Impact of chloroform exposures on reproductive and developmental outcomes: A systematic review of the scientific literature

Amy Lavin Williams1 | Christopher A. Bates1 | Nelson D. Pace2 | Megan J. Leonhard3 | Ellen T. Chang4 | John M. DeSesso1,5

1Exponent, Inc., Alexandria, Virginia
2Exponent, Inc., Oakland, California
3Exponent, Inc., Bellevue, Washington
4Exponent, Inc., Menlo Park, California
5Georgetown University School of Medicine, Washington, District of Columbia

Correspondence
Amy Lavin Williams, Exponent, Inc. 1800 Diagonal Road, Suite 1800, Alexandria, VA 22314.
Email: awilliams@exponent.com

Funding information
American Chemistry Council

Aims: We assessed the animal and epidemiological data to determine if chloroform exposure causes developmental and/or reproductive toxicity.

Results and Discussion: Initial scoping identified developmental toxicity as the primary area of concern. At levels producing maternal toxicity in rats and mice, chloroform caused decrements in fetal weights and associated delays in ossification. In a single mouse inhalation study, exposure to a high concentration of chloroform was associated with small fetuses and increased cleft palate. However, oral exposure of mice to chloroform at a dose 4 times higher was negative for cleft palate; multiple inhalation studies in rats were also negative. Epidemiologic data on low birth weight and small for gestational age were generally equivocal, preventing conclusions from being drawn for humans. The animal data also show evidence of very early (peri-implantation) total litter losses at very high exposure levels. This effect is likely maternally mediated rather than a direct effect on the offspring. Finally, the epidemiologic data indicate a possible association of higher chloroform exposure with lower risk of preterm birth (<37 weeks gestation).

Conclusions: The available animal data suggest that exposures lower than those causing maternal toxicity should be without developmental effects in the offspring. Also, most studies in humans rely on group-level geographic exposure data, providing only weak epidemiologic evidence for an association with development outcomes and fail to establish a causal role for chloroform in the induction of adverse developmental outcomes at environmentally relevant concentrations.

KEYWORDS
birth defects, chloroform, developmental toxicity, epidemiology, teratogenicity

1 | INTRODUCTION

Chloroform, also known as trichloromethane or CHCl3, is a colorless, sweet-smelling, volatile liquid at room temperature (Agency for Toxic Substances and Disease Registry (ATSDR), 1997; National Toxicology Program (NTP), 2016; Office of Environmental Health Hazard Assessment (OEHHA), 2016a; World Health Organization (WHO), 2004). In the past, chloroform was used as an inhaled anesthetic, but this application has been largely phased out worldwide. Today, chloroform is commonly used in the production of various other chemicals, some of the most notable being fluorinated refrigerants and Teflon coatings. Chloroform also occurs as a byproduct of paper bleaching and is the most prevalent trihalomethane produced in water disinfection processes, which are regulated by the U.S Environmental Protection Agency (EPA). As described by the EPA, the routine use of chlorination to disinfect public water
supplies since the early 1900s has led to a substantial decline in the incidence of waterborne diseases in the United States and worldwide (EPA, 1999). The levels of chloroform present in water as a result of disinfection processes depend on the existing chlorine and organic matter concentrations as well as on the pH and temperature of the water. Because of its physicochemical properties (particularly, its high volatility) and varied uses, human exposure to chloroform can occur via multiple routes of exposure. Releases to the air and water may result due to the chemical’s use in various industrial manufacturing processes. Oral exposures may occur through consumption of chlorinated or chloraminated drinking water, as well as inhalation (and minor dermal) exposures due to showering and bathing. In the United States, the normal concentration of chloroform in air ranges from 0.02 to 0.05 parts per billion (ppb), and that in treated drinking water ranges from 2 to 44 ppb (ATSDR, 1997).

Chloroform has been previously evaluated for its potential to cause developmental and reproductive toxicity. In 2001, the EPA evaluated the available database of studies for chloroform, and concluded that although effects on reproduction and developmental were observed to occur in animals, “these effects occur at the same or higher doses as those that cause effects on the dam … suggesting that most of the effects are secondary to maternal toxicity” (EPA, 2001). Likewise, the WHO in 2004, evaluating much of the same data as the EPA, concluded that there were no indications of teratogenicity associated with chloroform exposure, and effects on the developing organism were seen only at doses shown to be maternally toxic. As such, neither organization has expressed significant concerns regarding developmental or reproductive toxicity due to chloroform exposure. In contrast, the State of California’s OEHHA has listed chloroform as a reproductive toxicant under Proposition 65 since 2009 (OEHHA, 2016a).

The purpose of this analysis is to conduct a systemic assessment for persons of reproductive potential (population of interest) exposed to chloroform via either ingestion or inhalation (exposures of concern)—comparing those highly exposed to those either not exposed or minimally exposed (comparators)—for developmental and/or reproductive toxicity (outcomes). We thus conducted an assessment of the animal and epidemiological data for chloroform to determine if the available evidence indicates that chloroform causes developmental and/or reproductive toxicity. As a first step in this effort, a scoping exercise was undertaken to identify the primary developmental/reproductive concern(s) for chloroform (i.e., female reproductive toxicity, male reproductive toxicity, or developmental toxicity). Based on the results of this scoping effort, a more focused analysis was conducted on the developmental toxicity potential of chloroform. This evaluation involved separate analyses of the animal and epidemiological studies, followed by an integrated assessment of the data. We also compared the results of our assessment with the conclusions of the EPA, WHO, and OEHHA.

2 | METHODS

2.1 | Literature search

We conducted a search of the published scientific literature for human or animal studies that involved chloroform exposures and assessment of developmental or reproductive outcomes (male or female) using the PubMed database, available through the National Center for Biotechnology Information (NCBI) at the U.S. National Library of Medicine (NLM) (https://www.ncbi.nlm.nih.gov/pubmed). Human epidemiological studies examining the relationship between chloroform exposure and developmental toxicity, male reproduction, or female reproduction were identified on May 2, 2017 using the following PubMed query:

(chloroform OR disinfection OR trihalomethane*) AND ("birth defect" OR "birth defects" OR hypospadia* OR "birth weight" OR birth-weight OR birth* OR fetal OR fetus OR pre-term OR pre-term OR spontaneous OR abortion* OR stillbirth* OR miscarriage* OR gestation* OR fertil* OR fecund* OR menstrual OR sperm OR semen).

Animal toxicity studies of chloroform exposure on developmental outcomes, male reproduction, or female reproduction were identified on May 23, 2017, using the following PubMed query:

(chloroform) AND ("reproduction" OR “gestation” OR “lactation” OR “teratogen” OR “prenatal” OR “postnatal”).

Search results were winnowed to a body of potentially relevant studies through a review of titles and abstracts. These were then compared against the body of studies reviewed in the August 2016 OEHHA assessment of chloroform for listing as a reproductive toxicant under the State of California Proposition 65 legislation (OEHHA, 2016a, 2016b) to ensure that no relevant studies previously reviewed by OEHHA were excluded or missed. The outcome of this literature search is discussed in greater detail below in the Results section.

2.2 | Scoping exercise

To narrow our assessment to those outcomes of greatest concern for a more extensive analysis, we undertook a scoping exercise whereby we reviewed the full-text articles of the epidemiologic and animal studies identified in our literature search to assess the primary developmental and reproductive adverse outcomes associated with chloroform exposure. The
results of this scoping effort, which identified developmental outcomes as the area of primary focus, are described below under the Results section.

## 2.3 Data quality assessment

The animal and human studies relevant to an assessment of developmental toxicity were subjected to assessments of data quality.

To assess the quality of the animal developmental toxicity studies, the Toxicological data Reliability Assessment Tool (ToxRTool) was utilized (Schneider et al., 2009; https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/toxtool). The ToxRTool, developed by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), is a software-based tool that allows one to apply a set of assessment criteria to a study to determine the reliability of the study’s data in terms of both relevance and adequacy for regulatory decision-making. The tool was established based on the Klimisch reliability categorization system (Klimisch, Andreea, & Tillmann, 1997) and assigns a study to 1 of 3 possible categories (1 = reliable without restrictions; 2 = reliable with restrictions; or 3 = not reliable) based on the reported study methods and level of data documentation. Studies were evaluated by two authors in unison to assign a preliminary score. These scores were then reviewed by a third author and any discrepancies in scoring were discussed among the three authors to come to an agreement on the overall score assigned to each study.

To assess the quality of human epidemiologic studies of developmental outcomes, we considered the potential for selection, information, and confounding bias and their possible roles in influencing reported study results. This assessment was guided by the use of the revised Research Triangle Institute (RTI) item bank (Viswanathan, Berkman, Dryden, & Hartling, 2013). As an assessment tool, the RTI item bank provides a set of 13 questions used to identify confounding and evaluate the risk of bias in observational study designs; these questions were answered for the 30 observational studies considered (see Supporting Information Table S1). Independent evaluations were performed by two authors, and any rating discrepancies were resolved through discussion.

## 2.4 Data synthesis

With regard to the laboratory animal data, our analysis emphasized the strengths and weaknesses of the studies, the influence of potential confounding factors on the study results (e.g., the appetite-suppressive effect of chloroform inhalation and the influence of excessive maternal toxicity), and the consistency of the findings both across the animal toxicity database and with results from the epidemiological literature.

In our overall assessment, the selected epidemiologic studies were evaluated to determine whether they indicated a consistent pattern of association between exposure to chloroform and adverse developmental outcomes in humans. The quality and risk of bias in each epidemiologic study was evaluated based on the following criteria: study population definition and recruitment approach; validity and reliability of exposure, outcome, and confounder assessment; potential for confounding and other types of bias, including selection and information bias; appropriateness of the statistical analysis; and presentation, reporting, and interpretation of results. In addition, the overall weight of the epidemiologic evidence regarding a causal association was considered in light of the Bradford Hill guidelines (Hill, 1965), with particular emphasis on strength, consistency, temporality, biological gradient, and plausibility. The Hill guidelines were used with the understanding that none of the guidelines, except for a temporal sequence where by a exposure precedes an outcome, is strictly required to establish causality, and that plausibility “depends on the biological knowledge of the day” (Hill, 1965) and may sometimes be controversial.

## 3 RESULTS

### 3.1 Literature search

#### 3.1.1 Animal studies

The search of the literature for animal studies of chloroform exposures and developmental or reproductive outcomes produced 227 results as of May 23, 2017. A review of titles and abstracts identified 6 potentially relevant articles in addition to the 17 articles already reviewed in the 2016 OEHHA assessment that listed chloroform as a reproductive toxicant under the State of California Proposition 65 legislation (OEHHA, 2016a); these 23 articles underwent a full-text review. No relevant articles published after the OEHHA review (post-2015) were identified. Following full-text review, one article from the OEHHA review was excluded because it did not report specific doses. The remaining 16 studies from the OEHHA review were included; a single study not previously considered by OEHHA (Brown-Woodman et al., 1998) was also included. In total, 17 of 23 articles that underwent full-text review were included and 6 were excluded. Areas of focus for each of the studies identified in the literature search and reasons for exclusion of the six studies removed from assessment are listed in Table 1. In particular, studies that assessed exposure to trihalomethanes in general, but not exposed specifically to chloroform, and articles that provided no original or additional information such as perspective, commentary, or state of the science articles, were excluded.

#### 3.1.2 Epidemiologic studies

The search of the literature for human epidemiologic studies of chloroform exposures and developmental or reproductive outcomes produced 1,126 results as of May 2, 2017. From the 2016 OEHHA assessment of chloroform for listing as a
reproductive toxicant under the State of California Proposition 65 legislation (OEHHA, 2016a), 100 references were identified, 92 of which were duplicates of those found in the PubMed literature search. From the two sources, 1,134 articles were identified. A review of titles and abstracts identified 84 potentially relevant articles that underwent a full-text review. Following full-text review, 35 studies considered in the 2016 OEHHA report were included; 2 articles eligible for but omitted from the OEHHA report were added; and 5 articles were included that were first published after the completion of the OEHHA review. In total, 42 of 84 articles that underwent full-text review were included and 42 were excluded (Table 2). In particular, studies that assessed exposure to trihalomethanes in general, but not exposure specifically to chloroform, were excluded, with the exception of two studies (also included by OEHHA) that specified that chloroform comprised 89% of the total trihalomethanes (Lewis et al., 2006; Lewis et al., 2007).

### 3.2 Scoping effort

Of the 17 animal studies included for consideration in the scoping exercise, 14 focused on developmental toxicity endpoints (e.g., fetal survival, fetal body weight, and fetal abnormalities) (Baeder & Hofmann, 1988, 1991; Balster & Borzelleca, 1982; Borzelleca & Carchman, 1982; Brown-Woodman et al., 1998; Burkalter & Balster, 1979; Lim et al., 2004; Murray et al., 1979; Newell & Dilley, 1978; NTP, 1988; Ruddick et al., 1983; Schwetz et al., 1974; Teixidó et al., 2015; Thompson et al., 1974). Of these, 11 studies reported positive findings, including increased incidences of resorptions/abortions, reduced numbers of implantations, reduced fetal/pup body weights or crown-rump length, and reduced ossification. These effects were typically accompanied by indications of maternal toxicity, including decreased body weights and food consumption.

Three of the studies that addressed developmental toxicity also examined female reproductive endpoints (Balster & Borzelleca, 1982; Murray et al., 1979; NTP, 1988). Additionally, one subchronic study (Jorgenson & Rushbrook, 1980) and one carcinogenicity study (Heywood et al., 1979) were reviewed because these studies examined female reproductive organ histopathology following treatment. A two-generation reproductive study (Borzelleca & Carchman, 1982) reported reduced mating indices; in contrast, an NTP reproductive toxicity study (NTP, 1988) did not report adverse effects on female reproductive parameters. In addition, no toxicity to the female reproductive organs

### Table 1

| Study                              | Developmental | Female reproductive | Male reproductive | Include/exclude | Reason for exclusion |
|------------------------------------|---------------|---------------------|-------------------|-----------------|----------------------|
| Schwetz, Leong, and Gehring (1974) | X             |                     |                   | Include         |                      |
| Thompson, Warner, and Robinson (1974) | X             |                     |                   | Include         |                      |
| Newell and Dilley (1978)            |               |                     |                   | Include         |                      |
| Burkalter and Balster (1979)        |               |                     |                   | Include         |                      |
| Heywood et al. (1979)              | X             | X                   |                   | Include         |                      |
| Murray, Schwetz, McBride, and Staples (1979) | X             |                     |                   | Include         |                      |
| Jorgenson and Rushbrook (1980)     |               |                     |                   | Include         |                      |
| Land, Owen, and Linde (1981)       |               |                     |                   | Include         |                      |
| Balster and Borzelleca (1982)       | X             |                     |                   | Include         |                      |
| Borzelleca and Carchman (1982)     | X             | X                   |                   | Include         |                      |
| Ruddick, Villeneuve, Chu, and Valli (1983) | X             |                     |                   | Include         |                      |
| Nelson et al. (1986)               |               |                     |                   | Exclude         | Review article; no new information |
| Smith, Zenick, and George (1986)   | X             | X                   |                   | Exclude         | Review article; no new information |
| Baeder and Hofmann (1988)           | X             |                     |                   | Include         |                      |
| NTP (1988)                         |               | X                   | X                 | Include         |                      |
| Garcia-Estrada et al. (1990)*      | X             |                     |                   | Exclude         | In Spanish; no specific doses reported |
| Baeder and Hofmann (1991)           |               |                     |                   | Include         |                      |
| Chapin et al. (1997)               | X             | X                   | X                 | Exclude         | Summary, no new information |
| Brown-Woodman et al. (1998)        |               |                     |                   | Include         |                      |
| Lim, Stals, Petrik, Foster, and Holloway (2004) | X             |                     |                   | Include         |                      |
| Guariglia et al. (2011)             |               |                     |                   | Exclude         | Chemical mixture     |
| Narotsky et al. (2011)             | X             | X                   |                   | Exclude         | Chemical mixture     |
| Teixidó et al. (2015)              |               |                     |                   | Include         |                      |

*Reviewed in OEHAA (2016a) report but excluded herein.
| Reference | Include/exclude | Brief reason for exclusion |
|-----------|-----------------|---------------------------|
| Aggazzotti et al. (2004) | Exclude | Trihalomethanes analyzed |
| Ali et al. (2014) | Exclude | Haloacetic acids analyzed |
| Borzelleca and Carchman (1982) | Exclude | Non-human subjects |
| Botton et al. (2015) | Include | -- |
| Bove et al. (1995) | Exclude | Trihalomethanes analyzed |
| Brecher, Shinde, Zhan, Gong, and Langlois (2014) | Include | -- |
| Brown (2007) | Exclude | News article |
| Cao et al. (2016) | Include | -- |
| Chang, Lin, Hsu, and Guo (2001) | Include | -- |
| Chisholm, Cook, Bower, and Weinstein (2008) | Exclude | Trihalomethanes analyzed |
| Costet et al. (2012) | Include | -- |
| Dahl, Sundby, Hensten-Pettersen, and Jacobsen (1999) | Include | -- |
| Danileviciute, Grazuleviciene, Vencloviene, Paulauskas, and Nieuwenhuijsen (2012) | Include | -- |
| Dodds et al. (2004) | Include | -- |
| Dodds and King (2001) | Include | -- |
| Dodds, King, Woolcott, and Pole (1999) | Exclude | Trihalomethanes analyzed |
| Fenster et al. (2003) | Exclude | Trihalomethanes analyzed |
| Gallagher, Nuckols, Stallones, and Savitz (1998) | Exclude | Trihalomethanes analyzed |
| Grazuleviciene, Kapustinskiene, Vencloviene, Buinauskiene, and Nieuwenhuijsen (2013) | Include | -- |
| Grazuleviciene et al. (2011) | Include | -- |
| Hassani et al. (2014) | Exclude | Organic solvent mixture analyzed |
| Hinckley, Bachand, and Reif (2005) | Include | -- |
| Hoffman et al. (2008a) | Include | -- |
| Hoffman et al. (2008b) | Exclude | Trihalomethanes analyzed |
| Horton et al. (2011) | Exclude | Trihalomethanes analyzed |
| Hwang, Jaakkola, and Guo (2008) | Exclude | Trihalomethanes analyzed |
| Hwang and Jaakkola (2012) | Exclude | Trihalomethanes analyzed |
| Ileka-Priouzeau et al. (2015) | Exclude | Trihalomethanes analyzed |
| Infante-Rivard (2004) | Include | -- |
| Iszatt, Nieuwenhuijsen, Bennett, and Toledano (2014) | Include | -- |
| Iszatt et al. (2013) | Include | -- |
| Iszatt, Nieuwenhuijsen, Nelson, Elliott, and Toledano (2011) | Include | -- |
| Joyce et al. (2008) | Exclude | Trihalomethanes analyzed |
| Kanitz et al. (1996) | Exclude | Chlorine dioxide/sodium hypochlorite analyzed |
| King et al. (2005) | Exclude | Trihalomethanes analyzed |
| King, Dodds, and Allen (2000) | Include | -- |
| Klotz and Pyrch (1999) | Exclude | Trihalomethanes analyzed |
| Kogevinas et al. (2016) | Include | -- |
| Kramer, Lynch, Isachsen, and Hanson (1992) | Include | -- |
| Kumar et al. (2014) | Exclude | Trihalomethanes analyzed |
| Levallois et al. (2016) | Include | -- |
| Levallois et al. (2012) | Include | -- |
| Lewis, Suffet, Hoggatt, and Ritz (2007) | Include | -- |
| Lewis, Suffet, and Ritz (2006) | Include | -- |
| Luben, Nuckols, Mosley, Hobbs, and Reif (2008) | Exclude | Trihalomethanes analyzed |
| Luben et al. (2007) | Exclude | Trihalomethanes and haloacetic acids analyzed |
| MacLehose et al. (2008) | Exclude | Chlorinated and brominated disinfection by-products analyzed |
| Malcom et al. (1999) | Exclude | Review for risk assessment |
| Nieuwenhuijsen et al. (2009) | Exclude | Health impact assessment methods |

(Continues)
(e.g., ovaries and uterus) was reported among the subchronic/chronic studies reviewed (Heywood et al., 1979; Jorgenson & Rushbrook, 1980). Overall, the studies reporting these data were limited in number, but provided generally consistent findings.

Five studies investigated male reproductive outcomes (e.g., sperm quality and/or motility, fertility, and reproductive organ weights). One study reported a significant increase in abnormal spermatozoa following 5 days of inhalation exposure (Land et al., 1981), but the fertility of these animals was not investigated. Another study reported increased absolute epididymal weights and degeneration of the epididymal ductal epithelium, but sperm motility, count, and morphology were unaffected (NTP, 1988). Independent calculation of the relative (to body weight) epididymal weights found no difference between these two groups. No toxicity to the male reproductive organs (e.g., testes) was reported among the subchronic/chronic studies reviewed (Heywood et al., 1979; Jorgenson & Rushbrook, 1980) or the other reproductive toxicity study (Borzelleca & Carchman, 1982). Overall, these studies generally reported equivocal or negative findings.

Of the included epidemiologic articles, the majority of studies (30 of 42) focused on developmental toxicity endpoints (e.g., low birth weight, fetal growth restriction, and/or birth defects). These endpoints are similar to those observed in the animal studies. Of the epidemiologic studies that examined female reproduction, outcomes assessed included altered menstruation (Windham et al., 2003), female fertility (Dahl et al., 1999), and pregnancy loss/stillbirth/spontaneous

| Reference                                           | Include/exclude | Brief reason for exclusion             |
|-----------------------------------------------------|-----------------|----------------------------------------|
| Nieuwenhuijsen et al. (2008)                        | Exclude         | Trihalomethanes analyzed               |
| Patelarou et al. (2011)                             | Exclude         | Trihalomethanes analyzed               |
| Pedersen et al. (2015)                              | Exclude         | Trihalomethanes analyzed               |
| Porter, Putnam, Hunting, and Riddle (2005)          | Include         | --                                     |
| Riederer, Dhingra, Blount, and Steenland (2014)     | Exclude         | Trihalomethanes analyzed               |
| Righi et al. (2012)                                 | Exclude         | Trihalomethanes, chlorite, and chlorate analyzed |
| Righi et al. (2003)                                 | Exclude         | Trihalomethanes analyzed               |
| Rivera-Núñez and Wright (2013)                      | Include         | --                                     |
| Savitz et al. (2006)                                | Include         | --                                     |
| Savitz et al. (2005)                                | Include         | --                                     |
| Savitz, Andrews, and Pastore (1995)                 | Exclude         | Trihalomethanes analyzed               |
| Shaw et al. (2003)                                  | Exclude         | Trihalomethanes analyzed               |
| Smith et al. (2016)                                 | Include         | --                                     |
| Summerhayes et al. (2012)                           | Include         | --                                     |
| Toledano et al. (2005)                              | Include         | --                                     |
| Villanueva, Cordier, Font-Ribera, Salas, and Levallois (2015) | Exclude | Review |
| Villanueva et al. (2011)                            | Include         | --                                     |
| Waller, Swan, Windham, and Fenster (2001)           | Exclude         | Trihalomethanes analyzed               |
| Waller, Swan, DeLorenze, and Hopkins (1998)         | Include         | --                                     |
| Wennborg, Bodin, Vainio, and Axelsson (2000)        | Include         | --                                     |
| Windham and Fenster (2008)                          | Exclude         | Review                                 |
| Windham et al. (2003)                               | Include         | --                                     |
| Wright, Evans, Kaufman, Rivera-Nunez, and Narotsky (2017) | Include | -- |
| Wright, Schwartz, and Dockery (2004)               | Include         | --                                     |
| Wright, Schwartz, and Dockery (2003)               | Exclude         | Trihalomethanes analyzed               |
| Xie et al. (2011)                                   | Exclude         | Trichloroacetic acid analyzed          |
| Yang et al. (2017)                                  | Include         | --                                     |
| Yang et al. (2016)                                  | Include         | --                                     |
| Yang, Xiao, Ho, Wu, and Tsai (2007)                 | Exclude         | Trihalomethanes analyzed               |
| Zeng, Zhou, et al. (2016)                           | Exclude         | Trihalomethanes analyzed               |
| Zeng, Chen, et al. (2014)                           | Include         | --                                     |
| Zeng, Wang, et al. (2014)                           | Exclude         | Trichloroacetic acid analyzed          |
| Zeng et al. (2013)                                  | Include         | --                                     |
| Zhou et al. (2012)                                  | Exclude         | Trichloroacetic acid analyzed          |
| Zhou, Li, and Xie (2010)                            | Include         | --                                     |
abortion (Waller et al., 1998; Savitz et al., 2006; King et al., 2000; Dodds et al., 2004; Wennborg et al., 2000; Savitz et al., 2005; Toledano et al., 2005; Iszatt et al., 2014). All studies of male reproduction examined sperm concentration, count, motility, and/or quality (Chang et al., 2001; Iszatt et al., 2013; Yang et al., 2016; Zeng et al., 2013; Zeng, Chen, et al., 2014). Although some positive statistical associations for male or female reproductive outcomes were detected in some (but not all) epidemiologic studies of chloroform exposure, the literature on these endpoints is relatively sparse compared with that for developmental outcomes. Thus, the preponderance of epidemiologic studies on chloroform exposure, like that of the animal data, is focused on developmental endpoints.

Based on the results of this scoping exercise, we determined that a more detailed analysis was warranted of the animal and human studies of chloroform exposure as they pertain to developmental outcomes. The decision to focus primarily on developmental outcomes is further supported by OEHHA’s recent determination to classify chloroform under Proposition 65 based on developmental toxicity.

The remainder of the analysis below focuses on detailed assessment of the animal studies and human epidemiologic studies of chloroform exposure as they pertain to developmental outcomes.

### 3.3 Data Synthesis

#### 3.3.1 Evaluation of the animal data

A total of 11 animal studies involving exposure to chloroform and assessment of developmental outcomes was identified for inclusion in this evaluation. All of these studies were previously considered in the assessment by Brown-Woodman et al. (1998). Seven of the studies involved in utero only exposures and assessment of offspring (Table 3); dosing was either by oral gavage (Ruddick et al., 1983; Thompson et al., 1974) or via inhalation (Baeder & Hofmann, 1988, 1991; Murray et al., 1979; Newell & Dilley, 1978; Schwetz et al., 1974). Most of these studies were conducted in the rat; however, Thompson et al. (1974) conducted separate experiments in both rats and rabbits and Murray et al. (1979) conducted in the mouse. Another four studies involved exposure of parental animals to chloroform prior to mating through gestation and lactation with postnatal assessment of offspring (Table 4); chloroform exposure was either by oral gavage or via drinking water. The data for one of these studies were reported in two separate publications (Balster & Borzelleca, 1982; Burkhalter & Balster, 1979). Lim et al. (2004) assessed rat pups after prenatal exposure only as well as after prenatal plus postnatal exposure. Information regarding the preparation of dosing solutions or analysis of dosing formulations or air concentrations of chloroform is provided in the tables. However, because of the age of these studies, few provided detailed information on dosing preparations. Only NTP (1988) reported analysis of oral dosing formation concentrations.

For the purposes of cross-study comparisons, inhalation and drinking water exposures were converted to mg/kg/day doses using standard body weights, water consumption values, and inhalation minute volumes. These calculated doses are shown in parentheses in Tables 3 and 4. Finally, two additional studies reported the results of experiments involving in vitro chloroform exposures on the development of either whole rat embryos (Brown-Woodman et al., 1998) or zebrafish (Teixidó et al., 2015) (Table 5). All of the studies included in this evaluation are individually summarized in Supporting Information to this article.

The overall quality scores derived by application of the ToxRTool (Schneider et al., 2009; https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/toxrtool) are presented in Tables 3–5. The in utero only exposure studies typically scored on the borderline between “1” (reliable without restrictions) and “2” (reliable with restrictions). The two studies by Baeder and Hofmann (1988, 1991) were judged to be the most reliable studies, meeting 20 and 21 of the 21 reporting criteria assessed, respectively. In contrast, the study by Newell and Dilley (1978) scored the poorest, meeting only 13 of the 21 reporting criteria. The most reliable reproductive toxicity studies that addressed developmental outcomes were deemed to be the NTP (1988) and Borzelleca and Carchman (1982) studies. In contrast, the studies by Burkhalter and Balster (1979), Balster and Borzelleca (1982), and Lim et al. (2004) were downgraded to “3” because the study designs were not appropriate for obtaining some of the substance-specific data needed for our analysis. Therefore, these studies were considered supporting information only.

Scoring of the in vitro studies used a different set of criteria from those used to assess the in vivo studies. Using these criteria, both in vitro studies were given scores of 1. However, despite these data being deemed reliable without restriction, the authors of the present analysis considered the data from these studies to be less informative in the overall assessment because they were obtained from organisms devoid of the placenta and maternal environment, the pharmacokinetics of which contribute greatly to the levels and time course of metabolites that may be experienced by a fetal organism. This issue is discussed in more detail below.

#### In utero-only developmental toxicity studies

The studies involving in utero-only exposures are listed in Table 3; specifics regarding dosing and duration are detailed therein. The most common maternal toxicity observations reported were significantly decreased maternal body weights and/or body weight gains in association with significantly decreased food consumption. Among the studies that used gavage dosing, these findings were observed at doses of ≥50 mg/kg/day in rats (Ruddick et al., 1983; Thompson et al., 1974).
### TABLE 3  In utero-only developmental toxicity studies of chloroform

| Reference (ToxRTool score)a | Species/strain (No./group) | Doses (mg/kg/day, unless otherwise noted) | Day of sacrifice | Main findings and NOAELs/NOAECs | Developmental toxicity | Additional comments |
|-----------------------------|---------------------------|------------------------------------------|------------------|---------------------------------|-----------------------|-------------------|
| Oral gavage                 |                           |                                          |                  |                                 |                       |                   |
| Thompson et al. (1974) (2)  | Rat (Sprague–Dawley) 25 dams/group | GD 6–15 0, 20, 50, 126 (Purity: NR) | GD 20 ≥50 mg/kg/day | Maternal toxicity |
|                             |                           |                                          |                  | 126 mg/kg/day                  | Maternal BWG |
|                             |                           |                                          |                  | Clinical signs (hair loss, and rough appearance) | |
|                             |                           |                                          |                  | Maternal FC, BW losses | |
|                             |                           |                                          |                  | Implantations (not toxicologically relevant) | |
|                             |                           |                                          |                  | (21, 23, 22, and 22 pregnant at sacrifice in the 0, 20, 50, and 126 mg/kg/day dose groups, respectively) | |
|                             |                           |                                          |                  | NOAEL: 20 mg/kg/day | 126 mg/kg/day |
|                             |                           |                                          |                  | Fetal BW | |
|                             |                           |                                          |                  | Extra bilateral lumbar ribs (fetal incidence) | |
|                             |                           |                                          |                  | No effect on resorptions, viable litter size, fetal sex ratio, external, visceral or skeletal malformations | |
|                             |                           |                                          |                  | NOAEL: 50 mg/kg/day | |
|                             |                           |                                          |                  | Dosing conducted twice daily; dosing vehicle: corn oil (volume not reported, but varied with dose) | |
|                             |                           |                                          |                  | Fetuses/litter: 1/3 examined viscerally; 2/3 examined skeletally; skeletal staining with Alizarin red only | |
|                             |                           |                                          |                  | Maternal BW and FC data not shown | |
|                             |                           |                                          |                  | Extra bilateral lumbar ribs not significantly increased based on litter incidence | |
|                             |                           |                                          |                  | Results of limited range-finding study also reported | |
| Ruddick et al. (1983) (1)   | Rat (Sprague–Dawley) 15 dams/group | GD 5–14c 0, 100, 200, 400 (Purity: 99%) | GD 21≥100 mg/kg/day | Maternal toxicity |
|                             |                           |                                          |                  | Maternal BWG (32%, 45%, and 58% reductions in the 100, 200, and 400 mg/kg/day dose groups, respectively) | |
|                             |                           |                                          |                  | 14, 12, 10, and 8 surviving pregnant dams at sacrifice in the 0, 100, 200, and 400 mg/kg/day dose groups, respectively. (15 pregnant/group) | |
|                             |                           |                                          |                  | NOAEL: 20 mg/kg/day | 400 mg/kg/day |
|                             |                           |                                          |                  | Fetal BW (19% reduction) | |
|                             |                           |                                          |                  | Runts | |
|                             |                           |                                          |                  | Sternebrae ossification delays, interparietal deviations | |
|                             |                           |                                          |                  | No significant effect on resorptions, litter size, external, visceral or skeletal malformations | |
|                             |                           |                                          |                  | NOAEL: 200 mg/kg/day | |
|                             |                           |                                          |                  | Dosing vehicle: corn oil (10 mL/kg) | |
|                             |                           |                                          |                  | Maternal hematologic and tissue histologic examinations conducted (data not evaluated herein) | |
|                             |                           |                                          |                  | 2 fetuses/litter fixed for histologic examination; ~1/3 of remaining per litter examined viscerally, ~2/3 of remaining per litter examined skeletally; skeletal staining with Alizarin red only | |
|                             |                           |                                          |                  | Maternal FC data not reported | |
| Thompson et al. (1974) (2)  | Rabbit (Dutch-belted) 15 does/group | GD 6–18 0, 20, 35, 50 (Purity: NR) | GD 29 ≥20 mg/kg/day | Maternal toxicity |
|                             |                           |                                          |                  | Maternal BWG | |
|                             |                           |                                          |                  | Clinical signs (anorexia, diarrhea) | |
|                             |                           |                                          |                  | 50 mg/kg/day | |
|                             |                           |                                          |                  | Maternal BWG | |
|                             |                           |                                          |                  | 2, 1, 0, and 4 maternal deaths in the 0, 20, 35, and 50 mg/kg/day dose groups, respectively. | |
|                             |                           |                                          |                  | Dosing vehicle: horn oil (volume not reported, but varied with dose) | |
|                             |                           |                                          |                  | Clinical signs (anorexia, diarrhea) | |
|                             |                           |                                          |                  | NOAEL: <20 mg/kg/day | |
|                             |                           |                                          |                  | w skeletally; skeletal staining with Alizarin red only | |
|                             |                           |                                          |                  | Complete ossification of skull bones not significantly increased based on litter incidence | |
|                             |                           |                                          |                  | Results of limited range-finding study also reported | |

(Continues)
### TABLE 3 (Continued)

| Reference | Species/strain (No./group) | Duration | Doses (mg/kg/day, unless otherwise noted) | Day of sacrifice | Main findings and NOAELs/NOAECs | Developmental toxicity | Additional comments |
|-----------|---------------------------|----------|------------------------------------------|------------------|--------------------------------|-----------------------|---------------------|
| **Inhalation** | | | | | | | |
| Murray et al. (1979) | Mouse (CF-1) 34–40 dams/group | GD 6–15, 1–7, or 8–15 (7 h/day) | 0, 100 ppm (Purity: NR) (0, 303 mg/kg/day) | GD 18 | **100 ppm, GD 6–15** (85% and 37% pregnant at sacrifice in control and treated groups, respectively) | **100 ppm, GD 6–15** | Nominal concentration of 100 ppm was analytically measured to be 99 ± 5 ppm. Fetuses/litter: 3/3 examined viscerally, all examined skeletal; skeletal staining with Alizarin red only. Maternal BW and FC data not shown. Pregnancy incidence among controls was low (62–85%). When empty uteri were stained, total pregnancies were 74% vs 44% (GD 1–7), 91% vs. 43% (GD 6–15), and 65% versus 60% (GD 8–15) in GD 8–15 treatment group, 6 cleft palates occurred in a single litter. 9 of 10 fetuses displaying cleft palate had BW (≥ 25% decrease) compared to controls (1.0–1.2) Data on skeletal variations, ossification delays not shown. |
| | | | | | [Maternal BW] | | |
| | | | | | [Pregnancy incidence] | | |
| | | | | | [100 ppm, GD 1–7] | | |
| | | | | | [Delayed ossification of skull bones] | | |
| | | | | | [Feal BW, crown-rump length] | | |
| | | | | | [Delayed ossification of skull bones, sternebrae] | | |
| | | | | | [62% and 45% pregnant at sacrifice in control and treated groups, respectively] | | |
| | | | | | [Maternal BW] | | |
| | | | | | [Pregnancy incidence] | | |
| | | | | | [100 ppm, GD 8–15] | | |
| | | | | | [Feal BW, crown-rump length] | | |
| | | | | | [Cleft palate incidence (10 fetuses in 4 litters; 6 from a single litter)] | | |
| | | | | | [Delayed ossification of skull bones, sternebrae] | | |
| | | | | | [Maternal BW] | | |
| | | | | | [NOAEL: <100 ppm] | | |
| Schwetz et al. (1974) | Rat (Sprague-Dawley) 20–31 dams/group (77 regular controls, 8 food-starved controls) | GD 6–15 (7 h/day) | 0, 30, 100, 300 ppm (Purity: 99.30%) (0, 35, 119, 354 mg/kg/day) | GD 21 | **30 ppm** [Corpora lutea (not toxicologically relevant)] | **NOAEL: <100 ppm** | Nominal concentrations were 30, 100, and 300 ppm; measured concentrations were reported to be 30, 95 and 291 ppm. Rats in the food-starved control group and in the 30 and 100 ppm groups were exposed in different experimental phase from the regular controls and 300 ppm group rats. Food-restricted control group included methods for fetal examinations not fully reported; only limited data on skeletal examinations reported. |
| | | | | | [Maternal FC @ GD 6–7; Maternal BW @ GD 13] | | |
| | | | | | [≥ 100 ppm] [FC, ]; Maternal BW throughout treatment] | | |
| | | | | | [300 ppm] [Marked anorexia] | | |
| | | | | | [Resorptions (1 of 3 litters completely resorbed)] | | |
| | | | | | [68/77, 22/31, 23/28, 3/20 pregnant animals at sacrifice in the 0, 30, 100 and 300 ppm exposure groups, respectively. 8/8 animals pregnant in food-starved controls.] | | |
| | | | | | [Maternal BWG (dose-related during treatment; significant at 20.1 mg/m3)] | | |
| | | | | | [Food-Restricted Controls] | | |
| | | | | | [Resorptions] | | |
| | | | | | [Maternal death] | | |
| | | | | | [10, 9, 10, and 8 pregnant dams in the control, low, mid and high dose groups, respectively] | | |
| Newell and DIlke (1978) | Rat (Sprague-Dawley) 10 dams/group | GD 7–14 (1 h/day) | 0, 4.6, 10.9, 20.1 mg/L (0, 950, 2700, 4100 ppm; Purity: NR) | GD 20 | ≥ 46 mg/L (950 ppm) [Maternal FC (dose-related during treatment; significant at 20.1 mg/m3)] | ≥ 46 mg/L (950 ppm) [Maternal BW (also decreased in food-restricted controls)] | Exposures at 4.6, 10.9, and 20.1 mg/mL reported to be equivalent to 950, 2300, and 4100 ppm or 110, 260 and 480 mg/kg/day. Food-restricted control group also included methods for fetal examinations not fully reported; 1/2 of each litter examined viscerally; 1/2 of each litter examined skeletally. Mean litter sizes, fetal sex ratio not reported; data on external and visceral examinations not reported; only limited data on skeletal examinations reported. |
| | | | | | [Maternal BW (dose-related during treatment; significant at 20.1 mg/m3)] | | |
| | | | | | [10.9 mg/L (2700 ppm)] | | |
| | | | | | Sleepiness during exposure | | |
| | | | | | 20.1 mg/L (4100 ppm) Marked lethargy, sleepiness during exposure | | |
| | | | | | Maternal death; resorptions | | |
| | | | | | Food-Restricted Controls | | |
| | | | | | [Resorptions] | | |
| | | | | | [10, 9, 10, and 8 pregnant dams in the control, low, mid and high dose groups, respectively] | | |
### TABLE 3 (Continued)

| Reference (ToxRTool score)* | Species/strain (No./group) | Duration | Doses (mg/kg/day, unless otherwise noted) | Day of sacrifice | Main findings and NOAELs/NOAECs |
|----------------------------|----------------------------|----------|------------------------------------------|-----------------|----------------------------------|
| Baeder and Hofmann (1988)  | Rat (Wistar) 20 dams/group | GD 6–15c (7 hrs/day) | 0, 30, 100, 300 ppm (Purity: 99.0%–99.4%) | GD 20c | ≥30 ppm
|                            |                            | (0, 35, 118, 354 mg/kg/day)b | | | Maternal FC during treatment (dose-related; more prominent at beginning of treatment)
|                            |                            |                                  | 20, 18, 17 and 12 surviving pregnant dams at sacrifice in the 0, 30, 100 and 300 ppm dose groups, respectively | | NOAEL: <30 ppm |
|                            |                            |                                  | | | Maternal BWG during treatment (dose-related)
|                            |                            |                                  | | | Crown-rump length (not dose-related)
|                            |                            |                                  | | | Fetal BW
|                            |                            |                                  | | | Thickered/wavy ribs
|                            |                            |                                  | | | No effect on resorptions, viable litter size, fetal sex ratio, placental weights, external, visceral or skeletal malformations
|                            |                            |                                  | | | NOAEL: <30 ppm |
| Baeder and Hofmann (1991)  | Rat (Wistar) 20 dams/group | GD 6–15c* (7 hr/day) | 0, 3, 10, 30 ppm (Purity: 99.0%–99.4%) | GD 20c | ≥10 ppm
| (1)                       |                            |                                  | | | Maternal FC during treatment (dose-related)
|                            |                            |                                  | | | Maternal BW/BWG during treatment
|                            |                            |                                  | | | One dam at 30 ppm had a full litter resorption
|                            |                            |                                  | | | 20 pregnant dams at sacrifice in all dose groups, respectively
|                            |                            |                                  | | | NOAEL: 3 ppm |

**NOAEL = no observed adverse effect level; NOAEC = no observed adverse effect concentration; GD = gestational day; NR = not reported; BW = body weight; BWG = body weight gain; FC = food consumption.**

*ToxRTool score: 1 = reliable without restriction; 2 = reliable with restriction; 3 = unreliable (https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/toxrtool)*

*Inhalation concentrations converted to mg/kg/day doses based on body weights and minute volumes from EPA (1988).*

*Corrected for day of insemination as GD 0.

Nominal concentrations were 30, 100, and 300 ppm; measured concentrations were reported to be 30–34, 115–121 and 308–318 ppm.

Fetuses/litter: 1/2 examined viscerally; 1/2 examined skeletally; skeletal staining with Alizarin red only

Report includes individual animal data.

Nominal concentrations were 3, 10, and 30 ppm; measured concentrations were reported to be 2.8–3.2, 9.8–11.6 and 29.2–31.2 ppm.

Fetuses/litter: 1/2 examined viscerally; 1/2 examined skeletally; skeletal staining with Alizarin red only

Maternal food consumption @ 3 ppm: statistical significance on GD 7–14 due to higher than normal control food consumption in that period

Percentage of fetuses with BW < 3 g: 3.2%, 14.2%, 24%, and 26.9% in the 0, 3, 10 and 30 ppm groups, respectively; only significant at 30 ppm when considered on litter basis

Reduced ossification in skull and sternebrae noted in treated groups compared to controls, but control values were below historical control range and treated were within historical control range

Report includes individual animal data.
## TABLE 4  Reproductive toxicity studies of chloroform with developmental toxicity outcomes

| Reference (ToxRTool score) | Study design | Species/strain (No./group) | Duration | Doses (mg/kg/day, unless otherwise noted) | Main findings | Maternal toxicity | Developmental toxicity | Additional comments |
|-----------------------------|--------------|-----------------------------|----------|------------------------------------------|---------------|------------------|----------------------|----------------------|
|                             | Oral gavage  |                             |          |                                          |               |                  |                      |                      |
|                             | Burkhalter and Balster (1979); Balster and Borzelleca (1982) (2/3) | Reproductive toxicity + neurobehavior | Mouse (ICR) | Parental animals: 3 week pre-mating through mating, gestation and lactation. Pups: Directly PND 7–23 | 0, 31.1 (Purity: NR) | No effect on litter size, righting reflex, forelimb placing response, forepaw grasp, rooting reflex, cliff drop aversion, auditory startle response, bar holding ability, eye opening, motor performance on an inverted screen or passive avoidance reported | Results of Burkhalter and Balster (1979) later reported again in Balter and Borzelleca (1982). Dosing vehicle consisted of Emulpbor, polyoxymethylated vegetable oil, and 0.9% saline. Group-housed by sex 6/cage during pre-mating period, then 3 females and 1 male per cage during breeding period; pregnant females housed singly. Litter culled to 8 pups each on PND 0. Daily testing using a neurobehavioral battery on 3 pups per litter on PND 1–15; motor performance testing of 15 mice/group on an inverted screen on PND 17; passive avoidance testing on 15 mice/group on PND 22 and 23. No parental toxicity data reported; limited pup data reported. No teratogenicity observed. |
|                             | NTP (1988) (1) | Reproductive toxicity (continuous breeding) | Mouse (CD-1) | 20 per sex/group (40 per sex in control group) | 0, 8, 20, 50 (Purity: > 99%) | No effect on parental BW, water consumption, fertility, litters produced per breeding pair, days between subsequent litters, female body weights at delivery. NOAEL: > 50 (41) mg/kg/day | 50 (41) mg/kg/day | NOAEL: > 50 (41) mg/kg/day | Dosing vehicle: corn oil (10 mL/kg). Nominal doses were 8, 20 and 50 mg/kg/day. Actual doses were 6.6, 16, and 41 mg/kg/day, respectively. Actual doses were calculated from aliquots analyzed via gas chromatography. Group-housed by sex 2/cage during pre-mating period, then as breeding pairs. Water consumption values determined per cage. Offspring used for second-generation fertility assessment were dosed directly beginning on PND 22. Limited parental and offspring toxicity data reported. F1 pup parameters assessed in high dose and control group animals only (low and mid-dose animals not evaluated). Crossover mating of first generation animals not conducted because no effect of treatment on fertility observed. Reproductive performance of F1 offspring assessed 64–84 days post-weaning. |

(Continues)
| Reference (ToxRTool score)a | Study design | Species/strain (No./group) | Duration | Doses (mg/kg/day, unless otherwise noted) | Maternal toxicity | Developmental toxicity | Main findings | Additional comments |
|----------------------------|-------------|---------------------------|----------|------------------------------------------|-------------------|----------------------|--------------|---------------------|
| **Drinking water**         |             |                           |          |                                          |                   |                      |              |                     |
| Borzelleca and Carchman (1982) (1) | Reproductive toxicity (2-gen) + dominant lethal test + teratology screen | Mouse (ICR) 10 males, 30 females/group | Parental animals: 5 weeks pre-mating through mating, gestation and lactation of 3 litters F1b Pups: 11 weeks pre-mating through mating gestation and lactation of 2 litters | 0, 0.1, 1.0, 5.0 mg/mL (Purity: ≥ 99%) (0, 24, 240, 1,200 mg/kg/day)b | ≥0.1 mg/mL | 0.1 mg/mL | | Survival: F1b (males only) ≥1.0 mg/mL Postnatal BW: F2b @ Days 7, 14, 21 | Terminal BW: F1b (females only) Viability Index: F1b Lactation Index: F1a ≥1.0 mg/mL | NOAEC: 0.1 mg/mL | Chloroform first dissolved in emulphor, then diluted 1:1000 in drinking water; drinking water solutions prepared fresh twice weekly. Females group-housed 3/cage & males housed singly during pre-mating period, then 3 females and 1 male per cage during breeding period; pregnant females housed singly. Included two control groups: naive control group and emulphor vehicle control group. Dominant lethality test: Treated F0 and F1b males (after generation of the F1b and F2a litters, respectively) mated with naive females, which are sacrificed for examination of uterine contents at ~GD 14. Teratology screen: 1/3 of treated F0 and F1b females (after generation of the F1b and F2a litters, respectively) mated with naive females, then sacrificed for examination of fetuses at GD 18; 1/3 of fetuses examined viscerally, rest examined skeletally with Alizarin red only. |
|                       |             |                           |          |                                          |                   |                      |              |                     |
| Lim et al. (2004) (2/3)     | Reproductive toxicity (assessment of diabetogenic parameters) | Rat (Wistar) 4 dams/group | (1) 2 weeks prior to mating + gestation (2) 2 weeks prior to mating + gestation + lactation | 0, 75 μg/L (Purity: NR) (0, 0.02 mg/kg/day)b | Not reported. |                      | GESTATION ONLY | Pup BW on PND 21, but not by 2 weeks post-weaning (data not shown) |
|                           |             |                           |          |                                          |                   |                      | GESTATION + LACTATION | Pup BW on PND 21 and at 26 weeks |                      |                      |                      |                      |

NOAEL = No observed adverse effect level; NOAEC = No observed adverse effect concentration; NR = Not reported; BW = Body weight.

ToxRTool score: 1 = reliable without restriction; 2 = reliable with restriction; 3 = unreliable (https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/toxrtool)

Drinking water concentrations converted to mg/kg/day doses based on water consumptions values from Table 107 of Derelanko (2000) and body weights from EPA (1988).
et al., 1974). In studies involving inhalation (Baeder & Hofmann, 1988, 1991; Murray et al., 1979; Newell & Dilley, 1978; Schwetz et al., 1974), these effects were observed at exposures of ≥10 ppm (approximately equivalent to 11.8 mg/kg/day in the rat). The reason for the greater sensitivity to the anorexic effects of chloroform via inhalation versus oral bolus dosing is not known but may reflect reduced first-pass metabolism in the liver due to inhalation compared to oral dosing. Alternatively, it is possible that chloroform volatilized from the oral dosing solutions if they were not made up fresh daily.

Clinical signs of toxicity in the chloroform-dosed maternal animals were reported rather inconsistently across the developmental toxicity studies. For example, in the oral gavage studies, Thompson et al. (1974) reported anorexia and/or diarrhea at all doses in the rabbit and hair loss and/or “rough appearance” in rats at 126 mg/kg/day, whereas Ruddick et al. (1983) did not report any clinical observations at doses of ≤400 mg/kg/day. Reductions in food consumption were also observed among the inhalation studies. Schwetz et al. (1974) reported marked anorexia in rats exposed to 300 ppm; Newell and Dilley (1978) also reported sleepiness and/or lethargy at exposures of ≥10.9 mg/mL (2,700 ppm). Consequently, both of these studies included food-restricted control groups to account for any effects related to reduced food consumption. In both studies, the food restricted groups had fetuses of reduced body weights. In Newell and Dilley (1978), the food restricted group had a more severe retardation in fetal body weights than in the high-dose chloroform group and also showed a greater amount of reduced ossification. In contrast, Baeder and Hofmann (1988, 1991) and Murray et al. (1979) did not report any clinical signs of toxicity associated with chloroform exposure. The reason that no clinical signs were noted is not clear, as clinical signs were likely to have been present at the exposures administered in these studies. The absence of clinical signs from these reports could simply reflect the relative age of the studies, which were conducted at a time when studies often did not report on such parameters. Finally, although some of the studies reported maternal mortalities (Newell & Dilley, 1978; Ruddick et al., 1983; Thompson et al., 1974), none were attributed to treatment, except those in rabbits dosed with 50 mg/kg/day of chloroform, which showed significant signs of hepatotoxicity (Thompson et al., 1974).

Evaluation of the in utero-only exposure studies also revealed that high-dose chloroform exposures (≥200 mg/kg/day) were associated with very early (peri-implantation) total litter losses. These findings were generally characterized as reduced pregnancy incidences in the affected studies. Ruddick et al. (1983) reported only 10 and 8 surviving pregnancies (out of 15 females per group) at oral gavage doses of 200 and 400 mg/kg/day, respectively; experiments done

### TABLE 5  In vitro culture developmental toxicity studies of chloroform

| Reference (ToxRTool score)* | Species/strain (No. group) | Duration | Exposure concentrations | Developmental toxicity findings | Additional comments |
|-----------------------------|---------------------------|----------|------------------------|---------------------------------|---------------------|
| Brown-Woodman et al. (1998) (1) | Whole rat embryos (Sprague–Dawley; GD 10.5, 12–15 somites) | 40 hrs | 0, 0.53, 1.05, 2.06, 3.71 μmol/mL (culture medium (Purity: NR)) | ≥2.06 μmol/mL | Chloroform dissolved in DMSO (final DMSO concentration in media = 0.1%) Cultured in 6.0–7.5 mL of heat-inactivated rat serum (4–5 embryos per flask) in 50 mL glass culture bottles. Concentration of chloroform in medium was only evaluated at initiation of culture; no indications that medium was changed over the culture duration. @ 3.71 μmol/mL: Marked cell death was observed throughout embryo with exception of the heart after at least 16 hours of culturing. |
| Teixidó et al. (2015) (1) | Zebrafish (4 hours post-fertilization) | 72 hours (sacrificed 76 hours post-fertilization) | 0, 0.14, 0.31, 0.63, 1.26 mM (Purity: NR) | ≥0.63 mM Hatching success 1.26 mM Reduced Motility Teratologic EC50 reported to be 0.85 mM; teratogenic effects on the eyes heart and tail reported Minimal concentration causing growth inhibition (decreased tail length) reported to be 1.26 mM | Chloroform dissolved in DMSO (final DMSO concentration in media = 0.1%) Cultured in 10 mL buffered embryo medium in 20 mL sealed glass vials; culture solutions renewed every 24 hours Chlorofom LC50 = 2.1 mM Data for vehicle control not shown |

NOAEC = no observed adverse effect concentration; NR = not reported.

*ToxRTool score: 1 = reliable without restriction; 2 = reliable with restriction; 3 = unreliable (https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/toxtool)
via gavage dosing at lower doses (Thompson et al., 1974) did not show similar effects. However, in the inhalation study by Schwetz et al. (1974), only 3 of 20 rats in the 300 ppm (~354 mg/kg/day) group were reported as pregnant (one of which had a completely resorbed litter); the lower exposure of 100 ppm (~118 mg/kg/day) had no effect on the number of pregnant dams. Similarly, Baeder and Hofmann (1988) reported only 12 surviving pregnant dams (of 20 dams in the dose group) at 300 ppm (~354 mg/kg/day); however, lower inhalation exposures in both this and their subsequent study (Baeder & Hofmann, 1991) were without substantial effect. Murray et al. (1979) reported reduced pregnancy incidences in the mouse at 100 ppm (~303 mg/kg/day) with either mid-gestation (GD 6–15) or early gestation (GD 1–7) exposures; however, exposures begun after implantation (GD 8–15) were without effect. In contrast, although resorptions were reported to be increased at the highest exposure in the study by Newell and Dilley (1978) (20.1 mg/L; ~839 mg/kg/day), no substantial effect on pregnancy incidences were reported at this or lower doses. However, exposure in this study was for only 1 hr per day (versus 7 hrs per day in the other inhalation studies of chloroform) and exposure started one day later than in the other inhalation studies (GD 7 versus GD 6).

These data suggest that two factors determined whether early total litter losses occurred as a result of chloroform exposure. First, peri-implantation losses were only evident at extremely high doses—doses well above those at which other indications of maternal toxicity (clinical signs of toxicity, reduced food consumption, and decreased body weight gains) were observed. Second, peri-implantation losses were only observed when dosing started on GD 6 or earlier. Studies for which dosing was initiated after GD 6 did not report reduced pregnancy incidences, indicating that the reduced pregnancy indices were likely due to disruption of the implantation process. Further, studies that included staining of the uterus to determine pregnancy status often found indications that the implantation process had begun, but was disrupted (Baeder & Hofmann, 1988, 1991; Murray et al., 1979). For example, Baeder and Hofmann (1988, 1991) describe 8 of 20 high-dose impregnated dams as having only “empty implantation sites” at necropsy. The mechanism by which chloroform mediates this effect, however, is not known, although chloroform has been shown to reach the uterine compartment shortly after inhalation exposure (Danielsson, Ghantous, & Dencker, 1986). Other studies have shown that chloroform affects vascular function (McCarthy, 1969), causing vasodilation. Thus, we hypothesize that chloroform is likely affecting blood vessels in the endometrium. Although the process of implantation in the rat is associated with increased uterine blood flow to the areas of implantation (McRae & Heap, 1988; Tawai & Rogers, 1992), how the increased uterine blood flow associated with chloroform treatment might affect the process of implantation is yet unknown.

Fetal body weights were reduced at doses as low as 300 ppm (35.4 mg/kg/day) in rats (Baeder & Hofmann, 1988; Schwetz et al., 1974). Reductions in crown-rump length were also observed, sometimes at lower doses, but this is a less sensitive measure and typically did not show a clear dose response. Fetal weight reductions were also observed in mice at the single exposure level tested, 100 ppm (approximately 303 mg/kg/day; Murray et al., 1979). In rabbits, reduced fetal weights were reported at 20 and 50 mg/kg/day, but not at 35 mg/kg/day (Thompson et al., 1974); thus, the relation of reduced fetal body weight to treatment in rabbits is unclear. It is important to note, however, that the lowest dose at which fetal weight reductions were observed in rats (300 ppm) (Baeder & Hofmann, 1988; Schwetz et al., 1974) was higher than the lowest dose at which maternal toxicity was reported (10 ppm) (Baeder & Hofmann, 1991). Further, this observation was consistent across the available rat studies. Reductions in fetal size have been reported as a result of maternal food restriction (Nitzsche, 2017) and were also observed in the food-restricted control groups of Schwetz et al. (1974) and Newell and Dilley (1978). This suggests that the fetal weight reductions are likely a consequence of maternal toxicity.

Alterations or delays in bone ossification were also observed in most, but not all, of these studies. In rats, these included delayed ossification of the skull bones and sternebrae, lumbar ribs, wavy ribs, and “interparietal deviations” (Baeder & Hofmann, 1988, 1991; Ruddick et al., 1983; Schwetz et al., 1974; Thompson et al., 1974). Delayed ossification of skull bones was observed in rabbits at the low and mid doses (but not the high dose) in Thompson et al. (1974) and in mice (in combination with sternebral ossification delays) at 100 ppm (Murray et al., 1979). None of these findings are considered skeletal malformations; rather, they are developmental delays commonly seen in fetuses with reduced body weights and considered transient findings (Carney & Kimmel, 2007; Kimmel, Garry, & DeSesso, 2014). Indeed, these skeletal findings were seen at the same or high doses than those at which reductions in fetal weights were observed and are commonly associated with reductions in fetal weight due to reduced maternal food consumption (Nitzsche, 2017).

Finally, only the studies by Schwetz et al. (1974) and Murray et al. (1979) reported any malformations in offspring. In Schwetz et al. (1974), some malformations related to caudal dysplasia were reported at 100 ppm, although the number of affected fetuses is not reported. Similar findings were neither reported at the next lower exposure of 30 ppm, nor at 300 ppm although the number of fetuses available for evaluation was severely limited. These findings were not confirmed in the other rat inhalation studies of chloroform (Baeder & Hofmann, 1988, 1991). In Murray et al. (1979),
offspring from the group of pregnant mice exposed to chloroform on GD 8–15 exhibited an increased incidence of cleft palate in combination with a substantial reduction in mean fetal body weight. Specifically, 10 fetuses from 4 litters (6 in a single litter) in the chloroform-exposed group presented with cleft palate. One cleft palate occurred in a concurrent control fetus and three cleft palates occurred in fetuses from one of the other control groups; however, no cleft palates occurred in fetuses from either of the other chloroform exposed groups (treatment on GD 1–7 or GD 6–15). Further, no cleft palate was seen in any of the rat developmental toxicity studies, nor in Borzelleca and Carchman (1982), which is a multigeneration study with a teratology screen. The latter study was conducted in mice at multiple doses, some above those administered in Murray et al. (1979). Thus, the cleft palate seen in Murray et al. (1979) was not confirmed by other studies at the same or higher exposures. Further, the malformations seen in Schwetz et al. (1974) and Murray et al. (1979) occurred only at high maternally toxic doses.

The no observed adverse effect levels (NOAELs) tended to be substantially higher in the gavage dosing studies compared to the inhalation studies. Based on the inhalation studies, the NOAEL for maternal toxicity is 4 mg/kg/day based on a chloroform inhalation exposure concentration of 3 ppm in the rat study of Baeder and Hofmann (1991). At the next higher exposure concentration (10 ppm; 12 mg/kg/day) and exposures above this level, maternal toxicity in the form of reduced food consumption, maternal body weights, and body weight gains were consistently observed. Other indications of maternal toxicity, including clinical signs and reduced pregnancy incidences, are seen only at the highest exposure levels used in these studies. The lowest LOAEL for fetal toxicity, from the studies of Schwetz et al. (1974) and Baeder and Hofmann (1988, 1991), is 35 mg/kg/day based on an inhalational exposure concentration of 30 ppm. Thus, the next lower dose at which no fetal effects were observed was 10 ppm. This dose, which is approximately equivalent to 12 mg/kg/day based on minute volume and breathing rate, would be considered the developmental NOAEL. For human health risk assessment, it is important to note that these fetal effects occurred at doses higher than those at which maternal toxicity was observed.

Reproductive toxicity studies

Four reproductive toxicity studies of chloroform were identified (Table 4). Two of these studies treated pregnant animals with chloroform and investigated potential effects in offspring; however, standard parameters regarding reproductive or developmental effects were not assessed (e.g., no parental toxicity data reported; limited pup data reported); as such, these studies were assigned ToxRTool scores of only 2/3, which restricts their use to supporting information only. The first study assessed developmental neurobehavior in pups following oral gavage of the parental animals to 31.1 mg/kg/day of chloroform beginning prior to mating though gestation and lactation (Balster & Borzelleca, 1982; Burkhalter & Balster, 1979); no adverse effects were reported. Further, endpoints related to parental effects were not assessed. The other study (Lim et al., 2004) focused on the assessment of diabetogenic parameters in offspring following exposure of parental animals to drinking water containing 75 μg/L chloroform (equivalent to 0.02 mg/kg/day) beginning prior to mating and continuing either through gestation only or through both gestation and lactation. Again, standard parameters relating to fertility, reproduction, and offspring development were not examined and parental systemic toxicity was not assessed. Although significantly decreased pup body weights were observed on PND 21 in both treated groups, the number of animals per dose group was very low (only 3 pups per litter from only 4 dams per group). Further, the investigators did not report whether these data were assessed based on the litter rather than the fetus. Thus, these data are not considered reliable for assessing the developmental toxicity of chloroform.

Two other studies investigated reproductive toxicity in mice after administration of chloroform via either gavage or drinking water exposure (Borzelleca & Carchman, 1982; NTP, 1988); in both of these studies, exposure began prior to mating, continued through generation of multiple litters, and included assessment of multiple developmental parameters in the offspring. In Borzelleca and Carchman (1982), although a few spurious findings were observed at a drinking water exposure of 0.1 mg/mL (equivalent to 24 mg/kg/day), this exposure level is considered the study NOAEL for both parental toxicity (based on enlarged livers) and developmental effects (based on reduced F2b postnatal weights, reduced F1b viability index, and reduced F1a lactation index). It should be further noted that the mouse pups in this study likely began consuming the drinking water prior to weaning. Due to their low body weights, the resulting direct doses to the offspring likely were considerably higher than those received by the parental animals; further, these direct exposures were in addition to those received by the offspring via lactation. Thus, the NOAEL for developmental effects established in this study is considered to be a conservative estimate.

In the NTP (1988) study at the nominal dose of 50 mg/kg/day (actual dose of 41 mg/kg/day), significantly increased epididymal weights were reported in F1 offspring of treated mice. Mild vacuolar degeneration of the epididymal epithelium was reported in some of these animals. However, no effects on fertility of the F1 offspring were reported at this dose. As such, the study investigators considered this finding to be possibly unrelated to chloroform treatment. Other changes at this dose were not considered adverse, and therefore not toxicologically relevant. Also, no parental toxicity was reported. Consequently, the NOAEL for parental toxicity was considered to be the highest dose tested—50
(41) mg/kg/day. The NOAEL for offspring effects was also considered to be 50 (41) mg/kg/day. This NOAEL for developmental effects is similar to that determined in the study of Borzelleca and Carchman (1982).

**In vitro developmental assays**

Only two studies that used in vitro alternative assays to address the potential developmental effects of chloroform were identified (Table 5). Brown-Woodman et al. (1998) reported effects of chloroform on the development of Sprague Dawley rat embryos in culture at 2.06 μmol/mL. Teixidó et al. (2015) reported altered development of zebrafish embryos in the presence of chloroform, albeit at effective concentrations (≥0.63 mM) significantly lower than that identified in the study by Brown-Woodman et al. (1998). Additionally, Teixidó et al. (2015) reported potential teratogenic effects associated with exposure.

While these studies are potentially useful from the perspective of generating hypothesis-driven mechanistic research data, they are of little use for the purposes of conducting a human health risk assessment to understand the potential for developmental toxicity due to chloroform exposure. For one, these investigations were conducted using in vitro systems devoid of the maternal organism and placenta, both of which play important roles in regulating the disposition of chloroform that is experienced by the developing organism. For example, studies have shown that, following a single oral exposure, most of the chloroform is eliminated from the human body via exhalation within 8 hours (Fry, Taylor, et al., 1972), and in the pregnant animal, chloroform only reaches the placenta and fetuses for short intervals following an inhalation exposure (Danielsson et al., 1986). Further, chloroform is extensively metabolized in rodents (Brown et al., 1974). Thus, because the metabolic processes and pathways of elimination for chloroform are almost completely absent from these in vitro systems, it cannot be concluded that the results obtained using these systems are necessarily indicative of what would be observed in vivo. In fact, the “teratogenic effects” reported in the study of Teixidó et al. (2015) are not seen in the whole animal studies.

The reductions in crown-rump length and somite number were observed in Brown-Woodman et al. (1998) occurred in the absence of the maternal animal and might falsely suggest that these effects are mediated directly on the offspring, independent of the maternal animal. However, it is more likely that the embryotoxic effects observed in Brown-Woodman et al. (1998) are due to a toxic effect of chloroform on vasculogenesis in the yolk sac, which would interfere with the delivery of nutrients to and removal of waste products from the developing organism in vitro. These reductions in yolk sac blood vessel development in the present of chloroform are consistent with our previously stated hypothesis that chloroform is affecting the blood vessels in the endometrium, resulting in reduced implantation success in the whole animal studies.

**Summary of the animal data**

A total of 14 animal studies that addressed developmental endpoints were included for consideration in our data evaluation. Most of the in utero developmental toxicity studies were conducted in the rat and involved the administration of chloroform either via gavage dosing or inhalation exposure. The most common findings of maternal toxicity reported were significantly decreased maternal body weights and/or body weight gains in association with significantly decreased food consumption. These observations were consistently observed at the same or lower doses than those at which fetal toxicity was reported. Higher exposures (≥200 mg/kg/day) were associated with very early (peri-implantation) losses, which may be due to the vasoconstrictive effects of chloroform. Based on data from multiple inhalation studies, we considered the NOAEL for maternal toxicity to be 3 ppm (4 mg/kg/day).

Fetal body weight reductions, as well as alterations or delays in fetal bone ossification, were consistently observed in the in utero developmental toxicity studies of chloroform. In all of the studies, such effects were observed at doses equal to or above those at which maternal toxicity was seen. Ultimately, the offspring effects of chloroform exposure occur at doses also associated with maternal toxicity. Based on review of these data, we consider the NOAEL for offspring effects to be 10 ppm or 12 mg/kg/day (Baeder & Hofmann, 1991).

The findings from the reproductive studies are generally consistent with those from the in utero developmental toxicity studies in that the offspring effects occurred at the same or higher doses than those at which maternal/parental toxicity was observed. The NOAELs identified in these studies are generally higher than those identified in the in utero developmental toxicity inhalation studies. The exception is the NOAELs from Lim et al. (2004), which are not considered reliable for assessing the developmental toxicity potential of chloroform because standard parameters related to offspring development were not assessed. Finally, the findings from the two in vitro developmental toxicity assays provide little to no value to the overall evaluation due to the absence of the maternal organism and placenta in the experimental system. In conclusion, the data from these studies show that the developmental effects of chloroform exposure occur at higher doses than those than cause maternal toxicity.

### 3.3.2 Evaluation of the epidemiologic data

Thirty human epidemiologic studies that investigated the potential relation between chloroform exposure and developmental outcomes underwent a detailed scientific assessment. Twenty-five had been reviewed previously by OEHHA
(2016a). There were five studies not included in the OEHHA review (Brender et al., 2014; Cao et al., 2016; Kogevinas et al., 2016; Levallois et al., 2016; Wright et al., 2017). Eleven studies were prospective cohort studies, 11 studies were retrospective cohort studies, and 8 were case–control studies (7 population-based and 1 hospital-based case–control study). The design and results of these studies are summarized in Tables 6–9 and Supporting Information Table S1.

To gauge the overall quality of the available epidemiologic data on developmental outcomes associated with chloroform exposure, the identified studies were assessed using the RTI revised item bank on risk of bias and precision of observational studies (Viswanathan et al., 2013). The RTI item bank provides a framework for evaluating the validity of observational epidemiologic studies based on 13 questions that are designed to assess potential areas of bias (selection bias, information bias, and confounding bias) and error that can limit the value of such studies as a basis for causal inference. Numerical grading and scoring systems to evaluate epidemiologic literature tend to provide inconsistent weighting and poorly predict study results (Greenland & O’Rourke, 2001; Sanderson, Tatt, & Higgins, 2007). The RTI item bank is not a grading or scoring system for epidemiologic studies. Rather, it is a systematic way of qualitatively examining each study that allows for the identification of key flaws and shortcomings of each study as well as patterns of bias across studies. These patterns are highlighted in this review. The results of this evaluation are shown in Supporting Information Table S1. Some of the factors considered to have the most impact on the reliability of the epidemiologic studies, and therefore of greatest relevance for this assessment, are discussed in more detail below. In this section, the term “significant” is used in reference to statistical significance (by convention, defined as p-value <.05), and the term “association” is used in reference to any relationship between chloroform and developmental outcomes that was found to be statistically significant, unless otherwise specified.

**Integrated exposure assessment**

Across the epidemiologic literature on chloroform exposure and developmental outcomes, studies can be distinguished based on their method of exposure assessment. One of the major methodological shortcomings of many of the studies included in this evaluation is the lack of adequate exposure assessment. Only 13 (Botton et al., 2015; Cao et al., 2016; Costet et al., 2012; Danileviciute et al., 2012; Grazuleviciene et al., 2011; Grazuleviciene et al., 2013; Iszatt et al., 2011; Kogevinas et al., 2016; Levallois et al., 2012; Levallois et al., 2016; Savitz et al., 2005; Smith et al., 2016; Villanueva et al., 2011) of the 30 studies assessed in this review incorporated individual-level exposure information (e.g., frequency and duration of showering, use of a water filter, and bottled versus tap water consumption) when estimating a mother’s exposure to chloroform during pregnancy; the other 17 studies relied solely on chloroform sampling data in local or regional water, and in one case, in air (Brender et al., 2014), for chloroform exposure estimation. Thus, these studies were semi-ecologic in nature, meaning that they relied on group-level geographic exposure data. Ecologic studies are generally considered by epidemiologists to be among the weakest of epidemiologic study designs because associations observed for groups cannot be assumed to hold for individuals (Morgenstern, 1995). That is, average exposure levels across a group may differ substantially from those for individuals within that group, leading to gross misclassification of exposure, and producing substantial bias of estimated associations in either direction.

As stated by the authors of one study that used ecologic data on chloroform exposure, “A potential limitation [is] the lack of individual exposure data. THMs are volatile, and exposure can occur from activities such as showering and bathing. The THM exposure for a 10-minute shower may be equivalent to drinking 2 L of water” (Summerhayes et al., 2012). Thus, failure to account for showering/bathing and water consumption habits could substantially reduce the accuracy of estimated chloroform exposure. Furthermore, only one study assessed individualized exposure to chloroform via a biomarker (Cao et al., 2016); the rest estimated exposure based on models that have not been validated by comparison with physiological biomarkers of chloroform exposure.

Overall, studies that incorporated individualized information into chloroform exposure assessment were considered more likely to obtain valid results because they had a more accurate exposure estimate. Thus, studies that included this information should be considered more pertinent for making any causal inferences. To this point, we considered all studies collectively as well as studies with individualized exposure assessment independently. Throughout the remainder of this evaluation, we refer to exposure measures that incorporate individual data on water use or biomarker information as “integrated chloroform exposure.”

Because exposure assessment is so fundamental to the validity of epidemiologic research, we classified study quality based largely on whether exposure assessment was integrated using individualized data. That is, we classified the studies with integrated exposure estimates as being relatively strong methodologically (Botton et al., 2015; Cao et al., 2016; Costet et al., 2012; Danileviciute et al., 2012; Grazuleviciene et al., 2011; Grazuleviciene et al., 2013; Iszatt et al., 2011; Kogevinas et al., 2016; Levallois et al., 2012; Savitz et al., 2005; Smith et al., 2016; Villanueva et al., 2011) compared with studies that lacked integrated exposure estimates (Brender et al., 2014; Dodds & King, 2001; Hinckley et al., 2005; Hoffman et al., 2008a; Infante-Rivard, 2004; Iszatt et al., 2014; Kramer et al., 1992;
| Author            | Location                                      | Study design     | Study dates   | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels                                                                                     | Main findings                                                                                     |
|-------------------|-----------------------------------------------|------------------|---------------|-------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|
| Botton et al.     | Gipuzkoa, Sabadell, and Valencia, Spain, and Crete, Greece | Prospective cohort | 2003–2008     | 1,474 mother–infant pairs | Integrated exposure assessment of total residential uptake (from ingestion, showering, and bathing) and ingestion uptake | Infant weight gain 0-6 months after birth                                                                                           | Maternal age, gender, gestational age, parity, maternal pre-pregnancy weight, paternal weight, paternal height, maternal education, and maternal smoking during pregnancy | IQR increase (not specified; <1 μg/d) in TCM total residential uptake or ingestion (μg/d) | Median total residential uptake at each site <0.5 μg/d | Beta regression coefficients of postnatal weight gain (grams) from 0 to 6 months associated with IQR increase in TCM uptake: Total residential uptake (μg/d) beta (95% CI): -11.49 (-87.8, 65.0) Ingestion uptake (μg/d) beta (95% CI): -40.3 (-122.4, 41.8) |
| Brender et al.    | Texas                                         | Case-control     | 1996–2008     | 3,245       | Non-integrated exposure assessment using emissions model for sources within a 10-km effective radius | Birth defects: NTD, anencephaly, spina bifida, any oral cleft defect, cleft palate alone, cleft lip with or without cleft palate, conotruncal heart defect, obstructive heart defect, septal heart defect, any limb deficiency, longitudinal limb deficiency, and transverse limb deficiency | Year of delivery, maternal age, education, race/ethnicity, and public health region of residence | Emissions-based modeled exposure risk value >0 TCM exposure risk value intensity: | Median among controls = 0 | Sepal heart defect OR (95% CI): exposure risk value >0: 1.10 (1.01, 1.19) Any NTD OR (95% CI): exposure risk value >0: 1.40 (1.04, 1.87) Spina bifida OR (95% CI): exposure risk value >0: 1.55 (1.10, 2.20) Spina bifida OR (95% CI): TCM exposure risk value intensity: (1) 1.0 (2) 1.74 (1.02, 2.99) (3) 1.23 (0.63, 2.40) (4) 1.66 (0.94, 2.91) p for trend: .027 |
| Cao et al.        | Wuhan and Xiaogan cities, Hubei, China        | Prospective cohort | 2011–2013     | 1,184 live births 60 SGA | THM concentration in whole blood | BW Birth length, Gestational age, SGA (birth weight <10th percentile for gestational age in national Chinese referent population) | Prenatal BMI, weight gain during pregnancy, infant sex, parity, study city, maternal age, gestational age, education, and household income | TCM tertiles (ng/L): (1) <38.2 (2) 38.2–67.1 (3) >67.1 | Median = 50.7 ng/L Geometric mean = 40.7 ng/L | Birth weight (grams) beta (95% CI): (1) 0.04 (2) 0.25 (0.10, 0.40) (3) 0.48 (1.03, 6.4) p for trend: .08 Birth length (cm) beta (95% CI): (1) 0.04 (0.23, 0.07) (3) 0.04 (0.11, 0.18) p for trend: .63 Gestational age (weeks) beta (95% CI): (1) 0 (2) 0.13 (0.05, 0.30) (3) 0.15 (0.03, 0.32) p for trend: 0.10 |

*Table 6* Description and key results of epidemiological studies included in the systematic review.
| Author et al.         | Location          | Study design         | Study dates       | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels                                                                                                       | Main findings                                                                                                       |
|----------------------|-------------------|----------------------|-------------------|-------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|
| Costet et al. (2012) | Brittany, France  | Prospective cohort   | 2002–2006 (enrollment) | 3,223 live births 105 PTB 171 SGA | Integrated exposure assessment to estimate total uptake from inhalation, ingestion, and dermal absorption | PTB SGA (fetal growth restriction; birth weight <5th percentile of expected distribution) | Parity, marital status, diabetes before and during pregnancy, hypertension before or during pregnancy, tobacco use, alcohol consumption, dietary habits, educational level, and obstetric history | TCM water concentration quartiles (μg/L): (1) <5 (2) 5–10 (3) 10–15 (4) ≥15 | PTB OR (95% CI) by quartile of T3 TCM water concentration (μg/L): (1) 1.0 (2) 0.7 (0.4, 1.2) (3) 0.5 (0.3, 0.9) (4) 0.8 (0.4, 1.4) |
|                      |                   |                      |                   |             |                                                                                       |                                                                                         |                                                                                                                        | Total integrated TCM uptake quartiles (μg/d): (1) <0.068 (2) 0.068–0.133 (3) 0.133–0.237 (4) ≥0.237 | OR (95% CI) by quartile of T3 total integrated TCM uptake (μg/d): (1) 1.0 (2) 1.8 (0.7, 4.8) (3) 0.7 (0.2, 2.1) (4) 1.0 (0.4, 2.5) |
|                      |                   |                      |                   |             |                                                                                       |                                                                                         |                                                                                                                        | Ingested TCM uptake quartiles (μg/d): (1) <0.001 (2) 0.001–0.006 (3) 0.006–0.015 (4) ≥0.015 | OR (95% CI) by quartile of T1 ingested TCM (measured via urinary TCM; μg/d): (1) 1.0 (2) 0.7 (0.3, 1.5) (3) 0.8 (0.4, 1.8) (4) 1.2 (0.6, 2.5) |

| Costet et al. (2012) | Kaunas, Lithuania | Prospective nested case-control | 2007–2009 | 59 LBW cases 96 SGA cases 682 controls | Integrated exposure assessment to estimate total uptake from inhalation, ingestion, and dermal absorption | LBW (<2,500 g) SGA (lowest decile of birth weight by gestational week within infant sex and maternal ethnic group) | Marital status, gestational age, maternal education, maternal smoking, paternal smoking, alcohol consumption, BMI, blood pressure, premature birth, infant gender, birth year, and parity | TCM internal dose dichotomized at median (μg/d): (1) <0.1424 (2) ≥0.1424 | LBW OR 95% CI, ≥ median TCM versus < median: Entire pregnancy: 1.24 (0.57, 2.68) T1: 1.15 (0.54, 2.48) T2: 1.29 (0.60, 2.76) T3: 1.45 (0.67, 3.13) |
|                      |                   |                      |                   |             |                                                                                       |                                                                                         |                                                                                                                        | Mean (SD) for all sites = 0.1434 μg/d | SGA OR 95% CI, ≥ median TCM versus < median: Entire pregnancy: 1.31 (0.82, 2.09) T1: 1.25 (0.78, 1.99) T2: 1.23 (0.77, 1.95) T3: 1.31 (0.82, 2.08) |

(Continues)
| Author          | Location                  | Study design    | Study dates | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels | Main findings  |
|-----------------|---------------------------|-----------------|-------------|-------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|-----------------|----------------|
| Dodds and King (2001) | Nova Scotia, Canada       | Retrospective cohort | 1988–1995  | 49,842 mother-infant pairs 77 NTD 430 cardiovascular anomalies 82 cleft defects 96 chromosomal abnormalities | Non-integrated exposure assessment based on routine monitoring of municipal water supply | NTD Cardiovascular defects Cleft defects Chromosomal anomalies | Maternal age, income level, parity, maternal smoking, and neighborhood family income | TCM concentration categories (μg/L): (1) <50 (2) 50–74 (3) 75–99 (4) ≥100 | NTD RR (95% CI): (1) 1.0 (2) 0.7 (0.4, 1.2) (3) 0.7 (0.3, 1.5) (4) 1.2 (0.7, 2.3) Cardiovascular anomalies RR (95% CI): (1) 1.0 (2) 1.2 (0.7, 2.0) (3) 0.9 (0.4, 2.0) (4) 1.5 (0.8, 2.8) Cleft defects RR (95% CI): (1) 1.0 (2) 1.2 (0.7, 2.0) (3) 0.9 (0.4, 2.0) (4) 1.5 (0.8, 2.8) Chromosomal abnormalities RR (95% CI): (1) 1.0 (2) 1.3 (0.8, 2.2) (3) 1.9 (1.1, 3.3) (4) 1.4 (0.8, 2.8) |
| Grazuleviciene et al. (2011) | Kaunas, Lithuania       | Prospective cohort | 2007–2009  | 3,341 mother-infant pairs 156 LBW 270 SGA | Integrated exposure assessment to estimate total uptake from inhalation, ingestion, and dermal absorption | LBW (<2,500 g) SGA (lowest decile of birth weight by gestational week within infant sex and maternal ethnic group) | Marital status, gestational age, maternal education, smoking, alcohol consumption, chronic diseases, BMI, blood pressure, premature birth, infant gender, and birth year | TCM internal dose tertiles (μg/d): (1) 0.0013–0.0249 (2) 0.0249–0.2868 (3) 0.2868–2.1328 Continuous (0.1 μg/d) Mean for all sites = 7.8 (10.2) μg/L | LBW OR (95% CI): Entire pregnancy: (1) 1.0 (2) 2.06 (1.03, 3.85) (3) 1.79 (0.95, 3.36) Cont. 1.10 (1.01, 1.19) Change in BW (grams): −59.1 (−114.8, −3.7) |

*Table 6 (Continued)*
| Author                  | Location          | Study design      | Study dates | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                                                                             | Covariates considered                                                                 | Exposure levels                                      | Main findings                                                                                                                                                                                                 |
|------------------------|-------------------|-------------------|-------------|-------------|--------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| Grazuleviciene et al.  | Kaunas, Lithuania | Prospective cohort| 2007–2009   | 3,074 pairs | Integrated exposure assessment to estimate total uptake from inhalation, ingestion, and dermal absorption | Heart, musculoskeletal, and urogenital anomalies                                                                                             | Age, BMI, chronic disease, alcohol consumption, fetus number, previous premature birth, and infant sex | TCM internal dose tertiles (μg/L): (1) 0.001–0.026 (2) 0.026–0.288 (3) 0.288–2.109                  | Heart defect OR (95% CI), T1: (1) 1.0 (2) 1.05 (0.53, 2.08) (3) 1.37 (0.72, 2.63) p for trend: 0.245 Cont. 1.97 (0.90, 4.35) |
|                        |                   |                   |             |             | Continuous (1 μg/L)                                                                     | Mean at lowest-level site = 0.9 μg/L Mean at highest-level site = 17.7 μg/L                                                                      | Musculoskeletal defect OR (95% CI), T1: (1) 1.0 (2) 0.61 (0.29, 1.32) (3) 0.51 (0.22, 1.14) p for trend: 0.111 Cont. 0.43 (0.11, 1.71) |                                                                                                   |                                                                                                   |
|                        |                   |                   |             |             |                                                                                        | Urogenital defect OR (95% CI), T1: (1) 1.0 (2) 2.21 (0.67, 7.23) (3) 2.50 (0.78, 8.06) p for trend: 0.118 Cont. 2.23 (0.69, 7.17) |                                                                                                   |                                                                                                   |
|                        |                   |                   |             |             |                                                                                        | ORs also available for 1st, 2nd, and 3rd months of pregnancy                                                                                     |                                                                                                   |                                                                                                   |
|                        |                   |                   |             |             |                                                                                        |                                                                                                   |                                                                                                   |                                                                                                   |
| Hinckley et al. (2005) | Arizona, U.S.     | Retrospective cohort | 1998–2003   | 48,119 pairs | Non-integrated exposure assessment based on sampling at local water facilities        | Term LBW (<2,500 g at ≥37 weeks) SGA (intrauterine growth retardation; <10th percentile of weight for gestational age and race/ethnicity) | Maternal age, parity, education, race, ethnicity, smoking, and Kessner index (prenatal care adequacy) | TCM water concentration categories (μg/L): (1) < 10 (2) 10–16 (3) ≥ 16                        | OR (95% CI): T3, TCM water concentration: Term LBW (1) 1.0 (2) 1.18 (1.00, 1.39) (3) 1.04 (0.88, 1.23) Cont. 1.00 (1.00, 1.01) |
|                        |                   |                   |             |             |                                                                                        | Continuous (1 μg/L)                                                                                                                               |                                                                                                   |                                                                                                   |
|                        |                   |                   |             |             |                                                                                        |                                                                                                   |                                                                                                   |                                                                                                   |
|                        |                   |                   |             |             |                                                                                        |                                                                                                   |                                                                                                   |                                                                                                   |
| Author          | Location            | Study design       | Study dates | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels                                                                 | Main findings                                                                 |
|-----------------|---------------------|--------------------|-------------|-------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Hoffman et al.  | U.S. communities    | Prospective cohort | 2000–2004   | 1,958 pairs | Non-integrated exposure assessment based on sampling of tap water at geographically representative locations | Term BW                                                                                | Maternal age, race/ethnicity, income, education, marital status, pre-pregnancy BMI, parity, caffeine intake, and all other disinfection byproduct species reported | TCM residential water categories (μg/L): (1) 19.9–44.2 (2) 44.3–49.0 (3) 49.1–94.0 Mean (SD) = 46.7 (13.3) μg/L | SGA OR (95% CI) in T3 water concentration: (1) 1.0 (2) 1.9 (0.5, 8.1) (3) 1.7 (0.4, 7.1) |
|                 |                     |                    |             |             |                                                                                      |                                                                                       |                                                                                     | Bhrominated site: (1) 1.0 (2) 2.4 (0.6, 3.7) (3) 3.6 (0.5, 3.0)                  | Term BW beta (grams /95% CI) in T3 water concentration: Chlorinated site: (1) 0.0 (2) 2.5 (0.5, 10.4) (3) 3.1 (0.2, 6.1) |
| Infante-Rivard  | Montreal, Quebec, Canada | Hospital-based case-control | 1998–2000   | 458 cases | Non-integrated exposure assessment based on regulatory data for drinking water treatment facilities | SGA (intrauterine growth restriction; birth weight <10th percentile for gestational age and sex, based on Canadian standards) | Race, sex, gestational age, weight gain during pregnancy, pre-pregnancy BMI, parity, history of preclampsia, prior history of SGA, primiparity, and smoking during pregnancy | TCM water concentration (μg/L): (1) < 23.7 (90th percentile) (2) >23.7 (>90th percentile) Mean (SD) for cases = 11.84 (18.19) μg/L Mean (SD) for controls = 11.58 (16.31) μg/L | SGA OR (95% CI) by water concentration, entire pregnancy: (1) 1.0 (2) 1.06 (0.63, 1.79) |
|                 | Southeast England   | Case-control       | 2000–2003   | 354 cases | Integrated exposure assessment to estimate total uptake from ingestion, inhalation, and dermal exposure based residential water zone monitoring data and individualized information on water use activities | Hypospadias                                                                             | Family income, low birth weight, folic acid supplement use, maternal smoking at weeks 6–18, maternal occupational exposure to phthalates, swimming, family history of hypospadias, history of previous stillbirth, gestational diabetes, and intake of cold tap water/bottled water | TCM water concentration quartiles (μg/L): (1) 10.0–9 (2) 2.0–2.9 (3) 1.3–6.5 (4) 1.7–9 Mean for cases = 2.84 (1.70) μg/L Mean for controls = 2.49 (1.36) μg/L | Hypospadias OR (95% CI) for TCM exposure in water: (1) 1.0 (2) 1.17 (0.67, 2.03) (3) 0.99 (0.57, 1.69) (4) 0.84 (0.49, 1.46) |
|                 |                     |                    |             | 336 controls |                                                                                      |                                                                                       |                                                                                     | TCM ingestion quartiles (μg/d): (1) 0.0 (2) >0.0–1.4 (3) 1.5–4.2 (4) 3.1–6.5 Mean for cases = 0.5 (0.3, 0.7) μg/d Mean for controls = 0.3 (0.1, 0.5) μg/d | Hypospadias OR (95% CI) for TCM ingestion: (1) 1.0 (2) 2.12 (0.79, 2.01) (3) 1.12 (0.70, 1.79) (4) 1.36 (0.84, 2.22) p for trend: .30 |
| Iszatt et al.   |                     |                    |             |             |                                                                                      |                                                                                       |                                                                                     | TCM total uptake quartiles (μg/d): (1) 1.0–1.37 (2) 3.8–4.78 (3) 9.8–13.38 (4) 13.99–∞ Mean for cases = 1.36 (1.36) μg/d Mean for controls = 1.36 (1.36) μg/d | Hypospadias OR (95% CI) for TCM total uptake: (1) 1.0 (2) 2.93 (0.56, 1.53) (3) 0.86 (0.52, 1.42) (4) 0.74 (0.45, 1.21) |
| Author               | Location                      | Study design       | Study dates          | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels                                                                                     | Main findings                                                                                                                                 |
|----------------------|-------------------------------|--------------------|----------------------|-------------|---------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| Iszatt et al. (2014) | Northwest England             | Retrospective cohort | 2000–2002, 2005–2007 | 429,599     | Non-integrated exposure assessment based on water sampling in random geographic locations within water zones | LBW (<2,500 g) VLBW (<1,500 g)                                                          | None (crossover comparison)                                                                 | Change in TCM water concentration (based on total THM) (μg/L): (1) Low: ≤10 increase or <10 decrease (2) Medium: 10 to <30 decrease (3) High: 30 to 65 decrease | Percent change (95% CI) for rates before and after enhanced coagulation: LBW (1) −5 (−9, −1) (2) −5 (−9, −1) (3) −9 (−12, −5) p for interaction between TCM change and before/after enhanced coagulation: 0.29 |
|                      |                               |                    |                      | 27,664 LBW  |                                                                                        | LBW                                      |                                                                                       | Mean (SD) for 2000–2002 = 38.6 (4.2) μg/L Mean (SD) for 2005–2007 = 19.4 (1.0) μg/L   | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.02 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.04 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.05 (−0.7, 0.7) |
|                      |                               |                    |                      | 4,209 VLBW  |                                                                                        | VLBW                                     |                                                                                       | Mean (SD) = 12.5 (38.7) μg/L                                                                   | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.06 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.08 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.10 (−0.7, 0.7) |
|                      |                               |                    |                      |             |                                                                                        | LBW                                      |                                                                                       | Mean (SD) = 12.5 (38.7) μg/L                                                                   | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.08 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.10 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.12 (−0.7, 0.7) |
|                      |                               |                    |                      |             |                                                                                        | VLBW                                     |                                                                                       | Mean (SD) = 12.5 (38.7) μg/L                                                                   | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.10 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.12 (−0.7, 0.7) TCM in drinking water, entire pregnancy: 0.14 (−0.7, 0.7) |
| Kogevinas et al. (2016) | 5 European cohorts: France, Greece, Lithuania, Spain, and the United Kingdom | Prospective cohort | 2002–2010 (enrollment) | 14,005      | Integrated exposure assessment to estimate total uptake from ingestion, inhalation, and dermal exposure | Term LBW (<2,500 g at ≥37 weeks) SGA (<10th percentile of cohort-specific reference of fetal growth, stratified by week of gestation and sex; customized model of fetal growth restriction for two studies) PTB (<37 gestational weeks) | Study center, infant sex, gestational age, mother’s ethnicity and parity, maternal age, maternal height, maternal prenatal care, maternal smoking during pregnancy | 10-μg/L increase in TCM concentration in drinking water 10% increase in TCM uptake | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.02 (−0.26, 0.2) TCM in drinking water, entire pregnancy: 0.04 (−0.28, 0.2) TCM in drinking water, entire pregnancy: 0.06 (−0.30, 0.2) |
|                      |                               |                    |                      | 14,005       |                                                                                        | LBW                                      |                                                                                       | Mean (SD) = 12.5 (38.7) μg/L                                                                   | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.04 (−0.28, 0.2) TCM in drinking water, entire pregnancy: 0.08 (−0.30, 0.2) TCM in drinking water, entire pregnancy: 0.12 (−0.30, 0.2) |
|                      |                               |                    |                      | 14,005       |                                                                                        | VLBW                                     |                                                                                       | Mean (SD) = 12.5 (38.7) μg/L                                                                   | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.08 (−0.30, 0.2) TCM in drinking water, entire pregnancy: 0.12 (−0.30, 0.2) TCM in drinking water, entire pregnancy: 0.16 (−0.30, 0.2) |
| Kramer et al. (1992) | 151 towns with 1,000–5,000 people, Iowa, U.S. | Case-control | 1989–1990 | 159 LBW 795 controls for LBW 342 PTB 1,710 controls for PTB 187 SGA 935 controls for SGA | Non-integrated exposure assessment based on 1987 municipal drinking water survey data | LBW (<2,500 g) PTB (<37 weeks completed gestation) SGA (intrathoracic growth retardation; <5th percentile of weight for gestational age according to California data on non-Hispanic whites) | Birth status, maternal age, number of previous children, prenatal care, maternal smoking, maternal education, and maternal smoking during pregnancy | TCM water concentration categories (μg/L): (1) Undetectable (2) 1–9 (3) ≥10 | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.02 (−0.03, 0.03) TCM in drinking water, entire pregnancy: 0.04 (−0.05, 0.05) TCM in drinking water, entire pregnancy: 0.06 (−0.07, 0.07) |
|                      |                               |                    |                      | 342 PTB 1,710 controls for PTB 187 SGA 935 controls for SGA |                                                                            |                                                                                       |                                                                                       | Mean (SD) = 1.0 (0.98, 1.07)                                                               | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.04 (−0.05, 0.05) TCM in drinking water, entire pregnancy: 0.08 (−0.07, 0.07) TCM in drinking water, entire pregnancy: 0.12 (−0.07, 0.07) |
|                      |                               |                    |                      | 159 LBW 795 controls for LBW 342 PTB 1,710 controls for PTB 187 SGA 935 controls for SGA |                                                                            |                                                                                       |                                                                                       | Mean (SD) = 1.0 (0.98, 1.07)                                                               | Birth weight beta (grams) (95% CI) per 10-μg/L increase in TCM: TCM in drinking water, entire pregnancy: 0.04 (−0.05, 0.05) TCM in drinking water, entire pregnancy: 0.08 (−0.07, 0.07) TCM in drinking water, entire pregnancy: 0.12 (−0.07, 0.07) |

(Continues)
| Author                  | Location                | Study design | Study dates | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                                                                                                                                                                      | Exposure levels                                                                                       | Main findings                      |
|------------------------|-------------------------|--------------|-------------|-------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|----------------------------------|
| Levallois et al. (2012)| Quebec City area,      | Case-control | 2006–2008   | 571 term SGA cases 1,925 controls | Integrated exposure assessment to estimate uptake by ingestion, inhalation, and all routes combined | Term SGA (<10th percentile of gestational age at ≥37 weeks, according to Canadian standards) | Maternal age, calendar week, education, annual household income, BMI, parity, history of LBW, maternal smoking, passive smoking at home, caffeine consumption during pregnancy, alcohol consumption during pregnancy, history of chronic disease, pre-eclampsia, maternal ethnicity, work status, marital status, medical problems during pregnancy, and risky occupational exposure | TCM water concentration quantiles (μg/L):                                                                                     | Term SGA OR (95% CI) for TCM exposure in T3: Water concentration (μg/L):                                                                                       |
|                        | Canada                  |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (1) <15.96                                                                                                                                         | (1) 1.0 (0.7, 1.3)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (2) 15.96–27.26                                                                                                                     | (2) 1.2 (0.9, 1.7)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (3) 27.27–51.07                                                                  | (3) 1.0 (0.8, 1.3)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (4) >51.07                                                                      | (4) 1.2 (0.9, 1.7)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | Mean (SD) for controls = 41.1 (39.2) μg/L                                                                                                               | Mean (SD) for cases = 43.3 (40.7) μg/L                                             | p for trend = .10 |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | Ingestion (μg/d):                                                                                     | Ingestion (μg/d):                                                                                     | p for trend = .10 |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (1) 1.0                                                                                      | (1) 2.0 (0.9, 1.6)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (2) 1.2 (0.9, 1.6)                                                                                     | (2) 1.1 (0.8, 1.5)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (3) 1.3 (1.0, 1.8)                                                                                     | (3) 1.0 (0.8, 1.4)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (4) 1.3 (1.0, 1.8)                                                                                     | (4) 1.0 (0.8, 1.4)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | p for trend = .10                                                                      | p for trend = .81 |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | TCM exposure quantiles by route (μg/d):                                                                                       | Total pathway (μg/d):                                                                                     |                      |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | Ingestion: (1) 1.72                                                                                       | (1) 1.0 (0.7, 1.3)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (2) 1.72–11.88                                                                                      | (2) 1.2 (0.9, 1.6)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (3) 11.89–34.30                                                                                      | (3) 1.1 (0.8, 1.5)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (4) >34.30                                                                                       | (4) 1.3 (1.0, 1.8)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | Mean (SD) for controls = 132.7 (151.6) μg/d                                                                                     | Mean (SD) for cases = 134.9 (149.8) μg/d                                            |                      |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | Inhalation/dermal:                                                                                       | Total pathway (μg/d):                                                                                     |                      |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (1) 0.8 (0.6, 1.1)                                                                                     | (1) 1.0 (0.7, 1.3)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (2) 0.9 (0.7, 1.2)                                                                                     | (2) 1.0 (0.8, 1.4)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | p for trend = .81                                                                      | p for trend = .67 |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | Total pathway:                                                                                       | Total pathway:                                                                                       |                      |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (1) 4.2.4                                                                                       | (1) 1.0 (0.7, 1.3)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (2) 4.2.4–8.21                                                                                      | (2) 1.0 (0.8, 1.4)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (3) 8.22–169.81                                                                                     | (3) 1.0 (0.8, 1.4)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | (4) >169.81                                                                                       | (4) 1.0 (0.8, 1.4)           |
|                        |                         |              |             |             |                                                                                      |                                                                                         |                                                                                                                                  | p for trend = .10                                                                      | p for trend = .10 |

Levallois et al. (2016) Quebec City area, Canada Case-control 2006–2008 293 term SGA cases 1,162 controls Integrated exposure assessment to estimate uptake by ingestion, inhalation, and all routes combined Term SGA (<10th percentile of gestational age at ≥37 weeks, according to Canadian standards) Maternal age, maternal education, annual household income, working status, marital status, pre-pregnancy BMI, parity, history of chronic disease, medical problems during pregnancy, active smoking during the third trimester and passive smoking throughout the pregnancy, coffee and alcohol consumption, and risky occupational exposure and DNA source TCM water concentration and internal dose quantiles (not specified) Mean (SD) water concentration for controls = 41.9 (40.3) μg/L Mean (SD) water concentration for cases = 43.6 (40.6) μg/L Mean (SD) internal dose for controls = 132.7 (151.6) μg/d Mean (SD) internal dose for cases = 134.9 (149.8) μg/d Term SGA OR (95% CI) for TCM exposure in T3: Water concentration (μg/L): (1, 2, 3, 4) 1.0 (0.8, 1.4)
| Author | Location | Study design | Study dates | Sample size | Exposure assessment | Outcomes of interest | Covariates considered | Exposure levels | Main findings |
|--------|----------|--------------|-------------|-------------|---------------------|---------------------|----------------------|-----------------|---------------|
| Lewis et al. (2006) | 27 communities in Massachusetts, U.S. | Retrospective cohort | 1999–2001 | 36,529 mother-infant pairs 3,451 term LBW | Non-integrated exposure assessment based on community-level water data from state records on a single water supplier | Term LBW (<2,500 g at ≥37 weeks) | Gestational age, infant sex, marital status, adequacy of prenatal care, maternal age, maternal race/ethnicity, maternal education, parity, maternal smoking, prenatal care source of payment, conception season, birth season, per capita income, previous preterm or SGA infant, previous trimester exposure, and maternal disease factors including anemia, cardiac disease, diabetes, hydramnios, chronic hypertension, pregnancy-related hypertension, Rh sensitivity, saddle cell anemia, and uterine bleeding | Total THM water concentration categories (μg/L): (1) <40 (2) 40 to <50 (3) 50 to <60 (4) 60 to <70 (5) ≥70 Continuous (10 μg/L) IQR = 59 (min: 28, max: 87) μg/L | Term LBW OR (95% CI), T1: (1) 1.00 (2) 0.82 (0.66, 1.03) (3) 0.84 (0.66, 1.08) (4) 0.88 (0.66, 1.17) (5) 0.87 (0.63, 1.12) Cont.: 0.97 (0.89, 1.04) Term LBW OR (95% CI), T2: (1) 1.00 (2) 1.10 (0.81, 1.49) (3) 1.08 (0.79, 1.49) (4) 1.24 (0.92, 1.67) (5) 1.50 (0.67, 2.0) Cont.: 1.08 (1.00, 1.17) Term LBW OR (95% CI), T3: (1) 1.00 (2) 0.87 (0.60, 1.26) (3) 0.79 (0.56, 1.12) (4) 0.84 (0.58, 1.21) (5) 0.74 (0.44, 1.22) Cont.: 0.94 (0.85, 1.05) Term LBW OR (95% CI), entire pregnancy: (1) 1.00 (2) 0.87 (0.60, 1.26) (3) 0.95 (0.56, 1.21) (4) 1.23 (0.92, 1.64) (5) N/A Cont.: 1.03 (0.93, 1.13) Change in BW (grams) (95% CI) per 10 μg total THM: T1: −7 (−14, 0) T2: −6 (−13, 1) Pregnancy: −12 (−21, −3) |
| Lewis et al. (2007) | 27 communities in Massachusetts, U.S. | Risk-set case-control | 1999–2001 | 37,498 mother–infant pairs 2,813 PTB | Non-integrated exposure assessment based on community-level water data from state records on a single water supplier | PTB (32 to <37 weeks of gestation) | Infant sex, marital status, adequacy of prenatal care, maternal age, maternal race/ethnicity, maternal education, parity, maternal smoking, prenatal care source of payment, conception season, birth season, per capita income, previous preterm or SGA infant, previous trimester total THM exposure, and the presence of one or more maternal disease factors including lung disease, diabetes, eclampsia, hydramnios, chronic hypertension, pregnancy-related hypertension, incompetent cervix, renal disease, uterine bleeding, and inappropriate maternal weight gain/loss | Total THM water concentration categories (μg/L): (1) <40 (2) 40 to <60 (3) >60 Continuous (10 μg/L) IQR = 59 (min: 28, max: 87) μg/L | PTB HR (95% CI), T1: (1) 1.0 (2) 1.02 (0.92–1.13) (3) 1.00 (0.88–1.14) Cont.: 0.98 (0.95–1.02) PTB HR (95% CI), T2: (1) 1.0 (2) 0.87 (0.77, 0.99) (3) 0.82 (0.71, 0.94) Cont.: 0.95 (0.92, 0.99) PTB HR (95% CI), 4-week risk sets: (1) 1.0 (2) 1.00 (0.87–1.15) (3) 1.13 (0.95–1.35) Cont.: 1.01 (0.97–1.06) PTB HR (95% CI), entire pregnancy: (1) 1.0 (2) 0.92 (0.82, 1.02) (3) 0.85 (0.74, 0.97) Cont.: 0.95 (0.91, 0.99) |
| Author                  | Location                      | Study design     | Study dates      | Sample size      | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels                                                                 | Main findings                                                                 |
|-------------------------|-------------------------------|------------------|------------------|------------------|-------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| Porter et al. (2005)    | Maryland County, Maryland, U.S. | Retrospective    | 1998–2002        | 15,315 mother−infant pairs | Non-integrated exposure assessment based on sampling data from local water utilities | SGA (intrauterine growth retardation; birth weight <10th percentile for gestational age with sex and race group according to U.S. census data) | Maternal, age, maternal weight gain, race/ethnicity of the child, adequacy of prenatal care (determined by the Kessner index), marital status, tobacco/cigarette use, and alcohol consumption | TCM water concentration quintiles (not specified) | SGA OR (95% CI) by quintile of TCM exposure in water, entire pregnancy:   |
|                         |                               | cohort           |                  | 1,114 SGA        |                                                                                      |                                                                                          |                                                                                     |                                                                                 | (1) 1.0 (2) 1.24 (1.02, 1.50) (3) 1.08 (0.88, 1.32) (4) 1.12 (0.92, 1.36) (5) 1.04 (0.85, 1.27) |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | OR (95% CI), T1:                                                                 |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | (1) 1.0 (2) 0.88 (0.73, 1.07) (3) 1.04 (0.86, 1.25) (4) 1.03 (0.85, 1.24) (5) 0.90 (0.74, 1.09) |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | OR (95% CI), T2:                                                                 |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | (1) 1.0 (2) 1.16 (0.95, 1.40) (3) 1.02 (0.84, 1.24) (4) 1.04 (0.85, 1.26) (5) 1.08 (0.89, 1.32) |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | OR (95% CI), T3:                                                                 |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | (1) 1.0 (2) 1.02 (0.84, 1.24) (3) 0.96 (0.79, 1.16) (4) 0.98 (0.81, 1.19) (5) 1.07 (0.88, 1.29) |
| Rivera-Núñez and Wright (2013) | Massachusetts, U.S. | Retrospective      | 1996–2004      | 672,120 mother-infant pairs | Non-integrated exposure assessment based on town-level water monitoring data | Term SGA (birth weight <10th percentile for gestational age at ≥37 weeks, specific to sex and maternal race) | Maternal age, race/ethnicity, education, smoking, parity, adequacy of prenatal care, prenatal source of payment, median household income, marital status, maternal medical history (e.g., diabetes, chronic hypertension, lung disease, and renal disease), reproductive health factors (e.g., pre-eclampsia, oligohydramnios, incompetent cervix), preterm birth, and weight gain during pregnancy, season, and total trihalomethane and haloacetic acid concentrations | TCM water concentration quintiles (μg/L): (1) ≤5 (2) >5–21 (3) >21–36 (4) >36–52 (5) >52 | PTB OR (95% CI), T1 (1) 1.0 (2) 0.99 (0.93, 1.05) (3) 1.06 (1.00, 1.13) (4) 1.03 (0.97, 1.10) (5) 1.00 (0.94, 1.06) |
|                         |                               | cohort           |                  | 37,136 PTB | Term BW (at ≥37 weeks)                                                                 |                                                                                          |                                                                                     |                                                                                 | PTB OR (95% CI), T2 (1) 1.0 (2) 1.00 (0.94, 1.06) (3) 1.08 (1.02, 1.14) (4) 1.06 (0.99, 1.12) (5) 1.00 (0.94, 1.07) |
|                         |                               |                  |                  | 68,409 term SGA | Term SGA (birth weight <10th percentile for gestational age at ≥37 weeks, specific to sex and maternal race) | Maternal age, race/ethnicity, education, smoking, parity, adequacy of prenatal care, prenatal source of payment, median household income, marital status, maternal medical history (e.g., diabetes, chronic hypertension, lung disease, and renal disease), reproductive health factors (e.g., pre-eclampsia, oligohydramnios, incompetent cervix), preterm birth, and weight gain during pregnancy, season, and total trihalomethane and haloacetic acid concentrations | TCM water concentration quintiles (μg/L): (1) ≤5 (2) >5–21 (3) >21–35 (4) >35–52 (5) >52 | PTB OR (95% CI), T2 (1) 1.0 (2) 0.99 (0.93, 1.05) (3) 1.06 (1.00, 1.13) (4) 1.03 (0.97, 1.10) (5) 1.00 (0.94, 1.06) |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | Term SGA OR (95% CI), T2 (1) 1.0 (2) 0.91 (0.79, 1.04) (3) 1.03 (0.90, 1.18) (4) 1.01 (0.88, 1.16) (5) 1.02 (0.89, 1.18) |
|                         |                               |                  |                  |                  |                                                                                      |                                                                                          |                                                                                     |                                                                                 | Term SGA OR (95% CI), T3 (1) 1.0 (2) 1.01 (0.96, 1.05) (3) 1.00 (0.95, 1.04) (4) 1.04 (1.00, 1.10) (5) 1.04 (0.99, 1.10) |
| Author | Location | Study design | Study dates | Sample size | Exposure assessment | Outcomes of interest | Covariates considered | Exposure levels | Main findings |
|--------|----------|--------------|-------------|-------------|---------------------|---------------------|-----------------------|-------------------|---------------|
| Rivera-Núñez and Wright (2013) | | | | | | | | | Change in term BW (grams) (95% CI); T2 (1) 0.0 (2) –0.7 (3) −1 (4) −1.1 (5) −1.2 |
| | | | | | | | | | Change in term BW (grams) (95% CI); T3 (1) 0.0 (2) –1 (3) −1.2 (4) −1.7 (5) −1.8 |
| Savitz et al. (2005) | | | | | Integrated exposure assessment to estimate intake from tap water, ingestion, and showering/bathing, and non-integrated exposure assessment based on tap water concentration, ingested amount, or showering/bathing | SGA (birth weight <10th percentile for gestational age given sex, maternal race/ethnicity, and parity) PTB (<37 weeks completed gestation) | Term BW (at ≥37 weeks) | Maternal caffeine intake, income, BMI, live birth history, maternal race, education, smoking, employment, diabetes status, and gestational age | TCM water concentration quartiles for SGA/T3 (μg/L): (1) >0.0 –0.2 (2) >0.2 –0.5 (3) >0.5 –0.8 (4) >0.8 –1.2 | SGA OR (95% CI) T3 TCM water concentration (μg/L): (1) 1.0 (2) 0.79 (3) 0.71 (4) 0.54 |
| | | | | | | | | | T3 total integrated TCM exposure (μg/d): (1) 0 (2) 0.63 (3) 0.68 (4) 0.62 |
| | | | | | | | | | PTB OR (95% CI) T3 TCM water concentration (μg/L): (1) 1.0 (2) 0.68 (3) 0.76 (4) 0.52 (5) 0.54 |
| | | | | | | | | | T3 total integrated TCM exposure (μg/d): (1) 0 (2) 0.63 (3) 0.76 (4) 0.52 (5) 0.54 |
| Savitz et al. (2005) | 3 communities in the U.S. | Prospective cohort | 2000–2004 | 1,934 mother-infant pairs | 102 SGA 196 PTB | | | TCM water concentration quintiles for PTB/BW, T3 (μg/L): (1) >0.0 –0.1 (2) >0.1 –0.2 (3) >0.2 –0.3 (4) >0.3 –0.4 (5) >0.4 –0.5 | Term BW, mean change (95% CI) T3 TCM water concentration (μg/L): (1) 1.0 (2) 0.68 (3) 0.76 (4) 0.52 (5) 0.54 |
| | | | | | | | | | T3 total integrated TCM exposure (μg/d): (1) 0 (2) 0.63 (3) 0.76 (4) 0.52 (5) 0.54 |
| | | | | | | | | | PTB OR (95% CI) T3 TCM water concentration (μg/L): (1) 1.0 (2) 0.68 (3) 0.76 (4) 0.52 (5) 0.54 |
| | | | | | | | | | T3 total integrated TCM exposure (μg/d): (1) 0 (2) 0.63 (3) 0.76 (4) 0.52 (5) 0.54 |

Additional results by other pregnancy time periods and estimates by ingestion and showering/bathing (tables 8.8, 8.20, and 8.32)
| Author | Location | Study design | Study dates | Sample size | Exposure assessment | Outcomes of interest | Covariates considered | Exposure levels | Main findings |
|--------|----------|--------------|-------------|-------------|---------------------|----------------------|----------------------|------------------|---------------|
| Smith et al. (2016) | Bradford, England | Prospective cohort | 2007–2010 | 7,438 term births | Integrated exposure assessment to estimated time-weighted update based on ingestion, inhalation, and dermal exposure | Term BW (≥ 37 weeks) | Caffeine intake, maternal education, fasting and post-load glucose, maternal ethnicity, smoking, parity, age, maternal BMI, socioeconomic status, gestational age, and infant sex | Total integrated TCM uptake terciles, entire pregnancy (μg/d): (1) <0.91 (2) ≥0.91 to <1.56 (3) ≥1.56 | Change in term BW (grams) (95% CI) for total integrated TCM uptake, entire pregnancy (μg/d): (1) 0.0 (2) –16.3 (−39.6, 6.5) (3) –20.9 (−44.6, 2.8) p for trend: 0.082 |
| Summerhayes et al. (2012) | New South Wales, Australia | Retrospective cohort | 1997–2004 | 314,982 mother-infant pairs | Non-integrated exposure assessment based on monitoring of water distribution systems | SGA (birth weight <10th percentile for gestational age) | Mothers age, indigenous status, country of birth, infant sex, smoking anytime during pregnancy, parity, hypertensive, maternal diabetes, pre-eclampsia, gestational diabetes, antenatal visit, year of birth, season, and socioeconomic status | Entire pregnancy: IQR increase: 25 μg/L (1st decile: 3.49–15.51 μg/L, 10th decile: 53.05–119.03 μg/L) | SGA RR (95% CI) for TCM exposure per 25 μg/L (IQR): (1) 1.0 (2) 1.03 (1.00, 1.06) (3) 1.04 (1.02, 1.07) p for trend: 0.094 |
| Toledano et al. (2005) | England (3 water regions) | Retrospective cohort | 1992–1998 | 920,571 mother-infant pairs | Non-integrated exposure assessment based on sampling in local water regions | LBW (<2,500 g) VLBW (<1,500 g) | Infant sex, maternal age, socioeconomic deprivation, and year of study | TCM categories (μg/L): (1) Low <20 (2) Med 20–40 (3) High >40 | T3 TCM exposure for all three water regions combined: LBW OR (95% CI): (1) 1.0 (2) 1.01 (0.96, 1.07) (3) 1.07 (0.99, 1.15) |
| Author                  | Location        | Study design   | Study dates | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels                                                                 | Main findings                                                                 |
|------------------------|-----------------|----------------|-------------|-------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| Villanueva et al. (2011) | Spain (5 regions) | Prospective cohort | 2000–2008  | 2,074 mother–infant pairs | Integrated exposure assessment to estimate total residential update and uptake via various behaviors | SGA (birth weight <10th percentile for gestational age and sex by national growth curves), PTB (<37 weeks completed gestation), LBW (<2,500 g), BW | Maternal age, height, parity, country of origin, paternal weight, smoking during pregnancy, maternal weight gain, gestational age, and maternal social class by occupation during pregnancy | 10% increase in TCM via ingestion, showering, swimming, or all routes | Change in BW (grams) (95% CI) per 10% increase in TCM uptake (μg/d) in T2: Ingestion: -0.44 (-1.01, 0.13) Showering: -0.39 (-1.30, 0.51) Swimming: -0.10 (-0.68, 0.47) Total: -0.45 (-1.36, 0.43) Change in BW (grams) (95% CI) per 10% increase in TCM total residential uptake (μg/d): T1: -0.36 (-1.26, 0.55) T2: -0.45 (-1.36, 0.45) T3: -0.07 (-1.00, 0.85) Entire pregnancy: -0.35 (-1.31, 0.60) PTB OR (95% CI) per 10% increase in TCM uptake (μg/d) in T2: Ingestion: 1.004 (0.995, 1.013) Showering: 0.992 (0.979, 1.006) Swimming: 0.989 (0.978, 0.999) Total: 0.993 (0.980, 1.007) PTB OR (95% CI) per 10% increase in TCM total residential uptake (μg/d): T1: 0.993 (0.979, 1.007) T2: 0.993 (0.980, 1.007) T3: 1.000 (0.985, 1.014) Entire pregnancy: 0.994 (0.979, 1.009) LBW OR (95% CI) per 10% increase in TCM total residential uptake (μg/d): T1: 1.000 (0.986, 1.014) T2: 1.001 (0.988, 1.016) T3: 1.004 (0.990, 1.018) Entire pregnancy: 1.002 (0.988, 1.017) SGA OR (95% CI) per 10% increase in TCM total residential uptake (μg/d): T1: 1.001 (0.993, 1.009) T2: 1.002 (0.994, 1.010) T3: 1.003 (0.995, 1.011) Entire pregnancy: 1.008 (0.994, 1.011) |
| Wennborg et al. (2000)  | Sweden          | Retrospective cohort | 1990–1994  | 654 term births | Non-integrated exposure assessment of occupational history via questionnaire to estimate work with TCM | Term BW (at ≥37 weeks) | Maternal age, consecutive pregnancy number, gestational age, infant sex, father’s laboratory work, maternal smoking, high blood pressure, other chronic diseases, gynecological diseases, sexually transmitted infectious diseases, presence of small children in the home, and previous spontaneous abortions | Laboratory work with TCM (yes or no) | Mean change in birth weight (grams) (95% CI): T2: 27 (-136, 190) |

(Continues)
| Author            | Location                  | Study design                       | Study dates | Sample size | Exposure assessment                                                                 | Outcomes of interest                                                                 | Covariates considered                                                                 | Exposure levels                                                                 | Main findings                                                                 |
|-------------------|---------------------------|------------------------------------|-------------|-------------|--------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Wright et al. (2004) | 109 towns with >10,000 people, Massachusetts, U.S. | Retrospective cohort               | 1995–1998   | 194,827 births | Non-integrated exposure assessment based on sampling of town-level drinking water sources from state data | Term BW (at ≥37 weeks) Gestational age Term SGA (lowest 10% of birth weight per gestational age at ≥37 weeks, sex, and race) PTB (<37 weeks completed gestation) | Diabetes, lung disease, renal disease, chronic hypertension, marital status, previous preterm delivery, previous birth of infant >4,000g, maternal education, parity, prenatal care, number of cigarettes smoked per day, maternal age, maternal race, median household income, and weight gain during pregnancy | TCM 50th, 50–90th, and >90th percentiles (μg/L): (1) 0.0 (2) >26.3 (3) >63 | Term BW (grams) beta (95% CI): T3 (1) 0.0 (2) −19 (3) −30 Cont. −19 (~25, 40) Gestational age (weeks) beta (95% CI): T3 (1) 0.0 (2) 0.0 (3) 0.4 Cont. (0.0, 0.03) Term SGA OR (95% CI): T3 (1) 1.0 (2) 1.05 (3) 1.11 (4) 0.4 (5) 0.1 PTB OR (95% CI): T3 (1) 1.0 (2) 0.95 (3) 0.90 | TCM quantiles (μg/L): (1) 1207 (2) >1207–29.99 (3) >29.99–42.17 (4) >42.17–55.41 (5) >55.41–98.99 Mean (SD): all births = 31.0 (23.6) μg/L Median, all births = 25 μg/L All cardiovascular defects (1) 1.0 (2) 0.73 (3) 0.68 (4) 0.60 (5) 0.77 Transposition of the great arteries (1) 1.0 (2) 2.12 (3) 0.62 (4) 0.53 (5) 0.63 Tetrology of Fallot (1) 1.0 (2) 0.50 (3) 0.76 (4) 0.81 (5) 0.33 (6) 0.07 |
| Wright et al. (2017) | 68 towns in Massachusetts, U.S. | Case-control                        | 1999–2004   | 9,040 controls | Non-integrated exposure assessment based on sampling of town-level drinking water sources from state and public water utility data | Congenital anomalies of the heart and circulatory system (all cardiovascular defects), conotruncal heart defect, tetralogy of Fallot, atrial septal defect, ventricular septal defect, pulmonary stenosis | Type of water source and treatment, infant sex, infant birth weight, maternal weight gain, maternal race, maternal age, maternal education, marital status, maternal smoking, parity, number of previous terminations, prenatal case source payment type, income, and various clinical factors (e.g., abruptio placenta, anemia, cardiomegaly, chronic hypertension, eclampsia, hemoglobinopathy, hepatitis, hydranemia, oligohydramnios, incompetent cervix, labor delivery complications, labor induction, lung disease, lupus, other maternal reproductive risk factors, pharmaceutical inhibition of labor, previous infant >4,000 g, previous infant with birth defect, previous premature or SGA infant, premature or prolonged rupture of membrane, renal disease, Rh sensitization, rubella infection, seizure disorder, sickle cell anemia, uterine bleeding) | TCM quintiles (μg/L): (1) ≤12.07 (2) 12.07–29.99 (3) >29.99–42.17 (4) >42.17–55.41 (5) >55.41–98.99 Mean (SD): all defects = 34.19 (23.56) μg/L Median = 36.07 μg/L | All cardiovascular defects (1) 1.0 (2) 0.73 (3) 0.68 (4) 0.60 (5) 0.77 Tetralogy of Fallot (1) 1.0 (2) 0.50 (3) 0.76 (4) 0.81 (5) 0.33 (6) 0.07 |
| Author          | Location | Study design         | Study dates | Sample size | Exposure assessment                                      | Outcomes of interest                              | Covariates considered                                                                 | Exposure levels                                                                 | Main findings                                                                 |
|-----------------|----------|----------------------|-------------|-------------|----------------------------------------------------------|--------------------------------------------------|----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Wright et al. (2017) |          |                      |             |             |                                                          | Atrial septal defect                             |                                                                                       | (1) 1.0                                                                                                                                   | (2) 1.38 (0.73, 2.60) (3) 1.67 (0.83, 3.38) (4) 1.50 (0.69, 3.26) (5) 1.42 (0.61, 3.27) |
| Zhou et al. (2010) | China    | Retrospective cohort | 2008–2009   | 1,385       | Non-integrated exposure assessment based on sampling of water supply systems | Term BW below median (at ≥37 weeks) | Gestational age, infant sex, maternal age, gravidity, education, number of prenatal visits, season of birth, occupation, prenatal residence, postpartum residence, last menstrual period, parity, illness during pregnancy, maternal height and weight, and congenital anomalies | TCM water concentration quartiles (not specified) | OR (95% CI) TCM water concentration, entire pregnancy: (1) 1.0 (2) 0.96 (0.60, 1.55) (3) 1.45 (0.88, 2.40) (4) 1.64 (0.90, 3.00) |
|                 |          |                      |             |             |                                                          |                                   |                                                                                       | T1: (1) 1.0 (2) 1.74 (1.10, 2.77) (3) 0.90 (0.47, 1.74) (4) 0.89 (0.44, 1.77)               |                                                                                      | (1) 1.0 (2) 1.15 (0.81, 1.63) (3) 0.93 (0.65, 1.32) (4) 1.10 (0.76, 1.57)                                  |
|                 |          |                      |             |             |                                                          |                                   |                                                                                       | T2: (1) 1.0 (2) 1.37 (0.99, 1.88) (3) 1.67 (0.98, 2.85) (4) 1.82 (1.10, 3.02)             |                                                                                      | (1) 1.0 (2) 1.07 (0.71, 1.68) (3) 1.62 (1.05, 2.50) (4) 0.93 (0.54, 1.60)                                  |

Abbreviations: μg/d = micrograms per day; μg/L = micrograms per liter; BMI = body mass index; BW = birth weight; CI = confidence interval; cont. = continuous measure; HR = hazard ratio; IQR = interquartile range; LBW = low birth weight; min = minimum; max = maximum; NA = not applicable; NTD = neural tube defects; OR = OR; PTB = preterm birth; RR = RR; SD = standard deviation; SGA = small for gestational age; T1, T2, T3 = trimesters 1, 2, and 3; TCM = trichloromethane, i.e., TCM; THM = trihalomethane; VLBW = very low birth weight; VPTB = very preterm birth.
| Developmental outcome | References | Relation to TCM exposure |
|-----------------------|------------|--------------------------|
| Congenital anomalies  |            |                          |
| Cardiovascular defect | Brender et al. (2014) | No association (conotruncal defect, P emissions risk value) |
|                       | Brender et al. (2014) | No association (obstructive heart defect, P emissions risk value) |
|                       | Brender et al. (2014) | Adverse association: OR = 1.10 (1,01, 1.19) (septal heart defect, P emissions risk value) |
|                       | Dodds and King (2001) | No association (P water concentration) |
|                       | Grazuleviciene et al. (2013) | No association (T1 integrated) |
|                       | Wright et al. (2017) | No association (any cardiovascular defect, P water concentration) |
|                       | Wright et al. (2017) | No association (conotruncal heart defect, P water concentration) |
|                       | Wright et al. (2017) | No association (transposition of the great arteries, P water concentration) |
|                       | Wright et al. (2017) | No association (tetralogy of Fallot, P water concentration) |
|                       | Wright et al. (2017) | No association (atrial septal defect, P water concentration) |
|                       | Wright et al. (2017) | No association (ventricular septal defect, P water concentration) |
|                       | Wright et al. (2017) | No association (pulmonary stenosis, P water concentration) |
| Neural tube defect    | Brender et al. (2014) | Adverse association: OR = 1.40 (1,04, 1.87) (any neural tube defect, P emissions risk value) |
|                       | Brender et al. (2014) | No association (anecephaly, P emissions risk value) |
|                       | Dodds and King (2001) | No association (P water concentration) |
| Spina bifida          | Brender et al. (2014) | Adverse association: OR = 1.55 (1,10, 2.20) (spina bifida, P emissions risk value) |
|                       | Brender et al. (2014) | Adverse association: OR = 1.74 (1,02, 2.99) for second intensity category, but not third or fourth (spina bifida, P emissions risk intensity) |
| Chromosomal anomaly   | Dodds and King (2001) | Adverse association: RR = 1.9 (1.1, 3.3) for second highest but not highest category (P water concentration) |
| Oral cleft defect     | Brender et al. (2014) | No association (any oral cleft defect, P emissions risk value) |
|                       | Brender et al. (2014) | No association (cleft palate, P emissions risk value) |
|                       | Brender et al. (2014) | No association (cleft lip with or without cleft palate, P emissions risk value) |
|                       | Dodds and King (2001) | No association (P water concentration) |
| Musculoskeletal defect | Brender et al. (2014) | No association (any limb deficiency, P emissions risk value) |
|                       | Brender et al. (2014) | No association (longitudinal limb deficiency, P emissions risk value) |
|                       | Brender et al. (2014) | No association (transverse limb deficiency, P emissions risk value) |
|                       | Grazuleviciene et al. (2013) | No association (T1 integrated) |
| Urogenital defect     | Grazuleviciene et al. (2013) | No association (T1 integrated) |
|                       | Iszatt et al. (2011) | No association (hypospadias, P water concentration) |
|                       | Iszatt et al. (2011) | No association (hypospadias, P ingestion) |
|                       | Iszatt et al. (2011) | No association (hypospadias, P integrated) |
| Gestation             | Costet et al. (2012) | Favorable association: OR = 0.5 (0.3, 0.9) for third quintile but not fourth (T3 water concentration) |
| Preterm birth         | Costet et al. (2012) | No association (T3 integrated) |
|                       | Costet et al. (2012) | No association (T1 ingested) |
|                       | Hinckley et al. (2005) | No association (T3 water concentration) |
|                       | Kugveinuas et al. (2016) | No association (P water concentration) |
|                       | Kugveinuas et al. (2016) | No association (P ingestion) |
|                       | Kugveinuas et al. (2016) | No association (P showers, baths) |
|                       | Kugveinuas et al. (2016) | No association (P integrated) |
|                       | Kramer et al. (1992) | No association (P water concentration) |
|                       | Lewis et al. (2007) | Favorable association: hazard ratio = 0.95 (0.91, 0.99) per 10 mcg/L (P water concentration) |
|                       | Lewis et al. (2007) | No association (T1 water concentration) |
|                       | Lewis et al. (2007) | Favorable association: hazard ratio = 0.95 (0.92, 0.99) per 10 mcg/L (T2 water concentration) |
|                       | Lewis et al. (2007) | No association (T1 integrated) |
|                       | Rivera-Núñez and Wright (2013) | Suggestive evidence of adverse association: OR = 1.06 (1.00, 1.13) for third quintile only, no clear trend (T1 water concentration) |
|                       | Rivera-Núñez and Wright (2013) | Adverse association: OR = 1.08 (1.02, 1.14) for third quintile only, no clear trend (T2 water concentration) |
|                       | Savitz et al. (2005) | Favorable association: OR = 0.54 (0.31, 0.92) for highest quintile; similar association in fourth quintile (T3 water concentration) |
|                       | Savitz et al. (2005) | Favorable association: OR = 0.56 (0.32, 0.96) for third quintile only; similar suggestive association in fifth quintile (T3 integrated) |
|                       | Villanueva et al. (2011) | No association per 10% increase in TCM (T2 ingestion, showering, swimming, or total) |
|                       | Villanueva et al. (2011) | No association per 10% increase in TCM (P integrated) |
|                       | Villanueva et al. (2011) | No association per 10% increase in TCM (T1 integrated) |
|                       | Villanueva et al. (2011) | No association per 10% increase in TCM (T2 integrated) |
|                       | Villanueva et al. (2011) | No association per 10% increase in TCM (T3 integrated) |
|                       | Wright et al. (2004) | Favorable association: OR = 0.80 (0.84, 0.97) for highest decile versus below median, possible trend (T3 water concentration) |
|                       | Wright et al. (2004) | No association (T3 water concentration) |
| Very preterm birth    | Hinckley et al. (2005) | No association (T3 water concentration) |
| Gestational age       | Cao et al. (2016) | Suggestive evidence of favorable association: 0.15 week (~0.03, 0.32) in highest tertile, P for trend = 0.10 (P whole blood TCM) |
|                       | Wright et al. (2004) | Suggestive evidence of favorable association: 0.1 week (0.0, 0.3) per 59 μg/L (T3 water concentration) |

(Continues)
| Developmental outcome | references | Relation to TCM exposure |
|-----------------------|------------|--------------------------|
| Weight                |            |                          |
| Low birth weight      | Danileviciute et al. (2012) | No association (P integrated) |
|                      | Danileviciute et al. (2012) | No association (T1 integrated) |
|                      | Danileviciute et al. (2012) | No association (T2 integrated) |
|                      | Danileviciute et al. (2012) | No association (T3 integrated) |
|                      | Grahuleviciene et al. (2011) | Adverse association: OR = 1.10 (1.01, 1.19) per 0.1 μg/d (P integrated) |
|                      | Grahuleviciene et al. (2011) | Adverse association: OR = 1.10 (1.02, 1.18) per 0.1 μg/d (T1 integrated) |
|                      | Grahuleviciene et al. (2011) | Adverse association: OR = 1.10 (1.01, 1.18) per 0.1 μg/d (T2 integrated) |
|                      | Grahuleviciene et al. (2011) | Adverse association: OR = 1.09 (1.01, 1.18) per 0.1 μg/d (T3 integrated) |
|                      | Izazt et al. (2014) | Adverse association: ~9% (~12, ~5%) decrease in low birth weight with high decline, smaller decrease with low/medium decline pre- vs. post-enhanced coagulation intervention (P ecologic) |
|                      | Kramer et al. (1992) | No association (P water concentration) |
|                      | Toledoano et al. (2005) | Adverse association: OR = 1.10 (1.07, 1.13) for highest (third) versus low category; middle versus low category shows similar but weaker results, possible trend (T3 water concentration) |
|                      | Villanueva et al. (2011) | No association per 10% increase (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T3 integrated) |
| Birth weight at term  | Hinkley et al. (2005) | No association (T3 water concentration) |
|                      | Kogevinas et al. (2016) | No association (P water concentration) |
|                      | Kogevinas et al. (2016) | No association (P ingestion) |
|                      | Kogevinas et al. (2016) | No association (P showers, baths) |
|                      | Lewis et al. (2006) | No association (P water concentration) |
|                      | Lewis et al. (2006) | No association (T1 water concentration) |
|                      | Lewis et al. (2006) | No association (T2 water concentration) |
|                      | Lewis et al. (2006) | No association (T3 water concentration) |
| Low birth weight      | Izazt et al. (2014) | Adverse association: ~16% (~24, ~8%) decrease in very low birth weight with high decline, not for low/medium decline pre- vs. post-enhanced coagulation intervention (P ecologic) |
|                      | Toledoano et al. (2005) | Suggestive evidence of adverse association: OR = 1.07 (0.99, 1.15) for highest (third) versus low category, possible trend (T3 water concentration) |
| Birth weight among term infants | Cao et al. (2016) | Suggestive evidence of adverse association: ~48.23 g (~103.64, 7.19) in highest tertile, p for trend = 0.08 |
|                      | Cao et al. (2016) | (P whole blood TCM) |
|                      | Grahuleviciene et al. (2011) | Adverse association: ~59.1 g (~144.8, ~3.7) per 0.1 μg/d (P integrated) |
|                      | Grahuleviciene et al. (2011) | Adverse association: ~52.8 g (~104.4, ~1.2) per 0.1 μg/d (T1 integrated) |
|                      | Grahuleviciene et al. (2011) | No association (T2 integrated) |
|                      | Grahuleviciene et al. (2011) | Adverse association: ~57.8 g (~111.6, ~4.0) per 0.1 μg/d (T3 integrated) |
|                      | Kogevinas et al. (2016) | No association (P water concentration) |
|                      | Kogevinas et al. (2016) | No association (P ingestion) |
|                      | Kogevinas et al. (2016) | No association (P showers, baths) |
|                      | Lewis et al. (2006) | No association (P integrated) |
|                      | Lewis et al. (2006) | Adverse association: ~12 g (~21, ~3) per 10 μg/L (P, water concentration) |
|                      | Lewis et al. (2006) | Suggestive evidence of adverse association: ~7 g (~14, 0) per 10 μg/L (T1, water concentration) |
|                      | Lewis et al. (2006) | No association per 10 μg/L (T2, water concentration) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T2 ingestion, showering, swimming, or total) |
|                      | Villanueva et al. (2011) | No association per 10% increase (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T3 integrated) |
| Birth weight          | Hoffman et al. (2008a) | No association (T3 water concentration) |
|                      | Rivera-Núñez and Wright (2013) | Adverse association: ~19 g (~25, ~12) for highest quintile, possible trend (T2 water concentration) |
|                      | Rivera-Núñez and Wright (2013) | Adverse association: ~15 g (~21, ~8) for highest quintile, possible trend (T3 water concentration) |
|                      | Savitz et al. (2005) | No association (T3 water concentration) |
|                      | Savitz et al. (2005) | No association (T3 integrated) |
|                      | Smith et al. (2016) | No association (P integrated) |
|                      | Smith et al. (2016) | Adverse association: ~26.5 g (~49.3, ~3.8) for middle tertile only, p for trend = 0.13 (T1 integrated) |
|                      | Smith et al. (2016) | No association (T2 integrated) |
|                      | Smith et al. (2016) | No association (T3 integrated) |
|                      | Summerhayes et al. (2012) | Adverse association: ~5.0 (~8.6, ~1.4) per 25 μg/L (P water concentration) |
|                      | Summerhayes et al. (2012) | Adverse association: ~5.0 (~8.6, ~1.4) per 25 μg/L (T1 water concentration) |
|                      | Summerhayes et al. (2012) | Adverse association: ~3.4 (~6.4, ~0.3) per 25 μg/L (T2 water concentration) |
|                      | Summerhayes et al. (2012) | Suggestive evidence of adverse association: ~2.6 (~5.8, 0.6) per 25 μg/L (T3 water concentration) |
|                      | Wennborg et al. (2000) | No association with occupational exposure (P occupational) |
|                      | Wright et al. (2004) | Adverse association: ~19 g (~25, ~14) per 59 μL (T3 water concentration) |
|                      | Zhou et al. (2010) | No association (P water concentration) |
|                      | Zhou et al. (2010) | Adverse association: OR = 1.74 (1.10, 2.77) for below-median birth weight in second quartile only (T1 water concentration) |
|                      | Zhou et al. (2010) | No association (T2 water concentration) |
|                      | Zhou et al. (2010) | Adverse association: OR = 1.82 (1.10, 3.02) for below-median birth weight in highest quartile (T3 water concentration) |
|                      | Zhou et al. (2010) | Adverse association: OR = 1.62 (1.05, 2.65) for below-median birth weight in third quartile only (T1 and T2 water concentration) |
| Birth length          | Cao et al. (2016) | No association (P whole blood TCM) |
| Postnatal weight gain | Botton et al. (2015) | No association (P integrated) |
|                      | Botton et al. (2015) | No association (P ingestion) |

(Continues)
TABLE 7 (Continued)

| Developmental outcome | References | Relation to TCM exposure |
|-----------------------|------------|--------------------------|
| Growth Small for gestational age | Cao et al. (2016) | No association (P whole blood; TCM) |
|                      | Costet et al. (2012) | No association (T3 water concentration)c |
|                      | Costet et al. (2012) | No association (T3 integrated)d |
|                      | Costet et al. (2012) | No association (T1 ingested)ε |
|                      | Danileviciute et al. (2012) | No association (P integrated) |
|                      | Danileviciute et al. (2012) | No association (T1 integrated) |
|                      | Danileviciute et al. (2012) | No association (T2 integrated) |
|                      | Danileviciute et al. (2012) | No association (T3 integrated) |
|                      | Graziuleviciene et al. (2011) | Suggestive evidence of adverse association: OR = 1.03 (0.99, 1.08) per 0.1 μg/d (P integrated) |
|                      | Graziuleviciene et al. (2011) | Suggestive evidence of adverse association: OR = 1.03 (0.99, 1.08) per 0.1 μg/d (T1 integrated) |
|                      | Graziuleviciene et al. (2011) | Suggestive evidence of adverse association: OR = 1.03 (0.99, 1.08) per 0.1 μg/d (T2 integrated) |
|                      | Graziuleviciene et al. (2011) | Suggestive evidence of adverse association: OR = 1.04 (1.00, 1.09) per 0.1 μg/d (T3 integrated) |
|                      | Hinckley et al. (2005) | No association (T3 water concentration) |
|                      | Hoffman et al. (2008a) | No association (T3 water concentration) |
|                      | Infante-Rivard (2004) | No association (P water concentration) |
|                      | Kogevinas et al. (2016) | No association (P water concentration) |
|                      | Kogevinas et al. (2016) | No association (P ingestion) |
|                      | Kogevinas et al. (2016) | No association (P showers, baths) |
|                      | Kogevinas et al. (2016) | No association (P integrated) |
|                      | Kramer et al. (1992) | Adverse association: OR = 1.8 (1.1, 2.9) for highest category (≥10 mcg/L) (P water concentration)ε |
|                      | Porter et al. (2005) | Adverse association: OR = 1.24 (1.02, 1.50) for second quintile only, no clear trend (P water concentration) |
|                      | Porter et al. (2005) | No association (T1 water concentration) |
|                      | Porter et al. (2005) | No association (T2 water concentration) |
|                      | Porter et al. (2005) | No association (T3 water concentration) |
|                      | Savitz et al. (2005) | No association (T3 water concentration) |
|                      | Savitz et al. (2005) | No association (T3 integrated) |
|                      | Summerhayes et al. (2012) | Adverse association: RR = 1.03 (1.02, 1.05) per 25 μg/L (P water concentration)ε |
|                      | Summerhayes et al. (2012) | Suggestive evidence of adverse association: RR = 1.03 (1.00, 1.05) per 25 μg/L (T1 water concentration)ε |
|                      | Summerhayes et al. (2012) | Adverse association: RR = 1.04 (1.02, 1.06) per 25 μg/L (T2 water concentration)ε |
|                      | Summerhayes et al. (2012) | Adverse association: RR = 1.04 (1.02, 1.07) per 25 μg/L (T3 water concentration)ε |
|                      | Villanueva et al. (2011) | No association per 10% increase (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase (T3 integrated) |
|                      | small for gestational age at term | No association (T3 water concentration)c |
|                      | Levallois et al. (2016) | No association (P water concentration) |
|                      | Levallois et al. (2016) | No association (T3 water concentration)c |
|                      | Levallois et al. (2016) | No association (T3 integrated)d |
|                      | Levallois et al. (2012) | No association (T3 water concentration)c |
|                      | Levallois et al. (2012) | No association (T3 ingested)ε |
|                      | Rivera-Núñez and Wright (2013) | No association (T2 water concentration) |
|                      | Rivera-Núñez and Wright (2013) | Suggestive evidence of adverse association: OR = 1.04 (1.00, 1.10) for fourth quintile; fifth quintile is similar (T3 water concentration) |
|                      | Wright et al. (2004) | Adverse association: OR = 1.11 (1.04, 1.17) for highest decile versus below median, possible trend (T3 water concentration) |

Abbreviations: CI = confidence interval; OR = odds ratio; NA = not applicable; P = pregnancy (entire); RR = relative risk; T1, T2, T3 = trimesters 1, 2, and 3; TCM = trichloromethane, i.e., chloroform

*When results are available for continuous and categorical classifications of the same exposure, continuous results are prioritized.

**Ecologic study; no individual measurements for exposure.

†Most studies used a 10 μg/L increase in exposure as the unit of analysis, so a 25 μg/L (Summerhayes et al., 2012) or 59 μg/L (Wright et al., 2004) increase may show more pronounced associations. Summerhayes et al. (2012) also compared the highest decile with the lowest decile, which showed similarly adverse associations of a slightly stronger magnitude (ORs between 1.07 and 1.12). Wright et al. (2004) also compared the highest decile with below-median values, resulting in similarly adverse associations (~18 [95% CI: −26, −10]).

‡Small for gestational age has cut-off at 5th percentile of adjusted birth weight, whereas most other studies use 10th percentile.

§Levallois et al. (2012) and Levallois et al. (2016) are based on the same data set.

Lewis et al., 2006; Lewis et al., 2007; Porter, PutnamD, Hunting, & Riddle, 2005; Rivera-Núñez & Wright, 2013; Summerhayes et al., 2012; Toledano et al., 2005; Wennborg et al., 2000; Wright et al., 2004; Wright et al., 2017; Zhou et al., 2010). While exposure assessment is one of many aspects that influence the quality of any study, given the critical importance of exposure assessment in this subject matter, we determined that the use of an integrated (i.e., individualized) exposure assessment method should be required of any study to be considered informative on the potential causal relationship between chloroform exposure and developmental outcomes (Koepsell & Weiss, 2003; Morgenstern, 1995).

Of note, outcome misclassification is seldom a major concern in epidemiologic research on birth outcomes, because such measures (e.g., preterm birth, small-for-gestational age, and low birth weight) are routinely and easily measured based on standard practices with little potential for systematic error (i.e., bias). When mismeasurement does occur, it is typically in estimating gestational age
### TABLE 8

Summary of epidemiological evidence for associations between integrated chloroform exposure and developmental outcomes. Bold font indicates adverse associations; underlined font indicates favorable associations. "Integrated" refers to exposure assessment that incorporates water concentration and behaviors to account for dermal absorption, ingestion, and inhalation.

| Developmental outcome | References | Relation to TCM exposure |
|-----------------------|------------|--------------------------|
| Congenital anomalies  | Grazuleviciene et al. (2013) | No association (T1 integrated) |
| Cardiovascular defect | Grazuleviciene et al. (2013) | No association (T1 integrated) |
| Neural tube defect    | Grazuleviciene et al. (2013) | No association (T1 integrated) |
| Spina bifida          | Grazuleviciene et al. (2013) | No association (T1 integrated) |
| Chromosomal anomaly   | Grazuleviciene et al. (2013) | No association (T1 integrated) |
| Oral cleft defect     | Grazuleviciene et al. (2013) | No association (T1 integrated) |
| Musculoskeletal defect| Iszatt et al. (2011) | No association (T1 integrated) |
| Urogenital defect      | Grazuleviciene et al. (2013) | No association (T1 integrated) |
|                        | Iszatt et al. (2011) | No association (T1 integrated) |
|                      | Costet et al. (2012) | No association (T3 integrated) |
|                      | Savitz et al. (2005) | Favorable association: OR = 0.56 (0.32, 0.96) for third quintile only; similar suggestive association in fifth quintile (T3 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T3 integrated) |
| Gestation             | Preterm birth | Costet et al. (2012) | No association (T3 integrated) |
|                      | Kogevinas et al. (2016) | No association (P integrated) |
|                      | Savitz et al. (2005) | Favorable association: OR = 1.10 (1.02, 1.18) per 0.1 μg/d (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T3 integrated) |
|                      | Very preterm birth | Cao et al. (2016) | Suggestive evidence of favorable association: 0.15 week (−0.03, 0.32) in highest tertile, p for trend = 0.10 (P whole blood TCM) |
| Weight                | Low birth weight | Danileviciute et al. (2012) | No association (P integrated) |
|                      | Danileviciute et al. (2012) | No association (T1 integrated) |
|                      | Danileviciute et al. (2012) | No association (T2 integrated) |
| Grazuleviciene et al. (2011) | Adverse association: OR = 1.10 (1.01, 1.19) per 0.1 μg/d (P integrated) |
|                      | Grazuleviciene et al. (2011) | Adverse association: OR = 1.10 (1.02, 1.18) per 0.1 μg/d (T1 integrated) |
|                      | Grazuleviciene et al. (2011) | Adverse association: OR = 1.09 (1.01, 1.18) per 0.1 μg/d (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T3 integrated) |
|                      | Low birth weight at term | Kogevinas et al. (2016) | No association (P integrated) |
|                      | Very low birth weight | Cao et al. (2016) | Suggestive evidence of adverse association: −48.23 g (−103.64, 7.19) in highest tertile, p for trend = 0.08 (P whole blood TCM) |
| Birth weight         | Cao et al. (2016) | Suggestive evidence of adverse association: −59.1 g (−114.8, 3.7) per 0.1 μg/d (P integrated) |
|                      | Grazuleviciene et al. (2011) | Adverse association: −59.1 g (−114.8, 3.7) per 0.1 μg/d (T1 integrated) |
|                      | Grazuleviciene et al. (2011) | Adverse association: −52.8 g (−104.4, −1.2) per 0.1 μg/d (T2 integrated) |
|                      | Grazuleviciene et al. (2011) | Adverse association: −57.8 g (−111.6, −4.0) per 0.1 μg/d (T3 integrated) |
|                      | Kogevinas et al. (2016) | No association (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (P integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T1 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T2 integrated) |
|                      | Villanueva et al. (2011) | No association per 10% increase in TCM (T3 integrated) |
| Birth weight among term infants | Savitz et al. (2005) | No association (T3 integrated) |
|                      | Smith et al. (2016) | No association (P integrated) |
| Smith et al. (2016)  | Adverse association: −26.5 g (−49.3, −3.8) for middle tertile only, p for trend = 0.13 (T1 integrated) |
|                      | Smith et al. (2016) | No association (T2 integrated) |
|                      | Smith et al. (2016) | No association (T3 integrated) |
| Birth length         | Cao et al. (2016) | No association (P whole blood TCM) |
| Postnatal weight gain| Botton et al. (2015) | No association (P integrated) |
TABLE 8 (Continued)

| Developmental outcome | References                     | Relation to TCM exposure |
|-----------------------|--------------------------------|--------------------------|
| Growth                | Cao et al. (2016)              | No association (P whole blood TCM) |
|                       | Costet et al. (2012)           | No association (T3 integrated) |
|                       | Danileviciute et al. (2012)    | No association (P integrated) |
|                       | Danileviciute et al. (2012)    | No association (T1 integrated) |
|                       | Danileviciute et al. (2012)    | No association (T2 integrated) |
|                       | Grazuleviciene et al. (2011)    | Suggestive evidence of adverse association: OR = 1.03 (0.99, 1.08) per 0.1 μg/d (P integrated) |
|                       | Grazuleviciene et al. (2011)    | Suggestive evidence of adverse association: OR = 1.03 (0.99, 1.08) per 0.1 μg/d (T1 integrated) |
|                       | Grazuleviciene et al. (2011)    | Suggestive evidence of adverse association: OR = 1.03 (0.99, 1.08) per 0.1 μg/d (T2 integrated) |
|                       | Grazuleviciene et al. (2011)    | Suggestive evidence of adverse association: OR = 1.04 (1.00, 1.09) per 0.1 μg/d (T3 integrated) |
|                       | Kogevinas et al. (2016)        | No association (P integrated) |
|                       | Savitz et al. (2005)           | No association (T3 integrated) |
|                       | Villanueva et al. (2011)       | No association per 10% increase (P integrated) |
|                       | Villanueva et al. (2011)       | No association per 10% increase (T1 integrated) |
|                       | Villanueva et al. (2011)       | No association per 10% increase (T2 integrated) |
|                       | Villanueva et al. (2011)       | No association per 10% increase (T3 integrated) |
| Small for gestational age at term | Levallois et al. (2016) | No association (T3 integrated) |
|                       | Levallois et al. (2012)        | No association (T3 integrated) |

Abbreviations: CI = confidence interval; OR = odds ratio; NA = not applicable; P = pregnancy (entire); RR = relative risk; T1, T2, T3 = trimesters 1, 2, and 3; TCM = trichloromethane, i.e., chloroform.

*a*When results are available for continuous and categorical classifications of the same exposure, continuous results are prioritized.

*b*Small for gestational age has cut-off at 5th percentile of adjusted birth weight, whereas most other studies use 10th percentile.

*c*Levallois et al. (2012) and Levallois et al. (2016) are based on the same data set.

Based on the date of last menstrual period (Dietz et al., 2007; Howards, Hertz-Picciotto, Weinberg, & Poole, 2006). The use of ultrasound measures taken early in pregnancy can minimizes this mismeasurement (Howards et al., 2006). Even if this correction does not occur, misclassification of gestational age is unlikely to differ systematically based on chloroform exposure level as prior research has shown chloroform not to be associated with menstrual cycle length and timing (Howards et al., 2006; Windham et al., 2003).

**Biomarker study**

Estimation of chloroform exposure in 29 of the 30 studies included in this review was highly dependent upon self-report of behaviors such as showering, bathing, hand-washing, use of a water filter, amount of tap water consumption, and/or non-individualized regional data on chloroform concentrations in water. Biological markers (biomarkers), by contrast, have the advantage of being independent of self-reporting error or bias and can potentially provide an objective measure of average exposure during a given time period. However, biomarkers can be limited for exposure assessment, for example, if they provide information on only recent exposure when long-term exposure is of etiologic interest; if they do not capture all sources or routes of exposure; or if they are not specific to the exposure of concern (Grandjean, 1995; Rothman, Greenland, & Lash, 2008).

Only one epidemiologic study assessed exposure to chloroform using a biomarker—namely Cao et al. (2016). This study used maternal whole blood samples taken at the time of delivery prior to the infant’s birth as a basis for estimating chloroform exposure. The median concentration of chloroform in the blood was 50.7 ng/L. Higher levels of chloroform in maternal prenatal blood were not significantly associated with birth weight, birth length, or gestational age at birth in this study (adjusted mean difference in birth weight between highest and lowest tertiles of chloroform = −48.23 g; 95% confidence interval (CI) = −103.64, 7.19).

**Occupational exposure**

Studying the risk of developmental outcomes among mothers exposed to chloroform as part of their occupation can be particularly insightful, because occupational exposure levels tend to be higher than those encountered in non-occupational settings. Any health effects may be more readily detected in highly exposed populations than those with lower levels of exposure that may be too subtle to cause overt effects. However, the occupational context is not always generalizable to the public and may better inform health and safety measures to be taken only for those exposed at the workplace. The single study that examined occupational chloroform exposure in relation to developmental outcomes (Wennborg et al., 2000) was conducted in Sweden. Study investigators identified potentially chloroform-exposed mothers using workplace information from government employment records, and contacted them with a request to complete a questionnaire to assess occupational exposure. Specifically, mothers who worked in biomedical research laboratories at major Swedish universities, along with a comparison group of mothers who worked in non-laboratory departments at the same universities, were
TABLE 9  Summary of epidemiological evidence for associations between chloroform exposure and developmental outcomes among studies conducted in the United States. Bold font indicates adverse associations; underlined font indicates favorable associations. "Integrated" refers to exposure assessment that incorporates water concentration and behaviors to account for dermal absorption, ingestion, and inhalation.

| Developmental outcome | References | Relation to TCM exposure |
|-----------------------|------------|--------------------------|
| Congenital Anomalies  |            |                          |
| Cardiovascular defect | Brender et al. (2014) | No association (conotruncal defect, P emissions risk value) |
| Brender et al. (2014) | Adverse association: OR = 1.10 (1.01, 1.19) (septal heart defect, P emissions risk value) |
| Wright et al. (2017) | No association (any cardiovascular defect, P water concentration) |
| Wright et al. (2017) | No association (conotruncal heart defect, P water concentration) |
| Wright et al. (2017) | No association (transposition of the great arteries, P water concentration) |
| Wright et al. (2017) | No association (tetralogy of Fallot, P water concentration) |
| Wright et al. (2017) | No association (atrial septal defect, P water concentration) |
| Wright et al. (2017) | No association (ventricular septal defect, P water concentration) |
| Wright et al. (2017) | No association (pulmonary stenosis, P water concentration) |
| Neural tube defect    | Brender et al. (2014) | Adverse association: OR = 1.40 (1.04, 1.87) (any neural tube defect, P emissions risk value) |
| Brender et al. (2014) | No association (anecephaly, P emissions risk value) |
| Spina bifida          | Brender et al. (2014) | Adverse association: OR = 1.55 (1.10, 2.10) (spina bifida, P emissions risk value) |
| Brender et al. (2014) | Adverse association: OR = 1.74 (1.02, 2.99) for second intensity category, but not third or fourth (spina bifida, P emissions risk intensity) |
| Chromosomal anomaly   |            |                          |
| Oral cleft defect     | Brender et al. (2014) | No association (any oral cleft defect, P emissions risk value) |
| Brender et al. (2014) | No association (cleft palate, P emissions risk value) |
| Brender et al. (2014) | No association (cleft lip with or without cleft palate, P emissions risk value) |
| Musculoskeletal defect| Brender et al. (2014) | No association (any limb deficiency, P emissions risk value) |
| Brender et al. (2014) | No association (longitudinal limb deficiency, P emissions risk value) |
| Brender et al. (2014) | No association (transverse limb deficiency, P emissions risk value) |
| Urogenital defect     |            |                          |
| Gestation             |            |                          |
| Preterm birth         |            |                          |
| Hinkley et al. (2005) | No association (T3 water concentration) |
| Kramer et al. (1992)  | No association (P water concentration) |
| Lewis et al. (2007)   | No association (T1 water concentration) |
| Lewis et al. (2007)   | Favorable association: hazard ratio = 0.95 (0.91, 0.99) per 10 mcg/L (P water concentration) |
| Lewis et al. (2007)   | Favorable association: hazard ratio = 0.95 (0.92, 0.99) per 10 mcg/L (T2 water concentration) |
| Lewis et al. (2007)   | No association (T3 four-week risk sets water concentration) |
| Rivera-Núñez and Wright (2013) | Suggestive evidence of adverse association: OR = 1.06 (1.00, 1.13) for third quintile only, no clear trend (T1 water concentration) |
| Rivera-Núñez and Wright (2013) | Adverse association: OR = 1.08 (1.02, 1.14) for third quintile only, no clear trend (T2 water concentration) |
| Savitz et al. (2005)  | Favorable association: OR = 0.54 (0.31, 0.92) for highest quintile; similar association in fourth quintile (T3 water concentration) |
| Savitz et al. (2005)  | Favorable association: OR = 0.56 (0.32, 0.96) for third quintile only; similar suggestive association in fifth quintile (T3 integrated) |
| Wright et al. (2004)  | Favorable association: OR = 0.90 (0.84, 0.97) for highest decile versus below median, possible trend (T3 water concentration) |
| Very preterm birth    |            |                          |
| Hinkley et al. (2005) | No association (T3 water concentration) |
| Gestational age       |            |                          |
| Weight                |            |                          |
| Low birth weight      |            |                          |
| Kramer et al. (1992)  | No association (P water concentration) |
|  | No association (T3 water concentration) |
| Lewis et al. (2005)   | No association (P water concentration) |
| Lewis et al. (2006)   | No association (T1 water concentration) |
| Lewis et al. (2006)   | Suggestive evidence of adverse association: OR = 1.08 (1.00, 1.17) per 10 μg/L (T2 water concentration) |
| Lewis et al. (2006)   | No association (T3 water concentration) |
| Very low birth weight |            |                          |
| Weight                |            |                          |
| Lewis et al. (2006)   | Adverse association: −12 g (−21, −3) per 10 μg/L (P, water concentration) |
| Lewis et al. (2006)   | Suggestive evidence of adverse association: −7 g (−14, 0) per 10 μg/L (T1, water concentration) |
| Lewis et al. (2006)   | No association per 10 μg/L (T2, water concentration) |
| Birth weight among term infants | Hoffman et al. (2008a) | No association (T3 water concentration) |
| Rivera-Núñez and Wright (2013) | Adverse association: −19 g (−25, −12) for highest quintile, possible trend (T2 water concentration) |
| Rivera-Núñez and Wright (2013) | Adverse association: −15 g (−21, −8) for highest quintile, possible trend (T3 water concentration) |
| Savitz et al. (2005)  | No association (T3 water concentration) |
| Savitz et al. (2005)  | No association (T3 integrated) |
| Wright et al. (2004)  | Adverse association: −19 g (−25, −14) per 59 μg/L (T3 water concentration) |

(Continues)
asked about whether they performed laboratory work with chloroform during pregnancy. This study found no significant difference in mean birth weight between the infants of mothers exposed to chloroform at work and those that were not exposed to chloroform at work (adjusted mean difference = 27 g; 95% CI = −136, 190).

No epidemiologic study to date has examined the relationship between paternal occupational or environmental exposure to chloroform and developmental outcomes. Additional research may be warranted to address this gap in the scientific literature.

### 3.3.3 Summary of the epidemiologic data by developmental endpoint

#### Birth defects

Five epidemiologic studies investigated the risk of birth defects in association with chloroform exposure (Brender et al., 2014; Dodds & King, 2001; Grazuleviciene et al., 2013; Iszatt et al., 2011; Wright et al., 2017) (Table 7). Of the 30 associations evaluated in these studies, 25 associations were statistically null, and only 5 statistically significant associations were detected. Dodds and King (2001) observed a significantly increased risk of chromosomal anomaly when comparing the second highest level of exposure to the lowest level (risk ratio (RR) = 1.9; 95% CI = 1.1, 3.3). No significant increase in risk was evident when comparing the highest exposure category to the lowest categories, such that no positive exposure-response gradient was observed. In light of the absent exposure-response trend, and because chloroform has not been found to be genotoxic (EPA, 2001), this association is likely spurious.

Brender et al. (2014) is unique among the epidemiologic studies, as it examined ambient air chloroform exposure attributed to industrial sources. This study estimated associations of airborne chloroform exposure with various limb defects, neural tube defects, and cardiovascular defects that were ascertained from the Texas Birth Defects Registry. Exposure assessment used a modified emission weighted proximity model that took into consideration emissions from all sources within a 10-km buffer around each maternal residence, and the amounts of chemical released from each source during each pregnancy. The model produced a risk value for each mother–infant pair. Having a risk value greater than 0 for ambient chloroform exposure was associated with a 10% increase in the odds of having an infant with a neural tube defect (odds ratio (OR) = 1.10; 95% CI = 1.01, 1.19). A risk value greater than 0 for ambient chloroform exposure was associated with 1.40 times the odds of having an infant with any form of a neural tube defect (95% CI = 1.04, 1.87) and with 1.55 times the odds of having an infant with spina bifida (95% CI = 1.10, 2.20). When parsing risk values into intensity levels (four categories of the risk value), Brender et al. (2014) found a significant positive association of high intensity ambient chloroform exposure with spina bifida occurrence (OR = 1.74; 95% CI = 1.02, 2.99); however, moderate and high levels of exposure were not significantly associated with spina bifida occurrence, indicating the absence of a positive exposure-response trend. Wright et al. (2017) examined seven different congenital heart defects, none of which were associated with chloroform concentrations in water. In contrast, two relatively high-quality studies of birth defects used integrated exposure measures of chloroform (Grazuleviciene et al., 2013; Iszatt et al., 2011); both showed no significant association between chloroform

### TABLE 9 (Continued)

| Developmental outcome | References | Relation to TCM exposure |
|-----------------------|------------|--------------------------|
| Growth                |            |                          |
| Small for gestational age |            |                          |
| Hincley et al. (2005)  | No association (T3 water concentration) |
| Hoffman et al. (2008a) | No association (T3 water concentration) |
| Kramer et al. (1992)  | Adverse association: OR = 1.8 (1.1, 2.9) for highest category (≥10 mcg/L) (P water concentration)* |
| Porter et al. (2005)  | Adverse association: OR = 1.24 (1.02, 1.50) for second quintile only, no clear trend (P water concentration) |
| Porter et al. (2005)  | No association (T1 water concentration) |
| Porter et al. (2005)  | No association (T2 water concentration) |
| Porter et al. (2005)  | No association (T3 water concentration) |
| Savitz et al. (2005)  | No association (T3 water concentration) |
| Savitz et al. (2005)  | No association (T3 integrated) |
| Small for gestational age at term |            |                          |
| Rivera-Núñez and Wright (2013) | No association (T2 water concentration) |
| Rivera-Núñez and Wright (2013) | Suggestive evidence of adverse association: OR = 1.04 (1.00, 1.10) for fourth quintile; fifth quintile is similar (T3 water concentration) |
| Wright et al. (2004)  | Adverse association: OR = 1.11 (1.04, 1.17) for highest decile versus below median, possible trend (T3 water concentration) |

Abbreviations: CI = confidence interval; OR = odds ratio; NA = not applicable; P = pregnancy (entire); RR = relative risk; T1, T2, T3 = trimesters 1, 2, and 3; TCM = trichloromethane, i.e., chloroform.

*When results are available for continuous and categorical classifications of the same exposure, continuous results are prioritized.

*Most studies used a 10 µg/L increase in exposure as the unit of analysis, so a 25 µg/L (Summerhayes et al., 2012) or 59 µg/L (Wright et al., 2004) increase may show more pronounced associations. Summerhayes et al. (2012) also compared the highest decile with the lowest decile, which showed similarly adverse associations of a slightly stronger magnitude (ORs between 1.07 and 1.12). Wright et al. (2004) also compared the highest decile with below-median values, resulting in similarly adverse associations (−18 [95% CI: −26, −10]).

*Small for gestational age has cut-off at 5th percentile of adjusted birth weight, whereas most other studies use 10th percentile.
exposure and cardiovascular, musculoskeletal, or urogenital birth defects (Table 8).

Considering the available published epidemiologic studies on chloroform exposure and birth defects, we conclude that the weight of the evidence does not demonstrate an association between chloroform exposure during pregnancy and a risk of birth defects. This conclusion is based on the paucity of statistically significant associations in light of the number of associations tested, the lack of evidence for a positive exposure-response relationship, and the lack of a significant association in both studies with integrated exposure. Given that birth defects comprise a highly heterogeneous group of outcomes that should be examined separately for etiologic effects, additional insight could potentially be gained from larger studies with integrated exposure assessment that examine specific birth defects individually.

Age: Preterm birth or gestational age at birth
Preterm birth (birth prior to 37 weeks’ completed gestation) is a commonly studied outcome measure in developmental epidemiology. Variations of this outcome measure include very preterm birth (birth prior to 32 weeks completed gestation) and gestational age (a continuous measure based on time since last menstrual period or on ultrasound guided estimation). Ten epidemiologic studies, representing seven independent populations, investigated the risk of preterm birth or shorter gestational age in association with exposure to chloroform (Table 7). Three of these studies were based on separate, but overlapping, samples of Massachusetts residents during different study periods (Lewis et al., 2007; Rivera-Núñez & Wright, 2013; Wright et al., 2004). Five of the independent studies found no significant association between chloroform exposure and preterm birth or shorter gestational age (Cao et al., 2016; Hinckley et al., 2005; Kogevinas et al., 2016; Kramer et al., 1992; Villanueva et al., 2011). Rivera-Núñez and Wright (2013) observed a significant association with higher risk of preterm birth when comparing the 3rd quintile of chloroform levels in water to the 1st (lowest) quintile (OR = 1.08; 95% CI = 1.02, 1.14), but no clear exposure-response trend was observed across the five quintiles. Four studies showed a significantly reduced risk of preterm birth associated with higher chloroform exposure (Wright et al., 2004 [OR = 0.90; 95% CI = 0.84, 0.97 comparing >90th to <50th percentile of exposure]; Savitz et al., 2005 [OR = 0.56; 95% CI = 0.32, 0.96 comparing 3rd to 1st quintile of exposure]; Lewis et al., 2007 [hazard ratio (HR) = 0.95; 95% CI = 0.91, 0.99 per 10 mcg/L increase]; Costet et al., 2012 [OR = 0.5; 95% CI = 0.3, 0.9]), including possible inverse exposure-response trends in three of these studies (Lewis et al., 2007; Savitz et al., 2005; Wright et al., 2004). Of the two studies that examined gestational age as a continuous measure, one found no significant association with chloroform exposure (Cao et al., 2016) and another found a significant positive association with the >90th percentile of exposure, but not with continuous exposure (Wright et al., 2004). No studies examined associations with post-term birth (≥42 weeks) explicitly.

Among the five relatively high-quality studies of preterm birth or continuous gestational age in association with integrated chloroform exposure (i.e., incorporates water concentration and behaviors from dermal absorption, ingestion, and inhalation; see Table 8), no study found a significant positive (i.e., adverse) association with preterm birth or gestational age, and one study (Savitz et al., 2005) found a significant inverse (i.e., protective) association.

The preponderance of the epidemiologic evidence on chloroform exposure and preterm birth suggests no association, or even a potential protective association, between chloroform exposure and preterm birth. Evidence in favor of a protective effect includes a significant inverse association detected in four of nine studies, including three studies (one with integrated exposure assessment) that provided evidence in support of an inverse exposure-response relationship. A biologically plausible mechanism to explain a beneficial effect of chloroform exposure on preterm birth or gestational age is lacking, although the absence of such a mechanism does not preclude a causal effect. Therefore, the observed inverse association detected in multiple studies could be due to confounding by a protective factor (e.g., higher socioeconomic status and healthier lifestyle) associated with chloroform exposure or due to a beneficial effect of chloroform itself (e.g., decreased risk of maternal waterborne infection).

Fetal growth: Low birth weight or continuous birth weight
Low birth weight, defined as weight at birth less than 2,500 g, has historically been used as a straightforward and well-recorded indicator of poor health at birth and increased risk of infant mortality. Nine studies from eight independent populations examined chloroform exposure as it relates to the risk of low birth weight (including low birth weight at term, i.e., restricted to full-term births) or very low birth weight (<1,500 g) (Table 7). Six studies showed no significant association of low birth weight (in various forms) with the level of chloroform exposure during pregnancy (Danileviciute et al., 2012; Hinckley et al., 2005; Kogevinas et al., 2016; Kramer et al., 1992; Lewis et al., 2006; Villanueva et al., 2011). Three studies, in contrast, found a significantly increased risk of low birth weight in association with chloroform exposure (Grazuleviciene et al., 2011; Iszatt et al., 2014; Toledano et al., 2005); all three of these studies provided evidence for a positive exposure-response relationship.

Toledano et al. (2005) found an increased risk of low birth weight in association with higher chloroform water concentrations in a U.K. sample of mothers (OR = 1.10; 95% CI = 1.07, 1.13). The association with very low birth weight was somewhat attenuated and of borderline statistical
of low birth weight for an increase in chloroform exposure were less likely to have an infant born with very low birth weight, including after the analysis was restricted to infants born at term (16% decrease; 95% CI = 24, 8%); similar results were observed among low-birth-weight infants (9% decrease; 95% CI = 12, 5%).

Grazuleviciene et al. (2011), a study based in Lithuania, observed a statistically significant 10% increase in the odds of low birth weight for an increase in chloroform exposure of 0.1 μg/day during pregnancy (1.10 per 0.1 μg/d increase; 95% CI = 1.01, 1.19) across all routes of exposure, that is, drinking, showering, and so forth. By contrast, a study using a subset of data from the same population (Danileviciute et al., 2012), including participants who had provided samples for genetic analysis, did not observe a significant association between chloroform exposure and risk of low birth weight overall. However, Danileviciute et al. (2012) observed an increased odds of low birth weight in the absence of a certain maternal polymorphism (GSTM1) that could lend an individual more susceptible to environmental exposures.

Four studies of low birth weight employed integrated exposure methods (Table 8; Danileviciute et al., 2012; Grazuleviciene et al., 2011; Kogevinas et al., 2016; Villanueva et al., 2011), including two studies of the same population (Danileviciute et al., 2012; Grazuleviciene et al., 2011). Only Grazuleviciene et al. (2011) found a significant positive association between chloroform exposure and increased risk of low birth weight overall.

Considering the available epidemiologic studies on chloroform exposure and low birth weight, we found that three studies found a positive association between chloroform exposure and increased risk of low birth weight, along with a possible exposure-response trend. However, two of the three studies had core limitations: Toledano et al. (2005) had insufficient control for confounding factors in their analysis and Iszatt et al. (2014) was an ecologic study. The majority of studies—including three of four studies with integrated exposure assessment—did not find any clear association. Given the methodological limitations of most of the available studies, and the small number of epidemiologic studies with higher quality exposure assessment, no firm conclusion can be drawn regarding any causal effect of chloroform on low birth weight.

Six studies observed no significant association between chloroform exposure and birth weight when it was considered as a continuous variable (Cao et al., 2016; Hoffman et al., 2008a; Kogevinas et al., 2016; Savitz et al., 2005; Villanueva et al., 2011; Wennborg et al., 2000), whereas seven studies found some significant associations of chloroform exposure with decreases in birth weight (Grazuleviciene et al., 2011; Rivera-Núñez & Wright, 2013; Smith et al., 2016; Summerhayes et al., 2012; Wright et al., 2004; Zhou et al., 2010). Most of these studies assumed a linear relationship between chloroform exposure and birth weight; assuming this to be the correct form of the relationship, a significant association implies an exposure-response trend.

Of the 13 studies, 6 had integrated exposure measures (Table 8). Of these, four studies showed no significant association (Cao et al., 2016; Kogevinas et al., 2016; Savitz et al., 2005; Villanueva et al., 2011) and two showed a significant inverse association between chloroform exposure and birth weight (Grazuleviciene et al., 2011; Smith et al., 2016). Grazuleviciene et al. (2011) found decrements of slightly more than 50 g in birth weight per 0.1 μg/d increase in chloroform exposure, whereas Smith et al. (2016) showed a decrement in birth weight of 2.65 g in association with an increase in chloroform exposure from tertile one (<0.91 μg/d) to tertile two (0.91–1.56 μg/d). Other studies with less-rigorous exposure assessment generally showed relatively small decrements (3.4–19 g) in birth weight when comparing the higher to lower categories of chloroform exposure; these decreases would likely be considered clinically insignificant. The larger decrement detected by Grazuleviciene et al. (2011) thus appears to be an outlier in terms of the magnitude of the association between chloroform exposure and birth weight. Further, compared with other integrated studies, Grazuleviciene et al. (2011) had a lower average chloroform exposure level (mean = 7.8 μg/L; SD = 10.2).

Considering the epidemiologic evidence on chloroform exposure and continuous birth weight, majority of studies with weaker exposure assessment methods suggest an inverse association, but the majority of studies with higher quality exposure assessments do not. In light of the reliance on poor-quality ecological exposure estimation in most of the available studies, as well as the paucity of studies with more rigorous integrated exposure assessment, no reliable conclusion can be reached on the potential causal relationship between chloroform and birth weight.

Infant growth: Postnatal weight gain

A single study by Botto et al. (2015) examined the relationship of chloroform exposure with differences in postnatal weight gain during the first six months of life. The study showed no significant association of chloroform exposure—whether estimated based only on ingestion or based on integrated exposure data—with postnatal weight gain. Thus, as the only available study, it does not provide evidence in favor of a causal relationship between chloroform exposure and reduced postnatal weight gain.

Growth: Small for gestational age

Small for gestational age is defined as an infant below the tenth percentile of weight for gestational age at birth, specific to the population into which the child is born. The United
States, Canada, and some other countries have population-specific designations for the tenth percentile cut-off. These designations can be further specified by the infant’s sex and race, and potentially by the parity of the mother.

More than half (17) of the epidemiologic studies examined the association between chloroform exposure and risk of small for gestational age, including either all infants collectively or only those born at term (Table 7). These studies represent 13 separate populations, with 43 associations estimated across different levels and periods of exposure. Thirteen of the 17 studies (37 associations) observed no clear increased or decreased risk of small for gestational age infants with increasing chloroform exposure (Cao et al., 2016; Costet et al., 2012; Danileviciute et al., 2012; Grazuleviciene et al., 2011; Hinckley et al., 2005; Hoffman et al., 2008a; Infante-Rivard, 2004; Kogevinas et al., 2016; Levallois et al., 2012; Levallois et al., 2016; Rivera-Núñez & Wright, 2013; Savitz et al., 2005; Villanueva et al., 2011). The other four studies (six associations) found a significant increase in risk of small for gestational age with increased exposure to chloroform during pregnancy (Kramer et al., 1992; Porter et al., 2005; Summerhayes et al., 2012; Wright et al., 2004). Wright et al. (2004) showed a possible positive exposure-response trend across three categories of exposure (<50th percentile, 50–90 percentile, >90th percentile). Mothers with chloroform exposure above the 90th percentile had 1.11 times the odds of having an infant born small for gestational age compared to mothers with chloroform exposure below the median value (95% CI = 1.04, 1.17) (Wright et al., 2004). The evidence observed in Kramer et al. (1992) (OR = 1.8 comparing highest to lowest categories of chloroform exposure; 95% CI = 1.1, 2.9) and Summerhayes et al. (2012) (RR = 1.03 for a 25 μg/L increase in chloroform water concentration; 95% CI = 1.02, 1.05) is also consistent with a positive exposure-response trend, whereas that from Porter et al. (2005) is not (OR = 1.24 comparing 2nd to 1st quintile; 95% CI = 1.02, 1.50). It should be noted that Kramer et al. (1992) used a slightly different small for gestational age definition (below the 5th instead of 10th percentile weight for gestational age). The use of this more extreme category for the outcome would result in a larger odds ratio than expected for the standard definition if the modeled relationship were a positive monotonically increasing trend between chloroform exposure and the risk of small for gestational age. Summerhayes et al. (2012) also reported a relatively large relative risk, but this was due in part to the fact that they used a relatively large incremental increase in chloroform exposure (25 μg/L) as the unit for analysis, compared with other studies that examined chloroform as a continuous variable (typically 10 μg/L).

Not a single study of small for gestational age that used an integrated approach for assessing chloroform exposure (of which there were nine) found a statistically significant association between chloroform exposure and risk of small for gestational age (Cao et al., 2016; Costet et al., 2012; Danileviciute et al., 2012; Grazuleviciene et al., 2011; Kogevinas et al., 2016; Levallois et al., 2012; Levallois et al., 2016; Savitz et al., 2005; Villanueva et al., 2011). In the absence of significant associations among studies with methodologically rigorous chloroform exposure assessment, combined with the inconsistency of observed associations across all available studies, the preponderance of the evidence does not establish a clear association between chloroform exposure during pregnancy and small for gestational age infants.

U.S.-based studies
For understanding the impact of disinfection by-products such as chloroform in the U.S. water supply, it is important to take transportability of study results into consideration by giving particular attention to studies conducted in the United States (Table 9). Of the 30 studies included in this systematic evaluation, 11 were conducted in the United States: 5 in Massachusetts, 1 in Arizona, 1 in Iowa, 1 in Maryland, 1 in Texas, and 2 in three U.S. communities of North Carolina, Tennessee, and Texas; none were conducted in California.

Two of the 11 U.S. studies examined congenital anomalies (Breder et al., 2014; Wright et al., 2017), and neither study integrated information on individual behaviors for more accurate exposure assessment estimates. Most models yielded no significant association with any type of birth defect, and no exposure-response relation was seen between chloroform exposure and the risk of spina bifida, which was significantly elevated only in the second versus first quartile of exposure.

Preterm birth and gestational age at birth were inconsistently associated with chloroform exposure in U.S. studies. Wright et al. (2004), Savitz et al. (2005), and Lewis et al. (2007) all observed significant inverse associations of gestational age with chloroform exposure, including one study (Savitz et al., 2005) that used integrated exposure assessment. The one significant positive association with preterm birth in a U.S. study showed no positive exposure-response trend (Rivera-Núñez & Wright 2013).

Seven U.S. studies considered small for gestational age as an outcome in association with chloroform exposure. Four studies found no significant association between chloroform exposure and small for gestational age births. Wright et al. (2004) found a significant positive association between median chloroform concentration in water and the odds of small for gestational age, with some evidence of a positive exposure-response gradient. Kramer et al. (1992) and Porter et al. (2005) also observed significant positive associations, but no clear exposure-response trends. Both of these studies were considered to be lower quality based on the use of ecologic exposure assessments.

Significant decreases in mean birth weight were observed in association with chloroform exposure in three
U.S. studies (Lewis et al., 2006; Rivera-Núñez & Wright, 2013; Wright et al., 2004). However, all of these studies had limited exposure assessment that lacked individual-level behavioral data.

Overall, U.S.-based studies indicate inconsistent associations of chloroform exposure at levels typically detected in various U.S. communities with the risks of adverse developmental outcomes. The one U.S.-based study with higher quality exposure assessment using individual-level behavioral data (Savitz et al., 2005) found a significant inverse (protective) effect of chloroform exposure during the third trimester on gestational age at birth. Thus, U.S. studies with high-quality exposure assessment are sparse, revealing a gap in the scientific database; nevertheless, the available U.S. studies show no consistent relationship between exposure to chloroform and adverse developmental outcomes in humans.

4 | CONCLUSIONS

The purpose of this analysis is to conduct a systemic assessment for persons of reproductive potential (population of interest) exposed to chloroform via either ingestion or inhalation (exposures of concern)—comparing those highly exposed to those either not exposed or minimally exposed (comparators)—for developmental and/or reproductive toxicity (outcomes). Thus, we conducted an assessment of the animal and epidemiologic data for chloroform to determine if the available weight of evidence showed that chloroform exposure causes developmental and/or reproductive toxicity.

An initial scoping exercise was undertaken, the results of which identified developmental toxicity as the primary area of concern for chloroform. Both the animal and epidemiology data for male and female reproductive endpoints were rather limited, but generally equivocal or negative for potential effects in association with exposure. The one exception is the finding of total litter loss, which is discussed in more detail below. This conclusion, that male and female reproductive toxicity are not areas of significant concern due to chloroform exposure, is consistent with that of OEHHA (2016c); therefore, the bulk of our analysis is focused on the potential for developmental toxicity in association with chloroform exposure.

In Schwetz et al. (1974), reported malformations related to caudal dysplasia at 100 ppm, but these findings were not confirmed in the other rat inhalation studies of chloroform at similar exposures (Baeder & Hofmann, 1988, 1991). A single animal inhalation study (Murray et al., 1979) reported an increased incidence of cleft palate in mice exposed during late gestation to 100 ppm chloroform in air. This finding was not replicated in the oral reproductive mouse study by Borzelleca and Carchman (1982), which included a teratogenic screen and used a higher oral dosage of chloroform than the inhalation dosage administered in Murray et al. (1979). Cleft palate also was not observed with chloroform exposure in rats. Further, the corresponding epidemiologic evidence does not demonstrate an association between chloroform exposure at environmentally relevant exposures and the occurrence of birth defects in humans (Brender et al., 2014; Dodds & King, 2001). A single epidemiologic study (Brender et al., 2014) suggested possible associations of chloroform exposure with septal heart defects and/or spina bifida/neural tube defects more broadly in humans. However, these findings did not demonstrate clear exposure-response trends and were not supported by the other epidemiologic studies evaluated, including those that used integrated exposure measures of chloroform.

The animal data showed consistent evidence of reduced fetal body weight (along with expected delays in fetal bone ossification) due to chloroform exposure. These findings were observed at the same doses at which indications of maternal toxicity (i.e., reduced food consumption, body weight, and/or body weight gains) were seen or at doses above which maternal toxicity was first observed. Fetal weight reductions and delays in ossification have been shown in other studies to be secondary to reduced maternal food consumption (Nitszche, 2017). EPA (2001) considered the fetal effects observed after chloroform exposure to be secondary to maternal toxicity. In contrast, OEHHA’s Developmental and Reproductive Toxicant Identification Committee (DARTIC) considered the fetal body weight reductions to be independent of maternal toxicity (OEHHA, 2016c).

In contrast to the relatively consistent nature of the animal data, the epidemiology data are generally equivocal on the issue of fetal growth. Multiple studies have evaluated parameters related to fetal growth, including low birth weight or weight at birth evaluated as a continuous variable. Approximately half of the epidemiology studies herein reviewed reported no association of chloroform exposure with birth weight. However, the other half of the studies suggested a possible inverse relationship between increasing chloroform exposure and birth weight; no association of chloroform exposure with increased birth weight was reported. The equivocal nature of the epidemiologic data on birth weight was also evident when the analysis was limited to those studies that employed integrated measures of chloroform exposure. Related epidemiologic data are available on the risk of small for gestational age. These data are less variable than those on birth weight: 13 of the 17 studies that examined this parameter showed no significant relationship between chloroform exposure and small for gestational age. The less-equivocal nature of these data may be due to the more specific definition of fetal growth. The remaining four studies found an increase in risk of small for gestational age with greater chloroform exposure, but the results were only marginally statistically significant, and some of these studies tested numerous hypotheses. Based on the equivocal results...
overall, including the lack of significant findings in the higher quality studies with integrated assessment of chloroform exposure, a causal relationship cannot be established between chloroform exposure and fetal growth.

The animal data also showed indications of very early (peri-implantation) total litter losses related to high-dose chloroform exposures. This finding was only observed at exposures well above those at which indications of maternal toxicity were seen. Further, because chloroform affects vascular function and blood flow in the uterus (which is important to the process of implantation) and the outcome affected the complete litter rather than individual fetuses, we believe that this effect was mediated on the maternal animal rather than directly on the offspring. In OEHHA’s decision to classify chloroform as a developmental toxicant (OEHHA, 2016c), the DARTIC discussed this finding of early litter losses in the animal studies. They considered that the finding could be indicative of either a female reproductive effect or mediated at the level of the developing offspring. For the purposes of classification, however, they appear to have categorized the finding as a developmental effect.

Finally, the epidemiologic data for chloroform suggest no or even an inverse association of exposure with preterm birth. These findings were also noted in OEHHA’s evaluation of the data (OEHHA, 2016c). The available animal data did not support the presence of an association of chloroform exposure with early birth; however, the available data to assess this endpoint were limited to only two reproductive toxicity studies (Borzelleca & Carchman, 1982; NTP, 1988). Further, the biological plausibility of a protective effect is not established, but could relate to cleaner drinking water sources with reduced bacterial or viral contamination.

In general, human epidemiology studies and animal toxicology studies have complementary strengths and limitations for assessing the reproductive toxicology of chloroform. The main strength of epidemiologic studies is the evaluation of human health outcomes in relation to real-world exposure levels and scenarios in a range of settings, making these studies directly relevant to public health. Epidemiologic studies of birth outcomes can take advantage of routinely collected objectively measured birth data that are available for nearly all newborns. Limitations include imprecise and probably inaccurate exposure assessment, especially in studies without integrated exposure assessment; confounding and other forms of bias, which can be present in all observational studies to a varying extent; and a paucity of information on reproductive health outcomes at the highest levels of human exposure such as those seen in occupational settings.

Strengths of animal studies include precise control over exposure doses and circumstances, and the potential for minimization of confounding and other forms of bias through randomization, allocation concealment, and blind outcome assessment. However, limitations include uncertainties about interspecies extrapolation of results to humans, possible evaluation of doses and exposure routes that are not relevant to most or all humans, and typically small sample sizes. For example, in some studies, the laboratory animals were exposed to air concentrations as low as 3 ppm while people are exposed to air concentrations that are 60,000–150,000-fold lower (0.02–0.5 ppb; ATSDR, 1997). In the oral dosing studies, the highest dose administered was 400 mg/kg/day. By comparison, water concentrations have been measured at 2–44 ppb (ATSDR, 1997). Assuming 2 L of water consumed per day and a default body weight of 70 kg, these concentrations translate to human chloroform exposures of <2 μg/kg/day (200,000-fold lower) due to water consumption.

In conclusion, based on our detailed analysis of the animal and epidemiology data associated with developmental chloroform exposure, we conclude that chloroform, if teratogenic, is only so at high maternally toxic doses. At levels sufficient to cause maternal toxicity chloroform may also cause decrements in fetal weights and associated delays in ossification. The available animal data related to this outcome are relatively consistent; the related epidemiologic data, in contrast, are generally equivocal, preventing any firm conclusions that such effects of chloroform exposure are seen in humans. The animal data also show evidence of early (peri-implantation) total litter losses, which we consider indicative of an effect that is mediated on the maternal animal rather than directly on the offspring. Based on this analysis, we find the available animal data demonstrate that chloroform perturbs development only at doses that typically cause maternal toxicity. Further, the available data in humans do not demonstrate a causal role for chloroform on the induction of adverse developmental effects at concentrations experienced generally in the environment.

ACKNOWLEDGMENTS

The authors wish to acknowledge the Foundation for Chemistry Research and Initiatives (FCRI) of the American Chemistry Council (ACC) for their financial support, which helped to fund in part the preparation of this article. The analysis and conclusions presented herein are those of the authors and not of their employer or the sponsor.

CONFLICT OF INTEREST

The authors are employees of Exponent, a scientific and engineering consulting firm. Exponent is a consulting firm that provides scientific analysis and advice in areas that include toxicology and risk assessment. Georgetown University School of Medicine is a provider of medical education.

Although the authors have consulted on issues (including litigation) related to developmental outcomes in the past, the authors have not previously addressed developmental outcomes related to chloroform exposures and thus have no
conflicts of interest to declare related to the analysis of chloroform and developmental outcomes, as presented herein. The authors are solely responsible for the analyses and preparation of this article; the opinions and conclusions are those of the authors and are not those of their employer or the sponsor.

**ORCID**

Amy Lavin Williams https://orcid.org/0000-0002-9499-7053

**REFERENCES**

Agency for Toxic Substances and Disease Registry (ATSDR). 1997. Toxicological profile for chloroform. U.S. Department of Health and Human Services. September 1997.

Baeder, C. & Hofmann. 1988. Inhalation embryotoxicity study of chloroform in Wistar rats. Hoescht Aktiengesellschaft and Dow Europe S.A. (Report No. 88-922000571). April 9, 1987.

Baeder, C. & Hofmann. 1991. Chloroform: Supplementary inhalation embryotoxicity study in Wistar rats. Hoescht Aktiengesellschaft and Dow Europe S.A. (Report No. 88-920000566). July 12, 1990.

Balster, R. L. & Borzelleca, J. F. (1982). Behavioral toxicity of trihalomethane contaminants of drinking water in mice. *Environmental Health Perspectives*, 46, 127–136.

Borzelleca JR, RA Carchman. 1982. Effects of selected organic drinking water contaminant on male reproduction (Report No EPA-6001-82-009). Health Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency.

Botton, J. M., Kogevin, M., Gracia-Lavedan, E., Roumeliotaki, T., Kharrazi, M., & Nieuwenhuijsen, M. J. (2015). Individual exposures to drinking water trihalomethanes, low birth weight and small for gestational age risk: a prospective Kaunas cohort study. *Environmental Health*, 10, 32.

Greenland, S., & O’Rourke, K. (2001). On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biosciences*, 4, 463–471.

Heywood, R., Sortwell, R. J., Noel, P. R. B., Streit, A. E., Prentice, D. E., Roe, F. J. C., ... Van Abbe, N. J. (1979). Safety evaluation of toothpaste containing chloroform III. Long-term study in Beagle dogs. *Journal of Environmental Pathology and Toxicology*, 2, 835–851.

Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, 58, 295–300.

Hincley, A. F., Bachand, A. M., & Reif, J. S. (2005). Late pregnancy exposures to disinfection by-products and growth-related birth outcomes. *Environmental Health Perspectives*, 113, 1808–1813.

Hoffman, C. S., Mendola, P., Savitz, D. A., Herring, A. H., Loomis, D., Hartmann, K. E., ... Olshan, A. F. (2008a). Drinking water disinfection by-product exposure and fetal growth. *Epidemiology*, 19, 729–737.

Howards, P. P., Hertz-Picciotto, I., Weinberg, C. R., & Poole, C. (2006). Misclassification of gestational age in the study of spontaneous abortion. *American Journal of Epidemiology*, 11, 1126–1136.

Infante-Rivard, C. (2004). Drinking water contaminants, gene polymorphisms, and fetal growth. *Environmental Health Perspectives*, 112, 1213–1216.

Iszatt, N., Nieuwenhuijsen, M. J., Bennett, J., Best, N., Povey, A. C., Pacey, A. A., ... Toledano, M. B. (2013). Chlorination by-products in tap water and semen quality in England and Wales. *Occupational and Environmental Medicine*, 70, 754–760.

Iszatt, N., Nieuwenhuijsen, M. J., Bennett, J. E., & Toledano, M. B. (2014). Trihalomethanes in public drinking water and stillbirth and low birth weight rates: An intervention study. *Environment International*, 73, 434–439.

Iszatt, N., Nieuwenhuijsen, M. J., Nelson, P., Elliott, P., & Toledano, M. B. (2011). Water consumption and use, trihalomethane exposure, and the risk of hypospadias. *Pediatrics*, 127, e389–e397.

Jorgenson TA, CJ Rushbrook. 1980. Effects of chloroform in the drinking water of rats and mice. *Ninety-day subacute toxicity study* (Report No EPA/600/1-80-030). SRI International. July 1980.

Kimmel, C. A., Garry, M. A., & DeSesso, J. M. (2014). Relationship between bent long bones, bent scapulae, and wavy ribs: Malformation or variations? *Birth Defects Research, Part B*, 101, 379–392.

King, W. D., Dodds, L., & Allen, A. C. (2000). Relation between stillbirth and specific chlorination by-products in public water supplies. *Environmental Health Perspectives*, 108, 883–886.

Klimisch, H.-I., Andreade, M., & Tillmann, U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regulatory Toxicology and Pharmacology*, 25, 1–5.
EXCLUDED EPIDEMIOLOGY STUDIES

Chapin, R., Gulati, D., Hope, E., Mounce, R., Russell, S., & Poonacha, K. B. (1997). Chloroform, CAS #67-66-3, Swiss CD-1 mice, at 0.0, 6.6, 15.9, and 41.2 mg/kg by gavage. Environmental Health Perspectives, 105(Supplement 1), 1–2.

García-Estrada J, A Navarro-Ruiz, J Bañuelos-Pineda, V Gómez, Zeng, Q., Li, M., Xie, S. H., Gu, L. J., Yue, J., Cao, W. C., Yang, P., Zeng, Q., Cao, W. C., Wang, Y. X., Huang, Z., Li, J., Chloroform. World Health Organization (WHO). 2004.

Zeng, Q., Chen, Y. Z., Xu, L., Chen, H. X., Luo, Y., Li, M., … Lu, W. Q. (2014). Evaluation of exposure to trihalomethanes in tap water and semen quality: A prospective study in Wuhan, China. Reproductive Toxicology, 46, 56–63.

Zeng, Q., Li, M., Xie, S. H., Gu, L. J., Yue, J., Cao, W. C., … Lu, W. Q. (2014). baseline blood trihalomethanes, semen parameters and serum total testosterone: a cross-sectional study in China. Environment International, 54, 134–140.

Zhou, W., Li, Y., & Xie, S. (2010). Relationship between exposure to disinfection by-products during pregnancy and term infants birth weight. Journal of Environment and Health, 27, 17–20.

EXCLUDED ANIMAL STUDIES

Aggazzotti, G., Ri gibi, E., Fantuzzi, G., Biasotti, B., Ravera, G., Kanitz, S., … Collaborative Group for the Study of Chlorinated Drinking Waters and Pregnancy. (2004). Chlorination by-products (CBPs) in drinking water and adverse pregnancy outcomes in Italy. Journal of Water and Health, 2, 233–247.

Ali, A., Kurzawa-Zegot, M., Najafzadeh, M., Gopalan, R. C., Plewa, M. J., & Anderson, D. (2014). Effect of drinking water disinfection by-products in human peripheral blood lymphocytes and sperm. Mutation Research, 770, 136–143.

Borzellina, J. F., RA Carchman. (1982). Effects of selected organic drinking water contaminants on male reproduction. Unpublished report provided to US EPA (PB82-259847):14.

Bove, F. J., Fulcomer, M. C., Klitz, J. B., Esmart, J., Duffey, E. M., & Savrin, J. E. (1995). Public drinking water contamination and birth outcomes. American Journal of Epidemiology, 141, 850–862.

Brown, V. J. (2007). Sperm quality and tap water: disinfection by-product effects not supported. Environmental Health Perspectives, 115, A416.

Chisholm, K., Cook, A., Bower, C., & Weinstein, P. (2008). Risk of birth defects in Australian communities with high levels of brominated disinfection by-products. Environmental Health Perspectives, 116, 1267–1273.

Dodd, L., King, W., Woolcott, C., & Pole, J. (1999). Trihalomethanes in public water supplies and adverse birth outcomes. Epidemiology, 10, 233–237.

Fenster, L., Waller, K., Windham, G., Hemmnen, T., Anderson, M., Mendola, P., … Swans, S. H. (2003). Trihalomethane levels in home tap water and semen quality. Epidemiology, 14, 650–658.

Gallagher, M. D., Nuckols, J. R., Stallones, L., & Savitz, D. A. (1998). Exposure to trihalomethanes and adverse pregnancy outcomes. Epidemiology, 9, 484–489.

Hassani, S., Namvar, M., Ghoreishvandi, M., Attarchi, M., Golabadi, M., Seyedi-mehdi, S. M., & Khodarahmian, M. (2014). Menstrual disturbances and hormonal changes in women workers exposed to a mixture of organic solvents in a pharmaceutical company. Medical Journal of the Islamic Republic of Iran, 28, 156.

Hoffman, S. N., Mendola, P., Savitz, D. A., Herring, A. H., Loomis, D., Hartmann, K. E., … Olshan, A. F. (2008b). Drinking water disinfection by-product exposure and duration of gestation. Epidemiology, 19, 738–746.

Horton, B. J., Lunen, T. J., Herring, A. H., Savitz, D. A., Singer, P. C., Weinberg, H. S., & Hartmann, K. E. (2011). The effect of water disinfection by-products on pregnancy outcomes in two Southeastern US communities. Journal of Occupational and Environmental Medicine, 53, 1172–1178.

Hwang, B. F., & Jaakkola, J. J. (2012). Risk of stillbirth in the relation to water disinfection by-products: A population-based case-control study in Taiwan. PLoS One, 7, e33949.

Hwang, B. F., Jaakkola, J. J., & Guo, H. R. (2008). Water disinfection by-products and the risk of specific birth defects: A population-based cross-sectional study in Taiwan. Environmental Health, 7, 23.

Ileka-Priouzeau, S., Campagna, C., Legay, C., Deonandan, R., Rodriguez, M. J., & Levallois, P. (2015). Women exposure during pregnancy to haloacetaldehydes and haloacetonitrides in drinking water and risk of small-for-gestational-age neonate. Environmental Research, 137, 338–348.

Joyce, S. J., Cook, A., Newhiham, J., Brenters, M., Ferguson, C., & Weinstein, P. (2008). Water disinfection by-products and pre-labor rupture of membranes. American Journal of Epidemiology, 168, 514–521.

Kanitz, S., Francos, Y., Patrone, V., Savitz, D. A., Singer, P. C., Weinberg, H. S., & Hartmann, K. E. (1996). Association between drinking water disinfection and somatic parameters at birth. Environmental Health Perspectives, 104, 516–520.

King, W. D., Dodds, L., Allen, A. C., Armson, B. A., Fell, D., & Nimrod, C. (2005). Haloacetic acids in drinking water and risk for stillbirth. Occupational and Environmental Medicine, 62, 124–127.

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

Kumar, S., Forand, S., Babcock, G., Richter, W., Hart, T., & Hwang, S. A. (2014). Total trihalomethanes in public drinking water supply and birth outcomes: A cross-sectional study. Maternal and Child Health Journal, 18, 996–1006.

Luben, T. J., Nuckols, J. R., Mosley, B. S., Hobbs, C., & Reif, J. S. (2008). Maternal exposure to water disinfection by-products during gestation and risk of hypospadias. Occupational and Environmental Medicine, 65, 420–429.

Luben, T. J., Olshansky, A. F., Herring, A. H., Jeffay, S., Strader, L., Buss, R. M., … Pereraut, S. D. (2007). The healthy men study: An evaluation of exposure to disinfection by-products in tap water and sperm quality. Environmental Health Perspectives, 115, 1169–1176.

MacLelohe, R. F., Savitz, D. A., Herring, A. H., Hartmann, K. E., Singer, P. C., & Weinberg, H. S. (2008). Drinking water disinfection by-products and time to pregnancy. Epidemiology, 19, 451–458.

Nieuwouhijsen, M. J., Smith, R., Golfinopoulos, S., Best, N., Bennett, J., Aggazzotti, G., … Kogevinas, M. (2009). Health impacts of long-term exposure to disinfection by-products in drinking water in Europe: HIWATE. Journal of Water and Health, 7, 185–207.

Nieuwouhijsen, M. J., Toledano, M. B., Bennett, J., Best, N., Hambl, P., de Hoogh, C., … Elliott, P. (2008). Chlorination disinfection by-products and
risk of congenital anomalies in England and Wales. *Environmental Health Perspectives*, 116, 216–222.

Patelaru, E., Kargaki, S., Stephanou, E. G., Nieuwenhuijsen, M., Sourtzi, P., Gracia, E., ... Kogevinas, M. (2011). Exposure to brominated trihalomethanes in drinking water and reproductive outcomes. *Occupational and Environmental Medicine*, 68, 438–445.

Pedersen, M., Mendez, M. A., Schoket, B., Godschalk, R. W., Espinosa, A., Landström, A., ... Kogevinas, M. (2015). Environmental, dietary, maternal, and fetal predictors of bulky DNA adducts in cord blood: a European mother-child study (NewGeneris). *Environmental Health Perspectives*, 123, 374–380.

Riederer, A. M., Dhingra, R., Blount, B. C., & Steenland, K. (2014). Predictors of blood trihalomethane concentrations in NHANES 1999-2006. *Environmental Health Perspectives*, 122, 695–702.

Righi, E., Bechtold, P., Tortorici, D., Lauriola, P., Calzolari, E., Astolfi, G., ... Aggazzotti, G. (2012). Trihalomethanes, chlorite, chlorate in drinking water and risk of congenital anomalies: A population-based case-control study in Northern Italy. *Environmental Research*, 116, 66–73.

Righi, E., Fantuzzi, G., Montanari, M., Bargellini, A., Predieri, G., & Aggazzotti, G. (2003). Exposure to water disinfection by-products and adverse pregnancy outcomes: results of a case-control study carried out in Modena (Italy). *Annali di Igiene: Medicina Preventiva e di Comunità*, 15, 649–662.

Savitz, D. A., Andrews, K. W., & Pastore, L. M. (1995). Drinking water and pregnancy outcome in central North Carolina: Source, amount, and trihalomethane levels. *Environmental Health Perspectives*, 103, 592–596.

Shaw, G. M., Ranatunga, D., Quach, T., Neri, E., Correa, A., & Neutra, R. R. (2003). Trihalomethane exposures from municipal water supplies and selected congenital malformations. *Epidemiology*, 14, 191–199.

Villanueva, C. M., Cordier, S., Font-Ribera, L., Salas, L. A., & Levallois, P. (2015). Overview of disinfection by-products and associated health effects. *Current Environmental Health Reports*, 2, 107–115.

Waller, K., Swan, S. H., Windham, G. C., & Fenster, L. (2001). Influence of exposure assessment methods on risk estimates in an epidemiologic study of total trihalomethane exposure and spontaneous abortion. *Journal of Exposure Analysis and Environmental Epidemiology*, 11, 522–531.

Windham G, L Fenster. 2008. Environmental contaminants and pregnancy outcomes. *Fertility and Sterility* 89(2 Suppl):e111–116; discussion e117.

Wright, J. M., Schwartz, J., & Dockery, D. W. (2003). Effect of trihalomethane exposure on fetal development. *Occupational and Environmental Medicine*, 60, 173–180.

Xie, S. H., Li, Y. F., Tan, Y. F., Zheng, D., Liu, A. L., Xie, H., & Lu, W. Q. (2011). Urinary trichloroacetic acid levels and semen quality: A hospital-based cross-sectional study in Wuhan, China. *Environmental Research*, 111, 295–300.

Yang, C. Y., Xiao, Z. P., Ho, S. C., Wu, T. N., & Tsai, S. S. (2007). Association between trihalomethane concentrations in drinking water and adverse pregnancy outcome in Taiwan. *Environmental Research*, 104, 390–395.

Yang, P., Zhou, B., Cao, W. C., Wang, Y. X., Huang, Z., Li, J., ... Zeng, Q. (2017). Prenatal exposure to drinking water disinfection by-products and DNA methylation in cord blood. *Science of the Total Environment*, 586, 313–318.

Zeng, Q., Wang, Y. X., Xie, S. H., Xu, L., Chen, Y. Z., Li, M., ... Lu, W. Q. (2014). Drinking-water disinfection by-products and semen quality: A cross-sectional study in China. *Environmental Health Perspectives*, 122, 741–746.

Zeng, Q., Zhou, B., He, D. L., Wang, Y. X., Wang, M., Yang, P., ... Lu, W. Q. (2016). Joint effects of trihalomethanes and trichloroacetic acid on semen quality: A population-based cross-sectional study in China. *Environmental Pollution*, 212, 544–549.

Zhou, W. S., Xu, L., Xie, S. H., Li, Y. L., Li, L., Zeng, Q., ... Lu, W. Q. (2012). Decreased birth weight in relation to maternal urinary trichloroacetic acid levels. *Science of the Total Environment*, 416, 105–110.

ENDNOTES

1Four articles reporting on pregnancy loss/stillbirth/spontaneous abortion also report developmental toxicity endpoints and are included in both study totals (Wennborg et al., 2000; Savitz et al., 2005; Toledano et al., 2005; Iuzzotti et al., 2014).

2When empty uteri were stained for implantation sites, total pregnancies in the control and 100 ppm groups treated on GD 8–15 were 65% and 60%, respectively.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Williams AL, Bates CA, Pace ND, Leonhard MJ, Chang ET, DeSesso JM. Impact of chloroform exposures on reproductive and developmental outcomes: A systematic review of the scientific literature. *Birth Defects Research*. 2018;110:1267–1313. https://doi.org/10.1002/bdr2.1382