9-1-2021

# Prediction of the Formation, Speciation, and Health Risks of Unregulated Disinfection Byproducts in Drinking Water using a Kinetic Binomial Model

Xian Ma
*University of Massachusetts Amherst*

# Prediction of the Formation, Speciation, and Health Risks of Unregulated Disinfection Byproducts in Drinking Water using a Kinetic Binomial Model

A Dissertation Presented

by

**XIAN MA**

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirement for the degree of

DOCTOR OF PHILOSOPHY

September 2021

Department of Civil and Environmental Engineering

**Prediction of the Formation, Speciation, and Health Risks of Unregulated Disinfection Byproducts in Drinking Water using a Kinetic Binomial Model**

A Dissertation Presented

by

**XIAN MA**

Approved as to style and content by:

_____

David A. Reckhow, Chair

_____

John E. Tobiason, Member

_____

Youssef Oulhote, Member

_____

J. Michael Wright, Member

_____

John E. Tobiason, Department Head

Civil and Environmental Engineering Department

I have gone away to come back.

For the ones I left behind.

For the ones cannot out.

我离开是为了回来

为了那些我留在身后的

为了那些从未离开的

## ACKNOWLEDGEMENTS

First, I would like to show my deepest gratitude to my advisor and mentor, Dr. David A. Reckhow, for the years of patience, guidance, enlightenment, and generosity throughout my research life. His trust and confidence in me allowed me to embrace the challenges and to enjoy this long but never been dull journey.

Appreciation also extend to the members of my committee: Drs. John E. Tobiason, Youssef Oulhote, and J. Michael Wright. It was an honor to work with them and to have them on my committee. The high standards in research, the professionalism in work, the selflessness in guidance, and the optimistic attitude in life that they taught me would always be treasures of my life.

I have been more than lucky to know and to work with outstanding people. I am grateful. Thanks to Sherrie Webb-Yagodzinski and Patrick Wittbold for all the help in the laboratory. Thanks to Yun Yu, Yanjun Jiang, Ran Zhao, Rassil El Sayess, Joseph Goodwill, and Arianne Bazillio for guiding me when I was young and naïve. Thanks to Aarthi Mohan, Soon-mi Kim, Julie Mullen, Yue Sun, Junqi Jia, and Xuyen Mai for the help and the tolerance. Thanks to all other colleagues and friends for having me a good time – it is simply not possible to fully list all of them. Special thanks to my girlfriend, Simin Zhai, for all the unconditional trust and support.

Finally, thanks to my parents, Wanhua Cao and Xinqiang Ma, for everything. I am indebted, forever.

## ABSTRACT

PREDICTION OF THE FORMATION, SPECIATION, AND HEALTH RISKS OF UNREGULATED DISINFECTION BYPRODUCTS IN DRINKING WATER USING A KINETIC BINOMIAL MODEL

SEPTEMBER 2021

**XIAN MA**

B.S., TIANJIN UNIVERSITY

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Dr. David A. Reckhow

Among the regulated disinfection byproducts (DBPs) with the current Stage 2 Disinfectant and Disinfection Byproduct Rule (the Stage 2 D/DBPR), trihalomethanes (THMs) and haloacetic acids (HAAs) are two DBP classes that have been widely studied. Currently, the summation of concentrations of all four THMs ($CHCl_3$, $CHCl_2Br$, $CHClBr_2$, and $CHBr_3$) were regulated as total trihalomethanes (THM4), but without control of individual species. The summation of concentrations of five out of nine HAAs (monochloroacetic acid, monobromoacetic acid, dichloroacetic acid, dibromoacetic acid, and trichloroacetic acid) were regulated as HAA5. The other four HAA species were known as unregulated HAAs. Recent studies have directed attention to the unregulated HAA species due to their higher potential for carcinogenicity. At present, it is hard to

evaluate the health risks from unregulated DBP species as they are mostly not measured or monitored.

The core of this research is a kinetic binomial model that predicts the unregulated HAA species. We used a simple precursor, acetone, to test the reaction rate constants of natural organic matter halogenation reactions and to build the chemistry foundation of the model. By using the reaction rate constants derived from acetone halogenation, we predicted the formation of all for THM species with high accuracy.

The kinetic model was then developed and resulted showed very high accuracy between predicted and measured unregulated HAAs ($R^2$>0.98). We also provided a set of equations for the application of the model. The model was then applied to utility data (i.e., DBP data collected by public water suppliers) to verify the reliability in more complicate conditions and received results with high accuracy ($R^2$>0.95).

Lastly, the model was used to predict the concentration of unregulated HAAs in the state of Ohio. The predictions were connected with collected birth records to evaluate the association between the DBP exposure and the adverse birth outcomes. Chlorinated DBPs showed higher exposure risks. We believe that the result reflected the limitation from the currently used DBP surrogates. We also compared the state-of-art DBP surrogates in the field of epidemiology and suggested the usage of a more toxicity based approach, instead of simply summation of concentrations approach.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

Figure Page

# CHAPTER 1
# OVERVIEW AND INTRODUCTION

Among the regulated disinfection byproducts (DBPs) with the current Stage 2 Disinfectant and Disinfection Byproduct Rule (the Stage 2 D/DBPR), trihalomethanes (THMs) and haloacetic acids (HAAs) are two DBP classes that have been widely studied. Under the current regulation in the US, the summation of the concentration of all four chlorinated and brominated THMs ($CHCl_3$, $CHCl_2Br$, $CHClBr_2$, and $CHBr_3$) were regulated as total trihalomethanes (THM4), but without detail of individual species. The summation of the concentration of five out of nine HAAs (monochloroacetic acid (MCAA), monobromoacetic acid (MBAA), dichloroacetic acid (DCAA), dibromoacetic acid (DBAA), and trichloroacetic acid (TCAA)) were regulated as HAA5. However, no individual species concentrations were regulated in the US. Recent studies have revealed that the brominated HAA species have a higher potential for carcinogenicity, such as bladder cancer and colon cancer. At present, it is hard to evaluate the health risks from unregulated DBP species as they are mostly not measured or monitored.

The goal of this research is to establish a kinetic-based binomial model to predict the unregulated DBP species from regulated species and to apply this model to utility data (i.e., DBP data collected by public water suppliers (PWSs)) to evaluate the exposure risks to human being. Before the establishment of the kinetic model, we first examined the formation and speciation of THMs from the halogenation of natural organic matter (NOM) and one simple target compound, acetone. Then we started with bench-scale experiments on representative surface water samples that cover typical water sources of

PWSs to create and to validate the model. The model was verified by NOM containing natural surface water. To further test the application of the kinetic model, we then expanded the scope to real distribution system samples and examined the ability of the model in predicting DBP formation in more close-to-reality scenarios. Lastly, we applied the calibrated model to utility data and applies these predictions to public health data (i.e., birth outcomes) to study the impact of unregulated HAA exposure on human being.

The first part of this dissertation (**CHAPTER 2**) mainly studied the formation kinetic of THM species from the target compounds. The study started with the halogenation tests on one simple precursor, acetone. From the study of acetone, we calculated the key rate constants in the formation process of THM and discussed the discrepancies between our work and previous research. Then we evaluated the competition between chlorination and bromination in the same bench-scale experiments. The competition between such halogenations under a more close-to-reality experiment setting, such as our work, was noticed to result in different conclusion than those in a more theoretical scenario. We also detailed the reason for such difference. This part of the dissertation provided the basic understanding and knowledge for the design of the binominal predictive model in the later chapters.

The second part of this dissertation (**CHAPTER 3**) examined the associations between THMs and HAAs, in terms of both formation and speciation. This study provided a relative simple binominal predictive model for the prediction of unregulated HAAs by using regulated HAA and THM measurements. The validation and verification of the binominal predictive model used the data collected from bench-scale experiments. Surface water samples from six public water sources in Massachusetts were collected in

this part of the research. Sodium bromide (NaBr) was spiked into each water sample to simulate scenarios with different bromide concentrations. Standard bench-scale chlorination experiments were then conducted on each water sample and THM and HAA data was collected. Some systematic error was noticed in this process and need additional calibration. After a certain level of modification and calibration, the model predictions were nearly identical to the measurements.

The third part of this dissertation (**CHAPTER 4**) focused on the further application of the kinetic predictive model to the actual DBP data from distribution systems. To compliment the laboratory experiments, we used collected facility data from an external database (WITAF database) for the purpose of verification of our binomial model. The main premise of this chapter was to test the accuracy and availability of our model when applied to DBP data with a higher level of measurement error and variability, such as DBP data collected from distribution systems. Also, we tried to provide some evidence to promote the use of our model for researchers in this field in the future. The results showed that the predicted DBP concentrations were quite similar to the actual measured DBP concentrations. Even though some random errors were detected between the predictions and the measurements, we found that our kinetic model succeed in predicting unregulated HAA levels in real distribution systems.

The fourth part of this dissertation (**CHAPTER 5**) redirected the focus from environmental chemistry to the application of the model to the epidemiological study. Most of the effort in this part focused on examining the relationship between DBP exposure and potential human health risks (e.g., birth outcomes). DBP records in Ohio from year 2005-2013 were provided from Ohio EPA as part of routine monitoring data

collected by public water suppliers (PWSs). After extensive data cleaning (e.g., extreme value verification, regulatory violation verification, distribution system spatial averaging, temporal averaging across quarterly samples, etc.), the binomial model was applied to collected DBP records to predict the unregulated HAA exposure. DBP records were then linked to birth records from the state of Ohio to evaluate the associations with birth weights and gestational age. It should be noted that due to the high sensitivity of the birth records and for the safety concerns, the information that could potential lead to the identification of the infants or the mothers was concealed in this dissertation. The result show unregulated HAA exposure has a negative impact on the birth weight of the infants.

# CHAPTER 2

# COMPETITION BETWEEN BROMINE AND CHLORINE IN THE HALOFORM REACTION AND IMPLICATIONS FOR NATURAL ORGANIC MATTER

## 2.1 Introduction

Disinfection of potable water with free chlorine is widely practiced in the US and across the world. Use of chlorine is a convenient and inexpensive way of protecting consumers from microbial pathogens. However, for nearly 50 years, water scientists and engineers have been aware that chlorine also produces toxic organic disinfection byproducts (DBPs) in treated drinking waters. The four chlorine and bromine-containing trihalomethanes (THMs) are the first of the DBPs to be detected and regulated. The DBPs are of concern due to toxicological and epidemiological evidence pointing to carcinogenicity, especially in the bladder and colon, and adverse birth outcomes (Muellner et al., 2007; Rahman et al., 2010; Regli et al, 2015; Villanueva et al., 2017; Wright et al, 2017; Atwood et al., 2019; Evlampidou et al., 2020).

During chlorination, ambient bromide ($Br^-$) is quickly oxidized to active bromine (mostly hypobromous acid and hypobromate, HOBr/$OBr^-$) which reacts with natural organic matter (NOM) forming DBPs containing covalently bonded bromine atoms (i.e., Br-DBPs). The Br-DBPs have generated special concern due to their perceived higher toxicity than the fully chlorinated DBPs (Plewa et al., 2004). It has long been known that chlorination DBP species shift to more brominated forms as the raw water bromide level increases (Rook, 1974; Arguello et al., 1979). Many studies have shown that this results in a decrease in concentration of the most highly chlorinated species and successively elevates concentrations of the more brominated species. Simple binomial models based

on bulk bromine incorporation have been used to describe these bromination shifts (Nokes et al., 1999).

Nevertheless, there remain some key gaps in our understanding of bromine incorporation along with some apparent contradictions. Despite nearly 50 years of research in this area, few quantitative models have been proposed that can accurately connect bromide levels and treatment conditions to the degree of bromine incorporation. Second, there remains uncertainties around the binomial model, especially with regard to bromine versus chlorine preference in the individual steps. Finally, data from tests with NOM and aromatic model compounds (e.g., phenol, resorcinol, etc.) seems to conflict regarding relative bromine-to-chlorine reactivity. The NOM-based studies indicate reactivity ratios between bromination and chlorination lie between 4 and 25 (Nokes et al., 1999; Westerhoff et al., 2004), whereas the model compound kinetic studies suggests this ratio really should be closer to 3000 (Heeb et al., 2014).

Many of these gaps can be attributed to the complex nature of NOM, the principal component of organic DBP precursors in raw drinking waters. NOM is composed of numerous organic molecules, each with its own reactivity with chlorine. Furthermore, each of these molecules may produce a more-or-less unique cascade of intermediates leading to regulated DBPs, such as the THMs. One helpful tool in understanding DBP formation has been laboratory testing with simple model compounds (De Laat et al., 1982; Reckhow & Singer, 1985; Deborde & von Gunten, 2008). Yet, these also have limitations. Until the present, none of these studies has used an unhalogenated precursor and followed the kinetics of each successive step through key intermediates to the full set of four trihalogenated products such as the THMs.

Our study used acetone as the model compound, because of the simple structure, the history of being studied, and the well-known haloform reaction featured in many organic chemistry textbooks. Despite its presumed minor role in THM formation, acetone is nevertheless a useful model for laboratory study because of its relative simplicity and the commercial availability of intermediates. Morris and Baum (1978) summarized the haloform mechanism as it relates to drinking water conditions (Figure 1). The key features are an initial slow ionization, forming an enolate anion (the rate limiting step), followed by rapid addition of an active halogen, resulting in monohaloacetone (MHAce). Next, the MHAce undergoes an analogous slow ionization and rapid halogenation to form the dihaloacetone (DHAce). Finally, the process repeats itself forming a trihaloacetone (THAce) which can under hydrolysis to form a THM. Because of the electron withdrawing nature of chlorine and bromine atoms, the rates of ionization are expected to increase with each stepwise addition of a halogen (e.g. Bell & Lidwell, 1940; Bell et al., 1949).

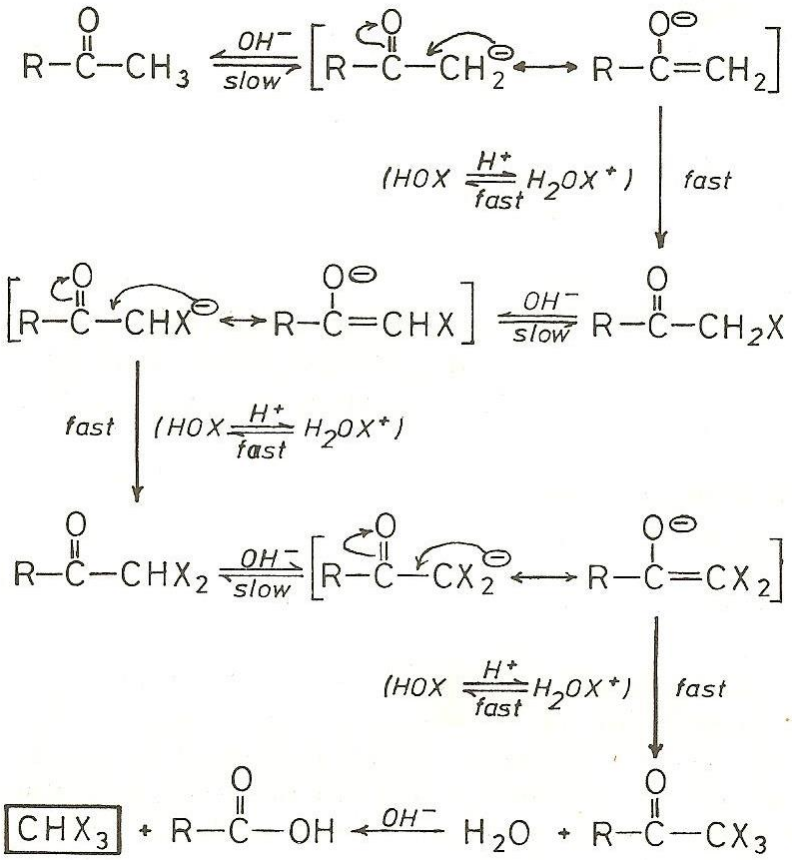



**Figure 1. Reaction Pathway of the Haloform Reaction as Depicted by Morris & Baum, 1978.**

Of course, as halogen concentrations diminish, the second step will eventually become limiting, but the literature indicates that this will not be the case at chlorine residual levels and pHs typically used in water treatment (De Laat, 1981). Under these conditions, the overall reaction is expected to be independent of halogen concentration. At high pH, the ionization occurs much faster and the dependence on halogen residual should start to become evident even at moderate chlorine residuals.

Guthrie and co-workers (Guthrie et al., 1984; Guthrie & Cossar, 1986) established a detailed kinetic model for the chlorination of acetone. Most of their work was conducted at elevated pH and halogen concentration. While not all is applicable to dilute chlorination at neutral pH, the bulk of their work provides valuable insight. Likewise, they based their kinetics on the enolization rate from Bell & Longuet-Higgins (1946) which was also determined under alkaline conditions (pH 11.8-12.6) and with elevated hypochlorite (5mM). Using the rate constants from Guthrie & Cossar (1986), first order rate constants between chlorine and acetone can be calculated as a function of pH (**Error! Reference source not found.**). Here, the compression of chlorine residual lines at pH 7 is evident, supporting the expected insensitivity to residual concentration.

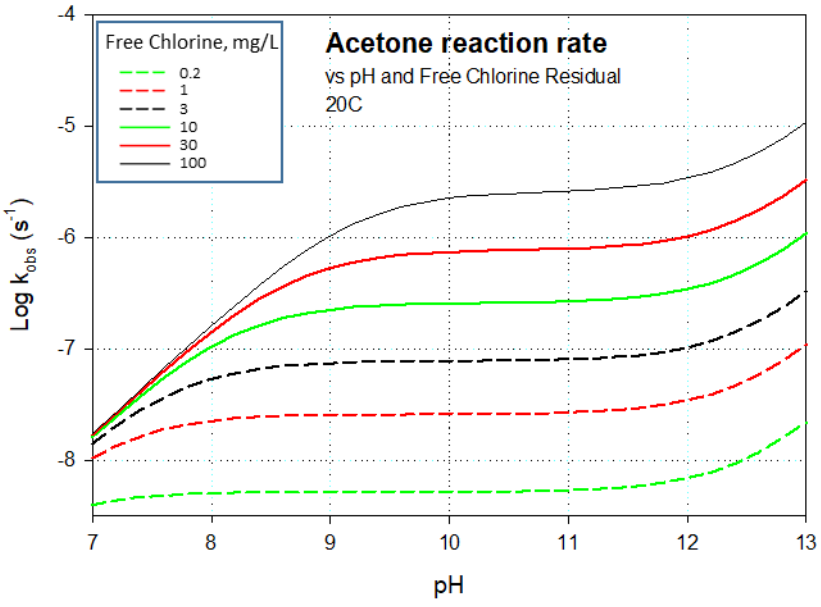



**Figure 2. Predicted First Order Rate Constants (kobs) For Chlorination of Acetone Based on Kinetics in Guthrie & Cossar (1986).**

While acetone is commonly used as an example for the haloform reaction, it was recognized early on that acetone itself cannot be an important THM precursor in raw drinking waters due to its low natural abundance and its slow reaction rate. However, many important THM precursors with functional groups that support faster carbanion formation, are likely to "react as ketones" (Morris & Baum, 1978). As the "R" group becomes more electron withdrawing, the methyl group is more acidic and the reaction proceeds more rapidly.

Acetone and resorcinol are often viewed as end members of the spectrum of THM precursors. Resorcinol reacts very quickly, mirroring the rapid THM formation seen immediately upon chlorine addition. Acetone reacts slowly over long periods of time much like the slow formation observed across distribution systems. Resorcinol and related activated phenolic precursors react initially via electrophilic aromatic substitution, mostly through the phenate form. Electron withdrawing substituents like the halogens themselves slow the reaction. However, they also result in greater tendency to form the phenate ion, which tends to compensate for the species-specific reduced reactivity. With acetone and similar aliphatic precursors, the rate of reaction is quite closely linked to extent or rate of ionization, and the resulting carbanion is also the site of attack. This means that the electron withdrawing properties of halogens cause much faster reaction and addition of subsequent halogens. For this reason, the acetone type precursors are thought to result in more gem polyhalide products such as the regulated DBPs.

Thus, the study of acetone presents a good opportunity to understand how one end-member precursor reacts competitively with chlorine in the presence of bromide to create

409 a mixture of THMs, and to better define the relationship between this THM blend and the
410 starting bromide concentration.

411

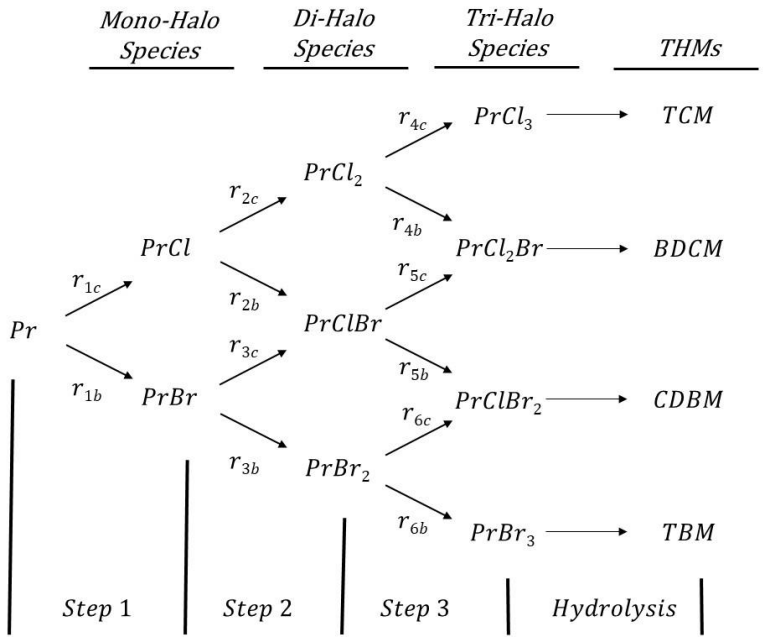
412 **2.2 Material and Methods**

413 **2.2.1 Experimental Methods**

414 Five DBP precursors were selected in this research: acetone, 1-chloriacetone
415 (MCAce), 1,1-dichloroacetone (DCAce), 1-bromoacetone (MBAce), and 1,1-
416 dibromoacetone (DBAce). Figure 3 illustrated the basic concept of the experiment. It
417 should be noted that 1-bromo-1-chloroacetone (BCAce) was currently not commercially
418 available and thus was not included in this research.

419 The reaction conditions were selected so that substantial THMs would be formed, yet
420 the reaction would not be complete. Reaction times (72 hours, 8 hours, 15 minutes, and
421 10 minutes; for acetone, MBA, MCA and both DCA and DBA, respectively) were
422 selected based on preliminary testing of the speed of the reaction under our experimental
423 conditions (e.g., close to neutral pH, dark storage settings, and 25℃).

424




425

426 **Figure 3. Conceptual Diagram for Simple Competitive Halogenation Pathway,**
427 **Resulting in the Binomial Model**

428

429 As these reactions are well known to be susceptible to both general and specific base
430 catalysis, we elected to conduct the reactions without added buffer ions, and to monitor
431 pH instead. Each precursor was dosed separately into 300-mL of deionized water in clean
432 BOD bottles to achieve a concentration of 200 µg/L. Then a concentrated sodium
433 bromide solution was spiked into each bottle to achieve the desired concentrations of
434 bromide (i.e., 0%, 25%, 50%, 75%, 100%, 150%, 200%, 250%, or 300% of the molar
435 concentration of the precursor). Finally, the sodium hypochlorite solution was spiked into
436 each bottle to achieve a dose of about 30 µM (25.9 to 33.2 µM or 1.84 to 2.36 mg/L as
437 $Cl_2$, Table 1). Once chlorine was added, the samples were moved to a 25°C incubator for
438 the desired reaction time. Reaction times were tailored to each precursor so that a
439 substantial amount of the precursor (20-95%) remained after quenching. Both pH and

chlorine residual were measured at the end of incubation. At the end of the reaction, 20 mL of sample was transferred into a 45-mL amber bottle for THM analysis, and immediately quenched with sodium arsenite. A sodium hydroxide solution was then added to bring the pH up to 13.0. This was done to hydrolyze all trihaloacetones (THAce) species to their corresponding THMs (Figure 3). Samples were kept at 25°C for 1 hour to ensure the completion of hydrolysis reactions. Complete hydrolysis and quantitative conversion to the THMs was assured in two ways. First a series of tests were run with TCAce at pH 11 and 13. After one hour of incubation at 25°C, the concentration of the hydrolysis product, TCM, was measured to calculate the recovery. Second, several samples within each set were run by GC prior to hydrolysis. This allows us to document the appearance of the THAce peaks and their subsequent disappearance following hydrolysis.

**Table 1. Summary of TCAce recovery in three scenarios[1].**

| pH | TCAce Concentration (μM) | Measured TCM Concentration (μM) | Recovery (%) |
|---|---|---|---|
| 7.8 | 0.31 | 0.17 | 54.0 |
| | 0.62 | 0.37 | 59.0 |
| | 1.24 | 0.77 | 62.0 |
| 11.0 | 0.31 | 0.31 | 99.8 |
| | 0.62 | 0.60 | 96.5 |
| | 1.24 | 1.21 | 98.0 |
| 13.0 | 0.31 | 0.30 | 98.0 |
| | 0.62 | 0.60 | 97.0 |
| | 1.24 | 1.20 | 96.9 |

[1] Note that without any additional base, the pH in TCAce solution (0.31, 0.62, and 1.24 μM) was measured at 7.7-7.9.

In these experiments, bromide was present prior to addition of chlorine much like full-scale potable water treatment systems. Thus, the reaction between chlorine and bromide that produces active bromine, HOBr and $OBr^-$, occurred simultaneously with reaction between chlorine and some of the fast organic precursors. Note that small amounts of BrCl, $Br_2O$ and BrOCl, will also be produced, but these all deemed inconsequential under these conditions (Heeb et al., 2014). These reactions are all fast compared to the halogenation reactions with the acetone and haloacetone precursors. Second order rate constants of 1500-6800 $M^{-1}s^{-1}$ have been proposed for the reaction of chlorine with bromide (Heeb et al., 2014), and this results in a bromide to bromine oxidation half-life of about 30 seconds, which is between 0.01% and 5% of the reaction times used.

### 2.2.2. Chemicals and Reagents

Sodium bromide (NaBr, > 99.0%), sodium hypochlorite (NaClO, 5.65-6.00%), reagent-grade pentane, and other chemicals were purchased from either Sigma-Aldrich (St. Louis, MO, US) or Fisher-Scientific (Fair Lawn, NJ, US). Acetone, MCAce, and TCAce were also purchased from Sigma-Aldrich. DCAce were purchased from AccuStandard (New Haven, CT, US). DBAce was purchased from Cansys Chemical (Toronto, ON, CA). MBAce (90%) was purchased from Win-win Chemicals (Shanghai, China). All aqueous solutions were prepared in acid-cleaned borosilicate glassware with ultrapure water generated from a Milli-Q system (Millipore, Billerica, MA, US). The calibration standards of THMs were prepared by diluting certified standard mixtures (Commercial Mix 551A and 551B, Absolute Standards, Hamden, CT, US) into acetone and were kept below 0 °C until use.

### 2.2.3 Analytical Methods

Chlorine residuals were measured by the DPD ferrous titrimetric method (APHA, 1989). pH was measured with the glass membrane electrode and laboratory meter (Accument AP85, Fisher-Scientific). Four THMs were analyzed by liquid/liquid extraction with pentane followed by gas chromatography and electron capture detection (Model 6890, Agilent) equipped with a DB-5 capillary column in accordance with US EPA Method 551.1. THM samples were extracted immediately after the completion of hydrolysis. After extraction, samples were kept below 0°C and analyzed within 2 days.

### 2.2.4 Simple Binomial Kinetic Model

In this and in prior work we have interpreted DBP formation in the context of a kinetic model featuring competitive reactions of halogens (active chlorine and bromine) for reactive sites on organic precursors, resulting in formation of new stable carbon-halogen bonds. This presumes that there are sufficient reactants (organic precursors, active chlorine and active bromine) so that the reactions continue for the timeframe of interest. In the case of trihalogenated DBPs such as THMs, the model could be represented as in Figure 3 below. Under this scheme the unhalogenated precursor (i.e. ‘Pr’) undergoes repeated halogen addition on one of its carbon atoms (the “reactive carbon”). Other portions of the molecule may undergo oxidation or halogenation, but that is not viewed as important to the halogen-specific competition for the reactive carbon.

For example, the rate of formation of $PrCl_3$ from $PrCl_2$ is represented as $r_{4c}$, which can be expressed as an apparent second-order kinetic law:

$$r_{4c} \equiv \frac{d[PrCl_2 \rightarrow PrCl_3]}{dt} = k_{4c}[Cl^+][PrCl_2] \qquad \textit{(Eqn. 1)}$$

We use the generic $[Cl^+]$ as the form of active chlorine might be either [HOCl], $[OCl^-]$, $[Cl_2]$ or $[Cl_2O]$ or some combination. Of course, we also recognize that in some cases, the intermediate precursor $[PrCl_2]$ may be in equilibrium with a more reactive form, as is the case for the haloform reaction. In this instance, the kinetics of conversion of $[PrCl_2]$ to its reactive form are incorporated into the rate constant ($k_{4c}$), which may also become inversely proportional to active halogen atoms so that the rate ($r_{4c}$) approaches zero order in active halogen. This is the case for diffusion-controlled halogenation of some enolate ions. Conceptually, this is identical to the approach used by Nokes and colleagues (1999) and later adopted by others (Cowman & Singer, 1996).

## 2.3 Result and Discussion

### 2.3.1 Extent of Reaction

As intended, all reactants (chlorine, inorganic bromine and organic precursors) were present throughout the reaction periods. Chlorine demands ranged from 0.08 to 0.54 mg/L, leaving 75-96% of the dosed chlorine as measureable residual. Except for DCAce, the remaining precursor was calculated to be in the range of 50% to 90%. In the case of DCAce, it was about 22%. The remaining presence of the haloacetones was confirmed through analysis of pre-hydrolysis samples. This also confirmed that final hydrolysis

resulted in loss of peaks attributed to trihaloacetones, and that no substantial unknown peaks remained. Inorganic bromine, when added, was only partly incorporated such that its concentration never decreased to below 75% of the dose. Post-reaction pH averaged 7.88 and ranged from a low of 7.57 to a high of 8.19.

### 2.3.2 Hydrolysis of Trihaloacetones

The hydrolysis of TCAce is well known to result in nearly 100% yield of TCM (Reckhow and Singer, 1985). While TCAce hydrolysis occurs under normal drinking water conditions, it is slow and inconvenient when the objective is quantitative conversion during short-term laboratory experiments. Accordingly, we tested and verified that conversion of TCAce to TCM could be done quantitatively using elevated pHs. Use of pH 11 or 13 for a period of 1 hour was found to be sufficient for complete loss of TCAce and 98% recovery of TCM (Table 1).

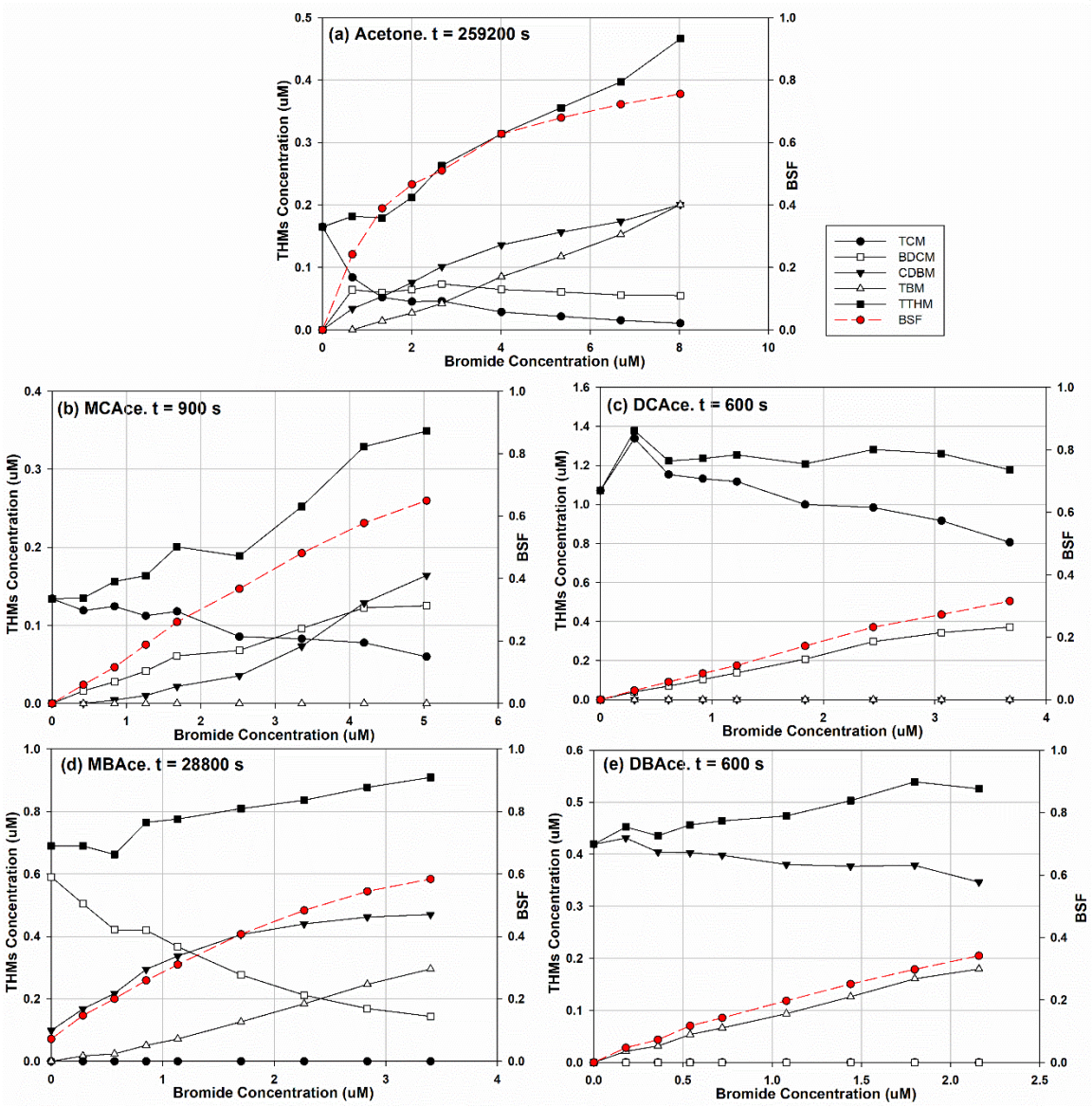
As the other three brominated trihaloacetones (bromodichloroacetone (BDCAce), chlorodibromoacetone (CDBAce), and tribromoacetone (TBAce)) were not commercially available at the time of this study, we could not repeat the same hydrolysis tests on these precursors. However, one would expect the brominated forms to be even more susceptible to alkaline hydrolysis than TCAce (Jeffers & Wolfe, 1997). In addition, we noted previously that the peaks thought to be these trihaloacetone compounds were entirely lost from the chromatograms upon hydrolysis.

### 2.3.3 Formation and Speciation of Trihalomethanes

A full set of halogenation experiments on acetone and the halogenated acetone precursors was conducted with different designated reaction time, as explained previously. As a part of quality assurance process, mass balance tests on bromide, chloride, and other precursors were conducted to confirm the reliability of the result. Figure 4 summarizes the formation and speciation of THMs from the final set of experiments, together with the Bromine Substitution Factor (BSF), a measure of the degree of bromination (Chow et al., 2007; Engelage and Stringfellow, 2009; Hua and Reckhow, 2012). Quantitatively, the BSF is the fraction of halogens added to molecule that are bromine atoms. Equation 2 shows the generic formula used to calculate BSF. Values a and b correspond with the number of chlorine or bromine atoms in each precursor, and both ranged from 0 to 2. The value n is the number of bromine atoms added in the final trihalogenated products, a number which ranges from 0 to 3.

$$BSF = \frac{\sum_{n=0}^{(3-a-b)} n \times [PrClBr_{b+n}]}{(3-a-b) \times \sum_{n=0}^{3} PrCl_{3-a-n}Br_{b+n}} \qquad \textit{(Eqn. 2)}$$

As anticipated, increasing bromide levels led to shifts in speciation toward the more brominated THMs. This is especially evident from the monotonic increase in BSF for all 5 precursors. In addition, some results showed increases in molar THM formation with increasing bromide. This was not expected as the overall rate limiting step for each was thought to be the halogen-independent enolate formation. Nevertheless, it will be shown that at least in the case of acetone, the increase is clearly a result of upward drifting pH in the unbuffered solutions.




561

562 **Figure 4. Summary of Formation of THMs from each Experiment with Different**
563 **Starting Precursors[2] (a), Acetone; (b) MCAce; (c) DCAce; (d) MBAce; and (e)**
564 **DBAce.**

565

566

[2] Note that t = reaction time; TCM = Trichloromethane; BDCM = Bromodichloromethane; CDBM = Chlorodibromomethane; TBM = Tribromomethane; TTHM = Total trihalomethanes. The axis for BSF (in red) is on the right and the axis for THM concentration (in black) is on the left.

### 2.3.4 Reaction with Acetone

In this study, we simplified the determination of the rate of reaction of acetone by using the formation of THMs, which was the final hydrolyzed product of acetone halogenation. This is in contrast to most prior model compound studies that relied on the rate of loss of chlorine for kinetic analysis. Under drinking water conditions, acetone's reaction with chlorine or bromine is thought to be first order in acetone, and possibly zero order in halogenation (Morris & Baum, 1978; Guthrie & Cossar, 1986). Because acetone was present in great excess, its concentration did not change significantly during the reaction period and its loss rate was nearly constant. And since we expect that all intermediate halogenated acetones react quickly forming trihaloacetones that are then hydrolyzed to THMs, the molar THM yield should be equal to the molar loss of acetone:

$$\frac{d[TTHM]}{dt} = -\frac{d[Ace]}{dt} = k_{obs}[Ace] \approx k_{obs}[Ace]_o \qquad \textit{(Equ. 3)}$$

Or:

$$k_{obs} \approx \frac{\Delta[TTHM]}{[Ace]_o t} \qquad \textit{(Equ. 4)}$$

As previously noted, under our conditions, the rate limiting step is expected to be an initial ionization, and not the subsequent halogen addition. The literature also suggests that the rate of ionization should be first order in hydroxide. So, using Guthrie's nomenclature[3]:

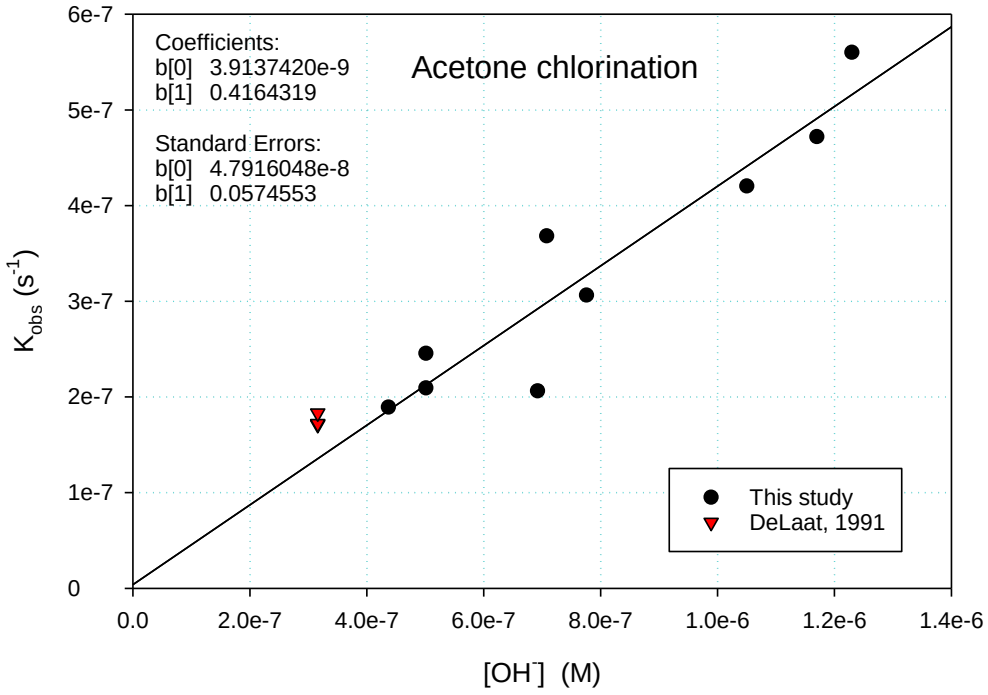
[3] Note that for the purpose of reducing difficulties for readers in comparing our calculation with others, we adopted the same naming system by Guthrie & Cossar (1986) and used the seemingly abrupt name of constant '$k_{12}$'.

585 $$k_{obs} = k_{12}[OH^-] \qquad (Equ.\ 5)$$

586 Where $k_{12}$ is the second order rate constant for the ionization of acetone.

587 De Laat (1981) and coworkers (1982) studied the chlorine-acetone reaction at pHs
588 from 7.5 up to 14. They confirmed that the reaction was indeed zero order in active
589 chlorine at circumneutral pH, supporting acetone ionization as the rate limiting step. At
590 pH 11.4, the reaction became first order in chlorine over the range of free chlorine
591 concentrations tested (2.6-60 mM; 200-4,000 mg/L).

592




593

594 **Figure 5. Observed THM Formation Rate from Acetone as a Function of Hydroxide**

595

Because the tests in our study were unbuffered, pH was not constant and it drifted upward with small additions of bromide. A least squares regression of the $K_{obs}$ values vs. $[OH^-]$ gives a straight line (Figure 5) with an intercept that includes zero ($0.4 \times 10^{-8} \pm 5 \times 10^{-8}$ $s^{-1}$). This supports the notion that it was slightly elevated pH that caused an increase in molar THM4 production with higher bromide, and it was likely unrelated to the bromide or hypobromous acid level itself. The data from De Laat (De Laat, 1981; De Laat et al., 1982) have a second order rate constant of 0.55 $M^{-1}s^{-1}$ and fall quite close to the regression line (Figure 5). The slope of our line ($0.42 \pm 0.02$ $M^{-1}s^{-1}$) is close to De Laat's and more than two times larger than the ionization rate constant determined by Bell and Longuet-Higgins (1946) at pH 12 (0.173 $M^{-1}s^{-1}$) and subsequently adopted by Guthrie. Chiang et al. (1992) criticized previous assessments of $k_{12}$ as having insufficient scavenging (i.e., low hypohalite), and they proposed 0.224 $M^{-1}s^{-1}$ as a better value. Note that all but our values and those of De Laat's were assessed at high pH and ionic strength. Since the accepted mechanism holds that the rate of acetone halogenation cannot exceed the rate of enolate formation (i.e., ionization), we propose that the apparent value for $k_{12}$ is slightly above 0.42 $M^{-1}s^{-1}$ when evaluated under conditions closer to drinking water practice.

While Guthrie's model and rate constants reasonably match data collected under conditions appropriate to drinking water systems (e.g., our data and DeLaat's), some adjustments in the rate constants give significantly better agreement while still retaining the core features. In re-casting the graphical model from Error! Reference source not found., we selected a value for 0.65 $M^{-1}s^{-1}$ for $k_{12}$ and adjusted $k_{21}$ accordingly based on the acetone pKa of 18.9 as determined by Tapuhi & Jenks (1982). We also elevated the

diffusion-controlled estimate for $k_{23o}$ to a higher value, but still within the range expected for diffusion-controlled reactions in water. This is in recognition of the large uncertainty for this value as noted by many (e.g., Tapuhi & Jencks, 1982; Guthrie & Cossar, 1986). This is also in recognition that estimated rates for these reactions are interdependent and many are likely to have been affected by general base catalysis from buffers used to control pH (e.g., see Tapuhi and Jencks, 1982, on uncertainty in estimating $k_{21}$). Figure 6 **s**how a revised version of Error! Reference source not found. including the experimental data from this work and that reported by De Laat.

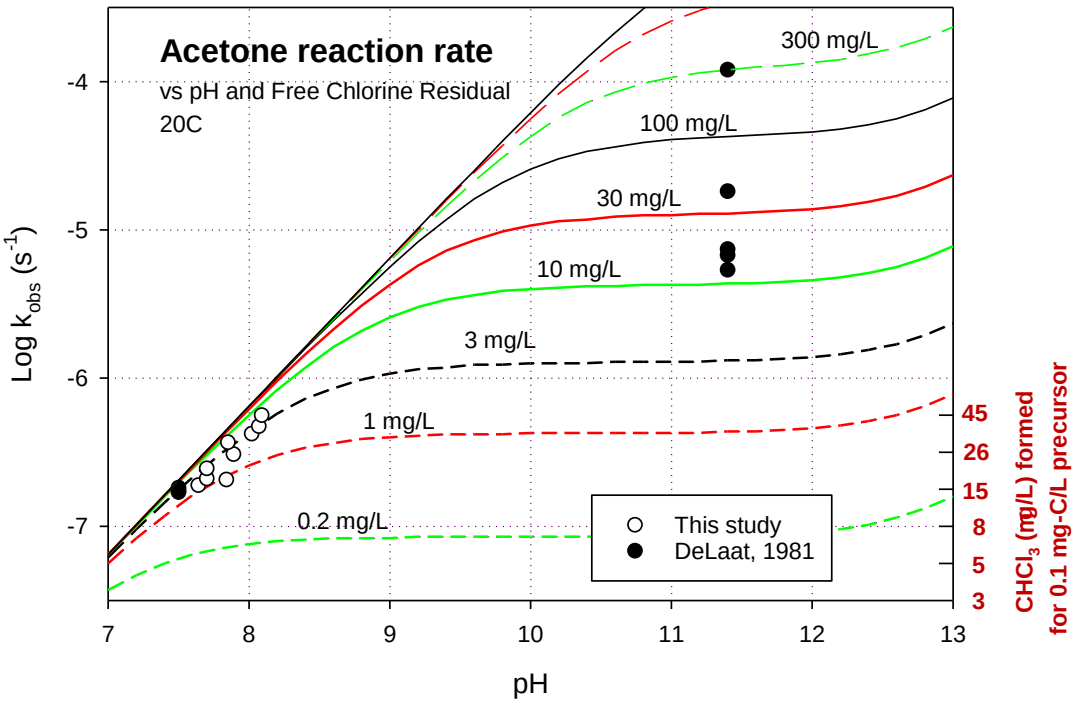



**Figure 6. Standard Acetone Halogenation Model[4].**

[4] Adapted to Drinking Water Conditions (using Tapuhi & Jenks's pKa).

631 **Table 2. Rate Constants used in Model with pKa from Tapuhi & Jenks**

| Rate constant | Fitted value | Guthrie values | Units |
|---|---|---|---|
| $k_{12}$ | 0.65 | 0.173 | $M^{-1}s^{-1}$ |
| $k_{21}$ | $5.1 \times 10^4$ | $4.1 \times 10^4$ | $s^{-1}$ |
| $k_{23o}$ | $7.0 \times 10^9$ | $1.7 \times 10^9$ | $M^{-1}s^{-1}$ |
| $k_{23-}$ | $2.0 \times 10^4$ | $2.0 \times 10^4$ | $M^{-1}s^{-1}$ |

632

633 Figure 6 also offers an opportunity to visualize the potential for "acetone-like"
634 precursors to play a role in THM formation. For example, if 0.1 mg-C/L (i.e., about 4%
635 of the bulk DOC of a typical finished water) is in the form of poorly-activated methyl
636 ketones or compounds that oxidize to methyl ketones of this type, then the right-side
637 scale indicates the amount of THM4 expected from the $k_{obs}$ values on the left-side scale.
638 This calculation is based on a 3-day water age. So at pH 8, between 5 and 26 µg/L
639 chloroform would originate from these types of precursors at the typical range of chlorine
640 residuals (0.2 to 1 mg/L as free chlorine). This is not unlike the slow THM formation that
641 is commonly observed in municipal water systems after the initial fast reactions have
642 largely been exhausted. Thus, these slow-reacting precursors cannot be dismissed based
643 on kinetic arguments.

644

645 **2.3.5 Analysis of Data for Overall Reaction of Halogenated Acetones**

646 As is the case for acetone, the overall rate of halogenation for each of the remaining
647 precursors can be determined by the THMs formed. This is true because the subsequent
648 halogenations are faster than the initial halogenation, and final hydrolysis converts the
649 trihaloacetones to THMs. The following equations were derived with varied precursor
650 [P], based on the same assumptions from Equation 3-4:

$$\frac{d[TTHM]}{dt} = -\frac{d[P]}{dt} = k_{obs}[P] \approx k_{obs}[P]_{avg} \quad \textit{(Equ. 6)}$$

Or:

$$k_{obs} \approx \frac{\Delta[TTHM]}{[P]_{avg}t} \quad \textit{(Equ. 7)}$$

And finally:

$$k_{P-OH} = \frac{k_{obs}}{[OH^-]} \approx \frac{\Delta[TTHM]}{[P]_{avg}[OH^-]t} \quad \textit{(Equ. 8)}$$

Where $k_{P\text{-}OH}$ is the second order rate constant which we interpret as the rate for the ionization of the precursor, P.

Each of the 9 sets of tests for each precursor led to an independent estimate of $k_{obs}$, and $k_{P\text{-}OH}$ (Table 3). In the acetone case, the $k_{P\text{-}OH}$ gave more accurate values based on relative standard deviation, and this is clearly due to the strong relationship between $k_{obs}$ vs. $[OH^-]$ as shown in Figure 5. For MBAce, DCAce and DBAce, the error in pH measurement seemed to have overwhelmed any underlying drift in $K_{obs}$, so that the latter exhibited less uncertainty. The case for MCAce was in between these two.

Aside from chloroform's enolate formation rate as discussed above, the parallel rates ($k_{P\text{-}OH}$) match Guthrie and Cossar's values quite well. Our estimate for the rate of ionization of MCAce (139 ±47 $M^{-1}s^{-1}$) and DCAce (3360 ±1200 $M^{-1}s^{-1}$) is nearly identical to their $k_{34}$* (136 $M^{-1}s^{-1}$) and $k_{56}$* value (3010 $M^{-1}s^{-1}$), respectively.

**Table 3. Kinetic Analysis of Data from Acetone and Haloacetone Experiments (25°C).**

| Precursor | Observed Rate Constants $k_{obs}$ ($s^{-1}$) | Ionization Rate Constants $K_{P\text{-}OH}$ ($M^{-1}s^{-1}$) | Halogen Competition Ratios $\gamma_P \equiv k_{P\text{-}HOBr}/k_{P\text{-}HOCl}$ |
|---|---|---|---|
| Acetone | 3.3 (±1.2) x$10^{-7}$ | 0.42 ±0.06 | 6.51 ±1.45 |
| MCAce | 1.15 (±0.44) x$10^{-4}$ | 139 ±47 | 1.80 ±0.13 |
| MBAce | 2.33 (±0.37) x$10^{-5}$ | 18.4 ±3.3 | 3.39 ±0.66 |
| DCAce | 2.15 (±0.2) x$10^{-3}$ | 3360 ±1200 | 0.82 ±0.13 |
| BCAce | NA | NA | 2.23 ±0.07 |
| DBAce | 1.15 (±0.1) x$10^{-3}$ | 2190 ±500 | 1.79 ±0.20 |

### 2.3.6 Analysis of Data for Competition Ratios

The next step in our kinetic analysis was to distinguish the relative rates of halogenation, comparing bromination with chlorination (i.e., the γ-value). Again, each of the 9 sets of tests for each precursor led to an independent estimate of this ratio. For example, the rate of formation of TCM from DCAce is characterized by $r_{4c}$ (Figure 3), which can be expressed with the following kinetic law. We use the [DCAce*] in recognition that this particular step might not involve direct reaction of hypohalous acid and the neutral precursor, DCAce. As already proposed, the [DCAce*] is most likely the enolate, a much more reactive form.

$$r_{4c} \equiv \frac{d[DCAce \rightarrow TCM]}{dt} = k_{4c}[HOCl][DCAce^*] \qquad \textit{(Equ. 9)}$$

Using our prior assumptions of constant reactant concentrations, and applying this to the reaction of DCAce, as an example, we can determine the fraction of DCAce that leads to chloroform formation ($fr_{DCAce \rightarrow TCM}$) using the experimental data. This is the amount of TCM formed over any reaction period (i.e., following pathway 4c in Figure 3) divided by the THM4 formed during that same reaction time (i.e., the total amount of DCAce that has reacted including both pathways, 4c and 4b). We presume that HOX (i.e., HOCl or

HOBr in this case) is the key reactive halogen species, because the alternative reaction is between two negatively charged species (i.e., the enolate and halogens) which is unlikely.

$$fr_{DCAce \rightarrow TCM} = \frac{\Delta[TCM]}{\Delta[TTHM]} = \frac{k_{4c}[HOCl][DCAce^*]}{k_{4c}[HOCl][DCAce^*] + k_{4b}[HOBr][DCAce^*]} \qquad \textit{(Equ. 10)}$$

$$\frac{k_{4b}}{k_{4c}} = \frac{\alpha_{HOCl}[Cl^+]}{\alpha_{HOBr}[Br^+]} \left( \frac{\Delta[TTHM]}{\Delta[TCM]} - 1 \right) \equiv \gamma_4 \qquad \textit{(Equ. 11)}$$

It should be noted that free chlorine dose is used for $[Cl^+]$[5]. Also, recognizing the oxidation of bromide to active bromine to be sufficiently fast, we used average inorganic bromine (initial bromide minus bromine incorporated into the THMs) as the best proxy for $[Br^+]$.

And more generally, for any precursor or NOM mixture, the bulk rate ratios (averaged over all subsequent steps between the precursor and the final THMs) can be determined from the increase in THM-bound bromine and THM-bound chlorine along with estimates of the average free chlorine and free bromine. Further simplification used BSF as a proxy for DBP fraction and ignored the speciation of halogens, resulting in a 'conditional' γ-value ($^c\gamma$) that is defined for a particular pH.

$$\frac{k_{Br}}{k_{Cl}} = \frac{[Cl^+]}{[Br^+]} \left( \frac{BSF}{1 - BSF} \right) \equiv c_\gamma \qquad \textit{(Equ. 12)}$$

The remaining halogen competition ratios (γ-values) are determined in an analogous fashion and they are presented in Table 3. The value of the ratio for BCAce ($k_{BCA\text{-}HOBr}/k_{BCA\text{-}HOCl}$), was based on MCAce data. A check on these estimates using the MBAce was generally about 40% lower, which might be a result of the low purity of our MBAce

[5] $[Cl^+]$ is not typo. As described in Section 2.2.4, it stands for active chlorine, such as [HOCl], $[OCl^-]$, $[Cl_2]$, $[Cl_2O]$ or some combination.

standard (e.g., there was an apparent brominated product even without added bromide). The species-specific, HOX-based γ-values are all low and ranging from 0.81 to 6.5 ($M^{-1}s^{-1}$/$M^{-1}s^{-1}$). When calculated as an overall value for acetone, ignoring intermediate steps, the γ-value is 3.4. Of these species-specific values, DCAce is the lowest, followed by DBAce and MCAce, then BCAce, MBAce and finally acetone, itself.

Some model validation is possible with this dataset as not all values were used for calibration. The halogen competition ratios (γ-value) were determined using only two THM measurements (i.e., most highly chlorinated product and THM4), which leaves certain degrees of freedom that were not used. In addition, the lowest and in one case (MBAce), the highest bromide levels were excluded from determination of γ due to the high relative standard errors from the low THM yield. Figure 7 shows a binomial plot with all acetone THM data along with the model predictions using all 6 γ-values from Table 3 (solid lines). Figure 7 also shows the generic binomial model where all steps have identical γ-values (dashed lines). Use of the γ-values specific to the acetone system resulted in improved predictions of TCM and BDCM. They slightly over predicted the peak of the CDBM, while TBM was well predicted by both models.

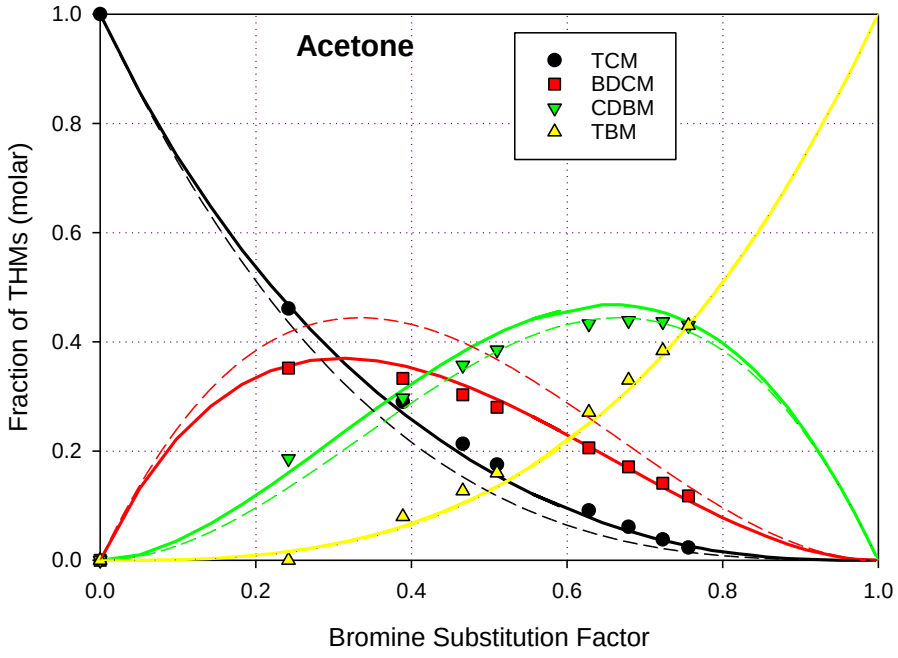



**Figure 7. Generic Binomial Plot, along with Acetone-specific Model[6], and Measured THMs from Acetone Tests.**

The full bromide-dependent competitive kinetic model allows testing of the ionization rate constants ($k_{P\text{-}OH}$) in Table 3 as well as the γ-values. Figure 8 presents the results of this model along with the measured data from experiments with acetone. The model predictions are not smooth functions because measured pHs were used instead of a single average pH value, and the measured values differed for each of the tests. Also, results are only affected by the acetone ionization rate which is considered as the rate limiting step. The model generally matches observed trends in THM concentrations. The greatest degree of divergence occurs at the highest bromide level (10.3 µM or 825 µg/L). This

[6] Dashed lines, same γ-value for each step - Generic Binomial Plot; Solid lines, Variable γ-values from Table 3 - Acetone-specific Model; Symbols- Measured THMs from Acetone Tests.

level of bromide concentration is not commonly seen (e.g., 95% of PWSs in the US have the average bromide concentration in MA below 200µg/L (Regli et al., 2015)).

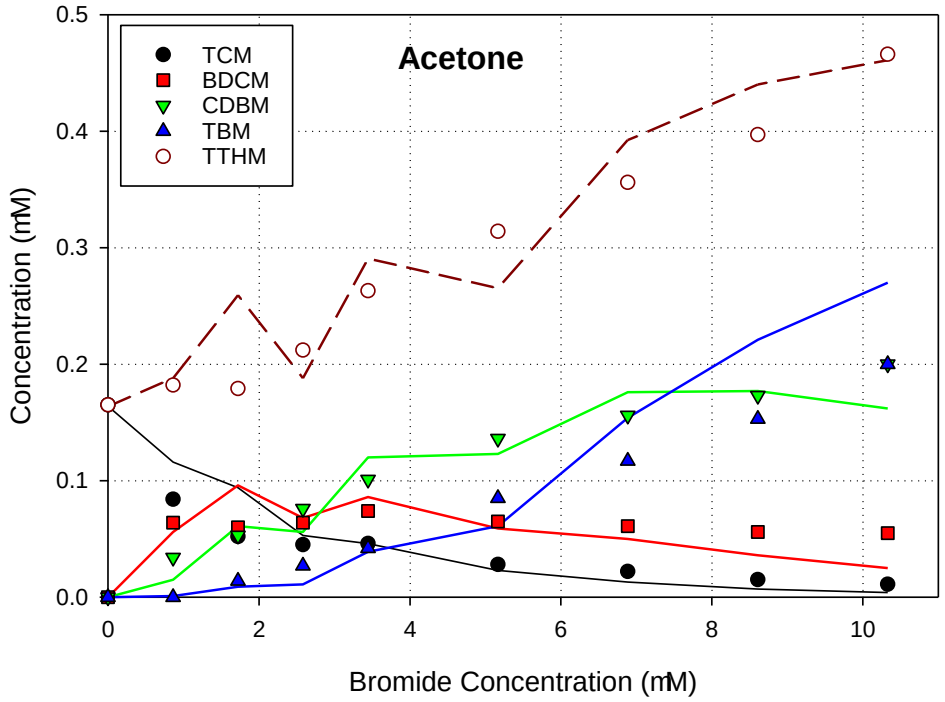



**Figure 8. Acetone Chlorination Model Predictions based on Measured pH, also Showing Measured THM Concentrations.**

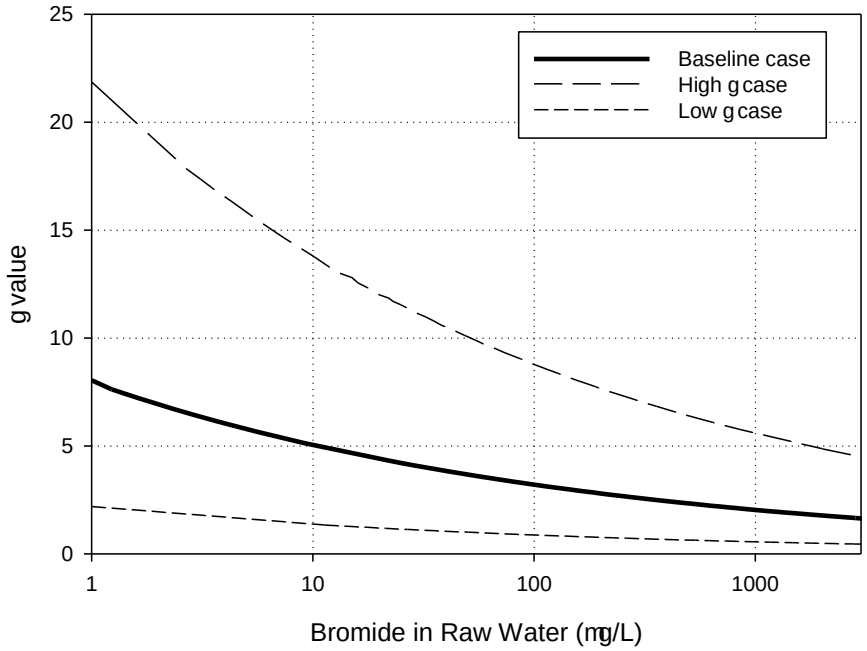
### 2.3.7 Reconciling Low γ-Values

On the surface, low γ-values in Table 3 appear to be in contrast to many publications that cite much more pronounced differences in the rates of bromine and chlorine with small molecules. Most often cited are the phenolic compounds that undergo activated aromatic substitution. Based on critical reviews of chlorination (Deborde & von Gunten, 2008) and bromination (Heeb et al., 2014), it is clear that the ratios of the 2$^{nd}$ order rate constants for bromine to those for chlorine average about 4000 (range: 2600- 5700) for

749 phenol and 10 of the mono and dihalogenated phenols. This is true whether the γ-values
750 are species-specific or "conditional", based on pH 7.

751 Many studies with natural waters and organic matter extracts have shown low γ-
752 values, many close to those noted here for acetone and its cascade of intermediates.
753 Nokes and colleagues (1999) studying 17 New Zealand waters found them to vary
754 between 4 and 15, with a best fit average of 9.1. Others have found best-fit γ-values in the
755 same range (10; Cowman & Singer, 1996) or slightly above (25; Chang et al., 2001).
756 Attempts to fit each of the 6 steps with a different γ-values have resulted in a range of 5
757 to 130 (Roccaro, et al., 2014). Multiple experiments with a single NOM extract spiked
758 with vary amounts of bromide resulted in γ-values mostly in the range of 17-20 (data
759 from Bond et al., 2014). When corrected for average pH (8.05), these decrease to about 6.
760 Analysis of THM formation in full-scale water treatment plants seems to support these
761 low γ-values. Francis and colleagues (2010) developed models for BSF from the
762 Information Collection Rule (ICR) database. From these models a range of γ-value
763 ranging from 2-7 would be expected for a typical treated surface water (Figure 9 and
764 Table 4). The range for low DOC waters more typical of groundwater systems ranges
765 from about 5-15.

766




767

768 **Figure 9. Model predictions based for ICR data based on Francis et al., 2010.**

769

770 **Table 4. Model input parameters for Figure 9.**

| Model value | Baseline | High γ | Low γ |
|---|---|---|---|
| Time in DS (hr) | 48 | 12 | 168 |
| FRC in DS (mg/L) | 0.5 | 1 | 0.1 |
| pH in DS | 7 | 7 | 7 |
| Alk influent (mg/L) | 41 | 300 | 5 |
| Temp (°C) | 10 | 10 | 10 |
| TOC @clpt (mg/L) | 1.75 | 1 | 4 |
| SUVA @clpt | 3 | 1.5 | 4 |
| Cl dose (mg/L) | 2.5 | 1.3 | 4 |

771

772 However, laboratory studies with NOM or full-scale treatment data can be difficult to
773 interpret in the context of model compound kinetic studies as NOM includes numerous of
774 different precursors that undergo successive exhaustion. Also most of the NOM studies
775 have been conducted with in-situ formed HOBr in the presence of a HOCl residual. Since

776 it's rare to measure HOBr or even inorganic bromine during these reactions, most
777 researchers have used the initial bromide concentration as a proxy for the extent of HOBr
778 exposure. It is also common to use chlorine dose as a proxy for HOCl exposure. Also,
779 bromine incorporation (e.g., BSF) is generally used as a measure of the extent of reaction
780 with reactive bromine. Since bromine is both an oxidant and a good leaving group, the
781 existence of C-Br bonds might lead to substantial underestimation of the amount of
782 reaction between NOM and active bromine species. All of this adds uncertainty when
783 comparing empirically-derived γ-values for NOM with those determined from laboratory
784 kinetics studies for a single precursor.

785 Some minor differences would also be expected from use of total residuals rather than
786 specific hypohalous concentrations (i.e., species-specific values vs conditional values).
787 The difference would be insignificant for acidic pHs, but would rise to a factor of 3 at pH
788 8, and to 8.5 at pH 9. This means that should the hypohalous acids prove to be the
789 reactive species, estimates based on the totals halogen residuals overestimate the species-
790 specific ratios by these pH-dependent values. In addition, there is a 2-3 order discrepancy
791 between the phenolic and the NOM γ-values.

792 We propose that there could be at least three reasons for the discrepancy in γ-values,
793 and have expressed them in the form of three hypotheses:

794 1. Electrophilic aromatic substitution (EAS) of the type that shows high bromine-to-
795 chlorine rate constant ratios (~3000) is not a major factor in the formation of C-X
796 bonds in regulated DBPs. Instead most of the regulated DBPs and their

intermediates (e.g., TOX) comes from reactions that do not show this apparent preference for HOBr.

2. The presence of bromine in DBPs is not a good indicator of the extent of reaction of bromine with NOM. Reasons might be bromine's ability to undergo electron transfer reactions, as well as its tendency to form metastable products that lead to hydrolysis and loss of TOBr (e.g., as a good leaving group)
3. By using initial bromide as a proxy for bromine exposure, we are substantially overestimating that exposure. Possibly because bromine is not immediately formed or because it becomes depleted long before chlorine does.

The first hypothesis has some strong support in the literature. There is much circumstantial evidence for the important role of activated aromatic structures in chlorine reactivity of NOM. This includes structural correlations (Reckhow et al., 1990; Croue et al., 2000; Weishaar et al., 2003) and correlations with UV absorbance (Edzwald et al., 1985; Korshin et al., 1997; Rouge at al., 2020). While studies of EAS for phenolic models shows γ-values of 3000 or more, the ratio may be much lower for other important precursors. For example, the next step in the reaction of trihalophenols is a slower ring cleavage where the γ-value decreases to 133 (Heeb et al., 2014). Resulting unsaturated aliphatic products and inactivated aromatics with such side chains (e.g., cinnamic acid) also react more slowly and are likely to have γ-values well below 100 and possibly as below 1 (Li et al., 2020).

At the extreme of reactivity are the polyphenols (e.g., tannic acid and phloroglucinol; Li et al., 2020) that can react with bromine at rates that are near diffusion controlled. This places a ceiling on the 2$^{nd}$ order rate constants (ca. $10^{10}$ $M^{-1}s^{-1}$) and has the effect of

masking inherent differences in reactivity of HOBr versus HOCl. As a result the γ-values for these extremely reactive precursors have been observed to decrease to as low as 10 (Criquet et al., 2015). This is also true for some reactive inorganic ions such as sulfite and iodide (Heeb et al., 2014). The small γ-values from our acetone tests may be due to similar diffusion limits on the reaction rate between the enolate ion and HOX. This may be generally true for aliphatic precursors that rely on formation of a cabanion as a preliminary step to halogenation.

The second hypothesis is also likely to play a role. Focus on substitution reactions in model compounds may artificially tip the halogen competition ratio in bromine's favor. For example, many have described the differences between chlorine and bromine, whereby the former tends to engage in oxidation reactions, whereas the latter tends more to substitution (e.g., Symons et al., 1993). Chlorine also is known to react primarily as an oxidant due to the electron donating capacity (Wenk et al., 2013). Comparisons of model compound rate constants indicate low ratios (i.e., γ~10) for phenolics that tend to undergo oxidation (forming quinones) as compared to the high ratios (i.e., γ=~3000) as already noted for those that undergo electrophilic aromatic substitution (e.g., Criquet et al., 2015).

Phenolic coupling products have been offered as evidence for rapid 1-electron transfers by chlorine forming phenate radicals (Xiang et al., 2021). Chlorination studies that cover the full timescale show that incorporation is favored at shorter water ages and at long water ages the reactions turn more toward oxidation (Reckhow, 1984; Li et al., 1998). This agrees with model compound work which shows that the fully-halogenated aromatic rings, if they are to produce THMs and HAAs, must undergo ring cleavage. This is generally a slower process (e.g., Lee & Morris, 1962), involving more oxidation

843 than incorporation (De Laat et al., 1982; Reckhow et al., 1990) often producing major
844 long-lived DBPs that are oxidized but not halogenated (e.g., butenedial; Prasse et al.,
845 2020).

846 Nevertheless, there is evidence that the fastest reactions between bromine and
847 activated aromatic compounds might first tend toward substitution, and later toward
848 oxidation (Echigo & Minear, 2006; Criquet et al., 2015). Working only with HOBr and
849 NOM, Echigo and Minear (2006) found 70-90% of the fast bromine demand (~1 second)
850 went to incorporation reactions whereas 10-30% were oxidation reactions. This fast
851 demand (0.3-0.8 µM-HOX/mg-C) corresponds to only about 2-6% of the eventual 3-day
852 chlorine demand[7]. Later work (Criquet et al., 2015) also focused on early stages of the
853 reaction with HOBr (0.28 µM-HOX/mg-C or about the first 2% of the 3-day chlorine
854 demand) showed much higher levels of oxidation vs incorporation (75-80% vs. 20-25%).
855 It might be possible to reconcile these two by recognizing immediate oxidation was
856 measured in the first whereas the second study measured delayed oxidation, which could
857 have included incorporation reactions that were followed by delayed hydrolysis and
858 bromide release. Observations of fast oxidation may be an indication that some of the
859 aromaticity of NOM is present as precursors that yield ortho and para-quinones (Criquet
860 et al., 2015). However, most of these studies pertain to very early reactions (e.g.,
861 seconds). There is ample evidence that in the longer term (i.e., hours to days) bromine is
862 released from C-Br bonds by reaction with hydroxide or chlorine. This is expected based
863 on known hydrolysis kinetics of halogenated compounds (Mabey & Mill, 1978) and has
864 been observed through loss of TOBr in chlorinated waters (Abusallout et al., 2017) or

[7] Assuming 1 mg-$Cl_2$ demand/ mg-C or 14 µM HOX/mg-C

865 increase in bromide (Tan et al., 2016). Once bromide is released, it is followed by rapid
866 re-oxidation to bromine and subsequent reaction with NOM. This bromine cycling is
867 driven by a persistent chlorine residual. The result is that each bromine atom may react
868 many times with NOM; something that does not occur with chlorine, nor does it occur for
869 bromine in studies using only HOBr. Bromine cycling invariably leads to
870 underestimation of the extent of bromine reaction in chlorine-containing waters when
871 gauged only by initial bromide concentration.

872 Regarding the third hypothesis, the use of initial bromide concentration as a proxy for
873 active bromine; and the use of chlorine dose for active chlorine is valid if bromide
874 oxidation is fast compared to reactions with organic precursors, and if the loss of active
875 bromine and chlorine parallel each other such that the ratios remain relatively constant.
876 There are reasons to believe that both of these conditions could be violated in many
877 cases. In a typical treatment system, the loss of chlorine is about 80-95% (e.g., dose of 2
878 mg/L leaving a residual of 0.1-0.4 mg/L). However, relatively little is known about the
879 loss of bromide or bromine residuals. Based on the limited published literature, free
880 chlorine systems with $[Br^-]$/TOC ratios $\geq$0.1 mg-Br/mg-C are expected to show excess
881 bromide upon addition of a quenching agent (Bond et al., 2014; Tan et al., 2016; Langsa
882 et al., 2017; Ackerson et al., 2020). This is a $[Br^-]$/TOC level that is near the median for
883 groundwater but above the 90%ile for surface water, as reported in the ICR database
884 (Obolensky et al., 2007). In at least one case, substantial bromide was found after
885 quenching a distributed water in a free chlorine system with a lower $[Br^-]$/TOC ratio
886 (0.05 mg-Br/mg-C; Tan et al., 2016). In a few cases, TOBr formation has been measured
887 and it can be used to reliably estimate inorganic bromine residual (Langsa et al., 2017).

For example, the calculated residual inorganic bromine and free chlorine were 29% and 14% of initial bromide and chlorine dose in tests run at 0.18 mg-$Br^-$/mg-C (Hua and Reckhow, 2007). Kim et al. (2020) worked at lower ratios (0.020 mg-Br/mg-TOC) and they noted increases in TOBr days after chlorine addition, suggesting a persistent bromine residual.

Heeb et al. (2014) performed some mathematical simulations of bromine and chlorine kinetics under conditions expected in water treatment. For many of these they assumed a $\gamma$-value of 3000, as derived from the phenolic model compound data. Not surprisingly, their predictions show nearly 100% bromide incorporation into organic matter in the pH range of 7-8.5.

In summary, published data can be used to support all three hypotheses. However, only the first hypothesis can by itself bridge the nearly 3 orders of magnitude difference in $\gamma$ values as noted. Therefore, we are proposing that non-EAS reactions are primarily responsible for the discrepancy in $\gamma$ values, but contributions from the other two hypotheses may contribute as well to a lesser degree in typical distribution system water.

### 2.3.8 Role of Slower Aliphatic Precursors in Water Treatment

Aside from advancing prior kinetic models on acetone chlorination, this study supports the importance of non-EAS reactions and aliphatic intermediates along the pathway from NOM to THMs. Most aliphatic intermediates will react slowly with active halogens and have only a modest preference for bromine addition over chlorine addition. There appears to be some impact of the number and type of halogens already

910 incorporated on reactive carbons affecting subsequent halogen addition reactions. This
911 tends to skew the binomial model describing bromine-chlorine competition in some
912 predictable ways. Results of the work with acetone may help to re-calibrate the binomial
913 model for THM formation from NOM.

914 This work and other recent literature also helps to shed light on the general way in
915 which THMs are formed in drinking water systems maintaining a free chlorine residual.
916 The weight of evidence suggests that chlorination of NOM is heavily influenced by
917 reactions with very reactive NOM (mostly phenolics; Figure 10), leading to extensive
918 substitution, oxidation and hydrolysis. In general, these molecules become rapidly
919 halogenated at free, unsubstituted ring carbons ortho- or para- to an activating substituent.

920 It is helpful to break out the first 1 minute of reaction, which is the period when free
921 chlorine rapidly oxidizes bromide to bromine (reaching about 95% completion after 60
922 seconds). For this purpose, we're focusing the discussion on uncontaminated surface
923 waters that are mostly free of reduced inorganic species. Reactions during that first
924 minute will be impacted by the mixing that occurs when a small flow (~0.1% of the bulk
925 flow) of concentrated chlorine solution[8] blends with the bulk water. Chlorine
926 concentrations at the expanding interface between the chlorine feed water and bulk flow
927 will be quite high and out-compete any oxidized bromine for reaction with NOM. During
928 this period, the most reactive NOM components such as polyphenols (e.g., tannins,
929 certain flavonoids) and some inorganic ions would react preferentially with chlorine until
930 the HOBr/HOCl ratio reaches a level where bromine reaction become predominant. Once

[8] In systems using gaseous chlorine that concentrated flow will be composed of treated water that has been mixed with gaseous chlorine for a timescale on the order of seconds. Byproducts from this tiny flow should not measurably impact the bulk treated water.

bromide oxidation is well along (a few seconds to a minute), it will out-compete chlorine for the remaining highly activated aromatics (e.g. some soluble lignin residues and other phenolic structures, reaction #4 in Figure 10) mostly through electrophilic aromatic substitution. During this period, the overall rate of substitution will be limited by the rate of bromide oxidation (reaction #3 in Figure 10). For example, according to the kinetic modeling data, the initial (1-minute) incorporation ratio would be about 7 M-Br/M-Cl for a typical phenolic ActAr ($k_{Cl}$=40$M^{-1}s^{-1}$, $k_{Br}$=1.2$e^5$ $M^{-1}s^{-1}$) in a water with typical levels of bromide and free chlorine (i.e., 100 µg-$Br^-$/L and 3 mg-$Cl_2$/L). Therefore, during this first minute, the existence of a prior rate limiting step for HOBr attack (i.e., bromide oxidation) has the effect of diminishing the γ down from a species-specific value of 3000 to an apparent value of 240 [species-specific γ= $k_{Br}/k_{Cl}$ =120000/40=3000; apparent γ= TOBr/TOCl * $Cl_2/Br^-$ = 7*($^3/_{71}$)/($^{0.1}/_{80}$)=240].

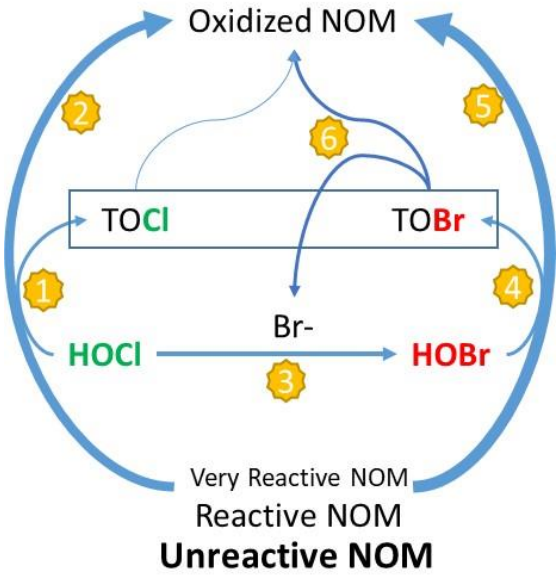


**Figure 10. Conceptual View of Chlorine-Bromide-NOM Reactions.**[9]

After that first minute, the rate of bromine substitution probably slows due to depletion of reactive aromatic structures or inorganic bromine or both. As noted above, inorganic bromine can become substantially depleted when the initial bromide to dissolved organic carbon ratio, $[Br_o]$/DOC, is low. After an hour, the available sites start to become exhausted, even without bromine and the slower process of ring cleavage and further reaction takes over. This is probably at about 30% of the eventual chlorine demand of phenolic structures (2-3 M/M of ring substitution as compared to about 8 M/M total demand). Ring cleavage and subsequent reactions are slower especially with reactive bromine, allowing inorganic bromine to persist at measureable levels if the [Br]/TOC ratio is high enough. Again, the relative incorporation is in the range of 1-10 M-Br/M-Cl, depending on residual inorganic bromine. At this point, the slower aliphatic reactions start to become more important. Examples include addition to activated olefins and the aforementioned enolates which depend on slow prior ionization. Both show little or no strong preference for reaction with bromine over chlorine (e.g., Li et al., 2020). In the case of the enolate reactions it is probably due to their near diffusion-controlled rate which has the effect of suppressing any inherent differences in the reaction rate of HOBr versus HOCl.

[9] In first few seconds chlorine reacts with the very reactive NOM producing TOCl (#1) and some oxidized products (#2). At the same time it oxidizes bromide (#3) and the resulting HOBr creates TOBr (#4) and more oxidized NOM (#5). Some of the TOBr hydrolyzes (#6) re-forming bromide and oxidized NOM. Eventually the VReNOM becomes exhausted and pool of Reactive NOM becomes the major reactants.

It is helpful to represent some of this in the form of simple kinetic rate laws and to parse NOM into reactive, very reactive [VReNOM] and unreactive components. Based on model compound studies the formation of organic chlorine by EAS from the very reactive aromatics should be a simple 2nd order process and the entire halogenation reaction can be expressed as[10]:

$$\frac{\Delta[TOBr]}{\Delta[TOCl]} \approx \frac{f}{k_1}\left(\frac{1}{\int[HOCl]t}\right)\left(\frac{[Br^-]_o}{[ReNOM]_o}\right) \sim \frac{1}{CT}\frac{[Br^-]_o}{DOC_o} \qquad \textit{(Equ. 13)}$$

Which implies that the degree of bromine incorporation decreases after bromide becomes largely depleted in relation to the chlorine exposure increases (i.e., chlorine CT).

[10] See Supplementary Information for the detail of the derivation.

# CHAPTER 3

# ACCURATE PREDICTION OF UNREGULATED HALOACETIC ACID SPECIES USING THE BINOMIAL MODEL

## 3.1 Introduction

Once built up certain level of basic understanding about the kinetic of THM formation, now we will discuss the actual design and the detail of the kinetic model that predicts the HAA formation and speciation.

One of the goals of this research was to predict the HAA9 concentration using only the regularly-measured DBPs (i.e., HAA5 and THM4). Several researchers (Cowman & Singer, 1995; Nokes et al., 1999; Roccaro et al., 2013; Roccaro et al., 2014) have established binomial models based indirectly on the ratio of reactivities of precursors with active bromine and chlorine to predict bromine substitution during chlorination. Yan et al. (2016) and Zhang et al. (2018) refined the simple binomial model adding specific reaction constants within each level of bromination or chlorination. Similar approaches were also applied by other researchers for treated drinking water collected in distribution systems (Samson et al., 2017). In this research, the binomial model was modified and tested against data from controlled bench-scale chlorination experiments.

This work focuses on the prediction of trihaloacetic acids (THAAs) and dihaloacetic acids (DHAAs) concentrations by the measurements of trichloroacetic acid (TCAA), dichloroacetic acid (DCAA), trichloromethane (TCM), and bromodichloromethane (BDCM). The monohaloacetic acids (MHAA, which include MCAA and MBAA were not considered because they are already measured as part of the HAA5 and they are

typically present at levels near or below method detection limits (MDLs). One main reason of selecting chlorinated DBP species for the prediction is for the potential less error in measurements and in further predictions.

The oxidation of bromide by reaction with chlorine was considered rapid, in comparison to the timeframe of interest of DBP formation (Heller-Grossman et al., 1993; von Gunten and Hoigne, 1996; Deborde and von Gunten, 2008). The concentration of bromide in most natural surface waters ranges from <5 to >400 µg/L (Amy et al., 1995) and is a crucial factor in both the formation and speciation of DBPs during chlorination (Chang et al., 2001; Liang and Singer, 2003; Jones et al., 2012; Pan and Zhang, 2013). To simulate the natural occurrence of bromide, high-purity sodium bromide was spiked into samples of six different raw surface waters in order to create scenarios with different bromide concentrations, holding the background NOM levels constant. Then the predicted speciation of THMs and HAAs from the model was compared with the experimental data. One important outcome of this research is a model and calibration protocol that can be used to predict HAA9 formation and speciation, using only the TCM, BDCM, TCAA, and DCAA concentrations that are regularly measured under the current D/DBP Rule.

## 3.2 Materials and Methods

### 3.2.1 Chemical and Reagents

Sodium Bromide (NaBr, >99.0%), pentane, sulfuric acid, Methyl tert-butyl ether (MtBE), methanol, sodium arsenite, and other reagents were purchased from either

Fisher-Scientific (Hampton, NH, US) or Sigma-Aldrich (St. Louis, MO, US) and used without further modification. All aqueous solutions were prepared with ultrapure water generated from a Milli-Q system (Millipore, Billerica, MA, US). Sodium hypochlorite solution was purchased as a nominal 5% solution (Fisher-Scientific) and was diluted by 10 times on the day of each experiment with concentration checked by DPD titration. The calibration standards of THMs and HAAs were prepared by diluting concentrated commercial products (Commercial mix 551A, 551B for THMs, and 552.2 for HAAs, Absolute Standards, Inc., Hamden, CT, US) into acetone and MTBE, respectively.

### 3.2.2 Raw Water

Six raw water samples were collected from surface water sources at six different locations in Massachusetts. Forge Pond, Metacomet Lake, and Swift River are located in or near the watershed of the Connecticut River and were chosen as the low alkalinity group. The high alkalinity group included samples from Lake Quannapowitt, Windsor Pond, and Lake Onota, located in the North Atlantic shore region, the Berkshire Mountains and the Housatonic River Valley, respectively. In each case 20 liters of water was collected in high-density polyethylene (HDPE) containers and, at the same time, selected physical and water quality parameters (i.e., pH, temperature, and hardness) were measured on site. The water samples were then transferred to the laboratory within 2 hours and were kept at 4°C until use.

A series of 3-day preliminary chlorine demand tests was conducted to determine the requisite chlorine dose to achieve the desired residual. It was necessary to do this for each

as the water samples contained a wide range of chlorine demanding substances as measured by dissolved organic carbon (DOC, 1.81-5.06 mg/L), and UV absorbance at 254 nm (UV254, 0.03-0.21 cm-1). THM4 and HAA9 concentrations after 3 days of chlorination (20°C) were measured along with the chlorine demand tests. All other measurements were made within 24 hours of the time of collection. Table 5 presents a summary of the general water quality for the six samples.

### 3.2.3 Experimental Methods

#### 3.2.3.1 Bench-scale Chlorination

Each of the six raw water samples was filtered (glass fiber filter (GF/F)) and distributed into ten 300 mL BOD bottles. The bottles were spiked with crystalline sodium bromide to achieve a range of concentrations (0, 10, 30, 50, 100, 150, 200, 250, 300, and 350 µg/L additional bromide). Then a small volume of sodium hypochlorite solution was added to achieve the target residual (2.5 ± 0.5mg/L as $Cl_2$ after a 72-hour incubation time). The exact chlorine dose was determined for each water using results from the preliminary 3-day chlorine demand tests. At the end of the incubation period, both pH and chlorine residual were measured. Then the chlorinated samples were quenched immediately with sodium arsenite (100 mg/L) and carefully transferred into amber 30-mL vials for further DBP extraction and analysis.

#### 3.2.3.2 Analytical Methods

In the field, hardness was measured by the EDTA titration using a commercial field kit (Hach Model 5-B, Hach). Water temperature and pH were measured with a portable sensor (Accumet AP85, Fisher-Scientific). In the laboratory, alkalinity was measured by titration using bromocresol green-methyl red indicator (Hach Model AL-AP, Hach). Chlorine residuals were measured by the DPD ferrous titrimetric method (APHA, 1989). Total organic carbon (TOC) and dissolved organic carbon (DOC) were analyzed on a high temperature combustion analyzer (Shimadzu 5000, Shimadzu Corporation, Kyoto, Japan) according to the Standard Method 5310B (AHPA, 1989). The UV absorbance was measured with a diode-array UV-Vis spectrophotometer (Agilent 8453, Agilent Technologies, CA, US). Bromide concentration was measured by ion chromatography (Metrohm IC 850, Metrohm AG, Switzerland).

Four THMs were analyzed by liquid/liquid extraction with pentane followed by gas chromatography and electron capture detection and using a DB-5 column (GC-ECD, Agilent 6890, Agilent Technologies) in accordance with USEPA Method 551.1. Nine HAAs were quantified by liquid/liquid extraction with MtBE followed by derivatization with acidic methanol and GC-ECD with a DB-1 column (GC-ECD, Agilent 6890, Agilent Technologies) in accordance with USEPA Method 551.2. Samples for both THMs and HAAs samples were quenched right after incubation and extracted within

**Table 5. Key Raw Water Quality Parameters.** [1]

| Water No. | Sampling Location | DOC (mg/L) | UV254 (cm-1) | SUVA (L/(mg-m)) | Bromide (µg/L) | Alkalinity (mg/L as $CaCO_3$) | Hardness (mg/L as $CaCO_3$) | pH | Chlorine Demand (mg/L) | THM4 (µg/L) | DHAA[2] (µg/L) | THAA[3] (µg/L) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **S1** | Forge Pond, Granby | 4.77 | 0.21 | 4.49 | 13.6 | 20.4 | 80 | 7.32 | 8.44 | 266.2 | 449.4 | 461.8 |
| **S2** | Metacomet Lake, Belchertown | 3.80 | 0.21 | 5.45 | 9.03 | 6.8 | 40 | 6.84 | 5.78 | 192.2 | 388.3 | 396.4 |
| **S3** | Swift River, Ware | 1.93 | 0.03 | 1.44 | 12.9 | 6.8 | <20 | 6.95 | 1.24 | 53.7 | 54.3 | 58.9 |
| **S4** | Lake Quannapowitt, Wakefield | 5.06 | 0.14 | 2.76 | 46.4 | 20.4 | 120 | 6.98 | 4.37 | 102.2 | 95.5 | 108.6 |
| **S5** | Windsor Pond, North Adams | 3.76 | 0.08 | 2.22 | 6.1 | 95.2 | 180 | 8.13 | 3.89 | 128.5 | 104.7 | 112.5 |
| **S6** | Lake Onota, Pittsfield | 1.81 | 0.12 | 6.85 | <5.0 | 81.6 | 180 | 7.98 | 3.08 | 63.7 | 60.5 | 64.6 |

[1], Chlorine demand and DBP concentrations were based on our standard chlorination conditions: 72 hours reaction time, 2.5 mg/L target residual, 20°C. [2], DHAA is the summation of concentrations of dichloroacetic acid, bromochloroacetic acid, and dibromoacetic acid. [3], THAA is the summation of concentrations of trichloroacetic acid, bromodichloroacetic acid, chlorodibromoacetic acid, and tribromoacetic acid.

5 hours. All extracts were kept under 0°C and analyzed by GC-ECD within 3 days. Once the GC-ECD analyses were finished, the recovery and linearity of each DBP species in the calibration standards were checked against historical data records to ensure the reliability. The measured slopes in this research mostly lied within 25-75% of the historical records and the results were reliable. Calibrations were based on peak area ratios to internal standards. In addition, to comparing calibration slopes of relative peak areas, direct calibration slopes and internal standard peak areas were also examined and compared to historical values (Table 6).

**Table 6. Details of Slopes of Calibration Standards for each DBP Species.** [11]

| | | Historical Records, Percentile | | | | | | This |
|---|---|---|---|---|---|---|---|---|
| | | 5% | 25% | 50% | 75% | 95% | mean | Research |
| HAAs | MCAA | 8.23E-05 | 0.000145 | 0.0002 | 0.000238 | 0.000698 | 0.000224 | 0.000169 |
| | MBAA | 0.001281 | 0.002625 | 0.003015 | 0.003718 | 0.007424 | 0.003778 | 0.003243 |
| | DCAA | 0.002512 | 0.005134 | 0.00584 | 0.0067 | 0.012505 | 0.006618 | 0.005552 |
| | BCAA | 0.006875 | 0.012935 | 0.0146 | 0.017425 | 0.036181 | 0.016876 | 0.016084 |
| | TCAA | 0.009525 | 0.015812 | 0.01912 | 0.022325 | 0.03098 | 0.020675 | 0.015816 |
| | DBAA | 0.00912 | 0.01539 | 0.017916 | 0.0211 | 0.045681 | 0.020749 | 0.021423 |
| | BDCAA | 0.007003 | 0.01379 | 0.019665 | 0.026026 | 0.029786 | 0.019436 | 0.01499 |
| | CDBAA | 0.000763 | 0.004023 | 0.007093 | 0.01203 | 0.0163 | 0.007998 | 0.007869 |
| | TBAA | 4.97E-05 | 0.00029 | 0.0009 | 0.004431 | 0.00636 | 0.002928 | 0.003368 |
| THMs | CHCl3 | 0.003835 | 0.007 | 0.0089 | 0.010648 | 0.016319 | 0.009274 | 0.013818 |
| | CHCl2Br | 0.027247 | 0.041889 | 0.054347 | 0.065735 | 0.095642 | 0.056258 | 0.084146 |
| | CHClBr2 | 0.027238 | 0.039684 | 0.048758 | 0.056907 | 0.089426 | 0.050523 | 0.071027 |
| | CHBr3 | 0.012097 | 0.016456 | 0.01965 | 0.022435 | 0.03689 | 0.020492 | 0.02546 |

10. Note that the 'historical records' in this table summarized all DBP analysis records from the year of 2011 -2019 at the University of Massachusetts, Amherst.

### 3.2.4 Model

The monotonic progression from fully chlorinated to fully brominated THMs (i.e., from TCM to BDCM to CDBM to TBM) with increasing bromide levels was noted by early DBP researchers (Lange and Kawczynski, 1978; Minear and Bird, 1980). The formation of brominated DBPs is inevitable in the presence of bromide as the oxidation of bromide is fast, and the rate constant for bromine addition is larger than for chlorine addition (Westerhoff et al., 2004). Nevertheless, until the mid-1990s models for mixed bromine and chlorine substitution were almost entirely in the form of empirical power functions. When studies of HAAs showed similar patterns to the THMs (Pourmoghaddas et al., 1993), the idea of a simple competitive kinetic model blossomed. MacNeill (1994) proposed that the BDCM/TCM ratio was a good indicator of that competition and that this ratio changed little with changing chlorine contact time. Cowman and Singer (1995) also published a complete probability-based model, what we are calling the binomial model. In its original form, it assumed a fixed probability of bromine addition versus chlorine addition (what we will call $\beta$) for any particular precursor at any point in time after addition of active chlorine. Furthermore, it presumed that the competition term, or probability ratio (p) is directly proportional to the ratio of the two halogenating agents (e.g., [HOBr/HOCl] or more generally [Br(+I)]/[Cl(+I)]). This is the likely outcome if one presumes that all halogenation reactions occur via similar mechanisms. For example, a presumption of simple $2^{nd}$ order kinetics leads to the following rate expressions for chlorination and bromination of any single precursor 'P'. It should be noted that P could

be any DBP precursor and that the rate constants $k_{Cl}$ and $k_{Br}$ will be different for different Ps.

$$rate\ of\ chlorination = k_{Cl}[P][Cl(+\mathrm{I})] \qquad \textit{(Equ. 14)}$$

$$rate\ of\ bromination = k_{Br}[P][Br(+\mathrm{I})] \qquad \textit{(Equ. 15)}$$

And then the ratio of the two is referred to β.

$$\frac{rate\ of\ bromination}{rate\ of\ chlorination} \equiv \beta = \frac{k_{Br}}{k_{Cl}}\frac{[Br(+\mathrm{I})]}{[Cl(+\mathrm{I})]} \qquad \textit{(Equ. 16)}$$

These authors used the initial bromide level as a proxy for the [Br(+I)] and the chlorine dose (corrected for loss due to the reaction with bromide) for [Cl(+I)]. With this information and an estimate of β, one can calculate γ-value, the ratio of the rate constants.

$$ratio\ of\ rate\ constants, \gamma \equiv \frac{k_{Br}}{k_{Cl}} = \beta\frac{[Cl(+\mathrm{I})]}{[Br(+\mathrm{I})]} \qquad \textit{(Equ. 17)}$$

They found that a γ-value of 10 best fit their HAA data, including THAAs, DHAAs and MHAAs. The analysis of the Minear and Bird (1980) data produced a good fit with a value of 20 for THM speciation from a commercial humic acid. This suggested that the model described both THMs and HAAs, but that γ-value was not the same for all waters. Nokes et al. (1999) used a calculated average chlorine residual (average of dose and the final measured residual) for [Cl(+I)] and found a γ of about 9 to best fit their data. As we mentioned in the previous chapter, we have also noticed γ-values in these reactions are close but not identical.

For the purposes of comparing DBP data within a single sample, we have simplified the original binomial model by reducing the competition term ($\gamma\frac{[Br(+I)]}{[Cl(+I)]}$) to a constant, β.

The justification for this is that all precursors present in a water sample undergoing chlorination are exposed to the same concentrations of active halogenating species, and therefore, the ratio of those species can be treated as a fixed value for that water sample. This also means that each water undergoing chlorination should have a unique β, and it will be shown below that the value of β can be easily calculated from just two of the measured THM species: TCM and BDCM.

The logic and inherent assumptions in this probabilistic model are best illustrated by taking binary steps across the three halogenation reactions during the formation of THMs or THAAs (Figure 11). Note that for DHAAs, the reaction stops at the second halogenation step. The value 'x' is the probability that the next halogen atom added to the molecule is a chlorine atom and '1-x' is the probability that it is a bromine atom (note that this model does not consider iodine or iodinated DBPs). Our value for 'β' is then defined as the ratio of these two probabilities.

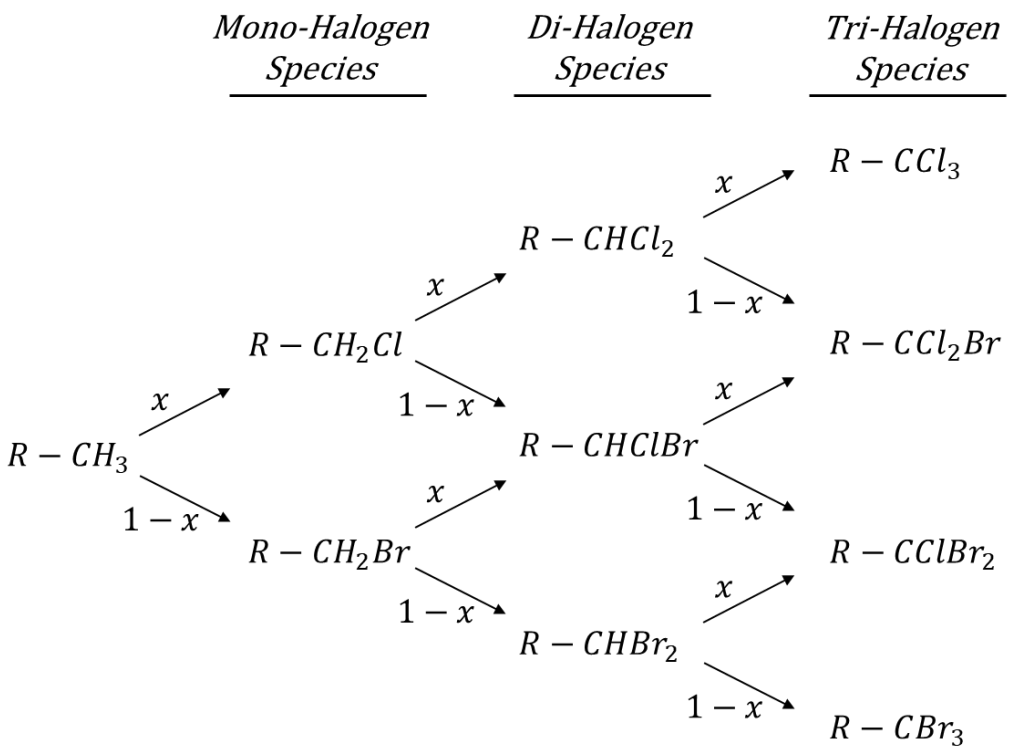



**Figure 11. Basic Framework for the Binomial Kinetic Model. The Probability of Chlorine Addition (x) was Considered Uniform across all 6 Steps. "R-" Represents any Organic Group.**

$$\beta = \frac{1-x}{x} \quad or \quad x = \frac{1}{1+\beta} \qquad \textit{(Equ. 18)}$$

Some researchers have calibrated models where each of the three halogenation steps has a different probability ratio (x/(1-x), where x may vary with each step in Figure 11, thus β and γ vary as well). Others have gone further to propose all 6 reactions could have unique ratios (Nokes et al., 1999). Given the purpose of our study and the lack of data on intermediate species, we have decided to adopt a single ratio for all 6 of the steps in Figure 11.

The value of 'x', and therefore, 'β', can then be easily calculated using readily available THM data. The most practical way is to first calculate the molar ratio of the fully chlorinated species to the species with one bromine atom. Then the alpha-value, the

fraction of each DBP species as compared to the total in their class, can be easily estimated by the product of the probabilities of each successive step (i.e., either x or 1-x). Equations 19a-19d show the result of this calculation for each of the THAA species. The corresponding values for each DHAA species are presented in Equations 20a-20c.

$$\alpha_{Cl3}\left(\equiv \frac{[TCAA]}{[THAA_T]}\right) = x^3 \qquad \textit{(Equ. 19a)}$$

$$\alpha_{BrCl2}\left(\equiv \frac{[BDCAA]}{[THAA_T]}\right) = 3x^2(1-x) \qquad \textit{(Equ. 19b)}$$

$$\alpha_{Br2Cl}\left(\equiv \frac{[CDBAA]}{[THAA_T]}\right) = 3x(1-x)^2 \qquad \textit{(Equ. 19c)}$$

$$\alpha_{Br3}\left(\equiv \frac{[TBAA]}{[THAA_T]}\right) = (1-x)^3 \qquad \textit{(Equ. 19d)}$$

$$\alpha_{Cl2}\left(\equiv \frac{[DCAA]}{[DHAA_T]}\right) = x^2 \qquad \textit{(Equ. 20a)}$$

$$\alpha_{BrCl}\left(\equiv \frac{[BCAA]}{[DHAA_T]}\right) = 2x(1-x) \qquad \textit{(Equ. 20b)}$$

$$\alpha_{Br2}\left(\equiv \frac{[DBAA]}{[DHAA_T]}\right) = (1-x)^2 \qquad \textit{(Equ. 20c)}$$

The value of x can be easily calculated from the molar ratio of the fully chlorinated DBP species over the DBP species containing one bromine atom (Equation 21). This model uses a single x or β value for all DBPs, which is based on the assumption that the degree of bromination for halogenated DBPs is the same across all geminal trihalogenated and dihalogenated DBPs. As a result, it only needs to be calculated from one DBP group.

$$x = \frac{3[TCM]/[BDCM]}{1+3[TCM]/[BDCM]} \qquad \textit{(Equ. 21)}$$

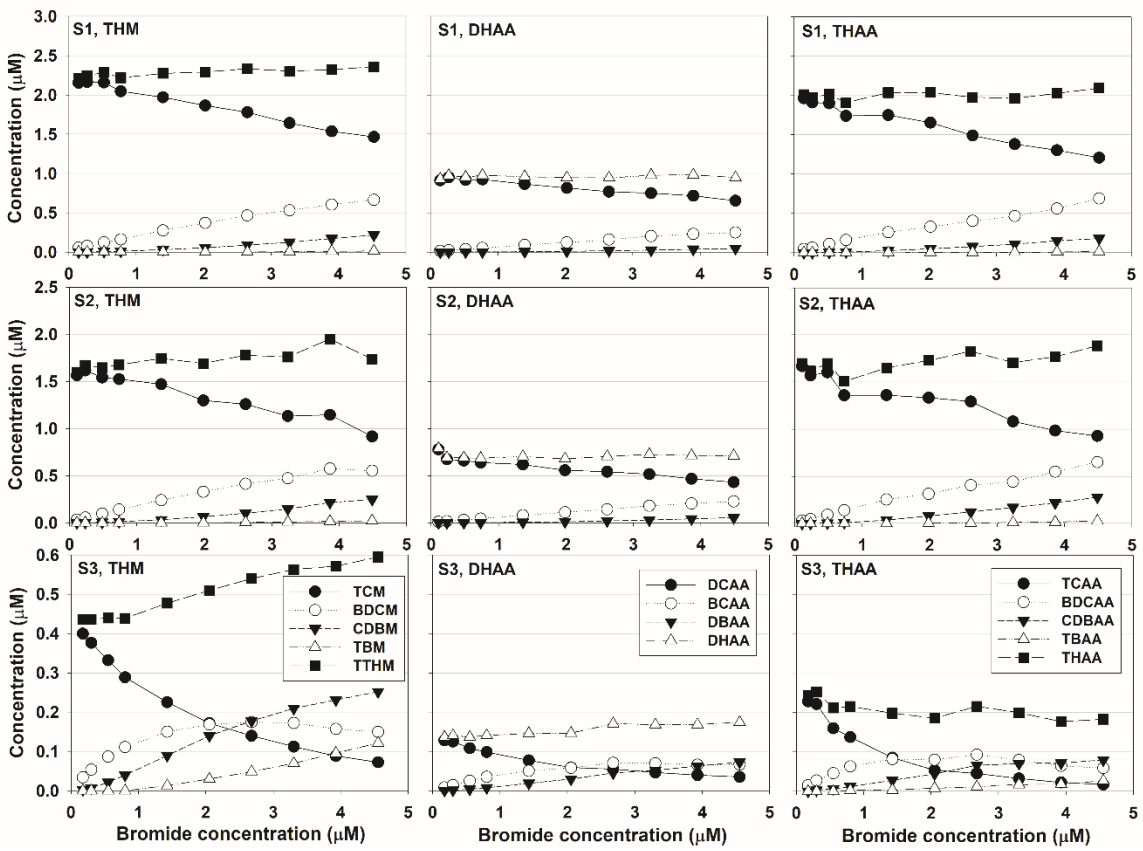
Since THM is the group that is most commonly measured, we recommend that the two most abundant THM (usually TCM and BDCM) be used to calibrate the model for any particular water sample.

## 3.3 Result and Discussion

### 3.3.1 Laboratory Test Waters

Free chlorine residuals for each bench-scale chlorination test were measured after incubation and samples were discarded when the residual did not fall within the range of 2.5 ±0.5 mg/L as $Cl_2$. The resulting DBP concentrations are presented in Figure 12. It should be noted that bromide concentrations used in Figure 12 are the sum of the background bromide concentrations and spiked bromide doses.

In addition to changes in speciation, the total molar concentration of DBPs was found to be affected by bromide addition. The molar yields of THMs and DHAAs generally showed an increasing trend with increases in bromide, whereas, the THAAs showed both increases and decreases. For example, addition of 350 µg/L of bromide resulted in increases of THMs, DHAAs, and THAAs ranging from 7% to 37%, -10% to 32%, and -37% to 22%, respectively. The magnitude of the increase in THM4 molar concentration in the presence of elevated bromide was higher than that of DHAA and THAA. This




115

116 **Figure 12. Molar DBP Concentrations from the Chlorination of all Six Raw Water**
117 **Samples.**

118

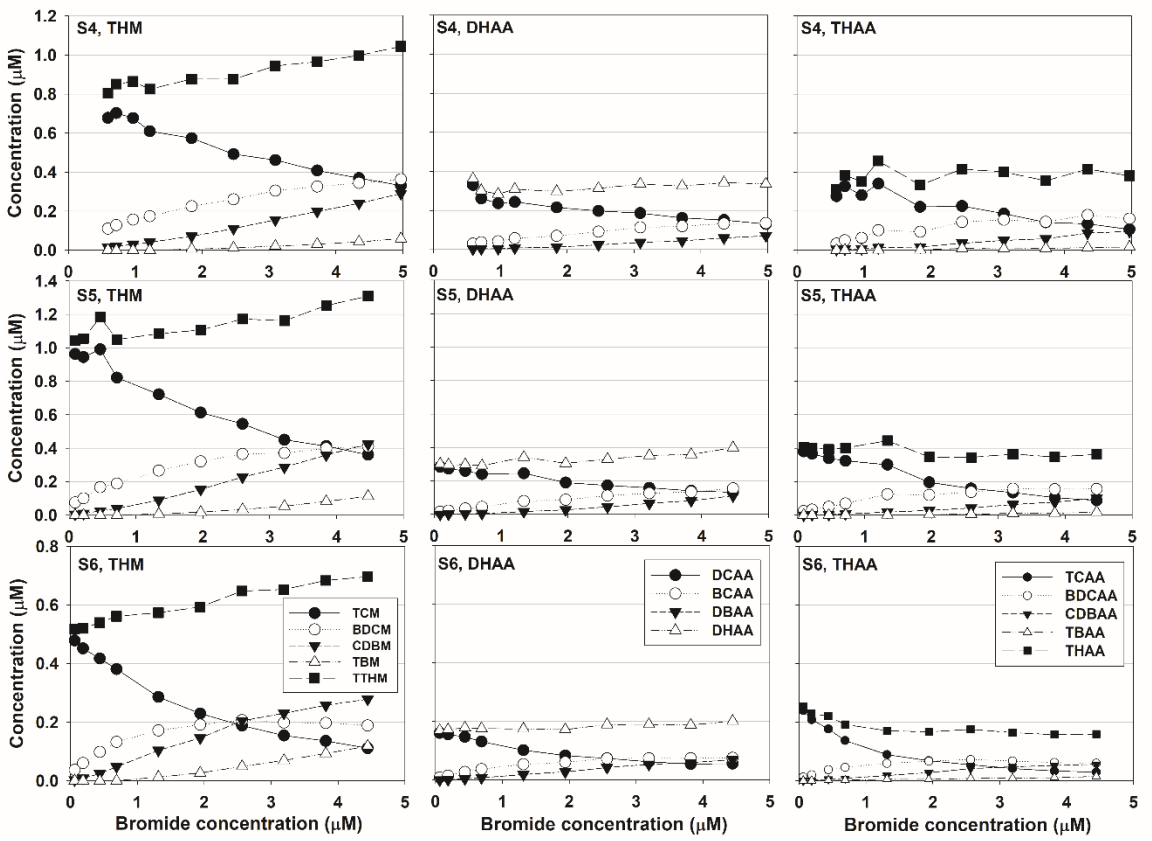



**Figure 12 (Cont.). Molar DBP Concentrations from the Chlorination of all Six Raw Water Samples.**

trend was also reported by Hua et al. (2006) and Wu and Chadik (1998). However, Cowman and Singer (1995) observed the opposite effect. We noticed THAAs often showed a decreasing trend with increasing bromide concentration, especially in low NOM containing waters (e.g., water S3 and S6, with corresponding DOC concentrations of 1.9 and 1.8 mg/L). These two waters were also the least productive as measured by DBP formation. It might be significant that Cowman and Singer (1995) used isolated humic substances in their experiments while Hua et al. (2006) and Wu and Chadik (1998) used a range of raw surface waters. The difference in humic component might have contributed to the different result. In addition to the influence of NOM chemistry,

differences in free chlorine contact times might also play a role. It was reported by Wu and Chadik (1998) that the increase in HAA molar concentration with elevated bromide was more obvious with longer chlorine contact time. Cowman and Singer (1995) used 24 h as their chlorine reaction time, which is shorter than in this present study (72 h), and the chlorine contact times in the work of Hua et al. (2006) and Wu and Chadik (1998) were 48 h and 168 h, respectively. The relative molar THM increase was found to be lower in water S1 and S2, which had higher SUVA values (4.49 and 5.45 L/m/mg, respectively). The largest increase in THM yield was found in water S3, which had the lowest SUVA (1.44 L/m/mg).

### 3.3.2 Testing the Binomial Model

The most obvious effect of increasing bromide level is the shift from fully chlorinated species to mixed bromochloro species and finally to the fully brominated species. BSFs[12] for each analyzed DBP class increase monotonically with increases in the bromide concentration (Figure 13). BSFs of THMs and THAAs were very similar within each sample and were only slightly lower than BSFs for DHAAs (0-57% for THM, 0-60% for DHAA, and 0-55% for THAA).

[12] The definition of BSF is shown in previous section.

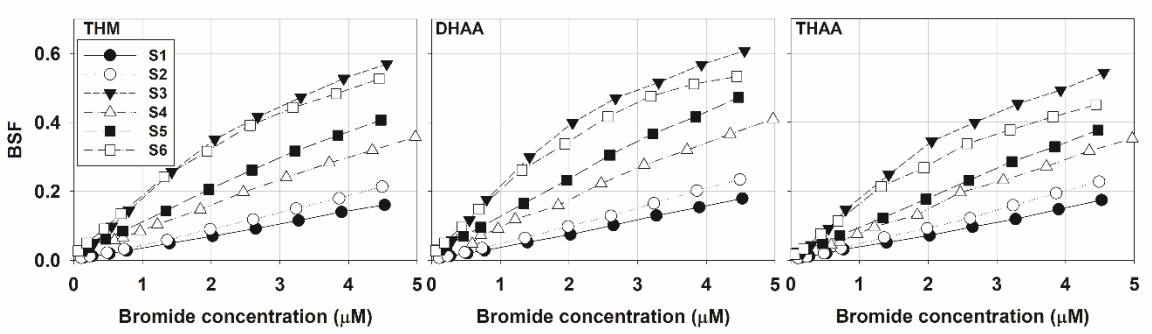



**Figure 13. Details of Impact of the Increase of Bromide Molar Concentrations on BSFs of THM, DHAA, and THAA.**

Note that BSF can also be calculated from the alpha values (Equations 22a-22b), and substituting Equations 19-20 into Equations 22a-22b. When applied to the binomial model, BSF is also a direct function of the probability "x" (Equation 23).

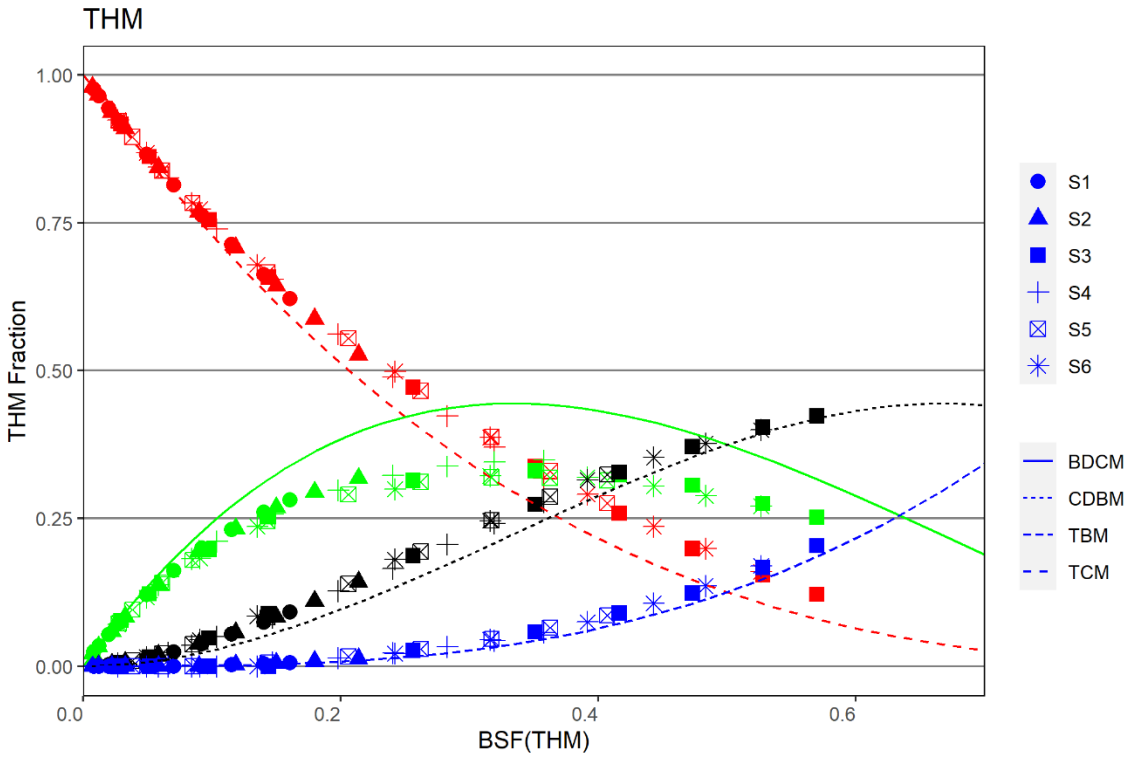
$$BSF = \frac{\alpha_{BrCl_2} + 2\alpha_{Br_2Cl} + 3\alpha_{Br_3}}{3}\text{, for THAAs} \qquad \textit{(Equ. 22a)}$$

$$BSF = \frac{\alpha_{BrCl} + 2\alpha_{Br_2}}{2}\text{, for DHAAs} \qquad \textit{(Equ. 22b)}$$

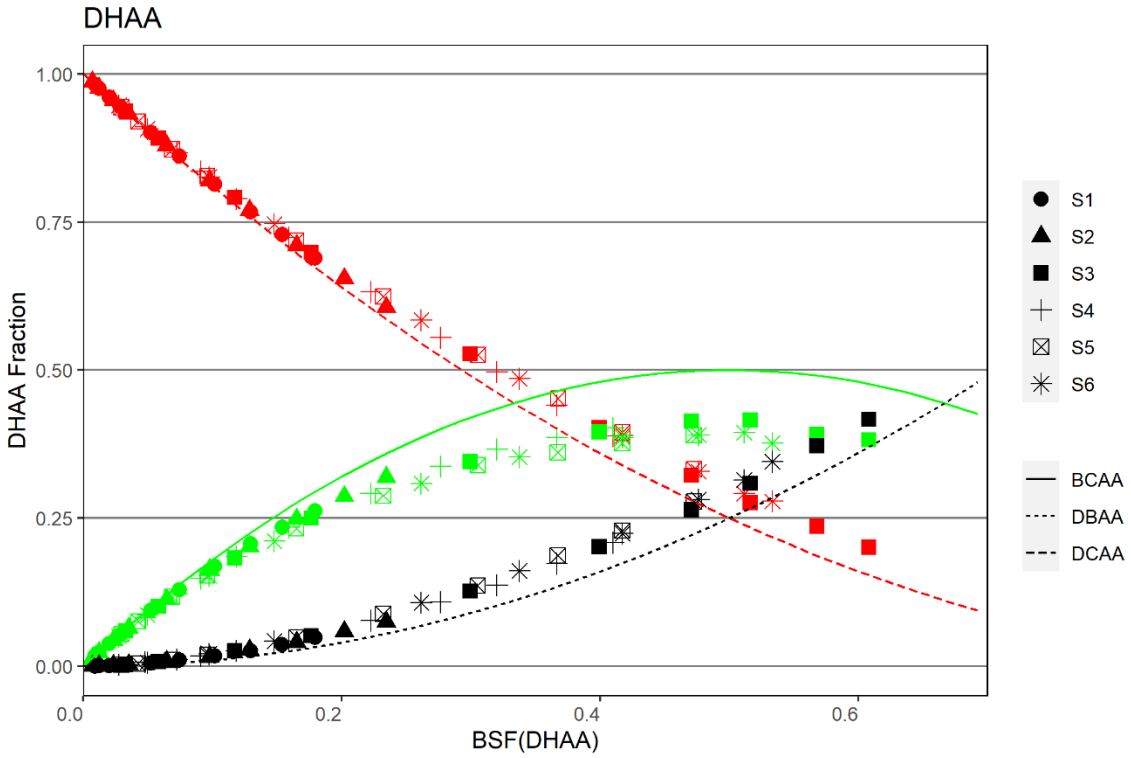
$$BSF = 1 - x \qquad \textit{(Equ. 23)}$$

Adherence to the binomial model may be most directly observed from BSF distribution diagrams where experimental data are transformed into alpha values and plotted versus BSF. Figure 14 illustrates this for the six test waters used in this study along with the model predictions (i.e., lines based on Equations S4a-S4d). This shows close agreement between the experimental data (points) and the model (lines), and it is evidence that the binomial model works reasonably well when based entirely on measured concentrations without independent estimates based on γ or β values. In some instances, however, systematic over-prediction and under-prediction was observed for

both THM and DHAA species. These deviations may be interpreted as being due to differences in relative rates of bromination to chlorination (i.e., β values) for the six halogenation steps delineated in Figure 11. This led us to adjust the binomial model as will be discussed in later sections and motivated us to suggest the model users to create their own practical 'patch' to more accurately adjust for the systematic error they may encounter. Others have used this same method of presentation with similar results (Chang et al., 2001; Reckhow & Singer, 2011). While not used to create Figure 14, β values could be calculated for each sample and for each DBP class. It should also be noted that models using different probability ratios for each reaction step will yield different predictions than shown in Figure 14. Thus, the strong model agreement evident in Figure 14 supports our decision to use only a single β or γ value for all six reactions as a good first approximation.


THM
1.00
0.75
0.50
0.25
0.00
THM Fraction
0.0
0.2
0.4
0.6
BSF(THM)
S1
S2
S3
S4
S5
S6
BDCM
CDBM
TBM
TCM

180


DHAA
1.00
0.75
0.50
0.25
0.00
DHAA Fraction
0.0
0.2
0.4
0.6
BSF(DHAA)
S1
S2
S3
S4
S5
S6
BCAA
DBAA
DCAA

181

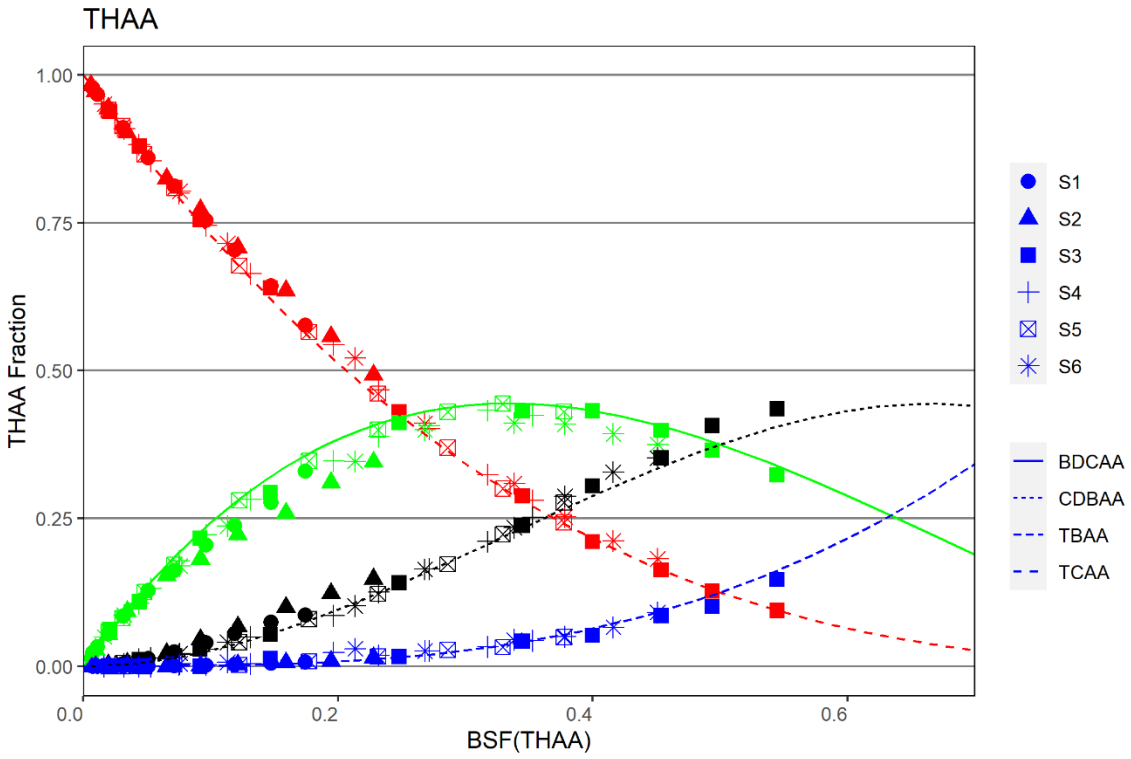



**Figure 14. Distribution of THMs, DHAAs, and THAAs Species as a Function of BSF. Comparison of Simple Binomial Model to Experimental Data.**

### 3.3.3 Testing the Assumption of a Common β and γ

A key premise of the simple binomial model is that the β values are the same for each DBP class in a given chlorinated sample. As mentioned in the previous sections, the calculation of γ follows the equations below. Equation 24 shows the calculation of γ based on the THMs. The calculation for γ based on DHAAs or THAAs follows the same pattern.

$$\gamma_{THM} = \beta_{THM}\left(\frac{[Cl(+I)]}{[Br(+I)]}\right) \approx \frac{1}{3[TCM]/[BDCM]} \cdot \frac{[Cl_2\ dose]}{[initial\ Br^-]} \qquad \textit{(Equ. 24)}$$

Because the ratio of active bromine to active chlorine (i.e., ($\frac{[Br(+I)]}{[Cl(+I)]}$)) is a unique value for each sample and shared across all of the DBP precursors in that sample, γ should be fixed in any given sample for all DBP classes, if β is fixed. Because γ is expected to be independent of bromide concentration whereas β is proportional to bromide level, it is easier to see deviations from the simple binomial model by graphing the relationship between γ and bromide concentration. Figure 15 shows the first of these plots where the three sets of γ values for waters S1 and S6 were determined by the ratio of the most abundant THMs (i.e., TCM/BDCM), THAAs (i.e., TCAA/BDCAA) and DHAAs (i.e., DCAA/BCAA). The remaining four waters are shown in Figure 16 and Figure 17. Note that while each water has its unique relationship, the curves for the three DBP families for each water cluster tightly together. This offers visible evidence in support of our hypothesis that β is indeed nearly the same across these three classes of DBPs. On the other hand it shows that γ is not a constant across the full range of bromide levels. Certainly, the $\frac{[Br(+I)]}{[Cl(+I)]}$ ratio is dynamic over the course of a reaction and cannot be perfectly represented by a simple ratio of initial bromide to chlorine dose. Several researchers have pointed out that bromine undergoes internal cycling as it forms C-Br bonds, many of which undergo hydrolysis, liberating more bromide, which is oxidized again by free chlorine and re-forms more C-Br. Our approach essentially uses the measured [TCM]/[BDCM] ratio as an internal chemical probe and proxy for [Cl(+I)]/[Br(+I)].

Figure 15 to Figure 17 offer some additional information that has helped to guide the binomial modeling approach. First, the data based on THMs seem to best capture the

central tendency for each water, and these data seem to be relatively free from random excursions (compared to DHAA data). We attribute this to the simpler nature of THM analysis, and the expected higher level of precision. Each step in a chemical analysis can introduce new random errors. The analytical method for THAAs and DHAAs have one additional step (derivatization) beyond that required for THMs. The required laboratory operation may add to the overall random error. Second, there is a tendency for the γ-value to be larger at low bromide levels. This is probably an artifact of how $\frac{[Br(+I)]}{[Cl(+I)]}$ is estimated. At low bromide levels, the HOBr becomes exhausted much earlier in the reaction and the effective $\frac{[Br(+I)]}{[Cl(+I)]}$ ratio is substantially lower than the calculated one (which is based on initial bromide and chlorine dose). Thus, the γ value has to increase to compensate. This is another reason why we view β as a simpler and less biased estimator of the tendency to add bromine atoms. Finally, the most obvious feature of these data is that each of the six waters exhibits different average γ-values. These range from about 4.5 (water S1 & S2) to 17 (Water S5 & S6). Others have applied a similar model and have found best-fit γ-values of 10, 9, and 25 (Cowman and Singer, 1995; Nokes et al., 1999; and Chang et al., 2001). As already mentioned, some researchers have tried to improve the binomial model fit by adopting different γ-values for each 'node of halogen incorporation' (Nokes et al., 1999; Roccaro et al., 2014; Zhang et al., 2018). While this can lead to better agreement with the model, it also increases the number of γ-values that need to be evaluated. Not surprisingly, as the model incorporates more node-specific γ-values, there is a wider range of fitted values (from 5 to 130, depending on the node; Roccaro et al., 2014).

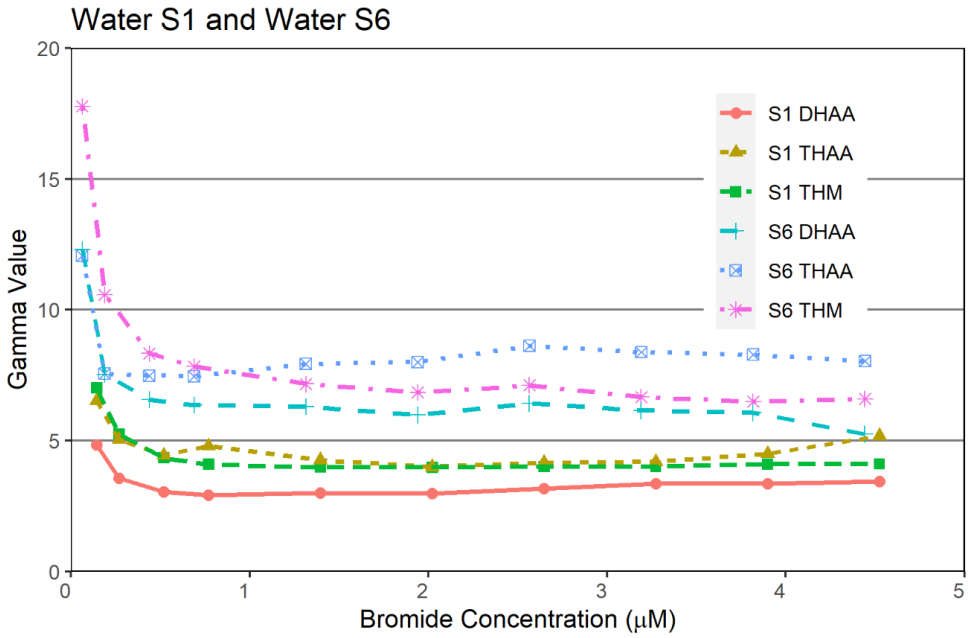



**Figure 15. γ-values based on each of the Three DBP Classes for Waters S1 & S6 versus Bromide Concentration.**

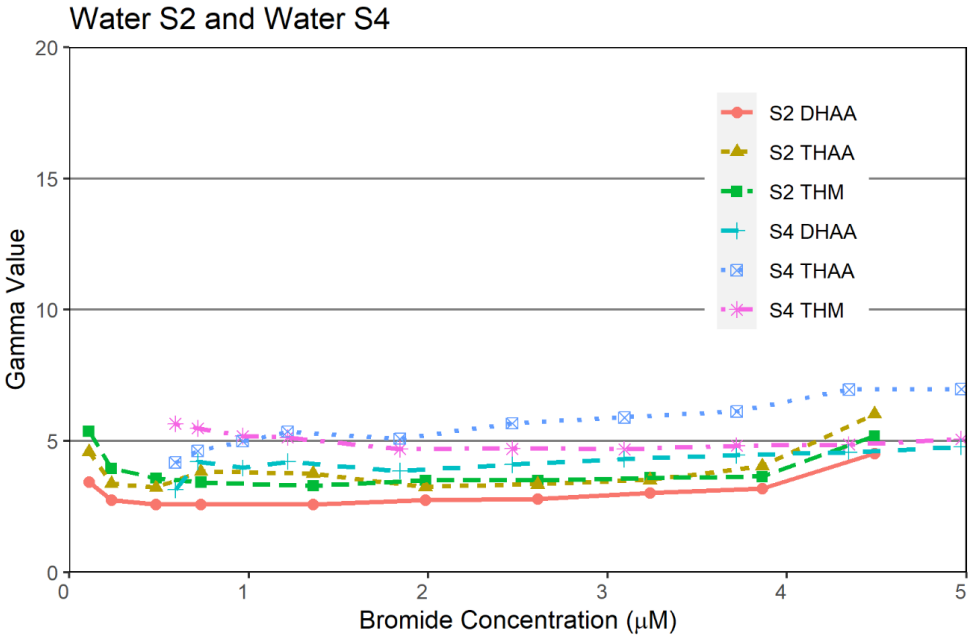



**Figure 16.** γ-values **based on each of the Three DBP Classes for Waters S2 & S4 versus Bromide Concentration.**

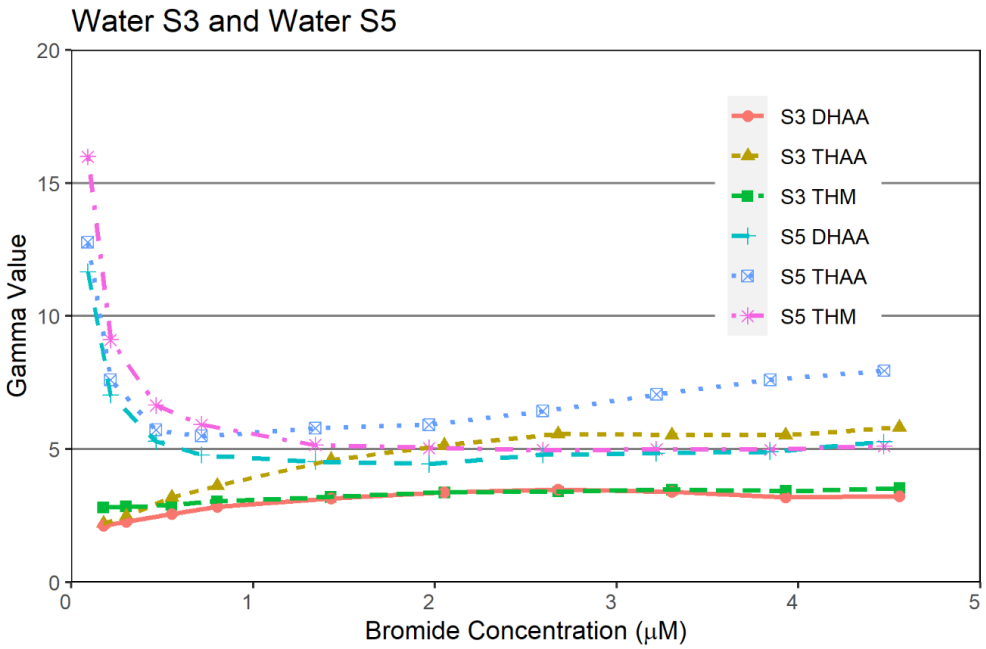



**Figure 17.** γ-values **based on each of the Three DBP Classes for Waters S3 & S5 versus Bromide Concentration.**

### 3.3.4 Adjustment for Model Error

As noted previously, there is a systematic error in the simple binomial model such that predictions of BDCM and BCAA are higher than observed (Figure 14). After carefully comparing our lab-generated DBP data with similar work from other researchers, we believe it is not unique to our dataset. We interpret this error to the assumption that reactive carbon centers show the same relative preference for bromine and chlorine addition regardless of the extent and nature of prior halogenation. For the purpose of simplifying the model calibration process, we decided to use an empirical value (E) to adjust the predicted concentrations for each DBP species. The revised predicted DBP concentrations are simply the product of the predicted values from the binomial model

and the empirical value (E) for the specific DBP species (Equation 25). A summary of the best-fit empirical values (E) is shown in Table 7.

$$Revised\ Con. = Predicted\ Conc. \times E \qquad \textit{(Equ. 25)}$$

**Table 7. Summary of Empirical Value (E) and Corresponding $r^2$ Values.**

| Target DBP | | | Empirical Value (E) | |
|---|---|---|---|---|
| **Abbreviation** | **Formula** | **Class** | **Value** | **$r^2$** |
| CDBM | $CHClBr_2$ | THM | 2.55 | 0.971 |
| TBM | $CHBr_3$ | THM | 4.86 | 0.965 |
| BCAA | CHClBrCOOH | HAA4 | 1.16 | 0.981 |
| DBAA | $CHBr_2COOH$ | HAA5 | 4.23 | 0.915 |
| BDCAA | $CCl_2BrCOOH$ | HAA4 | 1.11 | 0.977 |
| CDBAA | $CClBr_2COOH$ | HAA4 | 2.49 | 0.978 |
| TBAA | $CBr_3COOH$ | HAA4 | 3.78 | 0.948 |

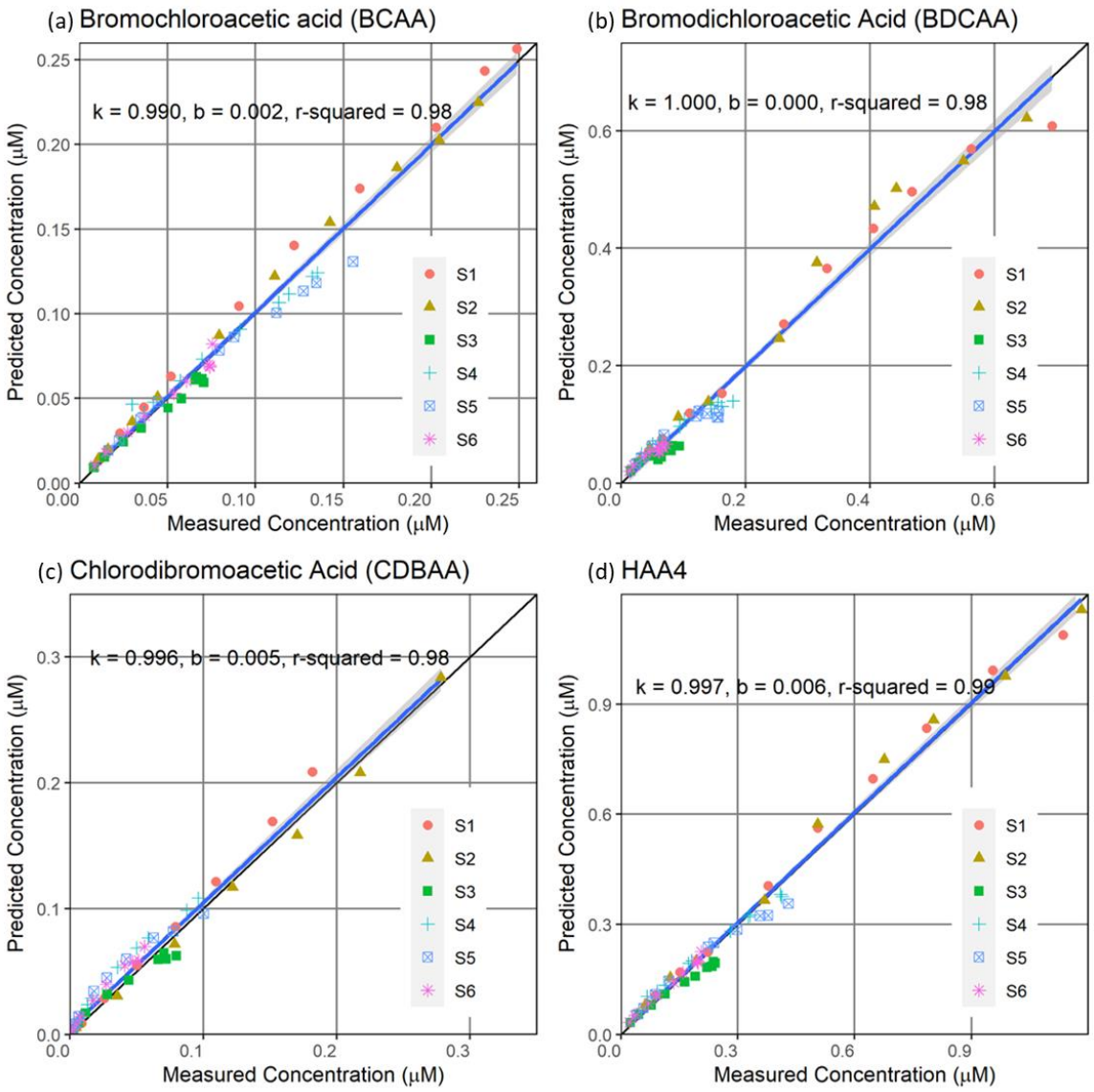



**Figure 18. Model Predictions vs. Measured Values for HAA Species: (a), BCAA; (b), BDCAA; (c), CDBAA; (d), HAA4.**

Figure 18 shows predicted versus measured concentrations for the three most abundant of the unregulated HAA4 species. The fourth (TBAA) is normally present at such low concentrations in many PWSs as to be of less interest. All three below show excellent agreement with the calibrated model predictions. As the extent of bromination

is established by the analogous THM species (BDCM), it's not surprising that BDCAA would be predicted with great accuracy. The model agreement for the unregulated HAA4 is probably of greatest interest, and our novel work provided a useful tool for such examination. Figure 18d shows that the adjusted model (by summing the HAA4 species in Table 7) does an excellent job of matching the measured values. An analysis of the absolute errors for each set of HAA4 measurements (Figure 19) indicates an average fixed error of 1.5 µg/L (intercept) plus a relative error of 2.9% (slope). This is quite low given that the analytical error alone for the HAA4 species has been documented in the range of 3-15% (EPA Method 552.2, Munch et al., 1995).

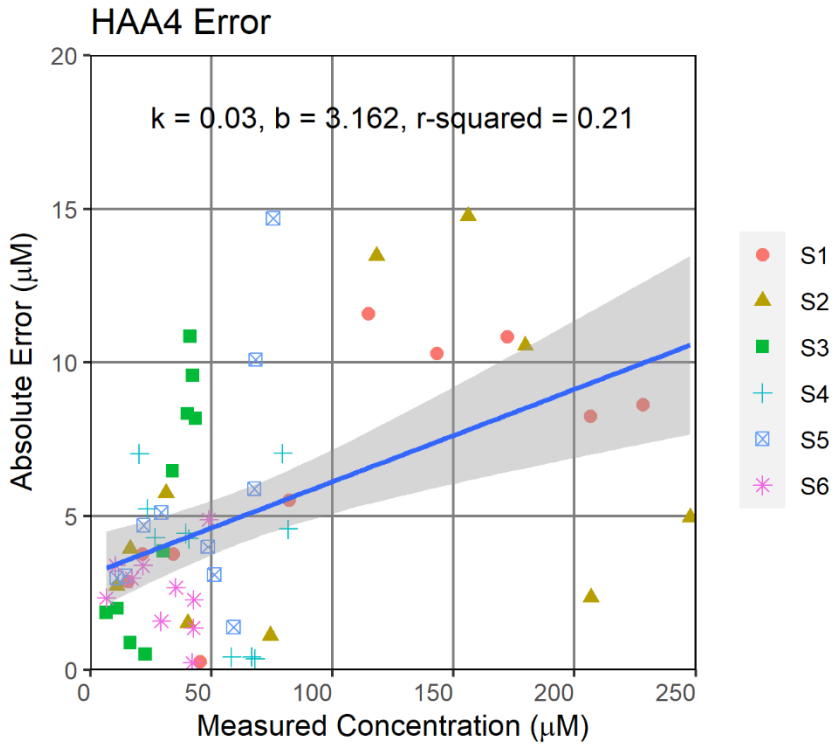



**Figure 19. Absolute Errors vs. Model Predictions for HAA4 in the Laboratory Dataset.**

### 3.3.5 Use of Model to Predict Missing DBPs

To make this model easier to use by utilities and other researchers, we have summarized the adjusted binomial model in a simpler form (Equations 26a-g). All of the intermediate model coefficients (i.e., β, γ, and x) were removed and replaced with the measured DBP concentrations upon which they are based. For the convenience of calculation, all DBP concentrations used in Equations 26a to 26g are in mass concentration units (µg/L). Accordingly, the predicted DBP species are also in mass concentration units (µg/L). A more detailed description of these final equations is presented in the Supporting Information (Table S1 and Table S2).

$$CDBM = 0.788 \cdot \left(\frac{BDCM \times BDCM}{TCM}\right) \quad \textit{(Equ. 26a)}$$

$$TBM = 0.146 \cdot \left(\frac{BDCM \times BDCM \times BDCM}{TCM \times TCM}\right) \quad \textit{(Equ. 26b)}$$

$$BCAA = 0.757 \cdot \left(\frac{DCAA \times BDCM}{TCM}\right) \quad \textit{(Equ. 26c)}$$

$$DBAA = 0.423 \cdot \left(\frac{DCAA \times BDCM \times BDCM}{TCM \times TCM}\right) \quad \textit{(Equ. 26d)}$$

$$BDCAA = 1.029 \cdot \left(\frac{TCAA \times BDCM}{TCM}\right) \quad \textit{(Equ. 26e)}$$

$$CDBAA = 0.680 \cdot \left(\frac{TCAA \times BDCM \times BDCM}{TCM \times TCM}\right) \quad \textit{(Equ. 26f)}$$

$$TBAA = 0.098 \cdot \left(\frac{TCAA \times BDCM \times BDCM \times BDCM}{TCM \times TCM \times TCM}\right) \quad \textit{(Equ. 26g)}$$

Of these seven predictive equations, three are for regulated DBPs (i.e., CDBM, TBM and DBAA), and as such they are regularly measured in US chlorinated public water supplies for compliance purposes. For these three, the equations above could be used as a check on laboratory data, should there be questions about a particular sample. Also they offer the opportunity to estimate concentrations when a result is lost or below detection limit. In

# CHAPTER 4

# APPLICATION AND FURTHER VERIFICATION OF THE BINOMIAL MODEL: UTILITY DATA STUDIES

## 4.1 Introduction

In the Chapter 3, the binomial model was developed and verified by bench-scale laboratory-generated data. The ultimate goal of the research is the application of the model to the collected DBP data from public water suppliers (PWSs) to provide insights into unregulated DBPs. The information would either be used as contaminant level quality assurance protocol or provide useful information for the development of future regulations.

Unlike the experimental conditions in the laboratory, where the reaction conditions were well-controlled, actual drinking water distribution systems have more complicated and complex conditions. The changes in temperature, pH, and other conditions could substantially affect the formation and, the speciation of the DBP. Additional chlorine dose and chlorine residual may increase the formation of DBP with the same reaction time, however, it may not change the speciation of HAA and THM. Both volatilization and biodegradation could also occur in the distribution systems and increase the uncertainty of DBP formation. Tung and Xie (2009) have evaluated the occurrence of the HAA biodegradation in water distribution systems and similar evidence were also reported by other researchers (Summers et al., 1996; Rossman et al, 2001; Hozalski et al., 2008). THM is unknown for high volatility and related studies have been widely conducted in different areas (Trussell et al, 1979; Golfinopoulos et al, 2001; Nikolaou et al, 2002; Rivera-Núñez et al, 2012). To better evaluate the adaptability of the binomial

model in the more complicated distribution systems, we further applied the binomial model to the collected DBP data from drinking water treatment utilities. The prediction of HAA concentration in distribution systems were was mostly evaluated in this chapter. The utility DBP database used in this chapter wascompiled by researchers from the University of Colorado, Boulder (Samson et al., 2017). 266 public water suppliers (PWSs) from the US participated in the DBP data collection process and both utility information and DBP concentrations were collected. After the application of the binomial model on the utility DBP database, observable random and systematic errors were noticed. In this chapter, we will further discuss the potential causes of the errors and evaluation of the binomial model prediction performance with utility DBP data.

## 4.2 Model Application on Utility DBP Data

The data collection was funded under the contract with AWWA's Water Industry Technical Action Fund (AWWA WITAF) and conducted by Samson et al. (2017). The utility DBP database was then often referred as the WITAF database. The utility information and the DBP data were collected directly from individual PWS and was very similar to the previous occurrence efforts including the Information Collection Rule (ICR, 1997-1998, n = 500 PWSs; McGuire and Graziano, 2002), the AWWA Research Foundation Trihalomethane Survey (AwwaRF, 1984-1986, n = 727 PWSs; McGuire and Meadow, 1988), and the National Organics Monitoring Survey (1975-1976, n = 113 PWSs; Brass et al., 1977). DBP data, includes all THM and HAA data, were collected from 266 PWSs with at least 100,000 population served over the period of 1997 to 2014. All DBP sample collections were conducted at pre-designated sampling locations across

each distribution system. We recognized that chemical reactions initiated by addition of chlorine wereone of a complex set of factors that affect the concentration and speciation of DBPs in treatment plants and in distribution systems. Nevertheless, we decided to use DBP concentrations from the WITAF database as a means of field-testing the binomial model.

### 4.3 Studied Utility and DBP Data Description

As the binomial model was only verified with chlorine disinfection, without any additional advanced oxidation process (e.g., ferrate, chlorine dioxide, and ozone), we then limited the study utility to those PWSs only applied chlorine disinfections as the primary disinfectant. We also noticed the unregulated DBP data were very limited in the WITAF database. Based on the requirement of the binomial model, only PWS that reported valid (1) TCM, (2) BDCM, (3) DCAA, and (4) TCAA were selected for the study. Furthermore, at least one of the four unregulated DBP species (i.e., BCAA, BDCAA, CDBAA, and TBAA) should also be reported to be used for prediction verification. Unfortunately, among all 266 PWS, only two facilities were qualified for these the requirements (Table 8).

**Table 8. Studied Utilities Profile**

| PWS No. | State | PWS Type | Water Source | Population Served | Recorded Period | Record Counts |
|---|---|---|---|---|---|---|
| 1 | Maryland | Community | Surface Water | 1,800,000 | 4/2/1999 - 10/21/2014 | 2866 |
| 2 | Missouri | Community | Surface Water | 175,000 | 2/25/1997 - 9/23/2014 | 375 |

A very limited amount of unregulated DBP concentrations were reported. PWS #1 reported concentrations of all nine regulated DBP species (i.e., 4 THMs and 5 HAAs) and one unregulated DBP species (BCAA). PWS #2 also reported all nine regulated DBP species and one unregulated DBP species (BDCAA). BCAA concentration data from PWS #1 are used to further verify the adaptability of the binomial model on BCAA formation prediction. Accordingly PWS #2 is used to evaluate the performance of the model on BDCAA formation prediction. The summary and the distribution of the DBP concentrations from two PWSs are shown in Table 9.

**Table 9. Distribution of DBP Concentrations in Two Studied PWSs.**

| | DBP Species | Concentration (µg/L) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Min. | 1st Quarter | 2nd Quarter | Mean | 3rd Quarter | Max. |
| **PWS #1** | | | | | | | |
| THM | TCM | 0.0 | 18.0 | 28.7 | 32.3 | 43.6 | 150.0 |
| | BDCM | 1.1 | 7.3 | 9.9 | 11.2 | 13.9 | 72.4 |
| | CDBM | 0.0 | 1.7 | 2.5 | 3.2 | 3.9 | 33.0 |
| | TBM | 0.0 | 0.0 | 0.0 | 0.1 | 0.2 | 27.0 |
| HAA | DCAA | 0.3 | 8.6 | 12.1 | 13.3 | 16.9 | 42.1 |
| | BCAA | 0.9 | 2.7 | 3.6 | 4.1 | 4.7 | 308.0 |
| | DBAA | 0.0 | 0.0 | 0.4 | 0.9 | 0.8 | 812.0 |
| **PWS #2** | | | | | | | |
| THM | TCM | 2.3 | 9.7 | 14.4 | 15.5 | 20.5 | 43.3 |
| | BDCM | 2.6 | 6.9 | 8.9 | 8.9 | 10.8 | 16.4 |
| | CDBM | 1.2 | 3.2 | 4.0 | 4.1 | 5.0 | 7.8 |
| | TBM | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 1.0 |
| HAA | TCAA | 1.0 | 5.4 | 7.6 | 8.9 | 10.2 | 113.0 |
| | BDCAA | 0.0 | 3.6 | 4.7 | 5.0 | 5.8 | 58.9 |

### 4.4 BCAA Prediction

Equation 26c from the Chapter 4 was applied to DBP concentration data collected from PWS #1 for the predictions of BCAA. Predicted and measured BCAA

concentrations are plotted for demonstration (Figure 20). Predicted BCAA concentrations show highly similarity to the measurements ($R^2$=0.94). Unlike the previous comparison between the predictions and laboratory measurements, the random errors are significantly increased. However, no significant systemic error was detected (slope (k) = 0.90). The close agreement between predictions and measurements provides the evidence for the application of the binomial model to specific utility DBP prediction. We recognize the random error was largely caused by (1) the less accurate detections from the volatility of THMs and (2) the unequal biodegradation rate between BCAA and DCAA.

$$BCAA = 0.757 \cdot \left(\frac{DCAA \times BDCM}{TCM}\right) \qquad \textit{(Equ. 26c)}$$

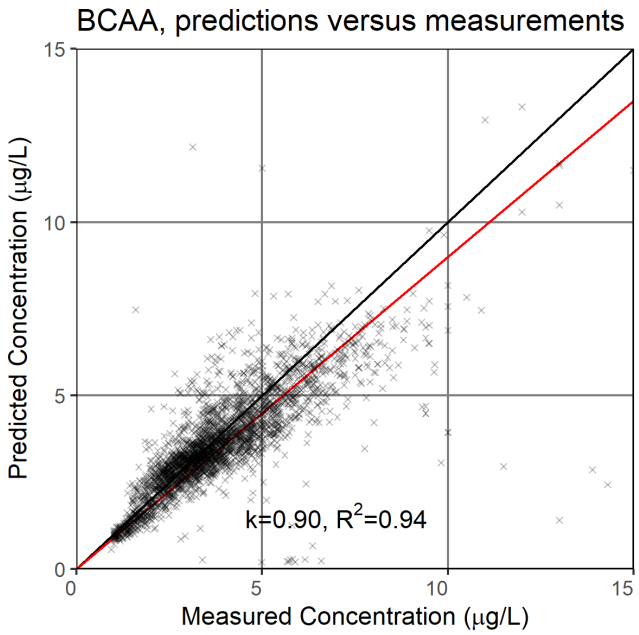



**Figure 20. Prediction of BCAA by the Binomial Model on Utility DBP data.**

In contrast to typical drinking water and storage facility conditions, under well-controlled reaction conditions, such as bench-scale DBP formation experiments, minimal

volatilization in THM would occur. The development of the binomial predictive model also did not take THM volatilization into consideration. The model prediction largely depends on the accurate measurement of two THM species: TCM and BDCM. Even though it is hard to estimate the loss in these two THM species by the volatilization, given the complicated reaction conditions in distributions, but it is reasonable to believe that the molar or mass ratio between TCM and BDCM concentrations are altered. This would certainly increase the level of uncertainty in BCAA prediction.

Another issue often being addressed is the biodegradation of HAA species, especially of DHAA species. Baribeau et al. (2005) reported that the inconsistent biodegradation rate between DHAA species followed the order of DCAA>BCAA>DBAA. The binomial model used measured DCAA concentrations as the proxy for BCAA concentration prediction and the biodegraded DCAA would later result in the underestimation of BCAA, which is also illustrated in Figure 20 as the slightly decreased slope (k=0.9).

One of the many challenges for the model application on utility DBP data is the variation in water quality parameters (e.g., temperature, pH, and TOC). In a relative well-maintained PWS, such as PWS #1, one good parameter that summarizes most of the water quality parameters is the season. Seasonal variation is noticed in both drinking water source and distribution systems that chemical reaction conditions (i.e., temperature and organic matter content) mostly show a repeating trend regardless of the relative minor changes between years (Baytak et al., 2008; Poudel et al., 2013; Kostyla et al., 2015). The increased activities of microorganisms in warmer months would also increase the biodegradation rate of DHAA species and result in seasonal variation.

To further evaluate the performance of the binomial model with seasonal variance, the BCAA prediction results are separated by four seasons by the sample collection data (Figure 21). Spring is defined from March 19 to June 20. Summer is defined from June 21 to September 22. Fall is defined as September 23 to December in 21. And winter is defined from December 22 to March 18. Linear regression relationships are developed between the predictions and the measurements for the quantitative evaluations of the performance of the binomial model. The seasonal regression coefficients for BCAA predictions are shown in Table 10.

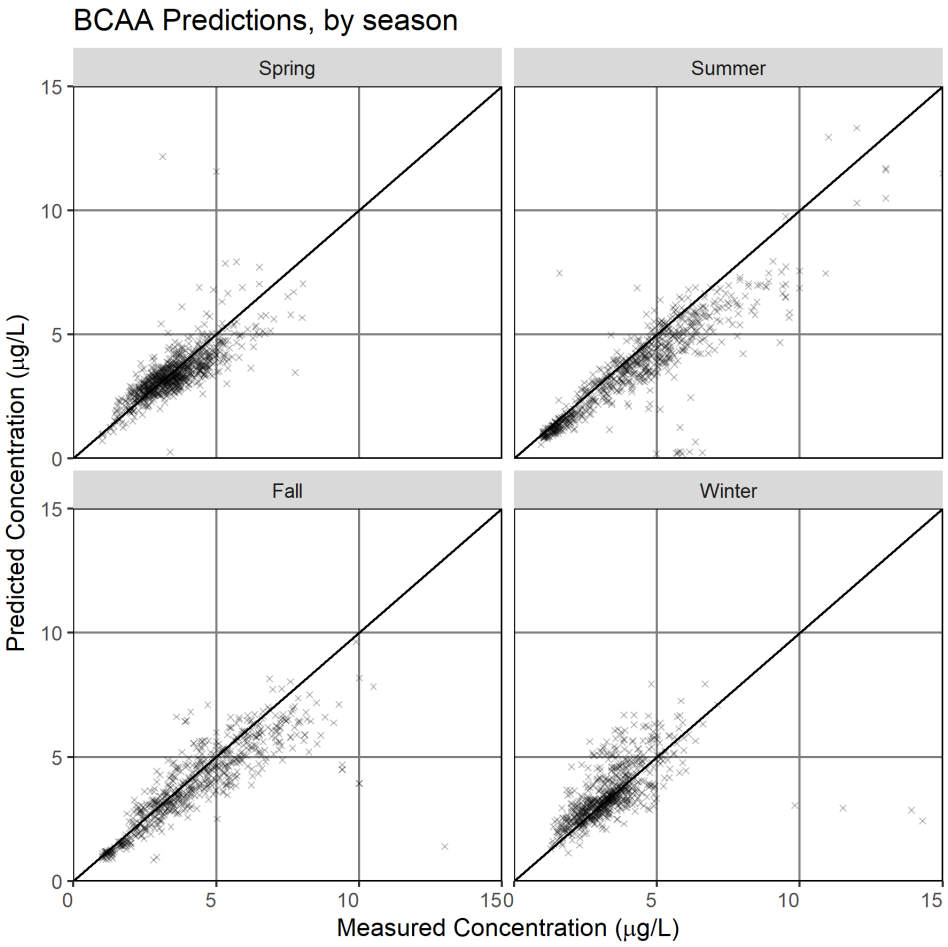


**Figure 21. BCAA Predictions by Seasons.**

**Table 10. BCAA Model Coefficients.**

| Season | Record Counts (n) | Slope (k) | Error ($R^2$) | Relative Error (%)[13] |
|---|---|---|---|---|
| Spring | 817 | 0.95 | 0.96 | 12.7 |
| Summer | 728 | 0.86 | 0.95 | 16.1 |
| Fall | 624 | 0.90 | 0.97 | 13.2 |
| Winter | 688 | 1.05 | 0.96 | 13.5 |

Despite the change in slope (k) due to seasonality, the binomial model works almost equally well in all four seasons with all $R^2 > 0.95$. The change in slope (k) is also evident to our previous assumption that seasonal variations have impacts on model performance. With the highest temperature in summer, the activities of microorganisms also peaked and resulted in higher loss of DCAA. The higher loss of DCAA is then reflected in the lowest slope (k = 0.86, Table 10), where the model underestimated the BCAA concentration. With moderate temperature in spring and fall, the slope also increased accordingly. In the coldest months, when the activity of microorganisms is at its minimal, we overserved the highest slope between model predictions and actual measurements (slope (k) = 1.05).

The binomial model does not take seasonal variation into consideration, however, it is logical for the model users (i.e., utility workers and research engineers) to create additional 'patches' to the model to improve the model adaptability in different seasons. For example, an additional coefficient should be applied to increase the summer BCAA

[13] Relative errors are calculated based on the mean and the error from each season.

predictions and to decrease the winter predictions. This practical coefficient only aims to make the prediction closer to measurement and should be determined individually by PWSs. Given the variance of seasonal temperature change by the different location of PWS, we could not recommend a set of values to address the seasonal variance, but one can simply compare the predictions and measurements to calculate the slope and adjust for the seasonal pattern by multiplying an additional coefficient.

Admittedly, there are other variables that could add more uncertainties to the prediction of BCAA, but after applying the binomial model to utility DBP data, we are confident to state the model works as well as expected.

### 4.5 BDCAA Prediction

Very similar to the approach we adopted for BCAA prediction, we applied the Equation 26f to the DBP data collected from PWS #2 to predict the formation of BDCAA. Due the lack of CDBAA and TBAA data from the utility database, only BDCAA can be used for model verification. The comparison between the predictions and the measurements is illustrated in Figure 22.

$$CDBAA = 0.680 \cdot \left(\frac{TCAA \times BDCM \times BDCM}{TCM \times TCM}\right) \quad \text{(Equ. 26f)}$$

Again, we notice close agreement between the predicted and the measured BDCAA concentrations (slope (k) = 1.02, $R^2$ = 0.96). Except for the negligible systematic error, the random errors are also to be less observable compared with BCAA prediction. THAA species are also reported to undertake certain level of biodegradation, however, the

biodegradation rate is lower than DHAAs (Xie and Zhou, 2002; Zhang et al., 2009). We believe the lower biodegradation rate in THAAs reduced the level of random errors in the binomial model application on BDCAA. The impact of biodegradation of THAAs on the model performance is similar to the impact on DHAAs, thus will not be further discussed.

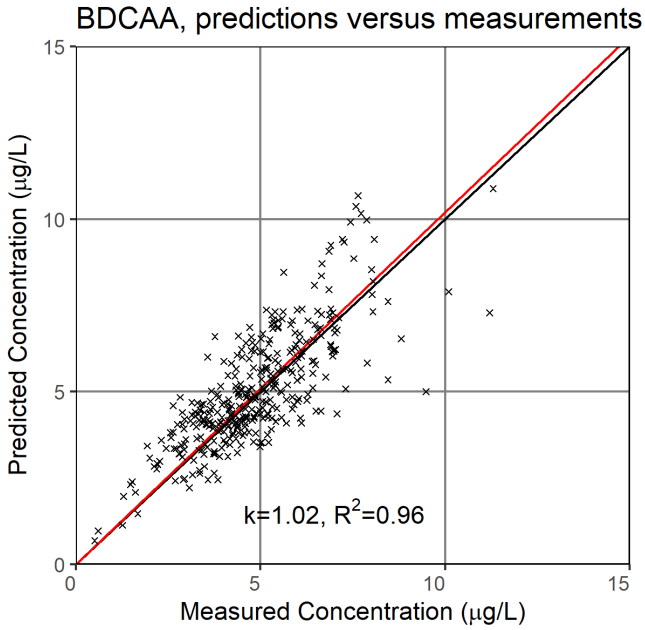



**Figure 22. Prediction of BDCAA by the binomial model on utility DBP data.**

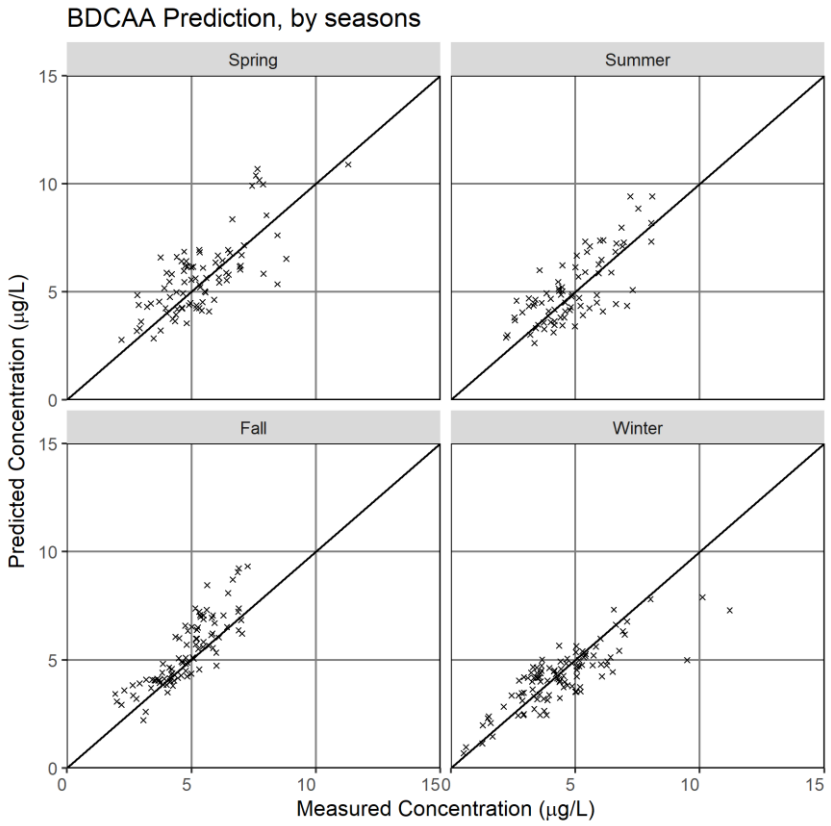



**Figure 23. BDCAA Predictions by Seasons.**

We also attempted to separate the BDCAA predictions by season to evaluate the impact of seasonal variation on model performance (Figure 23). Seasonal patterns are detected in BDCAA predictions, however, the patterns are inconsistent. Given the limited amount of collected BDCAA concentration data, we do not believe any definitive conclusions can be drawn. Regardless of the 'tight' correlation between the predicted and the measured BDCAA concentration ($R^2 > 0.95$, relative error = 13-16%), we encourage the actions of collecting more reliable BDCAA data to better evaluate the performance of the binomial model in different seasons.

492

493

494

# CHAPTER 5

# ASSOCIATION OF UNREGULATED DISINFECTION BYPRODUCT EXPOSURE WITH ADVERSE BIRTH OUTCOMES IN OHIO

## 5.1 Introduction

Compared with regulated HAA5, recently researchers have been focusing more on the four unregulated HAAs (i.e., HAA4; Bromochloroacetic acid (BCAA), bromodichloroacetic acid (BDCAA), chlorodibromoacetic acid (CDBAA), and tribromoacetic acid (TBAA)) due to their higher carcinogenicity potential, especially in bladder and colon (Villanueva et al., 2017; Rahman et al., 2010; Evlampidou et al., 2020; Atwood et al., 2019; Plewa et al., 2004; Muellner et al., 2007). Several studies also focused on potential adverse reproductive outcomes (e.g., reduce in birth weight, increase in birth defect, etc.) associated with regulated DBPs (Hoffman et al, 2008; Rivera-Núñez and Wright, 2013). Epidemiologic evidence suggests small but fairly consistent associations between DBP exposures (e.g., HAA5 and THM4) and small for gestational age (SGA) via meta-analyses (Grellier et al., 2010; Summerhayes et al., 2020). Few studies have been conducted on the unregulated HAAs (i.e., HAA4) due to the lack of reliable HAA4 from distribution system. Horton et al. (2011) examined the associations of HAA9 and brominated DBPs with SGA (Brominated DBPs vs. Term-SGA, aORs = 1.27, 95% CI: 0.93-1.75) and pre-term birth (PTB; Brominated DBP vs. PTB, aORs = 1.05, 95% CI: 0.77-1.45), but lacked data to examine the unregulated HAAs.

Kramer et al (1992) reported an association between THM4 concentrations and low birth weight (LBW), PTB, and SGA in Iowa. Bove et al (1995) carried out a retrospective cohort study in New Jersey and reported the mean birthweight (BWT) in high THM

519 exposed group (THM4>100 µg/L) was 70 g lower than low exposure group. Gallagher et
520 al (1998) reported higher risks of LBW in term births (OR 5.9, 95% CI 2.0 to 17.0) with
521 higher THM4 exposure (THM4 > 60 µg/L). Several other retrospective cohort studies
522 conducted on Massachusetts residents have reported associations between DBP exposure
523 and BWT, SGA, and PTB (Wright et al., 2004; Lewis et al., 2006). Rivera-Núñez and
524 Wright (2013) reported the increased risk of PTB between HAA exposures but showed
525 limited evidence for SGA.

526 One concern in DBP related epidemiologic research is that aggregate DBP metrics
527 such as HAA5 and THM4 are often used as the proxies of exposures and fail to capture
528 all of the most toxicologically relevant DBPs (Kaufman et al, 2020). This is also related
529 to the speciation of DBP differences between studies but also across systems in the same
530 study, in that even with the same aggregate DBP concentration, concentration of each
531 DBP species within the same class may vary (Hua et al., 2006). Depending on what the
532 DBP or DBP mixture is that presents the largest risk to adverse effects, if proxy measures
533 are less relevant this can reduce the sensitivity of a study to detect associations.

534 In this study, we applied the HAA predictive model (Ma and Reckhow, in review) to
535 the collected regulated DBP data and estimated the exposure level of unregulated HAAs.
536 We mainly focused on the impact of aggregate and individual unregulated HAA species
537 on potential adverse birth outcomes (e.g., SGA, PTB, and BWT). In addition to the
538 commonly used DBP exposure metrics in this field (such as THM4, HAA5), we also
539 attempted to explore the application of other metrics (such as DBP9, DBP13 and HAA-
540 Br, the sum of mass concentration of all brominated HAAs) in exposure risk estimations.
541 We have suspect that most of the mass concentration based metrics in the DBP-based

epidemiology studies may have insufficient aspects in describing the toxicity of DBP. This study discusses the advantage and limitations of each exposure metrics with actual examples and suggested the application of toxicity-based exposure metrics for accurate exposure risks evaluation.

## 5.2 Materials and Methods

### 5.2.1 Study Population and Outcomes

The Ohio Department of Health's (ODH) Center for Public Health Statistics and Informatics (CPHSI) provided individual-level birth records data extracted from birth certificates on 1,174,411 live infants born between 2006 and 2013 in the state of Ohio. For our analysis of fetal growth restriction and gestational duration, the inclusion criteria was the following[14]: (1), singleton births (excluded n = 17,237, 1.5%); (2), gestational age (GA) between 22 and 45 weeks (excluded n = 13,753, 1.2%); (3), infant birth weight (BWT) > 200g (excluded n = 15,797, 1.3%); (4), sex of the infant is not missing (excluded n = 802, 0.1%), and (5), infants that had no congenital anomalies recorded on the birth certificates (excluded n = 4,099, 0.3%). This resulted in 1,133,755 (96.5%) available birth records.

Gestational age was calculated from obstetric estimate derived from the birth certificates. Small for gestational age (SGA) was defined as infants with a BWT below the fifth and tenth percentiles for each gestation age stratified by sex and maternal

12. Record count (n) in parentheses is the count of records that do not meet each restriction. Note that there was overlap between each restriction and there are not mutually exclusive. For example, birth records with birth weight < 200 g, usually have higher chance that missed sex information too.

562 race/ethnic categories (Table 11). Due to insufficient number of births across some
563 gestational ages, this precluded stratification for certain race and ethnicity groups. We
564 combined limited data on Hispanic black and Hispanic white infants together in the SGA
565 analysis to allow for more participants to be included under the homogeneity assumption
566 as they tend to have more favorable developmental outcomes (Ribble and Keddie, 2001).
567 Mean BWT was also checked by sex and race/ethnicity as the confirmation.

568 Preterm birth (PTB, or preterm delivery) were defined as infants born prior to week
569 37. The comparison group used in the PTB analysis was selected as records of at least 37
570 gestational weeks and birth weight of at least 2500 g. We restricted the mean BWT
571 analysis to term births with gestational age from week 37 to 45.

572 The SGA analysis was restricted to term births reporting non-missing race/ethnicity
573 information with gestational ages from week 37 to 42 for all maternal race/ethnicity
574 categories except for Asian Indians and Native Americans, which was limited to weeks
575 37 to 41 due to the sparse birth numbers (Table 11). Pre-term birth were analyzed
576 separately from term births, (since birth weight is dependent on both the rate of fetal
577 growth rate (FGR) and gestational duration, and perturbations in each may arise from
578 different etiologies (Grantz et al., 2018). There is additional value in replicating the
579 restriction on gestational age for the BWT analysis to provide comparable results with
580 previous DBP studies.

581 In theory, SGA classifications based on smaller percentages will have less outcome
582 misclassification and be more likely to capture pathologically small infants. Therefore,
583 we calculated SGA based on the lowest 5 percentile or 10 percentile of BWT for each
584 stratification factor including: infant sex, maternal race/ethnicity, and gestational age.

585 **Table 11. Selection and distribution of gestational ages for each sex and**
586 **race/ethnicity[15].**

| No. | Maternal Race/Ethnicity | Sex of Infant | Total, Count | 5%tile SGA, Count | 10%tile SGA, Count |
|---|---|---|---|---|---|
| **1** | White | Male | 315,731 | 15,661 | 31,467 |
| | | Female | 303,356 | 15,132 | 30,069 |
| **2** | Hispanic | Male | 20,789 | 1,036 | 2,078 |
| | | Female | 20,381 | 1,021 | 2,018 |
| **3** | Black | Male | 78,066 | 3,879 | 7,778 |
| | | Female | 76,321 | 3,780 | 7,606 |
| **4** | Native American | Male | 2,098 | 105 | 208 |
| | | Female | 2,051 | 104 | 206 |
| **5** | Asian Indian | Male | 3,301 | 166 | 331 |
| | | Female | 3,196 | 162 | 318 |
| **6** | Other Asian | Male | 9,028 | 449 | 904 |
| | | Female | 8,452 | 424 | 845 |
| **7** | Others | Male | 4,403 | 222 | 441 |
| | | Female | 4,269 | 216 | 427 |
| | **Total** | Male | 433,416 | 21,518 | 43,207 |
| | | Female | 418,026 | 20,839 | 41,489 |
| | | Total | 851,442 | 42,357 | 84,696 |

587

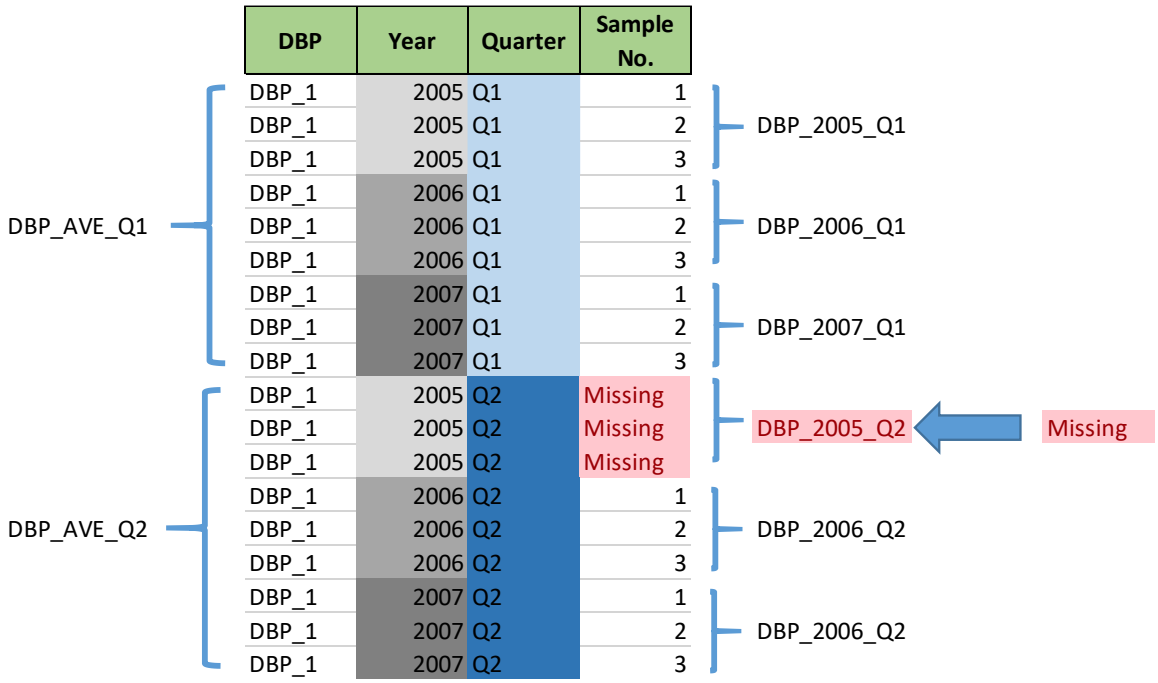
[15] SGA classification is based on weeks 37-42, except for Native American and Asian Indian, which were limited to weeks 37-41.

### 5.2.2 Exposure Data and Exposure Assessment

We acquired routinely-collected DBP exposure data from the Ohio Environmental Protection Agency (OH EPA). The fourth Unregulated Contaminant Monitoring Rule database from the US Environmental Protection Agency (US EPA) was used to identify the service zip codes by each public water system (PWS). Given that the available data were limited to four THMs but only five regulated HAA species (i.e., THM4 and HAA5), the unregulated HAA species (i.e., BCAA, BDCAA, CDBAA, and TBAA) were calculated by the kinetic binomial predictive model. Importantly, given the demonstrated variability and challenges related to DBP mixtures, we examined different individual DBPs as well as various DBP mixture measures (e.g. DBP13; HAA9; THMBr; HAABr; HAA4).

The sampling event is defined as sample collection at one location at one time. Most PWSs collect samples from multiple sampling locations during one sampling cycle (i.e., weekly, monthly, quarterly, or yearly). A total number of 42,993 DBP exposure records from 1,204 PWSs were collected from the years 2005-2013. The collection frequency is regulated that each PWS should collect samples quarterly. Still, we noticed this regulation was not completely conducted until approximately 2012. Multiple sampling events of different sampling sites within one PWS during one sampling cycle could happen and thus multiple exposure data was collected. The quarterly DBP average for each year were calculated from the records within the same quarter of each year. All records within the same PWS in the same quarter (regardless of the year) were then averaged. We used an imputation approach (Figure 24) to address missing quarterly exposure data. This involved replacing the missing quarter with the average exposure

concentration calculated based on the concentrations of the same quarter from the rest of the years with available data. For instance, for PWS_1, during the year of 2005, Q1, Q3, and Q4 averaged DBP exposure were successfully calculated from the exposure records, while one missing data occurred (i.e., Q2). Then this missing record was replaced with the averaged same quarterly data from all the other available years. The same method was applied to all missing records. If one PWS did not have any samples for a particular quarter, no data would be imputed for that quarter. DBP exposure concentrations of 0 were assigned to residents in zip codes served by private wells. The identification of private well users will be discussed later in this section.



| PWSID | Year | Quarter | DBP_1 Concentration | Average |
|---|---|---|---|---|
| PWS_1 | 2005 | 1 | DBP_2005_Q1 | DBP_AVE_Q1 |
| PWS_1 | 2005 | 2 | Missing | DBP_AVE_Q2 |
| PWS_1 | 2005 | 3 | DBP_2005_Q3 | DBP_AVE_Q3 |
| PWS_1 | 2005 | 4 | DBP_2005_Q4 | DBP_AVE_Q4 |
| PWS_1 | 2006 | 1 | DBP_2006_Q1 | DBP_AVE_Q1 |
| PWS_1 | 2006 | 2 | DBP_2006_Q2 | DBP_AVE_Q2 |
| PWS_1 | 2006 | 3 | Missing | DBP_AVE_Q3 |
| PWS_1 | 2006 | 4 | DBP_2006_Q4 | DBP_AVE_Q4 |

| PWSID | Year | Quarter | DBP_1 Concentration | Average |
|---|---|---|---|---|
| PWS_1 | 2005 | 1 | DBP_2005_Q1 | DBP_AVE_Q1 |
| PWS_1 | 2005 | 2 | DBP_AVE_Q2 | DBP_AVE_Q2 |
| PWS_1 | 2005 | 3 | DBP_2005_Q3 | DBP_AVE_Q3 |
| PWS_1 | 2005 | 4 | DBP_2005_Q4 | DBP_AVE_Q4 |
| PWS_1 | 2006 | 1 | DBP_2006_Q1 | DBP_AVE_Q1 |
| PWS_1 | 2006 | 2 | DBP_2006_Q2 | DBP_AVE_Q2 |
| PWS_1 | 2006 | 3 | DBP_AVE_Q3 | DBP_AVE_Q3 |
| PWS_1 | 2006 | 4 | DBP_2006_Q4 | DBP_AVE_Q4 |

**Figure 24. A Brief Illustration of the DBP Exposure Imputation Process.**

The method detection limits (MDLs) varied by analytical method and processing laboratories and over the 9-year sampling period (Table 12). DBP samples below detection limit (BDL) were replaced with a random value between 0 and the method detection limit (MDL) (Cohen and Ryan, 1989). Sensitivity analyses were also conducted on the primary birth outcomes (BWT, SGA, and PTB) for missing and BDL exposure data that were imputed (Table S3 and Table **S4** here need to add to SI).

**Table 12. Reported Method Detection Limits (MDL) from the Collected Exposure Records.**

| DBP Species | | Reported MDL (µg/L) |
|---|---|---|
| HAA | MCAA | 1.0-2.0 |
| | MBAA | 1.0-2.0 |
| | DCAA | 1.0-2.0 |
| | DBAA | 1.0-2.0 |
| | TCAA | 1.0-2.0 |
| THM | TCM | 0.5-1.0 |
| | BDCM | 1.0-2.0 |
| | CDBM | 0.5-1.0 |
| | TBM | 0.5-1.0 |

The quarterly exposure data was then linked to each birth records by zip code of the place of residence at birth. 952,190 out of 1,133,755 birth records (94.0%) were assigned with exposure data. The trimester-specific DBP exposures were calculated from the date

of birth based on the timing of quarterly averaged DBP data with the weighted exposure averages calculated proportionally for multiple quarters that overlapped (Figure 25).

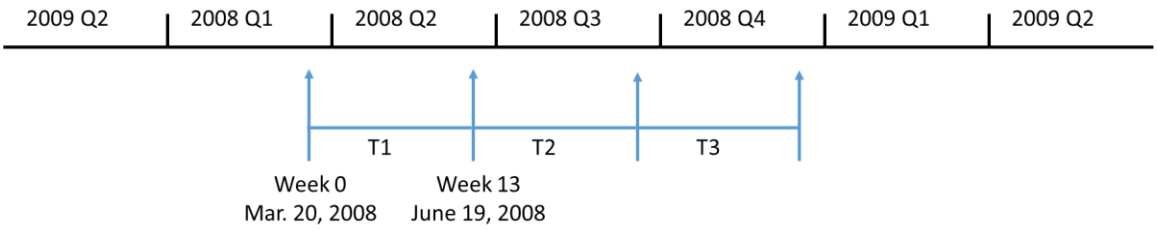



**Figure 25. Illustration of Trimester Exposure Weighted Averaging Process.**

For example, a birth record with the pregnancy on March 20 , 2008, will have the first 11 days of the first trimester (T1) assigned with quarterly DBP exposure of the 1st quarter of 2008 (DBP_2008_Q1), the remaining 80 days of the T1 will be assigned with quarterly DBP exposure of the 2nd quarter of 2008 (DBP_2008_Q2). Each trimester is set at 91 days (or 13 weeks) manually. Then, the weighted averaged DBP exposure of the T1 exposure will be: ("11/91) × DBP_2008_Q1 + (80/91) × DBP_2008_Q2".

The summary of THM4 and HAA5 exposure data before and after imputation is shown in Table 13. Participants served by private wells were included in the summary (i.e. were assigned DBP concentrations of 0). To contrast the extremes of the exposure distributions, we examined the lowest and highest 10% of each exposure as the referent and high category, respectively. The remaining participants were split evenly into quantiles (i.e. tertiles or quartiles)or  depending on the extent of the DBP distributions.

**Table 13. Summary of the Distribution of Selected DBP Species in (a), Non-imputed Records and (b), Imputed Records.**

**(a), Non-Imputed Records**

| Trimester | DBP | Count | DBP Exposure Concentration (µg/L) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Mean | Min. | 10% | 25% | 50% | 75% | 90% | Max. |
| T1 | THM4 | 744478 | 39.4 | 0.0 | 20.8 | 27.6 | 36.0 | 48.4 | 63.6 | 405.1 |
| | HAA5 | 744478 | 22.4 | 0.0 | 7.8 | 11.9 | 21.7 | 29.5 | 38.0 | 167.6 |
| T2 | THM4 | 753781 | 39.3 | 0.0 | 20.6 | 27.5 | 35.9 | 48.5 | 63.5 | 407.1 |
| | HAA5 | 753781 | 22.4 | 0.0 | 7.7 | 11.8 | 21.7 | 29.6 | 37.8 | 159.3 |
| T3 | THM4 | 766653 | 40.1 | 0.0 | 20.0 | 26.1 | 35.8 | 50.3 | 66.8 | 416.3 |
| | HAA5 | 766653 | 22.8 | 0.0 | 7.3 | 11.7 | 21.5 | 30.0 | 39.2 | 164.0 |

**(b), All Records after Imputation**

| Trimester | DBP | Count | DBP Exposure Concentration (µg/L) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Mean | Min. | 10% | 25% | 50% | 75% | 90% | Max. |
| T1 | THM4 | 952190 | 36.7 | 0.0 | 16.7 | 24.1 | 33.8 | 45.9 | 61.2 | 405.1 |
| | HAA5 | 952190 | 20.5 | 0.0 | 6.3 | 9.5 | 19.5 | 28.5 | 36.9 | 167.6 |
| T2 | THM4 | 952190 | 36.6 | 0.0 | 16.6 | 24.0 | 33.8 | 46.0 | 61.1 | 407.1 |
| | HAA5 | 952190 | 20.5 | 0.0 | 6.3 | 9.5 | 19.6 | 28.7 | 36.9 | 159.3 |
| T3 | THM4 | 952190 | 37.4 | 0.0 | 15.5 | 23.6 | 33.5 | 47.2 | 64.3 | 416.3 |
| | HAA5 | 952190 | 21.0 | 0.0 | 6.2 | 9.4 | 19.4 | 29.5 | 38.4 | 164.0 |

In this study, we used ZIP code as the geographic identifier to link birth records to DBP exposure data. As part of the exposure assessment (i.e., assign DBP exposure data to birth records), the study population was categorized into several groups (Table 14). The study population (n = 1,133,755 births) covered 3,713 ZIP codes. 936,915 births (82.6%) were identified to be served by one unique PWS, and 15,275 births (1.3%) were served by private wells. The identification of private well users were based on the ZIP code of the maternal residency and the private well files provided by Ohio Department of Natural Resources. The above 952,190 births (84.0%) were included for further analysis. For the remaining 181,565 births (16.0%), they were either (1) served by multiple PWS and thus the exposure cannot be determined, or (2) resided out of the State of Ohio during

669 pregnancy, or (3) assigned with one unique PWS but no DBP exposure was found during
670 pregnancy.

671 **Table 14. Water Designation for the Study Population.**

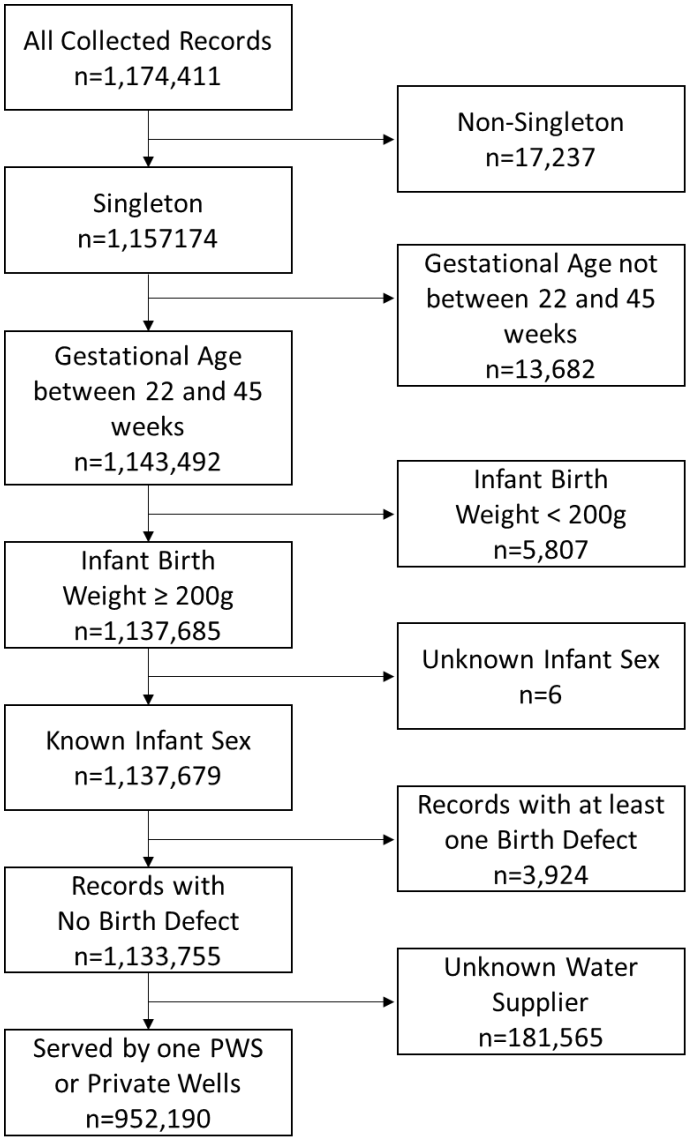
| Category | Definition | Birth Count | (%) | Decision | Birth Count | (%) |
|---|---|---|---|---|---|---|
| **1** | Private well | 15,275 | 1.3 | Included | 952,190 | 84.0 |
| **2** | Unique PWS[1] | 936,915 | 82.6 | | | |
| **3** | Unique PWS, no exposure data | 101,845 | 9.0 | Excluded | 181,565 | 16.0 |
| **4** | No PWS assigned/out-state | 33,645 | 3.0 | | | |
| **5** | Multiple PWS | 46,075 | 4.1 | | | |
| **Total** | | 1,133,755 | 100 | | 1,133,755 | 100 |

[1] PWS, Public water system.

672

673 The inclusion and exclusion criteria for the study population is presented in a flow
674 chart (Figure 26). As the BWT and SGA analyses required additional restriction, namely,
675 on gestational age and race/ethnicity, the study populations for these two endpoints are
676 shown in Figures 27 and 28. ).

677




678

679 **Figure 26. Study population flow chart of all birth for the PTB analysis (n=952,190)**

680

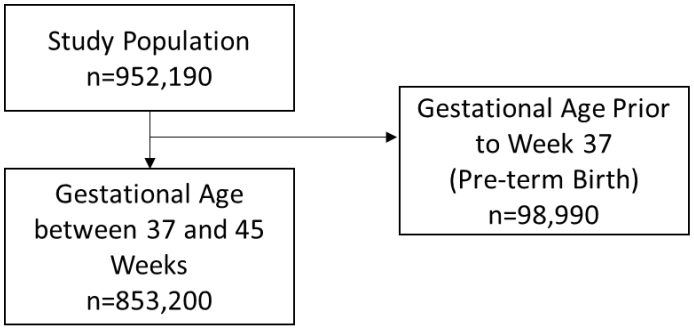



**Figure 27. BWT study population determination flow chart.**

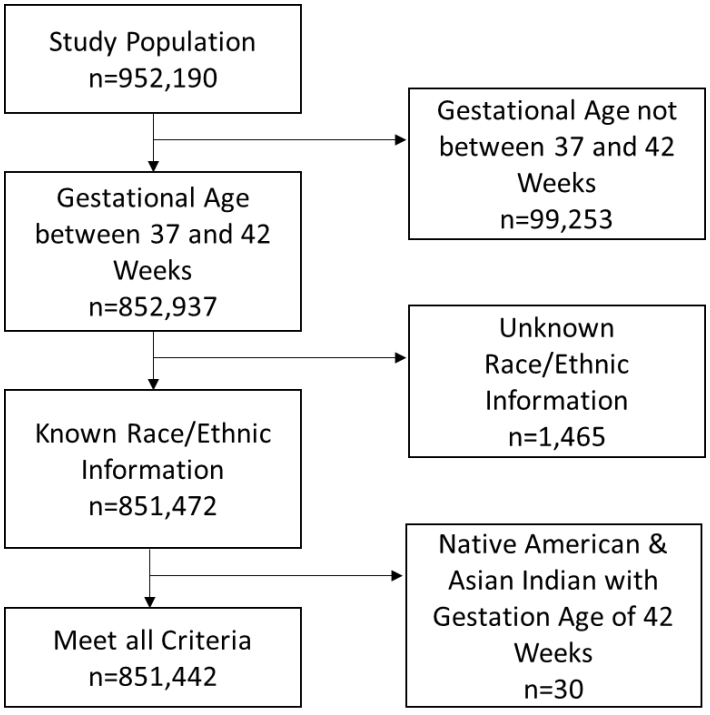



**Figure 28. SGA study population determination flow chart.**

### 5.2.3 Statistical Analysis

R (Version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria) and RStudio (Version 1.2.1335, RStudio, Inc., Boston, MA) were used for the statistical

analyses. Births in the lowest DBP exposure category were used as the referent for comparison with the upper categories. Due to the low or null exposure from the water, all the participants who were served by private wells and residents exposed to BDL levels of DBPs are included in the lowest category. Spearman correlation coefficients were used to compare individual DBP species and DBP classes (i.e., DBP13; DBP-Br, THM4, THM-Br, HAA5, HAA-Br, HAA9). Logistic regression was used to estimate ORs and 95% CIs for the associations between DBP exposures and SGA and PTB. Linear regression was used to evaluate the change in mean BWT relative to DBP exposure categories. Statistical significance was based on $\alpha = 0.05$ or less.

#### 5.2.3.1 Confounders Determination

Potential confounders were identified based on biological plausibility and previous literature including our earlier publications including a DBP meta-analysis and individual epidemiologic studies of FGR (Wright et al., 2004; Rivera-Núñez and Wright, 2013; Summerhayes et al., 2020). Potential confounders were factors known to be associated with birth outcomes and DBP exposures but were not intermediates on the causal pathway between DBPs and the adverse developmental endpoints. Most of the covariates were available directly from the collected original birth records. Household income data was acquired indirectly from the 2010 US Census (Geolytics, Inc., East Brunswick, NJ) and was assigned to birth records by maternal residential ZIP codes. Adequacy of prenatal care was evaluated as the Kessner Index, which was calculated indirectly and based on the start date of the first prenatal care visit and the number of total visits until delivery. The following confounders were noticed to have the most missing data:

maternal weight gain (8.6%), interpregnancy interval (8.4%), Kessner Index (6.7%), and maternal pre-pregnancy BMI (4.8%) and parity (2.7%).

Thirteen covariates were adjusted for in the multivariate models. The adjustment set for the BWT and PTB regression models included: maternal age, race/ethnicity, education, smoking, parity, adequacy of prenatal care, delivery source of payment, income, marital status, pre-pregnancy BMI, weight gain during pregnancy, season, and interpregnancy interval. The SGA models were adjusted for the same set of covariates with the exception of maternal race/ethnicity given it was used to define stratified SGA cut points. Modeled covariates were converted into categorical variables adjustment: maternal age (≤20, 21-25, 26-30, 31-35, 36-40, and >40), maternal race/ethnicity (non-Hispanic white, Hispanic, non-Hispanic black, Native American, Asian Indian, Asian, and Others), maternal education (below high school/GED, high school graduate/GED, bachelors (no degree), bachelor (graduate), and graduate or higher), smoking (no smoking, 1-5, 6-10, and >10 cigarette per day), parity (categorical: 0, 1, and ≥2), adequacy of prenatal care (adequate, intermediate, and inadequate), delivery source of payment (public, private, and others), income (≤32,494, >32,494-37,967, >32,967-43,768, >43,768-51,806, >51,806-66,989, and >66,989, USD), marital status (married and unmarried), maternal weight gain (<0, 0-25, >25-50, and >50), pre-pregnancy BMI ( ≤18, >18-24, >24-30, and >30), season (spring, summer, fall, and winter), and interpregnancy interval (0-1, >1-2, >2-3, >3-4, >4-5, and >5 years).

#### 5.2.3.2 Change in Estimate Analysis

To identify and gauge potential magnitude of key confounders, a change in estimate (CIE) analysis was employed for mean BWT in relation to third trimester THM4 and HAA5 exposures based on bivariate models. As part of the CIE analysis, we evaluated the change of linear regression beta coefficient prior to and after adjusting for each individual covariate compared to the beta from the univariate model. The modeling result between confounders selected via CIE analysis and the full adjustment set used in the primary analyses were also compared.

The CIE between univariate model ($\beta_1$, Equation 27) and adjusted bivariate model ($\beta_1$', Equation 28) was evaluated for each potential confounder. Any potential confounders result in larger than 10% change in CIE were identified as a stronger confounder for adjustment. This allowed for a sensitivity analysis for different adjustment sets.

$$Birth\ Outcome = \beta_0 + \beta_1 \times DBP\ Exposure \qquad \textit{(Eqn. 27)}$$

$$Birth\ Outcome = \beta_0' + \beta_1' \times DBP\ Exposure + \beta_2 \times Potential\ Confounder \qquad \textit{(Eqn.28)}$$

Exposure data were categorized into six categories by five cutpoints: 10, 30, 50, 70, and 90 percentile. For illustration purposes, we only examined the CIE analysis for the exposure of THM4 and HAA5 for the highest (>90%tile) exposure groups relative to the referent (e.g., lowest <10%tile).

The following covariates were identified by the CIE 10% analysis as confounders: (1) maternal race; (2) education; (3) delivery source of payment; (4) income; and (5) marital status in both THM4 and HAA5 mean BWT models (Table 15). Thus, for comparison as

shown in Table S5, we compared adjustment for these 5 covariates with the full adjustment set for the primary analysis.

**Table 15. Linear Regression Coefficient Estimations with Potential Confounder Adjustment [1].**

| | | Coefficient | | |
|---|---|---|---|---|
| Potential Confounder | Exposure | β, unadjusted | β, adjusted | CIE (%)[1] |
| Maternal Age | THM4 | -14.1 | -13.2 | 6.3 |
| Maternal Race | | | -10.2 | 27.7 |
| Maternal Education | | | -10.5 | 25.5 |
| Smoking | | | -11.3 | 19.8 |
| Parity | | | -13.3 | 5.6 |
| Adequacy of Prenatal Care [2] | | | -15.1 | 7.5 |
| Delivery Source of Payment | | | -10.4 | 26.2 |
| Income | | | -6.5 | 54.0 |
| Marital Status | | | -9.7 | 31.2 |
| Season | | | -15.4 | 9.7 |
| Maternal Weight Gain | | | -11.9 | 15.9 |
| Interpregnancy Interval | | | -12.2 | 13.1 |
| Pre-pregnancy Body Mass Index | | | -13.6 | 3.5 |

**Table 15. (continued) Linear Regression Coefficient Estimations with Potential Confounder Adjustment.**

| | | Coefficient | | |
|---|---|---|---|---|
| Potential Confounder | Exposure | β, unadjusted | β, adjusted | CIE (%)[2] |
| Maternal Age | HAA5 | -42.0 | -43.2 | 3.0 |
| Maternal Race | | | -12.6 | 70.1 |
| Maternal Education | | | -16.0 | 61.9 |
| Smoking | | | -45.8 | 9.1 |
| Parity | | | -40.3 | 4.0 |
| Adequacy of Prenatal Care [3] | | | -36.3 | 13.4 |
| Delivery Source of Payment | | | -36.0 | 14.3 |
| Income | | | -29.6 | 29.5 |
| Marital Status | | | -32.1 | 23.5 |
| Season | | | -40.1 | 4.5 |
| Maternal Weight Gain | | | -40.0 | 4.8 |
| Interpregnancy Interval | | | -38.0 | 9.4 |

| | | |
|---|---|---|
| Pre-pregnancy Body Mass Index | -40.6 | 3.3 |

[1]Linear regressions were based on third-trimester THM4 or HAA5 exposure and term birth BWT in the Ohio from 2006-2013.
[2] Coefficient estimations with a difference less than 10% are highlighted.
[3] Kessner Index was used as measure of adequacy of prenatal care.

760

761 **5.3 Results**

762 Among the study population (n = 952,190), 98,990 (10.4%) births were classified as
763 PTB. Among 851,442 term births, 42,357 (5.0%) and 84,696 (10.0%) births were
764 identified as SGA, based on the selection of SGA classification approach (5 percentile
765 and 10 percentile, respectively).

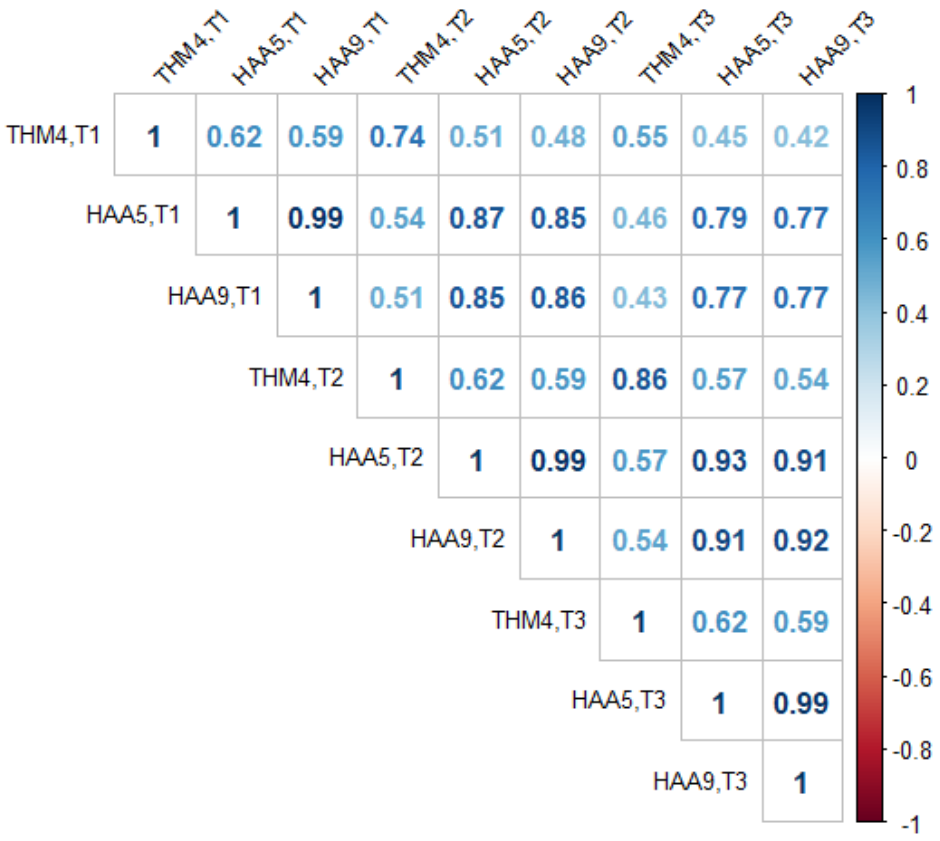
766 Mean BWT reductions were detected among the 853,200 term births records across
767 various maternal characteristics. The largest mean BWT reductions were found for
768 mothers who were unmarried (-148 g, vs. married mothers), smoking during pregnancy (-
769 240 g, vs. non-smoking mothers), had low pre-pregnancy BMI (-227 g, vs. normal BMI),
770 and inadequate prenatal care (-140 g, vs. adequate prenatal care) (Table 16).

771 Distribution of second-trimester (T2) and third-trimester (T3) DBP exposure is shown
772 in Table 17. Mean concentrations of THM4 exposure in T2 and T3 were 36.6 and 37.4
773 µg/L, respectively. Mean concentrations of HAA5 exposure in T2 and T3 were 20.5 and
774 21.0 µg/L, respectively. High Spearman correlation coefficients (>0.85) were detected
775 across adjacent trimesters for the following DBPs: HAA5 and DCAA, HAA5 and TCAA,
776 THM4 and TCM, HAA5 and HAA9. For example, across T2 and T3, high Spearman
777 coefficients (>0.85) were seen between the following DBP species: THM4 (T2) and

778 THM4 (T3), HAA5 (T2) and HAA5 (T3), HAA5 (T2) and HAA9 (T3), HAA9 (T2) and
779 HAA5 (T3), HAA9 (T2) and HAA9 (T3).

780 Table 3 shows moderate to strong correlations across each trimester within each DBP
781 class. For example, the HAA5 exposure correlation coefficients between T1-T2, T1-T3,
782 and T2-T3 are only 0.87, 0.79, and 0.93, respectively (Figure 29). The smaller correlation
783 (over windows T1 vs. T3) are likely indicative of the temporal variation demonstrated
784 during pregnancy.

785




786

787 **Figure 29. Correlation between Selected DBP Classes in all Three Trimesters.**

788

789

790

791

792

793

**Table 16. Maternal and Infant Characteristics of Study Population (n=952,190 [1]) in the state of Ohio, 2006-2013.**

| | Study Population, n (%) | BWT, g | 5% SGA, % [2] | 10% SGA, % [2] | PTB, % [3] |
|---|---|---|---|---|---|
| **Count, n** | 952,190 (100.0) | 853,200 | 851,442 | 851,442 | 952,190 |
| **Total Birth** | 952,190 (100.0) | 3,369 | 5.0 | 9.9 | 10.4 |
| **Infant Sex** | | | | | |
| Male | 486,578 (51.1) | 3,432 | 5.0 | 10.0 | 10.7 |
| Female | 465,612 (48.9) | 3,303 | 5.0 | 9.9 | 10.0 |
| **Maternal Age (Years)** | | | | | |
| ≤20 | 142,144 (14.9) | 3,267 | 6.7 | 13.2 | 10.8 |
| 21-25 | 253,433 (26.6) | 3,327 | 5.6 | 11.1 | 10.1 |
| 26-30 | 278,287 (29.2) | 3,401 | 4.4 | 8.9 | 9.9 |
| 31-35 | 192,344 (20.2) | 3,431 | 4.0 | 8.1 | 10.4 |
| 36-40 | 73,629 (7.7) | 3,427 | 4.3 | 8.4 | 12.1 |
| 41-62 | 12,329 (1.3) | 3,402 | 5.0 | 9.1 | 14.1 |
| Missing | 24 (0.0) | 3,290 | 10.0 | 10.0 | 41.7 |
| **Maternal Race** | | | | | |
| White | 684,836 (71.9) | 3,411 | 5.0 | 9.9 | 9.6 |
| Hispanic | 45,518 (4.8) | 3,355 | 5.0 | 9.9 | 9.5 |
| Black | 179,599 (18.9) | 3,225 | 5.0 | 10.0 | 14.0 |
| Native American | 4,774 (0.5) | 3,323 | 5.0 | 10.0 | 12.7 |
| Asian Indian | 7,182 (0.8) | 3,200 | 5.0 | 10.0 | 9.4 |
| Asian | 18,993 (2.0) | 3,292 | 5.0 | 10.0 | 8.0 |
| Others | 9,601 (1.0) | 3,313 | 5.1 | 10.0 | 9.6 |
| Missing | 1,687 (0.2) | 3,341 | 0.0 | 0.0 | 13.1 |
| **Maternal Education** | | | | | |
| Below high school/GED | 154,462 (16.2) | 3,255 | 7.5 | 14.2 | 12.0 |
| High school graduate/GED | 242,709 (25.5) | 3,326 | 6.0 | 11.6 | 10.9 |
| Some college, no degree | 212,922 (22.4) | 3,376 | 4.6 | 9.2 | 10.5 |
| Bachelor/associate degree | 241,185 (25.3) | 3,447 | 3.4 | 7.3 | 9.0 |

| | | | | | |
|---|---|---|---|---|---|
| Graduate or higher | 94,451 (9.9) | 3,446 | 3.4 | 7.2 | 9.2 |
| Missing | 6,461 (0.7) | 3,307 | 5.7 | 10.6 | 15.7 |
| **Marital Status** | | | | | |
| Married | 527,649 (55.4) | 3,434 | 3.8 | 7.9 | 9.3 |
| Unmarried | 423,421 (44.5) | 3,286 | 6.5 | 12.5 | 11.7 |
| Missing | 1,120 (0.1) | 3,309 | 6.3 | 12.3 | 12.1 |
| **Number of Previous Births** | | | | | |
| 0 | 308,743 (32.4) | 3,340 | 6.1 | 12.0 | 9.5 |
| 1 | 257,497 (27.0) | 3,389 | 4.3 | 8.8 | 9.3 |
| ≥2 | 360,226 (37.8) | 3,382 | 4.4 | 8.9 | 11.6 |
| Missing | 25,724 (2.7) | 3,332 | 5.8 | 10.9 | 15.3 |
| **Pay Source of Delivery** | | | | | |
| Public | 396,848 (41.7) | 3,292 | 6.4 | 12.4 | 11.6 |
| Private | 505,852 (53.1) | 3,428 | 3.9 | 8.0 | 9.4 |
| Others | 49,490 (5.2) | 3,367 | 5.1 | 10.1 | 10.8 |

**(continued)**

**Table 16. (Continued)**

| | Study Population, n (%) | BWT, g | 5% SGA, % [2] | 10% SGA, % [2] | PTB, % [3] |
|---|---|---|---|---|---|
| **Maternal Weight Gain [4]** | | | | | |
| ≤0 | 21,492 (2.3) | 3,265 | 7.0 | 13.6 | 14.3 |
| >0-25 | 270,136 (28.4) | 3,279 | 6.7 | 12.9 | 13.4 |
| >25-50 | 480,434 (50.5) | 3,392 | 4.3 | 9.0 | 8.3 |
| >50 | 97,933 (10.3) | 3,511 | 3.2 | 6.4 | 8.9 |
| Missing | 82,195 (8.6) | 3,370 | 5.1 | 9.9 | 13.8 |
| **Income [5]** | | | | | |
| ≤32,494 | 207,432 (21.8) | 3,278 | 6.2 | 11.9 | 12.5 |
| >32,494-37,967 | 192,554 (20.2) | 3,358 | 5.4 | 10.7 | 10.6 |
| >37,967-43,768 | 186,969 (19.6) | 3,383 | 4.9 | 9.9 | 9.8 |
| >43,768-51,806 | 176,374 (18.5) | 3,407 | 4.4 | 9.1 | 9.5 |
| >51,806-66,989 | 135,812 (14.3) | 3,426 | 3.9 | 8.1 | 9.3 |
| >66,989 | 50,706 (5.3) | 3,436 | 3.5 | 7.5 | 9.2 |
| Missing | 2,343 (0.2) | 3,387 | 5.6 | 10.3 | 9.3 |
| **Season [6]** | | | | | |
| Spring | 238,067 (25.0) | 3,376 | 4.8 | 9.6 | 10.5 |
| Summer | 250,912 (26.4) | 3,370 | 4.9 | 9.9 | 10.4 |
| Fall | 239,581 (25.2) | 3,367 | 5.2 | 10.2 | 10.2 |
| Winter | 223630 (23.5) | 3,363 | 5.0 | 10.0 | 10.6 |
| **Prenatal Care [7]** | | | | | |
| Adequate | 552,335 (58) | 3,399 | 4.4 | 9.0 | 9.2 |
| Intermediate | 306,640 (32.2) | 3,334 | 5.7 | 11.2 | 10.1 |
| Inadequate | 29,873 (3.1) | 3,259 | 7.8 | 13.9 | 18.5 |
| Missing | 63,342 (6.7) | 3,318 | 5.5 | 11.0 | 18.3 |
| **Pre-pregnancy BMI** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| ≤18 | 26391 (2.8) | 3,142 | 11.0 | 20.5 | 13.4 |
| >18-24 | 382736 (40.2) | 3,325 | 5.6 | 11.3 | 9.9 |
| >24-30 | 282861 (29.7) | 3,398 | 4.3 | 8.7 | 9.8 |
| >30 | 214285 (22.5) | 3,438 | 4.0 | 7.9 | 11.3 |
| Missing | 45917 (4.8) | 3,364 | 5.2 | 10.1 | 12.9 |
| **Maternal Smoking [8]** | | | | | |
| 0 | 777760 (81.7) | 3,403 | 3.9 | 8.2 | 9.9 |
| 1-5 | 55268 (5.8) | 3,265 | 7.6 | 14.3 | 12.1 |
| 6-10 | 62390 (6.6) | 3,196 | 10.6 | 19.1 | 12.3 |
| >10 | 50881 (5.3) | 3,163 | 12.2 | 21.9 | 12.6 |
| Missing | 5891 (0.6) | 3,281 | 6.0 | 11.9 | 16.6 |

BWT, birth weight. SGA, small for gestational age. PTB, pre-term birth.
[1]. Study population was restricted to singleton birth, gestational age between 22 to 45 weeks, BWT > 200 g, known infant sex, and served by either one PWS or private well.
[2]. SGA was defined as infants with a BWT below the 5 or 10 percentile for their gestation age at birth. The BWT cut point was specific to the combination of infant sex, maternal race/ethnicity, and gestational age.
[3]. PTB was defined as infants born at a gestational age < 37 weeks.
[4]. Maternal weight gain during pregnancy (lb).
[5]. Median household income of the maternal residency (US Dollars). From 2010 US Census.
[6]. Season was defined as the season when the delivery occurred (i.e., Spring, Summer, Fall, and Winter).
[7]. Adequacy of prenatal care was evaluated based on the Kessner Index.
[8]. Maternal smoking during pregnancy was categorized as cigarette smoking per day reported during the three trimesters, when applicable.

| Table 17. Maternal Third and Second Trimester Exposures to DBPs (µg/L). [1] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **DBPs** | **Mean** | **Min.** | **10%** | **25%** | **50%** | **75%** | **90%** | **Max.** |
| **Third Trimester** | | | | | | | | |
| THM4 [2] | 37.4 | 0.0 | 15.5 | 23.6 | 33.5 | 47.2 | 64.3 | 416.3 |
| TCM | 21.1 | 0.0 | 3.0 | 7.2 | 16.5 | 30.5 | 45.9 | 383.8 |
| BDCM | 8.9 | 0.0 | 3.8 | 6.3 | 8.5 | 11.0 | 14.3 | 54.1 |
| THM-Br [3] | 16.3 | 0.0 | 7.1 | 10.8 | 14.5 | 19.7 | 28.0 | 176.9 |
| HAA5 [4] | 21.0 | 0.0 | 6.2 | 9.4 | 19.4 | 29.5 | 38.4 | 164.0 |
| DCAA | 10.3 | 0.0 | 1.3 | 3.1 | 9.1 | 14.9 | 21.1 | 87.6 |
| BCAA | 2.4 | 0.0 | 1.0 | 1.5 | 2.2 | 2.8 | 3.6 | 38.4 |
| TCAA | 6.5 | 0.0 | 0.5 | 1.2 | 4.8 | 10.0 | 15.1 | 88.2 |
| BDCAA | 2.5 | 0.0 | 0.7 | 1.3 | 2.3 | 3.5 | 4.5 | 33.0 |
| CDBAA | 0.5 | 0.0 | 0.1 | 0.2 | 0.4 | 0.7 | 1.0 | 42.8 |
| TBAA | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 9.5 |
| HAA-Br [5] | 7.3 | 0.0 | 3.7 | 5.1 | 6.7 | 8.7 | 11.0 | 77.7 |
| HAA4 [6] | 5.5 | 0.0 | 2.3 | 3.4 | 4.9 | 7.0 | 9.0 | 76.1 |
| HAA9 [7] | 26.5 | 0.0 | 9.4 | 13.4 | 24.7 | 35.3 | 45.5 | 196.9 |
| DBP9 [8] | 58.4 | 0.0 | 24.0 | 37.2 | 53.4 | 75.9 | 97.2 | 475.0 |
| DBP13 [9] | 63.9 | 0.0 | 27.2 | 42.1 | 59.4 | 82.1 | 105.2 | 477.4 |
| **Second Trimester** | | | | | | | | |
| THM4 [2] | 36.6 | 0.0 | 16.6 | 24.0 | 33.8 | 46.0 | 61.1 | 407.1 |
| TCM | 20.6 | 0.0 | 3.2 | 7.2 | 16.8 | 29.9 | 43.5 | 375.4 |
| BDCM | 8.8 | 0.0 | 4.2 | 6.5 | 8.4 | 10.7 | 13.6 | 53.4 |
| THM-Br [3] | 16.0 | 0.0 | 7.7 | 10.9 | 14.2 | 19.1 | 27.1 | 173.0 |
| HAA5 [4] | 20.5 | 0.0 | 6.3 | 9.5 | 19.6 | 28.7 | 36.9 | 159.3 |
| DCAA | 10.1 | 0.0 | 1.5 | 3.0 | 9.3 | 14.6 | 20.1 | 85.7 |
| BCAA | 2.3 | 0.0 | 1.0 | 1.6 | 2.1 | 2.7 | 3.6 | 37.5 |
| TCAA | 6.4 | 0.0 | 0.5 | 1.2 | 4.9 | 9.9 | 14.5 | 85.9 |
| BDCAA | 2.5 | 0.0 | 0.8 | 1.3 | 2.3 | 3.4 | 4.3 | 32.3 |
| CDBAA | 0.5 | 0.0 | 0.1 | 0.2 | 0.4 | 0.6 | 1.0 | 40.5 |
| TBAA | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 9.2 |
| HAA-Br [5] | 7.1 | 0.0 | 3.8 | 5.1 | 6.6 | 8.7 | 10.9 | 74.1 |
| HAA4 [6] | 5.4 | 0.0 | 2.4 | 3.4 | 5.0 | 6.9 | 8.6 | 72.0 |
| HAA9 [7] | 25.9 | 0.0 | 9.4 | 13.8 | 25 | 34.6 | 43.8 | 191.2 |
| DBP9 [8] | 57.1 | 0.0 | 25.0 | 37.8 | 53.8 | 74.0 | 93.0 | 464.6 |
| DBP13 [9] | 62.5 | 0.0 | 28.3 | 42.9 | 59.7 | 79.8 | 99.6 | 466.9 |

[1]. Only records served by private wells or one identified PWS, regardless of gestation age or other restrictions, were included (n = 952,190).
[2]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[3]. THM-Br, the summation of BDCM, CDBM, and TBM.
[4]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[5]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[6]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[7]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

[8]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[9]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

### 5.3.1 Birth Weight Reduction

Statistically significant reductions in mean BWT were observed in third-trimester THM4 exposure categories in the unadjusted models (Table 18). Based on univariate models, BWT was 14 to 27 g lower in upper 5 THM4 exposure categories compared with the lowest exposure category (THM4 < 15 µg/L). After adjusting for all covariates, the THM4 results were null.

The largest reductions in the unadjusted mean BWT models were observed with DBP13 (range: -29 to -39 g), DBP9 (range: -27 to -39 g), HAA5 (range: -25 to -42 g), HAA9 (range: -23 g to -37 g), and DCAA (range: -23 to -50 g) exposures, compared to the referents (i.e., lowest exposure categories). Exposure-response relationships based on the univariate models were observed across HAA5 and DCAA categories with lower BWT in higher exposed groups.

Following statistical adjustment, the associations were attenuated and only a few of the DBP highest exposure categories showed significant associations with BWT decreases. For example, statistically significant deficits were seen for the upper four HAA5 categories (range: -4 to -8 g), upper three DCAA categories (range: -8 to -13 g), and upper three TCAA categories (range: -6 to -13 g) compared to the referents. Larger decreases were observed from multi-DBP models (i.e., models with further adjustments for THM4 or HAA5). After further adjustment for THM4 or HAA5 exposures, the reductions in BWT were slightly larger only in the upper four HAA5 categories (range: -

818 8 to -12 g) and upper three DCAA categories (range: -6 to -16 g). No evident deficits
819 were noticed in other exposure categories.

820

821 **5.3.2 Small for Gestational Age**

822 Very small increased risks of SGA (5%) were observed for third-trimester THM
823 exposures (OR range: 0.98 – 1.03, Table 19) in the unadjusted univariate model. Small
824 but statistically significant increased risks of SGA (5%) were detected for HAA5 (OR =
825 1.07, 95%CI: 1.03, 1.12), DCAA (OR = 1.06, 95%CI: 1.01, 1.11), TCAA (OR = 1.05,
826 95%CI: 1.01, 1.10), HAA9 (OR = 1.05, 95%CI: 1.01, 1.10), DBP9 (OR = 1.04, 95%CI:
827 1.00, 1.09), and DBP13 (OR = 1.05, 95%CI: 1.01, 1.10) in unadjusted models based on
828 the highest exposure categories. After adjusting for all the covariates, statistically
829 significant elevated risks were only observed in TCAA exposure categories (OR = 1.05;
830 95%CI: 1.00, 1.10). With additional adjustment for THM4 or HAA5, statistically
831 significant higher risks in SGA (5%) were observed in HAA5 (OR = 1.09, 95%CI: 1.02,
832 1.16), TCAA (OR = 1.08, 95%CI: 1.02, 1.15), and HAA9 (OR = 1.06, 95%CI: 1.00,
833 1.14) exposure categories. Similar results were observed based on second-trimester
834 exposures (Table S6-Table **S7**).

835 Although most THM4 results were null, higher increased risks of SGA (10%) were
836 noticed for THM4 exposure in the third trimester in unadjusted models (OR = 1.07,
837 95%CI: 1.03, 1.10, Table 20). Increased odds of SGA (10%) were also observed in upper
838 HAA5 (OR = 1.09, 95%CI: 1.05, 1.12), DCAA (OR = 1.07, 95%CI: 1.03, 1.11), TCAA
839 (OR = 1.05, 95%CI: 1.02, 1.08), HAA9 (OR = 1.06, 95%CI: 1.03, 1.10), DBP9 (OR =

1.08, 95%CI: 1.05, 1.12), and DBP13 (OR = 1.08, 95%CI: 1.05, 1.12) exposure categories. Following adjustment for confounding, results were largely null except for HAA5 (OR = 1.06; 95%CI: 1.01, 1.12). After adjusted for THM4 or HAA5, statistically significant increased risk of SGA (10%) was only observed for HAA5 exposures (OR = 1.06, 95%CI: 1.01, 1.12).

### 5.3.3 Preterm Birth

Statistically significant associations between second-trimester DBP exposure and PTB were observed in all HAA categories in unadjusted models (OR range: 1.05 - 1.18, Table 21). Higher range of ORs were observed in HAA5 (OR = 1.08 - 1.14), DCAA (OR = 1.07 - 1.18), HAA9 (OR = 1.06 – 1.14), DBP9 (OR = 1.07 – 1.13), and DBP13 (OR = 1.07 – 1.11) in unadjusted models. After adjustment for all covariates, only THM4 and HAA5 exposure categories showed statistically significant association to PTB OR (OR = 1.06 and 1.05, respectively). Results were largely null following additional adjustment for THM4 or HAA5 and PTB with the exception of THM4 (OR = 1.05) and DCAA (OR = 1.06). PTB results with first-trimester exposures were similar to results from second-trimester exposures (Table S8).

**Table 18. Association Between Third-Trimester DBP Exposure and Change in Mean BWT Among Term Births.**

| DBP Metrics [1] | Categorical Levels, µg/L | Births (n) | Unadjusted BWT (95% CI), g | Adjusted BWT [2] (95% CI), g | Adjusted BWT [3] (95% CI), g |
|---|---|---|---|---|---|
| **THM4** [4] | 0-15 | 80,170 | REF | REF | REF |
| | >15-25 | 162,624 | -23 (-27, -19) | 1 (-3, 5) | 5 (0, 10) |
| | >25-33 | 173,270 | -27 (-31, -23) | 2 (-2, 6) | 8 (2, 13) |
| | >33-43 | 165,094 | -22 (-26, -18) | 0 (-4, 4) | 7 (1, 12) |
| | >43-64 | 185,107 | -21 (-25, -17) | -3 (-7, 1) | 5 (-1, 11) |
| | >64-416 | 86,935 | -14 (-19, -10) | -2 (-7, 3) | 7 (0, 13) |
| **THM-Br** [5] | 0-7 | 77,066 | REF | REF | REF |
| | >7-11 | 145,866 | -12 (-16, -8) | 5 (1, 9) | 10 (6, 15) |
| | >11-14 | 171,219 | -6 (-10, -2) | 7 (3, 12) | 13 (8, 17) |
| | >14-18 | 195,976 | -10 (-14, -6) | 5 (1, 9) | 10 (5, 14) |
| | >18-28 | 176,532 | 2 (-2, 6) | 7 (3, 12) | 12 (7, 17) |
| | >28-177 | 86,541 | 1 (-3, 6) | 14 (9, 19) | 19 (13, 24) |
| **HAA5** [6] | 0-6 | 77,999 | REF | REF | REF |
| | >6-10 | 153,880 | -25 (-29, -21) | 0 (-4, 5) | -4 (-9, 1) |
| | >10-19 | 182,831 | -32 (-35, -28) | -4 (-8, 0) | -8 (-14, -3) |
| | >19-27 | 174,103 | -37 (-41, -33) | -7 (-11, -3) | -11 (-17, -6) |
| | >27-38 | 177,484 | -38 (-42, -34) | -5 (-9, -1) | -9 (-15, -4) |
| | >38-164 | 86,903 | -42 (-46, -37) | -8 (-13, -3) | -12 (-19, -6) |
| **DCAA** | 0-1 | 68,423 | REF | REF | REF |
| | >1-5 | 240,952 | -23 (-27, -19) | 1 (-3, 5) | -2 (-7, 3) |
| | >5-13 | 258,677 | -29 (-33, -25) | -8 (-12, -4) | -11 (-16, -6) |
| | >13-22 | 207,666 | -34 (-38, -30) | -4 (-8, 1) | -6 (-12, -1) |
| | >22-88 | 77,482 | -50 (-55, -45) | -13 (-18, -8) | -16 (-22, -10) |
| **TCAA** | 0-0.5 | 78,700 | REF | REF | REF |
| | >0.5-2 | 230,194 | -14 (-18, -11) | -4 (-8, 0) | -4 (-8, -1) |
| | >2-7 | 213,237 | -24 (-28, -20) | -13 (-17, -9) | -13 (-17, -9) |
| | >7-15 | 245,237 | -16 (-20, -12) | -2 (-6, 2) | -2 (-7, 2) |
| | >15-88 | 85,832 | -15 (-19, -10) | -6 (-11, -2) | -6 (-11, 0) |
| **HAA-Br** [7] | 0-3 | 55,658 | REF | REF | REF |
| | >3-6 | 279,734 | 0 (-5, 4) | 11 (7, 15) | 11 (7, 16) |
| | >6-8 | 226,233 | 0 (-5, 4) | 13 (9, 17) | 13 (9, 18) |
| | >8-11 | 206,477 | -12 (-16, -8) | 10 (6, 15) | 11 (6, 15) |
| | >11-78 | 85,098 | -3 (-8, 2) | 12 (7, 17) | 13 (8, 18) |
| **HAA4** [8] | 0-2 | 59,372 | REF | REF | REF |
| | >2-4 | 258,434 | 1 (-3, 5) | 12 (8, 16) | 12 (8, 17) |
| | >4-7 | 321,653 | 0 (-4, 4) | 12 (8, 16) | 13 (8, 17) |
| | >7-10 | 158,969 | -11 (-15, -6) | 7 (2, 11) | 7 (3, 12) |
| | >10-76 | 54,772 | -7 (-13, -2) | 12 (6, 17) | 12 (7, 18) |
| **HAA9** [9] | 0-9 | 73,905 | REF | REF | REF |
| | >9-15 | 175,151 | -23 (-27, -19) | 2 (-2, 6) | -1 (-6, 4) |

| | | | | | |
|---|---|---|---|---|---|
| | >15-24 | 164,391 | -31 (-35, -27) | -3 (-7, 1) | -5 (-11, 0) |
| | >24-33 | 178,214 | -37 (-41, -33) | -6 (-10, -2) | -8 (-13, -3) |
| | >33-45 | 172,288 | -35 (-39, -31) | -5 (-9, -1) | -7 (-13, -1) |
| | >45-197 | 89,251 | -34 (-39, -30) | -2 (-7, 3) | -4 (-10, 2) |
| **DBP9** [10] | 0-24 | 82,741 | REF | REF | - |
| | >24-41 | 172,178 | -30 (-34, -26) | -3 (-7, 1) | - |
| | >41-53 | 167,050 | -39 (-42, -35) | -3 (-7, 1) | - |
| | >53-71 | 173,428 | -36 (-39, -32) | -5 (-9, -1) | - |
| | >71-97 | 171,410 | -35 (-39, -31) | -8 (-12, -4) | - |
| | >97-475 | 86,393 | -27 (-32, -23) | -5 (-9, 0) | - |
| **DBP13** [11] | 0-27 | 82,702 | REF | REF | - |
| | >27-45 | 162,661 | -29 (-33, -25) | -3 (-7, 1) | - |
| | >45-59 | 170,350 | -39 (-43, -35) | -5 (-9, -1) | - |
| | >59-78 | 178,691 | -35 (-39, -31) | -5 (-9, -1) | - |
| | >78-105 | 171,246 | -33 (-37, -29) | -8 (-12, -4) | - |
| | >105-477 | 87,550 | -30 (-34, -25) | -5 (-9, 0) | - |

[1]. The categorization of each DBP class or species was based on the distribution of the DBP metrics. THM4, THM-Br, HAA5, HAA9, DBP9, and DBP13 were categorized into 6 categories: 0-10, 10-30, 30-50, 50-70, 70-90, and 90-100 percentile.
DCAA, TCAA, HAA-Br, and HAA4 were categorized into 5 categories: 0-10, 10-36, 36-64, 64-90, and 90-100 percentile.
[2]. Model adjusted for maternal age, race, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model also adjusted for THM4 or HAA5 exposures.
[4]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[5]. THM-Br, the summation of BDCM, CDBM, and TBM.
[6]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[7]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[8]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[9]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.
[10]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[11]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

860

861

**Table 19. Association Between Third-Trimester DBP Exposure and Change in SGA (5%) OR Among Term Births.**

| DBP Metrics [1] | Categorical Levels, µg/L | Births (n) | Unadjusted SGA OR (95% CI) | Adjusted SGA OR[2] (95% CI) | Adjusted SGA OR[3] (95% CI) |
|---|---|---|---|---|---|
| **THM4** [4] | 0-15 | 80,077 | REF | REF | REF |
| | >15-25 | 162,394 | 1.01 (0.98, 1.05) | 1.01 (0.97, 1.06) | 1.01 (0.96, 1.07) |
| | >25-33 | 172,822 | 1.00 (0.96, 1.04) | 0.99 (0.95, 1.03) | 0.98 (0.93, 1.04) |
| | >33-43 | 164,629 | 0.98 (0.94, 1.02) | 0.97 (0.93, 1.01) | 0.96 (0.90, 1.01) |
| | >43-64 | 184,706 | 1.01 (0.97, 1.04) | 0.98 (0.94, 1.03) | 0.96 (0.91, 1.02) |
| | >64-416 | 86,814 | 1.03 (0.99, 1.08) | 0.99 (0.94, 1.04) | 0.95 (0.89, 1.02) |
| **THM-Br** [5] | 0-7 | 76,993 | REF | REF | REF |
| | >7-11 | 145,434 | 1.01 (0.97, 1.05) | 1.03 (0.98, 1.08) | 1.04 (0.99, 1.10) |
| | >11-14 | 170,766 | 0.98 (0.95, 1.02) | 1.02 (0.98, 1.07) | 1.04 (0.99, 1.09) |
| | >14-18 | 195,579 | 1.00 (0.96, 1.03) | 1.01 (0.97, 1.06) | 1.02 (0.97, 1.07) |
| | >18-28 | 176,273 | 1.00 (0.96, 1.03) | 1.03 (0.99, 1.08) | 1.04 (0.99, 1.09) |
| | >28-177 | 86,397 | 0.92 (0.88, 0.96) | 0.93 (0.88, 0.98) | 0.94 (0.88, 0.99) |
| **HAA5** [6] | 0-6 | 77,926 | REF | REF | REF |
| | >6-10 | 153,676 | 1.00 (0.96, 1.04) | 0.99 (0.95, 1.04) | 1.00 (0.95, 1.06) |
| | >10-19 | 182,562 | 0.98 (0.94, 1.02) | 0.98 (0.93, 1.02) | 1.00 (0.95, 1.06) |
| | >19-27 | 173,677 | 0.98 (0.94, 1.02) | 0.97 (0.93, 1.02) | 1.00 (0.94, 1.06) |
| | >27-38 | 176,873 | 1.00 (0.96, 1.04) | 0.98 (0.94, 1.02) | 1.02 (0.96, 1.08) |
| | >38-164 | 86,728 | 1.07 (1.03, 1.12) | 1.04 (0.99, 1.10) | 1.09 (1.02, 1.16) |
| **DCAA** | 0-1 | 68,352 | REF | REF | REF |
| | >1-5 | 240,588 | 0.98 (0.94, 1.02) | 0.99 (0.95, 1.04) | 1.00 (0.95, 1.06) |
| | >5-13 | 258,187 | 0.98 (0.94, 1.02) | 1.00 (0.95, 1.04) | 1.01 (0.96, 1.07) |
| | >13-22 | 206,964 | 1.01 (0.97, 1.05) | 1.01 (0.96, 1.05) | 1.04 (0.98, 1.10) |
| | >22-88 | 77,351 | 1.06 (1.01, 1.11) | 1.01 (0.96, 1.07) | 1.04 (0.98, 1.11) |
| **TCAA** | 0-1 | 78,601 | REF | REF | REF |
| | >1-2 | 229,911 | 1.00 (0.96, 1.04) | 1.01 (0.97, 1.05) | 1.02 (0.97, 1.06) |
| | >2-7 | 212,905 | 1.00 (0.97, 1.04) | 1.02 (0.98, 1.06) | 1.03 (0.99, 1.08) |
| | >7-15 | 244,398 | 0.97 (0.93, 1.00) | 1.00 (0.95, 1.04) | 1.02 (0.97, 1.07) |
| | >15-88 | 85,627 | 1.05 (1.01, 1.10) | 1.05 (1.00, 1.11) | 1.08 (1.02, 1.15) |
| **HAA-Br** [7] | 0-3 | 55,615 | REF | REF | REF |
| | >3-6 | 279,204 | 1.00 (0.96, 1.04) | 1.01 (0.97, 1.06) | 1.01 (0.96, 1.06) |
| | >6-8 | 225,560 | 0.95 (0.91, 0.99) | 1.00 (0.96, 1.05) | 1.01 (0.96, 1.06) |
| | >8-11 | 206,075 | 0.95 (0.91, 1.00) | 1.00 (0.96, 1.05) | 1.01 (0.96, 1.06) |
| | >11-78 | 84,988 | 0.97 (0.92, 1.02) | 0.99 (0.94, 1.05) | 1.00 (0.95, 1.06) |
| **HAA4** [8] | 0-2 | 59,331 | REF | REF | REF |
| | >2-4 | 258,045 | 1.00 (0.96, 1.04) | 1.01 (0.97, 1.06) | 1.01 (0.97, 1.06) |
| | >4-7 | 320,676 | 0.96 (0.92, 0.99) | 1.00 (0.96, 1.05) | 1.01 (0.96, 1.06) |
| | >7-10 | 158,682 | 0.96 (0.92, 1.00) | 1.00 (0.96, 1.05) | 1.01 (0.96, 1.06) |
| | >10-76 | 54,708 | 1.02 (0.96, 1.07) | 1.02 (0.96, 1.09) | 1.03 (0.97, 1.09) |
| **HAA9** [9] | 0-9 | 73,833 | REF | REF | REF |
| | >9-15 | 174,930 | 1.01 (0.97, 1.05) | 1.01 (0.96, 1.05) | 1.02 (0.97, 1.08) |

| | | | | | |
|---|---|---|---|---|---|
| | >15-24 | 164,140 | 0.98 (0.94, 1.02) | 0.98 (0.94, 1.03) | 1.01 (0.95, 1.07) |
| | >24-33 | 177,721 | 0.99 (0.96, 1.03) | 0.98 (0.94, 1.03) | 1.01 (0.95, 1.07) |
| | >33-45 | 171,734 | 0.99 (0.95, 1.03) | 1.00 (0.96, 1.05) | 1.04 (0.98, 1.10) |
| | >45-197 | 89,084 | 1.05 (1.01, 1.10) | 1.02 (0.97, 1.08) | 1.06 (1.00, 1.14) |
| **DBP9** [10] | 0-24 | 82,652 | REF | REF | - |
| | >24-41 | 171,973 | 1.01 (0.97, 1.05) | 1.01 (0.97, 1.05) | - |
| | >41-53 | 166,683 | 0.97 (0.94, 1.01) | 0.97 (0.93, 1.02) | - |
| | >53-71 | 172,883 | 0.99 (0.96, 1.03) | 0.97 (0.93, 1.01) | - |
| | >71-97 | 171,009 | 1.01 (0.98, 1.05) | 0.99 (0.95, 1.03) | - |
| | >97-475 | 86,242 | 1.04 (1.00, 1.09) | 0.99 (0.94, 1.04) | - |
| **DBP13** [11] | 0-27 | 82,613 | REF | REF | - |
| | >27-45 | 162,463 | 1.03 (0.99, 1.07) | 1.03 (0.98, 1.07) | - |
| | >45-59 | 169,970 | 0.99 (0.96, 1.03) | 0.99 (0.94, 1.03) | - |
| | >59-78 | 178,158 | 1.00 (0.96, 1.03) | 0.98 (0.94, 1.03) | - |
| | >78-105 | 170,842 | 1.01 (0.98, 1.05) | 1.00 (0.96, 1.04) | - |
| | >105-477 | 87,396 | 1.05 (1.01, 1.10) | 1.00 (0.96, 1.06) | - |

[1]. The categorization of each DBP class or species was based on the distribution of the DBP. THM4, THM-Br, HAA5, HAA9, DBP9, and DBP13 were categorized into 6 categories: 0-10, 10-30, 30-50, 50-70, 70-90, and 90-100 percentile.
DCAA, TCAA, HAA-Br, and HAA4 were categorized into 5 categories: 0-10, 10-36, 36-64, 64-90, and 90-100 percentile.
[2]. Model adjusted for maternal age, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model also adjusted for THM4 or HAA5 exposures.
[4]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[5]. THM-Br, the summation of BDCM, CDBM, and TBM.
[6]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[7]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[8]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[9]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.
[10]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[11]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

862

863

864

**Table 20. Association Between Third-Trimester DBP Exposure and Change in SGA (10%) OR Among Term Births.**

| DBP Metrics [1] | Categorical Levels, µg/L | Births (n) | Unadjusted SGA OR (95% CI) | Adjusted SGA OR[2] (95% CI) | Adjusted SGA OR[3] (95% CI) |
|---|---|---|---|---|---|
| **THM4** [4] | 0-15 | 80,077 | REF | REF | REF |
| | >15-25 | 162,394 | 1.03 (1.00, 1.06) | 1.03 (1.00, 1.07) | 1.02 (0.98, 1.06) |
| | >25-33 | 172,822 | 1.02 (0.99, 1.05) | 1.02 (0.98, 1.05) | 1.01 (0.96, 1.05) |
| | >33-43 | 164,629 | 1.01 (0.98, 1.03) | 1.00 (0.97, 1.03) | 0.99 (0.94, 1.03) |
| | >43-64 | 184,706 | 1.04 (1.01, 1.07) | 1.01 (0.98, 1.04) | 0.99 (0.95, 1.04) |
| | >64-416 | 86,814 | 1.07 (1.03, 1.10) | 1.03 (0.99, 1.07) | 1.00 (0.96, 1.06) |
| **THM-Br** [5] | 0-7 | 76,993 | REF | REF | REF |
| | >7-11 | 145,434 | 1.02 (0.99, 1.05) | 1.03 (1.00, 1.07) | 1.02 (0.99, 1.06) |
| | >11-14 | 170,766 | 0.98 (0.95, 1.01) | 1.01 (0.98, 1.05) | 1.00 (0.97, 1.04) |
| | >14-18 | 195,579 | 1.01 (0.99, 1.04) | 1.03 (0.99, 1.06) | 1.02 (0.98, 1.05) |
| | >18-28 | 176,273 | 0.99 (0.97, 1.02) | 1.03 (1.00, 1.06) | 1.02 (0.99, 1.06) |
| | >28-177 | 86,397 | 0.94 (0.91, 0.97) | 0.95 (0.91, 0.98) | 0.94 (0.90, 0.98) |
| **HAA5** [6] | 0-6 | 77,926 | REF | REF | REF |
| | >6-10 | 153,676 | 1.03 (1.00, 1.06) | 1.03 (0.99, 1.06) | 1.02 (0.98, 1.07) |
| | >10-19 | 182,562 | 1.01 (0.98, 1.03) | 1.00 (0.96, 1.03) | 1.00 (0.96, 1.04) |
| | >19-27 | 173,677 | 1.02 (0.99, 1.05) | 1.03 (1.00, 1.06) | 1.03 (0.99, 1.08) |
| | >27-38 | 176,873 | 1.03 (1.00, 1.06) | 1.01 (0.98, 1.05) | 1.02 (0.97, 1.07) |
| | >38-164 | 86,728 | 1.09 (1.05, 1.12) | 1.06 (1.02, 1.10) | 1.06 (1.01, 1.12) |
| **DCAA** | 0-1 | 68,352 | REF | REF | REF |
| | >1-5 | 240,588 | 1.00 (0.97, 1.03) | 1.00 (0.97, 1.04) | 1.00 (0.96, 1.04) |
| | >5-13 | 258,187 | 1.01 (0.98, 1.03) | 1.02 (0.99, 1.06) | 1.02 (0.98, 1.06) |
| | >13-22 | 206,964 | 1.04 (1.01, 1.07) | 1.02 (0.99, 1.06) | 1.03 (0.99, 1.07) |
| | >22-88 | 77,351 | 1.07 (1.03, 1.11) | 1.04 (0.99, 1.08) | 1.04 (0.99, 1.09) |
| **TCAA** | 0-1 | 78,601 | REF | REF | REF |
| | >1-2 | 229,911 | 1.00 (0.97, 1.03) | 1.01 (0.98, 1.04) | 1.00 (0.97, 1.04) |
| | >2-7 | 212,905 | 1.02 (1.00, 1.05) | 1.03 (1.00, 1.06) | 1.03 (1.00, 1.07) |
| | >7-15 | 244,398 | 0.98 (0.95, 1.01) | 1.00 (0.97, 1.03) | 1.00 (0.97, 1.04) |
| | >15-88 | 85,627 | 1.05 (1.02, 1.08) | 1.04 (1.00, 1.08) | 1.04 (1.00, 1.08) |
| **HAA-Br** [7] | 0-3 | 55,615 | REF | REF | REF |
| | >3-6 | 279,204 | 1.01 (0.98, 1.05) | 1.02 (0.99, 1.05) | 1.01 (0.98, 1.05) |
| | >6-8 | 225,560 | 0.97 (0.94, 1.00) | 1.01 (0.97, 1.04) | 1.00 (0.96, 1.04) |
| | >8-11 | 206,075 | 0.97 (0.94, 1.00) | 1.00 (0.97, 1.04) | 1.00 (0.96, 1.04) |
| | >11-78 | 84,988 | 0.97 (0.93, 1.00) | 0.98 (0.94, 1.02) | 0.98 (0.94, 1.02) |
| **HAA4** [8] | 0-2 | 59,331 | REF | REF | REF |
| | >2-4 | 258,045 | 1.00 (0.97, 1.03) | 1.01 (0.98, 1.04) | 1.00 (0.97, 1.04) |
| | >4-7 | 320,676 | 0.96 (0.94, 0.99) | 0.99 (0.96, 1.02) | 0.99 (0.95, 1.02) |
| | >7-10 | 158,682 | 0.97 (0.94, 1.00) | 1.01 (0.97, 1.04) | 1.00 (0.97, 1.04) |
| | >10-76 | 54,708 | 0.99 (0.95, 1.03) | 0.99 (0.95, 1.03) | 0.98 (0.94, 1.03) |
| **HAA9** [9] | 0-9 | 73,833 | REF | REF | REF |
| | >9-15 | 174,930 | 1.04 (1.01, 1.07) | 1.04 (1.01, 1.07) | 1.04 (1.00, 1.08) |

| | | | | | |
|---|---|---|---|---|---|
| | >15-24 | 164,140 | 1.01 (0.98, 1.04) | 1.00 (0.97, 1.04) | 1.01 (0.97, 1.06) |
| | >24-33 | 177,721 | 1.03 (1.00, 1.06) | 1.02 (0.99, 1.06) | 1.03 (0.99, 1.07) |
| | >33-45 | 171,734 | 1.03 (1.00, 1.06) | 1.03 (1.00, 1.07) | 1.05 (1.00, 1.09) |
| | >45-197 | 89,084 | 1.06 (1.03, 1.10) | 1.04 (1.00, 1.08) | 1.04 (0.99, 1.10) |
| **DBP9** [10] | 0-24 | 82,652 | REF | REF | - |
| | >24-41 | 171,973 | 1.04 (1.01, 1.07) | 1.04 (1.01, 1.07) | - |
| | >41-53 | 166,683 | 1.01 (0.98, 1.04) | 1.01 (0.98, 1.04) | - |
| | >53-71 | 172,883 | 1.03 (1.00, 1.06) | 1.01 (0.98, 1.05) | - |
| | >71-97 | 171,009 | 1.04 (1.02, 1.07) | 1.02 (0.99, 1.06) | - |
| | >97-475 | 86,242 | 1.08 (1.05, 1.12) | 1.04 (1.00, 1.08) | - |
| **DBP13** [11] | 0-27 | 82,613 | REF | REF | - |
| | >27-45 | 162,463 | 1.05 (1.02, 1.08) | 1.05 (1.02, 1.09) | - |
| | >45-59 | 169,970 | 1.03 (1.00, 1.06) | 1.02 (0.99, 1.06) | - |
| | >59-78 | 178,158 | 1.03 (1.00, 1.06) | 1.03 (0.99, 1.06) | - |
| | >78-105 | 170,842 | 1.04 (1.02, 1.07) | 1.03 (1.00, 1.06) | - |
| | >105-477 | 87,396 | 1.08 (1.05, 1.12) | 1.04 (1.01, 1.08) | - |

[1]. The categorization of each DBP class or species was based on the distribution of the DBP metrics. THM4, THM-Br, HAA5, HAA9, DBP9, and DBP13 were categorized into 6 categories: 0-10, 10-30, 30-50, 50-70, 70-90, and 90-100 percentile.
DCAA, TCAA, HAA-Br, and HAA4 were categorized into 5 categories: 0-10, 10-36, 36-64, 64-90, and 90-100 percentile.
[2]. Model adjusted for maternal age, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model also adjusted for THM4 or HAA5 exposures.
[4]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[5]. THM-Br, the summation of BDCM, CDBM, and TBM.
[6]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[7]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[8]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[9]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.
[10]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[11]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

865

**Table 21. Association Between Second-Trimester DBP Exposure and Change in PTB OR Among Births.**

| DBP Metrics [1] | Categorical Levels, µg/L | Births (n) | Unadjusted PTB OR (95% CI) | Adjusted PTB OR[2] (95% CI) | Adjusted PTB OR[3] (95% CI) |
|---|---|---|---|---|---|
| **THM4** [4] | 0-16 | 89,276 | REF | REF | REF |
| | >16-26 | 191,959 | 1.09 (1.07, 1.12) | 1.06 (1.02, 1.09) | 1.05 (1.01, 1.09) |
| | >26-33 | 177,151 | 1.12 (1.09, 1.15) | 1.02 (0.99, 1.06) | 1.03 (0.99, 1.07) |
| | >33-42 | 191,667 | 1.09 (1.06, 1.12) | 1.00 (0.97, 1.03) | 1.01 (0.97, 1.05) |
| | >42-61 | 206,096 | 1.03 (1.01, 1.06) | 0.96 (0.93, 0.99) | 0.97 (0.93, 1.01) |
| | >61-407 | 96,041 | 1.04 (1.01, 1.07) | 0.98 (0.94, 1.02) | 0.99 (0.94, 1.03) |
| **THM-Br** [5] | 0-7 | 78,990 | REF | REF | REF |
| | >7-11 | 165,615 | 1.05 (1.03, 1.08) | 0.99 (0.96, 1.02) | 0.99 (0.96, 1.03) |
| | >11-14 | 214,349 | 1.05 (1.02, 1.08) | 0.99 (0.96, 1.02) | 0.99 (0.96, 1.03) |
| | >14-17 | 177,618 | 1.01 (0.99, 1.04) | 0.99 (0.96, 1.03) | 0.99 (0.95, 1.02) |
| | >17-27 | 219,052 | 1.02 (0.99, 1.04) | 0.99 (0.96, 1.02) | 0.98 (0.95, 1.02) |
| | >27-173 | 96,566 | 1.04 (1.01, 1.07) | 0.97 (0.93, 1.00) | 0.95 (0.91, 0.99) |
| **HAA5** [6] | 0-6 | 84,611 | REF | REF | REF |
| | >6-10 | 167,249 | 1.1 (1.07, 1.13) | 1.05 (1.02, 1.08) | 1.03 (0.99, 1.07) |
| | >10-19 | 211,400 | 1.08 (1.05, 1.11) | 0.97 (0.94, 1.00) | 0.97 (0.93, 1.01) |
| | >19-26 | 183,439 | 1.09 (1.06, 1.12) | 0.97 (0.94, 1.00) | 0.97 (0.93, 1.02) |
| | >26-36 | 202,454 | 1.12 (1.09, 1.15) | 0.98 (0.95, 1.01) | 1.00 (0.95, 1.04) |
| | >36-159 | 103,037 | 1.14 (1.11, 1.18) | 0.99 (0.95, 1.03) | 1.01 (0.96, 1.06) |
| **DCAA** | 0-1 | 66,909 | REF | REF | REF |
| | >1-6 | 294,576 | 1.12 (1.09, 1.15) | 1.04 (1.01, 1.08) | 1.03 (0.99, 1.07) |
| | >6-13 | 279,383 | 1.07 (1.04, 1.10) | 0.97 (0.94, 1.00) | 0.97 (0.93, 1.01) |
| | >13-21 | 230,249 | 1.12 (1.09, 1.16) | 0.99 (0.95, 1.02) | 1.01 (0.97, 1.06) |
| | >21-86 | 81,073 | 1.18 (1.14, 1.22) | 1.03 (0.99, 1.07) | 1.06 (1.01, 1.11) |
| **TCAA** | 0-1 | 80,440 | REF | REF | REF |
| | >1-2 | 269,972 | 1.05 (1.02, 1.08) | 0.99 (0.96, 1.02) | 0.99 (0.95, 1.02) |
| | >2-7 | 234,208 | 1.03 (1.01, 1.06) | 0.97 (0.94, 1.00) | 0.97 (0.94, 1.00) |
| | >7-14 | 264,299 | 1.07 (1.04, 1.09) | 0.96 (0.93, 0.99) | 0.97 (0.93, 1.00) |
| | >14-86 | 103,271 | 1.03 (1.00, 1.06) | 0.92 (0.89, 0.96) | 0.94 (0.90, 0.98) |
| **HAA-Br** [7] | 0-3 | 57,624 | REF | REF | REF |
| | >3-6 | 327,281 | 1.00 (0.97, 1.03) | 0.98 (0.94, 1.01) | 0.96 (0.92, 0.99) |
| | >6-8 | 256,420 | 1.03 (1.00, 1.06) | 0.96 (0.93, 1.00) | 0.95 (0.91, 0.98) |
| | >8-11 | 219,625 | 1.07 (1.04, 1.11) | 0.97 (0.94, 1.01) | 0.95 (0.92, 0.99) |
| | >11-74 | 91,240 | 1.05 (1.02, 1.09) | 0.97 (0.93, 1.01) | 0.95 (0.91, 0.99) |
| **HAA4** [8] | 0-2 | 60,888 | REF | REF | REF |
| | >2-5 | 417,277 | 1.01 (0.98, 1.04) | 0.99 (0.96, 1.02) | 0.97 (0.93, 1.00) |
| | >5-6 | 131,658 | 1.04 (1.01, 1.08) | 0.97 (0.94, 1.01) | 0.97 (0.93, 1.00) |
| | >6-9 | 263,452 | 1.05 (1.02, 1.08) | 0.96 (0.92, 0.99) | 0.94 (0.91, 0.98) |
| | >9-72 | 78,915 | 1.07 (1.03, 1.10) | 0.97 (0.93, 1.01) | 0.96 (0.92, 1.00) |
| **HAA9** [9] | 0-9 | 80,261 | REF | REF | REF |

| | | | | | |
|---|---|---|---|---|---|
| | >9-15 | 193,790 | 1.09 (1.06, 1.12) | 1.04 (1.00, 1.07) | 1.03 (0.99, 1.07) |
| | >15-24 | 182,329 | 1.08 (1.05, 1.11) | 0.99 (0.95, 1.02) | 1.00 (0.96, 1.04) |
| | >24-32 | 185,527 | 1.06 (1.03, 1.09) | 0.95 (0.91, 0.98) | 0.96 (0.92, 1.00) |
| | >32-43 | 207,701 | 1.14 (1.11, 1.17) | 1.00 (0.97, 1.04) | 1.03 (0.99, 1.08) |
| | >43-191 | 102,582 | 1.12 (1.08, 1.15) | 0.97 (0.93, 1.01) | 0.99 (0.95, 1.04) |
| **DBP9** [10] | 0-25 | 95,145 | REF | REF | - |
| | >25-41 | 188,939 | 1.08 (1.06, 1.11) | 1.02 (0.99, 1.05) | - |
| | >41-53 | 182,564 | 1.13 (1.10, 1.16) | 0.99 (0.96, 1.02) | - |
| | >53-69 | 196,726 | 1.08 (1.05, 1.11) | 0.96 (0.94, 1.00) | - |
| | >69-92 | 189,096 | 1.07 (1.04, 1.09) | 0.95 (0.92, 0.98) | - |
| | >92-465 | 99,720 | 1.07 (1.04, 1.11) | 0.97 (0.93, 1.00) | - |
| **DBP13** [11] | 0-28 | 93,564 | REF | REF | - |
| | >28-46 | 185,052 | 1.08 (1.05, 1.11) | 1.03 (0.99, 1.06) | - |
| | >46-59 | 188,181 | 1.11 (1.08, 1.14) | 0.97 (0.94, 1.00) | - |
| | >59-75 | 196,901 | 1.09 (1.06, 1.12) | 0.97 (0.94, 1.00) | - |
| | >75-99 | 190,429 | 1.07 (1.04, 1.09) | 0.95 (0.92, 0.98) | - |
| | >99-467 | 98,063 | 1.07 (1.04, 1.11) | 0.96 (0.93, 1.00) | - |

[1]. The categorization of each DBP class or species was based on the distribution of the DBP metrics. THM4, THM-Br, HAA5, HAA9, DBP9, and DBP13 were categorized into 6 categories: 0-10, 10-30, 30-50, 50-70, 70-90, and 90-100 percentile.
DCAA, TCAA, HAA-Br, and HAA4 were categorized into 5 categories: 0-10, 10-36, 36-64, 64-90, and 90-100 percentile.
[2]. Model adjusted for maternal age, race, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model also adjusted for THM4 or HAA5 exposures.
[4]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[5]. THM-Br, the summation of BDCM, CDBM, and TBM.
[6]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[7]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[8]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[9]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.
[10]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[11]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

### 5.3.4 Stratified Results

The following covariates were examined as effect-measure modifiers in this research: infant sex, maternal age, race, education, pay source of delivery, income, and adequacy of prenatal care. Infants from Hispanic mothers had the smallest mean BWT differences

of 5 to 10 g. Infants from mothers of other groups included the following mean BWT differences: -2 to 3 g from White mothers, -4 to 5 g from African American mothers, -18 to 18 g from Asian Indian mothers, -38 to -50 g from Asian mothers, and -57 to -96 g for Native American mothers (Table S10). Mothers who have bachelor/associate degrees had the largest infant mean BWT difference of -12 to 2 g. Maternal education level below high school/GED have the smallest mean BWT differences of 2 to 10 g. Infants from mothers with a public source of payment for delivery had the lowest mean BWT differences of 7 to 11 g. Infants from mothers with other/reported delivery source of payment had the largest mean BWT difference of -11 to 9 g. Infants with inadequate prenatal care visits had the largest mean BWT differences of -29 to 2 g. In the fully adjusted model of third-trimester DBP, we observed minor differences between infants' sex that male infants had a mean BWT difference of -4 to -21 g and -9 to -25 g with HAA5 and DCAA exposure, respectively. For female infants, the birth deficits are -6 to -23 g and -7 to -24 g for the same DBP exposure, respectively.

## 5.4 Discussion

### 5.4.1 Birth Outcomes

#### 5.4.1.1 Birth Weight Reduction

Based on our study population in Ohio, we saw little evidence of mean BWT reductions for THM exposures. This largely runs counter to the majority of evidence reported previously. For example, mean BWT deficits (Range: ~20−70 g) were seen in six (Bove et al., 1995; Wright et al., 2003; Wright et al., 2004; Smith et al., 2016; Grazuleviciene et al., 2011; Cao et al., 2016) out of nine (Bove et al., 1995; Hoffman et al., 2008; Wright et al., 2003; Wright et al., 2004; Rivera-Núñez and Wright, 2013; Kogevinas et al., 2016; Smith et al., 2016; Grazuleviciene et al., 2011; Cao et al., 2016) studies for high THM4 exposure groups and three (Cao et al., 2016; Smith et al., 2016; Rivera-Núñez et al., 2013) out of five (Cao et al., 2016; Smith et al., 2016; Rivera-Núñez and Wright, 2013; Kogevinas et al., 2016; Villanueva et al., 2011) studies of brominated THM exposures.

We detected statistically significant mean BWT reductions similar in magnitude for the highest DCAA (-16 g; 95%CI: -22, -10 g) and HAA5 (-12 g; 95%CI: -19, -6 g) exposure categories after adjusting for confounders and THM4 (Table 18). We found limited evidence of any differences across various DBP mixtures metrics such as DBP9 exposure categories (-5 g; 95%CI: -9, 0) or the estimates based on our HAA prediction model estimates for either HAA9 (-4 g; 95%CI: -10, 2) or DBP13 (-5 g; 95%CI: -9, 0) exposure categories.A study of maternal urinary levels (Zhou et al., 2012) showed a BWT reduction of -61.7 g for the highest exposure quartile (vs. quartile one referent) in

908 the fully adjusted model. Hoffman et al. (2008) reported consistent BWT deficits (-18 to -
909 43 g) across BDCAA in two study sites (one more chlorinated and one more brominated)
910 study sites. They also reported large deficits for the highest exposure categories for
911 BDCAA, DBAA and TBAA (-31 to -49 g) exposure which were only examined in the
912 brominated study site. A related community-based study by Horton et al. (2012) that was
913 largely null for mean BWT but provided one of the most comprehensive evaluations of
914 different exposure comparisons based on the underlying study populations examined by
915 Hoffman et al. (2008). For example, they reported a non-statistically significant BWT
916 reduction for HAA9 levels greater than the 90% percentiles (vs. referent) among the
917 chlorinated site which had a larger exposure contrast (-33 g; 95%CI: -85, 20 g) as well as
918 consistent small associations for the upper exposure categories for both chlorinated (–10
919 g) and one more brominated (-18 g) study sites. Other studies have shown association for
920 aggregate MX estimates (-18 g, Wright et al., 2004) and urinary chlorophenols (-31 to 38
921 g) as well as BWT reductions similar in magnitude across different aggregate mixture
922 measures such as TOX (-40 g; Hoffman et al., 2008) and DBP9 (-39 g; Rivera-Núñez and
923 Wright, 2013). One of the larger birth cohort studies did not show adverse associations
924 for THM4 and mean BWT in their pooled analysis (Kogevinas et al, 2016), but two of
925 their cohorts (Grazuleviciene et al., 2011; Smith et al., 2012) showed various associations
926 for different THM4 exposure metrics. Although they did not see any association with
927 HAA3 (sum of DCAA, TCAA, and BDCAA), Smith et al. (2012) reported a reduction (-
928 45 g) based on DBP7 (sum of THM4, DCAA, TCAA, and BDCAA) based on aggregate
929 concentration data only. Although results are not consistent across different DBP metrics,

these data generally suggest that there are some associations present in many drinking water systems.

Previous studies of aggregate residential concentration estimates have shown mixed results between HAA5 exposures and mean BWT, with two not showing any adverse associations (Wright et al., 2004; Hoffman et al., 2008). Unlike biomarker and uptake estimates, which measure the exposure level directly and individually, aggregate exposure estimates are based on the assumption that infants born on the same date to residents served by the same PWS have comparable levels of exposure. A more recent study neonates in Massachusetts neonates from 1996 to 2004 (Rivera-Núñez and Wright, 2013) showed statistically significant reductions between mean BWT and all HAA exposure categories (-28 to 36 g).

Compared to HAA5 results in our study, smaller non-significant mean BWT reductions were seen for HAA9 (-4 g; 95%CI: -10, 2), DBP9 (-5 g; 95%CI: -9, 0), and DBP13 exposure categories (-5 g; 95%CI: -9, 0). These data suggest that there was not much increased risk among these more comprehensive mixture measures nor the levels of brominated HAAs examined here. However, the previous toxicology studies suggested the higher potential health risks in brominated HAA, even though without providing reliable endpoints (Wang et al., 2017), seemed to disagree with our observations. We believe the limited HAA-Br concentration gradient in our DBP samples might have prevented the detection of associations between mean BWT and brominated HAAs due to the lack of enough sensitivity.

.

#### 5.4.1.2 Small for Gestational Age

Meta-analyses by Grellier et al. (2011) and by Summerhayes et al. (2020) have shown a small but consistent increase risk of SGA relative to DBP exposures based on aggregate-level concentration data. These meta-analyses were limited to studies using aggregate DBP concentrations data only. The evidence is mixed in those studies with more direct exposure measures although the literature is difficult to interpret, especially for biomarker-based studies, as some of the DBP species are rapidly metabolized. Elevated risks in SGA (5%) with THM4 exposure were detected in previous studies by Costet et al., (2012; OR = 1.80; 95%CI: 0.69, 1.75) and Summerhayes et al., (2012; OR = 1.10; 95%CI: 1.03, 1.18). Increased risks in SGA (10%) with THM4 were also reported by multiple studies with ORs between 1.03 to 1.30 with a pooled estimate by Summerhayes et al. (2021) of 1.05 (95%CI: 1.01-1.09). Increased in SGA (10%) with HAA5 exposure were reported by Horton et al., (2011; OR = 1.07; 95%CI: 0.62, 1.94), Levallois et al., (2012; OR = 1.40; 95%CI: 1.04, 1.88), and Rivera-Núñez and Wright (2013; OR = 1.04; 95%CI: 0.92, 1.17). Decreased risks of SGA (10%) were reported by Hoffman (2008; OR = 0.90; 95%CI: 0.50, 1.61). Hoffman et al. (2008) reported null associations between SGA (10%) and HAA5 or HAA9 in a study conducted on 2,766 pregnant women in three US communities.

In this study, we have detected increased SGA (5%) risks for the highest third-trimester TCAA (OR = 1.08; 95%CI: 1.02, 1.15), HAA5 (OR = 1.09; 95%CI: 1.02, 1.16), and HAA9 (OR = 1.06; 95%CI: 1.00, 1.14)) categories in multivariate and multi-DBP models adjusting for THM4 (Table 19). Similarly, higher risks in SGA (10%) were

noticed in HAA5 exposure with fully adjustments. Null associations were found between brominated HAA metrics (i.e., HAA-Br) and SGA (5%) or SGA (10%), given the less sensitivity from the relative narrow exposure range for HAA-Br (90%tile = 11 µg/L) compared with HAA5 (90%tile = 38 µg/L). Interestingly, compared with brominated HAA metrics, increased SGA (5%) risks were detected in exposure metrics with both chlorinated and brominated HAAs (i.e., HAA9).This is likely attributable to the commonly higher occurrence of fully chlorinated species (i.e., TCAA and DCAA) and less commonly occurred brominated species (i.e., HAA4 or HAA-Br).

We also evaluated the association between DBP exposure and both SGA (5%) and SGA (10%) to help discern whether smaller percentiles, like SGA5%, are less prone to misclassification which would decrease bias potential and increase study sensitivity. The result showed slightly higher ORs in SGA (5%) than in SGA (10%), which if true, may suggest that SGA (5%) is better capturing more of the pathologically growth retarded neonates. This could decrease the study sensitivity and potential for false negatives (i.e. Type 2 Error) in these semi-individual epidemiological studies like ours for large populations with considerable precision.

#### 5.4.1.3 Preterm Birth

Various associations between PTB and THM4 were reported. Seven out of ten publications reported inverse associations between PTB and THM4. A recent study in Sweden (Säve-Söderbergh et al., 2020) showed an inverse associations for PTB and THM4 (OR=0.82; 95%CI: 0.69, 0.98). Studies from Massachusetts (OR range: 0.85 to

997 0.92, Lewis et al., 2007; OR range: 0.88 to 1.00, Wright et al., 2003, 2004)) also reported
998 inverse associations for THM and other metrics. Similar inverse associations were
999 reported by Hoffman et al. (2008; RR range: 0.50 to 0.90), Gallagher et al. (1998; OR
1000 range: 0.7 to 1.0), and Dodds et al. (1999; OR range: 0.96 to 0.99). Yang et al. (2007) and
1001 Kramer et al. (1992) were the only two studies that showed any evidence of increased
1002 risks of PTB, although were all small in magnitude and not statistically significant
1003 (ORs=1.1).

1004 In our study, only second-trimester THM4 (OR = 1.05; 95%CI: 1.01, 1.09) and
1005 DCAA exposure (OR = 1.06; 95%CI: 1.01, 1.11) showed significant increased ORs in
1006 PTB after adjusting for confounding in multi-pollutant models. This is comparable to
1007 results reported by Rivera-Núñez and Wright (2013), in which they detected adverse
1008 associations for second-trimester THM4 exposure (OR range: 1.04 to 1.07), HAA5
1009 exposure (OR Range: 1.09 to 1.15), DCAA (OR Range: 1.04 to 1.10), TCAA (OR
1010 Range: 1.09 to 1.15), and DBP9 (OR Range: 1.20 to 1.29). The ORs from HAA9 and
1011 HAA5 were very similar and no association between them and PTB were detected. Null
1012 associations were detected for other exposures after fully adjustments. Similarly, the
1013 Horton et al. (2011) study did not detect associations for PTB but reported some evidence
1014 for increased risk for very PTB. For example, in the chlorinated site an elevated risk was
1015 seen in the highest HAA5 exposure category (OR=1.63; 95%CI: 0.94, 2.85). They
1016 reported larger risks in the brominated site (OR=4.17; 95%CI: 1.14, 15.32). They also
1017 reported increased risk for very PTB (OR=2.29; 95%CI: 1.01, 5.21) when examining the
1018 sum of all brominated DBPs in the brominated site. Wright et al., (2004) also showed
1019 adverse associations for TCA (OR = 1.33; 95%CI: 0.77, 2.30) and HAA5 when

examining very preterm infants (i.e., < 34 gestational weeks) and high trichloroacetic and HAA5 (OR =1.48; 95%CI: 0.84, 2.61) exposures.

#### 5.4.1.4 Birth Outcome Conclusion

In general, we found several null associations between DBP exposures and birth outcomes. This may be due to related to the considerable uncertainties as to the best exposure surrogate that best capture the most toxic species or combination of the DBPs. We proposed instead of using only brominated HAA surrogates (i.e., HAA4 or HAA-Br) as the exposures, a more comprehensive evaluation should be conducted with HAA9 or DBP13 as the exposures. Given the various HAA speciation across different PWSs, it is difficult to definitely isolate the association between adverse health outcomes and the individual DBPs, especially brominated HAAs, as the formation of each HAA species were not completely independent. The competition between chlorination and bromination during the halogenation of HAA changes the speciation and distribution of HAA species. With the same concentration of brominated HAA, the chlorinated HAA concentration (i.e., DCAA or TCAA) could vary tremendously (Ma and Reckhow, In Review). Thus, the almost negative associations between mean BWT reduction and HAA-Br or HAA4 were noticed in this study as chlorinated HAAs were not adjusted for. Even though we noticed higher mean BWT reduction in DCAA and TCAA exposures, these predominant HAA species in most PWSs with low bromine content in water source, likely swamp or make it difficult to differentiate risks from less prevalent DBPs. Although at times challenging due to collinearity concerns, adjustment of these predominant HAA species in HAA-Br models and evaluation of other HAA mixture models is needed.

As noted previously, the narrow range of the predicted brominated HAAs and use of indirect aggregate exposure measures subject to considerable measurement error can lead to misclassification may limit our ability to differentiate risks across different metrics especially when the concentrations didn't vary enough to move participants across categories. Similarly, we did not notice material differences for DBP9 and DBP13 but combining the volatile and non-volatile DBPs together may ultimately decrease our study sensitivity if it merely adds "noise" to our data and make it more difficult to detect associations. Since, for example, HAAs may be better reflected (i.e. less subject to misclassification) if water concentrations more closely mimic typical ingestion exposures. Unfortunately, we believe the occurrence of brominated DBP species is too low across these Ohio PWSs to allow for more definitive conclusions to be drawn based on these analyses, especially since we were not able to weight these mixture measures by anticipated toxicity. Future toxicological studies should be designed to address this limitation in DBP related epidemiological studies to reduce the uncertainty related to these exposure metrics and DBP mixture surrogacy challenges (see section 5.4.3 for additional discussion).

### 5.4.2 Strengths and Limitations of the Study

#### 5.4.2.1 Study Design/Analysis

The design of our retrospective cohort study offered some key strengths including the ability to examine not only multiple exposures, but several endpoints related to fetal growth restriction and gestational duration. The long study time period included an ample

number of years to allow for sufficient statistical power to detect associations that have been shown in the literature to be small in magnitude. It also allowed us to examine stratified results for some sociodemographic factors which is important given anticipated heterogeneity in risks across different groups. Similar to the main effects for mean BWT and THM4 exposures, the race-specific results for the highest exposure category were largely null. We did detect larger mean BWT differences for THM4 exposures among Asians (-49 g) and Native Americans (-96 g), although these results were quite imprecise results given the small sample sizes relative to other groups (Table S10).

#### 5.4.2.2 Outcome Data

A strength of our study was the accuracy of the health endpoints examined such as mean BWT and SGA which are based on very precise birthweight measurements and should minimize the potential for outcome measurement error and misclassification. There, however, is some potential outcome misclassification of SGA and PTB endpoints due to measurement error anticipated for clinical estimates of gestational duration. Clinical estimation of gestational age is based on the last menstrual period provided by the mothers and further estimated by early ultrasound when available according to the US Standard Certificate of Live Birth. As documented elsewhere (David 1980; Gjessing et al., 1999; Martin et al., 2015), both of gestational age dating approaches are subject to some error and may not be completely accurate. Although some measurement error is anticipated, it is expected to result in non-differential misclassification (i.e. not related to exposure assignments) for the dichotomous endpoints.

An additional study strength which should minimize the potential for selection bias was the comprehensive nature of the statewide vital records data as they are collected from different hospitals and areas in Ohio and maintained by ODH. We had to restrict the population to 81% of the total births occurring during our study period largely due to missing or ambiguous DBP data. Although we saw some evidence of differences in maternal race (91% vs. 72% of White) and median household income (22% vs. 8% in lowest group) in non-participants (Table S3) compared to participants, our data showed comparable mean BWT across these groups by participation status. As with other studies restricted to live births, it is not entirely clear if this restriction would result in bias (Raz et al, 2018). The likelihood of this type of bias in DBP studies seems minimal, since there is limited evidence of associations between DBP and fetal loss in most studies published to date. Therefore, given this and the large sample size and relatively small proportion of excluded participants, we expect that our results are not likely prone to selection bias.

#### 5.4.2.3 Confounding

Potential bias due to confounding is a challenge for environmental epidemiological studies given their observational study design. A strength of our retrospective cohort study is the ability to examine many potential confounders given the wealth of individual-level covariates available on birth records in Ohio. These following covariates were adjusted for in the BWT and PTB multivariate models: maternal age, race/ethnicity, education, smoking, parity, adequacy of prenatal care, delivery source of payment, income, marital status, pre-pregnancy BMI, weight gain during pregnancy, season of birth, and interpregnancy interval. Although several associations were denoted in the

unadjusted models, results were often attenuated and largely null following statistical adjustment. Based on a change-in estimate evaluation, we found that maternal race, education, smoking, and weight gain during pregnancy were the strongest confounders based on the mean BWT and THM4 model. Although we did developed additional THM models adjusted for HAA5 and HAA models adjusted for THM4, we saw minimal differences following this adjustment. In general, it remains unclear whether multipollutant models are warranted given the potential for amplification bias from highly correlated exposures (Weisskopf et al., 2018; Rivera-Núñez and Wright, 2013).

One concern is that some of the covariates that we examined may be subject to measurement error if based on maternal recall at time of birth (e.g. amount of smoking during early pregnancy). Previous studies reported high reproducibility and validity of maternal recall of pregnancy-related events even after 30 years and thus this is not anticipated to be a large source of error (Tomeo et al., 1999). Other covariates like maternal alcohol use during pregnancy were not collected as part of the vital records in Ohio. Similar to many other vital-records based studies, a limitation is that we did not have data on were diet and nutritional practices during pregnancy or reliable information on paternal education and paternal and second-hand smoke data.

Capturing the socio-economic contribution to fetal growth and gestational age is challenging, especially since many related variables can substantially overlap such as maternal education level, household income, marital status, and payment source for delivery. Many of these socioeconomic factors can be markers of access to health care, prenatal care, and nutrition as well other healthful behaviors (e.g., physical activity) of expectant mothers and fathers. We believe that adjustment for household income and

maternal education and other socioeconomic status measures examined here should largely capture the potential for confounding.

An additional study strength is that the birth records contain a small proportion of missing data in the covariates that we considered as confounders. Among the covariates we examined, maternal weight gain during pregnancy had the largest missing data (8.7%). We examined the distribution of missing data in all potential confounders (e.g., weight gain during pregnancy, maternal smoking, adequacy of prenatal care visit, etc.) and found no systematic differences in mean BWT and other measurements in participants with missing covariates (data not shown). These low proportions of missing data should minimize the potential for both selection bias as well as residual confounding that may result from statistical adjustment in the multivariate regression models.

#### 5.4.2.4 Exposure Assessment

The extensive statewide exposure data collected in this epidemiological study and estimated per our modeling allowed us to evaluate many individual DBPs and summary surrogate measures of DBP mixtures such as THM4, HAA5, HAA-Br, HAA4, HAA9, and DBP13. The varied exposure contrasts not only increased study sensitivity (i.e. ability to detect an association that may exist), it also enabled us to address limitations in previous studies which should minimize the potential for exposure misclassification. This includes consideration of multiple DBP summary surrogates, the use of low-exposed referent populations, and examination of multiple exposure categories (up to 6 categories) to assess exposure-response relationships. This also the first epidemiological study of

developmental or reproductive outcomes to examine DBP13, the second to examine DBP9, and only the third study to examine HAA9.

Our HAA9 measure was based on a recent publication of a kinetic prediction model to supplement exposure assessment efforts when only HAA5 samples are available (Ma and Reckhow, In Review). As noted previously, we used routinely-collected THM4 and HAA5 exposure data as inputs for the prediction of HAA9 concentrations. The prediction largely relied on the accurate measurement of individual THM and HAA species, but is considered accurate and robust across a variety of water quality parameters and some treatment options (Ma and Reckhow, in review). The kinetic predictive model is considered most accurate for chlorine disinfected water and average DBP levels in chloraminated PWSs could be overestimated. The kinetics of HAA and THM formation during chloramination are characterized by an initial rapid period of formation followed by a period of slower formation. However, treatment plants tend to have a significant period of free chlorination prior to ammonia addition for purposes of meeting the disinfection requirement and the DBP formation in this period is significantly more than the later period (Pope, 2006; Hua and Reckhow, 2007, 2008, 2012; Doederer et al., 2014; Manivannan and Borisover, 2020). We would anticipate that measurement error may be increased in chloraminated systems examined here due to this uncertainty in the formation, but we are still confident that the model estimates are able to capture the highest DBP formation potential for the unmeasured HAAs. Even though the predicted system average values may be lower in these PWSs, the low and high extreme values should be able to be differentiated in our categorical analysis. As described in Chapter 4, we also noticed seasonal differences in PWSs that may over- or under-predict the BCAA

and BDCAA concentration. These systematic errors are anticipated to be minor, however, they may still weaken our ability to detect any associations that may be present. Another potential source of uncertainty is the use of private well users who were assigned with zero for DBP concentrations. Since, private well users could potentially be exposed to other chemical constituents that may differ from public water drinking water systems and could potentially lead to adverse birth outcomes (Nielsen et al., 2001; Schempf and Strobino, 2008; Bradley et al., 2021). To the extent that other chemical contaminants could occur in either surface or groundwater sources, they would have to be associated with both DBPs and the health outcomes to serve as confounders.

Given the indirect nature of the DBP exposure estimates, there are many sources of measurement error that can add to uncertainty in our exposure and risk estimates, including spatial, temporal, and inter-individual variabilities. We calculated spatial averages across a large aggregate area (i.e., area served by one PWS) by ZIP code and assigned these exposure scores to each study participant in the area without distinguishing residential from non-residential (e.g. workplace) DBP concentrations nor the influence of specific water use activities. We also assumed potential risks among pregnant women having comparable levels of DBP exposure for birth occurring in the same birth week can be represented by the DBP concentrations in drinking water, regardless of different contribution of specific exposure routes (i.e., dermal, ingestion, and inhalation). One limitation of this assumption is the inability to characterize and integrate inter-individual water use behaviors based on the available aggregated DBP data. Similar to some earlier studies, we attempted to use finer geocodes (e.g., longitude and latitude) as the identifier and using spatiotemporal modeling to assign more specific

sampling data to individual study participants. However, the lack of specific sampling location information in most of the PWSs precluded development of more refined residential-based estimates. Another potential source of uncertainty is due to mobility of mothers during the pregnancy, of which our address at birth precluded examination of. Previous studies indicate that the mobility of mothers can impact the accuracy of exposure estimates when linking aggregate DBP averages to different exposure windows. This would only be a concern if these expectant mothers move to locations that had different water systems. Maternal mobility often occurs less frequently in later pregnancy and birth outcomes dependent on the 3$^{rd}$ trimester exposure may be less misclassified (Schulman et al., 1993). If this holds true for our study population, PTB risks associated with the 2$^{nd}$ trimester exposure data may be more subject some additional misclassification.

In the limited available research, there is some evidence that spatial variability might contribute the most to the uncertainty among analytical measurement, temporal variability, individual water use activities, and other sources of measurement error (Symanski et al., 2004; Luben et al., In Progress). As noted earlier, spatial variation in DBP concentrations could be large in some PWSs given complicated reaction conditions in distribution systems and intermediate facilities (e.g., pump stations and storage tanks). It has been demonstrated that spatial variability can be considerable especially for large systems that use chlorine as a secondary residual disinfectant and those with long residence times or dead spots that can result from decreased water usage/volume (Lahlou, 2002; Pereira et al., 2004; Liu and Reckhow, 2013). For example, the biodegradation of HAA is more sensitive to temperature and THMs may volatilize and decrease somewhat

upon movement and mixing of water during transmission (Pavon et al., 2008; Zhang et al., 2009). To the degree that the regulatory samples capture the range of concentrations that are averaged, our exposure assessment approach is anticipated to characterize extremes of exposure ranges (e.g. low, intermediate and high categories) and capture extreme contrasts across PWSs.

Temporal variability is another source of potential exposure measurement error and misclassification. This can occur if DBP formation extremes that can impact averages are not fully captured by routinely-collected data or other modeling estimates. Our exposure assessment approach based on quarterly averaged DBP concentrations may ignore the temporal variance if peak exposures are averaged or diluted out across a critical time-window during the pregnancy. In this study, we could not examine DBP exposure data with shorter intervals in most study sites due to the sparse sampling frequency. If the sampling frequency in future studies can be increased to monthly or even weekly, this finer resolution over time could allow for examination and reduction of exposure misclassification (Parvez et al., 2011).

Another exposure assessment challenge is having few available DBP surrogates for complex mixtures. The unregulated brominated HAAs are rarely evaluated in epidemiologic studies due to the trace level concentrations in many PWSs and due to the lack of monitoring data in most countries (Wang et al, 2015). Although no large or consistent associations were detected between individual brominated HAA species and adverse birth outcomes, we observed some statistically mean BWT reductions and ORs for SGA with both HAA9 and DBP13 exposure. The moderate Spearman correlations we detected between THM4 and DBP13 ($r = 0.56$) and between THM4 and HAA9 ($r = 0.59$)

suggest there may be need in examining more unique exposure surrogates that are not captured in routinely-collected monitoring data or considered in most epidemiologic studies. Examining multiple summary mixture methods such as THM4, HAA9, and DBP13 should allow for targeting of more specific combinations of toxicologically relevant DBPs. Given the dearth of available DBP data, DBP formation A predictive models, such as the kinetic model we used in this work, could provide some insight into these not often measured DBP species and enable more epidemiological studies to expand their scope. As the toxicities of each DBP species (i.e. components in a mixture) are not the same and may also vary across different endpoints, DBP surrogates should be developed to better target toxicologically relevant mixtures.

### 5.4.3 Exposure Surrogate Metrics

Admittedly, the most comprehensive DBP exposure assessment approach would be quantify risk from each and all DBP species. Resource limitations make this nearly impossible as over 700 DBPs have been identified, including a large proportion of these that have not been quantified in drinking water from both chlorine and chloramine disinfection (Richardson et al., 2007, Kristiana et al., 2020). As such, exposure assessment challenges will continue to be the main barrier for epidemiological studies of DBPs. As noted earlier (see Section 4.2 Strengths and Limitations), there is also increasing concern that the surrogate measures currently examined may not be sufficient that the occurrence of DBPs vary in mixtures of different compositions across PWSs, given various DBP precursors in raw water and different treatment methods (Summers et al., 1993; Hua and Reckhow, 2012; Villanueva et al., 2012; Kaufman et al, 2020; Hua and Reckhow, 2013; Mao et al., 2014; Jiang et al., 2015; Fang et al., 2018; Jiang et al., 2019).

The key of selecting proper surrogates is whether the chosen exposure metric can reflect the health risks caused by the DBPs representative of a certain area. Given the fact that the ability to detect a signal from the brominated DBPs is obviously stronger in more brominated systems, the key question is what constitutes sufficient contrasts needed for each DBP metric. One example of increased contrasts for summary DBP measures was shown by Horton et al. (2011) that showed some birthweight deficits were noticed in their upper two brominated DBP sum categories in a more brominated system (Range: -18 to -37 g vs. Range: -10 to -19 g). Other epidemiological studies to date that focused on adverse birth outcomes and brominated DBPs, such as THM-Br, are also often conducted

in less brominated water systems. This motived the attempt to use other type of approaches for higher study sensitivities, such as chemical structure similarity approaches (e.g., TOX, BSF, HAA-Br, THM-Br) and regulatory-based summary concentration approaches (e.g., THM4 and HAA5). Other less commonly used approaches includes physicochemical property similarity approaches (i.e. similar fate and transport considerations in environment), toxicokinetic similarity approaches (e.g., similar exposure routes), and toxicological similarity approaches (e.g., similar LD50 in vitro/vivo bioactivity). Dimension reduction techniques and other computational tools are developed by needs to estimate chemical toxicity, internal translate bioactivity, or external exposure pathways of data-poor DBPs (Hubal et al., 2019). These types of commonly used tools include principal components analysis (PCA), quantitative structure activity relationship (QSAR), Bayesian modeling, and physiologically-based toxicokinetic (PBTK) models (Feder et al., 2009b; Braun et al., 2014). Dimension-reducing statistical approaches (e.g., PCA, penalized modeling based on elastic net regression, etc.) are increasingly being used for screening large groups of chemical exposures and help to prioritize specific mixtures for further analysis. However, as noted by Meng et al. (2018), these approaches might be better suited as ‘prediction models to screen for a wide range of chemicals from different sources, and the interpretation of results might become less straightforward due to the necessary standardization of exposure values.’ Given these interpretation difficulties and potential for co-exposure amplification bias, the challenge of DBP mixtures for exposure assessment purposes remains.

The most commonly used approach in this field is grouping individual DBP species into THMs and HAAs, according to common monitoring requirements for THM4 and HAA5. For example, in the US, MCL-based grouped exposure measures which have regulatory import are often the only DBP surrogates included in epidemiological studies. Chemical structural similarity approaches usually group DBPs into one category by similar structure (e.g., number of halogenation substitution: HAA-Br and THM-Br; di- and tri-halogenated grouping; Krasner and Wright, 2005) to increase the exposure contrasts. This grouping approaches could also use summarized characteristics to distinguish exposures, such as using TOX for formation and BIF/BSF for speciation. TOX measures a large summary concentration metric, but similar to most common DBP surrogates, it does not incorporate or weight the relative toxicities from different DBPs and does not examine the unaccounted TOX contribution (i.e., UTOX, Hua and Reckhow, 2007). As noted in Chapter 2 and 3, BSF describes the relative proportion of all brominated species within one DBP class and only evaluates the level of bromination in this DBP class... We also used BSF to further characterize risks from the brominated proportions but results were largely null (Data not shown). Using BSF as the proxy separately does not provide much information as it does not account for the absolute concentrations or toxicities of each individual DBP species (Hinckley et al., 2005). Also, the relative narrow range of BSF, 0 to 2 or 0 to 3 for dihalogenated and trihalogenated DBPs, and the high occurrence of low BSF (e.g., usually below 1 in bromine poor area) noticed in our previous chapters suggest that the lack of contrast can limit its utility as a DBP proxy.

Toxicokinetic similarity approaches assemble and compare available information pertaining to the absorption, internal distribution, metabolism, and elimination (ADME) when data are available. This can be related to volatility of chemicals which dictate specific exposure routes for certain DBPs. Physicochemical property similarity approaches, which partly overlapped with toxicokinetic similarity approaches, usually group DBPs by the volatility or solubility-based exposure routes in human. These two approaches were considered as useful in certain scenarios (e.g., exposure estimated from showering-based duration measures (Silva et al., 2013) but only partially accounts for all exposure routes (). Toxicologically similar approaches, such as relative potency factors (RPF), use in vivo/vitro or computational toxicities and concentrations of studied DBPs to create unique exposure mixture metrics, such as ToxCast/Tox21 (Bull et al., 2009a, 2009b; Rice et al. 2009; Hubal et al., 2019). Treatment driven approaches group DBPs by the type of the applied disinfectant (e.g., chlorine or chloramine) may be important for differentiating key DBP formation across disinfectant groups. However, they are much less specific to identify important DBP surrogates that would allow actionable efforts by regulators.

Among these less common approaches, toxicologically similar approaches evaluate the contribution of individual DBPs by more directly considering toxicity data, compared with other concentration-based approaches. However, these toxicity data are often not available for comprehensive exposure evaluations for endpoints such as developmental effects. Hence, studies in this area often rely on regulatory-based summary concentration approaches and chemical structure similarity approaches, especially MCL-based, due to

its availability, even though these approaches may not be grounded well in toxicity data or the most relevant mixtures.

Statistical reduction or computational tools have been rarely used for the prediction of unknown toxicities of DBPs or the concentration of unmeasured DBPs. Villanueva et al. (2012) applied PCA to 233 water samples in Spain and suggested that THM-HAA mixture can be simplified to three main components that correlated with brominated DBPs, chlorinated DBPs, and chlorinated HAAs, respectively. These three components explained 80% of the variance but still have limited value as predictors of other DBPs. PCA uses simplified predictor variables that includes many individual exposure parameters to explain the exposure variabilities with multicollinearity but have limited ability to take known relationships of chemicals into consideration (Feder et al., 2009b; Govarts et al., 2016). Bull et al., (2009a, 2009b) suggested that variables, including disinfectant type, halogen concentrations, pH, temperature, and changes in the distribution system could all qualitatively and quantitatively affect the DBP formation, DBP speciation, and, subsequently, the toxicity of the mixture. However, their analysis suggested that more systematic studies on toxicology and chemical characteristic of DBP would be needed to complete such evaluation. QSAR models and other comparative quantitative toxicology models relate toxicities to physicochemical properties for prediction purposes (Qin et al., 2017; Wei et al., 2020). The usage of these tools have helped advance the DBP exposure field, but the broader application of them would still require more experimental toxicity data on DBPs, which is often not available currently.

Despite the advances noted above, a lot remains unknown on which DBP components or mixture measures have the most developmental toxicity. This may, in addition to the

low exposure contrasts, explain some of the inconsistency in exposure-response relationships reported in previous epidemiology studies. One of the challenges that preclude some comparisons across studies is due to inconsistency in DBP co-occurrence as reflected in the variable correlations between DBP species. Thus, to continue to enable comparisons with earlier research and opportunities to pool data for meta-analytical purposes, the commonly used surrogates, such as THM4 and HAA5, should continue to be examined in conjunction with other targeted individual DBPs or surrogates to allow for more definitive answers about characterizing human health risks. .

Correlation between individual DBP species and specific grouped classes are quite various across water systems and depend on treatment type, source water, and other characteristics. This is the key limitation of using summary concentration approaches for exposure assessment given that the DBP surrogates could represent completely different exposure mixtures in different study area or time period. We noticed a relatively high correlation between THM4 and HAA5 (r = 0.62) while other researcher reported even higher correlations (r = 0.90, Krasner et al., 1989; r = 0.78, Bond et al., 2011). Similar to other studies like Levallois et al. (2012), we observed very strong correlations between HAA5 and HAA9 (r = 0.99); these relationships may be less strong and offer a more unique exposure opportunity if PWSs with increased bromine concentration. The very high correlations between HAA5 and HAA9 suggested that in distribution system with low bromine concentration, HAA9 exposure might just be a marker of chlorinated HAAs, not brominated HAAs. For example, Levallois et al. (2012) reported comparable increased risks of SGA for TCAA (OR = 1.4; 95%CI: 1.0, 1.8), HAA5 (OR = 1.4;

95%CI: 1.0, 1.8), and HAA9 (OR = 1.4; 95%CI: 1.0, 1.8) exposure quartiles and no individual brominated HAA species was examined. The similar risks for HAA5 and HAA9 exposure suggest that these grouped HAA concentrations may not be sufficient to differentiate risks presumably due to the low formation of brominated species in the collected records. We also examined the occurrence of brominated DBPs in studies and detected similar narrow ranges of concentration indicating that in the Ohio PWSs that even grouped DBP exposure may still not provide enough contrast. Unlike the relatively consistent reported correlation between THM4 and HAA5, various correlation between other DBPs, such as HAN4 and HAA9 (r = 0.83, Bond et al., 2011), calls into question how the use THM4 or HAA5 as sole proxies of comprehensive DBP exposures. Ileka-Priouzeau et al. (2015) reported very poor correlations between THM4 and HAN4 (r = 0.3), THM4 and haloacetaldehydes (r = 0.2), and HAN4 and haloacetaldehydes (r = 0.4). Based on some limited sampling of non-regulated DBPs, Wright et al. (2002) reported weak correlations between MX and THM4 (r = 0.44) and between MX and HAA5 (r = 0.35). While these smaller corrleations allow for unique metrics to be examined, the source- and system-specific differences along with time-dependent NOM content makes it challenging to target similar DBP component proportions when comparing many systems. This lack of comparability complicates weight of evidence considerations when examining differences within and across studies.

As noted above, the commonly used concentration sum based DBP metrics in epidemiological studies are limited in ability to fully evaluate the risks of less correlated DBPs that may be deleterious, such as HAA-Br, HANs, and hydroxyfuranones (Christian et al., 2001a; Christian et al., 2001b; Bove et al., 2002; Klinefelter et al., 2004;

Richardson et al., 2007; Wigle et al., 2008; Ileka-Priouzeau et al. 2015). The exposure risks evaluation based on the commonly measured DBP metrics, such as THM4 or HAA5, cannot account for the exposure risks from all DBPs. This may also increase the noise in the study and potentially lead to the exposure misclassification and provide inadequate information to fully characterize risks (Teuschler and Simmons, 2003). The further application of TOX may provide more information as TOX contains three sub-groups, total organic chlorine (TOCl), total organic bromine (TOBr), and total organic iodine (TOI). Even though within each sub-groups the relative toxicity is hard to estimate, however, the comparison between three sub-groups can provide some information regarding the risks from different halogenated DBP classes and the control of halogenated DBP formation. Thus, there may be a need for studies that examine measures of complex mixtures to assess different risks across chemical sub-sets for evaluation in epidemiological studies. A more relevant approach, regarding both the formation and speciation of DBPs, is the weighted TOX. Weighted TOX approach is the weighted summation of all TOX sub-groups based on the relative toxicity. This is a compromise to numerous DBP species and their toxicities by assuming that DBPs with the same halogen substitution have the same level of toxicity. This might cause certain level of uncertainty, but it could still better describe the total toxicity than using TOX alone.

Except for using weighted TOX, the most essential research needs is additional developmental toxicity studies to better allow for more relevant toxicity-based approaches to capture the relative toxicity from target compounds (i.e. DBP mixture components). Toxicity-based summary approaches, based on the assumption that dose

additivity is feasible, would account for both concentrations of target individual DBP species and the relative or absolute toxicity of these species. As an earlier study of bladder cancer risk attempted this based on relative potential weighting to compare to a concentration-based approaches that are commonly used for different DBP classes (Salas et al, 2013). In this example, the total exposure of four THM species should not just be calculated as the summation of the concentrations of four THM species but would be based on the summation of the unit toxicity of each THM species times their concentrations. Unfortunately, the developmental toxicity of certain DBP species on specific endpoints is not often known (Wright et al, 2017; Kaufman et al, 2018; Kaufman et al, 2020). As more research is conducted, this approach can be iterative and modified by using the relative toxicity of each DBP species. Then the weighted toxicity would then act as a surrogate of the relative toxicity of each DBP species and allow for a more relevant toxicity index for the regulated THMs and HAAs (Itoh et al., 2011; Falandysz et al., 2014; Postigo et al., 2018; Suzuki et al., 2020). Future epidemiological studies could benefit from weighted-toxicity approaches if toxicological data can help prioritize individual or groups of DBPs that can increase the specificity of the exposure assessment.

# CHAPTER 6

# CONCLUSION

The halogen speciation of DBPs is largely controlled by competitive reactions of the two principal halogenating agents (i.e., free chlorine and oxidized bromine) with the organic precursors; a competition that is common to both the THMs and HAAs. By measuring just two THM species, we were able to characterize this competition for the di- and tri-halogenated HAAs with a relative high level of accuracy.

Given uncertainties in chemical analysis, mathematical predictions of the seven brominated DBP species using the adjusted binomial model may be nearly as accurate as actual measurements using EPA approved methods. These mathematical models require a direct measurement of a single-class DBP couple (e.g., TCM and BDCM) for characterization of the degree of bromination, as well as a single species measurement from the target group (e.g., TCAA and DCAA). With these four measurements, the model is capable of accurately predicting formation of the remaining species (in this case 7 compounds: CDBM, TBM, BCAA, DBAA, BDCAA, CDBAA, and TBAA). It produces data of high accuracy which are not constrained by method detection limits.

Currently, compliance with regulatory MCLs requires direct measurements using approved methodologies performed by certified laboratories. As a result, predictions from this model cannot be used in lieu of actual measurements. Furthermore, field samples may not appear to conform this model when selective DBP losses are occurring. For example, use of spray aerators to lower THMs at distribution system hot spots will distort the THM speciation by reducing the concentrations of lighter chlorinated species more

than the brominated ones. Similarly, biodegradation of HAAs may occur unevenly across the various THAA or DHAA species. Finally, thermal degradation of HAAs in home water heaters will cause conversion of some HAAs to THMs, potentially distorting both distributions. The model we propose could provide a tool for confirmation when these potential confounding factors occur.

The further evaluation and the verification of the binomial model application on utility BCAA and BDCAA data proved the broad adaptability of the binomial model with complicated reaction conditions in the distribution systems. Admittedly we are unable to verify the model with other two unregulated HAA species (CDBAA and TBAA), however, given the higher occurrence of BCAA and BDCAA species compared with other two unregulated HAA species, we believe the model is reliable in predicting HAA9. We also evaluated the relative error in BCAA and BDCAA predictions. The median relative errors in each season ranges from 13-16%. Keeping in mind that performance evaluation tests (representing lab error) for analysis of HAAs is considered acceptable if it falls within ±20%, an overall error that includes field sampling, quenching, transport, storage and analysis should be higher. Thus our model error of 13-16% is not substantially larger and may not represent more than just the actual error the BCAA data to which the model being compared.

We have also used the predicted unregulated HAA data in the field of epidemiology for the estimation of their potential adverse effects on birth outcomes. This is the first epidemiological study that applied kinetic predictive model to the formation of brominated DBPs to examine the adverse birth outcomes in the US. Although result showed these DBP surrogates had fairly low impact on birth outcomes, we believe it is

more due to the low DBP exposure contrast in the study area. Nonetheless, we believe the kinetic model we applied in this research can be useful in current epidemiological studies, especially for those have available THM4 and HAA5 data. This study also discussed the potential advantages and disadvantages of the currently-used DBP surrogates. Most of the summary concentration based DBP surrogates may be limited to examine the accurate health risks from the exposure of brominated DBPs, given the low formation of them in most of the low bromine water sources. The correlation between DBP surrogates were also discussed in this study and that highly correlated chlorinated DBPs and total DBPs may mask the exposure risks from brominated DBPs.

In summary, application of the model can bring value and insight into regulatory measurement as follows:

1. It can help provide insight into missing data, or data that below detection limit;
2. It can help identify anomalous DBP data (e.g., due to laboratory error, sampling error, custody error, etc.);
3. It can be used to identify samples that are likely to have undergone some analyte loss, either through volatilization (especially for THMs; i.e., at point of sample collection, or at some upstream point) or through biodegradation (especially for HAAs);
4. It can provide new DBP exposure data for future epidemiological studies that focus on the adverse effect of DBPs on human health.

We recommend that certified laboratories make use of this tool as a QC check and notify the utility client when data seem to deviate from the expected halogen speciation.

Again, there are good reasons to expect deviation in some cases, but it would be helpful for the utilities to be aware that such a deviation exists.

Beyond use with compliance data, we believe this model offers an excellent tool for those conducting research and treatability studies. Aside from offering a check on laboratory analysis of DBPs, it presents the possibility of reducing the analytical burden. It also opens up the possibility of acquiring data for analytes that are present at below detection limits or for analytes that cannot be measured for other reasons (e.g., chromatographic interference, lack of authentic standard).

## DISCLAIMER

The views expressed in this manuscript are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Ohio Department of Health data used in this study were obtained from the Bureau of Vital Statistics, Ohio Department of Health (ODH). Use of these data does not imply ODH agrees or disagrees with any presentations, analyses, interpretations or conclusions.

# APPENDICES: SUPPORTING INFORMATION

## Detail of Derivation of the Reaction of Very Reactive NOM.

It is helpful to represent some of this in the form of simple kinetic rate laws and to parse NOM into reactive, very reactive [VReNOM] and unreactive components. Based on model compound studies the formation of organic chlorine by EAS from the very reactive aromatics should be a simple 2nd order process:

$$\frac{d[TOCl]}{dt} = k_1[HOCl][VReNOM] \qquad \textit{(Equ. S1)}$$

And since oxidation of bromide (reaction #3, Figure 10. Conceptual View of Chlorine-Bromide-NOM Reactions.) is the rate limiting step during the earliest phases of the reaction:

$$\frac{d[TOBr]}{dt} = k_3[HOCl][Br^-] \qquad \textit{(Equ. S2)}$$

So that the incorporation ratio during that first minute is:

$$\frac{d[TOBr]}{d[TOCl]} = \frac{k_3}{k_1}\left(\frac{[Br^-]}{[VReNOM]}\right) \qquad \textit{(Equ. S3)}$$

If bromide dissipates over the same timescale as the most reactive NOM, and the amount of very reactive NOM correlates with DOC:

$$\frac{[TOBr]}{[TOCl]} \approx \frac{k_3}{k_1}\left(\frac{[Br^-]_o}{[VReNOM]_o}\right) \sim \frac{[Br^-]_o}{DOC_o} \qquad \textit{(Equ. S4)}$$

This indicates that the degree of bromine incorporation should be proportional to the bromide to DOC ratio, an empirical observation that has been documented many times.

During later phases of chlorine contact, the bromide level may become substantially depleted of the $Br^-$/TOC ratio is low, and the amount of bromine incorporation (i.e., TOBr) may approach a value that is close to the initial bromide (i.e., f~0.9-0.99). In that case:

$$\Delta[TOBr] \approx f[Br^-] \qquad \textit{(Equ. S5)}$$

And now combining:

$$\frac{\Delta[TOBr]}{\Delta[TOCl]} \approx \frac{f}{k_1}\left(\frac{1}{\int[HOCl]t}\right)\left(\frac{[Br^-]_o}{[ReNOM]_o}\right) \sim \frac{1}{CT}\frac{[Br^-]_o}{DOC_o} \qquad \textit{(Equ. S6)}$$

Which implies that the degree of bromine incorporation decreases after bromide becomes largely depleted in relation to the chlorine exposure increases (i.e., chlorine CT).

**Explanation of Generic Equations**

The generic equation for each of the seven DBPs is:

$$RCl_aBr_b = KE\frac{RCl_{a+b}}{(Ratio_T)^b} \qquad \textit{(Eqn. S7)}$$

Where $RCl_aBr_b$ is the concentration of a DBP with the halogen-substituted carbon atom bonded to 'a' chlorine atoms and 'b' bromine atoms; and $RCl_{a+b}$ is the concentration of the fully chlorinated member of that same group (i.e., TCM, DCAA or THAA; each having 'a+b' chlorine atoms and no bromine atoms). This equation pertains to THMs, DHAAs and THAAs. It has not been tested for accuracy against the MHAAs. The value 'K' is simply an uncalibrated binomial model coefficient (i.e., 1/3, 2/3, 1/9 or 1/27). Since the model is based on molar concentrations, the K values must be adjusted with gram formula weights when input and output data are in mass rather than molar concentrations. The mass-based K values are listed in Table S1 and Table S2.

**BSF deviation**

An Alternative equation for BSF uses the alpha values (Equations S10a-S10d).

$$BSF_T = \frac{\alpha_{BrCl2}+2\alpha_{Br2Cl}+3\alpha_{Br3}}{3} \quad \textit{(Eqn. S8a)}$$

$$BSF_D = \frac{\alpha_{BrCl}+2\alpha_{Br2}}{2} \quad \textit{(Eqn. S8b)}$$

$$x = 1 - BSF \quad \textit{(Eqn. S9)}$$

$$\alpha_{Cl_3} = (1 - BSF_T)^3 \quad \textit{(Eqn. S10a)}$$

$$\alpha_{BrCl_2} = 3BSF_T(1 - BSF_T)^2 \quad \textit{(Eqn. S10b)}$$

$$\alpha_{Br_2Cl} = 3(BSF_T)^2(1 - BSF_T) \quad \textit{(Eqn. S10c)}$$

$$\alpha_{Br_3} = 3(BSF_T)^3 \quad \textit{(Eqn. S10d)}$$

**Table S1. Prediction of unregulated DBPs from the binomial kinetic model.**

| Target DBP | | Prediction | | | | |
|---|---|---|---|---|---|---|
| | | Molar Concentration | | Mass Concentration | | |
| Abbreviation | Formula | Equation | E | Equation | K | KE |
| CDBM | $CHClBr_2$ | $E\frac{TCM}{3(Ratio_T)^2}$ | 2.55 | $KE\frac{TCM'}{(Ratio_T')^2}$ | 0.309 | 0.788 |
| TBM | $CHBr_3$ | $E\frac{TCM}{27(Ratio_T)^3}$ | 4.86 | $KE\frac{TCM'}{(Ratio_T')^3}$ | 0.030 | 0.146 |
| BCAA | CHClBrCOOH | $E\frac{2DCAA}{3Ratio_T}$ | 1.16 | $KE\frac{DCAA'}{Ratio_T'}$ | 0.653 | 0.757 |
| DBAA | $CHBr_2COOH$ | $E\frac{DCAA}{9(Ratio_T)^2}$ | 4.23 | $KE\frac{DCAA'}{(Ratio_T')^2}$ | 0.100 | 0.423 |
| BDCAA | $CCl_2BrCOOH$ | $E\frac{TCAA}{Ratio_T}$ | 1.11 | $KE\frac{TCAA'}{Ratio_T'}$ | 0.927 | 1.029 |
| CDBAA | $CClBr_2COOH$ | $E\frac{TCAA}{3(Ratio_T)^2}$ | 2.49 | $KE\frac{TCAA'}{(Ratio_T')^2}$ | 0.273 | 0.680 |
| TBAA | $CBr_3COOH$ | $E\frac{TCAA}{27(Ratio_T)^3}$ | 3.78 | $KE\frac{TCAA'}{(Ratio_T')^3}$ | 0.026 | 0.098 |

**Note that: [1] DBP species are either in molar concentration (TCM, DCAA, and TCAA, μM) or in mass concentration (TCM', DCAA', and TCAA', μg/L). [2], Similarly, $Ratio_T$ is the molar ratio of TCM to BDCM while $Ratio_T$' is the mass ratio of TCM to BDCM.**

**Table S2. Determination of Model Coefficients K (K values are based on µg/L concentrations for all Seven DBPs[16]).**

| Target DBP | | Predictive equation | DBP-Specific Constant (K) | |
|---|---|---|---|---|
| Abbreviation | Formula | | K Value (µg/µg) | Formulation |
| CDBM | $CHClBr_2$ | $K\frac{TCM'}{(Ratio_T')^2}$ | 0.309 | $\frac{MW_{CDBM}MW_{TCM}}{3(MW_{BDCM})^2}$ |
| TBM | $CHBr_3$ | $K\frac{TCM'}{(Ratio_T')^3}$ | 0.030 | $\frac{MW_{TBM}(MW_{TCM})^2}{27(MW_{BDCM})^3}$ |
| BCAA | CHClBrCOOH | $K\frac{DCAA'}{Ratio_T'}$ | 0.653 | $\frac{2MW_{BCAA}MW_{TCM}}{3MW_{DCAA}MW_{BDCM}}$ |
| DBAA | $CHBr_2COOH$ | $K\frac{DCAA'}{(Ratio_T')^2}$ | 0.100 | $\frac{MW_{DBAA}(MW_{TCM})^2}{9MW_{DCAA}(MW_{BDCM})^2}$ |
| BDCAA | $CCl_2BrCOOH$ | $K\frac{TCAA'}{Ratio_T'}$ | 0.927 | $\frac{MW_{BDCAA}MW_{TCM}}{MW_{TCAA}MW_{BDCM}}$ |
| CDBAA | $CClBr_2COOH$ | $K\frac{TCAA'}{(Ratio_T')^2}$ | 0.273 | $\frac{MW_{CDBAA}(MW_{TCM})^2}{3MW_{TCAA}(MW_{BDCM})^2}$ |
| TBAA | $CBr_3COOH$ | $K\frac{TCAA'}{(Ratio_T')^3}$ | 0.026 | $\frac{MW_{TBAA}(MW_{TCM})^3}{27MW_{TCAA}(MW_{BDCM})^3}$ |

**Note that $Ratio_T$' in this table is the mass ratio of TCM to BDCM. This is different from $Ratio_T$ used previous in the model, which is the molar ratio of TCM to BDCM.**

[16] $Ratio_T$' refers to the mass ratio of TCM to BDCM measured in the sample; and TCM, DCAA and TCAA are measured mass concentrations (µg/L) for each of the three DBPs. When using the constants in Table 3, the predicted DBP concentrations will also be in mass units (µg/L).

**Table S3. Maternal and Infant Characteristics of Study Population (n=952,190 [1]) and the Excluded Population (n=133,922[2]) in the state of Ohio, 2006-2013.**

| | Included in the Study | | Excluded from the Study | |
|---|---|---|---|---|
| | Study Population, n (%) | BWT, g | Study Population, n (%) | BWT, g |
| **Count, n** | 952,190 (100.0) | | 133,922 (100.0) | |
| **Total Birth** | 952,190 (100.0) | 3,369 | 133,922 (100.0) | 3,412 |
| **Infant Sex** | | | | |
| Male | 486,578 (51.1) | 3,432 | 68,215 (50.9) | 3,479 |
| Female | 465,612 (48.9) | 3,303 | 65,707 (49.1) | 3,344 |
| **Maternal Age (Years)** | | | | |
| ≤20 | 142,144 (14.9) | 3,267 | 16,948 (12.7) | 3,321 |
| 21-25 | 253,433 (26.6) | 3,327 | 36,748 (27.4) | 3,371 |
| 26-30 | 278,287 (29.2) | 3,401 | 41,411 (30.9) | 3,435 |
| 31-35 | 192,344 (20.2) | 3,431 | 26,923 (20.1) | 3,466 |
| 36-40 | 73,629 (7.7) | 3,427 | 10,067 (7.5) | 3,476 |
| 41-62 | 12,329 (1.3) | 3,402 | 1,823 (1.4) | 3,457 |
| Missing | 24 (0.0) | 3,290 | 2 (0.0) | 2,915 |
| **Maternal Race** | | | | |
| White | 684,836 (71.9) | 3,411 | 122,098 (91.2) | 3,423 |
| Hispanic | 45,518 (4.8) | 3,355 | 2,390 (1.8) | 3,410 |
| Black | 179,599 (18.9) | 3,225 | 6,648 (5.0) | 3,253 |
| Native American | 4,774 (0.5) | 3,323 | 461 (0.3) | 3,355 |
| Asian Indian | 7,182 (0.8) | 3,200 | 437 (0.3) | 3,228 |
| Asian | 18,993 (2.0) | 3,292 | 1,228 (0.9) | 3,325 |
| Others | 9,601 (1.0) | 3,313 | 552 (0.4) | 3,319 |
| Missing | 1,687 (0.2) | 3,341 | 108 (0.1) | 3,278 |
| **Maternal Education** | | | | |
| Below high school/GED | 154,462 (16.2) | 3,255 | 24,201 (18.1) | 3,366 |
| High school graduate/GED | 242,709 (25.5) | 3,326 | 35,967 (26.9) | 3,366 |
| Some college, no degree | 212,922 (22.4) | 3,376 | 27,988 (20.9) | 3,418 |
| Bachelor/associate degree | 241,185 (25.3) | 3,447 | 34,204 (25.5) | 3,469 |
| Graduate or higher | 94,451 (9.9) | 3,446 | 11,110 (8.3) | 3,481 |
| Missing | 6,461 (0.7) | 3,307 | 452 (0.3) | 3,288 |
| **Marital Status** | | | | |
| Married | 527,649 (55.4) | 3,434 | 90,210 (67.4) | 3,455 |
| Unmarried | 423,421 (44.5) | 3,286 | 43,550 (32.5) | 3,324 |
| Missing | 1,120 (0.1) | 3,309 | 162 (0.1) | 3,312 |
| **Number of Previous Births** | | | | |
| 0 | 308,743 (32.4) | 3,340 | 40,982 (30.6) | 3,368 |
| 1 | 257,497 (27.0) | 3,389 | 36,467 (27.2) | 3,424 |
| ≥2 | 360,226 (37.8) | 3,382 | 54,469 (40.7) | 3,440 |
| Missing | 25,724 (2.7) | 3,332 | 2,004 (1.5) | 3,364 |
| **Pay Source of Delivery** | | | | |
| Public | 396,848 (41.7) | 3,292 | 43,943 (32.8) | 3,332 |
| Private | 505,852 (53.1) | 3,428 | 83,328 (62.2) | 3,455 |
| Others | 49,490 (5.2) | 3,367 | 6,651 (5.0) | 3,405 |

**(continued)**

**Table S3. (Continued)**

| | Included in the Study | | Excluded from the Study | |
|---|---|---|---|---|
| | **Study Population, n (%)** | **BWT, g** | **Study Population, n (%)** | **BWT, g** |
| **Maternal Weight Gain [3]** | | | | |
| ≤0 | 21,492 (2.3) | 3,265 | 2,627 (2.0) | 3,324 |
| >0-25 | 270,136 (28.4) | 3,279 | 38,440 (28.7) | 3,334 |
| >25-50 | 480,434 (50.5) | 3,392 | 71,558 (53.4) | 3,431 |
| >50 | 97,933 (10.3) | 3,511 | 13,318 (9.9) | 3,547 |
| Missing | 82,195 (8.6) | 3,370 | 7,979 (6.0) | 3,428 |
| **Income [4]** | | | | |
| ≤32,494 | 207,432 (21.8) | 3,278 | 10,202 (7.6) | 3,344 |
| >32,494-37,967 | 192,554 (20.2) | 3,358 | 24,510 (18.3) | 3,416 |
| >37,967-43,768 | 186,969 (19.6) | 3,383 | 33,586 (25.1) | 3,401 |
| >43,768-51,806 | 176,374 (18.5) | 3,407 | 34,890 (26.1) | 3,420 |
| >51,806-66,989 | 135,812 (14.3) | 3,426 | 27,550 (20.6) | 3,436 |
| >66,989 | 50,706 (5.3) | 3,436 | 3,045 (2.3) | 3,436 |
| Missing | 2,343 (0.2) | 3,387 | 139 (0.1) | 3,303 |
| **Season [5]** | | | | |
| Spring | 238,067 (25.0) | 3,376 | 33,434 (25.0) | 3,423 |
| Summer | 250,912 (26.4) | 3,370 | 35,650 (26.6) | 3,410 |
| Fall | 239,581 (25.2) | 3,367 | 33,782 (25.2) | 3,411 |
| Winter | 223630 (23.5) | 3,363 | 31,056 (23.2) | 3,405 |
| **Prenatal Care [6]** | | | | |
| Adequate | 552,335 (58) | 3,399 | 83,820 (62.6) | 3,431 |
| Intermediate | 306,640 (32.2) | 3,334 | 42,843 (32.0) | 3,385 |
| Inadequate | 29,873 (3.1) | 3,259 | 3,252 (2.4) | 3,373 |
| Missing | 63,342 (6.7) | 3,318 | 4,007 (3.0) | 3,354 |
| **Pre-pregnancy BMI** | | | | |
| ≤18 | 26391 (2.8) | 3,142 | 3,547 (2.6) | 3,168 |
| >18-24 | 382736 (40.2) | 3,325 | 55,420 (41.4) | 3,360 |
| >24-30 | 282861 (29.7) | 3,398 | 40,642 (30.3) | 3,450 |
| >30 | 214285 (22.5) | 3,438 | 29,605 (22.1) | 3,485 |
| Missing | 45917 (4.8) | 3,364 | 4,708 (3.5) | 3,431 |
| **Maternal Smoking [7]** | | | | |
| 0 | 777760 (81.7) | 3,403 | 109,771 (82.0) | 3,453 |
| 1-5 | 55268 (5.8) | 3,265 | 6,933 (5.2) | 3,310 |
| 6-10 | 62390 (6.6) | 3,196 | 9,123 (6.8) | 3,213 |
| >10 | 50881 (5.3) | 3,163 | 7,631 (5.7) | 3,169 |
| Missing | 5891 (0.6) | 3,281 | 464 (0.3) | 3,360 |

BWT, birth weight (gram).

[1]. Study population was restricted to singleton birth, gestational age between 22 to 45 weeks, BWT > 200 g, known infant sex, and served by either one PWS or private well.

[2]. Excluded population was defined as no specific DBP exposure data.

[3]. Maternal weight gain during pregnancy, in lb.

[4]. Median household income of the maternal residency, USD. From 2010 US Census.

[5]. Season was defined as the season when the delivery occurred (i.e., Spring, Summer, Fall, and Winter).
[6]. Adequacy of prenatal care was evaluated based on the Kessner Index.
[7]. Maternal smoking during pregnancy was categorized as cigarette smoking per day reported during the three trimesters, when applicable.

1623

1624

1625

**Table S4. A Sensitivity Analysis of Imputations for Missing Values and Below Detection Limit Values in Exposure Data.**

| THM4* Metrics (µg/L) | Term Births (n) | Unadjusted BWT(g) (95% CI) | Adjusted BWT(g) (95% CI) | Adjusted BWT(g) (95% CI) | Term Births (n) | Unadjusted BWT(g) (95% CI) | Adjusted BWT(g) (95% CI) | Adjusted BWT(g) (95% CI) |
|---|---|---|---|---|---|---|---|---|
| **Prior to Imputation** | | | | | **After Imputation** | | | |
| 0-15 | 36440 | REF | REF | REF | 80170 | REF | REF | REF |
| >15-25 | 97566 | -25 (-31 to -20) | 4 (-2 to 10) | 8 (1 to 14) | 162624 | -23 (-27 to -19) | 1 (-3 to 5) | 5 (0 to 10) |
| >25-33 | 125783 | -20 (-26 to -15) | 5 (-1 to 10) | 10 (3 to 17) | 173270 | -27 (-31 to -23) | 2 (-2 to 6) | 8 (2 to 13) |
| >33-43 | 136712 | -13 (-18 to -8) | 4 (-2 to 9) | 11 (4 to 18) | 165094 | -22 (-26 to -18) | 0 (-4 to 4) | 7 (1 to 12) |
| >43-64 | 152273 | -11 (-16 to -5) | -1 (-7 to 4) | 7 (0 to 14) | 185107 | -21 (-25 to -17) | -3 (-7 to 1) | 5 (-1 to 11) |
| >64 | 77699 | -5 (-11 to 1) | -1 (-8 to 5) | 8 (1 to 16) | 86935 | -14 (-19 to -10) | -2 (-7 to 3) | 7 (0 to 13) |

| THM4* Metrics (µg/L) | Term Births (n) | Unadjusted SGA (5%) OR (95% CI) | Adjusted SGA (5%) OR (95% CI) | Adjusted SGA (5%) OR (95% CI) | Term Births (n) | Unadjusted SGA (5%) OR (95% CI) | Adjusted SGA (5%) OR (95% CI) | Adjusted SGA (5%) OR (95% CI) |
|---|---|---|---|---|---|---|---|---|
| **Prior to Imputation** | | | | | **After Imputation** | | | |
| 0-15 | 36389 | REF | REF | REF | 80077 | REF | REF | REF |
| >15-25 | 97415 | 1.02 (0.96 to 1.07) | 0.99 (0.93 to 1.05) | 0.98 (0.91 to 1.05) | 162394 | 1.01 (0.98 to 1.05) | 1.02 (0.97 to 1.06) | 1.01 (0.96 to 1.07) |
| >25-33 | 125435 | 1 (0.95 to 1.05) | 0.97 (0.92 to 1.03) | 0.95 (0.88 to 1.02) | 172822 | 1 (0.96 to 1.04) | 0.98 (0.94 to 1.03) | 0.98 (0.92 to 1.04) |
| >33-43 | 136310 | 0.98 (0.93 to 1.03) | 0.95 (0.9 to 1.01) | 0.92 (0.85 to 0.99) | 164629 | 0.98 (0.94 to 1.02) | 0.96 (0.92 to 1) | 0.95 (0.89 to 1.01) |
| >43-64 | 151954 | 1.02 (0.97 to 1.08) | 0.98 (0.92 to 1.04) | 0.93 (0.87 to 1.01) | 184706 | 1.01 (0.97 to 1.04) | 0.98 (0.93 to 1.02) | 0.96 (0.9 to 1.02) |
| >64 | 77586 | 1.03 (0.98 to 1.1) | 0.98 (0.92 to 1.05) | 0.92 (0.85 to 1) | 86814 | 1.03 (0.99 to 1.08) | 0.98 (0.93 to 1.03) | 0.94 (0.88 to 1.01) |

| THM4* Metrics (µg/L) | Term Births (n) | Unadjusted SGA (10%) OR (95% CI) | Adjusted SGA (10%) OR (95% CI) | Adjusted SGA (10%) OR (95% CI) | Term Births (n) | Unadjusted SGA (10%) OR (95% CI) | Adjusted SGA (10%) OR (95% CI) | Adjusted SGA (10%) OR (95% CI) |
|---|---|---|---|---|---|---|---|---|
| **Prior to Imputation** | | | | | **After Imputation** | | | |
| 0-15 | 36389 | REF | REF | REF | 80077 | REF | REF | REF |
| >15-25 | 97415 | 1.04 (1 to 1.08) | 1.03 (0.98 to 1.08) | 1.02 (0.97 to 1.08) | 162394 | 1.03 (1 to 1.06) | 1.03 (1 to 1.07) | 1.02 (0.98 to 1.06) |
| >25-33 | 125435 | 1.01 (0.97 to 1.05) | 0.99 (0.95 to 1.04) | 0.98 (0.93 to 1.04) | 172822 | 1.02 (0.99 to 1.05) | 1.02 (0.98 to 1.05) | 1.01 (0.96 to 1.05) |
| >33-43 | 136310 | 0.99 (0.95 to 1.03) | 0.98 (0.94 to 1.02) | 0.96 (0.91 to 1.02) | 164629 | 1.01 (0.98 to 1.03) | 1 (0.97 to 1.03) | 0.99 (0.94 to 1.03) |
| >43-64 | 151954 | 1.03 (0.99 to 1.07) | 1 (0.96 to 1.04) | 0.97 (0.92 to 1.03) | 184706 | 1.04 (1.01 to 1.07) | 1.01 (0.98 to 1.04) | 0.99 (0.95 to 1.04) |
| >64 | 77586 | 1.06 (1.01 to 1.1) | 1.01 (0.97 to 1.06) | 0.98 (0.92 to 1.04) | 86814 | 1.07 (1.03 to 1.1) | 1.03 (0.99 to 1.07) | 1 (0.96 to 1.06) |

| THM4 (µg/L) | All Births (n) | Unadjusted PTB OR (95% CI) | Adjusted PTB OR (95% CI) | Adjusted PTB OR (95% CI) | All Births (n) | Unadjusted PTB OR (95% CI) | Adjusted PTB OR (95% CI) | Adjusted PTB OR (95% CI) |
|---|---|---|---|---|---|---|---|---|
| **Prior to Imputation** | | | | | **After Imputation** | | | |
| 0-16 | 37467 | REF | REF | REF | 89276 | REF | REF | REF |
| >16-26 | 109040 | 1.11 (1.07 to 1.15) | 1.05 (1.01 to 1.1) | 1.07 (1.02 to 1.13) | 191959 | 1.09 (1.07 to 1.12) | 1.06 (1.02 to 1.09) | 1.05 (1.01 to 1.09) |
| >26-33 | 135104 | 1.11 (1.07 to 1.16) | 1.03 (0.99 to 1.08) | 1.06 (1.01 to 1.12) | 177151 | 1.12 (1.09 to 1.15) | 1.02 (0.99 to 1.06) | 1.03 (0.99 to 1.07) |
| >33-42 | 159726 | 1.07 (1.03 to 1.11) | 1.01 (0.97 to 1.06) | 1.04 (0.99 to 1.1) | 191667 | 1.09 (1.06 to 1.12) | 1 (0.97 to 1.03) | 1.01 (0.97 to 1.05) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| >42-61 | 174946 | 1 (0.97 to 1.04) | 0.98 (0.94 to 1.02) | 1 (0.95 to 1.06) | 206096 | 1.03 (1.01 to 1.06) | 0.96 (0.93 to 0.99) | 0.97 (0.93 to 1.01) |
| >61 | 84226 | 1.03 (0.99 to 1.07) | 1.03 (0.98 to 1.08) | 1.04 (0.98 to 1.1) | 96041 | 1.04 (1.01 to 1.07) | 0.98 (0.94 to 1.02) | 0.99 (0.94 to 1.03) |

**Table S5. Association Between Third-Trimester DBP Exposure and Change in Mean BWT Among Term Births.**

| THM4 Metrics (µg/L) | Birth Count (n) | Unadjusted BWT (95% CI), g | Adjusted BWT (95% CI), g | Adjusted BWT [1] (95% CI), g |
|---|---|---|---|---|
| **Model with All 13 Covariates [2]** | | | | |
| 0-15 | 80,170 | REF | REF | REF |
| >15-25 | 162,624 | -23 (-27 to -19) | 1 (-3 to 5) | 5 (0 to 10) |
| >25-33 | 173,270 | -27 (-31 to -23) | 2 (-2 to 6) | 8 (2 to 13) |
| >33-43 | 165,094 | -22 (-26 to -18) | 0 (-4 to 4) | 7 (1 to 12) |
| >43-64 | 185,107 | -21 (-25 to -17) | -3 (-7 to 1) | 5 (-1 to 11) |
| >64 | 86,935 | -14 (-19 to -10) | -2 (-7 to 3) | 7 (0 to 13) |
| **Model with CIE Identified 5 Covariates [3]** | | | | |
| 0-15 | 80,170 | REF | REF | REF |
| >15-25 | 162,624 | -23 (-27 to -19) | 1 (-3 to 5) | 7 (2 to 12) |
| >25-33 | 173,270 | -27 (-31 to -23) | 0 (-4 to 4) | 8 (3 to 13) |
| >33-43 | 165,094 | -22 (-26 to -18) | -3 (-7 to 1) | 6 (1 to 11) |
| >43-64 | 185,107 | -21 (-25 to -17) | -6 (-10 to -2) | 4 (-2 to 9) |
| >64 | 86,935 | -14 (-19 to -10) | -7 (-12 to -3) | 3 (-2 to 9) |
| **HAA5 Metrics (µg/L)** | **Birth Count (n)** | **Unadjusted BWT (95% CI), g** | **Adjusted BWT (95% CI), g** | **Adjusted BWT [1] (95% CI), g** |
| **Model with All 13 Covariates [2]** | | | | |
| 0-6 | 77,999 | REF | REF | REF |
| >6-10 | 153,880 | -25 (-29 to -21) | 0 (-4 to 5) | -4 (-9 to 1) |
| >10-19 | 182,831 | -32 (-35 to -28) | -4 (-8 to 0) | -8 (-14 to -3) |
| >19-27 | 174,103 | -37 (-41 to -33) | -7 (-11 to -3) | -11 (-17 to -6) |
| >27-38 | 177,484 | -38 (-42 to -34) | -5 (-9 to -1) | -9 (-15 to -4) |
| >38 | 86,903 | -42 (-46 to -37) | -8 (-13 to -3) | -12 (-19 to -6) |
| **Model with CIE Identified 5 Covariates [3]** | | | | |
| 0-6 | 77,999 | REF | REF | REF |
| >6-10 | 153,880 | -25 (-29 to -21) | -1 (-5 to 3) | -3 (-8 to 2) |
| >10-19 | 182,831 | -32 (-35 to -28) | -8 (-12 to -4) | -9 (-14 to -3) |
| >19-27 | 174,103 | -37 (-41 to -33) | -10 (-14 to -6) | -11 (-16 to -6) |
| >27-38 | 177,484 | -38 (-42 to -34) | -10 (-14 to -6) | -10 (-16 to -4) |
| >38 | 86,903 | -42 (-46 to -37) | -13 (-17 to -8) | -14 (-20 to -7) |

[1]. Model also adjusted for THM4 or HAA5 exposures.
[2]. Model adjusted for maternal age, race, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model adjusted for maternal race, education, marital status, delivery source of payment, and income.

**Table S6. Association Between Second-Trimester DBP Exposure and Change in SGA (5%) OR Among Term Births.**

| DBP Metrics [1] | Categorical Levels, µg/L | Births (n) | Unadjusted SGA OR (95% CI) | Adjusted SGA OR[2] (95% CI) | Adjusted SGA OR[3] (95% CI) |
|---|---|---|---|---|---|
| **THM4** [4] | 0-15 | 71,547 | REF | REF | REF |
| | >15-25 | 160,698 | 0.98 (0.94 to 1.02) | 0.98 (0.94 to 1.03) | 0.97 (0.92 to 1.03) |
| | >25-33 | 177,109 | 1.00 (0.96 to 1.04) | 0.99 (0.95 to 1.03) | 0.97 (0.91 to 1.03) |
| | >33-43 | 188,605 | 0.95 (0.92 to 0.99) | 0.94 (0.90 to 0.98) | 0.92 (0.86 to 0.98) |
| | >43-64 | 182,207 | 1.00 (0.96 to 1.04) | 0.98 (0.94 to 1.03) | 0.95 (0.90 to 1.02) |
| | >64 | 71,276 | 1.00 (0.95 to 1.05) | 0.98 (0.93 to 1.04) | 0.94 (0.87 to 1.01) |
| **THM-Br** [5] | 0-7 | 70,935 | REF | REF | REF |
| | >7-11 | 147,608 | 1.00 (0.96 to 1.04) | 1.01 (0.96 to 1.05) | 1.01 (0.96 to 1.06) |
| | >11-14 | 191,128 | 0.97 (0.93 to 1.01) | 1.02 (0.97 to 1.06) | 1.02 (0.97 to 1.07) |
| | >14-18 | 196,506 | 0.98 (0.94 to 1.02) | 1.00 (0.96 to 1.05) | 1.00 (0.95 to 1.05) |
| | >18-28 | 167,921 | 0.95 (0.91 to 0.99) | 0.99 (0.94 to 1.03) | 0.99 (0.94 to 1.04) |
| | >28 | 77,344 | 0.90 (0.86 to 0.94) | 0.92 (0.87 to 0.97) | 0.91 (0.86 to 0.97) |
| **HAA5** [6] | 0-6 | 76,409 | REF | REF | REF |
| | >6-10 | 149,529 | 0.98 (0.94 to 1.02) | 0.98 (0.94 to 1.02) | 1.01 (0.95 to 1.06) |
| | >10-19 | 189,413 | 0.98 (0.95 to 1.02) | 0.99 (0.95 to 1.03) | 1.03 (0.97 to 1.09) |
| | >19-27 | 188,830 | 0.98 (0.95 to 1.02) | 0.99 (0.95 to 1.04) | 1.03 (0.97 to 1.09) |
| | >27-38 | 173,417 | 0.97 (0.94 to 1.01) | 0.97 (0.93 to 1.02) | 1.02 (0.96 to 1.08) |
| | >38 | 73,844 | 1.06 (1.01 to 1.11) | 1.04 (0.99 to 1.09) | 1.09 (1.02 to 1.17) |
| **DCAA** | 0-1 | 60,470 | REF | REF | REF |
| | >1-5 | 238,302 | 0.99 (0.95 to 1.03) | 1.01 (0.96 to 1.05) | 1.05 (0.99 to 1.10) |
| | >5-13 | 275,515 | 1.01 (0.97 to 1.06) | 1.03 (0.99 to 1.08) | 1.09 (1.03 to 1.15) |
| | >13-22 | 213,979 | 1.02 (0.98 to 1.06) | 1.02 (0.97 to 1.07) | 1.08 (1.01 to 1.14) |
| | >22 | 63,176 | 1.08 (1.03 to 1.14) | 1.05 (1.00 to 1.11) | 1.11 (1.04 to 1.19) |
| **TCAA** | 0-1 | 72,301 | REF | REF | REF |
| | >1-2 | 241,488 | 0.99 (0.96 to 1.03) | 1.00 (0.96 to 1.04) | 1.02 (0.97 to 1.06) |
| | >2-7 | 209,767 | 0.99 (0.95 to 1.03) | 1.00 (0.96 to 1.05) | 1.03 (0.98 to 1.08) |
| | >7-15 | 250,528 | 0.98 (0.94 to 1.02) | 1.01 (0.96 to 1.05) | 1.04 (0.99 to 1.09) |
| | >15 | 77,358 | 1.07 (1.02 to 1.12) | 1.05 (1.00 to 1.1) | 1.09 (1.02 to 1.15) |
| **HAA-Br** [7] | 0-3 | 51,739 | REF | REF | REF |
| | >3-6 | 293,617 | 1.02 (0.97 to 1.06) | 1.02 (0.98 to 1.07) | 1.03 (0.98 to 1.08) |
| | >6-8 | 229,108 | 0.96 (0.92 to 1.00) | 1.00 (0.96 to 1.05) | 1.01 (0.96 to 1.07) |
| | >8-11 | 195,521 | 0.95 (0.91 to 0.99) | 1.00 (0.96 to 1.05) | 1.02 (0.97 to 1.07) |
| | >11 | 81,457 | 1.00 (0.95 to 1.06) | 1.01 (0.96 to 1.07) | 1.02 (0.97 to 1.09) |
| **HAA4** [8] | 0-2 | 54,689 | REF | REF | REF |
| | >2-4 | 250,178 | 1.00 (0.96 to 1.04) | 1.00 (0.95 to 1.05) | 1.00 (0.96 to 1.05) |
| | >4-7 | 345,682 | 0.94 (0.90 to 0.98) | 0.97 (0.93 to 1.02) | 0.98 (0.94 to 1.03) |
| | >7-10 | 153,474 | 0.96 (0.91 to 1.00) | 0.99 (0.95 to 1.04) | 1.01 (0.96 to 1.06) |
| | >10 | 47,419 | 1.03 (0.97 to 1.08) | 1.03 (0.97 to 1.09) | 1.03 (0.97 to 1.10) |
| **HAA9** [9] | 0-9 | 72,454 | REF | REF | REF |
| | >9-15 | 173,412 | 1.02 (0.98 to 1.06) | 1.00 (0.96 to 1.04) | 1.03 (0.98 to 1.09) |

| | | | | | |
|---|---|---|---|---|---|
| | >15-24 | 163,216 | 1.00 (0.96 to 1.04) | 0.99 (0.94 to 1.03) | 1.03 (0.97 to 1.09) |
| | >24-33 | 194,168 | 1.02 (0.98 to 1.06) | 1.01 (0.96 to 1.05) | 1.05 (0.99 to 1.11) |
| | >33-45 | 172,998 | 0.98 (0.94 to 1.02) | 0.98 (0.93 to 1.02) | 1.02 (0.96 to 1.09) |
| | >45 | 75,194 | 1.08 (1.03 to 1.13) | 1.05 (1.00 to 1.1) | 1.10 (1.02 to 1.17) |
| **DBP9** [10] | 0-24 | 80,117 | REF | REF | - |
| | >24-41 | 174,691 | 1.00 (0.96 to 1.04) | 1.00 (0.96 to 1.05) | - |
| | >41-53 | 162,435 | 0.96 (0.92 to 1.00) | 0.95 (0.91 to 1.00) | - |
| | >53-71 | 194,107 | 0.99 (0.96 to 1.03) | 0.99 (0.95 to 1.03) | - |
| | >71-97 | 169,749 | 1.00 (0.96 to 1.04) | 0.98 (0.94 to 1.03) | - |
| | >97 | 70,343 | 1.03 (0.98 to 1.08) | 1.01 (0.96 to 1.06) | - |
| **DBP13** [11] | 0-27 | 79,217 | REF | REF | - |
| | >27-45 | 157,903 | 1.02 (0.98 to 1.06) | 1.02 (0.98 to 1.06) | - |
| | >45-59 | 180,503 | 0.96 (0.92 to 0.99) | 0.95 (0.91 to 0.99) | - |
| | >59-78 | 204,640 | 1.01 (0.97 to 1.05) | 1.00 (0.96 to 1.05) | - |
| | >78-105 | 164,024 | 1.00 (0.96 to 1.04) | 0.99 (0.94 to 1.03) | - |
| | >105 | 65,155 | 1.03 (0.98 to 1.08) | 1.01 (0.96 to 1.06) | - |

[1]. The categorization of each DBP class or species was based on the distribution of the DBP. THM4, THM-Br, HAA5, HAA9, DBP9, and DBP13 were categorized into 6 categories: 0-10, 10-30, 30-50, 50-70, 70-90, and 90-100 percentile.
DCAA, TCAA, HAA-Br, and HAA4 were categorized into 5 categories: 0-10, 10-36, 36-64, 64-90, and 90-100 percentile.
[2]. Model adjusted for maternal age, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model also adjusted for THM4 or HAA5 exposures.
[4]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[5]. THM-Br, the summation of BDCM, CDBM, and TBM.
[6]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[7]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[8]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[9]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.
[10]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[11]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

**Table S7. Association Between Second-Trimester DBP Exposure and Change in SGA (10%) OR Among Term Births.**

| DBP Metrics [1] | Categorical Levels, µg/L | Births (n) | Unadjusted SGA OR (95% CI) | Adjusted SGA OR[2] (95% CI) | Adjusted SGA OR[3] (95% CI) |
|---|---|---|---|---|---|
| **THM4** [4] | 0-15 | 71,547 | REF | REF | REF |
| | >15-25 | 160,698 | 1.01 (0.98 to 1.04) | 1.01 (0.97 to 1.04) | 1.01 (0.97 to 1.05) |
| | >25-33 | 177,109 | 1.02 (0.99 to 1.05) | 1.01 (0.98 to 1.05) | 1.01 (0.97 to 1.06) |
| | >33-43 | 188,605 | 1.00 (0.97 to 1.03) | 0.99 (0.96 to 1.02) | 0.99 (0.94 to 1.04) |
| | >43-64 | 182,207 | 1.03 (1.00 to 1.06) | 1.01 (0.98 to 1.05) | 1.01 (0.96 to 1.06) |
| | >64 | 71,276 | 1.03 (1.00 to 1.07) | 1.02 (0.98 to 1.06) | 1.00 (0.95 to 1.06) |
| **THM-Br** [5] | 0-7 | 70,935 | REF | REF | REF |
| | >7-11 | 147,608 | 1.00 (0.97 to 1.03) | 1.01 (0.98 to 1.05) | 1.00 (0.97 to 1.04) |
| | >11-14 | 191,128 | 0.97 (0.94 to 1.00) | 1.01 (0.98 to 1.05) | 1.00 (0.97 to 1.04) |
| | >14-18 | 196,506 | 1.00 (0.97 to 1.03) | 1.02 (0.99 to 1.05) | 1.01 (0.97 to 1.05) |
| | >18-28 | 167,921 | 0.95 (0.92 to 0.98) | 0.99 (0.96 to 1.03) | 0.98 (0.95 to 1.02) |
| | >28 | 77,344 | 0.93 (0.90 to 0.96) | 0.95 (0.92 to 0.99) | 0.94 (0.90 to 0.98) |
| **HAA5** [6] | 0-6 | 76,409 | REF | REF | REF |
| | >6-10 | 149,529 | 1.00 (0.97 to 1.03) | 1.00 (0.97 to 1.04) | 1.00 (0.96 to 1.04) |
| | >10-19 | 189,413 | 1.00 (0.98 to 1.03) | 1.00 (0.97 to 1.03) | 1.00 (0.95 to 1.04) |
| | >19-27 | 188,830 | 1.02 (0.99 to 1.05) | 1.03 (1.00 to 1.06) | 1.03 (0.98 to 1.07) |
| | >27-38 | 173,417 | 0.99 (0.97 to 1.02) | 0.98 (0.95 to 1.01) | 0.98 (0.93 to 1.02) |
| | >38 | 73,844 | 1.08 (1.04 to 1.11) | 1.06 (1.02 to 1.10) | 1.05 (1.00 to 1.11) |
| **DCAA** | 0-1 | 60,470 | REF | REF | REF |
| | >1-5 | 238,302 | 1.01 (0.98 to 1.04) | 1.02 (0.98 to 1.05) | 1.03 (0.99 to 1.07) |
| | >5-13 | 275,515 | 1.03 (1.00 to 1.07) | 1.05 (1.01 to 1.08) | 1.06 (1.02 to 1.11) |
| | >13-22 | 213,979 | 1.04 (1.01 to 1.07) | 1.02 (0.99 to 1.06) | 1.04 (0.99 to 1.09) |
| | >22 | 63,176 | 1.08 (1.04 to 1.13) | 1.05 (1.00 to 1.09) | 1.06 (1.01 to 1.12) |
| **TCAA** | 0-1 | 72,301 | REF | REF | REF |
| | >1-2 | 241,488 | 0.99 (0.96 to 1.02) | 0.99 (0.96 to 1.03) | 0.99 (0.96 to 1.03) |
| | >2-7 | 209,767 | 1.01 (0.98 to 1.04) | 1.02 (0.99 to 1.05) | 1.02 (0.98 to 1.06) |
| | >7-15 | 250,528 | 0.97 (0.95 to 1.00) | 0.99 (0.96 to 1.03) | 1.00 (0.96 to 1.03) |
| | >15 | 77,358 | 1.06 (1.02 to 1.09) | 1.04 (1.00 to 1.08) | 1.05 (1.00 to 1.09) |
| **HAA-Br** [7] | 0-3 | 51,739 | REF | REF | REF |
| | >3-6 | 293,617 | 1.02 (0.99 to 1.05) | 1.03 (0.99 to 1.06) | 1.03 (0.99 to 1.06) |
| | >6-8 | 229,108 | 0.97 (0.94 to 1.00) | 1.01 (0.98 to 1.05) | 1.01 (0.97 to 1.05) |
| | >8-11 | 195,521 | 0.96 (0.93 to 0.99) | 1.01 (0.97 to 1.05) | 1.01 (0.97 to 1.05) |
| | >11 | 81,457 | 0.99 (0.95 to 1.03) | 1.00 (0.96 to 1.04) | 1.00 (0.96 to 1.04) |
| **HAA4** [8] | 0-2 | 54,689 | REF | REF | REF |
| | >2-4 | 250,178 | 1.01 (0.98 to 1.04) | 1.01 (0.97 to 1.04) | 1.00 (0.97 to 1.04) |
| | >4-7 | 345,682 | 0.96 (0.93 to 0.99) | 0.98 (0.95 to 1.01) | 0.98 (0.94 to 1.01) |
| | >7-10 | 153,474 | 0.97 (0.94 to 1.00) | 1.00 (0.96 to 1.04) | 1.00 (0.96 to 1.04) |
| | >10 | 47,419 | 0.99 (0.95 to 1.03) | 0.99 (0.94 to 1.04) | 0.98 (0.94 to 1.03) |
| **HAA9** [9] | 0-9 | 72,454 | REF | REF | REF |
| | >9-15 | 173,412 | 1.04 (1.01 to 1.07) | 1.02 (0.99 to 1.05) | 1.02 (0.98 to 1.07) |

| | | | | | |
|---|---|---|---|---|---|
| | >15-24 | 163,216 | 1.02 (0.99 to 1.05) | 1.01 (0.97 to 1.04) | 1.01 (0.97 to 1.06) |
| | >24-33 | 194,168 | 1.04 (1.01 to 1.07) | 1.03 (1.00 to 1.07) | 1.04 (0.99 to 1.08) |
| | >33-45 | 172,998 | 1.01 (0.98 to 1.04) | 1.00 (0.96 to 1.03) | 1.00 (0.96 to 1.05) |
| | >45 | 75,194 | 1.09 (1.05 to 1.12) | 1.05 (1.01 to 1.09) | 1.06 (1.00 to 1.11) |
| **DBP9** [10] | 0-24 | 80,117 | REF | REF | - |
| | >24-41 | 174,691 | 1.03 (1.00 to 1.06) | 1.04 (1.00 to 1.07) | - |
| | >41-53 | 162,435 | 1.02 (0.99 to 1.05) | 1.02 (0.99 to 1.05) | - |
| | >53-71 | 194,107 | 1.03 (1.00 to 1.06) | 1.03 (0.99 to 1.06) | - |
| | >71-97 | 169,749 | 1.04 (1.01 to 1.07) | 1.02 (0.98 to 1.05) | - |
| | >97 | 70,343 | 1.06 (1.03 to 1.10) | 1.05 (1.01 to 1.09) | - |
| **DBP13** [11] | 0-27 | 79,217 | REF | REF | - |
| | >27-45 | 157,903 | 1.04 (1.01 to 1.07) | 1.04 (1.01 to 1.08) | - |
| | >45-59 | 180,503 | 1.01 (0.98 to 1.04) | 1.01 (0.98 to 1.05) | - |
| | >59-78 | 204,640 | 1.04 (1.01 to 1.07) | 1.03 (1.00 to 1.07) | - |
| | >78-105 | 164,024 | 1.03 (1.00 to 1.06) | 1.01 (0.98 to 1.05) | - |
| | >105 | 65,155 | 1.06 (1.02 to 1.09) | 1.05 (1.00 to 1.09) | - |

[1]. The categorization of each DBP class or species was based on the distribution of the DBP metrics. THM4, THM-Br, HAA5, HAA9, DBP9, and DBP13 were categorized into 6 categories: 0-10, 10-30, 30-50, 50-70, 70-90, and 90-100 percentile.
DCAA, TCAA, HAA-Br, and HAA4 were categorized into 5 categories: 0-10, 10-36, 36-64, 64-90, and 90-100 percentile.
[2]. Model adjusted for maternal age, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model also adjusted for THM4 or HAA5 exposures.
[4]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[5]. THM-Br, the summation of BDCM, CDBM, and TBM.
[6]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[7]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[8]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[9]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.
[10]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[11]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

**Table S8. Association Between First-Trimester DBP Exposure and Change in PTB OR Among Births.**

| DBPs [1] | Metrics | Birth Count (n) | Unadjusted PTB OR (95% CI) | Adjusted PTB OR[2] (95% CI) | Adjusted PTB OR[3] (95% CI) |
|---|---|---|---|---|---|
| **THM4 [4], µg/L** | 0-16 | 88,218 | REF | REF | REF |
| | >16-26 | 191,681 | 1.13 (1.1 to 1.16) | 1.07 (1.04 to 1.11) | 1.04 (1 to 1.08) |
| | >26-33 | 177,842 | 1.11 (1.08 to 1.14) | 1 (0.97 to 1.03) | 0.97 (0.93 to 1.01) |
| | >33-42 | 193,133 | 1.09 (1.07 to 1.12) | 0.99 (0.96 to 1.02) | 0.96 (0.92 to 1) |
| | >42-61 | 205,254 | 1.03 (1 to 1.06) | 0.93 (0.9 to 0.96) | 0.9 (0.86 to 0.94) |
| | >61 | 96,062 | 1 (0.97 to 1.03) | 0.92 (0.89 to 0.96) | 0.88 (0.83 to 0.92) |
| **THM-Br [5], µg/L** | 0-7 | 78,600 | REF | REF | REF |
| | >7-11 | 162,870 | 1.1 (1.07 to 1.13) | 1.02 (0.98 to 1.05) | 1.02 (0.98 to 1.06) |
| | >11-14 | 211,083 | 1.05 (1.03 to 1.08) | 0.98 (0.95 to 1.02) | 0.98 (0.95 to 1.02) |
| | >14-17 | 181,103 | 1.01 (0.98 to 1.04) | 0.96 (0.93 to 1) | 0.95 (0.91 to 0.98) |
| | >17-27 | 220,592 | 1.03 (1 to 1.05) | 0.98 (0.95 to 1.01) | 0.96 (0.92 to 0.99) |
| | >27 | 97,942 | 1.01 (0.98 to 1.04) | 0.92 (0.89 to 0.96) | 0.89 (0.86 to 0.93) |
| **HAA5 [6], µg/L** | 0-6 | 84,019 | REF | REF | REF |
| | >6-10 | 168,523 | 1.12 (1.09 to 1.15) | 1.06 (1.03 to 1.1) | 1.07 (1.03 to 1.12) |
| | >10-19 | 214,108 | 1.08 (1.05 to 1.11) | 0.97 (0.94 to 1) | 1.02 (0.98 to 1.06) |
| | >19-26 | 185,162 | 1.09 (1.06 to 1.12) | 0.96 (0.93 to 1) | 1.02 (0.98 to 1.07) |
| | >26-36 | 196,859 | 1.13 (1.1 to 1.16) | 0.99 (0.95 to 1.02) | 1.07 (1.03 to 1.12) |
| | >36 | 103,519 | 1.15 (1.12 to 1.18) | 0.99 (0.96 to 1.03) | 1.1 (1.05 to 1.16) |
| **DCAA, µg/L** | 0-1 | 67,071 | REF | REF | REF |
| | >1-6 | 294,043 | 1.14 (1.11 to 1.17) | 1.08 (1.05 to 1.12) | 1.1 (1.06 to 1.15) |
| | >6-13 | 281,133 | 1.08 (1.05 to 1.11) | 1 (0.96 to 1.03) | 1.05 (1.01 to 1.09) |
| | >13-21 | 227,440 | 1.15 (1.12 to 1.18) | 1.02 (0.99 to 1.06) | 1.12 (1.07 to 1.17) |
| | >21 | 82,503 | 1.19 (1.15 to 1.24) | 1.07 (1.02 to 1.11) | 1.17 (1.12 to 1.23) |
| **TCAA, µg/L** | 0-1 | 79,597 | REF | REF | REF |
| | >1-2 | 272,806 | 1.06 (1.03 to 1.09) | 1.01 (0.98 to 1.04) | 1.02 (0.98 to 1.05) |
| | >2-7 | 235,130 | 1.05 (1.02 to 1.08) | 0.98 (0.95 to 1.01) | 1 (0.97 to 1.04) |
| | >7-14 | 262,702 | 1.07 (1.05 to 1.1) | 0.97 (0.94 to 1) | 1.01 (0.98 to 1.05) |
| | >14 | 101,955 | 1.03 (1 to 1.07) | 0.92 (0.89 to 0.96) | 0.99 (0.95 to 1.03) |
| **HAA-Br [7], µg/L** | 0-3 | 57,799 | REF | REF | REF |
| | >3-6 | 324,113 | 1.02 (0.99 to 1.05) | 1 (0.97 to 1.03) | 0.97 (0.94 to 1.01) |
| | >6-8 | 259,232 | 1.04 (1.01 to 1.07) | 0.97 (0.94 to 1.01) | 0.96 (0.92 to 0.99) |
| | >8-11 | 218,293 | 1.09 (1.05 to 1.12) | 0.99 (0.96 to 1.03) | 0.97 (0.94 to 1.01) |
| | >11 | 92,753 | 1.06 (1.02 to 1.1) | 0.97 (0.93 to 1.01) | 0.96 (0.92 to 1) |
| **HAA4 [8], µg/L** | 0-2 | 61,454 | REF | REF | REF |
| | >2-5 | 415,580 | 1.02 (0.99 to 1.05) | 0.99 (0.96 to 1.03) | 0.97 (0.94 to 1) |
| | >5-6 | 134,783 | 1.04 (1 to 1.07) | 0.96 (0.92 to 1) | 0.96 (0.92 to 0.99) |
| | >6-9 | 260,070 | 1.06 (1.03 to 1.09) | 0.97 (0.93 to 1) | 0.96 (0.92 to 0.99) |
| | >9 | 80,303 | 1.04 (1 to 1.08) | 0.95 (0.91 to 0.99) | 0.94 (0.9 to 0.98) |
| **HAA9 [9], µg/L** | 0-9 | 79,686 | REF | REF | REF |
| | >9-15 | 194,989 | 1.09 (1.06 to 1.12) | 1.04 (1.01 to 1.07) | 1.05 (1.01 to 1.09) |

| | | | | | |
|---|---|---|---|---|---|
| | >15-24 | 184,337 | 1.07 (1.04 to 1.1) | 0.97 (0.94 to 1) | 1.02 (0.98 to 1.07) |
| | >24-32 | 189,442 | 1.08 (1.05 to 1.11) | 0.96 (0.93 to 1) | 1.01 (0.97 to 1.06) |
| | >32-43 | 199,829 | 1.12 (1.09 to 1.15) | 0.98 (0.95 to 1.02) | 1.06 (1.02 to 1.11) |
| | >43 | 103,907 | 1.12 (1.09 to 1.15) | 0.97 (0.94 to 1.01) | 1.06 (1.01 to 1.12) |
| **DBP9 [10], µg/L** | 0-25 | 94,464 | REF | REF | - |
| | >25-41 | 190,596 | 1.13 (1.1 to 1.16) | 1.06 (1.03 to 1.1) | - |
| | >41-53 | 182,143 | 1.16 (1.13 to 1.19) | 1.01 (0.98 to 1.05) | - |
| | >53-69 | 198,356 | 1.09 (1.06 to 1.12) | 0.96 (0.93 to 0.99) | - |
| | >69-92 | 186,823 | 1.08 (1.05 to 1.11) | 0.96 (0.93 to 0.99) | - |
| | >92 | 99,808 | 1.06 (1.03 to 1.09) | 0.94 (0.91 to 0.98) | - |
| **DBP13 [11], µg/L** | 0-28 | 92,667 | REF | REF | - |
| | >28-46 | 186,899 | 1.13 (1.1 to 1.16) | 1.07 (1.04 to 1.1) | - |
| | >46-59 | 188,752 | 1.16 (1.13 to 1.19) | 1.01 (0.98 to 1.04) | - |
| | >59-75 | 197,671 | 1.1 (1.07 to 1.13) | 0.97 (0.94 to 1) | - |
| | >75-99 | 187,625 | 1.07 (1.05 to 1.1) | 0.95 (0.92 to 0.98) | - |
| | >99 | 98,576 | 1.07 (1.03 to 1.1) | 0.95 (0.91 to 0.98) | - |

[1]. The categorization of each DBP class or species was based on the distribution of the DBP.
THM4, THM-Br, HAA5, HAA9, DBP9, and DBP13 were categorized into 6 categories with cut points of 10, 30, 50, 70, and 90%tile:
DCAA, TCAA, HAA-Br, and HAA4 were categorized into 5 categories: 0-10, 10-36, 36-64, 64-90, and 90-100 percentile.
[2]. Model adjusted for maternal age, race, education, smoke, marital status, delivery source of payment, income, prenatal care adequacy (Kessner Index), interpregnancy interval, parity, season, pre-pregnancy BMI, and weight gain during pregnancy.
[3]. Model also adjusted for THM4 or HAA5 exposures.
[4]. THM4, the summation of four regulated THMs: TCM, BDCM, CDBM, and TBM.
[5]. THM-Br, the summation of BDCM, CDBM, and TBM.
[6]. HAA5, the summation of five regulated HAAs: MCAA, MBAA, DCAA, DBAA, and TCAA.
[7]. HAA-Br, the summation of five brominated HAAs: BCAA, TBAA, BDCAA, CDBAA, and TBAA.
[8]. HAA4, the summation of four unregulated HAAs: BCAA, BDCAA, CDBAA, and TBAA.
[9]. HAA9, the summation of all HAAs: MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.
[10]. DBP9, the summation of nine regulated DBPs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, DBAA, and TCAA.
[11]. DBP13, the summation of all THMs and HAAs: TCM, BDCM, CDBM, TBM, MCAA, MBAA, DCAA, BCAA, DBAA, TCAA, BDCAA, CDBAA, and TBAA.

**Table S9. Stratified Results between Term Birth mean BWT and the Highest Exposure Categories.**

| DBP | Sex Metrics | Change in mean BWT (g) | Note |
|---|---|---|---|
| **THM4** | Over all | 7 (0, 13) | Adjusted for HAA5 |
| | Male | 6 (-3 to 15) | |
| | Female | 8 (0 to 17) | |
| **THM Br** | Over all | 19 (13, 24) | Adjusted for HAA5 |
| | Male | 18 (11 to 26) | |
| | Female | 19 (12 to 27) | |
| **HAA5** | Over all | -12 (-19, -6) | Adjusted for THM4 |
| | Male | -13 (-21 to -4) | |
| | Female | -14 (-23 to -6) | |
| **DCAA** | Over all | -16 (-22, -10) | Adjusted for THM4 |
| | Male | -17 (-25 to -9) | |
| | Female | -16 (-24 to -7) | |
| **TCAA** | Over all | -6 (-11, 0) | Adjusted for THM4 |
| | Male | -9 (-16 to -2) | |
| | Female | -2 (-9 to 5) | |
| **HAA Br** | Over all | 13 (8, 18) | Adjusted for THM4 |
| | Male | 8 (1 to 16) | |
| | Female | 17 (10 to 25) | |
| **HAA4** | Over all | 12 (7, 18) | Adjusted for THM4 |
| | Male | 8 (0 to 16) | |
| | Female | 15 (7 to 23) | |
| **HAA9** | Over all | -4 (-10, 2) | Adjusted for THM4 |
| | Male | -5 (-13 to 4) | |
| | Female | -4 (-13 to 4) | |
| **DBP9** | Over all | -5 (-9, 0) | Not adjusted for HAA5 or THM4 |
| | Male | -3 (-10 to 3) | |
| | Female | -6 (-13 to 0) | |
| **DBP13** | Over all | -5 (-9, 0) | Not adjusted for HAA5 or THM4 |
| | Male | -4 (-11 to 2) | |
| | Female | -5 (-12 to 1) | |

**Table S10. Stratified Race-Specific Associations between Mean Birth Weight and THM4 Exposure Categories.**

| White | | | | Hispanic | | | |
|---|---|---|---|---|---|---|---|
| THM4 Metrics | Change in mean BWT (g) | Count | (%) | THM4 Metrics | Change in mean BWT (g) | Count | (%) |
| 0-15 | REF | 66,862 | 10.8 | 0-15 | REF | 3,286 | 8.0 |
| >15-25 | 2 | 114,720 | 18.5 | >15-25 | 8 | 6,822 | 16.6 |
| >25-33 | 3 | 115,284 | 18.6 | >25-33 | 10 | 9,352 | 22.7 |
| >33-43 | 1 | 116,068 | 18.7 | >33-43 | 7 | 8,084 | 19.6 |
| >43-64 | -1 | 136,005 | 22.0 | >43-64 | 5 | 9,895 | 24.0 |
| >64-416 | -2 | 70,298 | 11.4 | >64-416 | 8 | 3,741 | 9.1 |
| Total | | 619,237 | 100.0 | Total | | 41,180 | 100.0 |

| African American | | | | Native American | | | |
|---|---|---|---|---|---|---|---|
| THM4 Metrics | Change in mean BWT (g) | Count | (%) | THM4 Metrics | Change in mean BWT (g) | Count | (%) |
| 0-15 | REF | 7,531 | 4.9 | 0-15 | REF | 293 | 7.0 |
| >15-25 | 4 | 33,981 | 22.0 | >15-25 | -67 | 732 | 17.6 |
| >25-33 | 5 | 40,160 | 26.0 | >25-33 | -57 | 859 | 20.6 |
| >33-43 | 3 | 32,368 | 21.0 | >33-43 | -67 | 838 | 20.1 |
| >43-64 | -4 | 30,671 | 19.9 | >43-64 | -81 | 992 | 23.8 |
| >64-416 | 5 | 9,765 | 6.3 | >64-416 | -96 | 456 | 10.9 |
| Total | | 154,476 | 100.0 | Total | | 4,170 | 100.0 |

| Asian Indian | | | | Other Asian | | | |
|---|---|---|---|---|---|---|---|
| THM4 Metrics | Change in mean BWT (g) | Count | (%) | THM4 Metrics | Change in mean BWT (g) | Count | (%) |
| 0-15 | REF | 438 | 6.7 | 0-15 | REF | 1,075 | 6.1 |
| >15-25 | -16 | 1,196 | 18.4 | >15-25 | -40 | 3,033 | 17.3 |
| >25-33 | -18 | 1,418 | 21.8 | >25-33 | -38 | 3,771 | 21.6 |
| >33-43 | -7 | 1,540 | 23.7 | >33-43 | -50 | 4,055 | 23.2 |
| >43-64 | -15 | 1,493 | 22.9 | >43-64 | -47 | 3,985 | 22.8 |
| >64-416 | 18 | 425 | 6.5 | >64-416 | -49 | 1,563 | 8.9 |
| Total | | 6,510 | 100.0 | Total | | 17,482 | 100.0 |

| Others | | | |
|---|---|---|---|
| THM4 Metrics | Change in mean BWT (g) | Count | (%) |
| 0-15 | REF | 621 | 7.2 |
| >15-25 | -21 | 1,958 | 22.6 |
| >25-33 | -17 | 2,059 | 23.7 |
| >33-43 | 8 | 1,732 | 20.0 |
| >43-64 | -20 | 1,726 | 19.9 |
| >64-416 | -27 | 583 | 6.7 |
| Total | | 8,679 | 100.0 |

## BIBLIOGRAPHY


Abusallout, I., Rahman, S., & Hua, G. (2017). Effect of temperature and pH on dehalogenation of total organic chlorine, bromine and iodine in drinking water. *Chemosphere*, *187*, 11-18.

Ackerson, N. O. B., Liberatore, H. K., Plewa, M. J., Richardson, S. D., Ternes, T. A., & Duirk, S. E. (2020). Disinfection byproducts and halogen-specific total organic halogen speciation in chlorinated source waters–The impact of iopamidol and bromide. *Journal of Environmental Sciences*, *89*, 90-101.

American Public Health Association. (2005). American Water Works Association and Water Pollution Control Federation. 1995. *Standard methods for the examination of water and wastewater*, *19*.

Amy, G., Siddiqui, M., Zhai, W., DeBroux, J., & Odem, W. (1995). National survey of bromide ion in drinking water sources and impacts on disinfection by-product formation. *American Water Works association Research Foundation, Denver. CO*.

Arguello, M. D., Chrisweii, C. D., Fritz, J. S., Kissinger, L. D., Lee, K. W., Richard, J. J., & Svec, H. J. (1979). Trihalomethanes in water: a report on the occurrence, seasonal variation in concentrations, and precursors of trihalomethanes. *Journal-American Water Works Association*, *71*(9), 504-508.

Atwood, S. T., Lunn, R. M., Garner, S. C., & Jahnke, G. D. (2019). New Perspectives for Cancer Hazard Evaluation by the Report on Carcinogens: A Case Study Using Read-Across Methods in the Evaluation of Haloacetic Acids Found as Water Disinfection By-Products. *Environmental health perspectives*, *127*(12), 125003.

Backlund, P., Wondergem, E., Voogd, K., & De Jong, A. (1989). Mutagenic activity and presence of the strong mutagen 3-chloro-4-(dichloromethyl)-5-hydroxy-2 (5H)-furanone (MX) in chlorinated raw and drinking waters in the Netherlands. *Science of the total environment*, *84*, 273-282.

Beane Freeman, L. E., Cantor, K. P., Baris, D., Nuckols, J. R., Johnson, A., Colt, J. S., ... & Silverman, D. T. (2017). Bladder cancer and water disinfection by-product exposures through multiple routes: a population-based case–control study (new England, USA). *Environmental health perspectives*, *125*(6), 067010.

Bell, R. P., & Lidwell, O. M. (1940). The base catalysed prototropy of substituted acetones. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *176*(964), 88-113.

Bell, R. P., & Longuet-Higgins, H. C. (1946). 133. Kinetics of the halogenation of acetone in alkaline solutions. *Journal of the Chemical Society (Resumed)*, 636-638.

Bell, R. P., Gelles, E., & Möller, E. (1949). Kinetics of the base-catalyzed halogenation of some ketones and esters. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *198*(1054), 308-322.

Bond, T., Huang, J., Graham, N. J., & Templeton, M. R. (2014). Examining the interrelationship between DOC, bromide and chlorine dose on DBP formation in drinking water—a case study. *Science of the Total Environment*, *470*, 469-479.

Boorman, G. A. (1999). Drinking water disinfection byproducts: review and approach to toxicity evaluation. *Environmental health perspectives*, *107*(suppl 1), 207-217.

Bove, F. J., Fulcomer, M. C., Klotz, J. B., Esmart, J., Dufficy, E. M., & Savrin, J. E. (1995). Public drinking water contamination and birth outcomes. *American Journal of Epidemiology*, *141*(9), 850-862.

Bove, F., Shim, Y., & Zeitz, P. (2002). Drinking water contaminants and adverse pregnancy outcomes: a review. *Environmental health perspectives*, *110*(suppl 1), 61-74.

Bowen, H. J. M. (1979). *Environmental chemistry of the elements*. Academic Press..

Bradley, P. M., LeBlanc, D. R., Romanok, K. M., Smalling, K. L., Focazio, M. J., Cardon, M. C., ... & Wilson, V. S. (2021). Public and private tapwater: Comparative analysis of contaminant exposure and potential risk, Cape Cod, Massachusetts, USA. *Environment international*, *152*, 106487.

Brown, H. L., Chireau, M. V., Jallah, Y., & Howard, D. (2007). The "Hispanic paradox": an investigation of racial disparity in pregnancy outcomes at a tertiary care medical center. *American journal of obstetrics and gynecology*, *197*(2), 197-e1.

Bull, R. J., Rice, G., Teuschler, L., & Feder, P. (2009). Chemical measures of similarity among disinfection by-product mixtures. *Journal of Toxicology and Environmental Health, Part A*, *72*(7), 482-493.
Bull, R. J., Rice, G., & Teuschler, L. K. (2009). Determinants of whether or not mixtures of disinfection by-products are similar. *Journal of Toxicology and Environmental Health, Part A*, *72*(7), 437-460.

Chang, E. E., Lin, Y. P., & Chiang, P. C. (2001). Effects of bromide on the formation of THMs and HAAs. *Chemosphere*, *43*(8), 1029-1034.

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*(9), 1267-1273.

Chow, A. T., Dahlgren, R. A., & Harrison, J. A. (2007). Watershed sources of disinfection byproduct precursors in the Sacramento and San Joaquin Rivers, California. *Environmental science & technology*, *41*(22), 7645-7652.

Christian, M. S., York, R. G., Hoberman, A. M., Fisher, L. C., & Brown, W. R. (2002). Oral (drinking water) two-generation reproductive toxicity study of bromodichloromethane (BDCM) in rats. *International journal of toxicology*, *21*(2), 115-146.

Christian, M. S., York, R. G., Hoberman, A. M., Frazee, J., Fisher, L. C., Brown, W. R., & Creasy, D. M. (2002). Oral (drinking water) two-generation reproductive toxicity study of dibromoacetic acid (DBA) in rats. *International journal of toxicology*, *21*(4), 237-276.

Cohen, M. A., & Ryan, P. B. (1989). Observations less than the analytical limit of detection: a new approach. *JAPCA*, *39*(3), 328-329.

Costet, N., Garlantézec, R., Monfort, C., Rouget, F., Gagniere, B., Chevrier, C., & Cordier, S. (2012). Environmental and urinary markers of prenatal exposure to drinking water disinfection by-products, fetal growth, and duration of gestation in the PELAGIE birth cohort (Brittany, France, 2002–2006). *American journal of epidemiology*, *175*(4), 263-275.

Cowman, G. A., & Singer, P. C. (1995). Effect of bromide ion on haloacetic acid speciation resulting from chlorination and chloramination of aquatic humic substances. *Environmental science & technology*, *30*(1), 16-24.

Cozzie, D. A., Kanniganti, R., Charles, M. J., Johnson, J. D., & Ball, L. M. (1993). Formation and characterization of bacterial mutagens from reaction of the alternative disinfectant monochloramine with model aqueous solutions of fulvic acid. *Environmental and molecular mutagenesis*, *21*(3), 237-246.

Criquet, J., Rodriguez, E. M., Allard, S., Wellauer, S., Salhi, E., Joll, C. A., & Von Gunten, U. (2015). Reaction of bromine and chlorine with phenolic compounds and natural organic matter extracts–Electrophilic aromatic substitution and oxidation. *Water Research*, *85*, 476-486.

David, R. J. (1980). The quality and completeness of birthweight and gestational age data in computerized birth files. *American Journal of Public Health*, *70*(9), 964-973.

De Laat, J. (1981). Contribution à l'étude du mécanisme de formation des trihalométhanes-Incidence de l'azote ammoniacal et des traitements de préoxydation. *Doctorat de 3ème Cycle, Université de Poitiers*, (818).

Deborde, M., & Von Gunten, U. R. S. (2008). Reactions of chlorine with inorganic and organic compounds during water treatment—kinetics and mechanisms: a critical review. *Water research*, *42*(1-2), 13-51.

DeLaat, J., Merlet, N., & Dore, M. (1982). Chlorination of Organic Compounds: Chlorine Demand and Reactivity in Relationship to the Trihalomethane Formation. Incidence of Ammoniacal Nitrogen(Chloration De Composes Organiques: Demande En Chlore Et Reactivite Vis-a-Vis De La Formation Des Trihalomethanes. Incidence De L'Azote Ammoniacal). *Water Research*, *16*(10).

Dodds, L., & King, W. D. (2001). Relation between trihalomethane compounds and birth defects. *Occupational and Environmental Medicine*, *58*(7), 443-446.

Doederer, K., Gernjak, W., Weinberg, H. S., & Farré, M. J. (2014). Factors affecting the formation of disinfection by-products during chlorination and chloramination of secondary effluent for the production of high quality recycled water. *Water Research*, *48*, 218-228.

Donahue, S. M., Kleinman, K. P., Gillman, M. W., & Oken, E. (2010). Trends in birth weight and gestational length among singleton term births in the United States: 1990–2005. *Obstetrics and gynecology*, *115*(2 Pt 1), 357.

Echigo, S., & Minear, R. A. (2006). Kinetics of the reaction of hypobromous acid and organic matters in water treatment processes. *Water science and technology*, *53*(11), 235-243.

Edzwald, J. K., Becker, W. C., & Wattier, K. L. (1985). Surrogate parameters for monitoring organic matter and THM precursors. *Journal-American Water Works Association*, *77*(4), 122-132.

Engelage, S. K., Stringfellow, W. T., & Letain, T. (2009). Disinfection byproduct formation potentials of wetlands, agricultural drains, and rivers and the effect of biodegradation on trihalomethane precursors. *Journal of environmental quality*, *38*(5), 1901-1908.

Evlampidou, I., Font-Ribera, L., Rojas-Rueda, D., Gracia-Lavedan, E., Costet, N., Pearce, N., ... & Villanueva, C. M. (2020). Trihalomethanes in drinking water and bladder cancer burden in the European Union. *Environmental health perspectives*, *128*(1), 017001.

Falandysz, J., Fernandes, A., Gregoraszczuk, E., & Rose, M. (2014). The toxicological effects of halogenated naphthalenes: a review of aryl hydrocarbon receptor-mediated (dioxin-like) relative potency factors. *Journal of Environmental Science and Health, Part C*, *32*(3), 239-272.

Fang, C., Krasner, S. W., Chu, W., Ding, S., Zhao, T., & Gao, N. (2018). Formation and speciation of chlorinated, brominated, and iodinated haloacetamides in chloraminated iodide-containing waters. *Water research*, *145*, 103-112.

Feder, P. I., Ma, Z. J., Bull, R. J., Teuschler, L. K., & Rice, G. (2009). Evaluating sufficient similarity for drinking-water disinfection by-product (DBP) mixtures with bootstrap hypothesis test procedures. *Journal of Toxicology and Environmental Health, Part A*, *72*(7), 494-504.

Forssén, U. M., Herring, A. H., Savitz, D. A., Nieuwenhuijsen, M. J., Murphy, P. A., Singer, P. C., & Wright, J. M. (2007). Predictors of use and consumption of public drinking water among pregnant women. *Journal of exposure science & environmental epidemiology*, *17*(2), 159-169.

Francis, R. A., VanBriesen, J. M., & Small, M. J. (2010). Bayesian statistical modeling of disinfection byproduct (DBP) bromine incorporation in the ICR database. *Environmental science & technology*, *44*(4), 1232-1239.

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

Gjessing, H. K., Skjaerven, R., & Wilcox, A. J. (1999). Errors in gestational age: evidence of bleeding early in pregnancy. *American Journal of Public Health*, *89*(2), 213-218.Gould, J. P., Fitchhorn, L. E., & Urheim, E. (1983). Formation of brominated trihalomethanes: extent and kinetics. *Water chlorination: Environmental impact and health effects*, *4*.

Govarts, E., Remy, S., Bruckers, L., Den Hond, E., Sioen, I., Nelen, V., ... & Schoeters, G. (2016). Combined effects of prenatal exposures to environmental chemicals on birth weight. *International journal of environmental research and public health*, *13*(5), 495.

Grantz, K. L., Kim, S., Grobman, W. A., Newman, R., Owen, J., Skupski, D., ... & Albert, P. S. (2018). Fetal growth velocity: the NICHD fetal growth studies. *American journal of obstetrics and gynecology*, *219*(3), 285-e1.

Grellier, J., Bennett, J., Patelarou, E., Smith, R. B., Toledano, M. B., Rushton, L., ... & Nieuwenhuijsen, M. J. (2010). Exposure to disinfection by-products, fetal growth, and prematurity: a systematic review and meta-analysis. *Epidemiology*, 300-313.

Guthrie, J. P., & Cossar, J. (1986). The chlorination of acetone: a complete kinetic analysis. *Canadian journal of chemistry*, *64*(6), 1250-1266.

Guthrie, J. P., Cossar, J., & Klym, A. (1984). Halogenation of acetone. A method for determining pKas of ketones in aqueous solution, with an examination of the thermodynamics and kinetics of alkaline halogenation and a discussion of the best value for the rate constant for a diffusion-controlled reaction. Energetic requirements for a diffusion-controlled reaction involving heavy-atom bond formation. *Journal of the American Chemical Society*, *106*(5), 1351-1360.

Heeb, M. B., Criquet, J., Zimmermann-Steffens, S. G., & Von Gunten, U. (2014). Oxidative treatment of bromide-containing waters: formation of bromine and its reactions with inorganic and organic compounds—a critical review. *Water Research*, *48*, 15-42.

Heller-Grossman, L., Manka, J., Limoni-Relis, B., & Rebhun, M. (1993). Formation and distribution of haloacetic acids, THM and TOX in chlorination of bromide-rich lake water. *Water research*, *27*(8), 1323-1331.

Hinckley, A. F. (2003). *Disinfection by-products and prenatal development*. Colorado State University.

Hinckley, A. F., Bachand, A. M., Nuckols, J. R., & Reif, J. S. (2005). Identifying public water facilities with low spatial variability of disinfection by-products for epidemiological investigations. *Occupational and environmental medicine*, *62*(7), 494-499.

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

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

Horton, B. J., Luben, 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/American College of Occupational and Environmental Medicine*, *53*(10), 1172.

Hua, G., & Reckhow, D. A. (2007). Characterization of disinfection byproduct precursors based on hydrophobicity and molecular size. *Environmental science & technology*, *41*(9), 3309-3315.

Hua, G., & Reckhow, D. A. (2007). Comparison of disinfection byproduct formation from chlorine and alternative disinfectants. *Water research*, *41*(8), 1667-1678.

Hua, G., & Reckhow, D. A. (2008). DBP formation during chlorination and chloramination: effect of reaction time, pH, dosage, and temperature. *Journal-American Water Works Association*, *100*(8), 82-95.

Hua, G., & Reckhow, D. A. (2012). Evaluation of bromine substitution factors of DBPs during chlorination and chloramination. *Water Research*, *46*(13), 4208-4216.

Hua, G., & Reckhow, D. A. (2013). Effect of pre-ozonation on the formation and speciation of DBPs. *Water research*, *47*(13), 4322-4330.

Hua, G., Reckhow, D. A., & Kim, J. (2006). Effect of bromide and iodide ions on the formation and speciation of disinfection byproducts during chlorination. *Environmental Science & Technology*, *40*(9), 3050-3056.

Hubal, E. A. C., Wetmore, B. A., Wambaugh h, J. F., El-Masri, H., Sobus, J. R., & Bahadori, T. (2019). Advancing internal exposure and physiologically-based toxicokinetic modeling for 21st-century risk assessments. *Journal of exposure science & environmental epidemiology*, *29*(1), 11-20.

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

Iliodromiti, S., Mackay, D. F., Smith, G. C., Pell, J. P., Sattar, N., Lawlor, D. A., & Nelson, S. M. (2017). Customised and noncustomised birth weight centiles and prediction of stillbirth and infant mortality and morbidity: a cohort study of 979,912 term singleton pregnancies in Scotland. *PLoS medicine*, *14*(1), e1002228.

Itoh, S., Gordon, B. A., Callan, P., & Bartram, J. (2011). Regulations and perspectives on disinfection by-products: importance of estimating overall toxicity. *Journal of Water Supply: Research and Technology—AQUA*, *60*(5), 261-274.

Jeffers, P. M., & Wolfe, N. L. (1997). Hydrolysis of methyl bromide, ethyl bromide, chloropicrin, 1, 4-dichloro-2-butene, and other halogenated hydrocarbons.

Jiang, Y., Goodwill, J. E., Tobiason, J. E., & Reckhow, D. A. (2016). Comparison of the effects of ferrate, ozone, and permanganate pre-oxidation on disinfection byproduct formation from chlorination. In *Ferrites and Ferrates: Chemistry and Applications in Sustainable Energy and Environmental Remediation* (pp. 421-437). American Chemical Society.

Jones, D. B., Saglam, A., Song, H., & Karanfil, T. (2012). The impact of bromide/iodide concentration and ratio on iodinated trihalomethane formation and speciation. *Water research*, *46*(1), 11-20.

Kaufman, J. A., Wright, J. M., Evans, A., Rivera-Núñez, Z., Meyer, A., & Narotsky, M. G. (2018). Associations between disinfection by-product exposures and craniofacial birth defects. *Journal of occupational and environmental medicine*, *60*(2), 109.

Kaufman, J. A., Wright, J. M., Evans, A., Rivera-Núñez, Z., Meyer, A., & Narotsky, M. G. (2020). Disinfection by-product exposures and the risk of musculoskeletal birth defects. *Environmental Epidemiology*, *4*(1), e081.

Kim, D., Kanan, A., Sharma, N., Westerhoff, P., & Karanfil, T. (2020). Total organic halogen (TOX) species formation at different locations in drinking water distribution systems. *Environmental Science: Water Research & Technology*, *6*(9), 2542-2552.

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*(2), 124-127.

Klinefelter, G. R., Strader, L. F., Suarez, J. D., Roberts, N. L., Goldman, J. M., & Murr, A. S. (2004). Continuous exposure to dibromoacetic acid delays pubertal development and compromises sperm quality in the rat. *Toxicological sciences*, *81*(2), 419-429.

Komulainen, H., Vaittinen, S. L., Vartiainen, T., Tuomisto, J., Kosma, V. M., Kaliste-Korhonen, E., ... & K. Tuominen, R. (1997). Carcinogenicity of the drinking water mutagen 3-chloro-4-(dichloromethyl)-5-hydroxy-2 (5H)-furanone in the rat. *Journal of the National Cancer Institute*, *89*(12), 848-856.

Korshin, G. V., Li, C. W., & Benjamin, M. M. (1997). The decrease of UV absorbance as an indicator of TOX formation. *Water Research*, *31*(4), 946-949.

Kramer, M. D., Lynch, C. F., Isacson, P., & Hanson, J. W. (1992). The association of waterborne chloroform with intrauterine growth retardation. *Epidemiology*, 407-413.

Krasner, S. W., McGuire, M. J., Jacangelo, J. G., Patania, N. L., Reagan, K. M., & Aieta, E. M. (1989). The occurrence of disinfection by-products in US drinking water. *Journal-American Water Works Association*, *81*(8), 41-53.

Krasner, S. W., & Wright, J. M. (2005). The effect of boiling water on disinfection by-product exposure. *Water research*, *39*(5), 855-864.

Kristiana, I., McDonald, S., & Joll, C. A. (2020). The forest or the trees: a critical review on the analysis of total organic halogen (TOX) in drinking waters and its utility as a water quality parameter. *Environmental Science: Water Research & Technology*, *6*(9), 2313-2330.

Kumar, S., Forand, S., Babcock, G., & Hwang, S. A. (2014). Total trihalomethanes in public drinking water supply and birth outcomes: a cross-sectional study. *Maternal and child health journal*, *18*(4), 996-1006.

Lahlou, Z. M. (2002). Water quality in distribution systems. *A National Drinking Water, Clearinghouse, Fact Sheet*.

Lange, A. L., & Kawczynski, E. (1978). Controlling organics: the contra Costa County water district experience. *Journal-American Water Works Association*, *70*(11), 653-660.

Langsa, M., Allard, S., Kristiana, I., Heitz, A., & Joll, C. A. (2017). Halogen-specific total organic halogen analysis: Assessment by recovery of total bromine. *Journal of Environmental Sciences*, *58*, 340-348.

Lee, G. F., & Morris, J. C. (1962). Kinetics of chlorination of phenol-chlorophenolic tastes and odors. *Int. J. Air Wat. Poll*, *6*(567), 419-431.

Lee, S. C., Guo, H., Lam, S. M. J., & Lau, S. L. A. (2004). Multipathway risk assessment on disinfection by-products of drinking water in Hong Kong. *Environmental Research*, *94*(1), 47-56.

Levallois, P., Gingras, S., Marcoux, S., Legay, C., Catto, C., Rodriguez, M., & Tardif, R. (2012). Maternal exposure to drinking-water chlorination by-products and small-for-gestational-age neonates. *Epidemiology*, 267-276.

Lewis, C., Suffet, I. H., Hoggatt, K., & Ritz, B. (2007). Estimated effects of disinfection by-products on preterm birth in a population served by a single water utility. *Environmental health perspectives*, *115*(2), 290-295.

Lewis, C., Suffet, I. H., & Ritz, B. (2006). Estimated effects of disinfection by-products on birth weight in a population served by a single water utility. *American journal of epidemiology*, *163*(1), 38-47.

Li, C. W., Korshin, G. V., & Benjamin, M. M. (1998). Monitoring DBP formation with differential UV spectroscopy. *Journal-American Water Works Association*, *90*(8), 88-100.

Li, J., Jiang, J., Manasfi, T., & von Gunten, U. (2020). Chlorination and bromination of olefins: Kinetic and mechanistic aspects. *Water Research*, *187*, 116424.

Li, R., Hopke, P. K., Dozier, A., Thurston, S. W., Thevenet-Morrison, K., Croft, D., ... & Rich, D. Q. (2019). Term birth weight and ambient air pollutant concentrations during pregnancy, among women living in Monroe County, New York. *Journal of exposure science & environmental epidemiology*, *29*(4), 500-509.

Liang, L., & Singer, P. C. (2003). Factors influencing the formation and relative distribution of haloacetic acids and trihalomethanes in drinking water. *Environmental science & technology*, *37*(13), 2920-2928.

Liu, B., & Reckhow, D. A. (2013). DBP formation in hot and cold water across a simulated distribution system: effect of incubation time, heating time, pH, chlorine dose, and incubation temperature. *Environmental science & technology*, *47*(20), 11584-11591.

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*(6), 420-429.

Mabey, W., & Mill, T. (1978). Critical review of hydrolysis of organic compounds in water under environmental conditions. *Journal of Physical and Chemical Reference Data*, *7*(2), 383-415.

MacNeill, A. L. (1994). *Mechanistic modeling of trihalomethanes and other neutral chlorination by-products* (Doctoral dissertation, thesis, University of Massachusetts).

Manivannan, B., & Borisover, M. (2020). Strengths of correlations with formation of chlorination disinfection byproducts: effects of predictor type and other factors. *Environmental Science and Pollution Research*, *27*(5), 5337-5352.

Mao, Y., Wang, X., Yang, H., Wang, H., & Xie, Y. F. (2014). Effects of ozonation on disinfection byproduct formation and speciation during subsequent chlorination. *Chemosphere*, *117*, 515-520.

Martin, J. A., Osterman, M. J., Kirmeyer, S. E., & Gregory, E. C. (2015). Measuring gestational age in vital statistics data: transitioning to the obstetric estimate. *National Vital Statistics Reports: From the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, *64*(5), 1-20.

Martin, K. M. (2000). Are disinfection by-products in tapwater associated with intrauterine growth retardation? Master thesis. McGill University: Montreal, Quebec.

McTigue, N. E., Cornwell, D. A., Graf, K., & Brown, R. (2014). Occurrence and consequences of increased bromide in drinking water sources. *Journal-American Water Works Association*, *106*(11), E492-E508.

Meng Q, Inoue K, Ritz B, Olsen J, Liew Z. (2018). Prenatal Exposure to Perfluoroalkyl Substances and Birth Outcomes; An Updated Analysis from the Danish National Birth Cohort. *International journal of environmental research and public health*. 15(9):1832.

Minear, R. A., & Bird, J. C. (1980). Trihalomethanes: impact of bromide ion concentration on yield, species distribution, rate of formation and influence of other variables. *Water Chlorination, Environmental Impact and Health Effects*, *3*.

Morris, J. C., & Baum, B. (1978). Precursors and mechanisms of haloform formation in the chlorination of water supplies. *Water chlorination: environmental impact and health effects*, *2*.

Muellner, M. G., Wagner, E. D., McCalla, K., Richardson, S. D., Woo, Y. T., & Plewa, M. J. (2007). Haloacetonitriles vs. regulated haloacetic acids: are nitrogen-containing DBPs more toxic?. *Environmental science & technology*, *41*(2), 645-651.

Nielsen, G. L., Sorensen, H. T., Larsen, H., & Pedersen, L. (2001). Risk of adverse birth outcome and miscarriage in pregnant users of non-steroidal anti-inflammatory drugs: population based observational study and case-control study. *Bmj*, *322*(7281), 266-270.

Nieuwenhuijsen, M. J., Martinez, D., Grellier, J., Bennett, J., Best, N., Iszatt, N., ... & Toledano, M. B. (2009). Chlorination disinfection by-products in drinking water and congenital anomalies: review and meta-analyses. *Environmental Health Perspectives*, *117*(10), 1486-1493.

Nokes, C. J., Fenton, E., & Randall, C. J. (1999). Modelling the formation of brominated trihalomethanes in chlorinated drinking waters. *Water research*, *33*(17), 3557-3568.

Obolensky, A., & Singer, P. C. (2005). Halogen substitution patterns among disinfection byproducts in the information collection rule database. *Environmental science & technology*, *39*(8), 2719-2730.

Obolensky, A., Singer, P. C., & Shukairy, H. M. (2007). Information collection rule data evaluation and analysis to support impacts on disinfection by-product formation. *Journal of Environmental Engineering*, *133*(1), 53-63.

Pan, Y., & Zhang, X. (2013). Four groups of new aromatic halogenated disinfection byproducts: effect of bromide concentration on their formation and speciation in chlorinated drinking water. *Environmental science & technology*, *47*(3), 1265-1273.

Pérez-Pavón, José-Luis & Herrero-Martín, Sara & García Pinto, Carmelo & Cordero, Bernardo. (2008). Determination of trihalomethanes in water samples: A review. Analytica chimica acta. 629. 6-23. 10.1016/j.aca.2008.09.042.

Pereira, V. J., Weinberg, H. S., & Singer, P. C. (2004). Temporal and spatial variability of DBPs in a chloraminated distribution system. *Journal-American Water Works Association*, *96*(11), 91-102.

Plewa, M. J., & Wagner, E. D. (2009). Mammalian cell cytotoxicity and genotoxicity of disinfection by-products. *Water Research Foundation: Denver, CO*, 134.

Plewa, M. J., Wagner, E. D., Jazwierska, P., Richardson, S. D., Chen, P. H., & McKague, A. B. (2004). Halonitromethane drinking water disinfection byproducts: chemical characterization and mammalian cell cytotoxicity and genotoxicity. *Environmental science & technology*, *38*(1), 62-68.

Pope, P. G. (2006). *Haloacetic acid formation during chloramination: role of environmental conditions, kinetics, and haloamine chemistry* (Doctoral dissertation).

Porter, C. K., Putnam, S. D., Hunting, K. L., & Riddle, M. R. (2005). The effect of trihalomethane and haloacetic acid exposure on fetal growth in a Maryland county. *American journal of epidemiology*, *162*(4), 334-344.

Postigo, C., Emiliano, P., Barceló, D., & Valero, F. (2018). Chemical characterization and relative toxicity assessment of disinfection byproduct mixtures in a large drinking water supply network. *Journal of hazardous materials*, *359*, 166-173.

Pourmoghaddas, H., Stevens, A. A., Kinman, R. N., Dressman, R. C., Moore, L. A., & Ireland, J. C. (1993). Effect of bromide ion on formation of HAAs during chlorination. *Journal-American Water Works Association*, *85*(1), 82-87.

Prasse, C., Von Gunten, U., & Sedlak, D. L. (2020). Chlorination of phenols revisited: Unexpected formation of α, β-unsaturated C4-dicarbonyl ring cleavage products. *Environmental science & technology*, *54*(2), 826-834.

Qin, L., Zhang, X., Chen, Y., Mo, L., Zeng, H., & Liang, Y. (2017). Predictive QSAR models for the toxicity of disinfection byproducts. *Molecules*, *22*(10), 1671.

Raz, R., Kioumourtzoglou, M. A., & Weisskopf, M. G. (2018). Live-birth bias and observed associations between air pollution and autism. *American journal of epidemiology*, *187*(11), 2292-2296.

Rahman, M. B., Armstrong, B., Cowie, C., & Driscoll, T. (2007). Disinfection by-products in drinking water and colorectal cancer: A meta-analysis. *Epidemiology*, *18*(5), S211.

Rahman, M. B., Driscoll, T., Cowie, C., & Armstrong, B. K. (2010). Disinfection by-products in drinking water and colorectal cancer: a meta-analysis. *International journal of epidemiology*, *39*(3), 733-745.

Reckhow, D. (1985). Organic halide formation and the use of pre-ozonation and alum coagulation to control organic halide precursors.

Reckhow, D. A., & Singer, P. C. (1985). Mechanisms of organic halide formation during fulvic acid chlorination and implications with respect to preozonation. *Water chlorination: chemistry, environmental impact and health effects*, *5*, 1229-1257.

Reckhow, D.A. and Singer, P.C. (2011). Formation and control of disinfection by-products. Water quality and treatment: a handbook on drinking water. 6th ed. New York: McGraw-Hill/American Water Works Association.

Reckhow, D. A., Singer, P. C., & Malcolm, R. L. (1990). Chlorination of humic materials: byproduct formation and chemical interpretations. *Environmental science & technology*, *24*(11), 1655-1664.

Regli, S., Chen, J., Messner, M., Elovitz, M. S., Letkiewicz, F. J., Pegram, R. A., ... & Wright, J. M. (2015). Estimating potential increased bladder cancer risk due to increased bromide concentrations in sources of disinfected drinking waters. *Environmental science & technology*, *49*(22), 13094-13102.

Ribble, F., PhD, M., & Keddie, M. S. P. H. (2001). Understanding the Hispanic paradox. *Ethn Dis*, *11*(3), 496-518.

Rice, G. E., Teuschler, L. K., Bull, R. J., Simmons, J. E., & Feder, P. I. (2009). Evaluating the similarity of complex drinking-water disinfection by-product mixtures: overview of the issues. *Journal of Toxicology and Environmental Health, Part A*, *72*(7), 429-436.

Richardson, S. D. (2003). Disinfection by-products and other emerging contaminants in drinking water. *TrAC Trends in Analytical Chemistry*, *22*(10), 666-684.

Richardson, S. D., Fasano, F., Ellington, J. J., Crumley, F. G., Buettner, K. M., Evans, J. J., ... & Plewa, M. J. (2008). Occurrence and mammalian cell toxicity of iodinated disinfection byproducts in drinking water. *Environmental Science & Technology*, *42*(22), 8330-8338.

Rivera-Núñez, Z., Wright, J. M., Blount, B. C., Silva, L. K., Jones, E., Chan, R. L., ... & Savitz, D. A. (2012). Comparison of trihalomethanes in tap water and blood: a case study in the United States. *Environmental health perspectives*, *120*(5), 661-667.

Rivera-Núñez, Z., & Wright, J. M. (2013). Association of brominated trihalomethane and haloacetic acid exposure with fetal growth and preterm delivery in Massachusetts. *Journal of occupational and environmental medicine*, *55*(10), 1125-1134.

Roccaro, P., Chang, H. S., Vagliasindi, F. G., & Korshin, G. V. (2013). Modeling bromide effects on yields and speciation of dihaloacetonitriles formed in chlorinated drinking water. *Water research*, *47*(16), 5995-6006.

Roccaro, P., Korshin, G. V., Cook, D., Chow, C. W., & Drikas, M. (2014). Effects of pH on the speciation coefficients in models of bromide influence on the formation of trihalomethanes and haloacetic acids. *Water research*, *62*, 117-126.

Rook, J. J. (1974). Formation of haloforms during chlorination of natural waters *J Soc Water Treat Exam* 23: 234–243.

Rougé, V., von Gunten, U., & Allard, S. (2020). Efficiency of pre-oxidation of natural organic matter for the mitigation of disinfection byproducts: Electron donating capacity and UV absorbance as surrogate parameters. *Water Research*, *187*, 116418.

Samson, C. C., Seidel, C. J., Summers, R. S., & Bartrand, T. (2017). Assessment of HAA9 occurrence and THM, HAA speciation in the United States. *Journal-American Water Works Association*, *109*(7), E288-E301.

Säve-Söderbergh, M., Toljander, J., Donat-Vargas, C., Berglund, M., & Åkesson, A. (2020). Exposure to Drinking Water Chlorination by-Products and Fetal Growth and Prematurity: A Nationwide Register-Based Prospective Study. *Environmental health perspectives*, *128*(5), 057006.

Savitz, D. A., Singer, P. C., Herring, A. H., Hartmann, K. E., Weinberg, H. S., & Makarushka, C. (2006). Exposure to drinking water disinfection by-products and pregnancy loss. *American Journal of Epidemiology*, *164*(11), 1043-1051.

Schempf, A. H., & Strobino, D. M. (2008). Illicit drug use and adverse birth outcomes: is it drugs or context? *Journal of Urban Health*, *85*(6), 858-873.

Shaw, G. M., Wasserman, C. R., Lammer, E. J., O'Malley, C. D., Murray, J. C., Basart, A. M., & Tolarova, M. M. (1996). Orofacial clefts, parental cigarette smoking, and transforming growth factor-alpha gene variants. *American journal of human genetics*, *58*(3), 551.

Silva, L. K., Backer, L. C., Ashley, D. L., Gordon, S. M., Brinkman, M. C., Nuckols, J. R., ... & Blount, B. C. (2013). The influence of physicochemical properties on the internal dose of trihalomethanes in humans following a controlled showering exposure. *Journal of exposure science & environmental epidemiology*, *23*(1), 39-45.

Spencer, K., Cowans, N. J., Avgidou, K., Molina, F., & Nicolaides, K. H. (2008). First-trimester biochemical markers of aneuploidy and the prediction of small-for-gestational age fetuses. *Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, *31*(1), 15-19.

Summerhayes, R. J., Rahman, B., Morgan, G., Beresin, G., Moreno, C., & Wright, J. M. (2020). Meta-Analysis of Small for Gestational Age Births and Disinfection Byproduct Exposures. *Environmental Research*, 110280.

Summers, R. S., Benz, M. A., Shukairy, H. M., & Cummings, L. (1993). Effect of separation processes on the formation of brominated THMs. *Journal-American Water Works Association*, *85*(1), 88-95.

Suzuki, G., Michinaka, C., Matsukami, H., Noma, Y., & Kajiwara, N. (2020). Validity of using a relative potency factor approach for the risk management of dioxin-like polychlorinated naphthalenes. *Chemosphere*, *244*, 125448.

Symanski, E., Savitz, D. A., & Singer, P. C. (2004). Assessing spatial fluctuations, temporal variability, and measurement error in estimated levels of disinfection by-products in tap water: implications for exposure assessment. *Occupational and environmental medicine*, *61*(1), 65-72.

Tan, J., Allard, S., Gruchlik, Y., McDonald, S., Joll, C. A., & Heitz, A. (2016). Impact of bromide on halogen incorporation into organic moieties in chlorinated drinking water treatment and distribution systems. *Science of The Total Environment*, *541*, 1572-1580.

Tapuhi, E., & Jencks, W. P. (1982). Base-catalyzed halogenation of acetone. *Journal of the American Chemical Society*, *104*(21), 5758-5765.

Teuschler, L. K., & Simmons, J. E. (2003). Approaching DBP toxicity as a mixtures problem. *Journal-American Water Works Association*, *95*(6), 131-138.

Tomeo, C. A., Rich-Edwards, J. W., Michels, K. B., Berkey, C. S., Hunter, D. J., Frazier, A. L., ... & Buka, S. L. (1999). Reproducibility and validity of maternal recall of pregnancy-related events. *Epidemiology*, 774-777.

Vahala, R., Långvik, V. A., & Laukkanen, R. (1999). Controlling adsorbable organic halogens (AOX) and trihalomethanes (THM) formation by ozonation and two-step granule activated carbon (GAC) filtration. *Water science and technology*, *40*(9), 249-256.

Villanueva, C. M., Cantor, K. P., Cordier, S., Jaakkola, J. J., King, W. D., Lynch, C. F., ... & Kogevinas, M. (2004). Disinfection byproducts and bladder cancer: a pooled analysis. *Epidemiology*, 357-367.

Villanueva, C. M., Castaño-Vinyals, G., Moreno, V., Carrasco-Turigas, G., Aragonés, N., Boldo, E., ... & Kogevinas, M. (2012). Concentrations and correlations of disinfection by-products in municipal drinking water from an exposure assessment perspective. *Environmental research*, *114*, 1-11.

von Gunten, U., & Hoigné, J. (2017). Ozonation of bromide-containing waters: bromate formation through ozone and hydroxyl radicals. In *Disinfection By-Products in Water Treatment* (pp. 187-206). CRC Press.

Wang, Y., Small, M. J., & VanBriesen, J. M. (2017). Assessing the risk associated with increasing bromide in drinking water sources in the Monongahela River, Pennsylvania. *Journal of Environmental Engineering*, *143*(3), 04016089.

Wei, X., Yang, M., Zhu, Q., Wagner, E. D., & Plewa, M. J. (2020). Comparative Quantitative Toxicology and QSAR Modeling of the Haloacetonitriles: Forcing Agents of Water Disinfection Byproduct Toxicity. *Environmental science & technology*, *54*(14), 8909-8918.

Weinberg, H. S., Krasner, S. W., Richardson, S. D., & Thruston Jr, A. D. (2002). The occurrence of disinfection by-products (DBPs) of health concern in drinking water: results of a nationwide DBP occurrence study. *National Exposure Research Laboratory, Office of Research and Development, US Environmental Protection Agency*.

Wenk, J., Aeschbacher, M., Salhi, E., Canonica, S., Von Gunten, U., & Sander, M. (2013). Chemical oxidation of dissolved organic matter by chlorine dioxide, chlorine, and ozone: effects on its optical and antioxidant properties. *Environmental science & technology*, *47*(19), 11147-11156.

Westerhoff, P., Chao, P., & Mash, H. (2004). Reactivity of natural organic matter with aqueous chlorine and bromine. *Water Research*, *38*(6), 1502-1513.

Wigle, D. T., Arbuckle, T. E., Turner, M. C., Bérubé, A., Yang, Q., Liu, S., & Krewski, D. (2008). Epidemiologic evidence of relationships between reproductive and child health outcomes and environmental chemical contaminants. *Journal of Toxicology and Environmental Health, Part B*, *11*(5-6), 373-517.

Wright, J. M., Evans, A., Kaufman, J. A., Rivera-Núñez, Z., & Narotsky, M. G. (2017). Disinfection by-product exposures and the risk of specific cardiac birth defects. *Environmental health perspectives*, *125*(2), 269-277.

Wright, J. M., Schwartz, J., & Dockery, D. W. (2003). Effect of trihalomethane exposure on fetal development. *Occupational and environmental medicine*, *60*(3), 173-180.

Wright, J. M., Schwartz, J., & Dockery, D. W. (2004). The effect of disinfection by-products and mutagenic activity on birth weight and gestational duration. *Environmental health perspectives*, *112*(8), 920-925.

Wright, J. M., Schwartz, J., Vartiainen, T., Mäki-Paakkanen, J., Altshul, L., Harrington, J. J., & Dockery, D. W. (2002). 3-Chloro-4-(dichloromethyl)-5-hydroxy-2 (5H)-furanone (MX) and mutagenic activity in Massachusetts drinking water. *Environmental health perspectives*, *110*(2), 157-164.

Wu, W. W., & Chadik, P. A. (1998). Effect of bromide ion on haloacetic acid formation during chlorination of Biscayne aquifer water. *Journal of Environmental Engineering*, *124*(10), 932-938.

Xiang, W., Chang, J., Qu, R., Albasher, G., Wang, Z., Zhou, D., & Sun, C. (2021). Transformation of bromophenols by aqueous chlorination and exploration of main reaction mechanisms. *Chemosphere*, *265*, 129112.

Yan, M., Li, M., Roccaro, P., & Korshin, G. V. (2016). Ternary model of the speciation of I-/Br-/Cl-trihalomethanes formed in chloraminated surface waters. *Environmental science & technology*, *50*(8), 4468-4475.

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*(3), 390-395.

Zhang, C., Li, M., Han, X., & Yan, M. (2018). Quantum chemical examination of the sequential halogen incorporation scheme for the modeling of speciation of I/Br/Cl-containing trihalomethanes. *Environmental science & technology*, *52*(4), 2162-2169.

Zhang, P., LaPara, T. M., Goslan, E. H., Xie, Y., Parsons, S. A., & Hozalski, R. M. (2009). Biodegradation of haloacetic acids by bacterial isolates and enrichment cultures from drinking water systems. *Environmental science & technology*, *43*(9), 3169-3175.

Zhang, X., & Minear, R. A. (2002). Decomposition of trihaloacetic acids and formation of the corresponding trihalomethanes in drinking water. *Water research*, *36*(14), 3665-3673.

Zhang, X., Echigo, S., Minear, R. A., & Plewa, M. J. (2000). Characterization and comparison of disinfection by-products of four major disinfectants.

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.