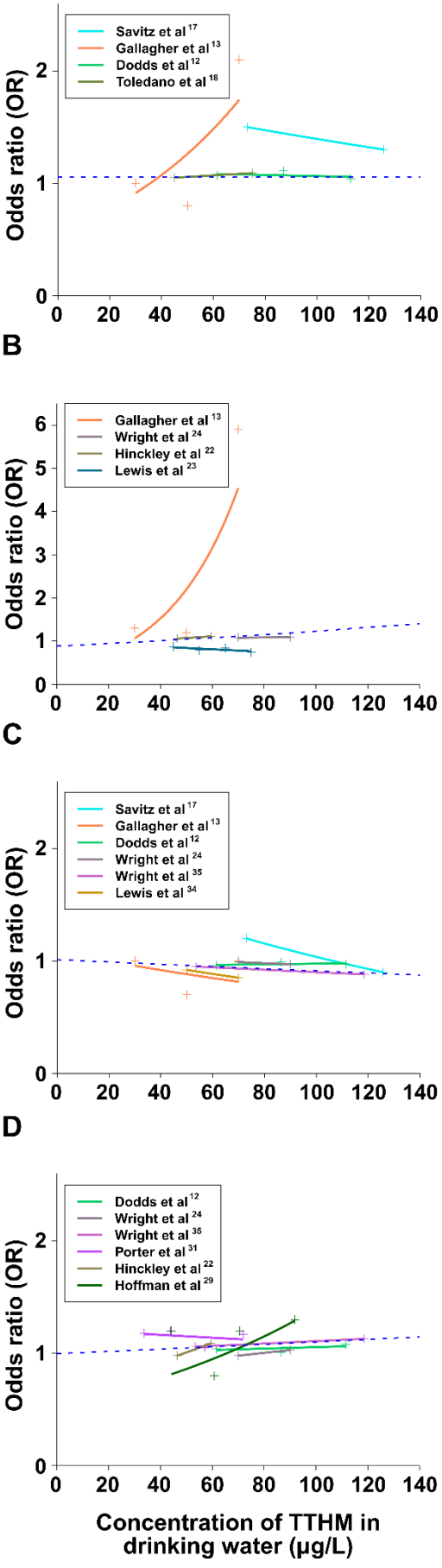
57. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-60.

58. Gavaghan DJ. An evaluation of homogeneity tests in meta-analyses in pain using simulations of individual patient data. *Pain.* 2000;85(3):415.

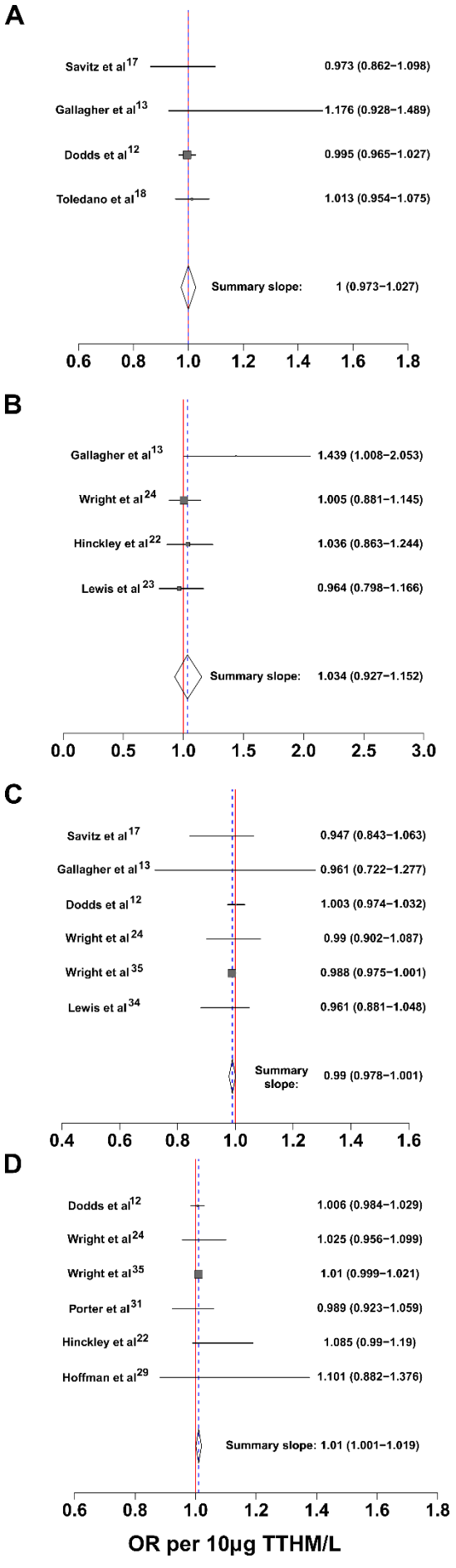
59. Hedges LV, Olkin, I. Statistical methods for meta-analysis. Orlando, FL: Academic Press, 1985.

60. 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(2):193-206.

# FIGURES




**Figures 1a-d** – Plots of individual study slopes (solid coloured lines) and the random-effects regression slope (dashed grey line) (both per 10µg/L TTHM) estimated from these for third trimester exposure to TTHM only, for (a) LBW, (b) TLBW, (c) PTD, and (d) SGA. Crosses indicate midpoints of exposure categories *vs.* OR in that category.




**Figures 2a-d** – Forest plots of OR slopes per 10µg/L TTHM for third trimester exposure to TTHM only for (a) LBW, (b) TLBW, (c) PTD, and (d) SGA. Study OR slopes are plotted with squares sized proportionally to their weight in the meta-analysis regression; horizontal lines indicate 95% CIs on these slopes. The red vertical line indicates no effect i.e. OR slope =1.0. The blue dashed line is the summary OR slope, with the tips of the diamond indicating 95% CIs around this estimate.

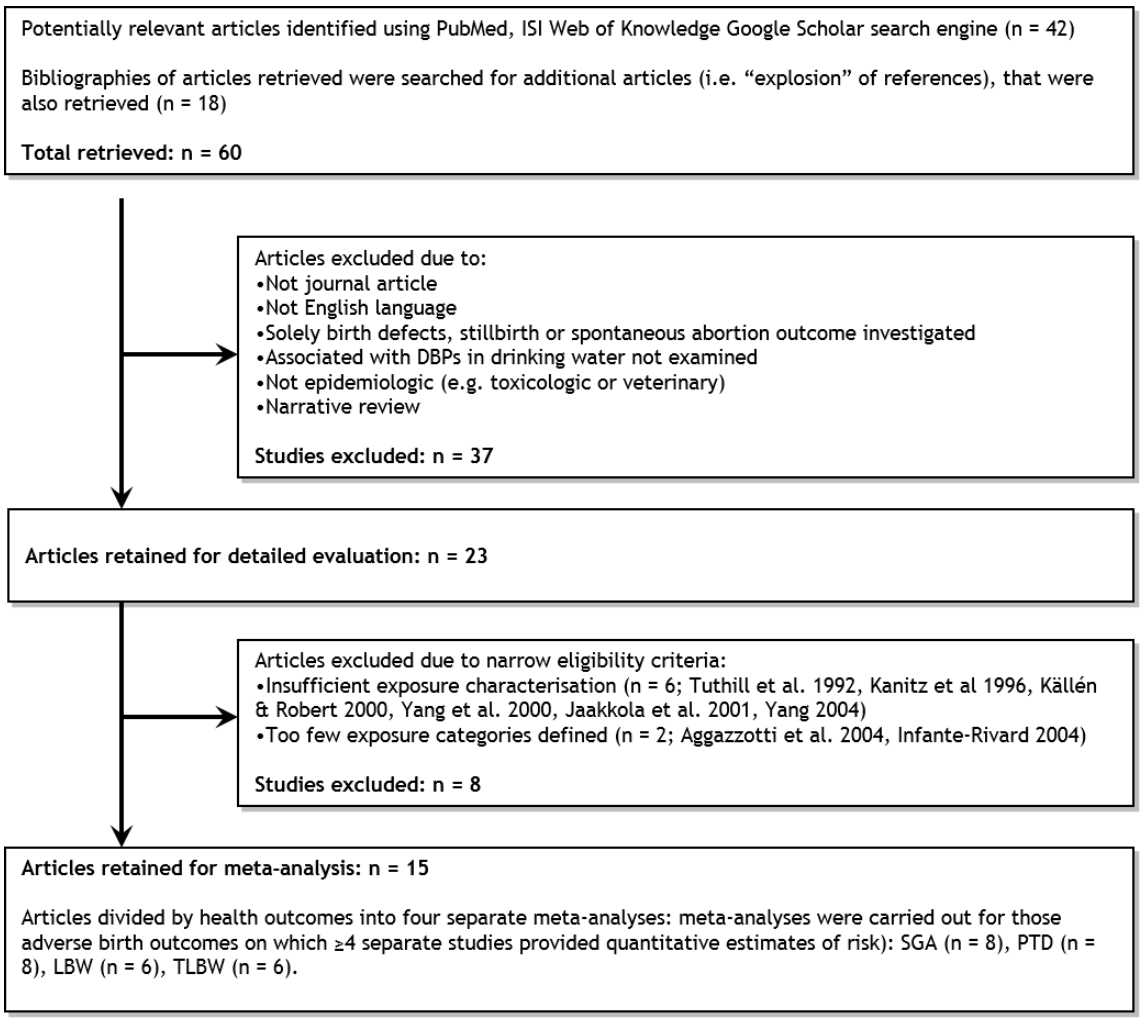



**Figure 3** –Summary QUOROM diagram showing how studies were identified and selected for inclusion.

# TABLES

## Table 1a – Characteristics of cohort studies included in the meta-analysis

| Author(s), year and location of study | Study period | Study design | Live birth outcomes investigated | Live birth outcomes relevant to meta-analysis | Outcome measure used | 1. Exposure data for TTHM and/or other chlorination by-products | 2. Exposure location | 3. Exposure timing |
|---|---|---|---|---|---|---|---|---|
| Bove et al. 1995<br><br>New Jersey, USA | 1985-1988 | Retrospective (cross-sectional) cohort | BW<br>PTD (n=7167)<br>SGA (n=4082)<br>TLBW (n=1853)<br>vLBW (n=905)<br><br>Total of 80,938 live births | PTD<br>SGA<br>TLBW (not reported in detail)<br>vLBW | OR | 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. 1998<br><br>Colorado, USA | 1990-1993 | Retrospective cohort | LBW (n=72)<br>TLBW (n=29)<br>PTD (n=68)<br><br>Total of 1244 births for which THM data available | LBW<br>TLBW<br>PTD | OR | Routine monitoring TTHM values from 1 year prior to study. | Individual exposures assigned according to maternal residence within a municipality by way of hydraulic modelling of TTHM values in the distribution network. | Exposure score estimated for third trimester. |
| Dodds et al. 1999<br><br>Nova Scotia, Canada | 1988-1995 | Retrospective cohort | SGA (n=4673)<br>LBW (n=2392)<br>vLBW (n=342)<br>PTD (n=2689)<br><br>Total of 49842 births for which TTHM data available | LBW<br>SGA<br>PTD | Relative Risk (RR) | Routine monitoring TTHM values and linear regression modelling 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. 2003<br><br>Massachusetts, USA | 1990 | Retrospective cohort | TLBW (n=1325)<br>SGA (n=5310)<br>PTD (n=3173)<br>BW | TLBW<br>SGA<br>PTD | OR | Maternal THM exposure for the 3rd trimester estimated from the quarterly average TTHM concentration | Maternal town of residence | Exposures estimated for all three trimesters and a total pregnancy average. Trimester |

| Author(s), year and location of study | Study period | Study design | Live birth outcomes investigated | Live birth outcomes relevant to meta-analysis | Outcome measure used | 1. Exposure data for TTHM and/or other chlorination by-products | 2. Exposure location | 3. Exposure timing |
|---|---|---|---|---|---|---|---|---|
| | | | Total of 56513 singleton infants | | | | | specific and pregnancy average exposures were assigned based on the month of birth |
| Wright et al. 2004<br><br>Massachusetts, USA | 1995-1998 | Retrospective cohort | BW<br>GA<br>SGA (n=17359)<br>PTD (n=11580)<br><br>Total of 196000 residents with singleton infants | SGA<br>PTD<br>BW | OR | Town-specific TTHM aggregate data based on quarterly monitoring samples | Maternal ZIP code at birth and infant month of birth used to assign third-trimester specific exposure data | Exposure estimated for third trimester (extrapolated from estimated exposure at month of birth)<br><br>Also averaged over whole pregnancy |
| Hinckley et al. 2005<br><br>Arizona, USA | | Retrospective cohort | IUGR (n=4346)<br>TLBW (n=1010)<br>PTD (n=4008)<br>vPTD (n=568)<br><br>Total of 48119 live births and fetal deaths | IUGR<br>TLBW<br>PTD<br>vPTD | OR | 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)<br><br>(For preterm and very preterm exposure only evaluated for specific time windows) |
| Porter et al. 2005<br><br>Maryland, USA | 1998-2002 | Retrospective cohort | IUGR<br><br>Total of 15315 births | IUGR | OR | Monthly TTHM concentrations for four sampling points in study area. | Maternal residence (zip code at birth) | Exposure estimated for both entire pregnancy and each trimester |
| Toledano et al. 2005<br><br>UK | 1992-1998 | Retrospective cohort | BW<br>LBW<br>vLBW<br>SB | LBW<br>vLBW | OR | Weighted average of modelled quarterly TTHM estimates for last 93 days before birth | Exposure modelled for each water zone using maternal residence (by | Exposure estimated for third trimester |

| Author(s), year and location of study | Study period | Study design | Live birth outcomes investigated | Live birth outcomes relevant to meta-analysis | Outcome measure used | 1. Exposure data for TTHM and/or other chlorination by-products | 2. Exposure location | 3. Exposure timing |
|---|---|---|---|---|---|---|---|---|
| | | | Total of approximately 1000000 birth records | | | | postcode) at time of birth | |
| Yang et al. 2007<br><br>Taiwan | 2000-2002 | Retrospective cohort | TLBW (n=2766)<br>SGA (n=8938)<br>PTD (n=2818)<br><br>Total of 90848 singleton births | TLBW<br>SGA<br>PTD | OR | Exposure established as an average of TTHM monitoring data over a two-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. 2008a<br><br>3 locations USA | 2000-2004 | Community-based prospective cohort study | SGA = 113<br><br>Total of 1958 live births (restricted to those born at 37 weeks 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. 2008b<br><br>3 locations USA | 2000-2004 | Community-based prospective cohort study | PTD = 185<br><br>Total of 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 HAA5 | Reported on second trimester |

## Table 1b – Characteristics of case-control studies included in the meta-analysis

| Author(s), year and location of study | Study period | Study design | Live birth outcomes investigated | Live birth outcomes relevant to meta-analysis | Outcome measure used | 1. Exposure data for TTHM and/or other chlorination by-products | 2. Exposure location | 3. Exposure timing |
|---|---|---|---|---|---|---|---|---|
| Kramer et al. 1992<br><br>Iowa, USA | 1989-1990 | Population-based case-control | IUGR (n=187)<br>LBW (n=159)<br>PTD (n=342) | IUGR<br>LBW<br>PTD | Odds ratio (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. 1995<br><br>North Carolina, USA | 1988-1989, 1988-1991 | Population-based case-control | PTD (586)<br>LBW (464)<br>MC (418) | LBW<br>MC<br>PTD | OR | Quarterly average THM values recorded by appropriate water supplier. | Women's addresses used to assign them to one of five public water supplies. | Dates of pregnancy used to assign reported nearest average quarterly THM value to each woman. Assignment time varied between outcomes: MC (4th week), PTD/LBW (28th week). |
| Lewis et al. 2006 | 1999-2001 | Population-based case-control | TLBW<br><br>Total of 36,529 births | TLBW | OR | Weekly TTHM monitoring data from four sampling sites | Each birth assigned to based on maternal residence at birth (from birth certificate) | TTHM exposure estimate calculated for each gestational period. (all trimesters). Pregnancy average also calculated |
| Lewis et al. 2007<br><br>Massachusetts, USA | 1999-2001 | Population-based case-control | PTD (n=2813)<br><br>Total of 37498 singleton births | PTD | Hazard Ratio (HR) | Weekly TTHM monitoring data from four 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 1st, 2nd trimesters, and 4-week risk sets, and pregnancy average) |

## Table 2a – Summary of published measures of effect (and their exposure characterisation) as used in the meta-analysis for LBW

| 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. 1992 | <2500g | Chloroform | OR (adjusted) | Entire pregnancy | *Non-detect* [a] | *1-9µg/L* | *≥ 10µg/L* | - |
| | | | | | 1.0 (reference) | 1.1 (0.7,1.6) | 1.3 (0.8,2.2) | - |
| Savitz et al. 1995 | <2500g | TTHM | OR (adjusted) | Third trimester | *40.8-63.3ppb* | *63.4-82.7ppb* | *82.8-168.8ppb* | - |
| | | | | | 1.0 (reference) | 1.5 (1.0,2.3) | 1.3 (0.8,2.1) | - |
| Gallagher et al. 1998 | ≤ 5 pounds, 8 ounces (≤ 2495g) | TTHM | OR (adjusted) | Third trimester | *≤ 20ppb* | *21-40ppb* | *41-60ppb* | *≥ 60ppb* |
| | | | | | 1.0 (reference) | 1.0 (0.6,1.8) | 0.8 (0.3,1.7) | 2.1 (1.0,4.8) |
| Dodds et al. 1999 | <2500g | TTHM | Relative risk (adjusted) | Third trimester | *0-49µg/L* | *50-74µg/L* | *75-99µg/L* | *≥ 100 µg/L* |
| | | | | | 1.00 (reference) | 1.07 (0.97,1.19) | 1.11 (0.97,1.26) | 1.04 (0.92,1.18) |
| Toledano et al. 2005 | <2500g | TTHM | OR (adjusted) | Third trimester | *<30µg/L* | *30-59µg/L* | *≥ 60µg/L* | - |
| | | | | | 1.00 (reference) | 1.05 (0.96,1.15) | 1.09 (0.93,1.27) | - |

a Where the exposure category was defined in Kramer et al. (1992) as below the limit of detection, it was assumed to be <1 µg/L

Notes:

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.

## Table 2b – Summary of published measures of effect (and their exposure characterisation) as used in the meta-analysis for TLBW

| Study | Definition of outcome | Exposure agent | Measure of effect | Exposure timing | Exposure categories | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Measures of effect | | | | |
| Gallagher et al. 1998 | ≥ 37 weeks of gestation and ≤ 5 pounds, 8 ounces (≤ 2495g) | TTHM | OR (adjusted) | Third trimester | *≤ 20ppb* | *21-40ppb* | *41-60ppb* | *≥ 61ppb* | |
| | | | | | 1.00 (reference) | 1.3 (0.5,3.3) | 1.2 (0.4,4.0) | 5.9 (2.0,17.0) | |
| Hinckley et al. 2005 | ≥ 37 completed weeks of gestation and weighing <2,500 g | TTHM | OR (adjusted) | Third trimester | *<40µg/L* | *40-53µg/L* | *>53µg/L* | - | - |
| | | | | | 1.00 (reference) | 1.06 (0.89,1.25) | 1.11 (0.94,1.31) | 1.00 (1.00,1.01) | - |
| Wright et al. 2003 | <2500g among term births | TTHM | OR (adjusted) | Third trimester, entire pregnancy average | *0-60µg/L* | *>60-80µg/L* | *>80µg/L* | - | - |
| | | | | | 1.00 (reference) | 1.08 (0.90,1.31)<br>0.97 (0.81,1.26) | 1.09 (0.91,1.31)<br>1.05 (0.85,1.29) | - | - |
| Lewis et al. 2006 | Birth weight <2,500g and born after 36 weeks of gestation | TTHM | OR (adjusted) | Third trimester, entire pregnancy average | *<40µg/L* | *40≤ 50µg/L* | *50≤ 60µg/L* | *60≤ 70µg/L* | *≥ 70µg/L* |
| | | | | | 1.00 (reference) | 0.87 (0.60,1.26)<br>1.07 (0.81,1.42) | 0.79 (0.56,1.12)<br>0.95 (0.75,1.20) | 0.84 (0.58,1.21)<br>1.23 (0.92,1.64) | 0.74 (0.44,1.22)<br>N/a |
| Yang et al. 2007 | 437 gestational weeks and <2500 gm | TTHM | OR (adjusted) | Entire pregnancy average | *<4.93µg/L* | *4.93-13.11g/L* | *>13.11g/L* | - | - |
| | | | | | 1.00 (reference) | 0.99 (0.90,1.08) | 1.00 (0.91,1.10) | - | - |

Notes:

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.

## Table 2c – Summary of published measures of effect (and their exposure characterisation) as used in the meta-analysis for PTD

| Study | Definition of outcome | Exposure agent | Measure of effect | Exposure timing | Exposure categories | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Measures of effect | | | | |
| Kramer et al. 1992 | <37weeks of gestation | Chloroform | OR (adjusted) | Entire pregnancy | *Non-detect* | *1-9µg/L* | *≥ 10µg/L* | - | - |
| | | | | | 1.0 (reference) | 1.1 (0.8, 1.4) | 1.1 (0.7, 1.6) | - | - |
| Savitz et al. 1995 | <37 weeks completed gestation | TTHM | OR (adjusted) | Third trimester | *40.8-63.3ppb* | *63.4-82.7ppb* | *82.8-168.8ppb* | - | - |
| | | | | | 1.0 (reference) | 1.2 (0.8, 1.8) | 0.9 (0.6, 1.5) | - | - |
| Gallagher et al. 1998 | <37 weeks completed gestation | TTHM | OR (adjusted) | Third trimester | *≤ 20ppb* | *21-40ppb* | *41-60ppb* | *>61ppb* | - |
| | | | | | 1.0 (reference) | 1.0 (0.6, 1.7) | 0.7 (0.3, 1.6) | 1.0 (0.3,2.8) | - |
| Dodds et al. 1999 | <37 weeks of gestation | TTHM | Relative risk (adjusted) | Third trimester | *0-49µg/L* | *50-74µg/L* | *75-99µg/L* | *>100µg/L* | - |
| | | | | | 1.00 (reference) | 0.96 (0.88, 1.06) | 0.99 (0.88, 1.12) | 0.97 (0.87, 1.09) | - |
| Wright et al. 2003 | <37 gestational weeks | TTHM | OR (adjusted) | Third trimester, entire pregnancy average | *0-60µg/L* | *>60-80µg/L* | *>80µg/L* | - | - |
| | | | | | 1.00 (reference) | 0.99 (0.87, 1,13)<br>1.00 (0.89, 1.12) | 0.97 (0.85, 1.11)<br>0.90 (0.77, 1.04) | - | - |
| Wright et al. 2004 | <37 gestational weeks | TTHM | OR (adjusted) | Third trimester | *0-33µg/L* | *>33-74µg/L* | *>74-163µg/L* | - | - |
| | | | | | 1.00 (reference) | 0.95 (0.91, 0.99) | 0.88 (0.81, 0.94) | - | - |
| Lewis et al. 2007 | Analyses were restricted to infants between 32 and 45 gestational weeks with a birth weight 500g-5000g | TTHM | HR (adjusted) | Entire pregnancy average | *<40* | *40<60µg/L* | *≥ 60µg/L* | - | - |
| | | | | | 1.00 (reference) | 0.92 (0.82, 1.02) | 0.85 (0.74, 0.97) | - | - |
| Yang et al. 2007 | <37 gestational weeks | TTHM | OR (adjusted) | Entire pregnancy average | *<4.93µg/L* | *4.93-13.11g/L* | *>13.11g/L* | - | - |
| | | | | | 1.00 (reference) | 1.03 (0.94, 1.13) | 1.08 (0.98, 1.18) | - | - |
| Hoffman et al. 2008b | Delivery before 37 weeks' gestation | TTHM | Risk ratio (adjusted) | Second trimester (assumed as entire pregnancy average in meta-analysis) | *2.2-4.6µg/L*[a] | *33.1-55.0µg/L* | *55.1-66.3µg/L* | *66.4-74.8µg/L* | *74.9-108.8µg/L* |
| | | | | | 1.00 (reference) | 0.89 (0.50, 1.30) | 0.90 (0.60, 1.40) | 0.70 (0.40, 1.10) | 0.50 (0.30, 0.90) |

a Exposure range at low DBP site; not contiguous with upper exposure categories but used as referent category. Notes: 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.

## Table 2d – Summary of published measures of effect (and their exposure characterisation) as used in the meta-analysis for SGA

| Study | Definition of outcome | Exposure agent | Measure of effect | Exposure timing | Exposure categories | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Measures of effect | | | | | |
| Kramer et al. 1992 [a] | <5th percentile of weight for gestational age (excluding unrealistic gestational ages of ≤ 22weeks or ≥ 46 weeks) | Chloroform | OR (adjusted) | Entire pregnancy | *Non-detect* | *1-9µg/L* | *≥ 10µg/L* | - | - | - |
| | | | | | 1.0 (reference) | 1.3 (0.9, 1.8) | 1.8 (1.1, 2.9) | - | - | - |
| Bove et al. 1995 | Live births at or below their race-, sex-, and gestational week-specific 5th percentile weight | TTHM | OR (unadjusted/ adjusted [b]) | Entire pregnancy average | *≤ 20ppb* | *>20-40ppb* | *>40-60ppb* | *>60-80ppb* | *>80-100ppb* | *>100ppb* |
| | | | | | 1.00 (reference) | 0.98 (0.84, 1.15) [d] | 1.33 (1.23, 1.44) [d] | 1.11 (1.01, 1.22) [d] | 1.22 (1.07, 1.39) [d] | 1.50 (1.15, 1.96) [d] |
| Dodds et al. 1999 | Lowest 10th of weight distribution among Canadian live births for each week of gestation for each sex | TTHM | Relative risk (adjusted) | Third trimester | *0-49µg/L* | *50-74µg/L* | *75-99µg/L* | *100 or more µg/L* | - | - |
| | | | | | 1.00 (reference) | 1.04 (0.97, 1.11) | 1.01 (0.92, 1.11) | 1.08 (0.99, 1.18) | - | - |
| Wright et al. 2003 | Lowest 10th centile of birth weight for each gestational week stratified by infant gender and maternal race (only term births) | TTHM | OR (adjusted) | Third trimester, entire pregnancy average | *0-60µg/L* | *>60-80µg/L* | *>80µg/L* | - | - | - |
| | | | | | 1.00 (reference) | 0.98 (0.89, 1.09)<br>1.00 (0.92, 1.09) | 1.03 (0.94, 1.14)<br>1.14 (1.02, 1.26) | - | - | - |
| Wright et al. 2004 | Lowest decile of birth weight for each gestational week stratified by infant sex and maternal race (only term births from 37-45 gestational weeks) | TTHM | OR (adjusted) | Third trimester | *0-33µg/L* | *>33-74µg/L* | *>74-163µg/L* | - | - | - |
| | | | | | 1.00 (reference) | 1.06 (1.02, 1.10) | 1.13 (1.07, 1.20) | - | - | - |
| Hinckley et al. 2005 [a] | Birth weight <lowest 10th percentile of birth weights by race, ethnicity, and gestational age (term or preterm) | TTHM | OR (adjusted) | Third trimester | *<40µg/L* | *40-53µg/L* | *≥ 53* | - | - | - |
| | | | | | 1.00 (reference) | 0.98 (0.90, 1.07) | 1.09 (1.00, 1.18) | - | - | - |

| Study | Definition of outcome | Exposure agent | Measure of effect | Exposure timing | Exposure categories | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Measures of effect | | | | | |
| Porter et al. 2005 [a] | Birth weight <10th percentile for gestational age (adjusted for sex and race) | TTHM | OR (adjusted) | Third trimester, entire pregnancy average | *≤ 29.48µg/L*<br>*≤ 37.38µg/L* | *>29.48µg/L ≤ 38.00*<br>*>37.38µg/L ≤ 43.67µg/L* | *>38.00µg/L ≤ 49.67µg/L*<br>*>43.67µg/L ≤ 49.13µg/L* | *>49.67µg/L ≤ 64.48µg/L*<br>*>49.13µg/L ≤ 59.00µg/L* | *>64.48µg/L*<br>*>59.00µg/L* | - |
| | | | | | 1.00 (reference) | 1.18 (0.97, 1.44)<br>1.01 (0.83, 1.22) | 1.20 (0.99, 1.46)<br>1.11 (0.92, 1.35) | 1.05 (0.86, 1.29)<br>0.98 (0.81, 1.20) | 1.17 (0.96, 1.42)<br>0.98 (0.81, 1.19) | - |
| Yang et al. 2007 | Birth weight ≤ 10th percentile for each gestational week stratified by infant sex | TTHM | OR (adjusted) | Entire pregnancy average | *<4.93µg/L* | *4.93-13.11g/L* | *>13.11g/L* | - | - | - |
| | | | | | 1.00 (reference) | 1.00 (0.94, 1.04) | 0.96 (0.91, 1.02) | - | - | - |
| Hoffman et al. 2008a | Birth weight <10th percentile for gestational age at birth, sex, maternal race and parity | TTHM | Risk ratio (adjusted) | Third trimester | *2.2-4.6µg/L* [c] | *33.1-55.0µg/L* | *55.1-66.3µg/L* | *66.4-74.8µg/L* | *74.9-108.8µg/L* | - |
| | | | | | 1.0 (reference) | 1.2 (0.7, 2.2) | 0.8 (0.4, 1.5) | 1.2 (0.7, 2.2) | 1.3 (0.7, 2.3) | - |

a Studies reporting on "IUGR" using definitions considered to be equivalent to SGA. b Adjusted results given when difference between adjusted & unadjusted >15% of unadjusted value. c Exposure range at low DBP site; not contiguous with upper exposure categories but used as referent category. d Quoted 95% CIs recalculated from data published at different levels of confidence Notes: 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.

**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 | Number of studies included | OR slope per 10µg/L | Lower 95% CI | Upper 95% CI | Q-statistic |
|---|---|---|---|---|---|---|---|
| *Only TTHM* | *Third trimester* | LBW | 4 | 0.9999 | 0.9735 | 1.0270 | 2.244 |
| | | TLBW | 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 | 0.974681 | 1.0286 | 2.495 |
| | | TLBW | 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 |
| | *Entire pregnancy only* | SGA | 4 | 1.0105 | 0.9712 | 1.0514 | 4.659 |
| *TTHM and chloroform* | *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 only* | PTD | 4 | 0.9696 | 0.9139 | 1.0286 | 1.441 |