DISSERTATION

DISINFECTION BY-PRODUCTS AND PRENATAL DEVELOPMENT

Submitted by

Alison Fay Hinckley

Department of Environmental and Radiological Health Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2003

COLORADO STATE UNIVERSITY

October 31, 2003

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY ALISON FAY HINCKLEY ENTITLED "DISINFECTION BY-PRODUCTS AND PRENATAL DEVELOPMENT" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Melvin Andersen, Ph.D.

Annette Bachand, Ph.D.

Henry Gardner, Jr., MSPH, DrPH

Rao Veeramachaneni, BVSc, MScVet, PhD

Advisor - John Reif, D.V.M., M.Sc.

Department Head – John Zimbrick, Ph.D.

ABSTRACT OF DISSERTATION

DISINFECTION BY-PRODUCTS AND PRENATAL DEVELOPMENT

While toxicologic studies have suggested modes of action for disinfection by-product (DBP) toxicity or teratogenicity, there is little evidence for a causal relationship between DBPs and induction of adverse outcomes in humans. This may be due, in part, to biases inherent to exposure assessment or analytical methodologies used in past studies. This dissertation presents three approaches designed to develop and evaluate new methods of minimizing exposure misclassification and improving risk estimation for studies of DBPs and prenatal development. These approaches are described below.

Past epidemiological studies of DBPs and birth outcomes have used exposure measurements for study participants that are susceptible to misclassification due in part to intra-system spatial variation. Spatial variability occurs when DBP contaminants in treated drinking water do not occur uniformly across individual public water distribution systems. We conducted a feasibility study to identify and test methods that can be used to select communities that are served by water distribution systems with low spatial variability for epidemiologic investigations. We also aimed to identify guidelines for characterization of brominated compounds. These methods were tested on quarterly DBP concentrations from the distribution system of 198 facilities, listed in USEPA's Information Collection Rule. We identified sites with low spatial variability that had high overall levels of brominated trihalomethanes (THMs), high overall levels of chlorinated

THMs and low levels of THM species. We classified spatial variation of THMs within these distribution systems using methods based upon: overall levels of chlorinated THMs and low levels of total THM species. We classified spatial variation of THMs within these distribution systems using methods based upon: (1) two-way analysis of variance, and (2) cutpoints deemed biologically important in epidemiologic literature. The method based upon epidemiologic literature was determined to be superior, as the two-way ANOVA method utilized arbitrary significance levels to ascertain low spatial variability. Using the epidemiologic literature method, we identified twenty sites with low spatial variability, one of which had consistently high THM levels that were primarily brominated. This study presents a simple method for *a priori* selection of sites with low spatial variability as a means to reduce misclassification in exposure assessment for epidemiologic studies of DBPs.

Previous epidemiologic studies of DBPs and growth-related birth outcomes such as preterm birth or low birth weight have not attempted to estimate exposure over the third trimester, the time period of biological importance to induction of these outcomes. We conducted a study to identify the existence and magnitude of biases that arise from use of variable third trimester lengths and regression techniques (Dodds et al., 1999; King et al., 2000; Dodds and King, 2001), when temporally variable environmental exposures, such as DBPs, are evaluated against time-dependent outcomes. Using examples from a simulated population modeled after the U.S. distribution of normal and preterm births, we evaluated how selection of participants and cutpoint use affected risk estimation. We found that failure to adjust for variable length of third trimesters and use of cutpoints with

regression can potentially lead to significant bias away from or towards the null, depending on peak or trough exposure and placement of cutpoint, and further complicate estimation of true risk in epidemiologic studies that likely contain biases from other sources such as inadequate or improper methods to assess exposure.

Toxicological studies have demonstrated associations between exposure to high concentrations of DBPs and several teratogenic outcomes. Exposure to high doses of haloacetic acids (HAAs), the second most commonly occurring group of DBPs, causes cardiovascular and orofacial malformations, such as cleft palate, in laboratory studies. Human studies have focused primarily on low-dose exposures to THMs. Most epidemiological studies have not evaluated exposure over time periods of pregnancy that are specific to induction of the specific outcome. We conducted a retrospective cohort study of cleft palate to determine whether a relationship exists with exposure to THMs or the five most prevalent individual haloacetic acids (HAA5), which include trichloroacetic acid (TCAA). The study population consisted of all singleton births in a large community served by three water treatment plants. This community was chosen since THMs and HAAs varied by season with levels ranging from low to moderate, but exhibited little spatial variability within the distribution system. Birth certificate data (n=51,717) were obtained and logistic regression techniques were used to estimate the relationship between mean exposures to DBPs over the first trimester, second month, and 6^{th}, 7^{th} and 8^{th} weeks of pregnancy and cleft palate.

We found no increased risk with categorical levels of exposure to THMs or HAAs during critical time windows. In addition, no associations were observed for any DBP evaluated as a continuous variable. However, the power of this study was low due to a small number of cleft palate cases.

Alison Fay Hinckley
Department of Environmental and Radiological Health Sciences
Colorado State University
Fort Collins, CO 80523
Fall 2003

ACKNOWLEDGEMENTS

This dissertation could not have been completed without the generous guidance, both intellectually and personally, of many individuals. In particular, I thank my advisor, Dr. John Reif, for his patience, persistence and candor. I also thank my other committee members, Drs. Annette Bachand, Mel Andersen, Hank Gardner and Rao Veeramachaneni, for their time, advice and good humor. I am indebted to Dr. Annette Bachand for her superb statistical and methodological counsel and unwavering cheerfulness.

I wish to acknowledge my surrogate family at 350 West Lake Street, the Department of Environmental Health. I thank Ms. Pat M. Brown for her expertise, gentle nature and for providing me sanctuary on days when I preferred not to be in my own office. Thanks to Mr. Tom Luben for his help in development of the literature review table and loquacious nature for one of my favorite topics, disinfection by-products. I also sincerely appreciate the computer knowledge of Mr. Philip Riggs and Mr. Ryan Miller during my simulation-modeling phase. I am grateful to have had support from my fellow students, particularly my officemates and workout partners, both past and present. I also owe Dr. Susan Hariri, my "partner in crime", immense gratitude for keeping me inspired, politically informed and, at times, sane.

Lastly, this endeavor could not have occurred without the unconditional support of my family. Accordingly, I thank my mother and father, Bernice and Charles Hinckley, for

being present in my thoughts every day of my life. I also thank my sister, Jackie Hinckley, for her unique perspective and encouragement. The last member of the Hinckley clan, Jennifer Hinckley, is tremendously important, for she provided pajamas and cookie dough and true sisterly love. Most of all, I thank my husband, Kamran Azad, without whose belief, incomparable friendship, and devotion none of this would have been possible. I dedicate this dissertation to him.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

CHAPTER 1

1 INTRODUCTION

1.1 BACKGROUND AND AIM

Chlorination is the primary method of disinfection of municipal water supplies in North America and constitutes a vital component in protecting the public from water borne infectious agents. The discovery that a large number of byproducts are formed by the reaction of chlorine with naturally occurring organic materials such as humic and fulvic acids raised concern about potential health effects that might result from exposures to these compounds. The chemical mixture of disinfection byproducts (DBPs) has not been fully characterized, but is known to contain trihalomethanes (THMs), haloacetic acids (HAAs), haloacetonitriles and other classes of chemicals, including some that are mutagenic or carcinogenic in laboratory animals. Concerns have been raised regarding the potential effects of by-products on reproductive outcomes, supported in part by the findings that some by-products cause reproductive and developmental toxicity in laboratory animals, albeit at higher doses than those encountered by humans (U.S. EPA/ILSI, 1993; Reif et al, 1996). However, due to the common exposure by humans to chlorinated water supplies, there is significant interest in any health outcome that may be associated with drinking water.

Limitations in exposure assessment have restricted the validity of many previous studies of DBPs (Reif *et al.*, 1996; Nieuwenhuijsen *et al.*, 2000). The quarterly or annual DBP average for the water utility has been used most often to generate an estimate for an individual's exposure at a specific point in time. Misclassification of exposure to DBPs has resulted from several sources, including: inappropriate selection of DBP markers, temporal and spatial variation in DBP concentrations, unmeasured use of bottled or filtered water, variations in individual showering/bathing and household water use, variability between residential and occupational exposures to water, residential mobility during pregnancy and mixing between distribution systems of multiple-plant utilities (Shimokura *et al*, 1998; Swan *et al*, 1998; Zender *et al,* 2001; Shaw *et al*, 1992; Khoury *et al*, 1988). In addition, most epidemiological studies have not evaluated exposure over time periods of pregnancy that are specific to induction of specific birth outcomes. Despite the knowledge and sources of these potential biases, few epidemiologic studies have attempted to account for or minimize their effects.

The overall goal of this dissertation was to develop and examine methods to improve exposure assessment and risk estimation in epidemiologic studies of DBPs and adverse birth outcomes. To that end, our specific objectives were to 1) conduct a feasibility study to identify and test methods that can be used to select communities for epidemiologic investigations that are served by water distribution systems with low spatial variability of DBPs and to identify guidelines for characterization of brominated compounds; 2) identify the existence and magnitude of biases that arise from use of

variable third trimester lengths, when temporally variable environmental exposures, such as DBPs, are evaluated against time-dependent birth outcomes, and; 3) conduct a retrospective cohort study of cleft palate, evaluating DBP exposure over specific times of gestation, to determine if a relationship exists and whether magnitude of risk depends upon timing of exposure.

1.2 DISSERTATION OVERVIEW

This dissertation is organized into three related projects regarding epidemiologic evaluation of DBPs and prenatal exposure. Therefore, each chapter is organized individually and includes an introduction, materials and methods, results, discussion and references section.

The literature review, chapter 2, presents all information related to epidemiologic investigation of DBPs and birth outcomes that is presently available. Included within this chapter are comprehensive evaluations of previous quantitative studies and relevant exposure assessment literature.

Chapter 3 presents the details of a feasibility study to minimize exposure misclassification due to spatial variability of THMs and HAAs within distribution systems. This study had the following hypotheses: (1) A simple method can be developed for *a priori* selection of sites with low spatial variability as a means to reduce misclassification in exposure assessment for epidemiologic studies of THMs and HAAs. (2) Standard measures can be identified and used to classify a site as having predominantly brominated DBPs.

Chapter 4 is an examination of the existence and magnitude of biases that arise from use of variable third trimester lengths for studies of DBPs and time-dependent outcomes. The specific hypothesis of this study was: Cutpoints of exposure over the variable-length third trimester and selection techniques for study participants can bias associations in studies of the relationship between preterm birth and temporally variable DBPs.

In the fifth chapter, exposure to THMs and HAAs is evaluated over time periods of pregnancy that are specific to induction of cleft palate. We investigated the following hypothesis: Women exposed to high levels of THMs and HAAs during specific critical periods of gestation are at increased risk of delivering babies with cleft palate compared to those exposed during less critical time periods.

Chapter 6 provides a summary of findings and a discussion of recommendations for future research efforts.

CHAPTER 2

2 LITERATURE REVIEW

2.1 INTRODUCTION

Chlorination is the primary method of disinfection of municipal water supplies in North America and is an essential action to protecting the public from water borne infectious agents. In the late 1970s, the discovery of a large number of by-products, resulting from disinfection techniques, raised concern about potential health effects that might result from exposures to these compounds (Rook, 1974; Bellar and Lichenberg, 1974). In particular, there are concerns over the potential ability of by-products to cause adverse birth outcomes. This fear is supported in part by the findings that some by-products cause reproductive and developmental toxicity in laboratory animals, albeit at higher doses than those encountered by humans (U.S. EPA/ILSI, 1993; Reif et al., 1996). Nonetheless, due to the *extensive* human exposure to chlorinated water supplies there is significant interest in any water-related health effects.

2.2 DISINFECTION BY-PRODUCTS

Several hundred organic disinfection by-products (DBPs), many of which are mutagenic or carcinogenic in test systems, are formed by the reaction of chlorine with naturally occurring organic materials such as humic and fulvic acids in surface waters. The chemical mixture of DBPs contains trihalomethanes, haloacetic acids, haloacetonitriles and a large number of unidentified compounds (Nieuwenhuijsen et al., 2000a).

Trihalomethanes (THMs) are the predominant class of DBPs formed during chlorination of surface waters and include chloroform, bromoform, bromodichloromethane (BDCM) and dibromochloromethane (DBCM). Non-volatile haloacetic acids (HAAs), the next most common subgroup of disinfection by-products, include monochloroacetic acid (MCAA), dichloroacetic acid (DCAA), trichloroacetic acid (TCAA), monobromoacetic acid (MBAA), and dibromoacetic acid (DBAA). DBP formation, concentration and speciation are a function of organic matter quantity, chlorine dose, pH, temperature, contact time and bromide ion concentration. High concentrations of bromide ion can also shift the chemical reaction towards production of the brominated species of trihalomethanes and other classes of DBPs (Krasner et al., 1989). The formation of DBPs by chlorination of surface water varies seasonally, with highest concentrations typically found during the summer months.

2.2.1 Exposure to DBPs

Human exposure to DBPs occurs through ingestion, by dermal absorption and, for the volatile THMs, through inhalation during showering, bathing and other water-related activities (Nieuwenhuijsen et al., 2000a). To monitor this exposure, the Safe Drinking Water Act (SDWA) mandated community water systems serving at least 10,000 people to sample their supplies quarterly for trihalomethanes and haloacetic acids. In 1992, the Environmental Protection Agency (EPA) proposed maximum contaminant levels (MCLs) for DBPs as amendments to the SDWA. These contaminant levels were based on available epidemiological and toxicological data and led to promulgation of the Stage 1 DBP Rule, finalized in November 1998 (U.S.EPA, 1998). Under these rules, maximum

contaminant levels (MCLs) were set at 80 μg/l for total trihalomethanes (TTHMs) and at 60 μg/l for the five most prevalent haloacetic acids (HAA5).

Exposure assessment weaknesses have limited the validity of many previous studies of DBPs and reproductive outcomes (Reif et al., 1996; Nieuwenhuijsen et al., 2000a). The quarterly or annual DBP average for the water utility has been used most often to generate an estimate for an individual's exposure at a specific point in time, as needed for study of certain reproductive outcomes. Misclassification of exposure to DBPs has resulted from: inappropriate selection of DBP markers, temporal and spatial variation in DBP concentrations, unmeasured use of bottled or filtered water, variations in individual showering/bathing and household water use, variability between residential and occupational exposures to water, residential mobility during pregnancy and mixing between distribution systems of multiple-plant utilities (Khoury et al., 1988; Shaw et al., 1992; Shimokura et al., 1998; Swan et al., 1998; Zender et al., 2001; Keegan et al., 2001). Depending upon the species of DBP and assay technique, laboratory methods of assessing DBP constituents also add uncertainty to exposure predictions. Collection of data regarding individual-level water ingestion and exposure to potential confounders is also important for minimization of bias (Reif et al., 1996).

2.2.2 Birth Outcomes Evaluated in Previous Studies of DBPs

Birth outcomes in humans that have been evaluated for associations with DBPs in previous studies include: fetal growth, somatic parameters, neonatal jaundice, preterm delivery, spontaneous abortion, stillbirth, and developmental anomalies. The outcomes

most relevant to this research, fetal growth, viability and malformations are defined below.

2.2.2.1 Fetal Growth

Fetal growth is an indicator of an infant's probability of survival, with smaller babies having increased mortality. In studies of DBPs, infants born weighing less than 2500 g have been categorized as low birth weight (LBW), which includes a large number of babies that are small due to a premature (or preterm) birth. In 1997, 7.5 percent of all U.S. births were classified as LBW. Very low birth weight (VLBW) describes births weighing less than 1,500 g. Small fetuses weighing less than 500 g seldom survive, while those weighing 500-1000 g have increased chances of survival within medical facilities (Sadler, 1995).

Preterm birth is the term used to describe infants born at less than 37 weeks gestation. The incidence of prematurity among pregnant women in the United States is about 7 to 9 percent of Caucasian pregnancies and about 17 percent of African-American pregnancies (Clayman, 1989).

Epidemiologists identify growth-retarded infants as those babies that are smallest at birth for each gestational age. Usually the smallest 5^{th} or 10^{th} percentile of birth weights by race and gender for a given population is characterized as small for gestational age (SGA) or intrauterine growth retarded (IUGR). Several studies of exposure to DBPs have included this percentile-based definition as an indicator of poor fetal growth.

2.2.2.2 Fetal Viability

Recognized pregnancy loss that occurs in the first 20 weeks of gestation is termed spontaneous abortion (SAB). Approximately 10 to 15 percent of recognized pregnancies are lost before 20 weeks gestation. Death of a fetus or loss of a pregnancy after 20 weeks gestation is referred to as a stillbirth. In the U.S., the incidence of stillbirth decreased from 19.2 stillbirths/1,000 total births in 1950 to 9.2 stillbirths/1000 total births in 1980 due to better prenatal and obstetric care. The precise cause of stillbirth is unknown in the majority of cases (O'Rahilly and Müller, 2001). However, this fetal death outcome frequently occurs in babies with severe developmental malformations, particularly anencephaly, spina bifida, or hydrocephalus. In addition, it may occur with maternal disorders limiting placental function and oxygen supply to the fetus, such as antepartum hemorrhage and hypertension (Clayman, 1989).

2.2.2.3 Malformations

The most common teratogenic outcomes evaluated with human exposure to DBPs have been oral cleft, major cardiac and neural tube defects.

Cleft lip and cleft palate are among the most prevalent congenital anomalies at birth. Cleft lip (with or without cleft palate) occurs in about 1:1000 Caucasian births and can be more frequent in males and babies of Asian descent. In addition, this defect is commonly associated with other anomalies. Cleft lip with cleft palate is primarily of genetic origin. Causes of cleft palate (not including the lip) can include combined genetic and environmental factors, mutant genes, chromosomal aberrations and specific

environmental agents (e.g. hydatoin). Cleft palate is slightly more common in females and presents in approximately 0.4:1000 births (O'Rahilly and Müller, 2001).

The cardiovascular system is the first major organ system to begin operating in the embryo. Anomalies of this system are not rare. Cardiac system defects are present in 5:1000 to 10:1000 live births, with nearly one-third of these conditions considered severe. The incidence of congenital cardiac anomalies present at birth is certainly diminished by spontaneous prenatal death, which is more likely to occur to cases that also possess chromosomal abnormalities. The four most common cardiac malformations at birth are: ventricular septal defects, atrial septal defects, persistent ductus arteriosus, and pulmonary stenosis (O'Rahilly and Müller, 2001).

Spina bifida and anencephaly are the two most common forms of neural tube defects (NTDs), comprising nearly 95 percent of all NTD births. NTDs occur in 1:1000 to about 5:1000 live births and are much more frequent in early stage pregnancies, with more than 90 percent of affected embryos lost before the fetal stages (Moore and Persaud, 1998).

2.3 EXPOSURE ASSESSMENT OF DBPS

Historically, DBP investigations have used both qualitative and quantitative exposure assessment methods to examine associations between DBPs and adverse birth outcomes. See Table 2.3.1 for a quick-reference of all studies.

2.3.1 Qualitative Studies of DBPs

Qualitative studies use surrogate parameters for exposure assessment, in lieu of measured DBP concentrations. Qualitative approaches began in the 1980's (Tuthill et al., 1982) with a comparison of risks after exposure to a variety of source waters (ground vs. surface), disinfection methods (chlorine, chlorine dioxide and chloramination), treatment (yes/no) or other DBP related parameters. Although these studies did not provide risk estimates for exposure to specific by-products or class of by-products, they may still be useful for indicating whether by-product(s), individually or grouped, may be present and responsible for adverse reproductive effects.

The first of the recent qualitative studies was a hospital based case-control study of drinking water quality with 1,039 congenital anomalies, 77 stillbirths, 55 neonatal deaths and 1,177 controls in Massachusetts (Aschengrau et. al., 1993). Exposure to tap water from sources in 155 cities was assessed for 2,348 women according to source and method of disinfection. Potential confounders included maternal age, pregnancy history, alcohol consumption, ethnicity, insurance status and other water contaminants. An increased, although not statistically significant, risk for stillbirth (adjusted OR = 2.6, 95% CI 0.9-7.5) was observed for exposure to chlorinated compared to chloraminated drinking water. Additionally, risk of any major congenital malformations (adjusted OR = 1.5, 95% CI 0.7-2.1) was increased. This was mainly attributed to increased risk of defects in the urinary tract (adjusted OR = 4.1, 95% CI 1.2-14.1) and respiratory tract systems (adjusted OR = 3.2, 95% CI 1.1-9.5). The lack of precise risk estimates was likely due to the absence of sufficient case subjects for analyses of specific defect groups. Strengths of

this study include use of the first trimester address to match water data and analysis of a large number of risk factors. Exclusions of women due to missing records, however, limited external validity to predominantly older, white, educated women with insurance. In addition, misclassification of pregnancy outcomes may have occurred for less recognizable or minor defects, biasing results towards the null. Lastly, adjustment for water parameters related to DBP formation may have been inappropriate.

A cross-sectional study of somatic parameters at birth and disinfection method was conducted at two hospitals in Genoa and Chiavari, Italy, during 1988 and 1989 (Kanitz et al., 1996). Data extracted from 676 hospital records were used to evaluate body length, head circumference, LBW and preterm birth. The exposed group was identified as having a maternal residential address in Genoa, which was served water from a treatment plants using either chlorine, chlorine dioxide or a combination of chlorine and chlorine dioxide as methods of disinfection. Residents of Chiavari, who received only untreated water, comprised the referent group. After adjustment for potential confounders including maternal age, education, smoking, income and gender of child, small increases in risk were found for exposure to chlorinated water. The estimated risk of LBW ($\leq$2500 g) was increased for women who drank water disinfected with chlorine dioxide, chlorine or both, although small numbers in the referent group likely biased the estimate. Odds ratios for infant births with shorter body lengths ($\leq$49.5 cm) were slightly increased and statistically significant for exposure to chlorine (adjusted OR = 2.3, 95% CI 1.3- 4.2) and chlorine dioxide (adjusted OR = 2.0, 95% CI 1.2- 3.3). The odds of smaller cranial circumferences in the exposed were also increased compared to the unexposed for water

treated with chlorine and chlorine dioxide (adjusted OR = 2.3, 95% CI 1.6- 5.3), chlorine alone (adjusted OR = 3.5, 95% CI 2.1- 8.5) and chlorine dioxide alone (adjusted OR = 2.2, 95% CI 1.4- 3.9).

Stratification by maternal age group for this Italian study showed that women giving birth that were older than 30 had more differences between exposed and unexposed for somatic parameters than younger women. In addition, the referent group for the older women included more births that were higher birth weight, longer body length and larger head circumferences than births to the younger referent group. Nonetheless, results for the exposed were similar for both age groups. While no information on the women excluded from the study was provided, an extremely low rate of LBW in the referent group (0.78 %) suggests selection bias. Residual confounding may also be present due to comparison of two distinct communities.

Magnus et al. (1999) performed a population-based cross-sectional study of birth defects and drinking water for births occurring between 1993 and1995 in Norway. In this study, the color of water was considered a surrogate for dissolved organic carbon, a precursor to DBP formation, and estimates of color and chlorination were available in the Norwegian Waterworks Registry. This information was coupled with 141,077 newborns from the national birth registry to assign exposure, where the referent group had low color and chlorination compared to the exposed. Potential confounders including sociodemographic (geographical placement, population density and industry profile) and reproductive history variables (maternal age and parity) were assessed via logistic

regression after independence was confirmed. Adjusted risk estimates for all malformations resulted in a slight increase in risk (adjusted OR = 1.14, 95% CI 0.99-1.31). Analysis of specific malformations yielded statistically significant increases for urinary tract defects (adjusted OR = 1.99, 95% CI 1.10- 3.57) and a slight, albeit non-significant, association for NTDs (adjusted OR = 1.26, 95% CI 0.61- 2.62). No increased risk was observed for cardiac, respiratory tract or oral cleft defects. This evaluation did not include low reliability diagnoses listed in the registry (e.g. congenital dislocation of hip), decreasing the potential for non-differential misclassification of disease.

Jaakola et al. (2001) conducted a follow-up study in Norway of birth weight, LBW, SGA and preterm birth, using the same study population and exposure assessment techniques as in Magnus et al. (1999). In this study, no positive associations with any outcomes were observed. Further, the odds of preterm birth for the population exposed to chlorinated water with high color were lower (adjusted OR = 0.91, 95% CI 0.84-0.99), compared to the unexposed group. According to the authors, this risk reduction was potentially due to the capacity of residual chlorine, present in drinking water, to reduce infections during pregnancy.

While no DBP concentrations specific to municipalities or waterworks were available in either Magnus et al. (1999) or Jaakola et al. (2001), national averages for 1995 were found to be low in both TTHMs (9.4 μg/l) and HAAs (14.6 μg/l), compared to U.S. national estimates. Low concentrations of DBPs may, therefore, account for lack of association with some specific birth defects and fetal growth outcomes.

Källén and Robert (2000) conducted another population-based study of birth outcomes and disinfection method in Sweden. Somatic parameters, multiple births, congenital malformations and childhood cancer were some of the outcomes compared for pregnant women exposed to water disinfected by chlorine, chlorine dioxide or no method between 1985 and 1994. Of the 115,801 births included in the study, the majority (75,187 births) was not exposed to any disinfection method. Mantel-Haenszel analysis was used to determine ORs and all analyses, except childhood cancer, were stratified by year of birth, and maternal age. In some cases, stratification by county of birth replaced control for maternal education or smoking. Slight, but significant, associations were obtained for exposure to chlorine disinfection and preterm birth, reduced body mass index, reduced body length, and small head circumference. Definitions for small head circumference and reduced body length, however, are different than those presented in the study by Kanitz et al. (1996), thus results may not be comparable. No significant associations were detected for exposure to chlorine dioxide. Associations for childhood cancer, although not significant, suggested increased risk for exposure to water treated with chlorine or chlorine dioxide. This study provided an adequate sample size for many birth outcomes and infant survival parameters. Although typical DBP concentrations were not provided, levels were considered to be similar to those in the study by Magnus et al. (1999), which presented significant associations with congenital malformations.

Another qualitative study of adverse birth outcomes and water treatment methods examined LBW and preterm birth rates for 28 Taiwan municipalities, 14 of which

provided chlorinated water to more than 90 percent of residents and 14 matched municipalities that provided chlorine-free water (Yang et al., 2000). Matching municipalities by degree of urbanization theoretically controlled confounding by socioeconomic status and additional adjustment by maternal age, marital status, maternal education and infant gender was performed in logistic regression. Of the18,025 births occurring from January 1994 to December 1996 included in this study, 238 cases of term LBW and 448 preterm births (< 37 weeks) were observed. A statistically significant association was found for chlorinating municipalities and preterm birth (adjusted OR = 1.34, 95% CI 1.15- 1.56), but not for term LBW. Typical DBP Concentrations for chlorinating municipalities in Taiwan were not provided.

The most recently published qualitative study was designed to specifically investigate associations between disinfection method and congenital cardiac defects in a single region of Sweden (Cedergren et al., 2002). The study population included all women giving birth between January 1982 and December 1996 with maternal residential address in a specific county in Sweden that had been characterized as having a high prevalence of congenital cardiac malformations. Vital statistics, birth defect and hospital discharge records from births between 1982 and 1996 were used to evaluate all congenital cardiovascular defects, except patent ductus arteriosus and single umbilical arteries. Exposure was assigned according to geographic location using Geographic Information Systems (GIS) during preconceptional period and method of disinfection (none, chlorine gas, hypochlorite, hypochlorite and chlorine dioxide). After adjustment for maternal age, parity, and education level, odds ratios for development of any cardiac defect were

moderately increased for hypochlorite and chlorine dioxide (adjusted OR = 1.61, 95% CI 1.00 – 2.59), compared to the group with no exposure. No increased risk was observed for hypochlorite alone. Interactions between chlorine dioxide and nitrates showed no increased risk for cardiac malformation induction. Additional quantitative analyses yielded a positive association (adjusted OR = 1.30, 95% CI 1.08 – 1.56) with average THM exposures greater than 10 μg/l. No other quantitative analysis for DBPs was performed, as these levels are not routinely measured in Sweden.

2.3.2 Quantitative Studies of DBPs

Quantitative studies of DBPs provide specific concentration data for use in dose-response and threshold analyses. These observational methods are more similar to toxicological assessment than qualitative studies and may, therefore, be more useful in determining causality. Presented below are twelve quantitative studies addressing human exposure to DBPs.

The first study to evaluate individual THMs was conducted in 151 small Iowa towns (Kramer et. al., 1992). This population-based case-control study of growth-related birth outcomes collected 588 cases of LBW, preterm birth or IUGR (5^{th} percentile) and 3,440 controls, selected randomly from 1989-1990 Iowa birth records. Exposure was designated for each birth according to maternal residence at time of birth according to trihalomethane concentrations from a 1987 municipal water survey. These contaminant levels were categorized as high, medium or low, or as detectable and undetectable, depending upon the chemical. Adjustment for potential confounders included: maternal age, parity, marital status, education, maternal smoking during pregnancy and adequacy

of prenatal care. Small associations were observed between chloroform concentrations ≥10 μg/l and LBW (adjusted OR = 1.3, 95% CI 0.8-2.2) and IUGR (adjusted OR = 1.8, 95% CI 1.1-2.9). BDCM concentrations ≥10 μg/l were also associated with IUGR (adjusted OR = 1.7, 95% CI 0.9-2.9), although were not statistically significant. The limitations of this study include use of a one-time survey to designate levels of exposure, when DBP concentrations are known to fluctuate over time. In addition, this 1987 survey was performed during a drought and may not be representative of THM levels observed during the relevant time of fetal development for 1989-1990 births. However, use of the 1989 birth records allowed adjustment by maternal smoking, a confounder in many growth-related assessments. Also, the cutpoints used to categorize exposure to chloroform and other byproducts in the Iowa study are skewed toward low exposure, compared to U.S. studies performed more recently. Chloroform concentrations for the 151 small towns were higher than 10 μg/l on average (mean = 12.5), although this was only attributable to 19 communities. With this wide range, it was possible to perform a few more analyses for use in approximating human dose-response. Limiting the study to these smaller towns did reduce the possibility of residual bias due to urbanization or socio-economic factors.

Bove et. al., (1995) conducted a study of growth and development related outcomes and DBPs in northern New Jersey residents between 1985 and 1988. This large study included 80,938 live births and 594 fetal deaths, with data obtained from birth certificate records and the New Jersey Birth Defects Registry. Quarterly samples were obtained for 49 utilities and a forward-backward averaging technique was used to estimate monthly

TTHM levels for 75 towns. Exposure data corresponding to critical time periods of development was generated for each mother through monthly estimates and maternal residential addresses. Estimated TTHM concentrations were averaged over the first trimester for analyses of birth defects and over the entire gestation for growth outcomes. Data were adjusted for certain potential confounders, including: maternal age, ethnicity, infant gender, adequacy of prenatal care, previous stillbirth and education. For growth-related analyses, positive associations for exposure to TTHM levels greater than 100 μg/l were found for term LBW (adjusted OR = 1.4, 50% CI 1.2-1.7) and small for gestational age (SGA) at the fifth percentile weight by gestational week (adjusted OR= 1.5, 50% CI 1.4-1.7). In addition, birth weight was found to decrease an average of 70.4 g among term births exposed to TTHM concentrations greater than 100 μg/l, as compared to the (≤ 20 μg/l) referent group. No associations were observed between preterm birth, VLBW, or fetal deaths and exposure to TTHM.

Statistically significant associations were found for several categories of birth defects and exposure to TTHM levels greater than 80 μg/l, including: all surveillance birth defects (adjusted OR=1.57, 90% CI 1.23-1.99), central nervous system defects (adjusted OR=2.59, 90% CI 1.53-4.30) and NTDs (adjusted OR=2.96, 90% CI 1.26-6.62). In addition, a significant association was observed for oral cleft defects (adjusted OR=3.17, 90% CI 1.18-7.26) at TTHM exposures greater than 100 μg/l. Although based upon five cases and not statistically significant, an association was also observed between exposure to TTHM concentrations greater than 80 μg/l and major cardiac defects (adjusted

OR=1.83, 90% CI 0.97-3.29). Overall, no monotonic trends were found for these growth or development outcomes and levels of exposure to TTHM.

This cross-sectional study had a very large sample size with which to assess DBP associations, although only exposure to TTHMs was evaluated. Another major strength of this study was use of a birth defect registry to identify cases. Birth certificates underreport congenital anomalies, mainly due to the difficulty of recognizing a defect in the first few days after birth (O'Rahilly and Muller, 2001). The registry enabled Bove et al. (1995) to confidently assess multiple outcomes, many of which were found positively associated with TTHM exposure. Positive associations may be due, in part, to the limitations involved with this method of assigning exposure. For example, averaging over an entire 9-month period for each subject would make virtually all of the exposures in a single community indistinguishable. This may be acceptable if the study intent is to compare several communities with DBP concentrations that are relatively different. However, use of multiple communities could introduce multiple potential confounders that may be difficult to identify or control. In addition, this method could introduce substantial exposure misclassification with regard to the relevant time period of pregnancy.

Savitz et al. (1995) conducted a case-control study of births at six North Carolina hospitals from 1988-1991. This study evaluated 126 SABs, 244 preterm births, 178 LBWs and 333 controls in relation to TTHM levels, water source and amount ingested. Controls were restricted to term, normal weight births, matched by race and hospital.

Exposure to TTHMs was assigned to women by maternal residential address and corresponding TTHM, as determined through quarterly municipal water surveys, and was contingent upon dates of pregnancy and outcome. For SAB cases and controls, the nearest quarterly average was assigned to the fourth week of gestation. For preterm and LBW outcomes, nearby quarterly averages were assigned to the 28th week of pregnancy. Risk estimates were determined for exposure to surface water, TTHM levels, water consumption during pregnancy and TTHM dose (concentration multiplied by amount). All odds ratios were adjusted by maternal age, ethnicity, hospital, education, marital status, poverty level, smoking, alcohol consumption, employment and presence of nausea. No increased risks were observed for water source, amount, TTHM concentration or TTHM dose and any of the outcomes. SAB, however, was significantly associated with TTHM, when analyzed as a continuous variable, in 50 μg/l increments (adjusted OR =1.7, 95% CI 1.1-2.7).

Missing data in this population-based North Carolina study could have contributed to bias. In addition to the high rate of non-respondents (approximately 30 percent), SAB cases may have been missed, as women undergoing SAB are not necessarily treated at a hospital or private clinic. Also, recognition of pregnancy is a potential bias in studies of SAB (Weinberg and Wilcox, 1998). There is no mention in the Savitz et al. paper of any questions in the survey or variables in the statistical analysis regarding pregnancy recognition date. Also, while these researchers correctly tried to assign exposure to critical time periods relative to fetal development, estimated exposure was still based upon the closest quarterly sample. Temporal variability of DBPs and imprecise methods

of assigning exposure are likely to increase exposure misclassification. Exposure assessment of these study subjects was based upon personal surveys, making it superior to previous studies. However, quantifying amount ingested may be inaccurate, as it involved a one-time survey used to assess average exposure over the entire pregnancy. In addition, the survey lacked questions about water consumed away from home and drinks prepared with cold water. Lastly, outcome ascertainment was based upon hospital records, rather than birth certificates.

Waller et al. (1998) addressed SAB and individual trihalomethanes in a prospective study of women from three regions of California, distinguishable by source of drinking water (surface, ground or mixed). From 1989 to 1991, women who were pregnant for $\leq$13 weeks with a prepaid health plan were recruited from these regions and were matched to one of 78 residential drinking water utilities. THM levels were estimated by averaging all distribution system measurements taken within each subject's 1st trimester. This method approximated exposure for 77 percent of the cohort. The remainder of the cohort was assigned the TTHM average within 30 days of their first trimester (4%) or the annual average (9%). In addition, estimates of water consumption type (cold, hot and total) and quantity during the eighth week of gestation were obtained for 5,144 women by telephone interview. All multivariate statistical models were adjusted for gestational age at interview, maternal age, smoking, history of pregnancy loss, maternal race and employment during pregnancy. A slight association was observed for TTHM exposure to concentrations $\geq$75 µg/l and SAB (p-value = 0.16). Stratification of the exposed group by level of cold tap water consumption, however, produced a moderate association that

was statistically significant for those consuming 5 or more glasses of water a day (adjusted OR=2.0, 95% CI 1.1-3.6). When cold tap water consumption and TTHM exposure were combined into a personal TTHM exposure variable, the OR was reduced slightly (adjusted OR=1.8, 95% CI 1.1-3.0). When the same analysis was performed with total tap water consumption (cold + hot), the risk estimates were lower and were no longer significant (adjusted OR=1.2, 95% CI 0.8-1.9). SAB rate was not influenced by letting water stand before drinking (allowing THMs to volatilize) or by filter use. Other routes of exposure, including showering and swimming, were not associated with increased risk of SAB, when included into a model with personal TTHM exposure.

BDCM was the only chemical associated with SAB, in analyses of personal exposure to individual THM constituents (adjusted OR=2.0, 95% CI 1.2-3.5). The high personal exposure variable for BDCM was defined as consumption of $\geq$5 glasses of cold tapwater per day with concentrations $\geq$ 18 μg/l BDCM. When stratified by region these associations were slightly stronger, except for Region I, where the association was no longer significant.

A subset analysis was also conducted for women not employed outside the home, where exposure assessment was likely to be more accurate. Risk estimates were stronger for high personal exposure in women who were unemployed (adjusted OR=3.0, 95% CI 1.2-7.9) compared to employed women with high exposure (adjusted OR=1.5, 95% CI 0.8-2.8), although smaller sample sizes may have decreased precision. Stratifying women according to employment status assumed that at least some water was consumed outside

the home. Although not mentioned in the paper, this method could also have been used to assess the "reproductively unhealthy worker bias", where women who are able to successfully reproduce are likely to stay home and those who are less successful will not (Weinberg, 1993). In this paper, there is no discussion of this bias, however, employment status was controlled for in the analyses.

The prospective design and personal interview method used in this California study was an improvement over exposure quantification methods used in previous investigations. For example, recent information was obtained on frequency and quantity of use for both hot and cold water, and for activities involving water. However, as in the Savitz et al. study, this information only pertained to the eighth week of pregnancy, which may not be indicative of exposures throughout pregnancy or during the most critical gestation time. The study had several limitations with respect to assigning exposure and analysis. For example, DBP samples were only available quarterly. In addition, all quarterly samples in a single distribution system were averaged to get a single exposure metric, disregarding the spatial variability inherent to most water distribution systems. Within the analysis, personal exposure calculations inappropriately allowed a person with TTHM concentrations significantly greater than 75 μg/l who drank less than 5 glasses of water to be in the low exposure category. Also, the variable named "previous history of SAB" may have been adjusted for incorrectly. This is because DBPs, if truly causal for SAB, could have influenced both current and past SABs (Weinberg, 1993). Selection bias was also possible due recruitment of women through an HMO, restricting the populations to potentially lower SES women. However, women were also selected by phoning during

early pregnancy to schedule a prenatal appointment, indicating a sample of higher SES women. No information on the magnitude or direction of selection bias was available for this study.

Waller et al. published a follow-up paper to the California study in 2001, which included an improved exposure assessment methodology. The data from the previous study was reevaluated systematically and geographic information systems were used to link subjects with appropriate water utilities. Exposure for each subject was designated by 1) a distribution system average for all samples taken during the first trimester, and 2) the nearest sampling point value(s) for that time period. To address spatial variability within the distribution system, individual exposures were weighted according to the statistical variance of all distribution system measurements for each quarter. In addition, closest site exposure values were weighted according to distance from site location. Subset analyses were also performed, as in the original analysis, for women with more accurate exposures. In this case, women from the same utilities that were within 20 μg/l TTHM of each other and women receiving only groundwater were included. Quantification of personal exposure to TTHM was amended to include groups with high DBP concentration exposure, but low water consumption. This analysis involved the product of DBP concentration and quantity of water intake and the stratification method shown in the previous publication.

Of the 5,144 women in the original analysis, a large number of women (n=932) with missing data were excluded. Regardless, the overall estimation of risk for the utility-

wide measure was similar to, if not stronger, than those presented in Waller et al., 1998. For the highest exposure category in the stratified analyses (≥75 μg/l TTHM and ≥5 glasses of water/day), the adjusted OR for the unweighted data was 1.7 (95% CI 0.9-3.4) compared to the referent group of <75 μg/l TTHM and ≥5 glasses of water/day. The OR increased to 2.4 (95% CI 1.1-5.3) for the weighted analysis and 5.1 (95% CI 1.8-14.7) for the subset analysis. Analysis of the product term yielded slightly decreased risk estimates, with an OR=1.3 (95% CI 0.9-1.9) for unweighted data and OR=2.0 (95% CI 1.1-3.4) for the subset data. In general, analyses for exposure by closest-site showed no association between DBP dose and SAB, regardless of analytical method. Weighted and subset analyses generally suffered from smaller sample sizes and were more influenced by a single facility.

The results of the 2001 study were only explored for exposure to the general TTHM class and not for the specific THMs, as in the previous study. The exposure assessment techniques are generally improved in regard to individual quantification of exposure and by adjustment for spatial variability. However, no additional information was made available to determine if, on average, samples taken more frequently than once per quarter existed. If this is the case and all samples in that quarter are included in the analyses, temporal variability may still influence the exposure assessment.

Differential misclassification of exposure was detected in the reanalysis (Waller et al., 2001). One particular zone in region I, had been previously designated as receiving surface water, when, in fact, it received groundwater. SAB rates in that area also happen

to be higher (14.6%) than in the remaining sample population (approximately 9%). Although this bias was expected to deviate risk estimates away from the null, adjustment or exclusion of the subjects in this zone did little to influence the risk estimates. Nevertheless, the results of these analyses are useful in validating current DBP methodologies and suggest pathways to establish more accurate exposure assessment techniques.

Gallagher et al. (1998) contributed significantly to current exposure assessment methodologies through a retrospective cohort study of third trimester exposure to TTHMs and growth related outcomes. This Colorado study examined exposure specific to the census blocks containing maternal residential addresses of infants born between 1990 and 1993. Birth records (n=1,244) containing 72 LBW, 29 term LBW and 68 Preterm births were obtained and linked to 28 census blocks in two water districts (distribution systems) and assigned TTHM exposure, estimated by hydraulic characterization of water distribution systems for the specific trimester of interest. Adjustment for potential confounders in the logistic model included maternal age, smoking, marital status, parity, education, employment and quality of prenatal care. At TTHM $\geq$61 μg/l, the highest exposure category, odds ratios for LBW (adjusted OR=2.1, 95% CI 1.0-4.8) and term LBW (adjusted OR=2.6, 95% CI 1.1-6.1) were elevated and statistically significant, compared to the referent category of $\leq$20 μg/l. However, small case numbers may have contributed to imprecise estimates. No relationship was found between exposure to TTHMs and preterm birth.

GIS techniques to better understand spatial variability of DBPs within distribution systems, Gallagher et al. ushered in a new generation of exposure assessment methods for studies of DBPs. This study made the requisite connection between proximity and tap water sampling prior by use of hydraulic characterization, although this limited the number of census tracts and power available for the study. However, the study had several methodological limitations. For example, small numbers and imprecise estimates may have stemmed from the 6,214 births excluded from 58 of 86 census block groups. Representation of the sample is further questioned because the prevalence of LBW (5.8%)and preterm (5.5%) is low compared to national standards (7.6% and 11.6%, respectively) (National Vital Statistics Report, 2002). Lastly, the exposure estimate for this sample was again based upon quarterly sampling; no attempt was made to account for temporal variability or to interpolate exposure to a specific date.

Dodds et al. (1999) performed a retrospective cohort study to assess TTHM exposure and birth outcomes for 49,842 births occurring between 1988 and 1995 in Nova Scotia, Canada. To account for temporal variability, regression of TTHM values was performed, for all available samples specific to a facility, to predict DBP concentrations for times with no data. Perinatal and Fetal Anomaly databases were used to obtain information on potential confounders, maternal residence, and birth outcomes, including: SGA (10^{th} percentile), LBW, VLBW, preterm birth, stillbirth, NTDs, cleft lip and palate, major cardiac defects, and chromosomal anomalies. Exposures were evaluated over specific gestation periods according to birth outcome. Predicted TTHMs were averaged over: (1) the last 3 months of pregnancy for fetal growth outcomes; (2) the first 2 months of

pregnancy for cleft and cardiac defects; (3) the month before conception to the month after conception (2 months) for NTDs; (4) 3 months before pregnancy for chromosomal anomalies; and (5) the entire pregnancy for stillbirth. TTHM values were categorized into three levels (50-74, 75-99 and ≥100 μg/l) with a referent category of ≤49 μg/l.

Prevalence ratio and relative risk estimates were estimated using poisson regression. Maternal age, parity, maternal smoking, attendance at prenatal classes, neighborhood family income, gender of baby, and weight (before pregnancy and parturition) were variables considered as potential confounders. Risk estimates for TTHM exposure and SGA (10^{th} percentile), LBW, VLBW, preterm birth and congenital anomalies were not significantly different from unity. However, the highest level of exposure to TTHMs was slightly and significantly associated with stillbirth (adjusted RR=1.66, 95% CI 1.1-2.5). Also, the estimated risks for chromosomal abnormalities were all elevated for TTHM exposure over the referent level, but none of the estimates were statistically significant nor was there evidence of a dose-response trend.

This large Nova Scotia study was the first quantitative study to associate stillbirth and exposure to high concentrations of TTHMs and it is the first paper to account for temporal fluctuations in DBPs by regressing data through time. A clear strength of this study is thorough case ascertainment of stillbirths and fetal anomalies, including anomalies of terminated pregnancies, through perinatal databases. However, comparability to previous studies is low due to fairly high concentrations of TTHM in Nova Scotia. In addition, the unexposed referent group in this study (exposure ≤49 μg/l)

includes estimates for groups considered exposed in other studies. Intra-system spatial variability may also lead to significant exposure misclassification, as could the time periods used to assign exposure. For example, analyses of preterm births, where exposure was assigned to the last three months of pregnancy, could mean comparing exposure during the second trimester for preterm births with exposure during the third trimester for term births. Also, no data pertaining to individual exposures were collected, such as quantity and type of water consumed, bathing habits or other water activities.

A population-based case control study in New Jersey was conducted to specifically evaluate NTDs and DBP exposure among births between 1993 and 1994 (Klotz and Pyrch, 1999). This study was a follow-up to the previous New Jersey study of DBPs (Bove et. al., 1995). NTD cases (n=112) were selected from the New Jersey Birth Defects Registry and Fetal Death Registry and confirmed by a CDC dysmorphologist. Controls were healthy births, weighing $\geq$2500 g, selected randomly from among New Jersey births. Exposure was estimated by sampling residential taps one year after the critical time period of neural tube development. Sample locations were based on the address of each pregnant mother during this critical time period, her first month of pregnancy. In addition to TTHMs, haloacetic acids and haloacetonitriles were measured in most samples. Records from each water utility were also used to estimate TTHM exposure for the critical time period and for that time one year later. Potential confounding data were obtained through birth certificates and/or personal interview, including: sociodemographic factors, pregnancy and medical history, family medical

history, parental occupation, tap water ingestion, bathing and showering, time at swimming pools, use of water filters, and use of vitamins.

Risk estimates for the association between NTDs and the highest exposure category of TTHM (≥80 μg/l) were slightly elevated (adjusted OR=1.6, 95% CI 0.7-3.6). In addition, risk increased when analyses were restricted to subjects with known residency at conception and isolated defect cases (adjusted OR=2.1, 95% CI 0.8-5.3), although smaller case numbers decreased precision. Risk of exposure to TTHM concentrations <80 μg/l were also increased after the aforementioned restriction, except for the exposure category of 20-<40 μg/l where no risk was observed. Thus, no trend with increasing dose was observed. For analyses with tap water sampling, no appreciable risk or trend was observed with respect to TTHM exposure. Surface water was associated with NTDs (adjusted OR=1.8, 95% CI 1.0-3.4), although only when analyses were restricted to isolated defects. Analyses with individual THMs yielded results similar to those including TTHMs. Also, no increased risks were observed for tertiles of haloacetic acids and haloacetonitriles. The interaction of "no vitamin use" and the highest TTHM tertile of exposure (≥40 μg/l) did, however, produce moderate prevalence odds ratios (adjusted OR=2.6, 95% CI 1.2-6.0) when analyzed for association with NTDs.

The strengths of this study include many improvements to exposure assessment and statistical analysis. For example, the authors use individual surveys and attempts to assess DBP classes other than THMs. However, this study is most novel in the attempt to collect tap water sampling at a surrogate critical time period. Unfortunately, there is no

way to confirm that these methods are any more accurate than use of distribution system estimates, especially since DBP concentrations are sensitive to changes in flowrate and meteorological events. Use of sensitivity analyses, however, is warranted and is a useful method for understanding and reducing bias. In this case, restricting analyses to subjects with known residency at conception and isolated defects produced stronger risk estimates, indicating that bias was toward the null and misclassification was likely non-differential. Also, basing case selection on a registry, rather than birth records, was another method likely to reduce bias due to disease misclassification. Fewer cases, however, were available for analysis at higher levels of exposure, limiting the power of those results. In addition, case ascertainment was potentially biased, as women with a lower quality of prenatal care or those with religious or ethnic influence were more likely to maintain a pregnancy and, thus, be selected for study. Electively terminated pregnancies with prenatal diagnoses were not included among cases. Differential participation rates may have led to another source of selection bias, as younger, less educated, women were less likely to participate.

King et al. (2000) performed a follow-up study to the Nova Scotia study by Dodds et al. (1999) using the same population. This study was designed to further investigate the association observed between stillbirth and total and specific THMs. To improve analysis of this outcome, information pertaining to the underlying cause of stillbirth was added. Stillbirth was then categorized into intrauterine asphyxia, congenital anomalies, death related to immaturity, other specified death and unexplained death. As in the previous study, exposure was estimated by linking maternal residence to the geographic

area served by each water facility and averaging predicted monthly values over the length of pregnancy. Additionally, this study was restricted to people receiving only treated surface water. The potential confounders obtained from the perinatal database were the same as those used in the previous study, except for attendance at prenatal classes and weight (before pregnancy and parturition), which were left out of analyses.

All results in this retrospective cohort study were adjusted for smoking and maternal age. Relative risks for stillbirth and exposure to ≥100 μg/l TTHMs (adjusted OR=1.69, 95% CI 1.11-2.58) and exposure to ≥100 μg/l chloroform (adjusted OR=1.59, 95% CI 1.06-2.37) were elevated and significant, when compared to the referent (<50 μg/l). Increased risks were also seen for exposures to ≥20 μg/l BDCM (adjusted OR=1.99, 95% CI 1.24-3.21) when compared to exposures of <5 μg/l BDCM. In addition, there was evidence of a dose-response trend for increasing concentrations of TTHMs, chloroform and BDCM. Analysis of THMs on a continuous scale yielded a 5% increase in risk for every 10 μg/l increase in TTHM and a stronger increase in risk (29%) for every 10 μg/l increase in BDCM. Analyses for specific causes of stillbirth were restricted to subjects with either an unknown cause of stillbirth or asphyxia-related deaths, due to small numbers in the other categories. Although both outcomes had small sample sizes in the highest categories of exposure, only risk estimates for asphyxia-related deaths were elevated, with an adjusted RR= 4.57 for exposure ≥100 μg/l TTHM (95% CI 1.93-10.77). Based on the dose-response trends and stronger risk estimates, this analysis is likely to stimulate further research into analysis of this outcome and exposure to individual byproducts. The

strengths and limitations of this study are similar to those mentioned for Dodds et al. (1999).

Dodds and King (2001) conducted a retrospective cohort study of births in Nova Scotia (NS) to evaluate the relationship between exposure to chloroform or BDCM and NTDs, major cardiac defects, cleft defects and chromosomal anomalies. Monthly estimates of exposure were predicted from water utility data using the same regression techniques described for the two previous studies and averaged over the critical exposure windows for each outcome (described in the Dodds section). At chloroform ≥100 μg/l, the highest exposure category, relative risk estimates for NTDs (adjusted OR=1.2, 95% CI 0.7-2.3) and oral cleft defects (adjusted OR=1.5, 95% CI 0.8-2.8) were elevated but not statistically significant, compared to the referent category of ≤50 μg/l. Risk estimates for exposure to BDCM ≥20 μg/l and NTDs were stronger, with an RR=2.5 (95% CI 1.2-5.1), compared to the referent category of ≤5 μg/l BDCM. There was no evidence of an association between exposure to either chloroform or BDCM and cleft defects at these levels. Overall, risk estimates for congenital anomalies from this study differ from those for individual THMs, presented in Dodds et al.(1999) for Total THMs. Risk of NTDs increased from RR=1.18 (95% CI 0.67-2.0) for exposure to TTHMs to RR=2.5 (95% CI 1.2-5.1) in this analysis of exposure to BDCM. Moreover, the risk of NTDs is similar for TTHMs and chloroform, appropriately, since chloroform is the predominant component of TTHMs in Nova Scotia water. This paper clearly demonstrates how analysis of a group of chemicals, e.g. TTHMs, may disguise the effects specific to a component chemical (e.g. BDCM).

Wright et al. (2003) performed a cross-sectional analysis of THMs during specific trimesters of pregnancy and fetal development for births occurring in 99 Massachusetts towns during 1990. Quarterly TTHM values were assigned to the 56,513 births according to city-specific aggregate data, which included annual average data from 10 towns with groundwater as their primary source water. For the 64 towns reporting multiple sampling sites within a distribution system, a quarterly average was assigned. Birth records were used to obtain information on potential confounders, maternal residence, and birth outcomes, including: birth weight, LBW, IUGR in term births (10^{th} percentile), gestational age and preterm birth. TTHM values were categorized into three levels (0-60 [referent], >60-80 and >80 μg/l).

Odds ratios were estimated using linear and logistic regression techniques. Gestational age, infant gender, adequacy of prenatal care maternal race, maternal education, maternal smoking, maternal age, parity, median household income, and several other pregnancy complications and maternal health concerns were considered to be potential confounders. Risk estimates for IUGR and pregnancy average (adjusted OR=1.14, 95% CI 1.02-1.26) and second trimester (adjusted OR=1.13, 95% CI 1.03-1.24) TTHM exposures were increased. In addition, incremental reductions in birth weight were associated with pregnancy average and second trimester exposures to TTHMs at levels greater than 60 μg/l. No association was observed for preterm delivery, although some increase in risk was observed for an *increase* in gestational length.

This large Massachusetts study was the first quantitative study to associate growth-related outcomes to TTHM exposure during the second trimester. However, the exposure assessment methods, as with the methods of Bove et al. (1992), do not reflect change in exposure due to temporal variability and may be subject to exposure misclassification. Like the Dodds et al. (1998) study, the unexposed referent group (exposure $\leq$60 μg/l) includes estimates for groups considered exposed in other studies. In addition, averaging data from multiple sampling sites within a distribution system without knowledge of spatial variability could contribute to bias. Lastly, further bias may be present owing to reliance on birth record data. For example, maternal residence listed on birth certificates and reported at the time of birth was used to assign exposure, although it was not necessarily the most relevant time of exposure,. Nonetheless, in the absence of interview data, birth records and hospital worksheets did provide a large quantity of information that had never before been investigated with respect to DBPs.

In 1991, Shaw et al. published a letter with results from a very early investigation of chlorinated water exposures and congenital cardiac anomalies. This qualitative study found little evidence of any association with cardiac outcomes, even when grouped by pathogenetically related conotruncal defects (Shaw et al., 1991). A more recent study by Shaw et al. (2003) evaluated the effects of THM exposure around conception and induction of neural tube, orofacial and conotruncal defects. The study population, livebirths and fetal deaths born between June 1989 and May 1991 in California, was formed from two case-control cohorts used in previous investigations of smoking. Cases were ascertained from medical records and controls were randomly selected from

proportionally-sized area hospitals. Maternal residence information for the three months before to the three months following conception was collected by interview. TTHM and individual THM exposure was estimated by linking data provided by municipal water sources to corresponding census tracts. In addition, water company personnel were asked to use quarterly monitoring data and knowledge of distribution characteristics to estimate THM levels for specific street addresses during given time windows. For one of the cohorts, TTHM dose (concentration multiplied by amount) was also estimated. Potential confounding variables for both cohorts included: maternal race and ethnicity, education, body mass index and use of multivitamins containing folic acid. Characterization of the folate-related gene, methylenetetrahydrofolate reductase (MTHFR), was also performed using infant blood collected from some cases and controls.

Risk estimates for the association between NTDs and TTHM exposures were below 1.0 for cohort 1 and showed no association in cohort 2. Non-significant risk increases were observed for oral cleft and conotruncal defects, but only for the lowest level of exposure. No trend with increasing dose was observed. Analyses for individual THMs, no appreciable risks or trends were observed for any outcome. Polymorphisms of the MTHFR genotype were not associated with any outcome; however, analyses of TTHM were restricted to only one cutpoint, TTHM $\geq$ 25 μg/l.

The strengths of this study included increased detail in exposure assessment methods, although these methods were not validated, and evaluation of a potential gene-environment interaction. Evaluating cases by specific defect, rather than birth defect

group, was likely to reduce bias due to disease misclassification, although power may have been lost for statistical analysis. However, subject enrollment was potentially prone to selection bias, as women unable to speak English or Spanish were excluded, a potentially large number of subjects in California. Also, women with previous NTD births were excluded. If previous NTDs were caused by earlier exposure to THMs, removing those women may have distorted the risk estimates for THMs.

2.3.3 Individual Quantification of Exposure

Exposure assessment in studies of (DBPs) in drinking water and adverse birth outcomes is often based upon assumptions (e.g. all pregnant women consume the same amount of water each day). However, heterogeneity in exposure exists between people and for a single person through time. Examples of the differences in relation to DBP exposure between people include ingestion of bottled or filtered water, and variability in tap water consumption at home and at work. For a single person, several variables may change during pregnancy, including: quantity of water consumed, rates of metabolism and even residency. In epidemiologic studies of DBPs, one of more of these differences may lead to exposure misclassification (Nieuwenhuijsen et al., 2000b; Arbuckle et al., 2002).

In order to understand the magnitude of these potential sources of misclassification, several investigations have been performed to assess frequency, route and type of water exposures (Shimokura et al., 1998; Swan et al., 1998; Zender et al., 2001; Barbone et al., 2002). Studies conducted in both the United States (Ershow et al., 1991; Burmaster et al., 1998) and Great Britain (Hopkins and Ellis, 1980) have evaluated tap water and total water consumption for pregnant and lactating women, compared to a control group. In a

Table 2.3.1. Summary of Epidemiologic Studies on Disinfection By-Products and Adverse Reproductive Outcomes.

Author Journal Date	Study Details	Number of Cases	Spatial Variability	Exposure assigned by:	Restrictions	Exposure Assessment	Other Risk Factors Included	Main Findings OR (95% CI)
Kramer et al Epidemiology 1992	Iowa, US 151 towns with a single water source. 1989-90 Sample population: 4028	588 (total) 159 LBW 342 PTB 187 IUGR	No mention (although, this was a 1 time survey for 151 included towns)	Maternal residence at time of birth (i.e. town-wide estimate from a 1 time water survey taken 1-2 years prior to births. Also, estimates were taken during a drought year)	Singleton white infants born to non-hispanic white women >19 years of age, residents of Iowa towns with 1000-5000 inhabitants, excluded infants with unrealistic gest ages ≥46 or ≤ 22	Based on maternal residential address and one municipal water survey to estimate individual THM levels (2 or 3 categories)	Maternal age, Parity, Marital Status, Education, Smoking, Prenatal care	No Vs. Med (1-9 µg/l) Vs. High (≥10 µg/l) **Chloroform:** LBW: 1 Vs. 1.1(0.7-1.6) vs. 1.3 (0.8-2.2) IUGR: 1 Vs. 1.3(0.9-1.8) Vs. 1.8 (1.1-2.9) **Dichlorobromomethane:** IUGR: 1 Vs. 1.2(0.8-1.7) Vs. 1.7 (0.9-2.9)
Aschengrau et al Archives of Environmental Health 1993	Mass, US 2 hospitals 1977-1980 Sample population: 2348	2316(total) 1039 congenital UTD, RTD 77 Stillbirths 55 Neonatal Deaths	Some sampling on same day in single plant, no idea how much, or how often	If available, address of res. During first trimester used to match to town water data.	Excluded women w/ no record, diabetes, epilepsy, prenatal herpes, toxo, rubella or a history of drug abuse – also, women outside MA w/ no public water supply. & if women were pregnant >1 time during study	Based on maternal residential address to ascertain type of water supply, chlorination vs. chloramination, and ground/mixed water vs. surface water	Maternal age, Pregnancy history, Alcohol, Ethnicity, Hospital payment, Other water contaminants	**Chlorinated Vs. Chloraminated:** Still birth 2.6 (0.9-7.5) Neonatal deaths 1.1 (NA) MCM 1.5 (0.7-2.1) RTD 3.2 (1.1-9.5) UTD 4.1 (1.2-14.1)
Bove et al American Journal of Epidemiology 1995	New Jersey, US 75 towns w/ public water supply 1985-1988 Sample populations: 81602	29268 (total) Live births: 1853 LBW 905 VLBW 4082 IUGR 7167 PTB 594 Fetal Death All births: defects 669 Surveillance 118 CNS defects 83 OCD 56 NTD 108 MCD	49 water companies, quarterly sampling, 4 sampling points	Forward-backward averaging technique for quarterly samples. Exposures aved over 1st trimester were used for BD, aved over entire preg for all other birth outcomes. Based gestation length on LMP.	Excluded plural births, therapeutic abortions and chromosomal anomalies. Excluded gestation <20 or >50 weeks.	Based on maternal residential address and municipal water surveys to estimate monthly TTHM levels (5 or 6 exposure categories)	Maternal age, Ethnicity, Baby's gender, Primipara, Prenatal care, Education, Previous still or miscarriage, Other contaminants	**TTHM levels > 100 µg/l Vs. ≤ 20 µg/l** LBW: 1.4 (50% CI 1.2-1.7) IUGR: 1.5 (90% CI 1.4-1.7) Both at 50th Confidence interval TTHM levels > 80 µg/l Vs. ≤ 20 µg/l Surveillance Register defects: 1.6 (90% CI 1.2-2.0) CNS system defects: 2.6 (90% CI 1.5-4.3) NTD: 3.0 (90% CI 1.3-6.6) MCD: 1.8 (90% CI 1.0-3.3) OCD: 3.2 (90% CI 1.2-7.3)

Author Journal Date	Study Details	Number of Cases	Spatial Variability	Exposure assigned by:	Restrictions	Exposure Assessment	Other Risk Factors Included	Main Findings OR (95% CI)
Savitz et al Environmental Health Perspectives 1995	North Carolina, US 6 Hospitals 1988-91 Sample population: 1003	548 (total) 126 SAB 244 PTB 178 LBW	5 plants, quarterly estimates, ? sampling points	Dates of pregnancy were used to assign quarterly DBP values. For miscarriage, the 4th week was time period of assgnmt. For preterm, LBW 28th week of pregnancy.	Controls restricted to term, normal weight births. 10-15% of cases and controls lost due to subject refusal. Restricted study to women served by public water and those drinking ≥1 glass of water.	Based on maternal residential address and quarterly municipal water surveys to estimate avg TTHM a. surface vs. ground b. TTHM levels (3 categories) c. Consumption during preg. d. Water source * amount e. TTHM dose (level * amt)	Maternal age, Ethnicity, Hospital, Education, Marital status, Poverty level, Smoking, Alcohol consumption, Employment, Nausea	**40.8-59.9 vs. 81.1-168.8 µg/l TTHM** SAB: 1.2 (0.6-2.4) **40.8-63.3 vs 82.8-168.8 µg/l TTHM** LBW: 1.3 (0.8-2.1) Per 50 µg/l TTHM increment change SAB: 1.7 (1.1-2.7)
Kanitz et al Environmental Health Perspectives 1996	Liguria, Italy 2 hospitals 1988-1989 Sample population: 676	548 live births in 'exposed area' 50 PTB 141 C sections 133 Neonatal jaundice 20 LBW 288 SBL 370 SCC	Two plants in Genoa: 1 public (ClO2 or Sodium hypo), 1 private (Sodium hypo). One plant in Chiavari (no treatment).	Exposure was assigned according to women's address.		Based on maternal residential address to ascertain type of water source (chlorine dioxide &/or hypochlorite vs. no treatment)	Maternal age, Education, Smoking, Alcohol, Gender of child	**Na hypochlorite (8-16 µg/l TTHM) vs no treatment:** Neonatal jaundice: 1.1 (0.7-2.8) LBW: 6.0 (0.6-12.6) SBL: 2.3 (1.3-4.2) SCC: 3.5 (2.1-8.5)
Waller et al Epidemiology 1998	3 regions of California, US surface, ground and mixed DW 1989-1991 Sample population: 5144	499 Spontaneous Abortion	78 Utilities.	For 77% of cohort, THM levels estimated by ave. of all dist. syst. Measurements taken within 1st trimester. For 4%, ave TTHM within 30 days of 1st trimester. For 9%, annual average used.	Excluded ectopic, molar and electively terminated pregnancies.	Based on maternal residential address and quarterly municipal water surveys to estimate average TTHM and individual THM levels. Analysis based on: a. THM levels b. Consumption during first trimester	Maternal age, Gestational age, Smoking, History of pregnancy loss, Ethnicity, Employment	**High TTHM dose (≥5 glasses/day + ≥75 µg/l) vs. low dose (<5 glasses/day + <75 µg/l)** SAB: 1.8 (1.1-3.0) **High BDCM (≥ glasses/day + ≥18 µg/l vs. low dose (<5 glasses/day + <18 µg/l):** SAB: 3.0 (1.4-6.6)

Author Journal Date	Study Details	Number of Cases	Spatial Variability	Exposure assigned by:	Restrictions	Exposure Assessment	Other Risk Factors Included	Main Findings OR (95% CI)
Gallagher et al Epidemiology 1998	Colorado, US 28 census blocks in 2 water districts 1990-1993 Sample population: 1244 live births	72 LBW 29 TLBW 68 PTB	2 water districts	Hydraulic characterization. Assigned exposure in 3rd trimester by calc. median of all representative [ ]'s in dist system.	Excluded 6,214 births from 58 of 86 census blk groups. Also, births <400g and those w/gest ages outside of 28-42 weeks. Excluded multiple births.	Based on maternal residential address and municipal water surveys. Estimate of household TTHM level during last trimester based on hydraulic modeling (4 exposure categories)	Maternal age, Smoking, Marital status, Parity, Education, Employment, Pre-natal care	**High TTHM level (≥61 µg/l) vs. Lowest (≤20 µg/l)** LBW: 2.1 (1.0-4.8) TLBW: 5.9 (2.0-17.0)
Dodds et al Epidemiology 1999	Nova Scotia, Canada 1988-1995 Sample population: 49842 births	4673 SGA 2393 LBW 342 VLBW 2689 PTB 77 NTD 82 OCD 430 MCD 197 Stillbirth 96 chromosomal abnormalities	>1 plant 3 sampling points, quarterly sampling – Regression of THM values over time performed to predict missing values.	TTHM average exposures over last 3 months of pregnancy for fetal growth outcomes. First 2 months of preg for cleft/cardiac defects. Month before to month after conception for neural tube defects. 3 mos before pregnancy for chromosomal anomalies. Ave TTHM over pregnancy for stillbirth.	Excluded women on private wells, municipal groundwater source or when water source unknown.	Based on maternal residential address and TTHM levels for public water facilities (3 sampling locations) modeled using linear regression on the basis of observation by year, month and facility (4 exposure categories)	Maternal age, Parity, Maternal smoking, Attendance prenatal classes, Neighborhood family income, Gender, Pregnancy and pre-delivery weight	**0-49 µg/l vs. >100 µg/l TTHM** Stillbirth: 1.66 (1.09-2.52) Chromosomal ab.: 1.38 (0.73-2.59) SGA 1.08 (0.99-1.18) NTD: 1.18 (0.67-2.10)
Klotz and Pyrch Epidemiology 1999	New Jersey, US 1993-1994 All births, of which 112 cases, 248 controls selected	112 NTD	>1 plant tap water sampling	Interview and tap water sampling performed 1 year after critical time period of development. Estimated ug/day for TTHMs and HAAs and HANs	Excluded controls: term births that were LBW, births w/other defects.	Based on residential address and public water facility TTHM data, and tap water sampling for TTHMs, HANs, and HAAs (3-5 exposure categories)	Sociodemographics, Pregnancy and medical history, Parental occupation, Use of vitamins	**<5 µg/l vs >40 µg/l TTHM** NTD: 2.1 (1.1-4.0)

Author Journal Date	Study Details	Number of Cases	Spatial Variability	Exposure assigned by:	Restrictions	Exposure Assessment	Other Risk Factors Included	Main Findings OR (95% CI)
Magnus et al Epidemiology 1999	Norway Sample Population: 141,077	2608 all birth defect 62 NTD 250 MCD 91 RTD 122 UTD 143 OCD	Included 366 municipalites with 1 (28%), 2 (26%), 3 (16%) or 4 (30%) waterworks (plants).	Estimated weighted mean color from waterwork (weight by pop. Served) Also made chlorination proportion for municipalities w/ both types of wws.	Did not include low reliability diagnoses from registry (e.g. congenital dislocation of hip.	Chlorination (yes vs no) Color (high vs low) In chorinated water, avg TTHM=9.4 µg/l, avg HAAs=14.6 µg/l	Maternal age, Parity, Geographical placement, Population density, Industry profile	**No chlorination low color vs chlorination high color** All Birth Defects : 1.14 (0.99-1.31) UTD: 1.99 (1.10-3.57) NTD: 1.26 (0.61-2.62) MCD: 1.05 (0.76-1.46) RTD: 1.07 (0.52-2.19)
King et al Environmental Health Perspectives 2000	Nova Scotia, Canada 1988-1995 Sample population: 49756	214 Stillbirths	Quarterly sampling of most Nova Scotia Facilities, no mention of sampling points	Estimated exp by linking mother's residence at birth to geographic area served by each water facility. Averaged predicted monthly values over the mother's pregnancy.	Restricted to municipalities where >90% of households were served by PWS. Restricted to people served by surface water. Excluded women with unknown gestational age.	Performed a least-square regression on DBP values for year, month and facility to obtain predicted values. Based on monitoring data from each facility (4 sampling events/year) Mother's residence linked to geographic area served by each facility. Averaged predicted values of DBP level for months covering duration of mother's pregnancy	Maternal age, Parity, Smoking, Infant's sex, neighborhood family income.	Total THM (µg/l) – Stillbirth (RR) **50-74 vs < 50** 1.27 (0.88-1.85) **75-99 vs < 50** 1.28 (0.81-2.03) **≥100 vs < 50** 1.66 (1.09-2.54) Chloroform (µg/l) – Stillbirth (RR) **50-74 vs < 50** 1.20 (0.85-1.68) **75-99 vs < 50** 1.35 (0.87-2.08) **≥100 vs < 50** 1.56 (1.04-2.34) BDCM (µg/l) – Stillbirth (RR) **5-9 vs < 5** 1.07 (0.77-3.19) **10-19 vs < 5** 0.44 (0.90-2.27) **≥20 vs < 5** 1.98 (1.23-3.49)
Kallen & Robert Reproductive Toxicology 2000	Sweden 1985-1989 114484 singleton births 1293 twin deliveries 24 triplet + deliveries	5615 Congenital Malformations	Multiple municipalities with disinfection types including: (1) no disinf. (2) ClO2, (3) NaClO3.	Exposure by disinfection source was assigned according to municipality.	Births were included if disinfection method was the same before and after delivery.	Used chlorination method used for DW and contained no info on the actual amounts of DBPs in DW. Based on municipality and DW treatment and composition.	Year of birth, Maternal age, Parity, Maternal Education, Maternal Smoking, County	NA Hypochlorite vs. Cl Dioxide or None PTB: 1.22 (1.00-1.48) SBL: 1.97 (1.30-2.97) BMI>16 kg/m3: 1.27 (1.19-1.37) SCC: 1.46 (1.07-1.98)

Author Journal Date	Study Details	Number of Cases	Spatial Variability	Exposure assigned by:	Restrictions	Exposure Assessment	Other Risk Factors Included	Main Findings OR (95% CI)
Yang et al Environmental Health Perspectives 2000	Taiwan 1994-1996 18025 primiparous, singleton births	238 TLBW 448 PTB	14 chlorinating municipalities; 14 nonchlor municipalities	Exposure by disinfection source was assigned according to municipality.	Excluded nonrepresentative munic. Also, births w/ gest ages >50 or <20 weeks. No twins or mult pregnancies. Excluded invalid or missing gest age subjects and births w/missing variables.	Mother's residence linked with Chlorinating municipality (CHM) or Non-chlorinating municipality (NCHM)	Maternal age, Marital status, maternal education and Sex of infant.	CHMs vs NCHMs PTB: 1.34 (1.15-1.56)
Waller et al Journal of Exposure Analysis and Environmental Epidemiology 2001	3 Regions of California 1990-1991 Sample Population: 4212 Births in re-analysis	397 SAB	78 Utilities, 4 sampling points typical for facilities used.	Utility-Wide Average:, THM estimated by ave. of all dist. syst. (Adj. By system var). Analysis for subset of women with more accurate exposure.	Excluded pregnancies included prev. that had exposure values extrapolated beyond a quarterly window of sampling.	1. Utility-wide avg during 1st trimester 2. Closest site avg during 1st trimester 3. Utility-wide weighted avg during 1st trimester 4. Closest site weighted avg during 1st trimester	Maternal Age, Gest age at interview, Pregnancy History, Race, Employment during pregnancy, Education, Alcohol consumption, Smoking, Water source, Residence	**Utility-wide avg during 1st trimester**SAB: 1.1 (0.8-1.6) **Closest site during 1st trimester** SAB: 1.1 (0.8-1.6) **Utility-wide weighted avg 1st trimester**SAB: 1.1 (0.8-1.6) **Closest site weighted 1st trimester** SAB: 1.1 (0.8-1.6)
Dodds and King Occupational and Environmental Medicine 2001	Nova Scotia, Canada 1988-1995 Sample population: 49,842	77 NTD 430 MCD 82 OCD 96 Chromosomal Abnormalities	>1 plant routine sampling – Regression of THM values over time performed to predict missing values.	First 2 months of preg for cleft/cardiac defects. Month before to month after conception for neural tube defects. 3 mos before pregnancy for chromosomal anomalies.	Singleton births to residents of Nova Scotia. Among those living in an area of municipal water supply. Included therapeutic pregnancy terminations.	BDCM and Chloroform (4 categories each) concentrations from routine monitoring. No information on individual patterns of consumption, showering or bathing.	Maternal age, parity, smoking, neighborhood family income	BDCM **5-9 µg/l vs < 5** NTD: 1.4 (0.8-2.3) **10-19 µg/l vs < 5** OCD: 1.1 (0.6-2.1) **≥20 µg/l vs < 5** NTD: 2.5 (1.2-5.1) Chloroform **50-74 µg/l vs < 50** OCD: 1.2 (0.7-2.0) Chrom. ab: 1.3 (0.8-2.2) **75-99 µg/l vs < 50** Chrom. ab: 1.9 (1.1-3.3) **>100 µg/l vs < 50** NTD: 1.2 (0.7-2.3) OCD: 1.5 (0.8-2.8) Chrom. ab: 1.4 (0.8-2.8)

Author Journal Date	Study Details	Number of Cases	Spatial Variability	Exposure assigned by:	Restrictions	Exposure Assessment	Other Risk Factors Included	Main Findings OR (95% CI)
Jaakola et al. Occupational and Environmental Medicine 2001	Norway Sample Population: 137,145 live births in 1993-95.	6249 LBW 5372 SGA 7886 PTB	Included 233 chlorinated waterworks, 1084 waterworks with no chlorination.	Estimated weighted mean color from waterwork (weight by pop. served)	Did not include subjects with no corresponding water information (n=40,284), children with birth defects (n=2,608), children with no birth weight info (n=1,324).	Chlorination (yes vs no) Color (high vs low) In chorinated water, avg TTHM=9.4 µg/l, avg HAAs=14.6 µg/l	Maternal age, Parity, Geographical placement (centrality), Population density, Industry profile	**No chlorination low color vs chlorination high color** LBW : 0.97 (0.89-1.06) SGA: 1.00 (0.91-1.10) PTB: 0.91 (0.84-0.99)
Cedergren et al. Environmental Research 2002	Sweden 1982- 1996 71,978 births with geo-coded maternal residential addresses in a specific county in Sweden	753 Cardiac Defects	Multiple municipalities with disinfection types including: (1) no disinf. (2) Cl2 gas, (3) NaClO3, (4) NaClO3 and ClO2.	Exposure by disinfection source was assigned according to municipality.	n=3854, excluded due to inability to geocode.	Used chlorination method used for DW and contained no info on the actual amounts of DBPs in DW. Based on municipality and DW treatment and composition.	Maternal age, Parity, Maternal Education, Maternal Smoking.	NA Hypochlorite vs. Cl Dioxide or None Cardiac malf (any): NaClO3 and ClO2- 1.61 (1.00-2.59)
Wright et al. Occupational and Environmental Medicine 2003	Mass. 99 towns w/ public water supply 1990 Sample populations: 56,513	Live births: 1325 TLBW 5310 SGA-term 3173 PTB	89 towns w/ quarterly sampling; 10 towns w/ annual measurements	Month of birth was used to assign quarterly DBP values to trimester specific or pregnancy average exposures. Preterm infants were not assigned a 3rd trimester exposure.	Excluded plural births, therapeutic abortions and chromosomal anomalies. Excluded gestation <22 or >45 weeks and those weighing less than 200g. 3 towns excluded for missing data. – missing quarters were imputed.	Based on maternal residential address and average TTHM concentrations for the towns of residence. Trimester exposure was based upon quarterly averages. If child was born in the first month of a particular quarter, it was assigned the previous quarter conc. If child was born in the 2nd or 3rd month, it was assigned that quarterly average.	Gestational age, gender, maternal race Maternal age, Prenatal care, Education, smoking, parity, median household income, maternal repro and medical history	**TTHM levels > 80 µg/l Vs. ≤ 60 µg/l (pregnancy average)** BW: -32g (-47 to -18) LBW: 1.05 (95% CI 0.85-1.29) SGA: 1.14 (95% CI 1.02-1.26) PTB: 0.90 (95% CI 0.77-1.04) **TTHM levels > 80 µg/l Vs. ≤ 60 µg/l (2nd trimester)** BW: -23g (-36 to -10) LBW: 1.14 (95% CI 0.95-1.38) SGA: 1.13 (95% CI 1.03-1.24) PTB: 0.90 (95% CI 0.79-1.03)

Author Journal Date	Study Details	Number of Cases	Spatial Variability	Exposure assigned by:	Restrictions	Exposure Assessment	Other Risk Factors Included	Main Findings OR (95% CI)
Shaw et al. Epidemiology 2003	California, US 1989-1991.	Two studies: 1st: 538 NTD cases, 539 controls 2nd: 265 more NTDs, 207 conotruncal heart defects, 409 OCDs, and 481 controls.	147 municipal water companies	Interview and genotyping performed. In study 1, estimated ug/day for TTHMs. Average THM concentration for each residence lived in for the 3 months before to 3 months after conception.	Excluded non-english or Spanish speaking women and women with previous NTD-affected pregnancies. Excluded women whose address could not be linked to municipal water co.	Based on residential address and expertise of water company personnel, who were asked to use quarterly monitoring data and knowledge of distribution characteristics to estimate THM levels for specific street addresses and dose during given time windows. Mutations in infant MTHFR gene identified.	Maternal race and ethnicity, education, body mass index and use of multivitamins containing folic acid.	**0 µg/l vs ≥75 µg/l TTHM** 1st study: NTD: 0.62 (95%CI 0.26-1.5) 2nd study: NTD: 1.8 (95%CI 0.85-3.7) MTHFR*THM >25 µg/l 1st study: NTD: 1.4 (95%CI 0.35-5.5) 2nd study: NTD: 1.6 (95%CI 0..29-9.7)

group of pregnant women and their male partners, one study by Shimokura et al. (1998) described tap water use and bathing and showering patterns. In this study of well-educated, health-conscious women, 39 percent consumed bottled or filtered water. In contrast, Zender et al. (2001) found that only 25 percent of lower socioeconomic status women drank filtered or bottled water in a study of pregnant and non-pregnant women attending public health clinics. In addition, 50 percent of the women in this sample reported working outside the home where nearly one-third of their daily water ingestion occurred. In a study of pregnant women in California, 59 percent percent of participants consumed some bottled water daily during the eighth week of gestation and 19 percent reported use of a filter (Swan et.al., 1998).

Recent research suggests that quantification of exposure to THMs may depend more heavily on amount of vapor inhaled, rather than volume of water ingested or absorbed through dermal exposure, while ingestion is likely to be the important exposure route for non-volatile HAAs. In one study of 31 people exposed to generally low THM levels, Backer et al. (2000) found that after exposure to DBPs by showering, bathing or ingestion, the highest levels of THMs in blood observed were in those who had been exposed via inhalation through showering. Higher DBP concentrations in blood after showering, as compared to bathing, have also been observed in a similar study by Kerger et al. (2000). In two other studies, blood from 50 recently pregnant women was analyzed and exposure was correlated to residence in a community with high concentrations of chlorinated DBPs (n=25) or in a community with high concentrations of brominated

DBPs (n=25) (Miles et al., 2002; Lynberg et al., 2001). These results, with respect to inhalation exposure and showering, were similar to the study by Backer et al. (2000). According to Weisel and Jo (1996), the difference in DBP blood levels between these two routes is likely the result of higher speeds of metabolism following ingestion. Unfortunately, exposure to volatile compounds is not easily quantifiable; it is contingent upon water temperature, concentration of these compounds in water and length of time exposed to steam (Weisel et al., 1999; Kerger et al., 2000). This is likely one exposure pathway that, when modeled for exposure, must be based upon some knowledge of habit.

Residential mobility during pregnancy introduces another source of misclassification. For example, in the study by Zender et al. (2001), thirty-two percent of women had moved during their current pregnancy. Similarly, Shaw et al. (1992) found that approximately 25 percent of women moved between the time of conception and delivery, and that moving was more common among younger and less educated whites and Hispanics. The study by Khoury et al. (1988) found that moving was more common among white women and those aged 20 to 24 years, and that 20 percent of mothers of children with birth defects changed address during pregnancy. These data suggest the value of obtaining individual-level data, considering the potential for bias due to exposure misclassification.

2.3.3.1 Biological Markers of Exposure

Additional efforts to improve exposure assessment have been described following analyses of DBPs in breath, blood and urine samples after exposure to water through consumption, inhalation and bathing. Weisel et al. (1999) conducted the first study of DBPs as biomarkers by analyzing post-exposure breath and urinary levels of THMs and

HAAs, respectively. In these analyses, post-shower breath levels of THMs were strongly correlated with THM concentrations in water and THM exposure due to showering. Also, urinary excretion of TCAA, a non-volatile HAA, was correlated to quantity of HAAs ingested; with stronger correlations when the analysis was limited to a subset of individuals with mainly in-home exposures. Analyses for DCAA, however, were not correlated with amount of water ingested. Froese et al. (2002) determined concentrations of TCAA in urine for ten subjects over five weeks with a bottled water intervention for a subset (n= 3) of participants. In this study, TCAA was determined to be dependent upon the spatial and temporal variability of the chemical in tap water and the amount of water ingested. In addition, the results of a study by Calafat et al. (2003) suggest that TCAA levels may vary by gender, where women exposed to urban supplies of chlorinated water have higher urinary TCAA levels. These findings, although not validated, support the use of non-invasive urinary or breath biomarkers for improved exposure assessment of specific DBPs, particularly in pregnant women.

In two of the studies of DBPs in blood, referred to previously (Lynberg et al., 2001; Miles et al., 2002), participants resided in a community with high concentrations of chlorinated DBPs or high concentrations of brominated DBPs. Study subjects were found to differ with respect to baseline level of THMs in blood, corresponding to the predominant class of DBPs in residential drinking water. These studies show that, regardless of biomarker, levels of DBPs in treated water are generally predictive of exposure in people utilizing that water.

Two studies have attempted to determine whether infants with polymorphisms in antioxidant genes are at greater risk of developing certain health outcomes when exposed to DBPs. Shaw et al. (2003) categorized cases with NTDs and control infants by TTHM exposure status and methylenetetrahydrofolate reductase (MTHFR) genotype. This gene, which is specifically related to maternal folate intake, was not found to increase the TTHM-NTD risk estimate. Another study by Infante-Rivard et al. (2002) evaluated the effects of polymorphisms in the GSTT1 and CYP2E1 genes on the relationship between childhood acute lymphoblastic leukemia (ALL) and THMs. This case-only study to estimate the interaction odds ratios (IORs) found significant positive associations with the CYP2E1 genotype when high average exposure (>95th percentile) occurred during pregnancy.

2.3.4 Spatial Variability

With respect to formation of DBPs, it is widely understood that several conditions (e.g. organic matter quantity, chlorine dose, pH, temperature, contact time and bromide ion concentration) can cause variability of these contaminants in the waters produced by one treatment plant versus another (Reif et al., 1996; Nieuwenhuijsen et al., 2000a; Bove et al., 2002; Golfinopoulos and Arhonditsis, 2002). In addition, biologic and hydrodynamic conditions in piping can cause DBP formation to vary spatially within a single plant distribution system (Rossman et al., 2001).

Spatial variability can lead to misclassification, which occurs when persons living at different locations within a distribution system are exposed to significantly different levels of DBPs but are assigned the same exposure levels. This occurs frequently in

epidemiologic studies when exposure assessment is based upon DBP levels measured at the plant. Since systems that rely on free chlorine in the distribution system are characterized by continuous, non-uniform DBP occurrence, concentrations of by-products measured at the water treatment plant may not accurately reflect those that exist throughout the distribution system (Chen and Weisel, 1998; Keegan et al., 2001). In addition, plants are not required to sample at a large number of locations in the distribution system and, due to concern over regulatory compliance, oversampling may be performed in areas of historical DBP excess (Waller et al., 2001).

Several studies have been performed recently to specifically address DBP formation and spatial variability in water distribution systems. In a study by Golfinopoul and Arhonditsis (2002), a multiple regression model was developed to predict THM production within a treatment plant system. The multiple linear equation determined for TTHM and individual species production was dependent upon several raw water characteristics, including: pH, temperature, chlorine dose, chlorophyll and bromide. In general, moderate correlations ($r \cong 0.50$) between these parameters and production of THMs were observed. Rossman et al. (2001) compared DBP formation in a simulated pipe environment to formation for that same water in a glass bottle. This experiment is important to understanding how pipe characteristics may influence DBP production towards production of THMs or HAAs. For example, some pipe distributions are known to possess biofilm, which can consume the chlorine available for DBP formation and, thereby, reduce DBP production. In this experiment, water in the pipe had 15 percent more THM production than water in bottles after 24 hours of contact time. The increased

THM production was attributed to the existence of THM precursor material attached to the pipe biofilm prior to the experiment. However, no significant increase in HAA formation occurred in the pipe as compared to the bottled water. Therefore, knowledge of DBP formation can be inferred, in part, from knowledge of raw water organic precursors and biofilm growth in specific pipe sections of the distribution system matrix.

Methods used to account for spatial variability in previous studies of DBPs and adverse birth outcomes have improved for studies utilizing quantitative exposure assessments. Initially, exposure assessments were based on utility data without accounting for the variability within the distribution system (Bove et al., 1995; Savitz et al., 1995; Dodds et al., 1998; King et al., 2000; Dodds and King, 2001). Klotz and Pyrch (1998) conducted residential tap water sampling for DBPs which corresponded to the address for the subject under study (Klotz and Pyrch, 1999). Gallagher et al. (1998) also attempted to account for spatial variability by incorporating a hydraulic model of the distribution system into a geographic information system, which assigned exposures to individual census block groups. While this method shows promise for prediction of DBP concentrations throughout the system, validation of the models for each class of byproduct is necessary. Most recently, Waller et al. (2001) used a weighting procedure to reduce the influence on risk estimates of individuals with less accurate exposure values, e.g. those that draw water from networks with high spatial variability in measured THM levels. Weighted analyses were based upon the within-utility variance in TTHM measurements.

One simple method that may be used to reduce misclassification would be to begin epidemiologic investigations at sites previously characterized as having low spatial variability. At present, however, there are no known criteria that can be used to classify water distribution systems as having low spatial variability.

2.3.5 Temporal Variability

Previous epidemiologic studies of disinfection by-products (DBPs) and the growth-related birth outcomes of preterm birth or LBW have used a variety of methods to assess exposure, none of which have specifically considered the third trimester. These earlier investigations have largely been based upon concentrations of DBPs measured in water distribution systems during quarterly sampling events. For instance, Bove et al. (1995) used quarterly values to extrapolate monthly exposure estimates over each woman's pregnancy. Three studies used quarterly DBP values to approximate exposure to the third trimester. Savitz et al. (1995) used the quarterly value nearest the 28th week of pregnancy, while Gallagher et al. (1998) used the median of all quarterly measurements taken during the third trimester. Wright et al. (2003) used the quarterly average that was current for the third trimester if the child was born in the 2nd or 3rd month of that quarter, all else were assigned the preceding quarterly average. In recent years, investigators have used spline regression techniques to estimate more frequent (e.g. daily or monthly) DBP values by taking year, month and individual distribution system characteristics into account (Dodds et al., 1999; King et al., 2000; Dodds and King, 2001). This method, which interpolates between sample measurements, uses a polynomial function to join adjacent data points with a smoothed line (Greenland, 1998) and enabled Dodds et al. (1999) to evaluate average exposure over the last three months of pregnancy.

While analytical techniques are becoming more refined for studies of DBPs and adverse birth outcomes, exposure assessment techniques are still considered the greatest weakness of epidemiologic investigations (Reif et al., 1996; Nieuwenhuijsen et al., 2000a). Moreover, these new analytical techniques evaluated over the third trimester may increase bias when temporally variable environmental exposures, such as disinfection by-products (DBPs), are evaluated against time-dependent outcomes. This is because time itself is used to define or is highly correlated with many adverse birth outcomes, such as preterm birth (Hertz-Piccioto et al., 1996).

2.4 BIRTH RECORDS

The original birth certificate developed by the Census Bureau in 1900 was meant for legal purposes and as a means to study population change. Later revisions, issued by the National Center for Health Statistics (NCHS) increasingly included medical data and the latest revision taking effect in 1989 included several major changes (Freedman et al., 1988). First, while medical information had previously been asked in an open-ended format, the new birth certificates assess medical information in a check-box format. Second, specific questions identifying Hispanic origin were added (Freedman et al., 1988). Regardless, the accuracy of new birth certificates depends upon the variable of interest. For demographic variables, all checks of birth certificate validity indicate high concordance with "gold standard" medical records. Variables that require highly trained individuals, such as an obstetrician, for identification of birth defects/birth outcomes/complications/procedures are more likely to be inaccurate. The type of birth

defects considered in this study are likely to be identified through ultrasonography, echocardiography or upon visual inspection at birth (O'Rahilly and Müller, 2001).

2.4.1 Internal Reliability of Birth Certificate Information

Several studies of the validity of state birth records, relevant to the studies of reproductive outcomes in this dissertation, are published or readily available through state health departments.

Zollinger et al., at Indiana University Purdue University Indianapolis (IUPUI), analyzed 1996 birth certificate data quality for an unpublished Indiana State Department of Health report. The purpose of this study was to assess data quality and determine variation in quality by data element and hospital size. Using a 2-stage stratified cluster sample of 1050 Indiana births at 18 hospitals, Zollinger et al. matched medical records to the US standard birth certificate used in Indiana for 127 variables. The correlation (Kappa/ Pearson's r) among categories of variables extracted from birth and medical records was above the median agreement (0.364) for variables relating to parents' demographics, birth outcome, delivery and cesarean indications and prenatal care, pregnancy history and risk. The correlation was below the median agreement for congenital anomalies and abnormalities, complications of pregnancy, concurrent illnesses, obstetric procedures and labor and delivery complications. The researchers concluded that improvements should be made to the completeness of reporting infrequent events (such as complications of pregnancy). Also, categories of continuous variables were found to be more reliable, although more specific information is needed on mother's risk factors (e.g. pre-pregnancy

weight). It should be noted that this information was for 1996 births only and the study was limited by the accuracy of the medical records (Zollinger, personal communication).

In Colorado, an unpublished report was prepared to explain discrepancies in data for 515 birth certificate records from 1989. Data collected from hospitals were compared against birth certificate data in files maintained by the Heath Statistics Section of the Colorado Department of Public Health and Environment (CDPHE). Selected variables were analyzed under four categories: child information, parent information, prenatal information and checkbox items. Birth weight, a variable under the category of child information, was found to be in error for 9.4 percent of births, although nearly all births differed from hospital records by one ounce (2.4 percent of birth weights differed by greater than one ounce). Demographic variables (parent information) on birth certificate records were not found to differ greatly from medical records, except for the variables called Hispanic origin of Father (25 percent discrepant), Residence City (15.4 percent discrepant) and "other" terminations (9 percent of birth certificate records were discrepant). A large proportion of medical records lacked data on alcohol and tobacco use, while birth certificate records did not. Also, checkbox items (e.g. medical risk factors) were found to vary considerably between record sources (unpublished data, CDPHE).

One published study in California investigated the validity of racial/ethnic information of birth certificate data. In this study, 7,428 interviews with California mothers (postpartum) were used to validate corresponding birth certificate data with respect to

racial or ethnic information. The findings indicated that racial/ethnic reporting was 94 – 99 percent accurate for nearly all ethnicities, with the exception of Native Americans (54 percent sensitivity). The results validated the techniques of birth clerks in identifying the maternal racial/ethnic information for California birth certificate data, except in the case of Native Americans (Baumeister et al., 2000).

Another study by the same group of researchers assessed the validity of health insurance information on California birth certificates (Braveman et al., 1998). The results indicated that insurance information would be appropriate to use to determine the extent of coverage of maternity care and to investigate any relations between prenatal care use and insurance status.

A study conducted by Emery et al. (1997) in California sought to determine whether gestational age estimates are improved by the use of medical record data as compared to birth certificates. This stratified case-control study compared 172 singleton births with cerebral palsy to 472 randomly selected controls by birth weight. The records agreed for 60 percent of cases and controls in each of the birth weight groups. However, the records were most concordant for the highest birth weight strata.

Hexter et al. (1990) published results comparing hospital discharge diagnoses index (DI) for newborns and California birth certificates as sources of information for birth defects. Both records were evaluated against their corresponding entry in the case registry of the California Birth Defects Monitoring Program (CBDMP). Hexter et al. determined that

birth defect reporting on birth certificates was very inaccurate. The DI was most concordant with the CBDMP for oral clefts and chromosomal defects. Using the CBDMP as the gold standard, the DI contained a high proportion of false positives and for some defect groups a high proportion of false negatives.

Adams (2001) investigated the sensitivity of four variables relating to a previous pregnancy (the outcome of previous delivery, history of LBW, history of high birth weight and infant death resulting from preceding pregnancy) on Georgia birth certificates. Linking birth and death certificates, the sensitivities found for all outcomes, except infant death, were low. Sensitivities were higher for more severe adverse outcomes or when an outcome was repeated in the second delivery.

Dobie et al. (1998) compared Washington state birth certificates to clinic and hospital records for women that, at entry into prenatal care, were thought to be of low risk for pregnancy complications. The researchers determined sensitivity and percent agreement for pregnancy characteristics. Birth certificates lacked up to 24 percent of data depending on the variable. The concordance between hospital records and birth certificates was high for gravidity and parity; low for prenatal or intrapartum complications. Overall, Washington birth certificates grossly underestimated complications of pregnancy and the number of interventions, procedures and prenatal visits (Dobie et al., 1998).

Overall, the accuracy of birth certificates depends upon the variable of interest. Demographic variables seem to be recorded most accurately. Variables that require

highly trained individuals, such as an obstetrician, for identification (e.g. birth defects/birth outcomes/complications/procedures) are more likely to be inaccurate. However, the "gold standard" birth record should be questioned as a source of accurate information. For example, the concordance between records in the Colorado study indicated that birth certificate variables have a high percentage of completed answers, as in the case of alcohol intake, while the medical records were more likely to be deficient.

Except in the California studies of race/ethnicity and insurance status, the majority of validity studies involved a comparison of birth records to hospital or prenatal medical records. These studies were likely to contain several types of information bias. It is difficult to determine the accuracy of the medical records, although case registries (e.g. birth defect registries) may help resolve the validity of hospital record and birth certificate data.

2.5 RISK ESTIMATION

2.5.1 Study Designs

2.5.1.1 Inter-community Comparison

Inter-community analysis can be used to compare the incidence of adverse reproductive outcomes in communities with high exposure to chlorinated or brominated DBPs with the incidence in communities with low exposure. This design is appropriate for exposures that are spread throughout a community, such as drinking water contaminants. Ideally, a full-scale study would involve covariates at the individual level in order to decrease confounding bias inherent in an ecologic design. The inter-community comparisons are

most useful when the distributions of potential confounders such as socioeconomic status are similar in the two communities (Hertz-Piccioto, 1998). This variety of study design was used in several previous quantitative studies, such as in Kramer et al. (1992) and Bove et al.(1995).

2.5.1.2 Intra-system Comparison

Another method of conducting studies is through an intra-community approach. The intra-community comparison would be the most adequate method of assessing health effects that have well-defined latency periods (e.g. birth defects), or where there is a defined line between the unexposed and exposed time periods (e.g. a change in water treatment processes). However, this method can only be used to evaluate health effects due to threshold DBP exposures. If cumulative low-level exposures were expected to lead to health effects (e.g. reduced sperm production) then the inter-community analysis would be more useful (Hertz-Piccioto, 1998).

Exposure to DBPs can also be evaluated for associations with reproductive outcomes via time-series analysis. This method of analysis compares health effects from exposures that fluctuate over time in a single community and makes use of the dramatic temporal variability that occurs in some communities. For example, in Arizona, a woman can experience very high to low exposures to DBPs at a single location, depending on the critical time period of exposure. For birth defects, this is the first trimester and for birth weight and other developmental delays, this is most likely the last trimester of pregnancy. The cumulative dose of exposure to a specific contaminant could be calculated for a subject from her gestational history, using monthly exposure data, by one of several

statistical approaches. In addition, this method minimizes the effect of confounders that vary between individuals but generally remain constant over the time period of interest in a subject, e.g. occupation. Generally, this analysis would be most useful to evaluate birth outcomes that have well-known critical time periods of development, such as specific birth defects (Hertz-Piccioto, 1998).

2.5.2 Relevant Time Periods of Exposure

During fetal development, the sensitivity to teratogens is contingent upon the developmental stage at the time of exposure. For example, the 14th day (gastrulation) of gestation is particularly sensitive to induction of cleft palate (O'Rahilly and Müller, 2001). Therefore, with respect to assigning exposure, it is appropriate to determine exposure as close as possible to the second or third week of pregnancy, if not on day 14. However, for some adverse outcomes of pregnancy, e.g. LBW or preterm birth, there is no time period during human embryologic or fetal development that is known to be a period of sensitivity, although the third trimester is most often suspected and investigated (Savitz et al., 1995; Gallagher et al., 1998; Dodds et al., 1999; Wright et al., 2003).

Hertz-Picciotto et al. (1996) has shown that for studies of environmental epidemiology, collecting exposure by month of gestation is feasible. Analyses may depend, however, on whether there is knowledge of the critical time window or induction time. For example, when knowledge of the critical time period is uncertain, but exposure data is relatively accurate, comparison of different induction times is appropriate. This can aid in establishing an induction period as related to the exposure, without dilution of exposure data over the entire gestation. This method of analysis is entirely appropriate

for studies of DBPs where the exposure is variable between subjects and through time (Hertz-Piccioto, 1996).

2.5.2.1 Growth Outcomes

As mentioned previously, the fetal period extends from the 9th week of gestation until birth and is characterized by rapid growth of the body and maturation of tissues and organ systems. At the 5th month of gestation, fetal weight is only 500 g. During the 2nd half of intrauterine life, weight gain is substantial. The most important weight addition is during the last 2-½ months of a term pregnancy, with approximately 50 percent of full term weight achieved (Sadler, 1995).

2.5.2.2 Malformations

The severity of abnormal development depends upon the dose and duration of exposure to teratogens. The sensitivity to teratogens is contingent upon the developmental stage at the time of exposure (Moore and Persaud, 1998). The period of embryogenesis, third to eighth week of gestation, is considered the most sensitive period for producing birth defects (O'Rahilly and Müller, 2001). Each organ system may have multiple periods of sensitivity to an exposure, as in the case of cleft palate (Larsen, 1997). This malformation is inducible during the 6th day (blastocyst stage), 14th day (gastrulation), during the 5th week (early limb bud stage), or during the 7th week when palatal shelves are developing. While most abnormalities tend to occur during early pregnancy (embryogenesis), there is no period of gestation that is completely resistant to teratogenesis (Sadler, 1995).

Cleft lip defects are most likely to occur in the 2nd month of pregnancy, with major malformations in weeks 5 and 6 and minor malformations in weeks 7 and 8. Major palate defects will occur during the 7th or 8th weeks of gestation, while minor defects are likely in week 9 (O'Rahilly and Müller, 2001).

Major cardiac malformations (Truncus arteriosus, atrial septal defects and ventricular septal defects) are thought to occur between weeks 2 and 6. Minor defects are induced in the 7th and 8th weeks of gestation (O'Rahilly and Müller, 2001).

Nearly all neural tube defects are considered major and are most likely induced between 2 and 6 weeks gestation (O'Rahilly and Müller, 2001).

2.6 DBPS AND SPECIFIC BIRTH OUTCOMES

2.6.1 Epidemiologic Data

A summary of estimated risk effects for outcomes evaluated in previous quantitative studies of DBPs is presented below.

2.6.1.1 Fetal Growth

The epidemiological evidence for an association between TTHMs and indicators of fetal growth is inconsistent. The potential confounders of education or socioeconomic status and race were adjusted for in all analyses. Smoking status was controlled for in all but the Bove et al. (1995) study. These studies may not be completely comparable, though, due to the considerable variation with which exposure was assigned.

2.6.1.1.1 Birthweight

Two studies examined birth weight as a continuous variable. In the first study among term births exposed to TTHM concentrations greater than 100 μg/l, birth weight was found to decrease an average of 70.4 g compared to the (≤ 20 μg/l) referent group (Bove et al., 1995). The more recent study by Wright et al. (2003) found that birth weight decreased by 32 g for term births exposed to an average of greater than 80 μg/l TTHM during pregnancy, as compared to a referent group of ≤ 60 μg/l.

2.6.1.1.2 Low birthweight

Four studies have evaluated LBW (term and preterm births) and exposure to DBPs. Kramer et al. (1992) found weak non-significant associations (adjusted OR=1.3, 95% CI 0.8-2.2) for exposure to chloroform ≥10 μg/l as compared to a referent group of non-detect exposures. Notably, the chloroform concentrations in this study were very low compared to other studies of DBPs, and a critical time period of exposure during gestation was not taken into account. For studies of DBP exposure during the third trimester, the time period of rapid fetal weight gain, results for LBW were inconsistent. Savitz et al. (1995) found no association at the highest level of TTHM exposure or dose (μg/l x glasses/day) for LBW. Gallagher et al. (1998) found moderate associations, with an adjusted OR=2.1 (95% CI 1.0-4.8) for TTHM >61 μg/l. However, this study was subject to bias through small sample sizes and exclusion of many participants. In addition, the cutpoint for LBW in this Colorado study was slightly different than the cutpoint used in the other three studies, as it included 2,500 g. The most recent study by

Dodds et al. (1999) found no association at the highest level of exposure (TTHM ≥100 μg/l).

Term LBW was analyzed in three studies of DBPs. For the Gallagher et al. study, the OR for LBW was stronger (adjusted OR=5.9, 95% CI 2.0-17.0) when limited to term births, although only 6 cases were analyzed. The study with the largest sample size (Bove et al., 1995) also observed a positive, albeit smaller, association for TLBW with an OR=1.42 (50% CI 1.22-1.65), although no dose-response trend was observed. The study by Bove et al. (1995) also differed methodologically from the other studies by analyzing exposure over the entire gestation and by not adjusting results for smoking. The study by Wright et al. (2003) reported no associations between TLBW and trimester-specific exposures or entire pregnancy exposures to TTHM.

2.6.1.1.3 Very low birthweight

Two studies have assessed VLBW and exposure to DBPs (Bove et al., 1995; Dodds et al., 1999). Despite their large sample sizes, neither study found any association for VLBW and DBPs.

2.6.1.1.4 Preterm birth

Of the six studies that evaluated preterm birth, none found an association with DBPs (Kramer et al., 1992; Bove et al., 1995; Savitz et al., 1995; Gallagher et al., 1998; Dodds et al., 1999; Wright et al., 2003).

2.6.1.1.5 Intrauterine growth retardation

In a study by Kramer et al. (1992) a dose-related trend was observed for IUGR at the 5th percentile for exposure to chloroform and BDCM. Exposure to chloroform ≥10 μg/l and BDCM ≥4 μg/l, produced ORs of 1.8 (95% CI 1.1-2.9) and 1.7 (95% CI 0.9-2.9), respectively, for risk of IUGR. Results from the study by Bove et al. (1995) also showed increased risk (adjusted OR=1.50, 90% CI 1.19-1.86) with SGA (5th percentile) and exposure to greater than 100 μg/l TTHMs. The results from the study by Dodds et al. (1999) were not consistent with the above findings. No association was found for SGA and TTHM ≥100 μg/l. In a Massachusetts cohort, Wright et al. (2003) found increased risk of SGA for pregnancy average exposures (adjusted OR=1.14, 95% CI 1.02-1.26) and second trimester exposures (adjusted OR=1.13, 95% CI 1.03-1.24) to TTHMs greater than 80 μg/l. These studies may not be comparable, however, due to use of different referent values and exposure time periods. For example, the referent group for the Massachusetts study had fairly high DBP exposures ranging from 0-60 μg/l, while the referent group from New Jersey had exposures of 20 μg/l or less. In addition, the three studies with positive associations did not assess exposure during the third trimester, while the study with equivocal results, Dodds et al. (1999), approximated third trimester exposure by evaluating the last three months of pregnancy. Lastly, residual confounding may still exist for studies that did not adjust for smoking and alcohol (Bove et al., 1995).

2.6.1.2 Fetal Viability

As with the growth-related outcomes, statistical associations between DBPs and SAB or stillbirth have been observed but are inconsistent. While both endpoints indicate

unsuccessful pregnancies, their mechanisms of induction may differ, as may their potential for susceptibility to DBPs.

2.6.1.2.1 Spontaneous abortion

Savitz et al. (1995) found a statistically significant relationship with SAB and increasing concentration of TTHM in North Carolina residents with the highest sextile of exposure (adjusted OR=2.8, 95% CI 1.1-2.7), but no relationship was determined for ingested dose or water source. In a study in California by Waller et al., (1998), increased rates of SAB (adjusted OR=1.2, 95% CI 1.0-1.5) were found for concentrations $\geq$ 75 μg/l TTHM. The highest risk reported (adjusted OR=2.0, 95% CI 1.2-3.5) was for consumption of water containing BDCM >18 μg/l. In addition, improvements in exposure assessment methods in a later analysis of the same cohort yielded higher risk estimates (adjusted OR=3.3, 95% CI 1.3-8.6) for SAB and the highest levels of TTHM exposure in a subset population (Waller et al., 2001).

2.6.1.2.2 Stillbirth

In New Jersey, little evidence was found for an association between stillbirth and TTHM exposures greater than 80μg/l (Bove et al., 1992). An increased risk of stillbirth (adjusted OR=1.66, 95% CI 1.09-2.52) was reported for Nova Scotia women exposed to water having greater than 100 μg/l TTHMs (Dodds et al. 1999). In further analyses of these data, King et al. (2000) found dose dependent increases in adjusted risk for stillbirth with exposure to TTHM, chloroform and BDCM. High BDCM exposure (> 20 μg/l) had the strongest association with stillbirth, with an adjusted OR=1.98 (95% CI 1.23-3.49). In

analyses of specific causes of stillbirth, the strongest relationship was found for exposure to TTHMs, BDCM and chloroform and asphyxia-related fetal deaths.

2.6.1.3 Malformations

The literature presents an inconsistent pattern of association between collective or specific congenital anomalies and exposure to DBPs. Adequate sample size for statistical assessment of risk for congenital anomalies is rare as small numbers of cases are typically available for study. In addition, selection bias may occur due to elective terminations of pregnancy (Reif et al., 1996). Grouping of specific defects into organ system defects may also lead to investigation of etiologically unrelated defects and, subsequently, diluted risk estimates (Reif et al., 1996). Further, if higher doses of a teratogen induce multiple or severe defects, outcomes could be expressed as stillbirth, SAB, or unrecognized fetal loss ((O'Rahilly and Müller, 2001). Thus far, there have been few studies of THMs and congenital defects in humans and only one study has been published evaluating exposure to HAAs (Klotz and Pyrch, 1999). However, identification and reporting of neonatal defects is improving each year (O'Rahilly and Müller, 2001) and improvements in case selection will naturally occur with future studies of DBPs and congenital anomalies.

2.6.1.3.1 Oral cleft defects

In the 1995 study by Bove et al., a threefold increase in risk (adjusted OR = 3.17, 90% CI 1.18-4.89) for oral clefts were observed with exposure to TTHM concentrations >100 μg/l, as compared to the referent group. No trend, however, was apparent with increasing TTHM dose. In contrast, analyses of the individual THM species of chloroform and BDCM in the study by Dodds et al. (2001) yielded no increased risk of oral cleft defects.

Shaw et al. (2003), which analyzed cleft lip and cleft palate cases separately, also found no clear associations with TTHM exposures.

2.6.1.3.2 Cardiac malformations

Major cardiac malformations have been investigated for association with exposure to DBPs in one study in New Jersey (Bove et al., 1995), in the Nova Scotia cohort (Dodds et al., 1999; Dodds and King, 2001) and in the California study by Shaw et al. (2003). Exposure to TTHMs was not significantly associated with cardiac defects in any cohort. BDCM exposures ≥20 μg/l, however, were negatively associated (adjusted RR=0.3, 95% CI 0.2-0.7) with cardiovascular anomalies in the study by Dodds and King (2001).

2.6.1.3.3 Neural tube defects

In the New Jersey study by Bove et al. (1995), exposure to the highest level (>80 μg/l) of TTHMs was positively associated with NTDs with an OR=2.96 (90% CI 1.26-6.62). No consistent trend, however, was observed for increasing TTHM concentrations. In addition, studies by Klotz and Pyrch (1999), Dodds et al. (2001) and Shaw et al. (2003) reported several negative associations for varying TTHM exposure levels, with no consistent trends.

Recent analyses of NTDs and the individual THMs suggest that BDCM may be responsible for previously observed associations. In the Nova Scotia population studied by Dodds et al. (2001), exposure to BDCM at levels greater than 20 μg/l resulted in a statistically significant associations (OR=2.5) with NTDs (Dodds and King, 2001).

2.6.2 Toxicologic Data

Animal evidence suggests that DBP exposure can contribute to reduced fetal growth and increased malformations *in utero*. Reductions in body weight, reduced crown-rump length and growth retardation have been shown in experimental animals treated with chloroform at high doses (Thompson et al., 1974; Schwetz et al., 1974; Murray et al., 1979; Ruddick et al., 1983). Several toxicologic studies have also shown THMs to be fetotoxic (Murray et al., 1979; Ruddick et al., 1983; Borzelleca et al., 1982). Narotsky et al. (1997) demonstrated fetal resorptions at 50 and 75 mg/kg/day bromodichloromethane and fetotoxicity at 20 mg/kg/day.

BDCM exposures have reportedly reduced serum progesterone levels after single dose administrations (Bielmeier et al., 2001). The mode of action for induction of full litter resorption after exposure to BDCM potentially involves disruption of luteal responsiveness to LH. This has been evidenced in F344 rats by adverse outcomes following exposures during the LH-dependent period of pregnancy and reduction in progesterone levels without corresponding reductions in LH levels (Bielmeier et al, 2001).

Several halogenated acetic acids have also shown the capacity to induce embryonic resorption and fetal death. Smith et al. (1992) demonstrated that both di- and tri-chloroacetic acid had the ability to cause these effects in rodent experiments. HAAs have also contributed to teratogenic effects in rats. DCAA increased the rate of cardiac malformations at 140 mg DCAA /kg day and higher doses (Smith et al., 1989). While all

the toxicity data for DBPs in animal studies are several orders of magnitude higher than human exposures, they still support the biological plausibility of DBP effects on embryonic and fetal development.

Reproductive toxicity in toxicologic assessment for DBPs is dependent upon species of animal and time of exposure. For example, Bielmeier et al. (2001) found that BDCM exposure in early gestation is associated with full litter resorption in F344 rats, but has no effect upon Sprague-Dawley (SD) rat strains. In addition, BDCM administration affected resorption in F344 rats after exposure in gestational days 6-10, but not days 11-15.

Few studies using rabbits have been performed to evaluate DBP toxicity, although they are considered the most appropriate animal model to approximate human reproductive and developmental risk (Christian et al., 2001). Christian et al. (2001) used the rabbit model to study BDCM at doses similar to low-dose human exposures and at higher levels and found only a decrease in maternal weight gain. A related study found no significant effects in two generations of SD rats (Christian et al., 2002).

2.6.2.1 Modes of Toxicity

Both chronic and acute toxicity can provide a mode of action for increases in congenital anomalies. Chronic exposures are most likely required to perturb biochemical and endocrine events associated with gamete toxicity, potentially leading to increases in chromosomal anomalies. Also, congenital anomalies may result from single exposures causing mutation events in the early embryonic stage (Klassen et al., 1996). These

effects have been observed after exposures to environmental and chemical toxins (Rutledge et al., 1997).

The mode of action for DBP toxicity is dependent upon the class of DBPs and the type of halogens involved. Trihalomethane toxicity is thought to occur through two modes of action, both involving reactions with regulatory proteins (Ross and Pegram, 2003). Chloroform is most likely oxidized by Cytochrome P4502E1 to form a phosgene metabolite, the product responsible for toxic and carcinogenic activity. Chloroform is toxic at high levels, with little toxic effects at low dose human exposures (Pegram et al., 1997). Low dose brominated THMs are potentially metabolized through another pathway, in which the enzyme glutathione-S-transferase-theta (GST) catalyzes a conjugation reaction between glutathione (GSH) and the THM (Pegram et al., 1997). This metabolic pathway is favored by brominated THMs because they have a better-leaving group than chlorinated THMs. Generally, the better the leaving group the faster the metabolic reactions. The metabolic reaction catalyzed by GST yields mutagenic compounds that are likely to have carcinogenic effects.

Potential modes of action for some HAAs include production of mutagenic metabolites and alterations of proteins that are involved in signaling pathways. Chemicals known to induce cardiac malformations, such as trichloroethylene, also metabolize to HAAs, suggesting that these DBPs may take part in a toxic pathway (Johnson, 1998). Also, as with the brominated THMs, brominated haloacetic acids have showed greater mutagenic and genotoxic potential than chlorinated HAAs (Plewa et al., 2002).

Metabolism of haloacids is thought to form acetylated protein adducts that could accumulate through chronic dosing. This accumulation could permanently change important signaling proteins and, ultimately, disrupt embryonic development (Ross and Pegram, 2003). Both retinoic acid and dioxin have demonstrated teratogenicity through disruption of signaling. Altered response to morphogen signals after exposure to these chemicals has resulted in interrupted growth and differentiation of epithelial cells during mesenchymal development. Modulation of epithelial cell responses resulting from exposures during critical developmental time periods can lead to malformations such as cleft palate (Abbott and Birnbaum, 1990; Abbott and Birnbaum, 1989).

2.7 REFERENCES

Abbott BD. Birnbaum LS. Rat embryonic palatal shelves respond to TCDD in organ culture. Toxicology & Applied Pharmacology. 1990.103(3):441-51.

Abbott BD. Birnbaum LS. TCDD-induced altered expression of growth factors may have a role in producing cleft palate and enhancing the incidence of clefts after coadministration of retinoic acid and TCDD. Toxicology & Applied Pharmacology. 1990.106(3):418-32.

Adams M. Validity of birth certificate data for the outcome of the previous pregnancy, Georgia, 1980-1995. American J of Epidemiology. 2001: 154(10); 883-888.

Agency for Health Care Policy and Research. Low Birthweight in Minority and High-Risk Women. Final Report of the Low-Birthweight Patient Outcomes Research Team. (Abstract). Rockville, MD. 1998. [cited 7 December 2000] Available on World Wide Web:(http://www.ahrq.gov/clinic/lobrabs.htm)

Arbuckle TE. Hrudey SE. Krasner SW. Nuckols JR. Richardson SD. Singer P. Mendola P. Dodds L. Weisel C. Ashley DL. Froese KL. Pegram RA. Schultz IR. Reif J. Bachand AM. Benoit FM. Lynberg M. Poole C. Waller K. Assessing exposure in epidemiologic studies to disinfection by-products in drinking water: report from an international workshop.Environmental Health Perspectives. 2002:110 Suppl 1;53-60.

Aschengrau A. Zierler S. Cohen A. Quality of community drinking water and the occurrence of late adverse pregnancy outcomes. Archives of Environmental Health. 1993:48(2);105-13.

Backer LC. Ashley DL. Bonin MA. Cardinali FL. Kieszak SM. Wooten JV. Household exposures to drinking water disinfection by-products: whole blood trihalomethane levels. Journal of Exposure Analysis & Environmental Epidemiology. 2000:10(4);321-6.

Barbone F. Valent F. Brussi V. Tomasella L. Triassi M. Di Lieto A. Scognamiglio G. Righi E. Fantuzzi G. Casolari L. Aggazzotti G. Assessing the exposure of pregnant women to drinking water disinfection byproducts. Epidemiology. 2002:13(5);540-4.

Baumeister L. Marchi K. Pearl M. Williams R. Braveman P. The validity of information of "race" and "Hispanic ethnicity" in California birth certificate data. Health Services Research. 2000: 35(4); 869-83.

Bellar TA, Lichenberg JJ. Determining volatile organic compounds at microgram-per-litre levels by gas chromatography. J Am Water Works Assoc. 1974:66;739-744.

Bielmeier SR, Best DS, Guidici DL, Narotsky MG. Pregnancy loss in the rat caused by bromodichloromethane. Toxicological Sciences. 2001:59; 309-15.

Borzelleca JF and Carchman RA. Effects of selected organic drinking water contaminants on male reproduction. Research Triangle Park, NC, US Environmental Protection Agency, 1982 (EPA 600/1-82-009; NTIS PB82-259847; Contract No. R804290).

Bove FJ, Fulcomer MC, Koltz JB, Esmart J, Dufficy EM, Zagraniski RT. Report on phase IV-A: Public drinking water contamination and birthweight fetal deaths, and birth defects, a cross-sectional study. New Jersey Dept. of Health, 1992.

Bove FJ, Fulcomer MC, Klotz JB, Esmart J, Dufficy EM, Savrin JE. Public water contamination and birth outcomes. Amer J Epidemiol. 1995:141;850-862.

Bove F. Shim Y. Zeitz P. Drinking water contaminants and adverse pregnancy outcomes: a review. Environmental Health Perspectives. 2002:110 Suppl 1;61-74.

Braveman P. Pearl M. Egerter S. Marchi K. Williams R. Validity of insurance information of California birth certificates. Am J of Public Health. 1998: 88(5);813-6.

Burmaster DE. Lognormal distributions for total water intake and tap water intake by pregnant and lactating women in the United States. Risk Analysis. 1998:18(2);215-9.

Calafat AM. Kuklenyik Z. Caudill SP. Ashley DL. Urinary levels of trichloroacetic acid, a disinfection by-product in chlorinated drinking water, in a human reference population. Environmental Health Perspectives. 2003:111(2);151-4.

Cedergren MI. Selbing AJ. Lofman O. Kallen BA. Chlorination byproducts and nitrate in drinking water and risk for congenital cardiac defects. Environmental Research.2002: 89(2);124-30.

Chen WJ, Weisel CP. Concentration changes of halogenated disinfection by-products within a distribution system. Am Water Works Assoc J 1998:90:151-163.

Christian MS, York RG, Hoberman AM, Fisher LC, Brown WR. Oral (Drinking Water) two-generation reproductive toxicity study of bromodichloromethane (BDCM) in Rats. International Hournal of Toxicology. 2001:20;225-237.

Christian MS, York RG, Hoberman AM, Diener RM, Fisher LC. Oral (Drinking Water) developmental toxicity studies of bromodichloromethane (BDCM) in Rats and Rabbits. International Hournal of Toxicology. 2002:21;115-146.

Clayman CB. *The American Medical Association Encyclopedia of Medicine.* New York: Random House; 1989: 941-942.

Dobie SA. Baldwin LM. Rosenblatt RA. Fordyce MA. Andrilla CH. Hart LG. How well do birth certificates describe the pregnancies they report? The Washington State experience with low-risk pregnancies. Maternal and Child Health Journal. 1998: 2(3); 145-54.

Dodds L, King W, Wolcott C, Pole J. Trihalomethanes in public water supplies and adverse birth outcomes. Epidemiology. 1999:443-446.

Dodds L, King W. Relation between trihalomethane compounds and birth defects. Occup Environ Med. 2001:58;233-237.

Emery ES. Eaton A. Grether JK. Nelson KB. Assessment of gestational age using birth certificate data compared with medical record data. Paediatric and Perinatal Epidemiology. 1997: 11(3); 313-21.

Ershow AG, Brown LM, Cantor KP. Intake of tapwater and total water by pregnant and lactating women. American Journal of Public Health. 1991: 81(3);328-34.

Freedman MA, Gay GA, Brockert JE, Potrzebowski PW, Rothwell CJ. The 1989 revisions of the US Standard Certificates of Live Birth and Death and the US Standard Report of Fetal Death. American Journal of Public Health 1988;78:168-172.

Froese KL. Sinclair MI. Hrudey SE. Trichloroacetic acid as a biomarker of exposure to disinfection by-products in drinking water: a human exposure trial in Adelaide, Australia. Environmental Health Perspectives. 2002:110(7);679-87.

Gallagher MD, Nuckols JR, Stallones L, Savitz DA. Exposure to trihalomethanes and adverse pregnancy outcomes. Epidemiology. 1998:9;484-489.

Geenland S. Introduction to Regression Models. In: Rothman KJ, Greenland S. eds. *Modern Epidemiology,* 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 1998.

Golfinopoulos SK. Arhonditsis GB. Multiple regression models: a methodology for evaluating trihalomethane concentrations in drinking water from raw water characteristics. Chemosphere. 2002:47(9);1007-18.

Hertz-Picciotto I., Pastore L., Beaumont JJ. Timing and patterns of exposures during pregnancy and their implications for study methods. Amer J Epidemiol 1996:143 (6);597-607.

Hertz-Picciotto I. Environmental Epidemiology. In: Rothman KJ, Greenland S. eds. *Modern Epidemiology,* 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 1998.

Hexter AC. Harris JA. Roeper P. Croen LA. Krueger P. Gant D. Evaluation of the hospital discharge diagnoses index and the birth certificate as sources of information on birth defects. Public Health Reports. 1990: 105(3); 296-307.

Hopkins SM, Ellis JC. Drinking Water Consumption in Great Britain. Marlow, Bucks: Water Research Centre.1980.

Infante-Rivard C. Amre D. Sinnett D. GSTT1 and CYP2E1 polymorphisms and trihalomethanes in drinking water: effect on childhood leukemia. Environmental Health Perspectives. 2002:110(6);591-3.

Jaakkola JJ. Magnus P. Skrondal A. Hwang BF. Becher G. Dybing E. Foetal growth and duration of gestation relative to water chlorination. Occupational & Environmental Medicine. 2001:58(7);437-42.

Johnson PD. Dawson BV. Goldberg SJ. A review: trichloroethylene metabolites: potential cardiac teratogens. Environmental Health Perspectives. 1998:106 Suppl 4; 995-9.

Kallen BA. Robert E. Drinking water chlorination and delivery outcome-a registry-based study in Sweden. Reproductive Toxicology. 2000:14(4);303-9.

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

Keegan T. Whitaker H. Nieuwenhuijsen MJ. Toledano MB. Elliott P. Fawell J. Wilkinson M. Best N. Use of routinely collected data on trihalomethane in drinking water for epidemiological purposes. Occupational & Environmental Medicine.2001:58(7);447-52.

Kerger BD. Schmidt CE. Paustenbach DJ. Assessment of airborne exposure to trihalomethanes from tap water in residential showers and baths. Risk Analysis. 2000:20(5);637-51.

Khoury M, Stewart W, Weinstein A, Panny S, Lindsay P, Eisenberg M. Residential mobility during pregnancy: implications for environmental teratogenesis. J Clin Epidemiol. 1988:41;15-20.

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

Klassen, C D. Casarett & Doull's Toxicology: The Basic Science of Poisons. Thomas J (ed). New York. McGraw-Hill, 1996. Pg. 567.

Klotz JB, Pyrch LA. Neural tube defects and drinking water disinfection by-products. Epidemiology. 1999:10;383-390.

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

Krasner S, McGuire M, Jacangelo J, Patania N, Reagan K, Aieta E. The occurrence of disinfection by-products in US drinking water. J Am Water Works Assoc. 1989: 91; 41-53.

Larsen, WJ. *Human Embryology* 2nd Edition. New York: Churchill Livingstone. 1997: 471-486.

Lynberg M. Nuckols JR. Langlois P. Ashley D. Singer P. Mendola P. Wilkes C. Krapfl H. Miles E. Speight V. Lin B. Small L. Miles A. Bonin M. Zeitz P. Tadkod A. Henry J. Forrester MB. Assessing exposure to disinfection by-products in women of reproductive age living in Corpus Christi, Texas, and Cobb county, Georgia: descriptive results and methods. Environmental Health Perspectives. 2001. 109(6):597-604.

Magnus P. Jaakkola JJ. Skrondal A. Alexander J. Becher G. Krogh T. Dybing E. Water chlorination and birth defects. Epidemiology. 1999:10(5);513-7.

Miles AM. Singer PC. Ashley DL. Lynberg MC. Mendola P. Langlois PH. Nuckols JR. Comparison of trihalomethanes in tap water and blood. *Environmental Science & Technology.*2002: 36(8);1692-8.

Moore KL. Persaud T.V.N. *Before we are born. Essentials of embryology and birth defects.* 5th edition. W. B. Saunders Company. Philadelphia, PA. 1998.

Murray FJ, Schwetz BA, McBride JF and Staples RE. Toxicity of inhaled chloroform in pregnant mice and their offspring. Toxicol Appl Pharmacol. 1979:50;515-522.

Narotsky MG, Pegram RA, Kavlock RJ. Effect of dosing vehicle on the developmental toxicity of bromodichloromethane and carbon tetrachloride in rats. Fundament Appl Toxicol. 1997: 40; 30-36.

National Vital Statistics Report. Number and percent low birthweight and number of live births by birthweight, by age and race and Hispanic origin of mother: United States, 1997. CDC. Vol. 47, No. 18, April 29, 1999. [cited 11 September 2000] Available on World Wide Web:(http://www.cdc.gov/nchswww/data/nvs47_187.pdf)

Nieuwenhuijsen MJ. Toledano MB. Eaton NE. Fawell J. Elliott P. Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: a review. Occupational & Environmental Medicine. 2000a:57(2);73-85.

Nieuwenhuijsen MJ. Toledano MB. Elliott P. Uptake of chlorination disinfection by-products; a review and a discussion of its implications for exposure assessment in epidemiological studies. *Journal of Exposure Analysis & Environmental Epidemiology.* 2000b: 10(6 Pt 1);586-99.

O'Rahilly R. Muller F. *Human Embryology and Teratology.* 3rd Edition. John Wiley & Sons, Inc. New York, NY, 2001.

Pegram RA, Andersen ME, Warren SH, Ross TM, Claxton LD. Glutathione S-transferase-mediated mutagenicity of trihalomethanes in Salmonella typhimurium: contrasting results with bromodichloromethane off chloroform. Toxicology & Applied Pharmacology. 1997:144(1);183-8.

Plewa MJ, Kargalioglu Y, Vankerk D, Minear RA, Wagner ED. Mammalian cell cytotoxicity and genotoxicity analysis of drinking water disinfection by-products. Environmental and Molecular Mutagenesis. 2002: 40(2);134-42.

Reif JS, Hatch MC, Bracken M, Holmes L, Schwetz B, Singer PC. Reproductive and developmental effects of disinfection byproducts in drinking water. Environ Health Perspect. 1996:104;1056-1061.

Reif JS, Bachand A, Andersen ME. Risk Assessment of Reproductive Effects of Disinfection By-Products. Bureau of Reproductive and Child Health, Health Canada. Tunney's Pasture, Ottawa, Ontario, Canada. April, 2000.

Rook JJ. Formation of haloforms during chlorination of natural waters. J soc Water Treat Exam. 1974: 23;234-243.

Ross MK. Pegram RA. Glutathione transferase theta 1-1-dependent metabolism of the water disinfection by-product bromodichloromethane. Chem Res in Tox. 2003:16(2);216-26.

Rossman LA. Brown RA. Singer PC. Nuckols JR. DBP formation kinetics in a simulated distribution system. Water Research. 2001:35(14);3483-9.

Ruddick JA, Villeneuve DC, Chi I, Valli VE. A teratological assessment of four trihalomethanes in the rat. J Environ Sci Health, Part B: Pesticides, food contaminants, and agricultural wastes. 1983:B18;333-349.

Rutledge JC. Developmental toxicity induced during early stages of mammalian embryogenesis. Mut Res. 1997:396;113-127.

Sadler, TW. *Langman's Medical Embryology* 7th Edition. Philadelphia, PA: Williams and Wilkins. 1995: 90-99.

Savitz DA, Andrews KW, Pastore LM. Drinking water and pregnancy outcome in central North Carolina: Source amount and trihalomethane levels. Environ Health Perspect. 1995:103;592-596.

Schwetz BA, Leong BKJ and Gehring PJ. Embryo- and fetotoxicity of inhaled chloroform in rats. Toxicol Appl Pharmacol. 1974:28;442-451.

Shaw GM, Milea A., Swan SH. Chlorinated water exposures and congenital cardiac anomalies. Epidemiology. 1991:6;459-60.

Shaw GM, Malcoe LH. Residential mobility during pregnancy for mothers of infants with or without congenital cardiac anomalies: A reprint. Arch Environ Health. 1992:47;236-238.

Shaw GM, Ranatunga D, Quach T, Neri E, Correa A, Neutra RR. Trihalomethane exposures from municipal water supplies and selected congenital malformations. Epidemiology. 2003:14;191-199.

Shimokura GH, Savitz DA, Symanski E. Assessment of water use for estimating exposure to tap water. Environ Health Perspect. 1998:106;55-59.

Smith M, Randall J, Stober J. Teratogenic activity of trichloroacetic acid in the rat. Teratology. 1989: 40; 445-451.

Smith MK, Randall JL, Read JL, Stober JA. Developmental toxicity of dichloroacetate in the rat. Teratology. 1992:46;217-223.

Swan SH, Waller K, Hopkins B, Windham G, Fenster L, Schaefer C, Neutra R. A prospective study of spontaneous abortion; relation to amount and source of drinking water consumed in early pregnancy. Epidemiology. 1998:9;126-133.

Thompson DJ, Warner SD, Robinson VB. Teratology studies on orally administered chloroform in the rat and the rabbit. Toxicol Appl Pharmacol. 1974:29;348-357.

Tuthill RW, Giusti RA, Moore GS, Calabrese EJ. Health effects among newborns after prenatal exposure to ClO_2-disinfected drinking water. Environ Health Perspect. 1982:46;39-45.

U.S. Environmental Protection Agency, International Life Sciences Institute, Risk Sciences Institute. Report of the Panel on Reproductive Effects of Disinfection By-Products in Drinking Water, Research Triangle Park, NC, 1993.

U.S. Environmental Protection Agency. Part IV National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts; Final Rule. U.S. Federal Register December 16, 1998. 1998:63;69390-69476.

Waller K, Swan SH, DeLorenze G, Hopkins B. Trihalomethanes in drinking water and spontaneous abortion. Epidemiology. 1998:9;134-140.

Waller K, Swan SH, Windham GC, Fenster L. Influence of exposure assessment methods on risk estimates in an epidemiologic study of total trihalomethane exposure and spontaneous abortion. Journal of Exposure Analysis & Environmental Epidemiology. 2001:11;522-531.

Weinberg CR, Wilcox AJ. Reproductive Epidemiology. In: Rothman KJ, Greenland S. eds. *Modern Epidemiology,* 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 1998.

Weisel CP, Jo WK. Ingestion, inhalation, and dermal exposures to chloroform and trichloroethene from tap water. Environmental Health Perspectives. 1996:104(1);48-51.

Weisel CP. Kim H. Haltmeier P. Klotz JB. Exposure estimates to disinfection by-products of chlorinated drinking water. Environmental Health Perspectives. 1999:107(2);103-10.

Wright JM. Schwartz J. Dockery DW. Effect of trihalomethane exposure on fetal development. Occupational & Environmental Medicine. 2003:60(3);173-80.

Yang CY. Cheng BH. Tsai SS. Wu TN. Lin MC. Lin KC. Association between chlorination of drinking water and adverse pregnancy outcome in Taiwan. Environmental Health Perspectives. 2000:108(8);765-8.

Zender R, Bachand AM, Reif JS. Exposure to tap water during pregnancy. Journal of Exposure Analysis & Environmental Epidemiology. 2001:11(3); 224-30.

Zollinger TW, Przybylski M. Birth Certificate Data Quality Study. (Unpublished Data). Presentation Received on July 31, 2001.

CHAPTER 3

3 SELECTION OF SITES HAVING DISINFECTION BY-PRODUCTS WITH LOW SPATIAL VARIABILITY

3.1 INTRODUCTION

Disinfection by-products (DBPs) are common contaminants in treated drinking water that can be classified into several subgroups, including: trihalomethanes (THMs) and haloacetic acids (HAAs) (Nieuwenhuijsen et al, 2000). Studies in laboratory animals have shown that brominated trihalomethanes and haloacetic acids are relatively more effective than chlorinated species in inducing reproductive effects (Pegram et al., 1997).

Several conditions (e.g. organic matter quantity, chlorine dose, pH, temperature, contact time and bromide ion concentration) can cause variability in formation of DBPs. In particular, due to increased disinfection contact time, DBP concentrations increase with longer times spent in a distribution system leading to spatial variability of contaminant levels. Residences situated furthest away from the source of treatment or at dead-end points in the distribution system are likely to have much higher (2 to 3 times higher) DBP levels than those measured at the treatment plant (Rossman et al., 2001). In addition, biologic and hydrodynamic conditions in piping can vary DBP formation spatially within a single plant distribution system (Rossman et al., 2001). Distribution system conditions

also create variability in DBP concentrations from one plant to another (Golfinopoulos and Arhonditsis, 2002).

Spatial variability can lead to misclassification if persons living at different locations within a distribution system are exposed to significantly different levels of DBPs but are assigned the same exposure level. This occurs frequently in epidemiologic studies when exposure assessment is based upon DBP levels measured at the plant. Since systems that rely on free chlorine in the distribution system are characterized by continuous, non-uniform DBP occurrence, concentrations of by-products measured at the water treatment plant may not accurately reflect those that exist throughout the distribution system (Chen and Weisel, 1998). In addition, plants are required to sample at only a small number of locations in the distribution system. Due to concern over regulatory compliance, oversampling or undersampling may be performed in areas of historical DBP excess (Waller et al., 2001).

Methods used to account for spatial variability have been introduced into recent studies of DBPs and birth outcomes. Initially, exposure assessments were based on utility data without accounting for the variability within the distribution system (Bove et al., 1995; Savitz et al., 1995; Dodds et al., 1998; King et al., 2000; Dodds and King, 2001). Klotz and Pyrch (1999) used residential tap water sampling to validate exposure for each subject. Gallagher et al. (1998) incorporated a hydraulic model of the distribution system into a geographic information system, and assigned exposures to individual census block groups. While this method shows promise for prediction of DBP concentrations

throughout the system, extensive validation of the models for each class of byproduct is necessary. Most recently, Waller et al. (2001) used a weighting procedure to reduce the influence on risk estimates of individuals with less accurate exposure values. Weightings were contingent upon degree of spatial variability in individual distribution systems and were assigned to each subject by residential location.

In order to reduce misclassification in epidemiologic studies of DBPs, selection of spatially consistent distribution systems could be undertaken. However, at present, there are no known criteria that can be used to classify water distribution systems as having low spatial variability.

In this study, we report on the occurrence of DBPs (THMs and HAAs) in 198 public water distribution systems reported in the U.S.EPA's Information Collection Rule (ICR) for the period July 1997 through December 1998. We evaluate two methods for selecting study populations that are supplied by distribution systems with spatially consistent THM levels. We also examine the occurrence of HAAs and brominated compounds within the selected systems.

3.2 METHODS

3.2.1 Data

Data collected under the Information Collection Rule (ICR), a utility database system designed by the U.S. Environmental Protection Agency (EPA), were used to supply information on water parameters related to pathogen occurrence and DBP formation

(U.S. EPA, 1998). The data were collected in quarterly sampling events from July 1997 to December 1998 at 198 public water distribution systems across the U.S. that served at least 100,000 people. We limited our study to those sites with complete reporting, i.e. sites that had collected total trihalomethane (TTHM) samples for at least four consecutive quarters and at four different sampling points in the distribution system. Only 52 sites made TTHM data available for all six consecutive quarters and all four sampling points; 32 additional sites reported complete data for one full year (October 1997 through September 1998). Data extraction from the ICR for each site included the total trihalomethane (TTHM) class and the sum of five haloacetic acids (HAA5) and individual trihalomethane species (chloroform, bromodichloromethane, dibromochloromethane and bromoform).

Samples were collected from the distribution system at the locations defined by the ICR guidelines: (1) samples from two locations of the "average" residence time of water in the distribution system (Avg1 and Avg2); (2) a sample from the point farthest along in the distribution system or at a dead-end representing the maximum residence time of water in the distribution system (MAX); and (3) a sample, termed the distribution system equivalent (DSE), taken at a location with a known retention time within the distribution system (USEPA, 1996). Collection of samples for each quarterly "sampling event" was performed during a single 24-hour window (U.S.EPA, 1996).

3.2.2 Methods used to Evaluate Occurrence

Descriptive statistics were obtained for individual and total THMs by sampling point and quarter for the 84 ICR sites with complete reporting. These statistics describe the spatial

and seasonal variability by THM for this broad selection of U.S. water treatment plant distribution systems.

Risks associated with human exposure to brominated byproducts have been assessed in five studies (Kramer et al., 1992; Waller et al., 1998; Klotz and Pyrch, 1999; King et al., 2000; Dodds and King, 2001) and future epidemiology studies are likely to continue to investigate populations residing in areas that have a high proportion of brominated byproducts. Therefore, we also attempted to identify sites with a high proportion of brominated THMs and HAAs. Three criteria were used to classify a site as having predominantly brominated DBPs after being identified as having low spatial variability. First, the site was considered to be predominantly brominated if the proportion of brominated DBP species (BDCM, DBCM and bromoform) exceeded 50% at all sampling points during all four quarters. Second, the sites' DBPs were considered to be predominantly brominated if the mean concentration of brominated species exceeded 50 μg/l. Third, we evaluated bromine predominance using the the Bromine Incorporation Factor (BIF),

$$\frac{0\times[CHCl_3]+1\times[CHBrCl_3]+2\times[CHBr_2Cl]+3\times[CHBr_3]}{[CHCl_3]+[CHBrCl_3]+[CHBr_2Cl]+[CHBr_3]},$$ which describes the molar contribution of all brominated species. Therefore, the higher the BIF, the greater the speciation shifts in favor of more brominated compounds (Nuckols et. al., 2001).

3.2.3 Tests of Spatial Variability

Two methods were initially used to define low intra-system variability of TTHM levels at each site: (1) a method based on two-way analysis of variance; and (2) a method based on the TTHM cutpoints reported in the literature for epidemiologic studies of birth outcomes. We compared these methods with respect to the number of sites deemed eligible for study and limiting factors, which were based upon the validity of using these methods.

We identified the single method of identifying sites with low spatial variability that was most appropriate, considering ease of use, validity and applicability to other sites, and used it to perform the final site selection. To demonstrate the applicability of this method, we selected sites for three profiles: 1: high trihalomethanes, high proportion of brominated DBPs; 2: high trihalomethanes, low proportion of brominated DBPs; and 3: low trihalomethanes, low proportion of brominated DBPs. To ensure that sites selected using either method would be expected to have similar exposure profiles in subsequent years, we also compared DBP values for the same quarters in successive years using the sites with complete reporting for 6 consecutive quarters.

3.2.3.1 Two-way ANOVA

We performed two-way analysis of variances (ANOVAs) on the complete randomized block designs our data presented. Our goal was to compare individual DBP levels between the four sampling points in the presence of an extraneous variable, season; assuming that each quarter reflects a season. Samples had been taken at each sampling point during each quarter, a procedure called matching (by quarter) or blocking. The

two-way ANOVA allowed the block (i.e. seasonal variation) to be independently assessed and removed from the error term. By applying this method, we were able to evaluate spatial variability between sampling points independently from the temporal variability between quarters. When seasonal differences are present, blocking not only allows for the independent evaluation of spatial variability, but also improves the power of the significance (F) test. The assumptions for our analysis were: (i) the 4 (sample points) x 4 (quarters) =16 observations that are independent random samples of size 1 from 16 populations; (ii) the 16 populations are normally distributed and have equal variances σ^2, which implies that the error components are normally distributed with mean 0 and variance σ^2; and (iii) there is no interaction between space (sampling points) and time (season).

3.2.3.2 Epidemiologic Literature

Epidemiologic studies investigating the effects of TTHMs have consistently suggested dose-response effects for certain reproductive outcomes (Bove et al., 1995; Waller et al., 1998; Dodds et al., 1999; Klotz and Pyrch, 1999; King et al., 2000; Waller et al., 2001), with increased risk at TTHM levels above 80 μg/l, whereas TTHM levels below 40 μg/l appear not to be associated with adverse effects. Based on this information, sites were selected as follows: if, in a given quarter, all four sample values fell within the same exposure group, low (<40 μg/l), medium (40-80 μg/l) or high (> 80 μg/l), and if this occurred for all four quarters, then the site was eligible.

3.3 RESULTS

3.3.1 Overview of ICR Data

The descriptive statistics (mean, standard deviation, median, and range) by THM for the 84 sites with complete reporting for four quarters are presented in Table 3.6.1 for individual and total THMs by sampling point. Descriptive statistics by quarter are shown in Table 3.6.2. Variability between the four sampling points tended to be small; the high upper limit of the range of TTHMs at the DSE sampling point (Table 3.6.1) was due to only one site. Stratification by quarter revealed seasonal fluctuation in THM concentrations with highest levels from July to September (mean TTHM concentration ≅ 60 μg/l) and lowest levels from January to March (mean TTHM concentration = 33 μg/l). Chloroform comprised the largest component of TTHMs.

3.4 COMPARING TESTS OF SPATIAL VARIABILITY

Table 3.6.3 presents the number of sites eligible for study by method of selection. In addition, we list the initial and secondary criteria for the ANOVA and epidemiologic literature methods.

3.4.1.1 Two-way ANOVA

Twenty-five sites remained eligible for study after applying the primary selection criteria of non-significant p-values ($p>0.05$) for all DBPs by two-way ANOVA. We evaluated the variability between sampling points independently from the variability between quarters, both calculated relative to the random variability. This enabled us to evaluate relative rather than absolute changes and allowed us to utilize all four sampling points. However, although no arbitrary TTHM level was established as a cutpoint, an arbitrary

alpha level for site inclusion was chosen ($\alpha = 0.05$). Further, we did not take exposure cutpoints suggested in the epidemiological literature into consideration in this approach. As an example, consider the following hypothetical scenario. The four TTHM levels measured during one quarter in the distribution system of plant A are 150, 170, 180, and 200 (variance = 433). In the distribution system of plant B, levels 70, 70, 90, and 90 were measured (variance=133). Although the variance in plant B is much lower, plant A is a better choice since all four values exceed 80 μg/l, indicating high exposure.

The model assumptions for using a two-way ANOVA were also evaluated. Since samples had been taken randomly during each quarter at each sampling point, independence (assumption i) was satisfied. Using the Shapiro-Wilk test for normality, assumption (ii) was also satisfied for all sites ($p>0.05$). However, for assumption (iii), Tukey's test for non-additivity suggested the presence of an interaction between temporal and spatial variability for eight sites. Therefore, eight sites violated the assumptions for ANOVA analyses and were excluded from the pool of eligible sites.

3.4.1.2 Epidemiologic Literature

As described above, a potential site was defined as high exposure when TTHM levels were greater than 80 μg/l for 4 quarters and low when TTHM levels were less 40 μg/l for 4 quarters. We identified eighteen sites based on these cutpoints, 5 with high exposure and 13 with low exposure. As shown in Table 3.6.4 these designations remained consistent over two successive years of sampling.

The second level of criteria for this epidemiologic literature method increased the number of sites eligible for study. This criterion was based upon the knowledge that exposure periods for adverse birth outcomes are approximately 9 months (or in some cases of birth defects, three months). Therefore, a 9-month period of consistently high TTHM levels would be sufficient to conduct an epidemiologic study in which high exposures are assigned to every woman served by the utility that is pregnant during those high exposure months. By adding sites with three consecutive quarters of data, two additional sites were identified which reported with TTHM values above 80 μg/l. These additional sites were added to the original 18 identified through primary cutpoint criteria and are reflected in the secondary criteria column of Table 3.6.3. Using both sets of criteria for this epidemiologic method, we ultimately identified twenty total sites, 7 with high exposure and 13 with low exposure

Additionally, using this test of spatial variability, we found that dose characterization of HAAs (corresponding to the U.S. regulatory limit of 60 μg/l) was generally consistent with TTHM categorization. For those sites with high TTHM levels, 5 out of 7 also had high average (mean HAA concentration > 60 μg/l) levels of HAAs and two had moderate levels (mean HAA concentration ≅ 30 μg/l), of which, one only reported these by-products during two (non-summer) quarters. With the exception of two sites having moderate HAA levels (NY506 and WA686), sites with TTHM levels less than 40 μg/l also had consistently low HAAs (mean HAA concentration < 20 μg/l).

3.4.2 Final Site Selection

Table 3.6.5 exhibits the three sites selected by both methods of evaluation. These sites were located in Arizona, Indiana and North Carolina and all were considered to have characteristically medium to high DBP levels, as defined by the epidemiologic method of selection.

Overall, we found little overlap between selection methods for number and characteristics of sites. After considering the primary and secondary criteria (Table 3.6.3) the method based upon epidemiologic literature was determined to be superior, as the two-way ANOVA method presented with more rigid limiting factors and fewer eligible sites.

3.4.2.1 Brominated Classification

Among the thirteen distribution systems with TTHM levels less than 40 μg/l, five reported approximately 50% or more brominated compounds and one plant had TTHM values too low to calculate meaningful proportions. The remaining sites also had low means and BIFs and thus exhibited characteristics consistent with profile 3 (low TTHM levels, low proportion of brominated DBPs). In Table 3.6.6 the proportion of brominated compounds, the mean concentration of the three brominated species and the Bromine Incorporation Factor (BIF) are presented for the seven sites with TTHM levels greater than 40 μg/l. Among these sites, only one reported high proportions of brominated compounds (mean>50%, proportion>50% and BIF>1.5) and thus was the only site that met exposure profile 1. The six other sites from this group had TTHM>40 and low proportions of brominated DBPs, fitting the exposure profile of low brominated, high overall TTHM exposure (profile 2).

3.4.2.2 Classification by Exposure Profile

The number of sites eligible for each exposure profile after applying our rigorous site-selection method is presented in Table 3.6.7. As shown, more sites were deemed eligible for study when the criteria for complete reporting required four quarters rather than six. Using the four-quarter criteria, the ICR database provided us with one site displaying profile 1 (high trihalomethanes, high proportion of brominated DBPs), six sites with profile 2 (high trihalomethanes, low proportion of brominated DBPs), and 7 sites displaying exposure profile 3 (consistently low TTHM levels and a low proportion of brominated compounds).

3.5 DISCUSSION

In this study, we describe the occurrence of DBPs (THMs and HAAs) in 84 public water distribution systems having complete reporting within the USEPA's Information Collection Rule (ICR) database for four quarters of time and four sampling points. As expected, we found that individual and total THM levels observed were consistent both overall and seasonally with those reported elsewhere in the U.S.

We evaluated two methods for *a priori* site selection of populations served by water distribution systems characterized as having spatially consistent disinfection by-product concentrations and examined the occurrence of HAAs and brominated compounds within the selected systems. For the *Two-way ANOVA* method, our goal was to compare individual DBP levels between the four sampling points controlling for the seasonal influence. Using this method, we found that statistical results were questionable. For example, this ANOVA method utilized arbitrary significance levels to ascertain low

spatial variability and presented with more rigid limiting factors. Specifically, 8 sites could not be considered part of the selection pool because the assumption of no interaction was not satisfied.

The second method used to select sites, which based selection upon cutpoints used in epidemiological literature, identified a moderate number of eligible sites and was only limited by degree of seasonal influence due to temporal variability. This limitation could be easily minimized through selection of sites with low temporal variability, leaving us with twenty eligible sites (out of 84) with complete reporting for at least three consecutive quarters. However, because this method is contingent upon previous epidemiologic findings, the cutpoints may be subject to change as future studies identify more meaningful threshold values. Nonetheless, we recommend that future researchers of DBPs employ this epidemiologic literature method presently, using TTHM and HAA data, and further determine whether the proportions of individual species remain fairly consistent throughout time. This is especially important as specific cutpoints for individual species are yet to be designated.

In using the test of spatial variability based upon epidemiologic literature, we found that the level of variability for TTHMs was generally consistent with the level of variability among the sum of 5 HAAs. This information is consistent with previous literature (Villanueva et al., 2003) and is useful for identification of populations receiving water from distribution systems with spatially consistent HAA levels.

Using the epidemiologic literature method as the preferred technique of *a priori* selection of low spatial variability sites, we selected sites for three profiles: 1: high trihalomethanes, high proportion of brominated DBPs; 2: high trihalomethanes, low proportion of brominated DBPs; and 3: low trihalomethanes, low proportion of brominated DBPs. However, we found it difficult to locate sites that could be characterized by profile 1. Whether this disadvantage stemmed from the limitations of our criteria or the rarity of this profile, we are unsure.

Most previous epidemiologic studies of reproductive outcomes have relied on total THMs as the relevant exposure metric, while recent studies suggest that the composition of the mixture and the concentration of specific DBPs, especially brominated DBPs, may be of critical importance. While we are not recommending a single method for use in identification of sites with a predominance of brominated DBPs in this paper, we have presented methods that may be used. Methods for identification of brominated sites provided consistent results, although it must be noted that our selection of cutpoints (e.g. mean concentration> 50 μg/l) was not based upon biological reasoning. Classification of a site as having predominantly brominated DBPs may also depend upon temporal variability. For example, a water treatment plant may rely on multiple water sources throughout a year, depending upon population demand and season of use, leading to significant seasonal differences in levels of brominated compounds with no change in overall DBP levels. This practice, however, can be mitigated during analysis by in-depth evaluation of data across all seasons. We note that this classification of sites was not

based upon individual HAA data because this information was not generally available in the ICR database.

Improvements in exposure assessment can be difficult to achieve (Reif et al., 1996). For example, performing comprehensive water sampling for multiple DBP chemicals can be time and cost intensive. Additional exposure assessment methods, such as distribution system modeling and biomarker monitoring are yet to be developed and/or validated. Our simple selection methods will aid future epidemiology studies utilizing prospectively or retrospectively collected data on DBPs by helping to identify appropriate sites for study before undertaking the study. Ultimately, techniques such as these will improve both the quality of data and our understanding of the true risks associated with exposure to DBPs.

3.6 TABLES

Table 3.6.1. Descriptive statistics for individual and total THMs (in µg/l) by sampling point among 84 sites with complete reporting for six quarters.

Sampling Point	Chemical	Mean	Standard Deviation	Median	Range
DSE:	CHCl3	30.6	23.0	24.0	*ND – 140.0
	BDCM	10.3	8.0	8.2	*ND – 50.3
	DBCM	4.2	5.6	1.9	*ND – 36.0
	CHBr3	0.6	1.4	ND	*ND – 9.8
	TTHM	45.6	30.1	35.65	1.1 – 193.0
AVG1:	CHCl3	32.0	21.5	27.25	*ND – 140.0
	BDCM	11.0	8.3	8.6	*ND – 48.0
	DBCM	4.5	6.4	1.9	*ND – 40.0
	CHBr3	0.5	1.4	ND	*ND – 9.7
	TTHM	48.0	28.8	39.9	3.2 – 188.0
AVG2:	CHCl3	33.8	24.4	28.0	2.0 – 160.0
	BDCM	10.9	8.1	8.7	*ND – 50.0
	DBCM	4.2	5.8	1.9	*ND – 36.0
	CHBr3	0.6	1.4	ND	*ND – 11.0
	TTHM	49.5	30.8	41.65	2 – 222.0
MAX:	CHCl3	37.6	24.8	32.0	1.9 – 140.0
	BDCM	11.6	8.3	8.95	*ND – 44.0
	DBCM	4.6	6.0	2.3	*ND – 31.0
	CHBr3	0.7	1.6	ND	*ND – 14.0
	TTHM	54.5	31.2	46.35	2.1 – 204.7

ND=Non-Detect

* Range contains a non-detect value

Table 3.6.2. Descriptive statistics for individual and total THMs (in μg/l) by quarter among 84 sites with complete reporting for four sampling points.

Time Period	Chemical	Mean	Standard Deviation	Median	Range
Quarter 1 (July – Sept 97):	CHCl3	39.75	23.03	35.00	*ND – 120.0
	BDCM	13.39	9.49	11.00	1.4 – 48.0
	DBCM	5.93	7.84	2.70	*ND – 40.0
	CHBr3	0.89	2.00	ND	*ND – 14.0
	TTHM	59.91	30.35	54.10	4.9 – 149.8
Quarter 2 (Oct – Dec 97):	CHCl3	28.22	17.28	24.00	1.9 – 104
	BDCM	10.70	7.59	8.65	*ND – 44.0
	DBCM	4.81	7.11	1.90	*ND – 37.0
	CHBr3	0.81	1.88	ND	*ND – 11.0
	TTHM	44.48	24.83	38.70	2.0 – 130.8
Quarter 3 (Jan – Mar 98):	CHCl3	21.23	12.78	19.00	1.9 – 97.9
	BDCM	8.18	6.08	6.85	*ND – 33.0
	DBCM	3.51	5.22	1.40	*ND – 24
	CHBr3	0.55	1.36	ND	*ND – 9.8
	TTHM	33.40	18.02	30.20	5.9 – 106.1
Quarter 4 (Apr – Jun 98):	CHCl3	37.14	25.23	31.15	*ND - 130.0
	BDCM	9.20	6.03	8.15	*ND – 34.0
	DBCM	2.58	3.45	1.60	*ND – 19.0
	CHBr3	0.24	0.80	ND	*ND – 5.9
	TTHM	49.09	29.31	42.60	1.1 – 136.4
Quarter 5 (Jul – Sept 98):	CHCl3	45.74	31.40	39.00	*ND – 160.0
	BDCM	12.52	9.08	9.90	*ND – 50.0
	DBCM	4.16	5.13	2.05	*ND – 22.0
	CHBr3	0.37	0.84	ND	*ND – 4.7
	TTHM	62.73	38.13	55.55	2.1 – 222.0
Quarter 6 (Oct – Dec 98):	CHCl3	28.86	18.37	26.00	1.9 – 95.1
	BDCM	11.00	9.02	8.25	*ND – 50.3
	DBCM	4.67	5.41	2.35	*ND – 27.0
	CHBr3	0.69	1.11	ND	*ND – 5.2
	TTHM	45.18	28.37	37.20	4.4 – 163.2

* Range contains a non-detect value

Table 3.6.3. Comparison of methods by criteria, number of sites deemed eligible (out of 84) and limiting factors.

Method	Basis	Primary Criteria*	Secondary Criteria*	Limiting Factors	Distribution Systems (state, id#)
1	Two-way ANOVA	P-value<0.05 for all DBPs N = 25	Presence of interaction N = 17	1.Arbitrary alpha level 2.Model assumptions (interactions)	N=17 **AZ126**,CA163,CO251,CO259,GA327, A333,IA340, **IN379**,MA401,MI413,MI423,**NC451**,NE453,NJ468, NY508,PA567,WV678
2	Epidemiologic Cutpoints	4 consecutive quarters N = 18	3 consecutive quarters N = 20	1.Temporal Variability	N=20 **AZ126**,CA139,CA206,CA209,CA214,CO259,ID343,IL355,**IN379**,**NC451**,MI413,MI415,MI416,NJ473,NY506, OH535,OK542,OR547, WA686,WV698

*Where N= number of sites eligible for study.

Table 3.6.4. Comparison of descriptive statistics for TTHM between the same quarters in consecutive years among the 52 sites with complete reporting for six quarters.

	Quarter 1 (Jul – Sep 97):	Quarter 5 (Jul – Sep 98):
Mean	59.91	62.73
Std. Dev	30.35	38.13
Median	54.10	55.55
Range	4.90 – 149.80	2.10 – 222.00
	Quarter 2 (Oct – Dec 97):	Quarter 6 (Oct – Dec 98):
Mean	44.48	45.18
Std. Dev	24.83	28.37
Median	38.70	37.20
Range	2.00 – 130.80	4.4 0– 163.20

Table 3.6.5. Sites common both methods of selection.

Methods	Distribution Systems (by state and id#)
1 and 2	AZ126, IN379, NC451

Table 3.6.6. Proportion of brominated compounds, the mean concentration of the three brominated species and the Bromine Incorporation Factor (BIF) are presented for the 7 remaining sites with TTHM levels greater than 40 μg/l.

Plant	Mean brominated	% Brominated	AVE BIF
AZ126	51.27	53.98	1.54
CA139	23.88	19.95	0.97
IN379	29.06	35.73	1.28
NC451	12.21	15.80	1.06
OK542	30.99	37.12	1.30
NJ473	17.53	21.77	1.13
WV698	13.61	16.37	1.12

Table 3.6.7. Number of eligible sites having one of the three exposure profiles for six consecutive quarters and/or one year (four quarters).

Profile 1: High TTHM, high brominated		Profile 2: High/medium TTHM, low brominated		Profile 3: Low TTHM, low brominated	
6 quarters	4 quarters	6 quarters	4 quarters	6 quarters	4 quarters
	AZ126*	IN379	IN 379		CA206
			NC451		CA209
			NJ473	CO259	CO259
			OK542	ID343	ID343
			WV698		NY506
			CA139 *	OR547	OR547
					WA686

.*= includes sites with data available for three consecutive quarters

3.7 REFERENCES

Bove FJ, Fulcomer MC, Klotz JB, Esmart J, Dufficy EM, Savrin JE. Public water contamination and birth outcomes. Amer J Epidemiol 1995:141:850-862.

Bove F, Shim Y, Zeltz P. Drinking water contaminants and adverse pregnancy outcomes: A review. Environmental Health Perspectives 2000:110:61-74.

Chen WJ, Weisel CP. Concentration changes of halogenated disinfection by-products within a distribution system. Am Water Works Assoc J 1998:90:151-163.

Dodds L, King W, Wolcott C, Pole J. Trihalomethanes in public water supplies and adverse birth outcomes. Epidemiology 1999:443-446.

Dodds L, King W. Relation between trihalomethane compounds and birth defects. Occup Environ Med 2001:58:233-237.

Gallagher MD, Nuckols JR, Stallones L, Savitz DA. Exposure to trihalomethanes and adverse pregnancy outcomes. Epidemiology 1998: 9:484-489.

Golfinopoulos SK, Arhonditsis GB. Multiple regression models: A methodology for evaluating trihalomethane concentrations in drinking water from raw water characteristics. Chemosphere. 2002: 47:1007-1018.

Hertz-Piccioto I. Environmental Epidemiology. In: Rothman KJ, Greenland S. eds. Modern Epidemiology, 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 1998.

King W, Dodds L, Allen A. Relation between stillbirth and specific chlorination by-products in public water supplies. Environ Health Perspect. 2000:108: 883-886.

Klotz JB, Pyrch LA. Neural tube defects and drinking water disinfection by-products. Epidemiology 1999:10:383-390.

Klotz JB, Pyrch LA, Trimbath L, Haltmeier P, Weisel C, Kim H. A case-control study of neural tube defects and drinking water contaminants. Atlanta: Agency for Toxic Substances and Disease Registry, Public Health Service. 1998.

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

Nieuwenhuijsen MJ. Toledano MB. Eaton NE. Fawell J. Elliott P. Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: a review. [Review] Occupational & Environmental Medicine. 2000:57(2):73-85.

Nuckols JR. Rossman L. Singer P. Speight V. Krapfl H. Miles A. Wilkes Cent of methods for predicting trihalomethane and haloacetic acid concentrations in exposure assessment studies. Final Report, Project 341. American Water Works Research Foundation. Denver. Colorado. 2001.

Pegram RA, Andersen ME, Warren SH, Ross TM, Claxton LD. Glutathione-S-Transferase-Mediated Mutagenicity of Trihalomethanes in Salmonella Typhimurium TA1535: Contrasting Results with Bromodichloro-methane and Chloroform. Toxicol. Appl. Pharmacol. 1997:144:183-188.

Reif JS, Hatch MC, Bracken M, Holmes L, Schwetz B, Singer PC. Reproductive and developmental effects of disinfection byproducts in drinking water. Environ Health Perspect 1996:104:1056-1061.

Rossman LA. Brown RA. Singer PC. Nuckols JR. DBP formation kinetics in a simulated distribution system. Water Research. 2001 Oct: 35(14):3483-9.

Savitz DA, Andrews KW, Pastore LM. Drinking water and pregnancy outcome in central North Carolina: Source amount and trihalomethane levels. Environ Health Perspect 1995:103:592-596.

U.S. Environmental Protection Agency. ICR Sampling Manual. April,1996: EPA 814-B-96-001.

U.S. Environmental Protection Agency. Part IV National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts; Final Rule. U.S. Federal Register December 16,1998:63:69390-69476.

Villanueva CM, Kogevinas M, Grimalt JO. Haloacetic acids and trihalomethanes in finished drinking waters from heterogeneous sources. Water Research 2003:37 (4): 953-958.

Waller K, Swan SH, DeLorenze G, Hopkins B. Trihalomethanes in drinking water and spontaneous abortion. Epidemiology 1998:9:134-140.

Waller K, Swan SH, Windham GC, Fenster L. Influence of exposure assessment methods on risk estimates in an epidemiologic study of total trihalomethane exposure and spontaneous abortion. Journal of Exposure Analysis & Environmental Epidemiology 2001:11:522-531.

CHAPTER 4

4 TEMPORAL VARIABILITY OF DISINFECTION BY-PRODUCTS AND TIME-RELATED BIRTH OUTCOMES

4.1 INTRODUCTION

The third trimester of pregnancy lasts from approximately the 26th week of gestation to parturition, with actual time length dependent upon individual pregnancy. This is the time period during human fetal development that is thought to be a period of sensitivity for several birth outcomes (Kline, 1989). These outcomes include intrauterine growth retardation (IUGR), term low birth weight, preterm birth and low birth weight. In epidemiologic investigations, the latter two outcomes are difficult to evaluate with various exposures because they are defined by or related to time-length of gestation. When analyzed, cases will usually have shorter gestation lengths than the typical comparison group (term births), increasing the potential for bias (Hertz-Picciotto et al., 1996).

Previous epidemiologic studies of disinfection by-products (DBPs) and preterm birth or low birth weight are characterized by large variability in analytical methods (Table 4.5.1). Quarterly measurements of DBP concentrations in water distribution systems have been the basis for determining exposure levels among subjects in previous investigations. For example, Bove et al. (1995) used quarterly values to extrapolate

monthly exposure estimates over each woman's entire pregnancy, while quarterly DBP values were used to approximate exposure to some time period specific to the third trimester in three other studies. Savitz et al. (1995) used the quarterly value nearest the 28th week of pregnancy, and Gallagher et al. (1998) used the median of all quarterly measurements taken during the third trimester. For children born in the 2nd or 3rd month of the quarter, Wright et al. (2003) used the average quarterly values for the third trimester; otherwise, the preceding quarterly averages were assigned. More recently, spline regression techniques including month and individual distribution system characteristics have been used to estimate more frequent (e.g. daily or monthly) DBP values (Dodds et al., 1999; King et al., 2000; Dodds and King, 2001). This method, which interpolates between sample measurements, uses a polynomial function to join adjacent data points with a smoothed line (Greenland, 1998) and enabled Dodds et al. (1999) to evaluate average exposure over the last three months of pregnancy. Other than these efforts, no study has attempted to evaluate exposure over the entire third trimester time period.

In this study, we demonstrate the potential bias arising from use of new analytical techniques when temporally variable environmental exposures, such as disinfection by-products (DBPs), are evaluated against time-dependent outcomes as a result of the high correlation of time with many adverse birth outcomes, including preterm birth (Hertz-Piccioto et al., 1996). Failure to adjust for variable length of third trimesters and use of these analytical methods can potentially lead to significant bias away from or towards the null and further complicate estimation of true risk in epidemiologic studies that likely

contain biases from other sources such as inadequate or improper methods to assess exposure. In this study, we use spline regression examples to demonstrate potential biases and to illustrate the magnitude of effect that these biases exert upon the relationship between temporally variable exposures and time dependent outcomes. Finally, we provide recommendations for avoiding or at least limiting the effect of these methodologic biases.

4.2 METHODS

4.2.1 Biases

4.2.1.1 Selection Bias

Epidemiologic sample selection allows investigators to model certain exposure or disease groups while facilitating the collection of complete covariate information (Hennekens and Buring, 1987). Using this general approach, studies of birth outcomes typically measure exposure within two discreet points in time, such that any births outside the given window are excluded (Kramer et al., 1992; Bove et al., 1995; Savitz et al., 1995; Gallagher et al., 1998; Dodds et al., 1999; Wright et al., 2003). This method can lead to biased selection of subjects in reproductive studies. For example, under this scheme, a full term birth occurring early in the study period would be excluded from participation due to exposure outside the study window, while a preterm birth occurring at the same time would be eligible on the basis that the shorter exposure window of interest fits within the designated study period (Hertz-Piccioto et al., 1996). Hence, the nature of exposure during this time period could dictate an increase or decrease in the effect estimate as compared to the null, as illustrated below.

Assume a study includes all births occurring on the 70th day of our investigation time period and that two participants are born on this day as shown in Figure 4.6.1 (where "day" represents individual days of births on the X-axis). In this example, participant "A" represents a full term birth whose third trimester exposure begins before the study time period (day 1); such that, exposure information is not available for part of the trimester. Participant "B" is a preterm birth with complete data for the third trimester. Using the inclusion criteria requiring exposure information for the entire third trimester, the full term birth would be ineligible due to incomplete data from early pregnancy whereas the preterm birth would qualify for inclusion in the study. This differential selection criterion would lead to an artificial increase in the association between preterm birth and average BDCM exposures greater than 18 μg/l.

4.2.1.2 Cutpoint Bias

If exposure varies over time and is estimated by averaging over the third trimester, which in turn can vary greatly in length, then short 3rd trimesters would more likely fall mostly or entirely into a shorter time period of high or low exposure relative to full length 3rd trimesters. Hence, exposure averages are likely to be more extreme (high or low) for preterm births. If the short time period is one of high exposure, preterm births will be artificially associated with high exposure and vice versa. The magnitude of the bias depends on the cutpoint selected to classify exposed and unexposed groups.

In figure 4.6.2, the cutpoint for average third trimester exposure is set at 10 μg/l , below which lie extremely low average exposure values over a relatively short period of time.

In this figure, the X-axis represents individual birthdays during the study time period. Based on the above information, more preterm births are likely to be categorized as unexposed (below 10 μg/l), resulting in odds ratios that are biased in a negative direction.

4.2.2 Sample Population

A hypothetical population of approximately 150,000 births was created for each scenario, with 26 to 42 week gestation periods proportionally assigned according to the U.S. incidence of births at each gestation week. Over a 365-day simulation period, the total population represented nearly 410 births per day. By having an equal distribution of preterm births we ensured no association (OR=1) between preterm birth (fixed) and DBP exposure (variable throughout the year). To illustrate parameters within realistic ranges, we based exposure on the bromodichloromethane (BDCM) species of DBP in four hypothetical communities. The example communities are variously characterized as having: (1) high temporal variability and high crest exposure; (2) high temporal variability and low trough exposure; (3) low temporal variability and high crest exposure, and; (4) low temporal variability and low trough exposure (Figures 4.6-1 to 4.6-4).

Quarterly averages for an individual DBP species were assigned to the sample year and daily values were extrapolated using a spline model. In this method, sampling averages for each quarter were used as join points of regression lines. Quarterly values for the community with low variability more closely resembled low-level exposures of bromodichloromethane (BDCM) documented in previous publications (Waller et al.,

1998; King et al., 2000) with a minimum value of 9 µg/l and a maximum of 18 µg/l before regression.

4.2.3 Sample Selection

Two methods were used to simulate sampling from the population. First, we performed simple random sampling (SRS) using SURVEYSELECT, a SAS procedure that allows use of random number seeds to approximate evenly distributed sample sets over the one-year time period. We also performed a stratified analysis where samples were selected by month according to the monthly frequency of U.S. births. This selection method was used to imitate a prevalence study in which all samples were proportionally representative of an actual U.S. investigation.

Replicated sampling is used during simulations to provide a simple method of variance approximation for risk estimates. In this study, the appropriate number of replicates for each simulation was determined based on sample sizes of either 5000 or 15,000. For smaller sample sizes, we processed 100 sample replicates from the original population. Replicate numbers for the larger sample size were limited to 30 or 50, depending upon the simulation technique and length of time required for models to converge. Sample size and replicate quantities were chosen to ensure good confidence interval coverage and to approximate actual sample sizes used in previous epidemiologic investigations of DBPs.

4.2.4 Exposure Assignment and Statistical Analyses

We approximated the average exposures for the entire third trimester as well as for the first month (28 days) of the third trimester for each birth using the mean of predicted daily DBP values over each time period. The start date for each third trimester was found by subtracting the length of the third trimester from the designated date of birth, where week 27 was the assumed beginning of a third trimester.

Exposures were dichotomized by cutpoint for DBP distributions after average exposures were tabulated (Figure 4.6.2). Cutpoints are commonly used in epidemiologic investigations when exposures do not follow any simple parametric distribution and to reduce the influence of misclassification (Waller et al., 1995). Because the cutpoint for the highest level of exposure for BDCM in Waller et al. (1998) was 18 µg/l, We designated a cutpoint of 18 µg/l for the community having high temporal variability to match the highest category estimated by Waller et al. (1998) based on actual BDCM values. We also chose cutpoints of 15 and 24 µg/l to illustrate degree and magnitude of exposures at low and high extreme values. Births with average third-trimester (or first 28 days of the third trimester) exposures greater than these cutpoint concentrations were considered exposed; all others comprised the referent group.

Estimated coefficients and odds ratios were calculated and compared to the true OR of 1.0 for each replicate via logistic regression. Replicate odds ratios were computed and averaged, and 95 percent confidence intervals based on Walds confidence limits were

generated. Also, beta coefficients were summed and averaged and the results exponentiated to compare against the average of the odds ratios.

The combinations of scenarios, given methodologic changes involving both high and low temporal variability, are listed below:

Low Trough in Concentration Levels
- A. Early in Study Period
  - 1. Averaged over the Entire 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling
  - 2. Averaged over the first 28 days of the 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling
- B. Middle of Study Period
  - 1. Averaged over the Entire 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling
  - 2. Averaged over the first 28 days of the 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling

High Crest in Concentration Levels
- A. Early in Study Period
  - 1. Averaged over the Entire 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling
  - 2. Averaged over the first 28 days of the 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling
- B. Middle of Study Period
  - 1. Averaged over the Entire 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling
  - 2. Averaged over the first 28 days of the 3rd Trimester
    - i. Population selected by simple random sampling
    - ii. Population selected by stratified monthly sampling

4.3 RESULTS

Results for the average of exponentiated beta coefficients and the average beta coefficients exponentiated (b-OR), which are both used to approximate an odds ratio (OR), were nearly identical, thus, only b-OR values are presented in the following tables. Additionally, we observed no difference in results by varying sample size, and only present results from simulations with sample size of 5000. Finally, confidence intervals

are not presented since use of large numbers of replicates practically ensured tight bandwidths.

Table 4.5.2 shows the influence of each analytical method on the relationship between preterm birth and average exposure to DBP concentrations of greater than 15, 18 and 24 μg/l in the high temporal variability community. Risk estimates were highest for crest exposures experienced early in the study period, particularly when the stratified sampling method was used, indicating some influence due to the aforementioned selection bias. In these same strata, moderately strong odds ratios were made stronger by increasing the cutpoint from 18 to 24 μg/l, as expected. Crest exposures during the middle of the study period were not affected by selection bias and were, therefore, more consistent with true risk (OR≈1.0). Average third trimester exposure risk estimates for crest exposures in the middle of the study time period were lower than those observed for the fixed exposure time (28 days) analyses and were more susceptible to changes caused by sampling technique as a result of cutpoint bias, where risk estimates for 18 μg/l were much lower than those observed for 24 μg/l. Odds ratios for crest exposures during the middle of the study time period for the fixed time (28 day) exposures were not affected by either the selection or cutpoint bias making these results most consistent with true risk (OR≈1.0).

A low steep trough early in the study time period produced highly protective risk estimates which increased with increasing cutpoint levels for both fixed (28 day) and variable (3rd trimester) exposure time periods, indicating that results were affected by both biases. In addition, odds ratios obtained using the SRS methods were higher than

for stratified results, although they remained lower than 1.0. The random sample design used by the SRS method was not subject to the same selection biases as the stratified sampling technique.

An exposure trough occurring later in the study time period more closely approximated the unbiased risk estimate. This is again due to removal of selection bias, other than remaining bias due to the stratified method of sampling. As with the crest exposures, presence of cutpoint bias for third trimester averages led to slightly increased or decreased risk estimates, while 28 day fixed exposures were again unaffected (OR≈1.0). Comparisons of an increasing crest in exposure versus a decreasing trough in exposure characterized by the community representing high temporal variability resulted in generally reciprocal ORs. Also note that the bias due to the stratified method of sampling was only slight when the primary sources of bias, selection and cutpoint, were removed.

The pattern of risk estimates for a community displaying low temporal variability (not shown) and shallow sloping was similar to that described for the community having high temporal variability and more severe sloping. Due to the subtler sloping, risk estimates were not as high or low for extreme cutpoints, regardless of sampling method, when evaluated against averaged time periods. As with the high temporal variability profile, the most consistent unbiased risk estimates were observed for exposure over the fixed time interval (first 28 days of the third trimester) with an exposure crest or trough in the middle of the study time period.

4.4 DISCUSSION

Previous epidemiologic investigations suggest no association between preterm birth and exposure to DBPs (Table 4.5.1). However, to date, no studies have attempted to estimate third trimester exposure over the entire time length, perhaps to avoid the risk of increased bias due to variability in length of exposure times given the likely presence of other biases resulting from exposure assessment discrepancies. Data from this investigation illustrate that analytical bias can be minimized in studies of third trimester time-related outcomes and temporally variable exposures when certain guidelines are followed:

First, investigators must remain conscious of the effect of exclusions due to incomplete data. Modeling exposures that occur early in the study time period and use of the stratified sampling method can lead to risk estimates that are generally higher (in the community having crests) or lower (in the community having troughs) than true risk. By excluding participants born before the start of the study time window, there can be misappropriation of cases and controls according to the selection bias identified earlier. Trough or crest exposures occurring in the middle to late days of the study time period, however, avoid the influence of selection bias at extreme values as the regression procedure can be extended to cover the weeks prior to the start of the study time window, allowing estimation of exposure for the entire third trimesters for those born full-term on day 1 of the study. In practice, this requires collection of at least one quarterly sample prior to the start date or the time we begin to consider subjects eligible for study.

Second, when using regression techniques, the effect of variable length exposures and cutpoints must be considered. In this study, we examined the magnitude and direction of bias introduced by selection of cutpoints used to dichotomize exposure. As expected, more extreme cutpoints led to more extreme results, suggesting that use of a temporally variable exposure time period such as the third trimester can lead to extreme case exposure values thereby creating artificially risky or protective ORs. The degree of variability will also influence the amount of time above or below a cutpoint, i.e. a model with high temporal variability will offer less chance for full term third trimester exposures below a cutpoint of 18 μg/l, because the exposure is characterized as having steep temporal slopes. In Table 4.5.2, this trend is apparent among the exposure crests and troughs. The low variability model has subtle sloping, allowing for generally more time above or below the cutpoint (Figure 4.6.2) and less extreme average exposures, making the effect of cutpoint placement less obvious among third trimester exposure comparisons. That the influence of cutpoint bias is minimized for analyses of middle to late exposures, and nonexistent for analyses of fixed time intervals (the first 28 days of the third trimester) is noteworthy. By performing an analysis in which all births have the same time period of exposure, as in 28 days of exposure, we removed the cutpoint bias effect. Thus, no birth had an increased chance of being above or below the dichotomized cutpoints and these risk estimates are no different from an OR of 1.0 where the sample size coverage probability for 95 percent confidence intervals containing OR=1 was 95 percent.

It is possible that all previous investigations of preterm birth and DBPs (Table 4.5.1) having exposure assigned by quarterly data were subject to the selection bias resulting from data exclusion prior to study commencement as this study demonstrates. The single study by Dodds et al. (1999) may have avoided this bias by using estimated values based on spline data when no sample data was available. This same Nova Scotia study also used the last three months of pregnancy, a fixed exposure time, to avoid bias from use of cutpoints. However, the relevance of the assigned time period is suspect, as it is unclear whether exposure for preterm cases was assigned over the 2nd trimester for comparison with third trimester exposure for controls.

A major goal of this paper was to identify exposure assignment methods that could bias analyses of third trimester exposures and to recommend alternative techniques. As we have shown, there are several concerns with averaging time-related exposures including selection and cutpoint biases. In epidemiology, cutpoints provide flexibility and robustness in order to circumvent measurement uncertainty (Waller et al., 1995). However, comparison of variable length exposures is likely to introduce bias that is further magnified by use of cutpoints. A better alternative is to compare outcomes by gestation age (e.g. intrauterine growth retardation) or to use equal length exposures. In this paper we found no cutpoint bias from use of the first 28 days of the third trimester. Yet, we had no reason to believe that this time period (or any other time period) was meaningful with respect to fetal or placental development. We are confident that, due to the time-specific definition of preterm birth, the last weeks of the third trimester cannot be biologically important to induction of this outcome. With this knowledge, we can

attempt another method, namely, to compare the first half of the 3rd trimester for term births and the entire 3rd trimester for preterms. Although this method was beyond the original goals of this study and therefore was not evaluated, it can be used to compare fairly equivalent but variable time-lengths. Use of this method would considerably decrease the cutpoint bias due to flexible exposure time periods and allow for a better approximation of risk for shorter time intervals. There are limitations of this method, however, including: (1) only being able to evaluate exposure and preterm birth relationships in the first 6 to 7 weeks of the third trimester; (2) requiring at least some third trimester exposure for all births included in the study; and (3) not having complete control for bias through use of variable averaged time periods.

4.5 TABLES

Table 4.5.1. Risk estimates with 95 percent CI and exposure method used to assign total trihalomethane (TTHM) levels over time in preterm birth and low birth weight studies.

Study	Odds Ratio and 95 % CI at Highest Level of Exposure		Exposure Assignment Method	THM exposure group (µg/l)
	Preterm Birth	Low Birth Weight[1]		
Kramer et al., 1992	1.1 (0.7 –1.6)	1.3 (0.8 – 2.2)	Maternal residence at birth	>10 vs. Non-detect CHCl3
Bove et al., 1995	1.0 (0.9 – 1.1)	-	Monthly averages estimated and averaged over entire pregnancy	>80 - ≤20 TTHM
Savitz et al., 1995	0.9 (0.6 – 1.5)	1.3 (0.8 – 2.1)	Quarterly average nearest to 28th week of pregnancy	82.8 – 168.8 vs. 40.8 – 63.3 TTHM
Gallagher et al., 1998	1.0 (0.3 – 2.8)	2.1 (1.0 – 4.8)	Median concentration of TTHM samples during 3rd Trimester	≥61 vs. ≤ 20 TTHM
Dodds et al., 1999	0.97 (0.87 – 1.09)	1.04 (0.92 – 1.18)	Average TTHM exposure during the last 3 months of pregnancy	≥100 vs. <50 TTHM
Wright et al., 2003	0.97 (0.85 – 1.11)	-	Quarterly average nearest to specific month of 3rd trimester	≥80 vs. ≤60 TTHM

[1]Does not include Term Low Birth Weight Results

Table 4.5.2. Odds ratios according to type of environmental exposure, timing of exposure, time period of averaged exposure and sampling techniques for a community with high temporal variability.

Type of Environmental Exposure	Timing of "Extreme" Exposure	Time Period Averaged	Sampling Technique	Cutpoint	OR (average of betas)
Crest	Early	3rd Trimester	Stratified by month[B]	15	-
				18	1.189
				24	2.008
			SRS[B]	15	-
				18	0.837
				24	1.32
		First 28 Days of Third Trimester	Stratified by month[s]	15	-
				18	1.802
				24	2.321
			SRS[s]	15	-
				18	1.180
				24	1.539
	Middle/Late	3rd Trimester	Stratified by month[c]	15	-
				18	0.716
				24	0.947
			SRS[c]	15	-
				18	0.726
				24	0.955
		First 28 Days of Third Trimester	Stratified by month	15	-
				18	1.004
				24	0.994
			SRS	15	-
				18	1.024
				24	1.000
Trough	Early	3rd Trimester	Stratified by month[B]	15	0.229
				18	0.327
				24	0.559
			SRS[B]	15	0.202
				18	0.709
				24	0.837
		First 28 Days of Third Trimester	Stratified by month[s]	15	0.203
				18	0.324
				24	0.564
			SRS[s]	10	0.234
				18	0.733
				24	0.856
	Middle/Late	3rd Trimester	Stratified by month[c]	15	0.863
				18	1.141
				24	1.060
			SRS[c]	15	0.885
				18	1.146
				24	1.062
		First 28 Days of Third Trimester	Stratified by month	15	0.978
				18	1.011
				24	1.000
			SRS	15	1.015
				18	1.000
				24	1.000

[s]Selection Bias Effect, [c]Cutpoint Bias Effect, [B]Selection Bias & Cutpoint Bias Effects

[α] Sample Size n=5000 presented.

4.6 FIGURES

Figure 4.6.1. Including only births occurring on day 70 that fall entirely within our study time period.

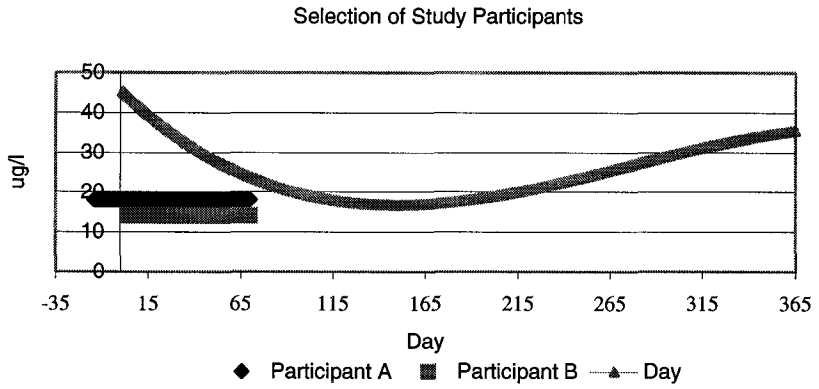
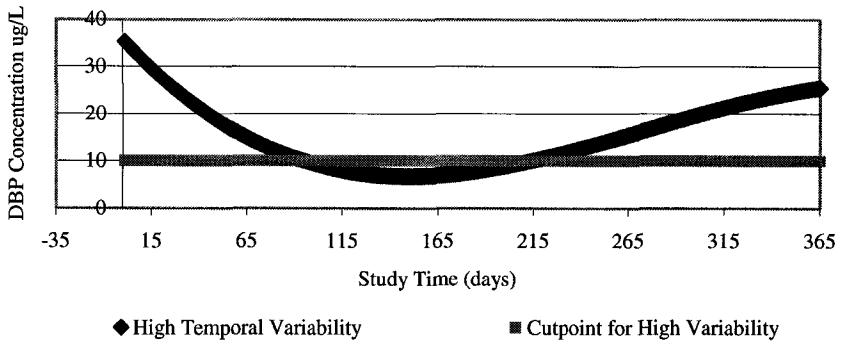
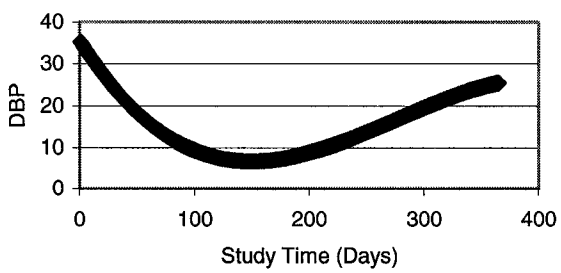


Figure 4.6.2. Daily DBP exposure levels and cutpoints for a hypothetical community.

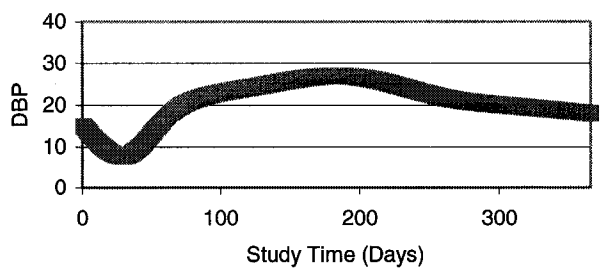


Figures 4.6-1 through 4.6-4. Four community profiles by exposure levels, degree of variability and "timing of extreme" exposure.

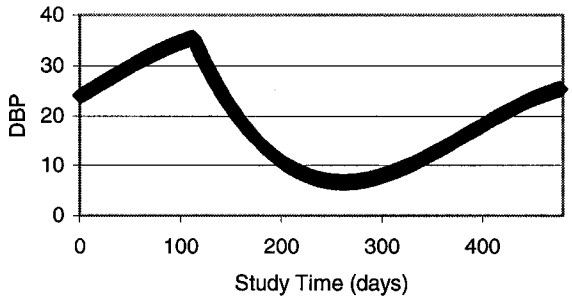
(4.7-1)



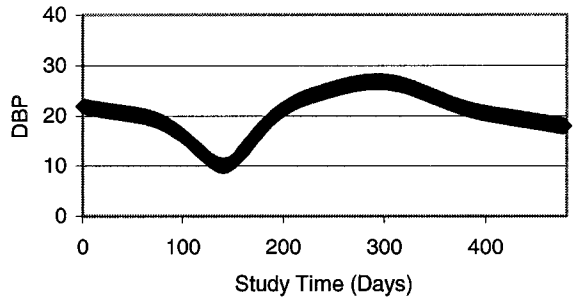
(4.7-2)



(4.7-3)



(4.7-4)



4.7 REFERENCES

Bove FJ, Fulcomer MC, Koltz JB, Esmart J, Dufficy EM, Zagraniski RT. Report on phase IV-A: Public drinking water contamination and birthweight fetal deaths, and birth defects, a cross-sectional study. New Jersey Dept. of Health, 1992.

Bove FJ, Fulcomer MC, Klotz JB, Esmart J, Dufficy EM, Savrin JE. Public water contamination and birth outcomes. Amer J Epidemiol. 1995:141;850-862.

Dodds L, King W, Wolcott C, Pole J. Trihalomethanes in public water supplies and adverse birth outcomes. Epidemiology 1999:443-446.

Dodds L, King W. Relation between trihalomethane compounds and birth defects. Occup Environ Med. 2001:58;233-237.

Gallagher MD, Nuckols JR, Stallones L, Savitz DA. Exposure to trihalomethanes and adverse pregnancy outcomes. Epidemiology 1998: 9:484-489.

Geenland S. Introduction to Regression Models. In: Rothman KJ, Greenland S. eds. Modern Epidemiology, 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 1998.

Hertz-Piccioto I., Pastore L., Beaumont JJ. Timing and patterns of exposures during pregnancy and their implications for study methods. Amer J Epidemiol 1996:143 (6);597-607.

Hertz-Piccioto I. Environmental Epidemiology. In: Rothman KJ, Greenland S. eds. Modern Epidemiology, 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 1998.

Hennekens C. Buring J. In: Mayrent S. ed. Epidemiology in Medicine. Little, Brown and Co. Boston, 1987.

Khoury M, Stewart W, Weinstein A, Panny S, Lindsay P, Eisenberg M. Residential mobility during pregnancy: implications for environmental teratogenesis. J Clin Epidemiol. 1988:41;15-20.

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

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

Nieuwenhuijsen MJ. Toledano MB. Eaton NE. Fawell J. Elliott P. Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: a review. Occupational & Environmental Medicine. 2000:57(2);73-85.

Reif JS, Hatch MC, Bracken M, Holmes L, Schwetz B, Singer PC. Reproductive and developmental effects of disinfection byproducts in drinking water. Environ Health Perspect. 1996:104;1056-1061.Savitz DA, Andrews KW, Pastore LM. Drinking water and pregnancy outcome in central North Carolina: Source amount and trihalomethane levels. Environ Health Perspect 1995:103:592-596.

Shaw GM, Malcoe LH. Residential mobility during pregnancy for mothers of infants with or without congenital cardiac anomalies: A reprint. Arch Environ Health. 1992:47;236-238.

Shimokura GH, Savitz DA, Symanski E. Assessment of water use for estimating exposure to tap water. Environ Health Perspect. 1998:106;55-59.

U.S. Environmental Protection Agency. Part IV National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts; Final Rule. U.S. Federal Register December 16,1998:63:69390-69476.

Wright JM. Schwartz J. Dockery DW. Effect of trihalomethane exposure on fetal development. Occupational & Environmental Medicine. 2003:60(3);173-80.

Zender R, Bachand AM, Reif JS. Exposure to tap water during pregnancy. Journal of Exposure Analysis & Environmental Epidemiology. 2001:11(3); 224-30.

CHAPTER 5

5 EARLY PREGNANCY EXPOSURES TO DBPS AND CLEFT PALATE

5.1 INTRODUCTION

Disinfection by-products (DBPs) are chemical contaminants in potable water resulting from current methods used for water treatment. Despite the benefits of water treatment, DBPs present a health concern due to their purported association with carcinogenic, reproductive, and teratogenic outcomes as described in previous epidemiologic studies (Reif et al., 1996; Nieuwenhuijsen et al., 2000). These compounds represent diverse chemicals that can be categorized into a number of potentially harmful groups. The two most common subgroups in the US include trihalomethanes (TTHMs) and five haloacetic acids (HAA5), which are currently regulated in drinking water at maximum contaminant levels (MCLs) of 80 μg/l and 60 μg/l, respectively. Total trihalomethanes (TTHMs) are the sum of four individual trihalomethane (THM) species, including chloroform, bromodichloromethane (BDCM), chlorodibromomethane (DBCM) and bromoform. The five most prevalent HAAs are monochloroacetic acid (MCAA), dichloroacetic acid (DCAA), trichloroacetic acid (TCAA), monobromoacetic acid (MBAA) and dibromoacetic acid (DBAA).

Cleft palate is a congenital malformation occurring in approximately one out of every 2,500 live births in the U.S. (O'Rahilly and Müller, 2001). This health outcome is

important because the cause, for most cases, is thought to stem from genetic and environmental stimuli (Moore and Persaud, 1998). Three prior epidemiological studies have evaluated exposure to DBPs in relation to the specific congenital malformation of cleft palate. The pattern of associationbetween THMs (both individual and total species) and cleft palate emerging from the studies has been inconsistent. Bove et al. (1995) found a threefold increase in risk for oral clefts defects with exposure to TTHM concentrations >100μg/l, as compared to a referent group of ≤20 μg/l. No trend, however, was apparent with increasing dose. Dodds and King (2001) conducted analyses of two individual species of THMs, chloroform and BDCM, but found no increased risk of oral cleft defects. Shaw et al. (2003) analyzed cleft lip and cleft palate cases separately and found no clear associations with TTHM exposures. The apparent discrepancies in the results could reflect inherent limitations of these past studies leading to biased results. For example, in studies by Bove et al. (1995) and Dodds and King (2001), grouping cleft lip and cleft palate into a general oral cleft defect (OCD) category may have produced etiologically dissimilar aggregations leading to a diluted risk estimate. In addition, most research has only evaluated exposure to TTHMs or, in the case of Dodds and King (2001), the two most common individual THMs.

Previous investigations have not evaluated exposure over critical time periods of embryonic development known for palatal formation and fusion or cleft palate induction. Cleft palate has several induction periods and is thought to initiate on the 14^{th} day of gestation and is again sensitive to exogenous insults in the later embryonic period (7^{th} – 9^{th} weeks) (O'Rahilly and Müller, 2001). However, despite knowledge of the specific

time periods of relevance, prior investigations evaluated exposure only during the first trimester (Bove et al., 1995), the first two months of gestation (Dodds and King, 2001) or the periconception period (Shaw et al., 2003). Risk estimates depend upon the magnitude of exposure over critical time windows and, accordingly, analyses over exposure windows that are too wide dilute risk estimates. In this case, where knowledge of the critical time period of cleft palate induction with respect to DBP exposure is somewhat uncertain, comparison of different exposure times can establish a window that provides a more precise risk estimate over a narrower, more relevant, period of time (Hertz-Piccioto et al., 1996). In this study, we examined the effect of exposure to four THMs and five HAAs over the first trimester, second month and 6^{th}, 7^{th} and 8^{th} weeks of gestation on the risk of cleft palate. Specifically, we evaluated whether time periods considered to be more specific and sensitive to cleft palate induction (e.g. 8^{th} week) produced more precise risk estimates.

5.2 SUBJECTS AND METHODS

We conducted a retrospective cohort study in a large Arizona community with three water treatment plants. The community was selected from the EPA's Information Collection Rule (ICR) database, because the distribution systems display large temporal fluctuations in TTHM levels. The study population included all live births and fetal deaths, identified through Arizona birth and fetal death records, from January 1998 through March 2003 (n=51,717) to women from 24 Arizona zip codes.

Case infants and fetuses with cleft palate (n=16) were identified through vital records. Unlike cardiac defects, diagnosis of cleft palate is more reliable in that it does not depend

on highly trained individuals or instruments. These isolated defects are identified at birth and documented on birth records. Information on date of last menstrual period was used to calculate specific weeks in gestation, corresponding to weeks when exogenous insults are known to induce oral cleft defects.

Exposure data for the years 1998 through 2002 were obtained from each facility. All sites measured individual THMs at least quarterly. With the exception of one treatment plant, HAA information was less consistently available and was only available for years 2000 and 2002. Supplemental biweekly TTHM and HAA5 data from 2002 were provided by one plant, while another provided monthly data for all THMs and HAAs. DBP concentrations were monitored at two to four locations within the distribution system waters of all three facilities. The quarterly and monthly data indicated presence of very low levels of bromoform, MBAA and MCAA; thus, these chemicals were not included in the statistical models. To estimate DBP values for monthly time periods (for years 1998-2001), we performed a spline regression for each water facility, similar to the procedures used by researchers in Nova Scotia (Dodds et al., 1999; King et al., 2000; Dodds and King, 2001). This technique was also utilized to estimate weekly DBP levels between 1998 and 2002 for use in analyses of specific exposure time windows.

Maternal residence at parturition was assumed to be the same as residence during the first trimester, the critical time period of oral cleft defect induction. Exposure estimates were obtained by assigning each maternal residence to designated geographic areas, defined as the service area of each water treatment plant. In cases where two plants shared the same

distribution system, knowledgeable treatment plant employees identified service boundaries without overlap. For each subject, exposure was averaged over all biologically appropriate time periods for palate development, including the first trimester, the second month of gestation and individual weeks of the second month (week 6, 7 and 8).

To evaluate DBP exposures, we used both continuous and categorized levels for each species of chemical according to the following methods: TTHMs and chloroform cutpoints were based upon categories used in previous DBP studies to enable comparison across studies. In situations where the low exposure category designated by the literature yielded few subjects and no cases, we combined the two lowest strata of exposure to establish a referent category. Categories for the remaining chemicals and the individual HAAs (TCAA and DCAA), which had never before been investigated in a large-scale epidemiologic study, were developed using rounded tertiles of species concentrations as calculated across the results of exposure from all three water treatment regions.

Information on potential confounders was abstracted from birth and fetal death records. These variables included maternal age, race, ethnicity, education, parity, month that prenatal care began, and, smoking; those significantly associated with cleft palate at the <0.20 level in univariate analyses were retained. By comparing the outcome of birth defects and different exposure time-windows within a single community, we expected to control for at least some of the many unknown and, thus, unobtainable confounding factors that may contribute to development of birth defects.

We used stratified chi-squares (exact method) and logistic regression analyses to determine if DBP exposures in the different months and weeks of early pregnancy increased the risk of cleft palate. We used logistic regression analysis to compare infants or fetal deaths with cleft palate to healthy births with the specific DBPs as the exposure variable of interest for each of the time windows. After adjustment for potential confounders, we obtained estimated odds ratios (ORs) and 95 percent confidence intervals (CIs) for the relationships between all individual THM species and the outcome of cleft palate.

5.3 RESULTS

The rate of cleft palate in these Arizona residents was slightly lower than national trends with approximately 0.78 per 2500 births. Table 5.5.1 summarizes characteristics of subjects and frequency of cleft palate. The majority of subjects were white, non-hispanic, nulliparous women, with some college education. Most mothers sought prenatal care in the first trimester and less than 10 percent smoked during pregnancy. Subjects were excluded for having no date for last menstrual period (LMP) or no estimated date of conception (EDC), which was used when data on LMP was missing or considered extreme (e.g. greater than 60 weeks before the birthday). Analyses were not different when using LMP, EDC or a combination of both methods; therefore, we used the combined method to minimize the number of subjects lost (n=36). Nine additional subjects were lost for having a date of LMP that was outside our range of dates for which we could calculate exposure. As well, we did not include fetal deaths (n=160) in

analyses because no cleft palate was observed among these subjects and record completeness, with respect to this outcome, was questionable.

Table 5.5.2 presents TTHM average exposures over the first trimester; individual months of the first trimester; and 6^{th}, 7^{th}, and 8^{th} weeks of pregnancy. As illustrated in the table, the majority of TTHM exposures occurred in the 40-59 μg/l category across all relevant time periods. The lower exposure categories were variable, as 23 percent of women had TTHM exposures less than 39 μg/l for the second month of pregnancy as compared to only 17 percent with this average level of exposure over the first trimester. Because the dilution effect of averaging was removed when looking at smaller time intervals, exposure values were generally higher with a fraction of exposure estimates over 80 μg/l for the individual weeks.

Table 5.5.3 presents ORs and 95 percent CIs for cleft palate according to individual DBP and time window of exposure. All ORs were adjusted for maternal race, which resulted in only slight changes to risk estimates. Approximately 9-10 percent of the subjects were lost for each analysis due to missing maternal race or exposure data. For TTHM, categorical analyses revealed exposures that were inversely associated with cleft palate, with significant protective risk estimates for all exposure windows except the second month. Under each exposure time window, ORs did increase toward the null with increasing levels of exposure. Only moderate levels of exposure were compared to the referent group (<40 μg/l), as no cases had exposure in the highest exposure level (>80 μg/l). When modeled as a continuous exposure, TTHM was not associated with induction

of cleft palate during any window of exposure. Results were similar when exposure to chloroform over varying time windows was evaluated against cleft palate. ORs were 0.3 for exposure to chloroform in the range 10 to 15 μg/l and 15 to 20 μg/l, when compared to the referent group of <10 μg/l, with the exception of one exposure window. Exposure to chloroform over the 8th week, the time period when sensitivity to cleft palate induction is greatest, yielded slightly higher risk estimates as compared to the other time windows. Continuous exposure to chloroform was not associated with cleft palate during any time window. We observed no increased risks of cleft palate with exposure to BDCM or DBCM during any time window. Risk estimates for tertiles of exposure ranged from 0.1 to 0.9, with the higher risks attributable to DBCM exposures. ORs for continuous exposures were typically not different from the null value, except for BDCM exposure during weeks 7 and 8, where the OR=0.8 (95% CI 0.65 – 0.99).

For the non-volatile HAAs, the ORs for the highest vs. the lowest tertiles of exposure were all less than 1.0. Analyses of the middle tertiles against the referent groups were less consistent and depended upon chemical and time window of exposure. Exposure to DBAA during the 7th week of exposure produced an OR of 1.1 (95% CI= 0.18-6.41). Risk estimates for all other times of exposure were less than 1.0, with CIs including 1.0. Results for average 2nd tertile exposure to TCAA (4-5.25 μg/l), as compared to the referent (<4 μg/l), were OR=1.2 for individual weeks and OR=2.1 for the second month, but confidence intervals were wide and all included the null value. The analyses of these variables as continuous variables revealed no increases in risk with increasing exposure during any time window.

5.4 DISCUSSION

Our results did not show strong evidence of an association between cleft palate and exposure to individual THMs and HAAs during specific critical time windows of gestation. However, timing of exposure did affect the results for analyses of HAAs at lower levels of exposure and for chloroform during the 8th week. For TCAA, there was a pattern of higher risk estimates for the 2nd tertile of exposure, although this was not observed across the trimester window of exposure; therefore, these observations are likely the result of small case numbers. Exposure to TCAA has been associated with induction of orofacial defects in animal studies, although at very high levels of exposure (Smith et al., 1989). The small increase in risk for chloroform during the 8th week of gestation, as compared to the entire first trimester, may reflect increased precision as a result of the slightly larger sample population retained for analysis of single weeks. Because the first trimester is a longer time period, it is more likely to fall outside of the study initiation and termination (or beginning and end) date than single weeklong periods; therefore, more trimesters are likely to be excluded.

Exposure data evaluated over different windows of gestation in this study indicate some variability across time. Over the time periods investigated, more subjects had higher levels of exposure when evaluated over individual weeks, as compared to the entire trimester. Similarly, proportionally more subjects fell into the lowest strata of exposure when evaluated by entire trimester than over any other time period. Thus, if exposure to individual DBPs had been associated with cleft palate in a dose-response pattern in this study, the stronger effect estimates would have been observed in the 6th, 7th and 8th weeks

of gestation. The effect of averaging a variable exposure over longer time periods, as in the case of the trimester, would have lead to misclassification over the critical time periods and diluted, lower, risk estimates (Hertz-Piccioto et al., 1996).

Windows of exposure have been historically important in epidemiologic investigations of thalidomide, retinoic acid (Vitamin A), maternal rubella, and radiation (O'Rahilly and Müller, 2001). For exposures during the first two weeks of gestation, few congenital abnormalities are observed because the teratogen either damages the majority of cells, resulting in cell and embryonic death, or affects only a few cells that can be repaired without resultant birth defects (Moore and Persaud, 1998). After the first two weeks, the tissue or organ that is most susceptible to malformation is the part undergoing critical development when the teratogen is active. Timing of insult and dose of teratogen will also determine the relative severity of a defect. For example, in utero exposure to retinoic acid during highly sensitive periods of development, the 3rd to 5th weeks of gestation, leads to a high incidence of spontaneous abortion and severe birth defects, including neuropsychological impairment. Conversely, exposures that occur later in gestation have a less drastic, if any, effect.

The biologic mechanisms to support an association between exposure to DBPs and cleft palate induction are not well understood. Toxicologic studies of congenital malformations more often support associations between exposure to DBPs, particularly HAAs, and heart malformations, although craniofacial defects have been observed after exposure to DCAA/TCAA in a study by Smith et al. (1989). The most consistent

reproductive effects resulting from exposure to high DBP doses has been reduced fetal weight and disrupted spermatogenesis and motility (Nieuwenhuijsen et al., 2001).

The literature presents an inconsistent pattern of association between DBPs and collective and specific oral cleft defects, which includes congenital cleft palate. As in this study, inadequate sample size has historically impeded statistical assessment of risk with this outcome. There is also potential for under-representation of cases resulting from elective termination of pregnancies, although this may be less likely with oral cleft defects as they respresent less serious conditions. In addition, with the higher exposure doses, multiple or more severe defects may have been expressed as stillbirth, spontaneous abortion or unrecognized fetal loss. Despite these weaknesses, the exposures and results presented in this paper are similar to two of the previous studies of DBPs and cleft defects that did not have these limitations (Dodds and King, 2001; Shaw et al., 2003). However, these studies only evaluated exposure over the first two months of pregnancy (Dodds and King, 2001) and periconception periods (Shaw et al., 2003), making direct comparison difficult. Our results did not agree with the increased risk estimates posed by the single study that evaluated first trimester exposure and TTHMs (Bove et al., 1995).

Available studies that have attempted to quantitatively evaluate DBP exposure and development of cleft palate are characterized by substantial variations in DBP exposure assessment. Bove et al. (1995) used a forward-backward averaging technique to estimate monthly exposures between quarterly DBP samples. Monthly exposure estimates were then averaged over each subject's first trimester in order to evaluate oral cleft defects.

Researchers in Nova Scotia (Dodds et al., 1999; Dodds and King, 2001) used a linear regression of quarterly samples to identify average exposures during the first two months of pregnancy for cleft defects. Shaw et al. (2003) relied upon water company personnel and quarterly monitoring data to estimate mean DBP levels for each residence lived in during the periconception period of all subjects. This estimate was applied to the period beginning either one or three months before through three months after conception for cleft palate alone and combined cleft palate/lip outcomes. Our method of assigning exposure estimated daily exposures for our study area through a spline regression technique of quarterly sampling values. Daily values were then averaged over each time window of interest. This method, which is similar to that used in the Nova Scotia studies, provides exposure estimates for intermittent time periods when data were missing or when sampling was not performed. Although the fit with our data was excellent (+/- 5%), the spline regression technique still requires validation against data from other distribution systems.

The strengths of this study include the large number of birth records, quantity of exposure data, intra-community study design and the ability to evaluate multiple time periods of exposure to several specific THMs and HAAs and to model exposure over more specific time intervals using the spline regression technique. By comparing subjects within the same community with respect to exposure levels, we also strived to control for multiple potential confounders. In environmental epidemiologic studies of this type, quantification of exposure can be difficult due to seasonal fluctuation of DBP chemicals, especially when exposure is often based upon quarterly data. As evidenced in this study,

investigation of continuous and threshold effects due to exposure to DBPs is possible in a single community where temporal variability is high, regression techniques are used, and the relevant time window is understood (Hertz-Piccioto et al., 1996).

The most significant limitation of this study is reduced power for statistical analyses, owing to the small numbers of cleft palate cases. Using our sample size for this cohort study, having approximately 17 percent (n=8,742) of the population unexposed during the first trimester, we had approximately 45 percent power to detect an odds ratio of 3.0. Sample sizes required for analyses with 80 percent power and odds ratios ranging from 1.5 to 4.0 are presented in Table 5.5.4. At 80 percent power, for an odds ratio of 3.0, we would have had to enroll 59,600 exposed and 15,025 unexposed subjects, approximately twice the sample size obtained for this study. To detect an odds ratio of 1.5 with 80 percent power for the association between TTHM exposure over the first trimester and cleft palate, we would have had to obtain information on 533,346 unexposed and 134,452 exposed subjects.

Another limitation stems from the use of birth records to ascertain individual exposure and outcome information. Although cases were identified through birth certificate data, we believe these records to be fairly accurate, as diagnosis of condition does not require a medical specialist. Birth records were also used to identify maternal residence for assigning the appropriate water service; however, this does not account for residential mobility, one potential source of exposure misclassification. Residential address reported

on birth certificates may not accurately represent location during the first critical months of pregnancy (Shaw et al., 1992; Zender et al., 2001).

Potential exposure misclassification could also result from lack of information regarding personal exposure. Exposure estimates were based upon distribution system DBP concentrations and do not account for variability in personal habits affecting ingested, inhaled or dermal exposures. Exposure misclassification would also result if maternal water exposures more often occurred outside the service area of the designated water treatment system. Efforts were taken in this study to minimize misclassification due to spatial variability within the distribution systems. This study was built upon an EPA feasibility study, where we utilized the EPA's national DBP database (US EPA, 1998) to select water treatment distribution systems characterized by low spatial variability. These biases are all most likely independent of case or control status and are, therefore, nondifferential.

Risk evaluation is necessary for DBPs in drinking water, due to the potential for widespread exposure. For congenital anomalies, toxicological evidence of an association with DBPs exists, yet few human studies have been conducted in this area. This work represented an effort to explore this relationship using seasonal variability and intra-community comparisons to define a natural experiment. We improved upon previous exposure assessments by considering two classes of DBPs (THMs and HAAs) and multiple time periods of exposure, although we found no associations between DBPs and cleft palate. While several byproducts were evaluated, other DBPs, potentially those at or

below detection levels, may be responsible for environmentally induced cleft palate. Future efforts are needed to confirm, using specific and critical exposure windows, any relationship between DBPs and cleft palate. By utilizing the methods presented in this study for this and other birth outcomes, the true effects of DBP exposures may become more apparent.

5.5 TABLES

Table 5.5.1. Characteristics of subjects and frequency of cleft palate.

	Study Population	%	Number of Cleft Palate Cases
Total Births	51,493		16
Maternal Age			
<20	5982	11.6	2
20-29	27654	53.7	11
≥30	15090	29.3	3
Unknown	2767	5.4	0
Maternal Race			
White	46037	89.4	11
Black	1461	2.8	2
Native American	1838	3.6	1
Other	1819	3.5	2
Unknown	338	0.7	0
Maternal Ethnicity			
Non-Hispanic	32790	63.7	11
Hispanic	16381	31.8	4
Unknown	2322	4.5	1
Maternal Education			
≥1 year of college	23892	46.4	8
High school graduate	14431	28.0	1
<12th grade	12026	23.4	7
Unknown	1144	2.2	0
Maternal Smoking			
No	47248	91.8	15
Yes	3641	7.1	1
Unknown	604	1.2	0
Parity			
0	20167	39.2	6
1	15360	29.8	5
2	8961	17.4	4
3	3975	7.7	0
≥4	2938	5.7	1
Unknown	99	0.2	0
Start of Prenatal Visits			
1st month	10961	21.3	1
2nd month	20966	40.7	8
3rd month	9396	18.3	3
4th-6th months	7363	14.3	1
7th month – birth	1405	2.7	2
Never	825	1.6	1
Unknown	577	1.1	0

Table 5.5.2. Frequency of maternal exposure to total trihalomethane (TTHM) concentrations during precise time periods of pregnancy.*

TTHM (µg/l)	1st Trimester	1st month	2nd month	3rd month	6th week	7th week	8th week
<20	316	99	86	100	16	10	8
	(0.6)	(0.2)	(0.2)	(0.2)	(0.03)	(0.02)	(0.02)
20-39	8464	12098	11960	9269	9353	9341	9302
	(16.5)	(23.5)	(23.2)	(18.0)	(18.2)	(18.1)	(18.1)
40-59	34887	30851	31399	34182	33321	33451	33570
	(67.8)	(59.9)	(60.9)	(66.4)	(64.7)	(64.9)	(65.2)
60-79	3471	3277	3264	3460	3469	3478	3449
	(6.7)	(6.36)	(6.3)	(6.7)	(6.7)	(6.8)	(6.7)
≥80	0	0	0	133	351	342	327
				(0.26)	(0.7)	(0.7)	(0.6)

*Due to exposure time start or end days outside our range of study dates, exposure columns do not total 100%.

Table 5.5.3. Odds ratios and 95 percent CIs for cleft palate according to DBP levels.

Exposure	Exposure Time Period (Adjusted OR and 95% CI)									
	Most Sensitive Week 8		More Sensitive Week 7		Less Sensitive Week 6		2nd Month		1st Trimester	
	OR	95%CI	OR	95% CI	OR	95%CI	OR	95% CI	OR	95%CI
TTHM (µg/l)										
<40*	-		-		-		-		-	
40-59	0.3	0.11-0.97	0.3	0.11-0.97	0.3	0.11-0.98	0.4	0.15-1.32	0.3	0.01-0.87
60-79	0.7	0.24-1.96	0.7	0.24-1.96	0.7	0.24-1.96	0.8	0.28-2.29	0.7	0.23-1.90
≥80	‡	-	‡	-	‡	-	†	-	†	-
Continuous	1.0	0.88-1.01	1.0	0.88-1.01	1.0	0.89-1.01	1.0	0.89-1.01	1.0	0.91-1.01
Chloroform (µg/l)										
<10*	-		-		-		-		-	
10-15	0.4	0.11-1.30	0.3	0.10-1.05	0.3	0.10-1.06	0.3	0.10-0.96	0.4	0.13-1.36
15-20	0.6	0.15-2.08	0.3	0.08-1.34	0.3	0.08-1.34	0.3	0.05-1.23	0.4	0.10-1.70
≥20	‡	-	‡	-	‡	-	‡	-	‡	-
Continuous	0.9	0.74-1.01	0.9	0.74-1.01	0.9	0.74-1.01	0.9	0.74-1.01	0.9	0.76-1.02
BDCM (µg/l)										
<15*	-		-		-		-		-	
15-17	0.2	0.05-1.05	0.4	0.10-1.47	0.4	0.10-1.46	0.2	0.05-1.01	0.1	0.01-0.74
≥17	0.5	0.14-1.52	0.5	0.15-1.77	0.5	0.15-1.77	0.6	0.15-2.10	0.5	0.14-1.49
Continuous	0.8	0.65-0.99	0.8	0.65-0.99	0.8	0.66-1.00	0.8	0.69-1.00	0.9	0.76-1.03
DBCM (µg/l)										
<14*	-		-		-		-		-	
14-15.5	1.0	0.28-3.34	0.6	0.18-2.27	0.7	0.18-2.30	0.5	0.12-1.82	0.9	0.27-3.19
≥15.5	0.7	0.20-2.73	0.6	0.18-2.20	0.6	0.18-2.22	0.9	0.27-3.16	0.8	0.21-2.96
Continuous	0.9	0.75-1.17	0.9	0.75-1.18	1.0	0.76-1.19	1.0	0.77-1.18	1.0	0.81-1.21
HAA5 (µg/l)										
<8*	-		-		-		-		-	
8-15.5	0.6	0.17-2.44	0.7	0.18-2.48	0.7	0.17-2.47	0.7	0.24-2.34	0.7	0.22-2.37
≥15.5	0.4	0.10-1.88	0.5	0.11-1.92	0.3	0.06-1.52	0.2	0.02-1.40	0.3	0.06-1.51
Continuous	1.0	0.89-1.03	1.0	0.89-1.02	0.9	0.87-1.01	0.9	0.87-1.01	0.9	0.88-1.02
DBAA (µg/l)										
<4 *	-		-		-		-		-	
4-5	0.6	0.09-4.58	1.1	0.18-6.41	0.6	0.09-4.50	0.5	0.08-3.06	0.4	0.07-2.19
≥5	0.6	0.81-4.19	0.3	0.02-2.93	0.3	0.02-2.84	0.3	0.03-2.66	0.2	0.02-1.91
Continuous	0.9	0.45-1.92	0.8	0.41-1.66	0.8	0.37-1.73	0.8	0.45-1.38	0.7	0.49-1.10
DCAA (µg/l)										
<6 *	-		-		-		-		-	
6-7	0.7	0.10-5.06	0.7	0.10-4.91	0.7	0.10-4.83	0.5	0.09-2.53	0.5	0.09-2.55
≥7	0.7	0.10-4.96	0.7	0.10-4.96	0.4	0.03-3.93	‡	-	0.2	0.03-2.18
Continuous	0.8	0.53-1.33	0.8	0.49-1.26	0.7	0.42-1.20	0.8	0.52-1.13	0.7	0.54-0.99
TCAA (µg/l)										
<4 *	-		-		-		-		-	
4-5.25	1.2	0.20-7.09	1.2	0.20-7.17	1.2	0.21-7.36	2.1	0.38-11.38	0.7	0.16-3.21
≥5.25	0.4	0.03-4.10	0.4	0.04-4.20	‡	-	‡	-	‡	-
Continuous	0.9	0.58-1.41	0.9	0.55-1.36	0.8	0.45-1.26	0.8	0.51-1.25	0.7	0.49-1.08

All analyses were adjusted for race.
† No subjects in category of exposure.
‡ No cases in category of exposure.
*Referent Category

Table 5.5.4. Number of unexposed and exposed subjects required for detection of each odds ratio with 80 percent power.

Odds Ratio[α]	Number of Unexposed Subjects	Number of Exposed Subjects	Total Number of Subjects
1.5	134,452	533,346	667,798
2.0	42,518	168,662	211,180
2.5	22,816	90,508	113,324
3.0	15,025	59,600	74,625
3.5	11,011	43,680	54,691
4.0	8,631	34,237	42,868

[α]Where incidence of cleft palate in the unexposed subjects is approximately 0.04%, as estimated from national birth data.

5.6 REFERENCES

Bove FJ, Fulcomer MC, Klotz JB, Esmart J, Dufficy EM, Savrin JE. Public water contamination and birth outcomes. Amer J Epidemiol. 1995:141;850-862.

Dodds L, King W, Wolcott C, Pole J. Trihalomethanes in public water supplies and adverse birth outcomes. Epidemiology. 1999:443-446.

Dodds L, King W. Relation between trihalomethane compounds and birth defects. Occup Environ Med. 2001:58;233-237.

Hertz-Picciotto I. Environmental Epidemiology. In: Rothman KJ, Greenland S. eds. Modern Epidemiology, 2nd edition. Philadelphia: Lippincott, Williams and Wilkins, 1998.

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

Moore KL. Persaud T.V.N. Before we are born. Essentials of embryology and birth defects. 5th edition. W. B. Saunders Company. Philadelphia, PA. 1998.

Nieuwenhuijsen MJ. Toledano MB. Eaton NE. Fawell J. Elliott P. Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: a review. Occupational & Environmental Medicine. 2000:57(2);73-85.

O'Rahilly R. Müller F. Human Embryology and Teratology. 3rd Edition. John Wiley & Sons, Inc. New York, NY, 2001.

Reif JS, Hatch MC, Bracken M, Holmes L, Schwetz B, Singer PC. Reproductive and developmental effects disinfection byproducts in drinking water. Environ Health Perspect. 1996:104;1056-1061.

Shaw GM, Malcoe LH. Residential mobility during pregnancy for mothers of infants with or without congenital cardiac anomalies: A reprint. Arch Environ Health. 1992:47;236-238.

Shaw GM, Ranatunga D, Quach T, Neri E, Correa A, Neutra RR. Trihalomethane exposures from municipal water supplies and selected congenital malformations. Epidemiology. 2003:14;191-199.

Smith M, Randall J, Stober J. Teratogenic activity of trichloroacetic acid in the rat. Teratology. 1989: 40; 445-451.

U.S. Environmental Protection Agency. Part IV National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts; Final Rule. U.S. Federal Register December 16, 1998. 1998:63;69390-69476.

Zender R, Bachand AM, Reif JS. Exposure to tap water during pregnancy. Journal of Exposure Analysis & Environmental Epidemiology. 2001:11(3); 224-30.

CHAPTER 6

6 CONCLUSIONS

Recent epidemiological studies have found associations between exposure to DBPs and growth-related birth outcomes, although these risks have not been consistent. This is potentially due to the spatial and temporal variability that exists using current epidemiologic methods. Variability with evaluation of specific time-windows of exposure has also been a problem, specifically for studies of select congenital malformations (Bove et al., 1995; Dodds et al., 1999; Klotz and Pyrch, 1999; Dodds and King, 2001; Shaw et al., 2003). This dissertation was undertaken to improve the body of information regarding the relationship between DBPs and adverse reproductive outcomes by developing, evaluating and implementing new methods of exposure assessment and analysis for these epidemiologic investigations.

Spatial variability can lead to exposure misclassification, which occurs when persons living at different locations within a distribution system are exposed to significantly different levels of DBPs but are assigned the same exposure levels. We performed a feasibility study to characterize water distribution systems as having spatially consistent DBP concentrations. In this study, we used the ICR database to obtain data on DBPs

from four sampling points per quarter, over at least four quarters between July 1997 and December 1998 at 84 water treatment plant distribution systems across the U.S. These data were tested using two methods to determine the degree of spatial variability. For the *Two-way ANOVA* method, we found that the assumptions were not always satisfied, and therefore statistical results were questionable. Overall, the method used to select sites according to cutpoints in the epidemiological literature seems to be most practical. In our attempt to select sites with low spatial variability, however, we found it difficult to locate sites with consistently high TTHM levels that were characteristically brominated. Whether this disadvantage stemmed from the limitations of our criteria or the rarity of this profile, we are unsure. Also, because this method is contingent upon previous epidemiologic findings, the cutpoints may be subject to change as future studies identify more meaningful threshold values. We recommend that future researchers of DBPs employ these methods presently, using TTHM and HAA data, and further determine whether the proportions of individual species remain fairly consistent throughout time. Our simple selection methods will aid future epidemiology studies utilizing prospectively or retrospectively collected data on DBPs by helping to identify appropriate sites for study.

Sensitivity to teratogens during embryonic and fetal development is contingent upon the developmental stage at the time of exposure and is important knowledge for analysis of epidemiologic information. Therefore, with respect to assigning exposure, it is appropriate to determine an exposure value that is as close as possible to the actual

exposure of concern during the specific gestation week of pregnancy, if not on the actual day.

For some adverse outcomes of pregnancy, e.g. low birth weight or preterm birth, there is no time period during human embryologic or fetal development that is known to be a period of sensitivity, although the third trimester is most often suspected (Kline, 1989). As stated previously, no current statistical method exists to evaluate time-dependent outcomes (e.g. preterm birth) with respect to temporally variable environmental exposures, such as disinfection by-products (DBPs). Prior investigations of disinfection by-products have relied upon a variety of techniques to evaluate these relationships. We modeled hypothetical populations to analyze a commonly used technique, the spline method, for the relationship between preterm birth and average exposure to BDCM. The results of this study demonstrate how time is related to both temporally variable exposures and time-related birth outcomes using the spline method and suggest the magnitude and direction of related biases. Cutpoint bias of temporally variable exposures influences risk estimates in either direction, depending upon degree of temporal variability and placement of cutpoint, but can be eliminated with use of fixed time periods. Selection bias has a greater affect on the risk estimate, but can be avoided with change in exclusion criteria. Overall, there are methods with which to assign third trimester exposure to temporally variable chemicals, however, they must be approached with great consideration and caution.

Among the studies quantifying exposure to DBPs, there are substantial variations with respect to evaluation of exposure over time windows of gestation. To evaluate birth outcome relationships with different time periods of sensitivity to DBP exposures, we conducted a retrospective cohort study, evaluating exposure through spline regression over very specific windows of exposure, to determine whether a relationship exists between exposure to total trihalomethanes (TTHM), chloroform, bromodichloromethane (BDCM), dibromochloromethane (DBCM) and the five most prevalent individual haloacetic acids (HAA5) in drinking water and cleft palate. Toxicologic data indicate a potential mode of action for specific THMs and HAAs via disruption of signaling proteins involved in morphogenesis, leading to altered epithelial cell differentiation and mesenchymal (palate) formation. The study population consisted of all singleton births in a large community served by three water treatment plants. This community was chosen since THMs and HAAs varied by season with levels ranging from low to moderate, but exhibited little spatial variability within the distribution system. Birth certificate data (n=51,717) were obtained to capture all births in which the critical time periods of exposure for each pregnancy occurred between January 1998 and March 2003. We approximated the average exposure over each time period for each birth by taking the mean of predicted daily DBP values. The predicted values were interpolated from the respective biweekly, monthly and quarterly THM or HAA concentrations reported by the utility. Logistic regression techniques were used to estimate the relationship between mean complete first trimester, second month of gestation, and 6^{th}, 7^{th} and 8^{th} weeks of pregnancy of each DBP and cleft palate. We found no association between categorical levels of exposure to individual THMs or TTHM during specific critical time windows

and cleft palate, although risks were slightly increased for exposure to chloroform during the 8th week of gestation, the time period of greatest sensitivity for induction of cleft palate, as compared to the other time periods. Results for average 2nd tertile exposure to TCAA (4-5.25 μg/l), as compared to the referent (<4 μg/l), were OR=1.2 for individual weeks and OR=2.1 for the second month, although confidence intervals were imprecise and all overlapped the null value. No associations were observed for any DBP evaluated as a continuous variable and cleft palate. In conclusion, we found a possible association between exposure to moderate levels of TCAA and cleft palate and no associations between exposures to any THM, regardless of exposure window. However, the small numbers of cleft palate cases, leading to reduced power, significantly limited this study. Adjustment for potential confounders did not change the conclusions. Future efforts are needed to confirm, with respect to specificity of exposure timing and analysis, any relationship between DBPs and cleft palate. By utilizing the methods presented in this study for this and other birth outcomes, the true effects of DBP exposures should become more evident.

The relationship between exposure to DBPs and adverse birth outcomes in humans is unclear. While toxicologic studies have shown DBP-induced toxic or teratogenic effects to be biologically plausible, there is little evidence for a causal relationship between DBPs and induction of adverse outcomes in humans. In this dissertation, by developing new methods of site selection for epidemiologic studies, evaluating sources of bias for analysis of growth-related birth outcomes during critical time periods of development, and implementing new techniques for more precise quantification of risk for birth defect

studies, we have improved our ability to investigate the epidemiologic relationship between DBPs and reproductive outcomes. Future research directions should include evaluation of spatial variability or efforts to build investigations upon spatially consistent sites, assessment of other forms of analytical bias and investigation of exposure over critical time periods of gestation. In particular, future studies should focus upon birth outcomes that have had significant findings in toxicologic studies of DBPs. The findings in this study can serve as a foundation to these studies and will provide some practical and theoretical guidance to many future epidemiologic investigations of DBPs and birth outcomes.

LIST OF ABBREVIATIONS*

ANOVA	Analysis of Variance
BDCM	Bromodichloromethane
BIF	Bromine Incorporation Factor
CNS	Central Nervous System Defects
DBAA	Dibromoacetic Acid
DBCM	Dibromochloromethane
DBP(s)	Disinfection By-Product(s)
DCAA	Dichloroacetic Acid
DI	Diagnoses Index
EDC	Estimated Date of Conception
EPA	United States Environmental Protection Agency
GST	Glutathione S Transferase
HAA5	The sum of five most prevalent haloacetic acids (MCAA, MBAA, DCAA, DBAA, TCAA).
HAAs	Haloacetic Acids
ICR	Information Collection Rule
IUGR	Intrauterine Growth Retardation
LBW	Low Birth Weight
LMP	Last Menstrual Period
MBAA	Monobromoacetic Acid
MCAA	Monochloroacetic Acid
MCDs	Major Cardiac Defects
MCLs	Maximum Contaminant Levels
NTDs	Neural Tube Defects
OCDs	Oral Cleft Defects
OR	Odds Ratio
PTB	Preterm Birth
RTD	Respiratory Tract Defects
SAB	Spontaneous Abortion
SDWA	Safe Drinking Water Act
SGA	Small for Gestational Age
SBL	Small Body Length
SCC	Small Cranial Circumference
SRS	Simple Random Sample
TCAA	Trichloroacetic Acid
THM	Trihalomethane
TLBW	Term Low Birth Weight
TTHMs	Total Trihalomethanes
UTD	Urogenital Defects
VLBW	Very Low Birth Weight

*List includes only those abbreviations used multiple times.