Haloacetic acids in public drinking water and risk of adverse birth outcomes in the Born in Bradford cohort

by

Susan C. Edwards

2014

A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy of Imperial College London
and Diploma of Imperial College

Department of Epidemiology and Biostatistics
Imperial College London

ABSTRACT

Disinfection of drinking water is vital to protect the public against disease. However disinfectants such as chlorine react with organic matter in drinking water to produce a wide range of chemical disinfection by-products (DBPs) of potential health concern including haloacetic acids (HAAs).

This thesis is an epidemiologic analysis investigating the relationship between prenatal exposure to HAAs in drinking water and adverse birth outcomes in "Born in Bradford", a large multi-ethnic prospective birth cohort study based in Bradford, England. It focuses on the understudied and as yet unregulated HAAs which are the second most prevalent class of chlorination DBPs in UK drinking waters.

To assess exposure, area-level concentrations to three select HAAs (measured in drinking water samples newly collected for this study, modelled in time and space, and weighted to each cohort woman's specific trimester of pregnancy by postcode of residence) were combined with individual water consumption information collected via questionnaire at recruitment to the cohort.

Despite the benefits of state-of-the-art exposure metrics and a large sample size, this study does not find any significant patterns of association between prenatal exposure to HAAs and either birth weight, being born term low birth weight or small-for-gestational age.

Water consumption over the course of late pregnancy was further studied in a subset of cohort women. A small but significant increase in water consumption was reported, bearing in mind that both behaviour change over the third trimester of pregnancy and measurement error likely contributed to this effect.

This research addresses some of the limitations of previous DBP studies in terms of exposure assessment and birth outcome definitions, and uniquely evaluates the variability of individual water consumption over time. It also identifies areas for future research and examines the importance of HAAs and birth weight-based outcomes in the larger research context.

DECLARATION OF ORIGINALITY

I hereby declare that the contents of this PhD thesis are all my own work and that where any material could be construed as the work of others it has been fully cited and referenced, and/or the appropriate acknowledgement given.

Between May 2010 and November 2010, I collected the HiWATE water samples in Bradford. This involved coordinating sampling dates with Yorkshire Water, sending instructions and clean labelled vials containing preservatives by courier, and picking up the samples after collection. I was then charged with driving those samples from Bradford to Cranfield University for laboratory analyses.

This work was made possible by the collaboration and expertise of various individuals. Dr Kees de Hoogh completed the automated geocoding of Born in Bradford participants' residence and work postcodes and addresses; I did the manual geocoding. Dr Hannah Slater advised me on linking women's addresses to water supply zones via ArcGIS (Chapter 4). Professor Nicky Best and Dr Juan Gonzalez Maffe helped me develop the WinBUGS models both for the HAA modelling (Chapter 4) and the repeat questionnaire study (Chapter 8), and advised me on the statistics relating to the main epidemiologic analysis (Chapter 7). Drs James Bennett and Léa Fortunato helped me in the initial stages of HAA modelling as well. The idea for the repeat questionnaire study came from Professor Mark Nieuwenhuijsen and Dr Mireille Toledano. I got ethics approval for the study in September 2010 and conducted the study between September and December 2010. No personal data were to leave the BiB office for confidentiality reasons. I therefore travelled to Bradford to access women's information, in order to prepare the mailings. I then collected the returned questionnaires, entered all data manually, and developed the statistical analysis plan. I kept in close contact with Dr Emily Petherick from the Born in Bradford team regarding the data extracts, and data cleaning issues.

On 7th June 2010, I collected 2 x 20L samples for a toxicology study ran jointly by Dr Susan Richardson (National Exposure Research Laboratory, U.S. Environmental Protection Agency, Athens, GA, USA) and Professor Michael Plewa (University of Illinois at Urbana-Champaign, IL, USA). I reviewed the study which came out of this effort and which was published (Appendix B).

I presented some of this work at the BiB Annual Open Day 2012 in Bradford, at the Gordon Conference (2012) on Drinking Water Disinfection By-Products at Mount Holyoke College, South Hadley, MA, USA, and at the International Society for Environmental Epidemiology (ISEE) conference in Barcelona, Spain (2011) and Basel, Switzerland (2013).

COPYRIGHT DECLARATION

ACKNOWLEDGEMENTS

I am grateful to my supervisors Dr Mireille B Toledano, Professor Nicky Best and Professor Mark J Nieuwenhuijsen for their guidance and support during these four years.

I would also like to collectively thank all my colleagues and friends at Imperial, including my fellow PhD colleagues in room 171. I am particularly grateful to Dr Rachel Smith for generously sharing her knowledge of disinfection by-products with me since my first day as a PhD student. I also thank Drs Léa Fortunato, James Bennett, Hannah Slater and Kees de Hoogh for their help with statistical and geocoding aspects of the thesis. Most of all, I am indebted to Dr Juan Gonzalez Maffe, who challenged me and taught me ever so much. This experience would not have been as valuable without him.

I am grateful to all those who work on the Born in Bradford study, in particular Dr Emily Petherick for the data provided, Shaheen Akhtar, Dr Pauline Raynor, and Dawn Wray for their assistance during recruitment to the repeat questionnaire study, and Professor John Wright for leading the project. Many thanks to all the Born in Bradford participants without whom this study would not have been possible. I also thank Yorkshire Water for providing us with data and for their assistance with water sampling in Bradford, Dr Emma Goslan from Cranfield University for the HAA analyses, and Carol Edwards for access to Imperial's Environmental & Water Resource Engineering laboratory.

My PhD was funded for three years by a Medical Research Council (MRC) Capacity Building Studentship. Field work was funded through my MRC Studentship's Research Training support grant 2010-11, as well as grant P10462 ("Health Risks", P.I.: Mireille Toledano).

My heartfelt thanks to Léa, Rachel, Tessa, Juan, Federico, Poppy and Peter for kindly proof-reading thesis chapters and giving me valuable feedback, and to Doris for priceless round-the-clock technical assistance during the writing-up phase.

I would like to thank my close friends for their indefatigable love and encouragement while I have been working on this thesis.

Finally, thank you to my parents, Doris and Peter, for their constant love and support. I would like to dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
ALSPAC	Avon longitudinal study of parents and children
ANOVA	Analysis of Variance
BCAA	bromochloroacetic acid
BDCAA	bromodichloroacetic acid
BDIMC	Bradford District Infant Mortality Commission
BiB	Born in Bradford
BMI	Body Mass Index
BRI	Bradford Royal Infirmary
CHO	Chinese hamster ovary
CI	Confidence Interval
DBAA	dibromoacetic acid
DBCAA	dibromochloroacetic acid
DBCM	dibromochloromethane
DBP	Disinfection By-Product
DCAA	dichloroacetic acid
DCBM	bromodichloromethane
DCBM	dichlorobromomethane
DIC	Deviance Information Criterion
DNA	Deoxyribonucleic Acid
DOC	Dissolved Organic Carbon
DWI	UK Drinking Water Inspectorate
EDD	Estimated Date of Delivery
FAQ	Frequently Asked Questions
FTU	Formazin Turbidity Unit
GC-ECD	Gas Chromatography-Electron Capture Detector
GC-HRMS	Gas Chromatography/High Resolution Mass Spectrometry
GIS	Geographical Information System
GTT	Glucose Tolerance Test
HAA	Haloacetic Acid
HiWATE	Health Impacts for Long-Term Exposure to Disinfection By-Products in Drinking Water
HPLC	High Performance Liquid Chromatography
ICC	Intra-Class Correlation
ICL	Imperial College London
IMD	Index of Multiple Deprivation
IQR	Inter-Quartile Range
IUGR	Intrauterine Growth Restriction
LBW	Low Birth Weight
LC-MS-MS	Liquid Chromatography – Tandem Mass Spectrometry
LLE	Liquid-Liquid Extraction
LMP	Last Menstrual Period
LMS	Method summarising the changing birth weight distribution by three curves representing the skewness (L), median (M) and coefficient of variation (S)
LOAEL	Lowest Observed Adverse Effect Level
LOD	Limit Of Detection
LOWESS	Locally Weighted Scatterplot Smoothing
LSOA	Lower Super Output Level
MBAA	bromoacetic acid
MCAA	chloroacetic acid
MCL	Maximum Contaminant Level

MCS	Millenium Cohort Study
MRC	Medical Research Council
MRL	Minimum Reporting Limit
mRNA	Messenger Ribonucleic Acid
MX	3-chloro-4-(dichloromethyl)-5-hydroxy-5H-furan-2-one (aka Mutagen X)
NHS	National Health Service
NIHR	National Institutes for Health Research
NOAEL	No Observed Adverse Effect Level
NOM	Natural Organic Matter
OR	Odds Ratio
P	Pakistani (Pakistani origin)
PCT	Primary Care Trust
PET	Polyethylene Terephthalate
POU	Point-Of-Use
PTD	Preterm Delivery
q1	quarter 1 (January through March)
q2	quarter 2 (April through June)
q3	quarter 3 (July through September)
q4	quarter 4 (October through December)
QC/QA	Quality Control/Quality Assurance
RQS	Repeat Questionnaire Study
RR	Relative Risk
SD	standard deviation
SDS	Standard deviation score
SES	Socio-Economic Status
SGA	Small-for-Gestational Age
TBAA	tribromoacetic acid
TBROM	Total Brominated Trihalomethanes
TCAA	trichloroacetic acid
THM	Trihalomethanes
tLBW	term Low Birth Weight
TOC	Total Organic Carbon
TOX	Total Organic Halogen
TTHM	Total Trihalomethanes
UK90	British 1990 reference birth centiles
UN	United Nations
US EPA	United States Environmental Protection Agency
UV254	Ultraviolet-Absorbance at 254nm
vLBW	Very Low Birth Weight
VPC	Variance Partition Coefficient
WB	White British
WHO	World Health Organization
WSZ	Water Supply Zone
WTW	Water Treatment Works
YW	Yorkshire Water

CHAPTER 1 INTRODUCTION

This chapter introduces the basic definitions and concepts on the subject of disinfection by-products and birth outcomes, as well as the experimental and human evidence to date on their association and possible mechanism of action. Justification for this thesis based on gaps in knowledge is established, and the structure of the thesis is outlined.

1.1 Background

1.1.1 Disinfection By-Products (DBPs) and Haloacetic Acids (HAAs)

1.1.1.1 Definitions

Water disinfectants such as chlorine—the most common disinfectant used in the United Kingdom—react with natural organic matter and bromide ions present in the water during the disinfection process to create disinfection by-products (DBPs).

Haloacetic acids (HAAs) are the second most prevalent chlorinated DBPs in drinking water after trihalomethanes (THMs), together comprising more than 50% on a weight basis (Singer et al. 2002), and 90% of the total volume of DBPs (Weinberg et al. 2002). There are nine chloro- and bromo-haloacetic acids, commonly referred to as "HAA9". They are MCAA, MBAA, DCAA, BCAA, TCAA, DBAA, BDCAA, DBCAA, and TBAA (Table 1.1). The five HAAs currently under regulations in the United States are referred to as "HAA5" are MCAA, MBAA, DCAA, TCAA, and DBAA.

1.1.1.2 Main route of exposure to HAAs: ingestion

Contrary to THMs, ingestion is thought to be the main exposure route to tap water HAAs (Nieuwenhuijsen et al. 2000b). Because HAAs are non-volatile at typical shower water temperatures, inhalation exposure to HAAs during showering is minimal (Kim and Weisel 1998; Xu and Weisel 2003). Reports on the estimated dermal dose from daily bathing activities conclude that it represents an insignificant proportion of daily HAA uptake (Kim and Weisel 1998; Xu et al. 2002), as polar molecules of HAAs do not readily penetrate the skin (Xu et al. 2002).

That being said, HAAs especially DCAA persist in boiled water and foods (Raymer et al. 2000; Wu et al. 2001). Thus, individual exposures to these compounds are determined by total ingestion of tap water (non-boiled and boiled), hot drinks, and tap water-containing foods (Egorov et al. 2003).

1.1.1.3 Regulations of HAAs and THMs

Current drinking water regulatory standards for DBPs were established based on evidence of their adverse human health effects, risk of cancer in particular (EPA 2006; WHO 2008).

Both THMs and HAAs are regulated in the US (maximum allowable contaminant level (MCL) for locational running annual averages of 80ug/l for TTHM (EPA 1979, 1998) and 60ug/l for HAA5 (EPA 2006; EPA 1998)). According to Krasner (2009), these regulations were based on the broad use of THMs and HAAs as 'surrogates' for the toxicity associated with chlorinated water. The control of THMs and HAAs results in an overall reduction in the concentration of many other DBPs (Reckhow and Singer 1984) which may themselves be associated with the adverse health effects.

Only THMs are currently regulated in the UK, with the MCL for total THM (TTHM, i.e. the sum of chloroform, bromodichloromethane, chlorodibromomethane and bromoform) at 100ug/l in any sample (DWI 2010). However, a standard of 80ug/L for HAA9 has been proposed for the future revision of the European Union's Drinking Water Directive (DHI 2008), renewing interest in levels of HAAs in UK drinking waters and control methods.

In Australia and New Zealand, the regulated limits of MCAA, DCAA and TCAA in drinking water are 150, 100 and 100ug/L, respectively (Golfinopoulos and Nikolaou 2005) (Table 1.2). Also listed in Table 1.2 are the World Health Organization (WHO) guidelines for DBPs (WHO 2008).

1.1.1.4 Methods used for exposure assessment to DBPs to date

Exposure assessment has been and remains the limiting factor in epidemiologic studies of drinking water DBPs and reproductive health (Arbuckle et al. 2002; Savitz 2012).

The following study design types reflect the trend in increasing precision of exposure assessment in this field over time (S. Cordier, Gordon Conference, 2012):

1. ecological studies
2. studies based on individual exposure
3. studies with area-level concentration data at water treatment level and some individual consumption data
4. studies with area-level concentration data at participants' homes and some individual consumption data
5. studies with biomarkers

Ecological studies (study design 1) typically require identification of a study population served by a wide range of water providers with varying concentrations of DBPs, and exposure assignment based

on residential location. Such studies tend to have the advantage of large contrasts in concentration that are certain to result in differences in exposure. However, they are susceptible to ecologic confounding in which other characteristics of the regions under study may well influence pregnancy outcomes (Savitz 2012).

Collecting individual level data (study design 2) necessitates querying people about their individual water use to distinguish exposure levels within one or a few water-services areas. Studies examining variation in water use within a single area tend to yield limited variation in tap water concentrations and therefore require accurate information on individual determinants of exposure to be informative. The study of volatile DBPs such as THMs through this approach is challenging, because a major component of the exposure comes from inhalation (Weisel and Chen 1994; Xu and Weisel 2005) and dermal absorption (Xu et al. 2002). Individual exposure assessment is more feasible for non-volatile DBPs such as HAAs, for which exposure occurs solely through ingestion (Nieuwenhuijsen et al. 2000b), but it still requires valid estimates of consumption at home and in other locations, information on whether the water is filtered at the tap or with a pitcher, and whether the water has been heated, as for coffee or tea (Savitz 2012).

A few studies have combined these approaches (study designs 3 and 4) in the past (Hoffman et al. 2008b; Infante-Rivard 2004; Savitz et al. 2006), which is the option chosen in this thesis.

As knowledge of exposure determinants expands (Arbuckle et al. 2002; Nieuwenhuijsen et al. 2000b), it is increasingly clear that the complexity of individual-level exposure determinants may go beyond what self-reports can address effectively—which is where biomarkers (study design 5) come into the picture (Savitz 2012).

1.1.1.5 Consequences of poor exposure assessment

The exposure assessment in Smith's thesis on THMs in Bradford concludes that failure to incorporate individual water use into exposure assessment and reliance simply on area-level THM concentrations results in exposure misclassification, particularly when spatial variability in DBPs is limited across the study area. This finding also suggests that exposure assessment resources would be best focussed on improving individual water use assessment, if spatial variability is likely to be limited in a study area (Smith 2011). This is consistent with previous studies which found that about half of subjects are classified differently by metrics representing THM concentrations at the tap (area-level or household level) and metrics incorporating THM concentrations and individual water use (King et al. 2004; Whitaker et al. 2003b; Wright et al. 2006).

Similarly, in epidemiologic studies where one or the other route of exposure has not been included in exposure assessment (e.g. a study by Savitz et al. (1995) in which exposure assessment reflects only

the ingestion route), exposure metrics may be a poor surrogate for total exposure, again resulting in exposure misclassification (Smith 2011).

1.1.1.5.1 Types of error

There are two types of measurement error to be expected under the ecological or individual study design types (Smith 2011).

Smith (2011) studied eleven studies which used area-level THM or HAA concentrations to estimate exposure for each mother (Bove et al. 1995; Dodds et al. 1999; Gallagher et al. 1998; Hinckley et al. 2005; Kramer et al. 1992; Lewis et al. 2006; Porter et al. 2005; Toledano et al. 2005; Wright et al. 2003, 2004). In these studies exposure measurement error would follow the Berkson error model. This type of error arises when the same approximate measure (a proxy) is used for many subjects, whilst the true exposures vary randomly about this proxy with their mean equal to it (Armstrong 1998). Berkson error causes loss of study power and loss of precision in effect estimates, but rarely causes bias (Armstrong 1998).

When exposure assessment was carried out at the individual level for each mother, as in a study by Aggazzotti et al (2004), exposure measurement error would follow the classical error model. This type of error occurs when the average of many replicate measurements of the same true exposure would equal the true exposure (Armstrong 1998). Classical error biases estimates towards the null (Armstrong 1998).

Finally, Smith (2011) reported that for three studies that have generated personalised exposure metrics by combining area-level THM or HAA concentration estimates with individual information on water use (Hoffman et al. 2008b; Infante-Rivard 2004; Savitz et al. 1995), exposure measurement error may combine elements of both Berkson and classical error models. Without a clear understanding as to the contribution of the various exposure errors, the health risk estimates may be more difficult to interpret (Nieuwenhuijsen et al. 2000b).

1.1.2 Birth outcomes

1.1.2.1 Common indicators of birth outcomes

The most common indicators of suboptimal growth during the fetal period include low birth weight (LBW), very LBW, term LBW, and small-size-for gestational age (SGA). More subtle definitions such as customised SGA, and fetal growth restriction (FGwR) have been developed in the last couple of decades (Table 1.3).

Birth weight is recognised globally as an indicator of perinatal and infant health. Low birth weight (LBW) is generally defined using WHO criteria as a live[1] baby born weighing less than 2500g, and very LBW as a live baby born weighing less than 1500g. A term LBW baby is one whose birth weight is under 2500g at 37 or more completed weeks of gestation. Preterm delivery (PTD) is formally defined solely by the gestational age at which delivery of a live birth occurs (less than 37 completed weeks) (Savitz et al 2002). Live births in the lowest decile ($\leq 10^{th}$ percentile) (Alexander et al. 1996) of weight for gestational age (occasionally in the lowest fifth or even third percentile (Lee et al. 2003)), or two standard deviations (Clausson et al. 2001) below the mean birth weight for gestational age are defined as small-for-gestational age (SGA). The standards used are either derived from the dataset being analysed or based on a published referent (standardised for ethnicity, sex, and/or parity if such referents are available) (Alexander et al. 1996; Wilcox 2010).

SGA is sometimes used interchangeably with intrauterine growth restriction (formerly "retardation") (IUGR). IUGR was first defined as LBW babies born at term, then took on the same definition as SGA in order to extend the IUGR definition to preterm births. More recently still, IUGR diagnosis considered clinical and ultrasonography-based presentation of suboptimal growth, in addition to a fetal growth ratio[2] or fetal length below the tenth percentile for gestational age (Choi et al. 2008; Smith et al. 1997). Other risk factors such as height and weight, ethnicity and even smoking can be included in its definition as well, muddying the picture further.

1.1.2.2 Growth restriction

This thesis aims to identify "pathologically" small (i.e. growth restricted) babies as opposed to identifying all small or light babies, some of which may be "constitutively" small (i.e. "naturally" so) (Nieuwenhuijsen et al. 2009c).

Birth weight without consideration for gestational age at birth is a limited marker of fetal growth (Wilcox 2001). Prematurity is correlated with lower birth weight regardless of pathology (the earlier a baby is born, the less time it has had to grow in the womb, and thus the smaller it is likely to be at birth), but growth restricted babies may technically be born at any gestational age. This means that babies categorised as LBW will represent a mix of those whose growth is suboptimal, those whose growth trajectory was normal but who were delivered early, and those who are small for genetic

[1] As opposed to a stillbirth, which refers to a fetus that is born at 24 or more completed weeks of gestation which does not show any signs of life such as a beating heart, breathing or voluntary movement (per ONS definition: http://www.ons.gov.uk)

[2] Birth weight divided by the mean birth weight for a given ethnic group, gestational age and sex Choi H, Perera F, Pac A, Wang L, Flak E, Mroz E, et al. 2008. Estimating individual-level exposure to airborne polycyclic aromatic hydrocarbons throughout the gestational period based on personal, indoor, and outdoor monitoring. Environ Health Perspect 116:1509-1518.

reasons unrelated to viability, i.e. constitutively so. The challenge is to identify which infants fall into each of the three groups (Savitz et al. 2002) (see Chapter 6 for more details on the pros and cons of each measure, section 6.4.3).

One way to identify growth restricted babies is to focus on birth weight as a continuous measure, avoiding an arbitrary dichotomy and potentially enhancing statistical power, and adjust for gestational age within statistical models. Another way is simply to exclude all premature babies, as with the term LBW measure, such that any babies born weighing less than 2500g despite adequate gestational age are likely growth restricted. Finally, thresholds of "statistically usual" weights for gestational age can be established in order to generate measures of deviation from the subgroup norms such as SGA (Wigle et al. 2008; Wilcox 2010). The presumption here is that those who are at the extreme low end of that distribution of birth weight are likely to have had their growth restricted in some manner (Savitz et al. 2002). But even a baby classified as SGA (i.e. unusually small in a statistical sense (below the tenth percentile) relative to other infants of the same duration of gestation, gender, parity, and race) may not be growth restricted. Indeed SGA being a statistical construct, some babies who are small at birth but grew normally in the womb will fall below the cut-off point, and some growth restricted babies will reach a weight above the cut-off point because SGA fails to distinguish between "constitutively" and "pathologically" smallness (Nieuwenhuijsen et al. 2009c). Succeeding in making this distinction will often depend on the growth charts of the referent population used. The reference used to derive the SGA measure studied in this thesis is described in Chapter 6 (see section 6.2.3).

1.1.2.2.1 Customised measures

One of the limitations of using population growth standards is that they include babies who seem to be of normal size but have in fact failed to reach their own growth potential. Indeed, there is evidence to show that the use of individually-adjusted fetal growth charts instead of standard population growth charts significantly reduces the proportion of false-positive and false-negative diagnoses of fetal growth restriction (Gardosi 2006; Gelbaya and Nardo 2005). The customised SGA developed by Gardosi et al (2006) uses an optimal growth curve for an individual fetus as the referent based on sonographic measurements (Gardosi 1998). Using a similar approach, Mamelle et al (2001) suggested a variable called fetal growth restriction (FGwR), which is based on the estimation of an infant's individualised birth weight limit, taking into account his/her genetic growth potential which depends on maternal characteristics (height, pre-pregnancy weight, smoking) and parental characteristics (age, weight, height).

Customised birth weight centiles assess birth weight against an individually calculated standard, which is based on the growth potential of each fetus (Gardosi et al. 1995). The customised standard adjusts for characteristics such as ethnic origin, parity, and maternal height and weight, excludes

pathological factors known to affect fetal growth, such as smoking, and uses a fetal weight-derived 'proportionality curve' to define normal weight at various gestational age points (Gardosi et al. 2007). This avoids limits based on preterm birth weights, which are heavily skewed because of the association between preterm birth and fetal growth restriction.

However, critics of this method have argued that the benefits of the "customised" method are marginal (Hutcheon et al. 2011), that it is unable to distinguish between the pathological and physiological influences of maternal characteristics on birth weight, and that it main value proposition is its recognition of the inappropriateness of using a birth weight-based standard at preterm ages (Hutcheon and Platt 2008) rather than because of its "customisation" for maternal characteristics (Hutcheon et al. 2008). Since data on maternal characteristics are often missing, a non-customised but intrauterine-based standard may indeed be the most parsimonious and practical standard for the prediction of perinatal mortality in clinical practice (Hutcheon et al. 2008; Zhang and Sun 2013).

In addition to the fact that no pre-pregnancy maternal weight—needed for the derivation of customised SGA—was available for the pregnant women from the Born in Bradford (BiB) cohort under study here, it was also not clear how comparable to the previously published literature customised SGA results would be. For these reasons, this method was not explored further in this thesis.

1.1.2.3 Consequences of poor birth outcomes on health

Growth restriction (not due to prematurity) is frequently linked to perinatal morbidity and mortality (MaCH 2001) as well as adverse effects in childhood and later life (Barker 1997; Wilcox 2010).

1.1.2.3.1 Consequences on early mortality and child health

Babies born LBW or SGA have been shown to have higher perinatal mortality and morbidity (Ashworth 1998; Branum and Schoendorf 2002; Gulmezoglu et al. 1997; IHDP 1990; Lee et al. 2003; Lira et al. 1996; McCormick 1985; Thomas et al. 2000; Wen et al. 2005; Yasmin et al. 2001) and more likely to experience developmental problems in childhood (Hediger et al. 2002; Rice and Barone 2000; Richards et al. 2002; Sizonenko et al. 2006; van Wassenaer 2005) than their counterparts born weighing 2500g or more, or who are not SGA. Indeed low birth weight has been associated with both cognitive and neurologic impairment (Paz et al. 1995; Taylor and Howie 1989).

However the causal role of birth weight in infant mortality is controversial; if the relationship is non-causal, it may be an unimportant endpoint in itself, and inconsequential in the analysis of infant mortality or other outcomes (Wilcox 2001). This would also imply that interventions to increase birth weight may be wasted.

1.1.2.3.2 Consequences on later mortality and morbidity

As per West (2011), a large number of studies have demonstrated inverse (sometimes 'J' shaped) associations of birth size (most commonly birth weight) with cardiovascular disease endpoints (Barker 2002; Huxley et al. 2007) and risk factors including type 2 diabetes (Whincup et al. 2008), fasting glucose (Al Salmi et al. 2008), insulin (Lawlor et al. 2003), total cholesterol (Lawlor et al. 2006), and triglycerides (Gluckman and Hanson 2004; Owen et al. 2003) as well as positive associations with high density lipoprotein cholesterol (Gluckman and Hanson 2004).

Few investigators suggest that size at birth *per se* matters, but rather that size is a proxy marker for other causal risk factor(s). For example and as eloquently summarised in West (2011), several hypotheses have been proposed to explain the association between birth weight and later cardiovascular disease outcomes:

a) Factors affecting intrauterine nutrition and growth provide signals to the developing fetus about the environment in which they will grow-up and develop, and as a result 'programme' the developing fetus for an environment of thrift or plenty (Bateson et al. 2004; Leon 2004). Under this hypothesis, poor intrauterine nutrition results in poor intrauterine growth and low birth weight, and programmes the offspring for a life of thrift. If the offspring subsequently experience a life of nutritional plenty, typified by high fat, energy dense diets, they are at a particularly increased risk of future cardiovascular disease. Epigenetic mechanisms are increasingly thought to mediate these processes (Waterland and Michels 2007). For example, significant changes to epigenetic states across the genome during intrauterine development present the opportunity for environmental stresses, such maternal malnutrition, to influence gene expression and thus phenotype (Li et al. 2010).

b) The fetal insulin hypothesis suggests that associations are largely due to the effects of genetic variants that have pleiotropic[3] effects influencing both fetal growth and later insulin resistance and hence cardiovascular risk (Hattersley and Tooke 1999).

c) Confounding (for example, by socio-economic position or shared familial behaviours such as smoking, physical activity and diet that could affect both perinatal outcomes and later disease risk in offspring), statistical artefact, and/or publication bias could also explain associations of birth weight with cardiovascular disease (Huxley et al. 2002; Tu et al. 2005). However, recent systematic reviews (Huxley et al. 2007; Whincup et al. 2008) together with intergenerational and sibling studies (Lawlor et al. 2009) have suggested that such bias and confounding is unlikely to fully explain these associations.

[3] when one gene influences multiple, seemingly unrelated, phenotypic traits

1.1.3 Relevant HAA literature

1.1.3.1 Experimental studies

Disinfection by-products were first discovered in drinking water in 1974 in the form of trihalomethanes (THMs) (Rook 1974). Toxicological evaluations were soon undertaken and first concern for public health arose. In 1976, a National Cancer Institute study was released in which chloroform was classified as a suspected human carcinogen (NCI 1976).

Kargalioglu et al. (2002) reported certain HAAs species (namely MCAA, DCAA, TCAA, BCAA, DBAA and TBAA) to be cytotoxic[4] and mutagenic[5] in *Salmonella typhimurium* strains, and that the brominated acetic acids (MBAA and DBAA) were more cytotoxic and mutagenic than their chlorinated analogues (MCAA and DCAA). They concluded that HAAs' mutagenic potency was inversely related to the number of halogen atoms of the molecule.

Itoh & Echigo (2008) performed chromosomal aberration tests using Chinese hamster lung cells and transformation tests using mouse fibroblast cells as indices to estimate the initiation and promotion, respectively, in the carcinogenesis process. They found that dichloroacetic acid (DCAA) and trichloroacetic acid (TCAA) contributed 2.9% of the chromosomal aberration-inducing activity and 1.4% to the transformation efficiency, while the contributions of MX (3-chloro- 4-(dichloromethyl)-5-hydroxy-2(5H)-furanone) and bromate ion for instance were almost negligible (less than 0.1%).

HAAs have also been shown to be cytotoxic and mutagenic in Chinese hamster ovary (CHO) K1 cells (SH Zhang et al. 2010), cytotoxic and genotoxic[6] in CHO AS52 cells (Plewa et al. 2002; Plewa et al. 2004; Plewa et al. 2010) and nontransformed human fetal hepatocyte (FH) cells (Attene-Ramos et al. 2010), and cytotoxic in human TK6 cells (Liviac et al. 2010).

In animal models, HAAs were teratogenic[7] in mice embryos (Hunter et al. 1996), and mixtures of THMs and HAAs were teratogenic in rats (Narotsky et al. 2011). High-dose prenatal exposure to HAAs has been found to cause fetal toxicity, including fetal resorptions and reduced fetal weight and survival (Graves et al. 2001; Nieuwenhuijsen et al. 2000a). Prenatal exposure to DCAA (Smith et al. 1992) and TCAA (Smith et al. 1989) in particular may have an adverse effect on fetal growth and development, as measured by fetal weight and length in Long-Evans rats exposed by oral intubation. However, these dose-dependent developmental effects occurred when the compound was administered at high doses, which can lead to treatment-related maternal effects such as substantial

[4] producing a toxic effect on cells
[5] inducing genetic mutation
[6] damaging to DNA and thereby capable of causing mutations or cancer
[7] causing malformations of an embryo or a fetus

inhibition to metabolically clear the compound (IARC 2004). This has important implications when considering the low doses present in drinking water.

Komulainen et al (2004) found that DCAA and TCAA are hepatocarcinogenic[8] in mice and that DCAA is hepatocarcinogenic in male rats. Their genotoxicity in this study was equivocal though, nongenotoxic mechanisms such as peroxisome proliferation and hypomethylation of DNA in the liver likely contribute to tumor development. DCAA has been classified as a probable human carcinogen (United States Environmental Protection Agency (US EPA) rating: B2) because of its cancer, reproductive and developmental effects, and TCAA as a possible human carcinogen (US EPA rating: C) because of effects on the liver, kidney, and spleen, and its developmental effects (Sadiq and Rodriguez 2004).

1.1.3.2 Human studies

Various thorough reviews have been conducted and have concluded that the relationship between HAA exposure and reproductive health outcomes remains unclear and inconsistent, mainly owing to limitations in exposure assessment (Nieuwenhuijsen et al. 2009a; Tardiff et al. 2006).

Klotz and Pyrch (1999) studied total HAA levels in water and neural tube defects and did not find any statistically significant association. Evidence for risk of hypospadias was inconclusive both in Kallen and Robert (2000) and in Luben et al. (2007), and a case-control study in Canada looking at the association between HAAs and stillbirth risk did not find any significant results after controlling for THM exposures (King et al. 2005).

Luben et al. (2007) studied the relation between exposure to classes of DBPs and sperm concentration and morphology, as well as DNA integrity and chromatin maturity, but found no association—or consistent pattern—of increased abnormal semen quality with elevated exposure to HAAs. Neither did Wright et al. (2004) find any statistically significant association between HAAs and a measure of mutagenicity (Nieuwenhuijsen et al. 2009a). Finally, on the basis of 258 pregnancy losses, Savitz et al. (2006) did not find an increased risk of pregnancy loss (spontaneous abortion) in relation to HAA concentrations—ingested amounts, or total exposure.

Studies on preterm delivery (PTD) have generally shown no statistically significant associations with HAAs (Hinckley et al. 2005; Hoffman et al. 2008a). However, elevated total HAAs and TCAA exposures were both associated with an increased risk of very preterm delivery (under 34 gestational weeks at birth) (Wright et al. 2004).

[8] linked to cancer of the liver

Study results on term low birth weight (LBW) have been mixed. A retrospective cohort study conducted in Arizona, USA, showed associations between term LBW and third trimester municipal drinking-water HAAs (third vs. first tertile total HAAs, adjusted OR=1.25 (95% CI: 0.96, 1.64)), especially dibromoacetic acid (DBAA) (adjusted OR=1.49 (1.09–2.04)) (Hinckley et al. 2005). Savitz et al. (2005) suggested that term birth weight was associated with HAA9, i.e. the sum of all nine HAAs (significance not stated). In contrast, not only did Wright et al. (2004) detect no associations for high third trimester HAA exposures and LBW in a retrospective cohort study in Massachusetts, USA, but also reported that birth weight increases were observed for intermediate total HAA and TCAA.

As for studies on small-for-gestational age (SGA) and/or intrauterine growth restriction (IUGR), Porter et al (2005) reported an increased risk of IUGR in the highest category of HAA exposure during the third trimester when comparing quintiles of exposure (fifth vs. first quintile HAA5, adjusted OR= 1.34 (1.04, 1.71)), but was unable to demonstrate any consistent statistically significant effect of HAAs on IUGR, nor find any indication of a dose-response relation. Levallois et al (2012) conducted a case-control study of 571 SGA births and 1925 controls and found that HAA levels above water standards (80ug/L in the US) were associated with an increased risk of term SGA. Wright et al (2004) found, if anything, a decreased risk of SGA with intermediate TCAA exposures (>18-27ug/l (second of 3 categories) vs. 0-18ug/l (reference), adjusted OR= 0.87 (0.76, 0.99)). Hoffman et al (2008b) did not report consistent evidence that exposure to HAA5 at residential concentrations below the current regulatory standards during the third trimester of pregnancy was associated with an increased risk of SGA. Finally, Horton et al (2011) found no association between HAA5 and term SGA, preterm birth, or very preterm birth (defined as less than 32 weeks of gestation at birth).

More recently, the first epidemiologic study relying on a biomarker of exposure to evaluate the impact of prenatal exposure to HAAs on pregnancy outcomes found that women with detectable TCAA in their urine (6.7% of the 611 maternal urines collected) had a higher risk of fetal growth restriction than those with TCAA levels below the detection limit (OR= 1.8 (0.9, 3.7)) and had an odds ratio for preterm birth below 1 (OR= 0.8 (0.3, 2.6)) (Costet et al. 2012). A cross-sectional study based in Wuhan, China, reported that subjects with creatinine-adjusted urinary TCAA concentrations in the top third and fourth quartiles had lower mean birth weights compared to those in other quartiles (only significant difference reported: -159.6 grams (-315.3, -4.0), when comparing the fourth to the first quartile on a subset of N=180) (Zhou et al. 2012).

In 2001, Graves et al (2001) reviewed all toxicological and epidemiologic to date and concluded that the weight of evidence suggested a positive association with DBP exposure for some measures of

growth restriction (such as IUGR or SGA) and for urinary tract defects, but none for low and very low birth weight, preterm delivery, congenital anomalies, or neonatal death.

1.1.3.3 Biological mechanisms

The biological molecule(s) with which the HAAs react and the mechanism(s) by which they induce toxicity, particularly at low concentrations, and ultimately poor birth outcomes remain unclear (Komulainen 2004; Pals et al. 2011).

One suggested mechanism is that HAAs interfere with folate metabolism. Dow & Green (2000) showed that TCAA interacts with vitamin B12, probably by a free radical mechanism, inhibiting both the methylmalonyl CoA[9] and methionine salvage[10] pathways in rats. As a result of the latter, a secondary folate deficiency develops, leading to a major impairment in formate metabolism (Alston 1991; Nieuwenhuijsen et al. 2009a). TCAA's role in inhibiting the vitamin B12-dependent methionine biosynthesis pathway could lead to vitamin B12 deficiency and consequently folate deficiency (Dow and Green 2000). Folate deficiency is known to be associated with increased risk of LBW and fetal growth restriction (Scholl and Johnson 2000) and is a probable risk factor for placental-mediated diseases such as pre-eclampsia, spontaneous abortion and placental abruption (Ray and Laskin 1999). Vitamin B12 deficiency has been shown to be associated with IUGR in humans (Muthayya et al. 2006).

Oxidative stress is another mechanism by which HAAs could have an effect on fetal growth. Maternal oxidative stress during pregnancy may be an important factor in adverse fetal growth (Karowicz-Bilinska et al. 2002; Kim et al. 2005; Matsubasa et al. 2002; Myatt and Cui 2004; Scholl and Stein 2001). DCAA and TCAA have been found to induce cellular death and oxidative stress in macrophage cells *in vitro* (Hassoun and Ray 2003), and to induce lipid peroxidation—a biomarker of oxidative stress—in mouse and rat livers *in vivo* (Larson and Bull 1992).

As for the cancer effects associated with HAAs, Pals et al (2011) highlights that HAAs have been considered direct-acting genotoxins because they are mutagenic in *S. typhimurium* without hepatic microsomal activation (Kargalioglu et al. 2002; Richardson et al. 2007). They induce genomic DNA damage and mutagenicity in CHO cells without exogenous cytochrome P_{450} activation (Plewa et al. 2004; Plewa et al. 2010; SH Zhang et al. 2010). However, new evidence is emerging that HAAs may not directly interact with genomic DNA. Elevated levels of 8-hydroxydeoxyguanosine (8-oxo-dG)[11]

[9] methylmalonyl CoA (CoA=coenzyme A) is an important intermediate in the biosynthesis of many organic compounds

[10] when homocysteine is recycled into methionine

[11] one of the major products of DNA oxidation (de Souza-Pinto NC, Eide L, Hogue BA, Thybo T, Stevnsner T, Seeberg E, et al. 2001. Repair of 8-oxodeoxyguanosine lesions in mitochondrial dna depends on the oxoguanine

in mice treated with chlorinated or brominated HAAs (Austin et al. 1995; Larson and Bull 1992; Parrish et al. 1996) suggest that HAA-mediated generation of reactive oxygen species is involved in the induction of toxicity and DNA damage. This means that a radical species may be involved in HAA-induced DNA damage (Cemeli et al. 2006; Pals et al. 2011).

1.1.3.4 Relevant windows of susceptibility

It is unclear when HAA exposure may have the most profound effect on a fetus' growth and development in the womb, whether it is during the first, second, or third trimester of pregnancy (Forssen et al. 2009). There is some evidence to suggest that adverse fetal growth outcomes may have their origins early in pregnancy (Smith 2004). Others believe that all three trimesters are critical (S. Cordier, personal communication, Gordon Conference 2012). As such I have considered the period when the data that make up this thesis' main exposure metric, the "combined metric" (see Chapter 5), is most robust because it corresponds to the trimester when the questionnaire was administered, i.e. the second trimester.

However, Hinckley et al (2005) suggests a critical window of exposure with respect to fetal development during weeks 33–40 for the effects of DBAA acid and during weeks 37–40 for the effects of DCAA. The rate of fetal growth and weight gain increases dramatically and reaches its peak at about week 33, i.e. during the third trimester of pregnancy (Owen et al. 1996; Williams et al. 1982). Average weight gain may reach almost 250g per week during this period (Williams et al. 1982). On this basis, as maternal exposure to high HAA levels during the third trimester may have an adverse effect on fetal growth, I have also tried to investigate water consumption variation over the course of the third trimester of pregnancy (see Chapter 8).

1.2 Gaps in knowledge

It is undeniable that disinfection of tap water is vital to ensure that water-borne diseases are eliminated from public drinking water supply. However, chlorine—the most common disinfectant used in the UK—is known to be toxic at high doses. Because vast segments of the population are directly exposed to these chemicals, it behoves environmental epidemiologists to address the health risks resulting from exposures to them and to provide guidance regarding appropriate regulations and societal investment in technology to reduce population exposure.

A lot of work has been done to date on exposure to trihalomethanes (THMs) because of their relative ease of analysis and the readily available data from routinely collected sampling for regulatory

---

dna glycosylase (ogg1) gene and 8-oxoguanine accumulates in the mitochondrial dna of ogg1-defective mice. Cancer research 61:5378-5381.)

purposes. But recent publications question the use of THMs as a proxy measure for DBP load, and highlight the lack of research on non-THM chlorination DBPs (Nieuwenhuijsen et al. 2009a). THMs are an insufficient proxy for DBPs as a whole because a) the metabolism of different DBP species varies (IPCS 2000), b) the toxicity of different DBP classes varies (Jeong et al. 2012), c) specific DBPs within a particular class have substantially different toxicities (Hunter et al. 2006), and d) the relationship of THM concentrations to other DBP concentrations varies and is more often than not unknown, as more than 50% of all organo-halogenated DBPs in water formed by chlorination remain chemically unidentified (Nieuwenhuijsen et al. 2009a; Richardson et al. 2007).

As a result, despite being the second most abundant DBP in drinking water (Singer et al. 2002), little is known about HAA occurrence in UK treated drinking waters (Bougeard 2009; Malliarou et al. 2005; Y Zhang et al. 2010), and even less about their possible health effects. More epidemiologic data on this subject are needed for policy purposes.

Exposure assessment is the Achilles' heel of environmental research. Exposure assessment in many of the epidemiologic studies published to date has been inadequate to definitively demonstrate an association of small magnitude. Exposure to DBPs has been primarily based on routine (i.e., quarterly) monitoring of public water supplies for trihalomethanes (THMs) matched to maternal residence. In order to determine whether an association exists between exposure to DBP and adverse birth outcomes, studies must consider both DBP concentration and the volume of water that each individual pregnant woman is individually exposed to (Graves et al 2001).

This thesis uniquely achieves these goals, combining information on individual water use (described in Chapter 3) with modelled area-level HAA concentration estimates based on HAA concentration data collected and analysed deliberately for this work (Chapter 4), and accounting for filtering and boiling to generate the most precise exposure assessment measure currently possible (Chapter 5) (Table 1.4). Because within-subject variability in questionnaire data may be substantial and may attenuate risk estimates, I further evaluated individual information for measurement error in a separate repeat questionnaire study (RQS) (Chapter 8). All with the goal of understanding exposure in ever finer detail in order to produce the best possible models to investigate the reproductive health effects of HAAs (Chapter 7).

Bradford was chosen as study site for the Born in Bradford (BiB) cohort because of its classification as the eighth most deprived health community in the UK, with high levels of morbidity and a standardised mortality rate above the UK average (see Chapter 2 for more details of the cohort). In addition, approximately half of the cohort is of South Asian origin, making this a unique study population. Investigating the role that environmental factors play is an important component of

addressing these poor health outcomes, and my expectation is that the findings of this thesis, in conjunction with others, could be of real benefit to this community. Chapter 6 is a descriptive chapter on BiB and birth outcomes.

In sum, this work is important because of its potential policy ramifications and for the benefit of the population of Bradford; it is novel because of its focus on understudied HAAs, presenting new concentration data from the recently completed HiWATE project; and it is methodologically innovative in its eagerness to go beyond what previous studies have achieved to date and improve exposure assessment by deriving a metric of exposure combining individual and areal-level information, and assessing possible individual-level measurement error in a repeat questionnaire study.

****

All tables are listed first in order of citation, followed by the figures listed in order of citation. Tables and figures are numbered including chapter number. Thus Table 2.8 is the eighth table to appear in Chapter 2. Any table or figure relating to a given chapter but deemed of secondary importance was relegated to the appendix, and numbered by chapter number preceded by an "A" for appendix, e.g. Appendix Figure 5 from Chapter 3 will be numbered Figure A3 – 5 and located in the Appendix section dedicated to Chapter 3. Once more, each chapter's tables precede its figures, and is listed in order of citation.

1.3 Tables

Table 1.1: Chemical and physical properties of nine HAAs (modified from a table in Bougeard (2009))

Abbreviation	Name	Molecular Formula	Boiling point (˚C)	Boiling point of ester (˚C)
MCAA *	Chloroacetic acid	$C_2H_3ClO_2$	189	130
MBAA *	Bromoacetic acid	$C_2H_3BrO_2$	206-208	132
DCAA *	Dichloroacetic acid	$C_2H_2Cl_2O_2$	194	143
BCAA	Bromochloroacetic acid	$C_2H_2BrClO_2$	215	174
TCAA *	Trichloroacetic acid	$C_2HCl_3O_2$	196	168
DBAA *	Dibromoacetic acid	$C_2H_2Br_2O_2$	128-130	NR
BDCAA	Bromodichloroacetic acid	$C_2HBrCl_2O_2$		NR
DBCAA	Dibromochloroacetic acid	$C_2HBr_2ClO_2$	NR	NR
TBAA	Tribromoacetic acid	$C_2HBr_3O_2$	245	225

* regulated as a sum in the US
NR: not reported

Table 1.2: Regulatory thresholds for THMs and HAAs in the European Union (incl. the UK), US, and Australia and New Zealand, as well as WHO guidelines
MCL: maximum allowable contaminant level
TTHM: total trihalomethanes, i.e. the sum of the four known individual THMs

	THM	HAA
European union standards (UK)	MCL for TTHM at 100ug/L in any sample	Not regulated (DCAA MCL: 50ug/L and TCAA MCL: 200ug/L under consideration)
US EPA regulations	MCL for locational running annual averages of 80ug/L for TTHM	MCL for locational running annual averages of 60ug/L for HAA5*
Australia and New Zealand	TTHM of 250ug/L (guideline value)	MCAA, DCAA and TCAA in drinking water are 150, 100 and 100ug/L, respectively
WHO guidelines	300ug/L for chloroform, 60ug/L for BDCM, 100ug/L for DBCM and 100ug/L for bromoform	50ug/L for DCAA (provisional) and 200ug/L for TCAA

* Sum of MCAA, MBAA, DCAA, TCAA, DBAA

Table 1.3: Definitions of common birth outcome measures described in Introduction (Chapter 1)

Birth outcome measure	**Abbr.**	**Definition**
Birth weight	BW	continuous measure of weight at birth (g)
Low birth weight	LBW	birth weight <2500g (irrespective of GA)
Very low birth weight	VLBW	birth weight <1500g (irrespective of GA)
Term low birth weight	LBW	birth weight < 2500g after at least 37 completed weeks of gestation
Gestational age	GA	Number of completed weeks at birth
Preterm delivery	PTD	gestational age at birth <37 completed gestational weeks
Very preterm delivery	VPTD	gestational age at birth <32 (or sometimes <34) completed gestational weeks
Small-for-gestational age	SGA	live infant born below the tenth percentile (or fifth or third) or at least two SD lower than the mean BW of BW for GA by sex (and/or other characteristics) in a referent population (Alexander et al. 1999; Choi et al. 2008; RCOG 2002)
Intrauterine growth restriction	IUGR	fetal growth ratio below the tenth percentile of a referent population (Smith et al. 1997)
Customised small-for-gestational age	cSGA	live infant born below the tenth percentile based on optimal growth curve for an individual fetus (Gardosi 2006)
Fetal growth restriction	FGwR	live infant born below the tenth percentile of the predicted BW (Mamelle et al. 2001)

Table 1.4: Sixteen exposure measures presented in this thesis' analyses

exposure type: (units)	**individual water consumption** (L/day)	**modelled area-level concentrations** (ug/L)	**combined exposure metric^** (ug/day)
trimester of pregnancy:			
first		DCAA ✓ TCAA ✓ BDCAA ✓	
second	cold tap water ✓ total tap water ✓ bottled water ✓ total water ✓	DCAA ✓ TCAA ✓ BDCAA ✓	total tap water x DCAA ✓ total tap water x TCAA ✓ total tap water x BDCAA ✓
third		DCAA ✓ TCAA ✓ BDCAA ✓	

^In addition to multiplying total tap water consumption by trimester-weighted area-level HAA concentrations, this measure incorporates factors for filtering and boiling, either at the home, the work place or both; as water consumption is enquired about at recruitment to BiB during women's second trimester of pregnancy, the combined metric uses second trimester area-level concentrations.

CHAPTER 2 BORN IN BRADFORD (BIB) & AIMS

This chapter describes the Born in Bradford (BiB) cohort which is at the heart of this thesis, the data available, as well as this thesis' main aims and objectives.

2.1 Born in Bradford cohort

The Born in Bradford (BiB) study was established in 2007 in Bradford, an industrial city in the North of England, which is the sixth largest city in the UK with a population of about half a million. Its aims are to examine how genetic, nutritional, environmental, behavioural and social factors impact on health and development during childhood and subsequently adult life in a deprived multi-ethnic population.

BiB is a longitudinal birth cohort study that involves research collaboration between Bradford Teaching Hospitals, Bradford & Airedale PCT, University of Leeds and the University of Bradford. Additional research partner include the University of Bristol, University of Loughborough, University of Edinburgh, London School of Hygiene and Tropical Medicine and Imperial College London. The project has received support from a number of funders including MRC, NIHR, Diabetes UK and the Department of Health.

The full study methodology is available at http://www.borninbradford.nhs.uk/[12]. The BiB study protocol and a detailed cohort profile have been published (Raynor 2008; Wright et al. 2012). Additional details on the study population and recruitment process are described in West (2011).

2.1.1 Why was the BiB cohort set up?

In response to rising concerns about the high rates of childhood morbidity and mortality in the city of Bradford, an independent Commission of bereaved mothers, politicians, members of voluntary organisations, health and other public service professionals was established in 2004. It commissioned an extensive analysis of local data, which was published as the Bradford District Infant Mortality Commission (BDIMC) report in 2006. Updates to this report have been published in 2008 and 2011.

Bradford was chosen as study site for the BiB cohort because of its classification as eighth most deprived health communities in the UK (APHO 2008; BDIMC 2006). Infant mortality in Bradford is

[12] last accessed 18/01/2014

consistently above the national average; it peaked at 9.4 deaths per 1000 live births in 2003, when the national average was 5.5 deaths per 1000 live births (Wright et al. 2012). A greater proportion of babies born in Bradford are of low birth weight (9.7%), an outcome typically associated with increased infant mortality and morbidity and an increased risk of developing various diseases in later life (see section 1.1.2.3)—compared with England and Wales as a whole (7.5%) (BDIMC 2006). Levels of congenital anomalies and childhood disability in Bradford are also among the highest in the UK (Wright et al. 2012).

2.1.1.1 Bradford's unique demographics

Around 20% of the population of Bradford is of South Asian origin (90% of whom are from Pakistan), a three-generation community that maintains close links with Pakistan (Small 2012). The relatively young age of the population of Pakistani origin and their higher fertility rates compared with the White British majority population explain why almost half of babies born in the city of Bradford have parents of Pakistani origin (BDIMC 2006).

Infant mortality rates among South Asians living in the UK are considerably higher than those of the UK White population. Between 1996 and 2003, infant mortality in Bradford for babies of Pakistani origin (12.9 per 1000) was substantially higher than for those of European origin (7.1 per 1000) (RR for Pakistani vs. White 1.83, 95% CI: 1.52, 2.20) (Raynor 2008).

Sixty percent of the babies born in the city are born into the poorest 20% of the population of England and Wales based on the British government's residential area Index of Multiple Deprivation (IMD) (see description of this indicator below in section 2.1.2.3.2).

2.1.2 Recruitment to BiB

Bradford has one maternity unit based at Bradford Royal Infirmary (BRI), which is located in the Girlington area. The unit is one of the busiest in the UK with over 6000 new births each year. Figure 2.1 shows a map of the catchment area of the BRI antenatal clinic, and the larger Bradford District which includes the city of Bradford and the smaller towns of Keighley, Ilkley, Bingley and Shipley.

Because incidence of gestational diabetes in Bradford is high, women in Bradford are offered an oral Glucose Tolerance Test (GTT) at 26-28 weeks gestation. All women who visited the hospital on this occasion were invited to join the BiB study (unless they were planning to move away from Bradford before the birth). Babies whose delivery was booked outside of the District or who were born before 26 weeks gestation were not captured in BiB. Additional criteria for exclusion from analyses are described in the relevant chapters, notably Chapters 4, 6 and 7.

Study information was handed out and full written consent was obtained for recruitment to BiB and for use of their data, including data specifically collected for the BiB study and other data obtained through linkage with medical records. The recruitment process was carried out by the BiB study team (PI: Dr John Wright).

More than 80% of the women who attended their GTT between March 2007 and December 2010 took up the invitation to participate and were recruited to the BiB cohort, for a total of 12,453 women with 13,776 pregnancies. (All babies born to women who agreed to participate in the cohort study were eligible for recruitment.) (Wright et al. 2012).

2.1.2.1 Interview questionnaire data

After consenting to participate, women were invited to complete a comprehensive face-to-face questionnaire administered by a trained interviewer (bilingual in Urdu, Punjabi or Mirpuri as necessary). This baseline questionnaire included a section on demographic information (residential and work addresses, age, ethnicity), as well as sections on employment, education, smoking/alcohol/drug consumption, physical exercise, family ancestry, diet, and tap water consumption (for the water consumption section of the baseline questionnaire, see Appendix A).

11,396 women (92% of total recruits) completed a baseline questionnaire (Wright et al. 2012).

Detailed information regarding smoking, alcohol and drug use during pregnancy was obtained via the questionnaire which included details of exposure to other people's cigarette smoke at home work or at work, and about other tobacco products such as Paan. Alcohol information included intake prior to pregnancy and during early and later stages of pregnancy.

During the questionnaire interview, trained project workers recorded the mother's height using the Leicester Height Measure (SECA Ltd., Birmingham, UK) and weight using SECA digital scales (SECA Ltd., Birmingham, UK) with outdoor clothing and shoes removed. The BMI calculation, defined as maternal weight in kilograms divided by maternal height in metres squared, is based on this measurement of maternal weight at the time of questionnaire completion, as pre-pregnancy weight was not available.

Here are a few more details on the derivation of select variables from the baseline questionnaire used in future chapters.

2.1.2.2 Ethnicity

Ethnic categorisation was based on self-defined ethnicity in the questionnaire. Throughout this thesis the term 'White British' includes those who originate either from the UK or Ireland. The term 'South Asian' refers to people originating from the Indian subcontinent i.e. India, Pakistan, Bangladesh and Sri Lanka, keeping in mind that the great majority of BiB participants of South Asian origin are from Pakistan. I chose to focus on the three most prevalent ethnic groups: White British women, women of Pakistani origin and women of any other ethnicity ("Other"). There were insufficient numbers of women in the Black, Indian or Bangladeshi subgroups to justify studying them individually. For details on how the ethnicity information collected in the baseline questionnaire was classified into these three most prevalent groups, see Table 2.1.

2.1.2.3 Measures of socio-economic status (SES)

2.1.2.3.1 Maternal education

Women were asked for details of their and their partner's education including the age at which they left full time education, their highest educational qualification and which country they received most of their education. Table 2.2 clarifies the qualifications required by each category of the derived variable used in analyses.

2.1.2.3.2 Index of Multiple Deprivation (IMD)

The English Indices of Deprivation 2010 are measures of deprivation calculated by the government at the Lower Super Output Area (LSOA) level. Seven distinct domains of deprivation (income deprivation, employment deprivation, health deprivation and disability, education skills and training deprivation, barriers to housing and services, living environment deprivation, and crime) can be measured separately or combined using appropriate weights into a single overall Index of Multiple Deprivation (IMD). The IMD 2010 can be used to rank each of the 32,844 LSOAs in England according to the deprivation experienced by the people living there, quintiles of which (from most deprived to least deprived) simplify its use in epidemiologic studies (Lad 2011).

While the IMD 2010 is a valuable area-level measure of deprivation frequently used in ONS statistics as the Carstairs Index (Carstairs and Morris 1991) was before it, maternal education is often used as an individual-level proxy measure for maternal socio-economic status.

2.1.2.3.3 Income data

The BiB team attempted to collect income data, but quickly discovered that these data were simply too unreliable to use. For one, 36% of Pakistani women simply did not know their household income, compared to 6% of White British mothers (West 2011). Secondly, given the complicated and changing nature of tax and benefits in England, different questions (pre-tax income, then post-tax

income) were asked in different phases of the questionnaire, making the responses difficult to compare.

2.1.2.4 Caffeine intake during pregnancy

In the food section of the questionnaire, women reported: daily cups of instant coffee (caffeinated and decaffeinated), filter/cafetiere coffee (caffeinated and decaffeinated), tea per day (caffeinated and decaffeinated), Kashmiri tea (caffeinated and decaffeinated), herbal tea (caffeinated), cola and diet cola (caffeinated and decaffeinated). The caffeine variable was calculated by linking women's responses to questions about caffeine ingestion in the baseline questionnaire to typical caffeine contents in beverages published by the Committee on Toxicity (COT) and summarised in Table 2.3.

COT values of 4mg per cup were used for decaffeinated beverages. Total caffeine was then used as a binary variable using the cut-off of 200mg proposed by the CARE study group (Konje et al. 2008). Derivation of this variable was done by Emily Petherick (Born in Bradford, Bradford Institute for Health Research) and Rachel Smith (Imperial College London).

2.1.3 Follow-up of BiB participants

2.1.3.1 Birth weight

Following delivery and prior to hospital discharge, neonatal anthropometric measurements were obtained, including birth weight. Birth weight is routinely recorded by the midwife at delivery using SECA digital scales as is standard clinical practice in Bradford, and entered by the midwife into a routine data maternity system called eClipse.

The total number of BiB pregnancies with eClipse data is 13,525 (98% of total pregnancies).

Babies in the BiB project are followed by a health worker throughout their childhood (starting at ~2 weeks, 7 weeks, and 8 months after birth).

2.1.3.2 Routinely collected maternity information

The Bradford NHS works with an electronic maternity care records system (eClipse) which was accessed by the BiB research team to obtain routine clinical data for participating mothers and babies and to validate information collected by the questionnaire. This included medical and obstetric information (maternal age, parity, maternal diabetes, and hypertensive disorders of pregnancy) and perinatal data (gestation at delivery, baby's gender and birth weight).

Maternal diabetes was categorised as existing diabetes or gestational diabetes based on clinical diagnosis. Gestational diabetes, formally defined as any degree of glucose intolerance with onset or first recognition during pregnancy and which continues beyond 24-28 weeks of gestation (Metzger and Coustan 1998), was determined based on the GTT conducted at ~26 weeks of pregnancy at BRI (see section 2.1.2). Where the test was not done, gestational diabetes status was assigned using the backfill data (<5% of cases). The 'backfilling' process was conducted by a physician who extracted these data from the written hospital notes. Within the Bradford Teaching Hospital NHS Trust to which the Bradford Royal Infirmary belongs, gestational diabetes was diagnosed using standard WHO thresholds for impaired glucose tolerance or impaired fasting glucose (fasting plasma glucose ≥6.0mmol/l and/or post challenge glucose ≥7.8mmol/l).

Hypertension (high blood pressure) and pre-eclampsia (high blood pressure with protein in the urine) data were also not routinely collected and therefore backfilled. Hypertension in pregnancy was classified as mild to moderate (≥140 systolic and 90 diastolic on 2 or more occasions) or severe (≥150 systolic and 105 diastolic on 2 or more occasions).

Gestational age at delivery was calculated by the attending midwife and entered into the eClipse data system. At BRI, the expected date of delivery (EDD) was based on the date of the mother's last menstrual period (LMP). This was then confirmed by a dating ultrasound scan at time of booking, around 12 weeks of gestation. If the ultrasound dates are within 7 days of the menstrual dates, date of LMP was used. If the difference is greater than 7 days, the ultrasound date is used (Bradford Teaching Hospitals NHS Foundation Trust 2005). If scan data were not available—most commonly when a woman booked her pregnancy late—gestational age was based on the date of LMP. Attempts to acquire ultrasound data themselves were made but were unsuccessful due to lack of resources to hire the medical professional needed to read in and digitize individual scans.

As part of a quality control exercise, out-of-range values were checked against the hospital notes by the BiB team. Previous work on child growth in this cohort showed that health workers are reliable in their routine antenatal (Johnson 2009; West et al. 2011) and postnatal anthropometric measurements (Johnson et al. 2009).

The codebook for variables considered in the model selection process and in the final models as well as their categorisation is available in the appendix (Table A2 - 1).

2.1.3.2.1 Data extracts

The BiB team sent us several versions of eClipse data extracts over time: I received the first eClipse dataset at Imperial College London (ICL) on February 1st, 2012; it was then updated on December 7th,

2012. This is the data extract used in the work presented in Chapters, 3, 4, 5 and in Chapter 8. A final data extract including updated maternal weight, height, and gestational diabetes status was sent to me on February 1st, 2013. This is the data extract used in Chapters 6 and 7.

2.2 Thesis Aims

This PhD thesis has three main aims:

1. To generate exposure estimates to ingested dichloroacetic acid (DCAA), trichloroacetic acid (TCAA) and bromodichloroacetic acid (BDCAA) for each trimester of pregnancy of each BiB participant for use in epidemiologic analyses of birth outcomes. This is done by combining information on:

    a. each individual woman's total tap water consumption during pregnancy collected via baseline questionnaire (at approx. 26-28 weeks of pregnancy), with

    b. area-level DCAA, TCAA and BDCAA concentration data (respectively) collected quarterly under the HiWATE project (June 2007-November 2010), modelled for the period February 2007 through February 2011 for Bradford's eight water supply zones using DBP determinants data provided by Yorkshire Water and ALcontrol, and weighted to each woman's specific trimester of pregnancy by postcode of residence (via GIS linkage)

2. To examine the epidemiologic association between prenatal exposure to each of DCAA, TCAA and BDCAA—as estimated by the combined metric generated in Aim 1—and birth weight, term low birth weight, and small-for-gestational age as measures of adverse fetal growth, adjusting the analyses for potential confounders

3. To investigate water use patterns in the third trimester of pregnancy by evaluating the agreement of individual water use values reported in BiB questionnaires at baseline (at approx. 26-28 weeks of pregnancy) and at two later time points in pregnancy in a subset of BiB women

2.3 Tables

Table 2.1: Ethnicity classification

Original classification	Classification used in models (3 most prevalent groups)
White British	White British
White Other	Other
Mixed White and Black	Other
Mixed White and South Asian	Other
Black	Other
Indian	Other
Pakistani	Pakistani
Bangladeshi	Other
Other	Other

Table 2.2: Maternal education classification (equivalency carried out by Rachel Smith)

Category	Grades included
No education	4 or fewer O-levels/Certificate of Secondary Education (CSEs)/ General Certificate of Education (GCEs) (any grades)
	National Vocational Qualification (NVQ) level 1
	Foundation General National Vocational Qualification (GNVQ)
School	5 or more O-levels
	5 or more CSEs (grade 1)
	5 or more General Certificate of Secondary Education (GCSEs) (grades A-C)
	School certificate
	NVQ level 2
	Intermediate GNVQ
Further education	1 or more A-levels/AS-levels
	2 or more A-levels
	4 or more AS-levels
	Higher School Certificate and NVQ Level 3
	Advanced GNVQ
Higher education	NVQ Levels 4-5
	Higher National Certificate (HNC)
	Higher National Diploma (HND)
	First degree (e.g. Bachelor of Arts, Bachelor of Science)
	Higher degree (e.g. Master of Arts/Science, PhD, Post-graduate certificate in Education (PGCE), Post-graduate certificates/ diplomas)
Other, don't know, and Unknown foreign	Overseas qualifications (whose equivalence could not be established)
	Any other unclassifiable entry

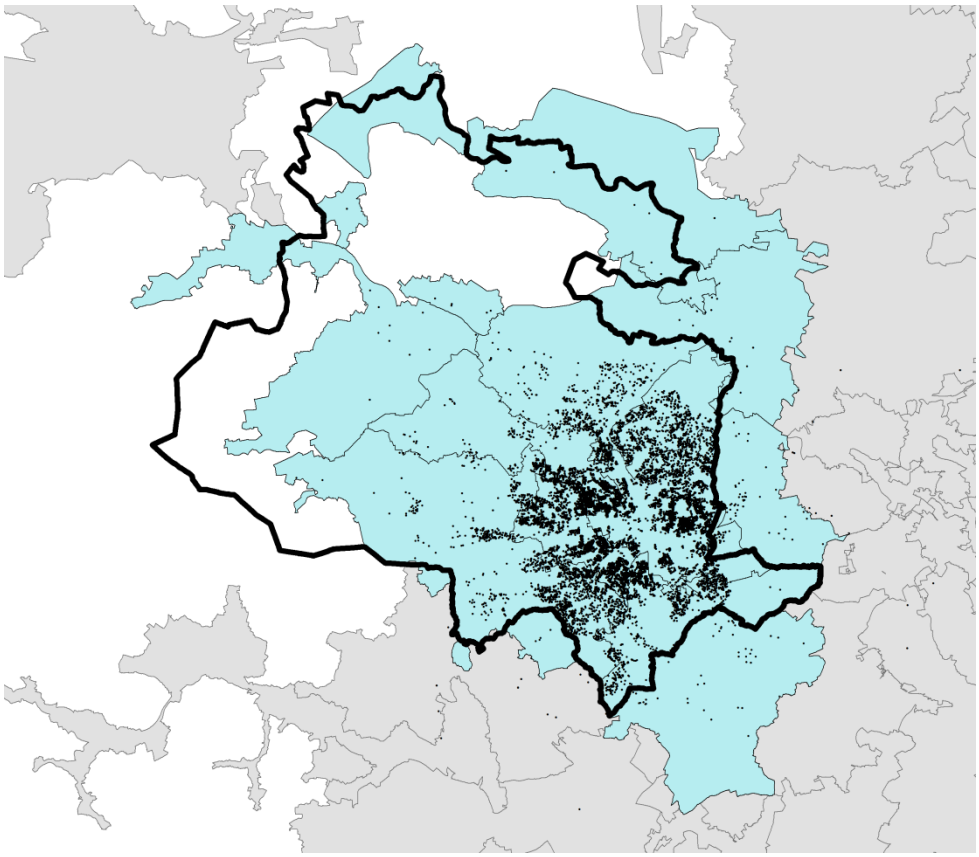
Table 2.3: Caffeine contents in beverages (COT: Committee on Toxicity) (conversions carried out by Emily Petherick and Rachel Smith)

Source	**Approximate typical caffeine content (mg) per serving**	**mg per COT unit (median if range given)**
Instant coffee	75 mg per 190 ml cup*	75
Brewed coffee (filter or percolated)	100-115 mg per 190 ml cup*	107.5
Decaffeinated coffee (brewed or instant)	4 mg per 190 ml cup*	4
Tea	50 mg per 190 ml cup*	50
Cola (regular and diet)	11 - 70 mg per 330 ml (one can) serving, i.e. 24.5 mg per 200 ml	40.5

* treated as equivalent to the 200ml cup volume referred to in the BiB questionnaire

2.4 Figures

Figure 2.1: Map of Bradford Metropolitan district and Bradford Royal Infirmary catchment area (within thick black border), water supply zones (8 blue areas), rest of Yorkshire (grey), BiB women's residences (black dots)



CHAPTER 3 INDIVIDUAL WATER CONSUMPTION

This chapter assesses the contribution of water consumption to total HAA exposure. As per previous studies pregnant women reported consuming cold tap water more than any other water type, with the greatest proportion of water consumption occurring in the home as opposed to other locations. By and large there is quite a bit of variability in individual water consumption.

3.1 Background

3.1.1 Water consumption during pregnancy

Consuming tap water (as well as beverages and foods prepared using tap water) allows for potential ingestion of HAAs contained in disinfected tap water. Estimates of average tap water consumption amongst pregnant women range from 0.6L/day (Egorov et al. 2003) to 1.8L/day (Zender et al. 2001) (Table 3.1).

Kaur et al. (2004) reported estimates of total tap water consumption for pregnant women in the UK at 1.3L/day (converted from reported weekly consumption by Smith in her PhD thesis (2011)). In the US, average tap water consumption was 0.9L/day among women of childbearing age (15-44 years) in the 1994–1996 Continuing Survey of Food Intakes by Individuals (EPA 2000) and 1.2L/day among pregnant women in the 1977–1978 National Food Consumption Survey (Ershow et al. 1991).

The proportion of tap water relative to total water consumed varies by country however, based on perception of water quality and cultural habit. For example, Barbone et al. (2002) examined water use during late pregnancy in relation to water use after delivery among pregnant women in Italy. They reported an average total ingested water estimate (tap and bottled) of 2.7 L/day, and an average tap water of only 0.6L/day (Table 3.1).

Drinking bottled water may act as a substitute for some or all tap water consumption. However, it is generally assumed that bottled water does not contain any HAAs (Savitz et al. 2005). Estimates of bottled water consumption amongst pregnant women vary from 0.6L/day to 0.9L/day (Forssen et al. 2007; Kaur et al. 2004), with Kaur et al. (2004) finding ~80% of pregnant women in a UK sample reported drinking some bottled water. The 2008 Drinking Water Inspectorate (DWI) survey on tap water consumption however found a fairly even spread of people who drink bottled water across England and Wales, with the exception of Yorkshire (where the BiB cohort is based) where only 26% consumed bottled water, compared to the average of 46% (DWI 2008).

Previous studies have also shown that tap water consumption in pregnant women tend to differ by ethnicity (Forssen et al. 2007), income (Forssen et al. 2009), and age (DWI 2008).

The aim of this chapter is to describe the individual water consumption component to the overall exposure metric I am building towards in Chapter 5.

3.2 Methods

3.2.1 Collection of demographic, lifestyle data, and water use data

At recruitment a detailed baseline questionnaire was administered by researchers to collect a wide range of data on the women's lifestyle, environment, ethnicity, and health (Raynor 2008).

Three versions of the baseline questionnaire have been used as updates were implemented over time: a pilot version administered between March 2007 and October 2007, a second version used from September 2007 to July 2009, and a final version used from May 2009 to December 2010. The data used in this study include versions 1, 2 and 3 of the questionnaire.

The baseline questionnaire includes a set of questions meant to ascertain typical daily consumption of tap water, bottled water, tea, coffee, and squash (which includes any other drinks made with tap water) at home, work/study, or elsewhere; and water filtering habits at home and work. The water use section of the questionnaire is available in Appendix A.

All relevant variables from the mother's baseline questionnaire, including from the water section of the questionnaire, were received from the BiB team in February 2012.

3.2.2 Deriving water consumption variables

Cups and glasses were converted to litres by multiplying by 0.2 (the relation 1 cup or glass = 200ml was printed in the water questionnaire).

3.2.2.1 Summary measures

To reduce the number of variables to more meaningful groupings, I applied the following to each of the three locations of interest:

Cold tap water (L/day) = tap water (L/day) + squash (L/day)

Hot tap water (L/day) = tea (L/day) + coffee (L/day)

Total tap water (L/day) = Cold tap water (L/day) + Hot tap water (L/day)

Total water (L/day) = Total tap water (L/day) + Bottled water (L/day)
(If two values were missing, their sum was marked as missing, not as zero.)

The locations of interest are the home, any other place besides the home (i.e. work/study + elsewhere locations), and all three locations combined ("all locations").

For comparative purposes, I was also interested in looking at HAA-free bottled water consumption.

The "All locations" water consumption measures for women who reported being employed at recruitment (whether already on maternity leave or not) (N=4,331) are sums of the volumes these women reported drinking at three locations: home, work and elsewhere. However for women who do not work, for full-time students, and for women for whom employment information was missing (N=7,597), they are only sums of the water reported to be consumed at two locations: home and elsewhere (Table 3.2).

Because the questionnaire enquires after water consumption habits during the full pregnancy to date, I decided to group the responses of employed women on maternity leave with those of the employed women who currently worked, in order to reflect a woman's status during the majority of her pregnancy, i.e. the period preceding leave. In other words, maternity leave (N=383) was assumed to be very recent by the time of recruitment to BiB (late second trimester).

While the water questionnaire did enquire after tap water consumption at work *or study*, it was deemed too uncertain where a student would be spending most of her work/study time as this may include a variety of locations, including the home. Students (N=339) were therefore grouped with the women out of work—who presumably do not spend all of their time in the home either. The impact of this decision is assumed to be minimal as only 20% of students reported drinking any tap water at work/study (and only 15% reported drinking any cold tap water specifically), and only 18% of these students (60 students, <0.1% of the total sample size of 11,928) reported any filtering habits at work/study.

3.2.2.2 Tap water filtering

Filtering of tap water is known to modify HAA concentrations affecting the ingested dose of HAAs and must therefore be taken into consideration.

Two questions in the baseline questionnaire specifically address the issue of filtering:

1. "Do you filter the water you drink at home?" (yes, no, don't know)
2. "Do you filter the water you drink at work?" (yes, no, don't know, NA (not available))

(The NA answer option to the filtering at work question was intended for the unemployed women, but some unemployed women nevertheless filled out the work filtering question leading to some inconsistent answers which were excluded.)

Both cold tap water types—tap water and squash—were considered to be filtered if a woman answered yes to either of the above questions.

Because cold tap water consumed outside the home or at all locations are aggregate measures—i.e. grouped answers to questions on different water types consumed at different locations— and given that filtering must be taken into account for two possible locations (home and workplace), the cohort was divided into six mutually exclusive subgroups: four groups for the employed, and two for the unemployed at recruitment as follows:

Employed women (on maternity leave or not, N=4,331) who filtered tap water:

1. at home and at work
2. at home only
3. at work only
4. Neither filtered at home nor at work

Women out of work, students, or women for whom employment information is missing (N=7,597) who:

5. Filtered tap water at home
6. Did not filter tap water at home

The various combinations of answers to the two filtering questions were then recognised for the employed women (Table 3.3, a) and for the unemployed women (Table 3.3, b).

Those who answered yes to filtering either at home, at work, or both places (dashed cells in Table 3.3) were defined as the filterers for the purposes of the summary table (Table 3.4), while those who answered no to either or both questions (grey cells) were non-filterers.

Summary Table 3.4 includes only women who actually reported consuming the given water type (i.e. zero values were excluded from each mean) because I was interested in average water consumption among those who consumed or reported consuming water. I did not want to include all those who women who do not drink, for instance, bottled water in the average bottled water consumption value as this would not be representative of average bottled water consumption among bottled water drinkers.

A number of women had inconsistent answers and were grouped in an “inconsistent” category. This category includes women (whether or not in employment) who reported filtering at home (and/or work) but didn't report drinking cold tap water at home (and/or at work) (consumption=0 or missing) (Table 3.5).

3.2.3 Exclusion criteria

In the following summaries and analyses, multiple births were systematically excluded. Of 13,199 singleton babies, only one of each multiple pregnancy registered in BiB (19% of BiB participants registered multiple births in BiB) was randomly selected so as to avoid having non-independent observations from the same mother: 1,199 mothers registered two separate singleton pregnancies and 36 women registered three separate pregnancies between 2007-2010. The total sample size left is N=11,928 (Figure 3.1).

There were 412 women for whom health data were not available. These were cases of women who did not give birth in Bradford. Though all women intend to give birth in Bradford at the time of recruitment, not all of them do.

3.3 Results

The majority (82%) of BiB women report consuming some tap water (Table 3.5). BiB cohort women drink on average 1.7 ± 0.9 L/day tap water and 1.9 ± 1.0L/day total water (Table 3.4). Approximately 60% of total water by mean volume is cold tap water and 30% hot beverages made from tap water. Variability in these means is non-negligible. The distribution of water consumption in all locations are right skewed as the histograms in Figure 3.2 (with tertile cut points which are used in the later analyses marked) show.

The majority of tap water consumption occurs in the home; 9,679 women reported consuming tap water at home vs. only 2,715 at any location outside the home (out of a total N=12,394).

The proportions of women who reported water consumption at Home vs. Outside the home are broken down in the appendix (Table A3 - 1, Table A3 - 2).

83% (N=5,510) of the 72% (6,639 of 9,193) of women who answered at least one of the filtering questions reported not filtering their tap water (Table 3.5, Figure 3.3). Table 3.5 is broken down by location in Table A3 - 1 and Table A3 - 2. In the US, 53% of the pregnant women attending North

Carolina obstetric clinics from 1994–1995 also reported drinking mostly unfiltered tap water (Shimokura et al. 1998), and 11% of the pregnant women from the Colorado study reported primarily using filtered water (Shimokura et al. 1998; Zender et al. 2001). Overall, BiB women who drink filtered tap water drink slightly more of it than those who drink unfiltered tap water (+0.24 L/day on average, equivalent to approximately 1 glass/day) (Table 3.4).

The proportion (28%) of women drinking bottled water was very low in the BiB cohort compared to consumption of bottled water by 80% of pregnant women as reported by Kaur et al. (2004). This being said, in the US, 24% of the pregnant women from the North Carolina study and 14% of the pregnant women in the Colorado study reported bottled water as their main source (Shimokura et al. 1998; Zender et al. 2001). This confirms that there are important geographic differences in self-reported water consumption habits.

3.4 Discussion

While in total BiB women drink on average less tap water than the women in the Kaur et al. (2004) central London study, the ratio of tap water to total water drunk is much greater in BiB than in the Kaur study. This may be due to the fact that these means only include the responses of women who actually reported consuming the given water type (i.e. zero values were excluded from each mean) such that mean water consumption of each water type is overestimated. Or perhaps this difference is due to the fact that Kaur et al. (2004) recruited women in their first trimester of pregnancy such that a greater proportion of women might still have been working and/or mobile and spending more time outside of the home where less tap water consumption occurs compared to BiB women. Life in a large city might be also more conducive to consuming water from sources other than the tap, but ultimately I conclude as did Kaur et al. that individual variability in water consumption is considerable (whether due to the measurement tool and recall bias, or to true differences in habits between women) and a critical factor to consider when assessing exposure to chlorinated water.

Of note, I refer to the sum of all water consumption categories from the water questionnaire as "total water", understanding that this does not include water included in other beverages which make up a woman's typical diet such as soda or juices. Also, no extreme values were excluded from the means (e.g. maximum cold tap water at all locations is 20L/day). It is unlikely that the select few women who reported such unrealistic consumption values will affect the conclusions of the epidemiologic analysis, but sensitivity analyses were conducted with and without potential outliers (>10L/day) in Chapter 7 to explore this further.

These data may not be indicative of an average (non-pregnant) person's water consumption. Based on studies to date, pregnancy itself is associated with significant increases in water consumption (Ershow et al. 1991; Forssen et al. 2009; Zender et al. 2001). Women's water consumption in late pregnancy is further investigated in Chapter 8 in a subset of women.

Differences in total water consumption were expected according to demographic characteristics and lifestyle factors, and both ethnicity and employment status have been shown to be related to varying water consumption (Forssen et al. 2007; Kaur et al. 2004). A detailed account of BiB women's patterns of tap water consumption and water use during pregnancy, including variability in water consumption according to various demographic and lifestyle factors, is in preparation for publication (Smith et al. in preparation). Among other findings, employed women are found to drink more on a daily basis than unemployed women and White women drink more than Asian women (Smith et al. 2009; Smith 2011). Another study showed that the most highly educated subjects were less exposed to volatile chlorination by-products such as THMs through ingestion but more exposed through dermal contact and inhalation in pools and showers/baths in a population in Spain (Castano-Vinyals et al. 2011). The authors concluded that health risk perception and economic capacity may affect patterns of water consumption and ultimately result in differences in exposure to water contaminants.

According to a small validation study within BiB of paired water questionnaires and TCAA biomarkers, employed women may be reporting their water consumption less accurately than the unemployed (Smith et al. 2012). It is therefore important to note when comparing consumption by strata that different groups either drink differential volumes of tap water per day, or report their water usage with different accuracy levels.

3.5 Tables

Table 3.1: Studies of water consumption during pregnancy, grouped by total water (all sources) and tap water only if specified (chronological order of publication)

reference	site	years	variable	time point	ingestion in L/day
TOTAL WATER					
Shimokura et al. (1998)	US	1994-1995	average total water consumed at home	pregnancy	1.9
Zender et al. (2001)	US	1996–1997	total water [75% reported tap water as their primary water source]	pregnancy	3.4
Barbone et al. (2002)	Italy	June-December 1999	average total water (tap and bottled)	late pregnancy	2.7
Forssen et al. (2007), Forssen et al. (2009)	3 US cities	2000-2004	-total water -total water	early pregnancy mid pregnancy	2.43 2.60
TAP WATER					
Ershow et al. (1991)	US	1977–1978	average tap water	pregnancy	1.2
Shimokura et al. (1998)	US	1994-1995	average tap water consumed at home [53% of the respondents reporting unfiltered tap water]	pregnancy	0.78
EPA (2000)	US	1994–1996	average tap water	women of childbearing age (15–44 years)	0.9
Zender et al. (2001)	US	1996–1997	cold tap water	pregnancy	1.8
Barbone et al. (2002)	Italy	June-December 1999	average tap water	late pregnancy	0.6
Egorov et al. (2003)	Russia	1999-2001	-average boiled tap water -average non-boiled tap water	pregnancy	0.81 0.01
Kaur et al. (2004)	UK	May-July 2002	total tap water	pregnancy	1.31

Table 3.2: Total tap water at all locations (the same applies to cold tap water, hot tap water, bottled water and total water consumption at all locations)

Definition	Notation	Calculation
For employed women:		
Total tap water consumed at Home	TTh	$cold_h + hot_h$
Total tap water consumed at Work	TTw	$cold_w + hot_w$
Total tap water consumed elsewhere	TTe	$cold_e + hot_e$
Total tap water consumed Outside the Home	TTwe	$TTw + TTe$
Total tap water consumed All locations	TTa	$TTh + TTwe$
For women not in employment:		
Total tap water consumed at Home	TTh	$cold_h + hot_h$
Total tap water consumed elsewhere	TTe	$cold_e + hot_e$
Total tap water consumed Outside the Home	TTwe	TTe
Total tap water consumed All locations	TTa	$TTh + TTe$

Legend:

w tap water
s squash
t tea
c coffee
cold tap water + squash
hot tea + coffee
subscripts: *h:* home; *w*: work; *e*: elsewhere

Table 3.3: Breakdown of filtering habits by location for (a) women in work and (b) women out of work

a) Employed women

		Filtering at Work				
		Yes	**No**	**Don't know**	**NA**	**missing**
Filtering at Home	**Yes**	A	B	B	B	B
	No	C	D	D	D	D
	Don't know	C	D	E	E	E
	missing	C	D	E	F	F

b) Women not in employment

Filtering at Home	**Yes**	B
	No	D
	Don't know	E
	missing	F

Legend:
A: Filtered both at Home and at Work
B: Filtered at Home only
C: Filtered at Work only
D: Filtered neither at Home nor at Work (whether certainly or probably never did)
E: Doesn't know her filtering at Home and/or Work
F: Information on filtering at Home and/or Work is not available or missing

Table 3.4: Water consumption (L/day) among women who reported consumption in any given water category (N_{max}=11,928)

		Percentile Distribution							
	n total = 11,928	mean	SD	min	25th %ile	Median	75th %ile	max	N^
	Cold tap water at Home	1.08	0.67	0.2	0.6	1.0	1.4	10.4	9112
	Cold FILTERED tap water at Home	1.03	0.68	0.2	0.6	1.0	1.4	9.0	909
	Cold UNFILTERED tap water at Home	1.09	0.67	0.2	0.6	1.0	1.4	10.4	8162
	Hot tap water at Home	0.53	0.50	0.2	0.2	0.4	0.6	12.0	7576
	Total tap water at Home	1.44	0.81	0.2	0.8	1.4	1.8	12.2	9679
	Bottled water at Home	0.72	0.55	0.2	0.4	0.6	1.0	4.4	1425
	Total water at Home	1.52	0.84	0.2	1.0	1.4	1.8	12.2	9788
	Cold tap water at Outside the Home	0.74	0.59	0.2	0.4	0.6	1.0	10.0	1749
	Cold FILTERED tap water Outside the Home for Employed women	0.82	0.7	0.2	0.4	0.6	1.0	10.0	633
	Cold UNFILTERED tap water Outside the Home for Employed women	0.76	0.5	0.2	0.4	0.6	1.0	4.0	833
	Hot tap water at Outside the Home	0.50	0.39	0.2	0.2	0.4	0.6	4.0	1661
	Total tap water at Outside the Home	0.78	0.63	0.2	0.4	0.6	1.0	10.0	2715
	Bottled water at Outside the Home	0.79	0.54	0.2	0.4	0.6	1.0	6.0	2093
	Total water at Outside the Home	0.99	0.69	0.2	0.4	0.8	1.2	10.0	3946
	Cold tap water at All Locations	1.22	0.8	0.2	0.8	1.0	1.6	20.0	9193
*	Cold FILTERED tap water at All Locations	1.51	1.0	0.2	1.0	1.2	2.0	20.0	1129
*	Cold UNFILTERED tap water at All Locations	1.27	0.7	0.2	0.8	1.2	1.6	10.4	5510
	Hot tap water at All Locations	0.61	0.6	0.2	0.2	0.4	0.8	12.8	7920
*	Total tap water at All Locations	1.65	0.9	0.2	1.0	1.4	2.0	21.0	9735
*	Bottled water at All Locations	0.83	0.6	0.2	0.4	0.6	1.2	5.6	3373
*	Total water at All Locations	1.91	1.0	0.2	1.2	1.8	2.4	21.0	9830

*signals variables which will be used in the epidemiologic analysis

^N=sample size of women who provided non-zero, valid responses to each component of the given summary measure (see Table A3 - 1, Table A3 - 2, Table 3.5 for details)

Table 3.5: Proportion who reported consumption at all locations with letter references to Table 3.3 (see Table A3 - 1/Table A3 - 2 for details on the proportion who report consumption at Home and Outside the home)

n total = 11,928	valid n	%	possible total for each water type	Filterers	% Filterers	Non Filterers	% Non Filterers	Filterers and Non Filterers	% Filterers and Non Filterers
Cold tap water at All Locations (L/day)									
Non zero, valid values	9,193	77.1	11,928						
zero values	43	0.4							
missing	2,692	22.6							
Filtered her water both at Home and at Work (A)[1]	90	1.0							
Filtered her water at Home only (B)[2]	531	5.8		1,129	12.3			6,639	72.2
Filtered her water at Work only (C)[3]	508	5.5							
Filtered her water neither at Home nor at Work (D)[4]	5,510	59.9				5,510	59.9		
Doesn't know her filtering at Home and/or Work (E)[5]	26	0.3							
Information on Filtering at Home and/or Work missing (F)[6]	7	0.1							
Inconsistent[7]	2,521	27.4							
Hot tap water at All Locations (L/day)									
Non zero, valid values	7,920	66.4	11,928						
zero values	102	0.9							
missing	3,906	32.7							
Total tap water at All Locations (L/day)									
Non zero, valid values	9,735	81.6	11,928						
zero values	6	0.1							
missing	2,187	18.3							
Bottled water at All Locations (L/day)									
Non zero, valid values	3,373	28.3	11,928						
zero values	356	3.0							
missing	8,199	68.7							
Total water at All Locations (L/day)									
Non zero, valid values	9,830	82.4	11,928						
zero values	0	0.0							
missing	2,098	17.6							

Legend:

[1] Employed, and Filtered cold tap water at Home and Work (Home is Yes; Work is Yes)

[2] Employed, and Filtered cold tap water at Home only (Home is Yes; Work is No, Don't know, NA or missing); Out of employment, and Filtered cold tap water at Home (Home=1)

[3] Employed, and Filtered cold tap water at Work only (Home is No, Don't know or missing; Work is Yes)

[4] Employed, and Did not filter cold tap water at all (Home is No; Work is No); Employed, and Did not filter cold tap water at Home (Home is No; Work is Don't know, NA or missing); Employed, and Did not filter cold tap water at Work (Home is Don't know or missing; Work is No); Out of employment, and Did not filter cold tap water at Home (Home=2)

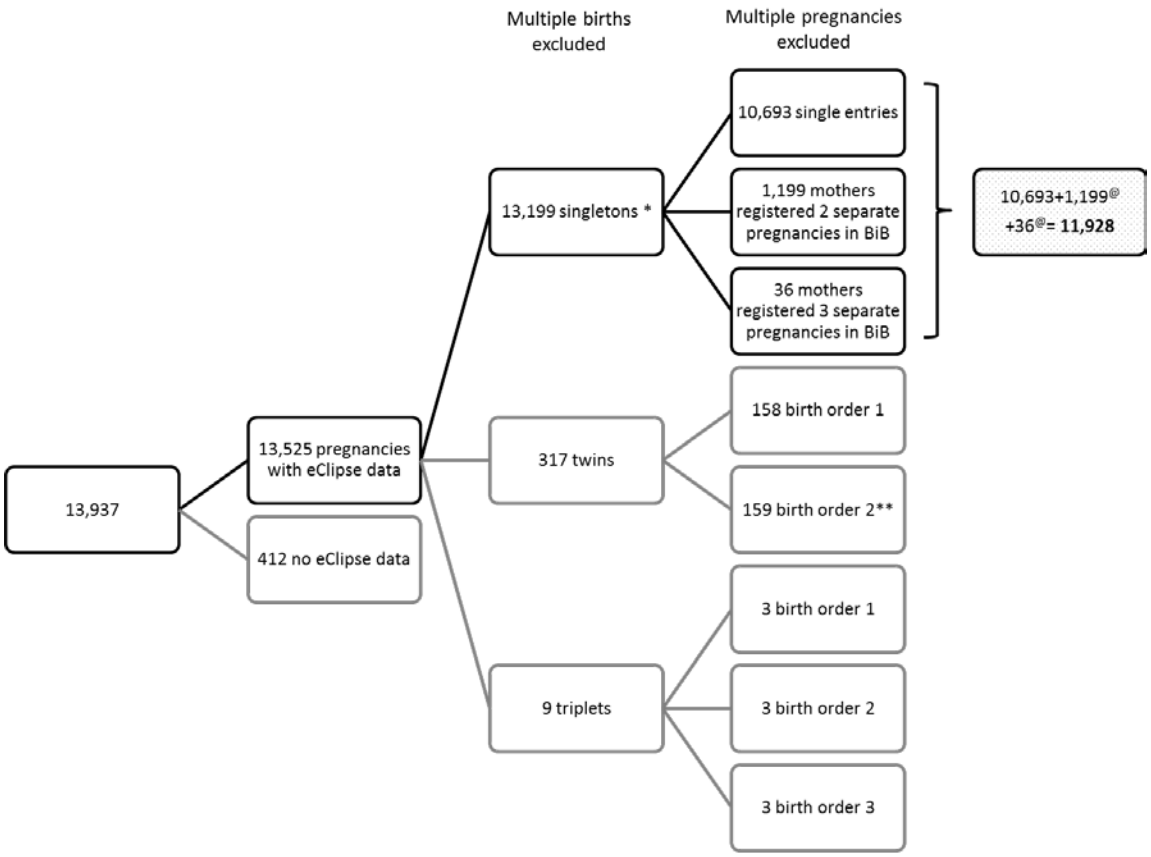
[5] Employed, and Doesn't know her filtering either at Home or Work (Home is Don't know, Work is Don't know); Employed, and Doesn't know her filtering at Home (Home is Don't know; Work is NA or missing); Employed, and Doesn't know her filtering at Work (Home is missing; Work is Don't know); Out of employment, and Doesn't know her filtering at Home (Home=3)

[6] Employed, and Information on filtering at Home missing (Home is missing; Work is NA or missing); Out of employment, and information on filtering at Home missing

[7] Employed, but didn't report drinking cold tap water at Home (=0); Employed, but didn't report drinking cold tap water at Home (=missing); Employed, but didn't report drinking cold tap water at Work (=0); Employed, but didn't report drinking cold tap water at Work (=missing); Out of employment, and didn't report drinking cold water at home (=0); Out of employment, and didn't report drinking cold water at home (=missing)

3.6 Figures

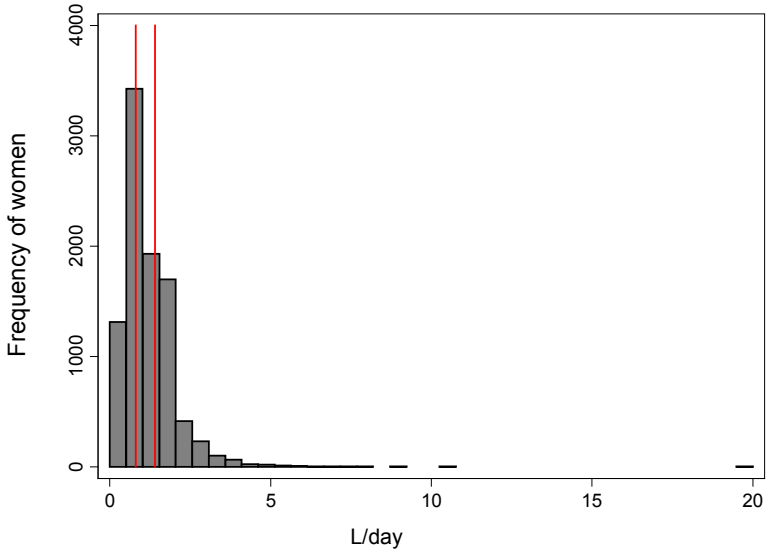
Figure 3.1: Flowchart of inclusion: definition of "eligibility" to epidemiology study



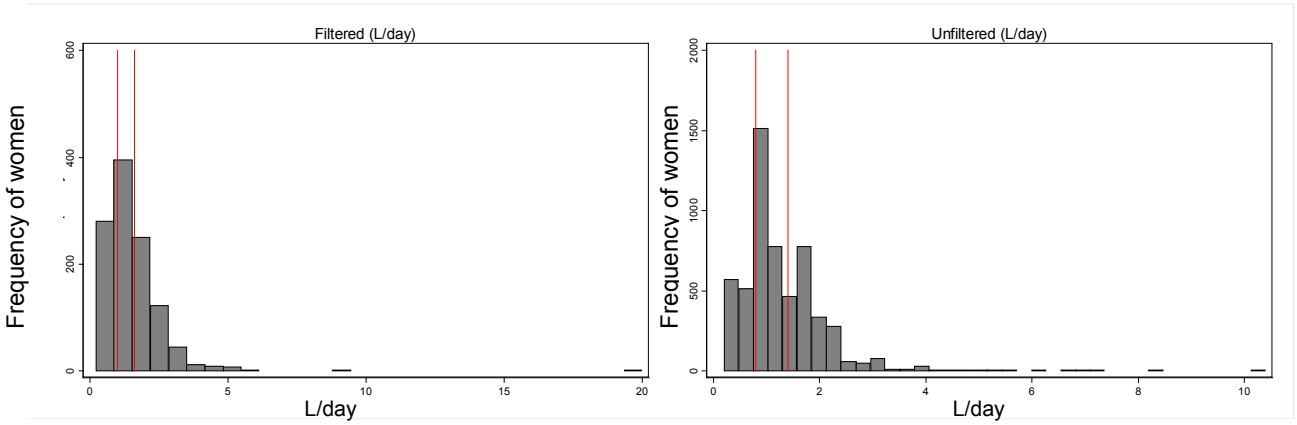
* All but 2 of 13,199 babies were born at the Bradford Royal Infirmary; **1 second-born twin baby died at birth; @ 1 pregnancy per mother was randomly selected

Figure 3.2: Histograms of water consumption with tertile cut points marked

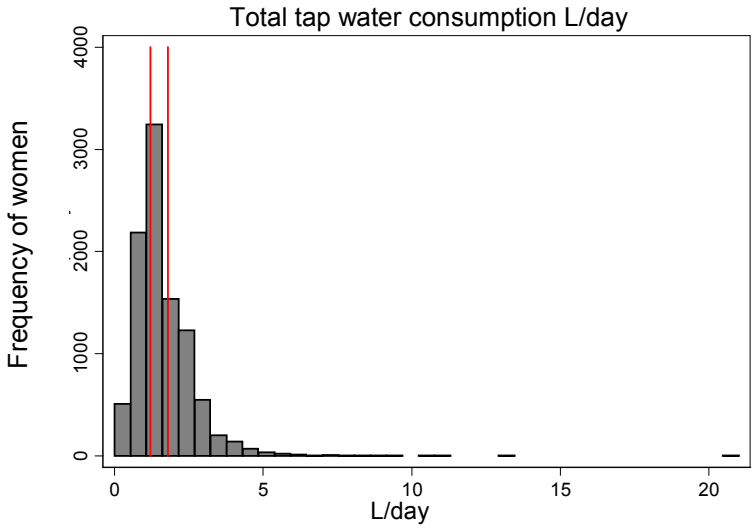
a) Cold tap water consumption
(Sample sizes: Tertile 1: 3,526; tertile 2: 2,838; tertile 3: 2,872)
(cut-off tertile 1: 0.8 L/day, cut-off tertile 2: 1.4 L/day)



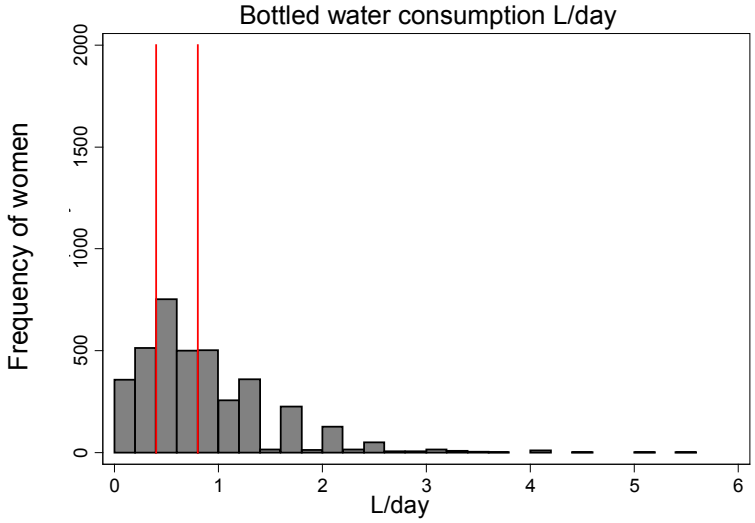
b) Cold filtered and unfiltered tap water consumption
(Sample sizes: Filtered: Tertile 1: 406; tertile 2: 378; tertile 3: 345; Unfiltered: Tertile 1: 2,596; tertile 2: 1,244; tertile 3: 1,670)



c) Total tap water consumption
(Sample size: Tertile 1: 3,834; tertile 2: 2,915; tertile 3: 2,992)
(cut-off tertile 1: 1.2 L/day, cut-off tertile 2: 1.8 L/day)



d) Bottled water consumption with tertile cut points marked
(Sample size: Tertile 1: 1,620; tertile 2: 1,000; tertile 3: 1,109)
(cut-off tertile 1: 0.4 L/day, cut-off tertile 2: 0.8 L/day)



e) Total water consumption with tertile cut points marked
(Sample size: Tertile 1: 3,764; tertile 2: 3,416; tertile 3: 2,650)
(cut-off tertile 1: 1.4 L/day, cut-off tertile 2: 2.0 L/day)

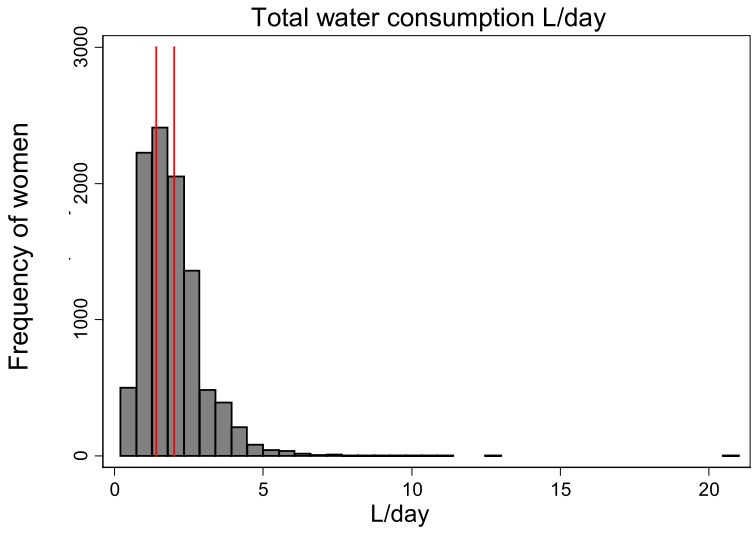
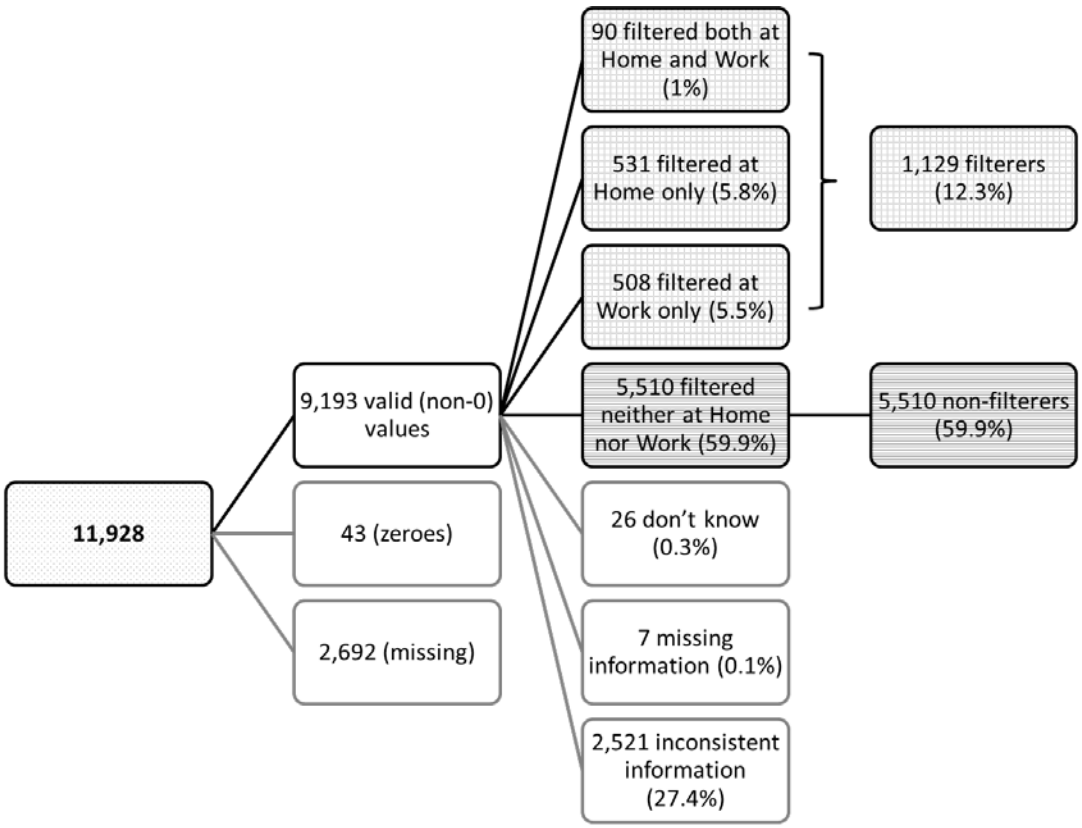


Figure 3.3: Cold tap water all locations organised by filtering



CHAPTER 4 AREA-LEVEL HAA CONCENTRATIONS

This chapter describes the raw HAA data from water samples collected in Bradford homes, their modelling over time and space, and their derivation into area-level quarterly HAA concentrations weighted for each woman's trimester of pregnancy.

4.1 Background

The Health Impacts of long-term exposure to disinfection by-products in drinking water (HiWATE) project funded the analysis of HAAs, as well as a number of HAA determinants in 5 cohorts in Europe.

The determinants available to me and described in this section were: ultraviolet absorbance at 254 nanometres (UV254), bromide, total organic carbon (TOC) (which is the primary ingredient of organic matter), water temperature at the tap, as well as free and total chlorine concentrations, water colour, water conductivity (which measures CO_2 by attributing the difference in sample conductivity before and after oxidization to the TOC of the sample), pH, and turbidity (a measure of water cloudiness), the last six of which were from Yorkshire Water's (YW) routine sampling efforts (see later Methods section).

4.1.1 Determinants of the composition and concentration of HAAs in drinking water

HAA formation in water is influenced by the quality of the raw water and conditions during water treatment and distribution (Chowdhury et al. 2010; Dion-Fortier et al. 2009; Sadiq and Rodriguez 2004; Singer et al. 2002). Water source, whether from ground water, upland (as in Bradford) or lowland reservoirs, largely determines raw water quality.

4.1.1.1 Organic precursors in water

Natural organic matter (NOM) consists of humic and non-humic substances, generally of terrestrial and biological origin respectively (Hwang et al. 2001). It provides the precursor material from which DBPs are formed. Total organic carbon and ultraviolet absorbance at 254 nanometres (UV254) have been widely used as surrogate parameters for monitoring HAA formation as they correlate well with HAA production (Singer et al. 2002). Water colour is also a proxy for NOM content.

4.1.1.2 Inorganic precursors in water

When chlorine is added to naturally-brominated water, the bromide ions are oxidised to hypobromous acid which reacts with NOM (faster than aqueous chlorine) to form brominated DBPs (Diehl et al.

2000; Westerhoff et al. 2004). Therefore, low bromide-containing waters disinfected with chlorine preferentially form DCAA and TCAA, whereas chlorinated waters containing levels of bromide > 100ug/L are dominated by brominated species (Ates et al. 2007; Bougeard 2009; Cowman and Singer 1996; Golfinopoulos and Nikolaou 2005; Heller-Grossman et al. 1993)

4.1.1.3 Water pH

While it is well established that the formation of THMs increases with increasing pH (Kim et al. 2002; Singer 1999; Xie 2003), effect of pH on the formation of HAAs is equivocal.

Overall, HAA formation increases with decreasing pH (Krasner 1999). Liang and Singer (2003) report that increasing pH from 6 to 8 had little effect on the formation of mono-halogenated HAAs (MBAA and MCAA), but significantly decreased the formation of tri-halogenated HAAs, in particular TCAA. DCAA, the other species the most affected by pH, was reported to be the highest at pH 7 on a pH ranges from 5.0 to 9.4 (Krasner 1999).

The speciation of HAAs and THMs at different pH values is determined by the formation mechanisms of the different species. Based on Reckhow and Singer's mechanism (1985), THMs and tri-halogenated HAAs have a common precursor structure (R-CO-CX3), and the relative formation of these species is determined by the nature of the R group and pH. Under alkaline conditions, base-catalysed hydrolysis prevails, yielding more THMs. While in acidic environments, if the R group is a readily oxidisable functional group, tri-halogenated HAAs are formed. If it is not a readily oxidisable functional group, hydrolysis might still prevail, resulting in THMs (Singer et al. 2002). Reckhow and Singer's model also showed that there might be more precursor structures and formation pathways for di-halogenated HAAs than for tri-halogenated HAAs, which may make the formation of di-halogenated HAAs exhibit more complex behaviour with respect to pH (Singer et al. 2002)(Bougeard 2009).

4.1.1.4 Water temperature

Dojlido et al. (1999) has reported that the level of HAAs was 0.63ug/mg C during the winter season (1°C), whereas concentration reached 7.4ug/mg C during summer (23°C). In Malliarou et al. (2005), water temperature is significantly correlated with the ratio of total THM and total HAA. The nature of NOM may also differ from summer to winter and could be responsible for the difference in measured HAA concentrations (Bougeard 2009).

4.1.1.5 Disinfectant dose

Increasing the chlorine dose results in increased DBP formation up to a point where the concentrations reach an equilibrium (Carlson and Hardy 1998; Fleischacker and Randtke 1983).

Free chlorine refers to the residual chlorine left over in the water system after chlorination, and which has not reacted either with NOM to produce DBPs or with ammonia or organic nitrogen to produce chloramines. Total chlorine refers to the sum of free chlorine and chloramines. Target chlorine dose refers to the initial chlorine dose injected into the system prior to reaction with raw water.

4.1.1.6 Reaction time and kinetics

HAAs and THMs form rapidly in the first few hours of the reaction and then the formation slows as the concentrations of the reactants (either NOM or residual chlorine) decreases with time (Gallard and von 2002; Liang and Singer 2003; Nikolaou et al. 2004; Reckhow et al. 1990; Singer 1999; Singer et al. 2002).

4.1.1.7 Residence time

Longer residence times in the distribution network result in greater formation of THM—which is why higher THM concentrations are observed at the extremities of water distribution systems (Rodriguez et al. 2004). However, concentrations of HAAs have been shown to decrease with increasing residence time (Chen and Weisel 1998), because HAAs degrade as they approach the system extremities (Rodriguez et al. 2004). Some hypothesize that this has to do with microbial activity (Hashimoto et al. 1998; McRae et al. 2004; Xie 2003).

The relationships between HAA levels and HAA determinants have been summarised in Table A4 - 1.

4.1.2 HiWATE data

4.1.2.1 Water source and water treatment

Yorkshire Water (YW) is the water company that supplies approximately 4.2 million Yorkshire residents. Each water supply zone (WSZ) covers a population of less than 50,000 people. The study area chosen for sampling covered 8 WSZs in and around Bradford (Figure 4.1). Water is largely supplied to these WSZs from two water treatment works (WTWs), Graincliffe and Chellow Heights, as summarised in Table 4.1 but the specific contribution from each WTW is unknown. The raw water feeding the WTW is drawn from a mixture of upland surface water sources.

Disinfection is via dosing of either chlorine or sodium hypochlorite.

Only about half of the areas of Graincliffe, Airedale and Bradford South East overlap with the study catchment area (see Figure 2.1 in Chapter 2). The areas most populated by BiB cohort mothers are Shipley/Bingley followed by Bradford Central, which together make up 64% of the cohort women's residences.

4.1.3 Predictive models for HAAs

Generating predictive models requires a large database of existing results. Sadiq and Rodriguez (2004) reviewed the existing models for the prediction of DBPs in raw, drinking and treated water, most of which concern THMs. Table A4 - 2 which is based on Bougeard's PhD thesis, summarises selective predictive HAA models for application in raw and treated waters. The HAA models identified in this summary are of three forms:

- DOC/UV based models (Amy et al. 1998; Sohn et al. 2004; Watson 1993)
- chlorine demand models (Gang et al. 2003; Gang et al. 2002)
- linear regression models (Serodes et al. 2003; Villanueva et al. 2003)

4.2 Methods

4.2.1 HiWATE data collection

4.2.1.1 Sampling strategy

All UK water companies are required to monitor levels of THMs in each WSZ at least 4 times a year (but not HAAs, as they are unregulated in the UK).

HAA sampling was undertaken as part of the HiWATE sampling campaign of a suite of DBPs in Bradford and 7 other European partner cohorts (Nieuwenhuijsen et al. 2009c). Three people, Rachel Smith (June 2007 until March 2009), Nina Iszatt (May 2009 through March 2010), and myself (May 2010 until November 2010) were charged with sampling for HAA analysis, as well as sampling for UV254 (abs/m), bromide (mg/l) and TOC (mg/l) analyses.

Taking advantage of YW's routine THM sampling efforts, 15 HiWATE water sampling rounds in at least one of the 8 WSZs in Bradford were carried out on a quarterly basis (Table 4.2). For each round of sampling, the exact sampling site within each WSZ was determined by random sampling of their customer address database. If the sampling officer could not collect a sample from the specified address for any reason, then a sample was collected from a neighbouring property. THM data and data on levels of free and total chlorine (mg/L), colour (mg/l), turbidity (FTU), pH, and conductivity (uS/cm) which are routinely sampled by YW anyway were also released to me for the period January 3rd 2006 and March 31st 2011.

4.2.1.2 Water sampling preparation and collection

Sample vials for HAAs were cleaned and prepared in accordance with a specified HiWATE protocol.

Briefly, 40ml amber glass vials with polypropylene screw caps and polytetrafluoroethylene-faced silica septa were used. They were washed 3 times each with detergent, followed by tap water, then reverse osmosis water and then acetone. The vials were placed in a drying oven at 150°C for 2 hours. Caps and septa were allowed to air dry overnight. Once the vials were cool and dry, preservatives were added 1 or 2 days before the vials were sent by courier to the YW sampling officers. These steps were carried out at Imperial College London's Environmental & Water Resource Engineering laboratory (Civil and Environmental Engineering department).

Approximately 4mg ammonium chloride was added to each vial to act as a preservative for HAAs. Samples collected after 31st March 2008 followed an updated protocol to reflect the use of new quenching agents and pH buffering: a phosphate buffer was added the 4 mg of ammonium chloride to achieve a pH in the range 4.8 to 5.5. The buffer was prepared as a dry homogeneous mixture of 1% sodium phosphate dibasic (NaH2PO4)/99% potassium phosphate monobasic (K2HPO4) by weight. Two vials were prepared to allow duplicate samples to be collected at each sampling site.

Samples for bromide, TOC, and UV254 did not need to be collected headspace-free, and no preservatives were required. ALcontrol Laboratories supplied the bottles for their sampling (clear glass bottles for TOC, PET bottles for bromide and UV254).

Water temperature is not a routine monitoring requirement for YW, but the sampling officer was requested to measure and record temperature when collecting HiWATE samples (in degrees Celsius, °C).

Both pH and temperature were recorded with a Hanna HI-98128 waterproof pHep pH/c meter. Other parameters such as conductivity, turbidity, and colour were measured by YW as part of their routine sampling at the same sampling site and date as the HiWATE samples.

Free and total residual chlorine (mg/L) was measured with a portable colorimeter (Lovibond Comparator 2000+ Test Kit AF 112 A). This kit has a range of 0.1-1.0mg/l measured in 0.1mg increments. Liquid reagents DPD No. 1 (Lovibond) were added to a vial of water, and the colour of the solution inspected visually using the comparator. If a sample fell between divisions in the range it was assigned an intermediate value (e.g. 0.15mg/l if between 0.1 and 0.2mg/l).

At ALcontrol Laboratories, TOC was determined by chemical oxidation and infrared spectrometry. Bromide was determined by ion chromatography. UV-absorbance was determined by UV-visible spectroscopy at 254nm.

Immediately after collection all tap water samples were stored in a refrigerator (4ºC) in the sampling officer's van, until they were returned to the sampling depot. The day after sampling, the vials were sent by courier to the University of the Aegean for laboratory analysis. Courier shipment to University of the Aegean laboratory (located in Mytilene on the island of Lesbos) took 3.5 days on average. As of May 2009, HAA samples were driven up to Cranfield University instead (Table 4.2). TOC, UV254 and bromide samples were submitted to ALcontrol Laboratories in Rotherham, UK (Wakefield, UK as of 2010).

4.2.1.3 HAA laboratory analyses

Both the University of the Aegean and Cranfield University analysed HAAs using a modified form of US EPA Method 552.3 (EPA 2003; Tung et al. 2006). At Cranfield University, all samples were dechlorinated and analysed in duplicate. The derivatized HAAs (methyl esters) were measured using gas chromatography (GC) with micro electron capture detection (μECD) (Agilent 6890, Santa Clara, CA, USA).

The limit of detection (LOD) is defined as the statistically calculated minimum amount that can be measured with 99% confidence that the reported value is greater than zero (Glaser et al. 1981). LOD is 2ug/l for MCAA and 1ug/l for all other HAAs. This is based on the lowest standard used for those analyses, as in Jeong et al (2012). Samples below the LOD were set to 2/3 of LOD, as was done in Whitaker et al (2003a). Samples below the minimum reporting limit (MRL), which is the threshold expected for accurate quantification in an unknown sample and has to be at least three times the limit of detection (Bougeard 2009; EPA 2003), were set to 1/3 of MRL (MRL = 3xLOD, so 1/3 of MRL= LOD) (Table 4.3).

Between 10 and 100% of the data were missing or undetectable for different HAAs (Table 4.3) such that modelling the nine HAAs as a sum total (e.g. HAA9, HAA5) was not feasible.

Consequently, I focused my investigation on three HAAs: dichloroacetic acid (DCAA), trichloroacetic acid (TCAA) and bromodichloroacetic acid (BDCAA). DCAA and TCAA together represent the largest proportion of all HAAs by mass (Figure 4.2) and had a good detection rate (90%)—when DCAA was detectable typically so was TCAA (Table 4.3). BDCAA was the brominated species in this sample with the highest number of valid samples and was therefore chosen to represent brominated species.

4.2.2 HAA Data & Modelling

158 DCAA, 158 TCAA and 143 BDCAA observed data points, sampled between June 2007 (second quarter of 2007) and November 2010 (fourth quarter of 2010), were used. Eight data points from the second quarter of 2009 were excluded from each analysis and interpolated from the model, over concerns that the data (which represent the first batch of data generated after transition to a new laboratory at Cranfield University) were unreliable. The models also extrapolates to one quarter on either end of the observed data period, in order to cover as many women's pregnancies as possible but to avoid overstretching the model.

My modelling approach was in two stages: model selection was first performed in a frequentist framework (section 4.2.2.1), and the best model was then run in a Bayesian framework in order to take every parameter's uncertainty into account and easily impute missing covariate data (section 4.2.2.2).

Bayesian hierarchical models were fit in WinBUGS (Lunn et al. 2000) to predict mean DCAA, TCAA, and BDCAA concentration levels (and 95% credible intervals[13]) by WSZ and time (on a year-quarterly basis between the first quarter of 2007 and the first quarter of 2011 inclusive).

Modelling on a year-quarterly basis was necessary because there appear to be differences in HAA levels between time points which cannot be explained by the other measured explanatory variables.

Linear regression model were used because the transformed data are approximately normal and data are sparse. This is similar to the approaches used by Villanueva et al. (2003), Serodes et al. (2003) and Smith and Bennett for the THM modelling conducted on routinely monitored YW data (Smith 2011).

4.2.2.1 Model building in frequentist framework (using R)

DCAA and TCAA were square root transformed, and BDCAA was natural log-transformed to better approximate the normal distribution (Figure 4.3).

The final models to be implemented in the Bayesian framework were first selected in R using a Generalised Additive Model (GAM) model (mgcv package). Each model included a factor for WSZ and a spline on time in order to enable WSZ- and time-specific predictions. Following the principle that the more information is added to the model, the more likely the assumption of missing at random holds true (Gelman et al. 2004), each HAA model included any predictor which meaningfully reduced

[13] The Bayesian 95% credible interval can be interpreted as an interval that has a high probability (95%) of containing the unknown quantity of interest (Gelman A, Carlin JB, Stern HS, al e. 2004. Texts in statistical science In: Bayesian data analysis 2nd ed Boca Raton, FL: 2004.:CRC Press, LLC.)

the AIC (i.e. a minimum reduction of 5, assuming that smaller errors might be due to sampling error). Only one proxy variable for natural organic matter (either TOC, UV254, or colour), and one variable for residual chlorine in the water distribution system (free or total chlorine) were used, because of these variables' high correlations (see Table A4 - 3, Table A4 - 4, and Table A4 - 5).

Water temperature and conductivity remained in the DCAA model; temperature, conductivity and TOC remained in the TCAA model; and conductivity and total chlorine in the BDCAA model.

4.2.2.2 Bayesian model

For each observation i (i=1, …, N), time j (j=1, …, T), and water supply zone z (z=1, 2, …, 8)

$$Y_{ij} \sim N(\mu_{ij}, \sigma^2)$$

$$\mu_i = \alpha + \theta_{j(i)} + \gamma w_{z(i)} + \sum_{k=1}^{k} \beta_k X_i$$

$$\theta \sim RW1(\sigma_0^2)$$

Y_{ij}	HAA concentration (ug/L)
σ^2	measurement error variance ($1/\sigma^2 = \tau$ is the precision)
μ_{ij}	mean HAA concentration (ug/L)
α	overall mean of the HAA concentrations
θ_j	non-linear time function (random walk of order 1, RW1)
σ_0^2	random walk variance
w_{ij}	Water supply zone (WSZ)
γ	the coefficient for WSZ factor
X_{ij}	covariates
β_k	regression coefficient for covariate k
T	total number of equally-spaced time points (T=17)
N	total number of observations (see section 4.2.2.2.3)

I assumed a normal distribution for natural log-transformed BDCAA. DCAA and TCAA however were specified using a normal distribution truncated at 0 to avoid negative predictions (N_T). All covariates were standardised to have a mean of 0 and a standard deviation of 1 to facilitate model convergence (see Figure 4.4, left, for a graphical representation). The mean HAA concentrations depend on an overall mean, α, non-linear time effects θ, the WZS, and covariates X.

In WinBUGS, the spline on time was replaced by a random walk on time, order 1 (using a car.normal distribution with temporal neighbours (t-1) and (t+1) for θ[2], ..., θ [T-1], and temporal neighbours (t+1) for θ [1] and (t-1) for θ [T]).

DCAA, TCAA and BDCAA models differ from the each other as specified in Table 4.4.

4.2.2.2.1 Covariate imputation model

As missing data are treated as an additional parameter within the Bayesian formulation (Gelman et al. 2004), covariates were imputed within the model assuming a missing at random mechanism. To take into account the correlation between the covariates, a multivariate normal model for the full set of covariates to be imputed was specified. Within it, temperature, TOC and total chlorine (all of which are predicted by season) were modelled with a factor for WSZ and a factor for quarter (and an intercept); conductivity was imputed only with a factor for WSZ (and an intercept) (see graphical representation in Figure 4.4, right).

The imputation model for the missing covariates for the DCAA model is as follows.
For each observation i (i=1, …, N) and time j (j=1, …, T),

$$X_{ij} = \begin{pmatrix} x_{1ij} \\ x_{2ij} \end{pmatrix} \sim \text{MVN} \left(\begin{pmatrix} \mu_{1ij} \\ \mu_{2ij} \end{pmatrix}, \Sigma \right)$$

$$\mu_{1ij} = \alpha_1 + \varphi q_{ij} + \gamma_1 w_{ij}$$

$$\mu_{2ij} = \alpha_2 + \gamma_2 w_{ij}$$

Where x_1 and x_2 are the temperature and conductivity, respectively, and are modelled using a multivariate normal distribution (MVN) with mean μ and covariance matrix $\sum$. α_k (k=1,2) represents the overall mean for each variable x_k; φ denotes the coefficient for quarter factor (q_{ij}) and γ_k (k=1,2) is the coefficient for WSZ factor for each variable x_k.

4.2.2.2.2 Priors

Non-informative priors were specified for both fixed parameters and variances (Gelman 2006; Gelman and Hill 2007). The coefficients for the covariates' effects and intercepts, for WSZ and for quarter factors were given normal priors (with a mean of 0 and precision of 0.01). For DCAA and TCAA, the coefficient for WSZ was truncated at 0 to avoid unrealistic negative values (specifying a mean of 0 and a precision of 0.01 for DCAA and a mean of 0 and a more informative precision of 0.1 for TCAA, in accordance with the truncated model).

The priors on measurement error (σ) and random walk standard deviations (σ_0) were given uniform distributions (between 0 and 100). The prior on the multivariate model precision was specified by a Wishart distribution. A prior guess at the covariance matrix of the covariates Σ was made by calculating the empirical variance of each covariate and the empirical covariance between each pair of covariates.

4.2.2.2.3 Computation

The software WinBUGS, based on Markov Chain Monte Carlo (MCMC) algorithms, was used to implement all models. After a burn-in period of 10,000, 10,000 more iterations per chain were performed. To reduce autocorrelation, every tenth iteration was monitored. The results are based on a sample of 2,000 iterations. Diagnostic tests were performed to check convergence of the parameters to the posterior distribution. Convergence was checked using the Gelman and Rubin statistic (Gelman and Rubin 1992) and by visual examination of plots (trace, quantiles).

Eight entries per WSZ and year-quarter for DCAA and TCAA (for a total of 1088 (=8x17x8) each), and 6 entries (for a total of 816 (=6x17x8)) for BDCAA were loaded into the model. This was done to ensure that final means per WSZ and time were taken on the same number of observations, and that observed data usage was maximised.

Predicted outcomes (also truncated for DCAA and TCAA) were back-transformed within the WinBUGS model. The MCMC output was analysed using STATA 12.1.

After calculating the means of the 8 (or 6) predictions per WSZ and year-quarter (2000 iterations of 136 means remain), I took the overall mean for DCAA and TCAA and the median for BDCAA (due to skew in the data) over the 2000 iterations. Daily predictions were generated by dividing the difference between two consecutive quarterly values by the number of days between them (starting mid-quarter). The modelled period therefore covers 14/02/2007 through 14/02/2011, inclusive.

4.2.3 Time-weighted area-level concentrations

4.2.3.1 Calculations

The duration of the first trimester was defined as 93 days from the date of conception, which itself was calculated with available BiB data as follows:

Date of Conception = Date of delivery - Estimated no. days of gestation (excluding the date of conception)

The second trimester was defined as the 93 days following the first trimester, and the third trimester as the remaining number of days from the end of the second trimester until the day preceding delivery. Date of conception was considered the first complete day of exposure (day 1).

Time-weighted area-level HAA concentrations were calculated for each woman by averaging the daily HAA predictions corresponding to each woman's first, second or third trimester of pregnancy (to the day). Because HAAs were modelled both in time and space, these averages were either based on the predictions in:

- the WSZ corresponding to the woman's home (Method 1)
- a time-weighted combination of the WSZ corresponding to the woman's workplace if available (for the days estimated to be spent at work) and home (for the other days) (Method 2)

This methodology has previously been used for THMs by Toledano et al. (2005), Nieuwenhuijsen et al. (2008) and Smith (2011) and similar time-weighted averages have also been used in other studies (Hwang et al. 2008; Patelarou et al. 2011; Smith 2011). Only concentrations for trimesters entirely contained within the modelled period were derived.

4.2.3.2 Time-weighted area-level concentrations based on residence WSZ: Method 1

The main epidemiologic analysis focuses on area-level concentrations based on the WSZ of the women's residence at recruitment to BiB. In sensitivity analyses, results were compared using time-weighted area-level concentrations derived from the combination work and residence WSZ if available

4.2.3.3 Time-weighted area-level concentrations based on a combination of work and residence WSZ if available: Method 2

In sensitivity analyses in Chapter 7, I compare the effects on birth outcomes of the area-level exposures of women based on their residence postcodes only (Method 1), to those of the area-level exposures of those same women using weighted residence and work place postcodes. Figure A4 - 1 shows the BiB women's work places which were geocoded with this goal.

Only 22% the cohort women under consideration a) reported working at BiB recruitment, b) provided both valid work and home addresses which were within Bradford and c) stated working for a given number of days a week (Figure A4 - 2, Table A4 - 6). Not included in this flowchart, sample sizes are further restricted by the requirement for pregnancies to overlap in full with the HAA modelled time period. The other 78% of the women kept their area-level exposures calculated by Method 1.

1257 workplace addresses could not be geocoded (only 1160 after accounting for women who stated working mostly from home, such that work place was now known) because of missing address or postcode information, or address information which were too imprecise to use ("Street level accuracy" or above was required (=6), www.spatialepidemiology.net). For addresses obviously outside of Bradford [14] , Town (city, village) level accuracy (or above) on www.spatialepidemiology.net 's accuracy scales was considered sufficient.

[14] i.e. stated city: Batley, Brighouse, Bury, Dewsbury, Guiseley, Halifax, Harrogate, Heckmondwike, Hipperholme, Huddersfield, Ilkley, Leeds, Manchester, Morley, Normanton, Ossett, Otley, Pool in Wharfedale, Pontefract, Rawdon, Sheffield, Wakefield, Yeadon

For Method 2, several assumptions had to be made.

A) The time spent at work was based on answers to a question in the baseline questionnaire on number of days per week spent at work. Work days were arbitrarily assigned to weekdays first, starting with Monday. If a woman spends 1 day at work per week, she was therefore assumed to work every Monday. If she spent 2 days/week, she was assumed to work every Monday and Tuesday, 3 days/week every Monday through Wednesday etc. (Table A4 - 7).

B) If a woman was categorised as working on a given day, her entire water consumption for that day was based on the modelled HAA concentration in that (work) WSZ.

C) The number of work days reported at BiB recruitment applies throughout pregnancy including the full third trimester.

D) If the main place of work is reported to be the home (one of the questions in the baseline questionnaire), then any work place information on record was overruled and the WSZ corresponding to the woman's home only was used.

All residence and available workplace addresses were geocoded and mapped in ArcGIS to their corresponding WSZs. Dr Kees de Hoogh and I were in charge of the geocoding for BiB women's home addresses, while I geocoded workplace addresses manually using www.spatialepidemiology.net because those data's quality was poorer.

4.2.4 Exclusion criteria

In the following summaries and analyses, the same exclusion criteria were applied as described in Chapter 3 (section 3.2.3).

I did not consider exposure over the full pregnancy time period, because of my interest in critical windows of exposure during pregnancy (first, second, and third trimesters) (see Introduction (Chapter 1)). There is also the concern over adding too many comparisons to this analysis.

4.3 Results

4.3.1 HiWATE data

Table 4.5 describes the nine (non-normally distributed) HAAs from the HiWATE sampling effort, excluding the second quarter of 2009. See Figure 4.2 for a boxplot representation of this table. The concentrations of HAAs in Bradford ranged from 1.80ug/l to 47.70ug/l with a median of 25.65ug/l.

As expected from other upland surface water studies with low-bromide source waters (Bougeard 2009; 2000), DCAA and TCAA were the species preferentially formed by mass, their medians

representing together more than 74% of the sum of the medians of the nine HAAs (27.68ug/l). But their ranges are quite large: 0.30-21.70ug/l for DCAA, and 1.00-25.40ug/l for TCAA (Table 4.5, Figure 4.2).

The range of HAA concentrations reported in drinking water worldwide have been summarised in Table 4.6.

4.3.1.1 Spatial variation of DCAA, TCAA, BDCAA, and determinants

There is little spatial variation between the 8 WSZs based on 11 to 26 data points per WSZ (Figure 4.5). This is perhaps not surprising since the treatment processes were broadly the same and the raw water used was from similar surface water sources in a relatively compact geographical area.

As expected from the literature, mean HAA concentrations are lower in the three high bromide water supply zones (Airedale, Graincliffe, and Keighley, marked as * in Figure 4.5) compared to the other WSZs.

Free and total chlorine levels also varied spatially, being slightly higher in Bradford Central (BDC), Idle/Pudsey (IPY) and Shipley/Bingley (SPY). Conductivity was higher in Airedale and Graincliffe. No distinctive spatial patterns were noted for colour, turbidity, temperature, UV254, and TOC (data not shown). Differences in HAA or HAA determinant levels (particularly free chlorine) in two WSZs which were supplied by the same treatment plant may have been due to differences in the positions of the WSZs in the distribution network but there was insufficient data to test this influence directly.

4.3.1.2 Temporal variation of DCAA, TCAA and BDCAA

DCAA concentrations are on average lowest in the hot months and highest in the cold months of the year, but there are no strong, consistent seasonal trends in HAAs over the HiWATE collection period (Figure 4.6).

The only noticeable temporal trend is temperature (highest in the third quarter), as well as colour, UV254 and TOC, which are higher on average in the third and fourth quarters as compared to the first and second.

HAA determinants' descriptive statistics and data availability are summarised in Table A4 - 8.

4.3.1.3 Relationship between HAAs

TCAA and DCAA, the two dominant HAAs in this sample, are significantly correlated ($r=0.34$, $p<0.001$, $N=158$) (see Table A4 - 9b and Figure A4 - 3). Their relationship appears to be roughly

inverse over time in the first two years of sampling (DCAA high when TCAA is lower, and vice versa), but to track each other in 2009 and 2010 (Figure 4.7).

4.3.1.4 Using THM determinants to predict HAAs

Correlations between HAAs and determinants confirmed expectations that low bromide, high temperature and high NOM waters lead to greater HAA concentrations on average (Diehl et al. 2000; Krasner et al. 1989; Krasner 1999).

Bromide is inversely correlated with temperature, and positively correlated with pH, UV254, and TOC. Surrogate measures of NOM (colour, UV254, and TOC) are significantly correlated with temperature; along with temperature, their levels tend to be higher in the summer-autumn seasons. Conductivity is inversely associated with pH and with colour, UV254 and TOC, likely denoting its association with less cloudy NOM-free waters. Free and total chlorine are significantly inversely correlated with pH, bromide, and all surrogate variables for NOM.

4.3.2 HAA Modelling

4.2.2.1 Frequentist model outputs

Here are the detailed outputs on best DCAA, TCAA, and BDCAA models selected in R.

4.3.2.1.1 Model selected for DCAA

The best model for DCAA, in addition to a smoothing spline on time and a factor for water supply zone, includes temperature and conductivity (N=115, k=13, df=7.1, AIC=229.5, R^2=0.57, and explains 63% of deviance) (Figure 4.8). Temperature is highly inversely significant and conductivity is highly significant as well (with a very low coefficient) (data not shown).

$$Y_{ij} \sim N(\mu_{ij}, \sigma^2)$$

$$\mu_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 C_{ij} + \sum_{p=1}^{8} \beta_{3[p]}\, zone_i + f(t) + \varepsilon_{ij}$$

f(t) is represented using rank 13 thin plate regression spline bases (Wood 2003).
Y_{ij} = sqrt(DCAA) (ug/l)
T_{ij} = temperature (degC)
C_{ij} = conductivity (uS/cm)
$zone_i$ = zone indicator (1 or 0)

Temperature and conductivity are not significantly correlated: r=0.11 (p=0.224, N=135).

Temperature follows a predictably marked seasonal trend: high in the summer, low in the winter. In fact, I tried replacing temperature by season—for which there were no missing data—in the models but this covariate did not come out significant in any model tried (data not shown). The information contained by the variable season was probably redundant with year-quarter. The observed seasonal trend is consistent across water supply zones. The relationship between sqrt(DCAA) and temperature is negative (r=-0.35, p<0.001, N=134), as comes out in the models (Table A4 - 3).

The conductivity data do not suggest a seasonal pattern. No consistent time pattern were found by water supply zone either. There is one observation where both conductivity and DCAA value are available for which conductivity is very high (conductivity at year-quarter = 2009q2 and wsz=4 = 615.0 uS/cm). But this does not seem to be an outlier, as values for conductivity in the full dataset (N=7,727) span from 8 to 630 uS/cm, and the coefficients for a linear regression on sqrt(DCAA) with and without this outlier are the same. The relationship between sqrt(DCAA) and conductivity is negative (r=-0.22, p=0.010, N=139) (Table A4 - 3).

After stratifying by WSZ and year-quarter in the complete dataset spanning 2007q1 to 2011q1, a difference in pattern by zone emerges: two of the three high bromide zones ADL (=1) and GCF (=5) have the highest average conductivity levels, while the third of the three high bromide zones KLY (=7) has the lowest. Water supply zone is therefore an important predictor of conductivity level.

4.3.2.1.2 Model selected for TCAA

The best model for TCAA includes temperature, conductivity and TOC (N=115, k=13, df=9.6, AIC=588.5, R^2= 0.71 (new 0.65 by cross-validation), 76% of deviance explained) (Figure 4.9).

$$Y_{ij} \sim N(\mu_{ij}, \sigma^2)$$

$$\mu_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 C_{ij} + \beta_3 W_{ij} + \sum_{p=1}^{8} \beta_{3[p]}\, zone_i + f(t) + \varepsilon_{ij}$$

f(t) is represented using rank 13 thin plate regression spline bases (Wood 2003).
Y_{ij} = sqrt(TCAA) (ug/l)
T_{ij} = temperature (degC)
C_{ij} = conductivity (uS/cm)
W_{ij} = total organic carbon (mg/L C)
$zone_i$ = zone indicator (1 or 0)

Temperature and TOC are significantly positively correlated (r=0.30, p<0.001, N=156), TOC and conductivity are significantly negatively correlated (r= -0.33, p<0.001, N=162), and temperature and conductivity are not correlated (r=0.11, p=0.224, N=135).

Because TCAA and DCAA are available on the same subset of 158 HiWATE data, the same associated dataset on temperature and conductivity is available for them both as well. As with the DCAA models, temperature is therefore found to follow a seasonal trend, while conductivity does not. While it was negatively correlated with sqrt(DCAA) (r=-0.35, p<0.001, N=134), temperature is positively correlated with sqrt(TCAA) (r=0.31, p<0.001, N=134); conductivity remains inversely related to sqrt(TCAA) (r=-0.25, p=0.003, N=139) (Table A4 - 4).

TOC levels are also somewhat seasonal: higher concentrations are observed in the summer (and to a lesser extent in autumn) seasons compared to winter and spring. The observed seasonal trend is consistent across water supply zones, GCF (=5) and BSW (=4) behaving a bit differently at 2007q4 and 2009q3, respectively. TOC is positively and significantly correlated with sqrt(TCAA) (r=0.33, p<0.001, N=158).

4.3.2.1.3 Model selected for BDCAA

Best model include conductivity and total free chlorine: N=132, k=15, df=14, AIC=159.8, R^2=0.802, and explains 84% of deviance (Figure 4.10).

$$Y_{ij} \sim N(\mu_{ij}, \sigma^2)$$

$$\mu_{ij} = \beta_0 + \beta_1 C_{ij} + \beta_2 L_{ij} + \sum_{p=1}^{8} \beta_{3[p]}\, zone_i + f(t) + \varepsilon_{ij}$$

f(t) is represented using rank 15 thin plate regression spline bases (Wood 2003).
Y_{ij} = ln(BDCAA) (ug/l)
C_{ij} = conductivity (uS/cm)
L_{ij}=total chlorine (mg/L)
$zone_i$ = zone indicator (1 or 0)

Again, I do not exclude the possible conductivity outlier, as the effects do not change with and without it.

4.3.2.2 Bayesian model outputs

The output of my Bayesian modelling exercise is summarised in Figure 4.11, and in Table 4.7, as well as in map form in Figure 4.12, Figure 4.13, and Figure 4.14. For details on the imputation models (for missing temperature, conductivity, TOC or total chlorine) of the analysis, see Table A4 - 10.

4.3.3 Time-weighted area-level concentrations

Women's mean time-weighted area-level concentrations (based on residence postcode, i.e. Method 1) ranged from 8.64-8.86ug/L, 11.98-12.41ug/L, and 1.32-1.33ug/L across trimesters for DCAA, TCAA and BDCAA, respectively. Figure 2.1 (in Chapter 2) presents the residence locations of the BiB women and their distribution within Bradford's eight WSZs. Figure 4.15 describe women's time-weighted area-level concentrations based on residence postcode (Method 1), which were categorised by tertiles for the purposes of the epidemiologic analysis. The sample size for all three HAA concentrations linked to women's first, second, and third trimester was 10,521, 11,312, and 11,585, respectively (Table 4.8, Figure 4.16).

Women's mean time-weighted area-level concentrations (based on residence and work postcode, i.e. Method 2) ranged from 8.72-8.92ug/L, 11.85-12.34ug/L, and 1.33ug/L across trimesters for DCAA, TCAA and BDCAA respectively (Table A4 - 11).

4.3.3.1 Comparison of HAA concentrations by Methods 1 and 2

Summary statistics by the two methods different for TCAA and BDCAA ($p<0.001$), justifying the sensitivity analyses carried out. Method 2 lead to smaller sample sizes than Method 1, but higher average concentrations for DCAA, lower average concentrations for TCAA and the same for BDCAA (Table A4 - 12, Figure A4 - 4). Requiring a woman's trimester of interest to overlap completely with the modelled period (Feb 14th 2007 through Feb 14th 2011) in order to use her area-level HAA concentration measure meant that quite a few women were lost.

4.4 Discussion

4.4.1 HiWATE data

The concentrations of HAAs in Bradford were within a similar range to those reported previously, particularly from other studies on upland surface water. For example, Zhang et al (2010) reported an average total HAA concentration of 21.3ug/L and maximum of 41ug/l.

Several national HAA occurrence studies have been undertaken in the US (Krasner et al. 1989; Krasner et al. 2006; McGuire et al. 2002; Weinberg et al. 2002). Krasner et al. (1989), in a survey of 35 water treatment utilities, found median total HAA5 concentrations ranging from 13 to 21ug/L, with MCAA, MBAA, DCAA, TCAA and DBAA at <1-1.2, <0.5, 5.0-7.3, 4.0-6.0 and 0.9-1.5ug/L respectively. These values are lower than Bradford's.

Malliarou et al. (2005) reported HAA9 means of 35, 52 and 95ug/L in finished waters from three regions in England and Wales, and a maximum concentration of 244ug/L. Still in the UK, Zhang et al (2010) reported low HAA levels (<2ug/L total HAA) in a groundwater-based system, and higher HAA formation in a system alimented by upland surface water (~40ug/L total HAA). In her PhD thesis, Bougeard observed considerable variation between 11 chlorine-treated waters with HAA9 levels ranging from 5.0 to 69ug/L, and an average value of 37ug/L (Bougeard 2009).

Of note when considering these levels, there are differences in chlorination practices and residuals in the UK compared to the US: for example, the maximum contaminant level (MCL) for residual chlorine in the US was set at 4.0mg/l, while the drinking water inspectorate (DWI) in the UK states that typical disinfectant levels were maintained at 0.5mg/l or less (www.dwi.gov.uk) (Goslan et al. in prep). Indeed, according to YW, the normal target range for free chlorine (pre manganese contactors) in the Chellow Heights WTW was 0.85-2.0 mg/l, a dose which was normally set automatically and was flow paced. The normal target for free chlorine in the clean water tanks was 0.35 mg/l. In the Graincliffe WTW, the normal target range for free chlorine in treated water was 0.6-1.1 mg/l.

Jeong et al (2012) found that HAA9 concentrations over a 24 hours monitoring period ranged from 11.6 to 51.5ug/L in 7 European cities (Appendix C). Each site in Jeong et al (2012)'s study corresponded to a HiWATE study site, including two sites in Bradford: HAA9 concentrations averaged 13.3ug/L in Shipley (a low bromide district of Bradford) and 11.6ug/L in Airedale (high bromide). This agrees with the expectation that higher bromide areas will have lower HAA levels (Table A4 - 1), but these levels are lower than those recorded in our tap water samples. This difference may be due to differences in analysis methods, or else it is possible that the water samples may have degraded in the process of being shipped to the US EPA laboratories where the samples were analysed, based in Georgia, USA.

Most of the surveys cited in Table 4.6 found TCAA and DCAA to be the major species formed in treated waters exposed to formation potential (FP) tests using chlorine. In particular, the major species formed in Bougeard's work (2009) were TCAA (ranging from 1.0 to 40ug/L) and DCAA (ranging from 2.5 to 22ug/L) followed by BDCAA, BCAA, MCAA and DBCAA. For all waters, MBAA, DBAA and TBAA were the least concentrated, with TBAA not always detected. This is the same as what I found in Bradford.

4.4.1.1 Spatial variation of DCAA, TCAA and BDCAA, and determinants

Typical concentrations of bromide in natural waters range from 30 to 200ug/L, with an average of 100ug/L (Amy et al. 1994). Most of the waters disinfected with chlorine will therefore have the potential to form brominated DBPs. The Bradford water only had a median of 10ug/L, but spatial

variation in HAAs was still noticeable. Consistent with literature predictions, the three high bromide areas also had the highest pH levels, and lowest free (and total) chlorine levels. Previous studies would also have predicted lower temperatures in these three areas compared to the others, but these data are not consistent with this prediction (data not shown).

Goslan et al. (in prep) warns that bromide levels may be falsely low in this study because no quenching agent was added to the bromide bottle. This could mean that any chlorine residual present could have continued to react with the bromide during shipping and storage. However, I suspect that this may have affected WSZs equally (causing lower levels in absolute terms but not in relative terms).

4.4.1.2 Temporal variation of DCAA, TCAA and BDCAA, and determinants

Krasner et al (1989) and Williams et al. (1997) reported highest HAA concentrations in summer likely due to the nature of natural organic matter (NOM) at that time of year (Dojlido et al. 1999). Zhang et al. (2010) reported an approximate twofold difference between the highest and the lowest concentrations over the four seasons in the WSZs (based on the highest measured level occurring in autumn, 42.2ug/L, and the lowest value obtained in spring, 19.2ug/L). Nissinen et al. (2002) found low concentrations in May and October but no marked seasonal trend, while Serodes et al. (2003) surprisingly found that the average concentrations of HAAs were not highest at the highest incubation temperature.

Relatively few data points (between 7 and 29 per time point) were collected to assess temporal variation. A seasonal trend could have emerged with larger samples, as temperature, colour, UV254 and TOC which all predict higher HAA levels all seemed to have higher levels in the last two quarters of each year. Increased biological activity in summer and leaf fall in autumn will contribute to higher NOM in raw water in summer and autumn, which is thought to increase HAA formation. This is what has been found for THMs (Chen and Weisel 1998; Garcia-Villanova et al. 1997; Golfinopoulos 2000; Krasner et al. 1989; Rodriguez et al. 2004; Whitaker et al. 2003a).

4.4.1.3 Using THMs (or other determinants) to predict HAAs

Determining whether THM levels can be used as a surrogate indicator for HAA levels has been the goal of many studies to date (see Table A4 - 13). This would be a useful method for quality control and monitoring in water utilities given that laboratory analyses for HAAs are more resource- and time-consuming than for THMs. However not only can the correlation rates be poor, varying widely by location even within the same study, but by chance only a few sampling time points that the HiWATE study used for its analyses coincided with routine THM sampling time points. This means that the correlations of observed THM and HAA data in Bradford cannot be assessed here.

4.4.2 HAA Modelling

Variables included in these models (temperature, conductivity, TOC and total chlorine) should not necessarily be viewed as predictors outside of the context of this dataset (Bennett, personal communication).

Because the factors that affect levels of HAAs in drinking water (such as amount of disinfectant, nature and concentration of NOM, pH of water, bromide ion, water temperature, contact time between the disinfectant and the water, seasonal and regional variability and stagnation of water in the plumbing pipes and hot water tanks) vary temporally and spatially, the characterisation of HAAs and their exposure analysis are complex (Chowdhury et al. 2010) .

I had relatively few HAA data points per WSZ and per time point to work with in this dataset, some of which had low detection rates, which is why I chose to focus on modelling DCAA, TCAA and BDCAA.

There were no duplicate samples that were both analysed by Cranfield University and by the University of the Aegean, for inter-laboratory comparison purposes. And because of scarcity of data, 2007-2008 and 2009-2010 data based on University of the Aegean and Cranfield University-generated data, respectively were not modelled separately. As such, the default assumption made is that laboratory and methodological differences are negligible, which may not hold true.

4.5 Tables

Table 4.1: Water supply zones (WSZ) and BiB population (N=11,928, including 24 missing addresses and 22 out of area)

	WSZ name (2004 nomenclature)	WSZ abbr.	Supplying WTW	BiB population (%)
1	Airedale	ADL	Graincliffe	607 (5.1)
2	Bradford Central *	BDC	Chellow Heights / Graincliffe	3172 (26.6)
3	Bradford SE	BSE	Chellow Heights	1187 (10.0)
4	Bradford SW	BWS	Chellow Heights	1474 (12.4)
5	Graincliffe	GCF	Graincliffe	15 (0.1)
6	Idle/Pudsey	IPY	Chellow Heights	961 (8.1)
7	Keighley	KLY	Graincliffe /Embsay /Oldfield /Sladen Valley†	9 (0.1)
8	Shipley/Bingley	SPY	Chellow Heights	4457 (37.4)

* 2008 (in 2008, the Bradford City WSZ merged with Peel Park/Laisterdyke to form Bradford Central); †Keighley is mainly supplied by the Graincliffe WTW, with small contributions from three other WTWs (Embsay WTW, Oldfield WTW and Sladen Valley WTW)

Table 4.2: HiWATE sampling by laboratory

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
2007						Aeg			Aeg		Aeg	
2008			Aeg		Aeg			Aeg			Aeg	
2009			Aeg		Cran				Cran		Cran	
2010			Cran		Cran			Cran			Cran	

Aeg: samples analysed at the University of the Aegean
Cran: samples analysed at Cranfield University

Table 4.3: Number of missing data for each of the nine HAAs (Nmax=184) MRL: minimum reporting limit; LOD: limit of detection (see section 4.2.1.3 for details)

	MCAA	MBAA	DCAA	BCAA	TCAA	DBAA	BDCAA	DBCAA	TBAA
N *(% of total)*	42 (*43*)	15 (*27*)	157 (*95*)	121 (*86*)	165 (*99*)	52 (*58*)	113 (*75*)	44 (*46*)	0 (*0*)
<MRL	24	16	9	8	1	7	14	24	24
<LOD	32	25	0	11	0	30	24	27	32
total (% of 184)	**98 (*53*)**	**56 (*30*)**	**166 (*90*)**	**140 (*76*)**	**166 (*90*)**	**89 (*48*)**	**151 (*82*)**	**95 (*52*)**	**56 (*30*)**

Table 4.4: Transformations, distributions and best model covariates for the three modelled HAAs: DCAA, TCAA and BDCAA

	transformation	distribution	covariates remaining in best frequentist model and included in Bayesian model*
DCAA	square root	truncated Normal	temperature, conductivity
TCAA	square root	truncated Normal	temperature, conductivity, TOC
BDCAA	ln	Normal	conductivity, total chlorine

*in addition to a factor for WSZ and a spline/RW(1) on time

Table 4.5: Summary of HiWATE HAAs in ug/L: 2007q2 through 2010q4, excluding 2009q2

	MCAA	MBAA	DCAA	BCAA	TCAA	DBAA	BDCAA	DBCAA	TBAA
median	2.00	0.67	9.90	1.10	10.70	0.40	1.20	1.04	0.67
IQR	1.67	0.33	9.05	0.85	7.05	0.27	1.43	0.83	0.33
min	0.80	0.28	0.30	0.10	1.00	0.11	0.10	0.67	0.67
max	5.30	2.62	21.70	3.62	25.40	1.38	4.60	1.90	1.11
N	90	48	158	132	158	81	143	87	48

Table 4.6: Reported levels of HAAs in studies worldwide (adapted from a table in Bougeard's PhD thesis) (chronological order)

Reference	Location	Water source	HAA measured	Range (ug/l)
Krasner et al. (1989)	US †	Drinking water	HAA5	13.0-21.0
Peters et al. (1991)	Netherlands	Drinking water	HAA9	0.0-14.7
Williams et al. (1997)	Canada	Drinking water	MCAA	0.3-10.0
			MBAA	0.01-9.0
			DCAA	0.2-163
			TCAA	0.04-473
			DBAA	0.01-2.0
Cancho et al. (1999)	Spain	Drinking water	HAA9	11.0-32.0
Dojlido et al. (1999)	Poland	After chlorination	HAA6	10.0-120
Nissinen et al. (2002)	Finland	Drinking water	HAA6	6.00-261
Serodes et al. (2003)	Canada	After experimental chlorination		
Malliarou et al. (2005)	UK	Drinking water	HAA9	NR-244
Krasner et al. (2006)	US †	Drinking water	HAA9	5.0-130
Ates et al. (2007)	Turkey	Filtered surface water	HAA9	6.0-177
Wang et al. (2007)	China	Drinking water	HAA6	0.4-14.0
Bougeard (2009)	UK	After chlorination	HAA9	5.0-69
Jeong et al. (2012)	UK	Drinking water	HAA9	11.6-13.3

† Goslan et al. (in prep): there are difference in chlorination practices and residuals in the UK vs. US

Table 4.7: Parameters for DCAA, TCAA, and BDCAA models

parameters	categories	DCAA model Mean (95% Cred. Int.)	TCAA model Mean (95% Cred. Int.)	BDCAA model Mean (95% Cred. Int.)
WSZ factor γ	zone 1	REF	REF	REF
	zone 2	0.94 (0.50, 1.38)	0.75 (0.42, 1.09)	0.21 (-0.02, 0.43)
	zone 3	0.99 (0.49, 1.45)	1.08 (0.71, 1.44)	0.31 (0.08, 0.55)
	zone 4	0.54 (0.09, 1.01)	1.09 (0.75, 1.46)	0.27 (0.05, 0.49)
	zone 5	0.20 (-0.27, 0.69)	0.05 (-0.33, 0.42)	0.05 (-0.18, 0.29)
	zone 6	1.03 (0.61, 1.47)	0.91 (0.58, 1.26)	0.24 (0.02, 0.47)
	zone 7	0.58 (0.04, 1.06)	0.28 (-0.12, 0.67)	0.20 (-0.07, 0.48)
	zone 8	1.05 (0.62, 1.48)	0.89 (0.56, 1.25)	0.26 (0.04, 0.49)
Temperature effect β_1		-0.37 (-0.54, -0.21)	0.12 (-0.03, 0.28)	
Conductivity effect β_2		0.03 (-0.09, 0.15)	-0.03 (-0.12, 0.06)	0.13 (0.07, 0.19)
TOC effect β_3			0.31 (0.17, 0.47)	
Total Chlorine effect β_4				0.07 (0.00, 0.13)
random walk variance σ_0		0.24 (0.08, 0.67)	0.32 (0.11, 0.86)	2.21 (0.99, 5.32)
measurement error variance σ_2		0.37 (0.29, 0.48)	0.20 (0.16, 0.26)	0.10 (0.08, 0.13)

zone 1=Airedale (ADL), zone 2=Bradford Central (BCE), zone 3=Bradford South East (BSE), zone 4=Bradford South West (BSW), zone 5=Graincliffe (GCF), zone 6=Idle/Pudsey (IPY), zone 7=Keighley (KLY), zone 8=Shipley/Bingley (SPY)

The credible intervals for the BDCAA model extend at the extremities from 0.06-18.49ug/L at 2007q1, and from 0.10-33.60ug/L at 2011q1.

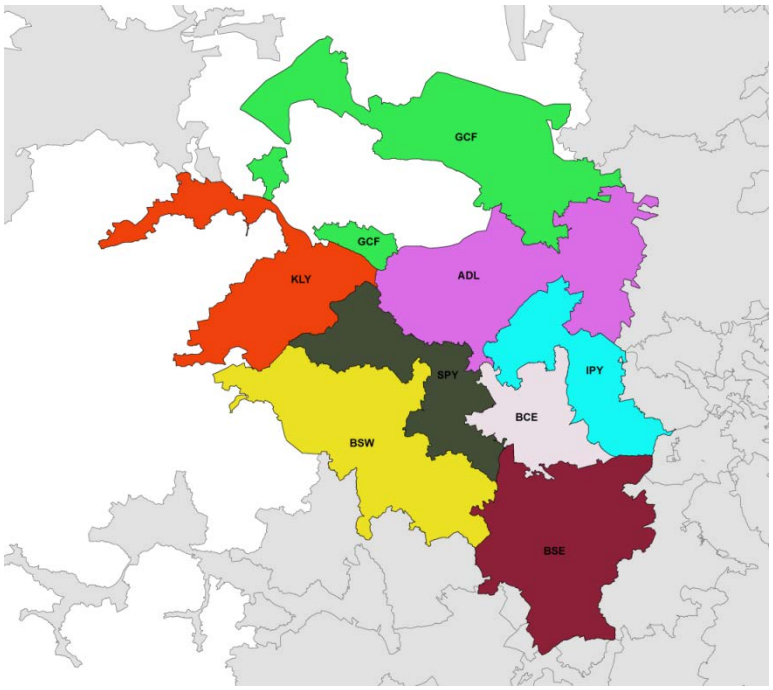
(Red means the 95% credible interval does not cross 0.)

Table 4.8: Summary statistics of 3 trimester-weighted HAA concentrations for each woman (in ug/L) (includes only women whose given trimester completely overlaps with the modelled period)

In ug/L		Mean	SD	Min	Max	N	% of N=11,928
Average [DCAA]	trimester 1	8.64	2.94	1.95	16.04	10,521	88.2
	trimester 2	8.68	2.88	1.95	16.04	11,312	94.8
	trimester 3	8.86	2.90	1.78	16.86	11,585	97.1
Average [TCAA]	trimester 1	11.98	3.45	3.10	20.67	10,521	88.2
	trimester 2	12.26	3.57	2.86	20.67	11,312	94.8
	trimester 3	12.41	3.44	3.32	21.08	11,585	97.1
Average [BDCAA]	trimester 1	1.32	0.68	0.35	3.34	10,521	88.2
	trimester 2	1.32	0.67	0.35	3.34	11,312	94.8
	trimester 3	1.33	0.68	0.26	3.58	11,585	97.1
Average no of days in third trimester		91.53	12.64	1.00	128.00	11,585	97.1

4.6 Figures

Figure 4.1: Map of Bradford's 8 water supply zones



Map generated using Arc Map 10.
GCF: Graincliffe (green)
ADL: Airedale (purple)
KLY: Keighley (red)
IPY: Idle/Pudsey (blue)
SPY: Shipley and Bingley (dark green)
BCE: Bradford Central (pink)
BSW: Bradford South West (yellow)
BSE: Bradford South East (burgundy)

Figure 4.2: Box plots of the nine HAA concentrations (ug/L)

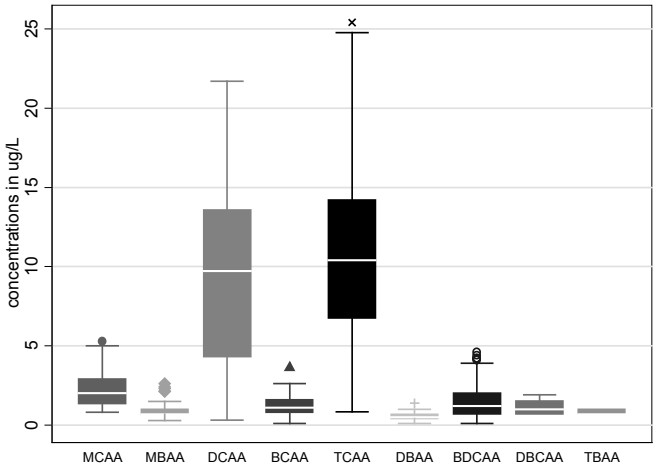


Figure 4.3: Distributions before and after transformation

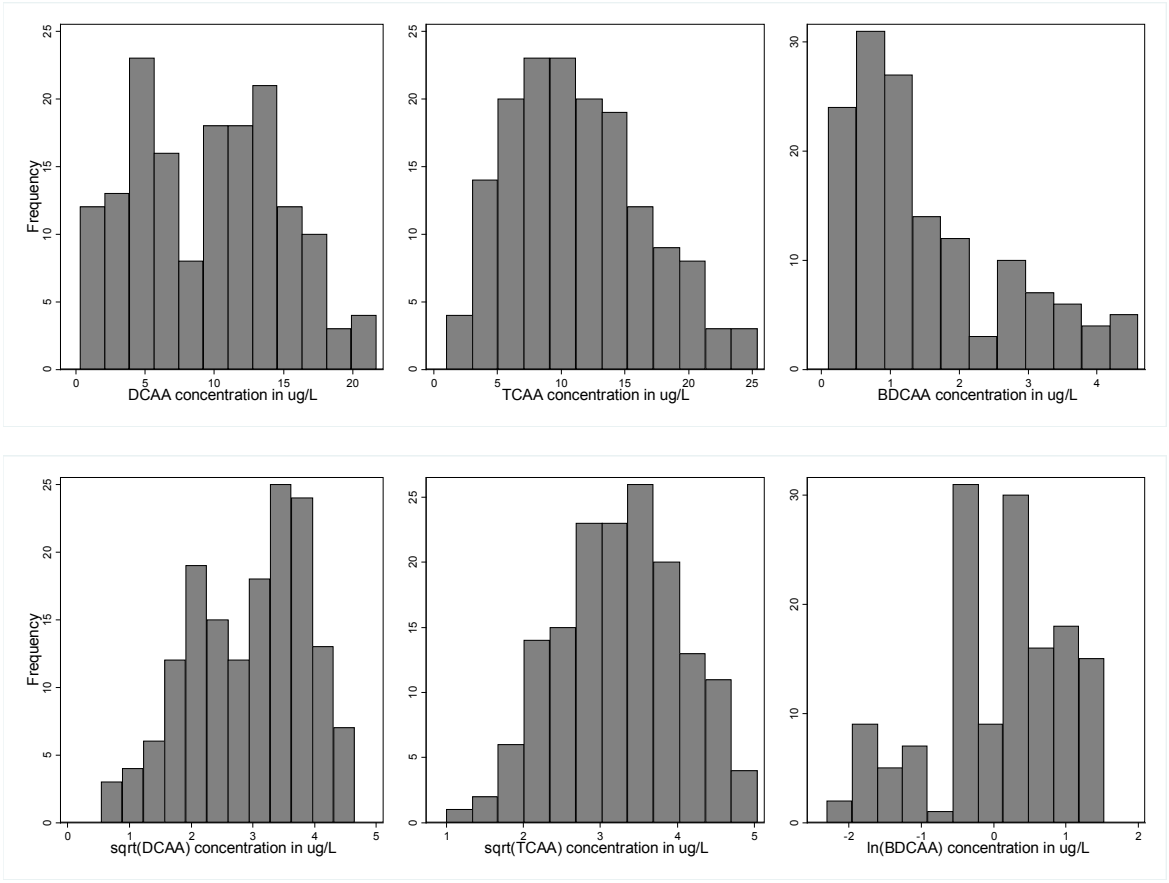


Figure 4.4: Directed acyclic graph for the DCAA model; $Z_{ij} = (w_{ij}, q_{ij},)$ (square shape because it is fully observed as opposed to the circle random variables). For a description of the variables, see section 4.2.2.2 (p. 68)

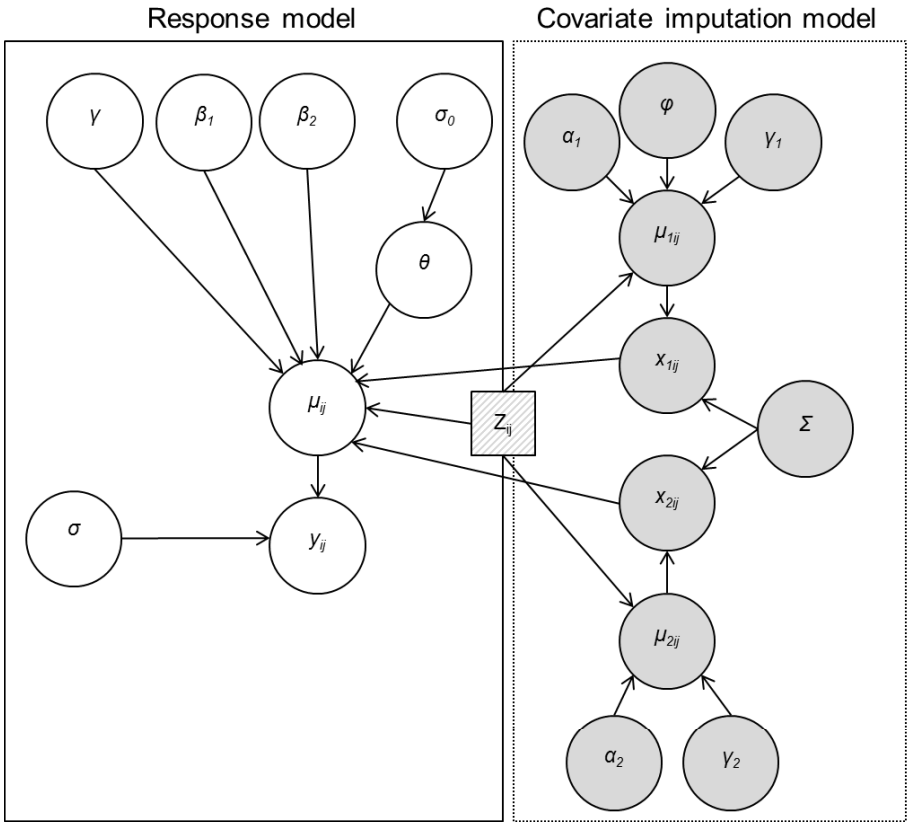
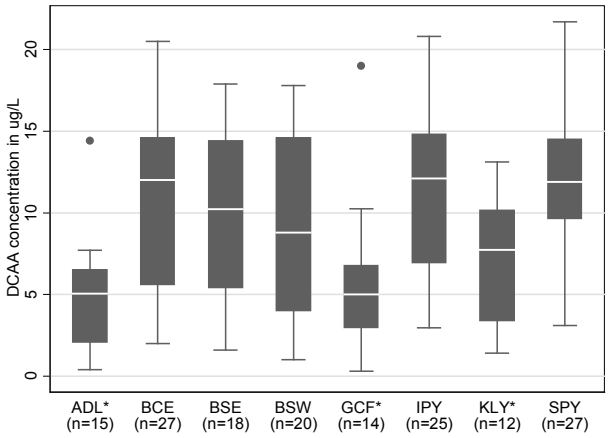
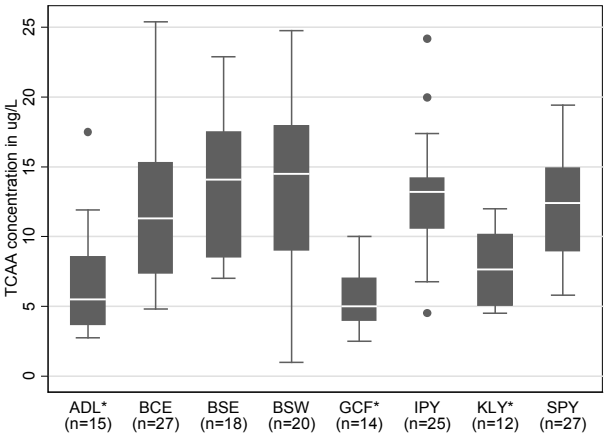


Figure 4.5: HAA concentration by water supply zone (WSZ) – HiWATE data: 2007q2 through 2010q4, excluding 2009q2 (*indicates high bromide WSZ) for a) DCAA, b) TCAA, and c) BDCAA

a)



b)



c)

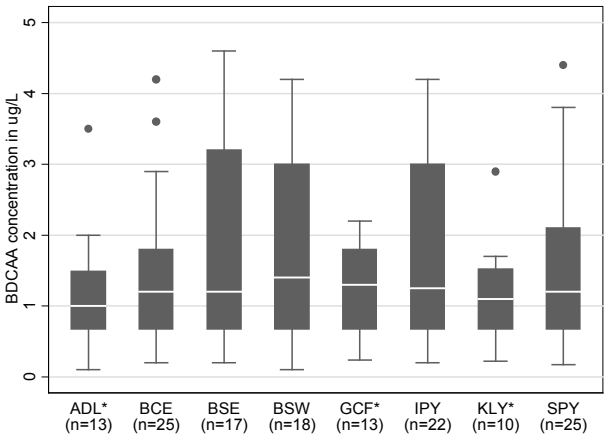
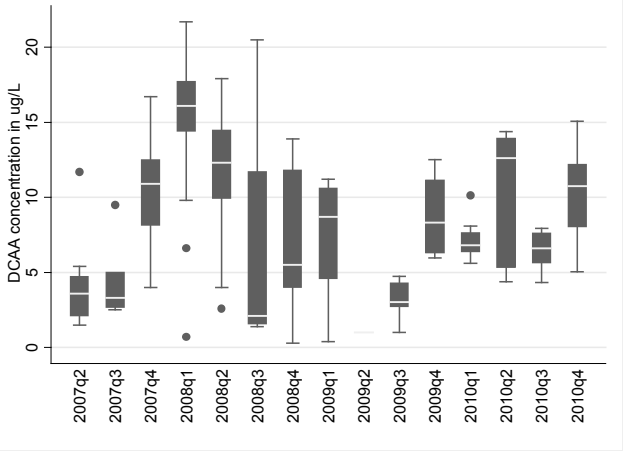
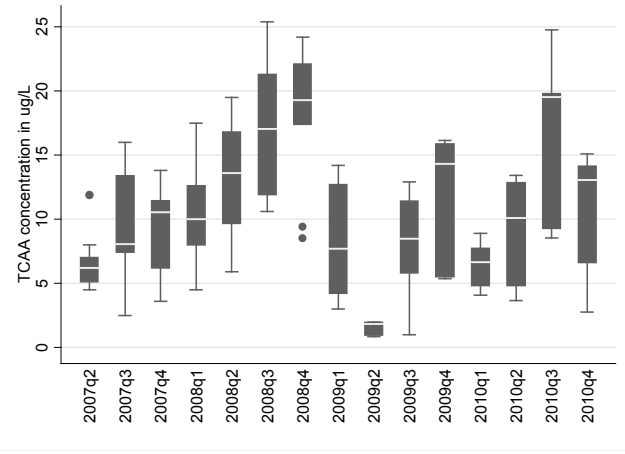


Figure 4.6: HAA concentration by year and quarter – HiWATE data: 2007q2 through 2010q4 (including 2009q2 which was deemed unreliable) for a) DCAA, b) TCAA, and c) BDCAA

a)



b)



c)

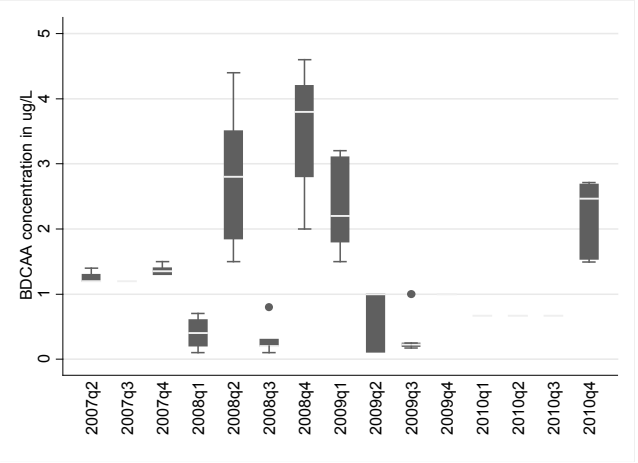
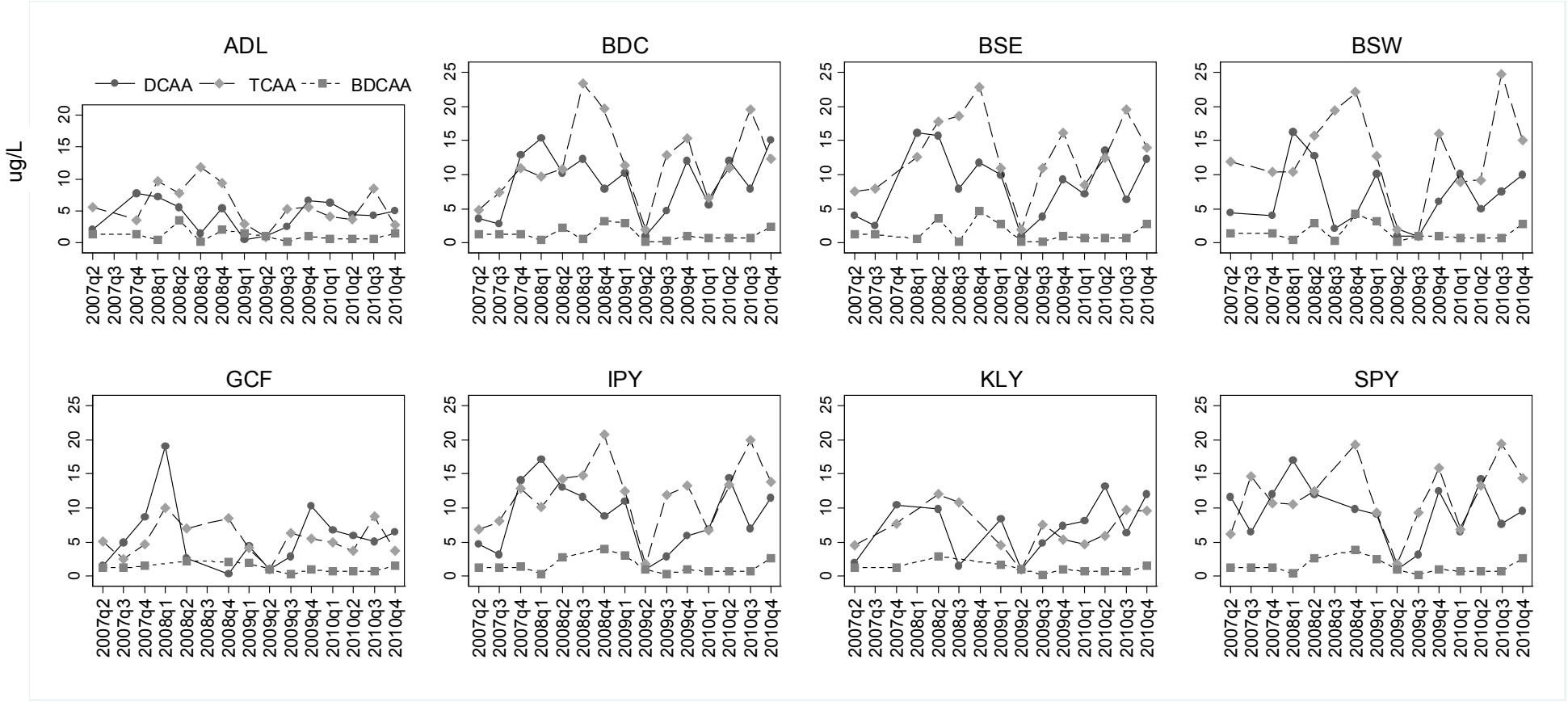


Figure 4.7: Mean DCAA, TCAA and BDCAA over time (ug/L): 2007q2 through 2010q4, including 2009q2 (for continuity of x-axis)



ADL=Airedale, BDC=Bradford Central, BSE= Bradford South East, BSW=Bradford South West, GCF=Graincliffe, IPY=Idle/Pudsey, KLY=Keighley, SPY=Shipley/Bingley

Figure 4.8: Model fit for best model for DCAA selected in R (dotted line: 95% CI)

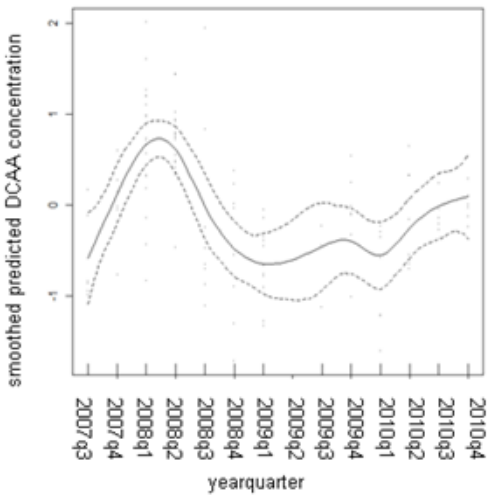


Figure 4.9: Model fit for best model for TCAA selected in R (dotted line: 95% CI)

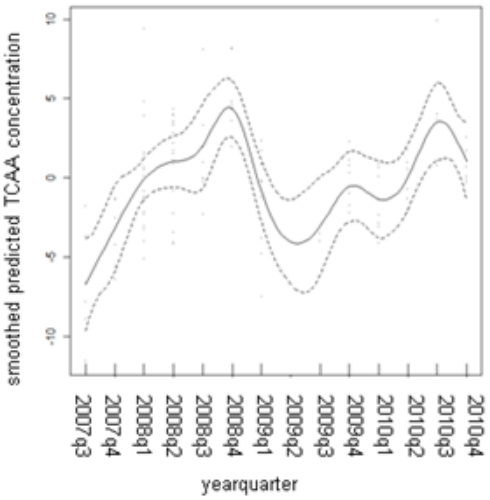


Figure 4.10: Model fit for best model for BDCAA selected in R (dotted line: 95% CI)

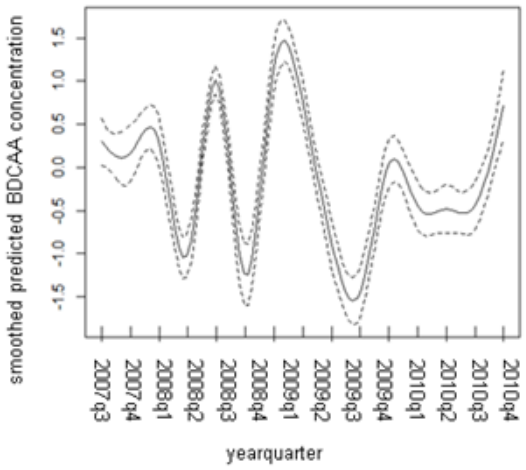


Figure 4.11: Modelled DCAA concentration (ug/L) over time for WSZ 1 (Airedale or ADL) (95% credible intervals), and a plot of posterior mean modelled DCAA concentration in all 8 WSZs. Same pair of plots for TCAA, and BDCAA
solid line = posterior mean for DCAA and TCAA, and posterior median for BDCAA
dashed line= 95% credible intervals
BDC=Bradford Central, BSE= Bradford South East, BSW=Bradford South West, GCF=Graincliffe, IPY=Idle/Pudsey, KLY=Keighley, SPY=Shipley/Bingley

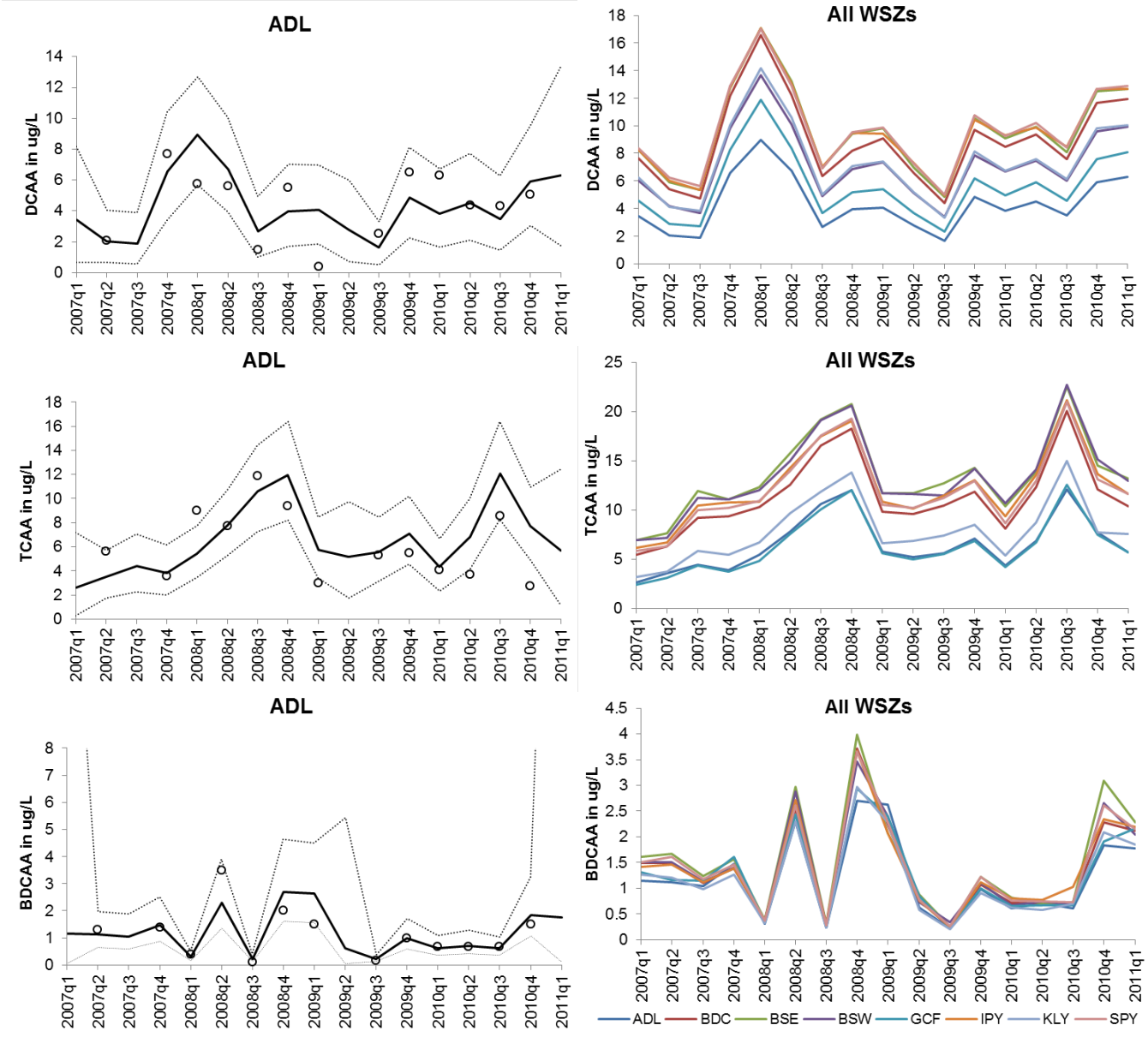


Figure 4.12: Modelled DCAA concentrations over time and space (by quartiles) for 8 water supply zones in Bradford
(Quarter 1 for years 2007 and 2011 as well as quarter 2 for year 2009 were extra/interpolated)

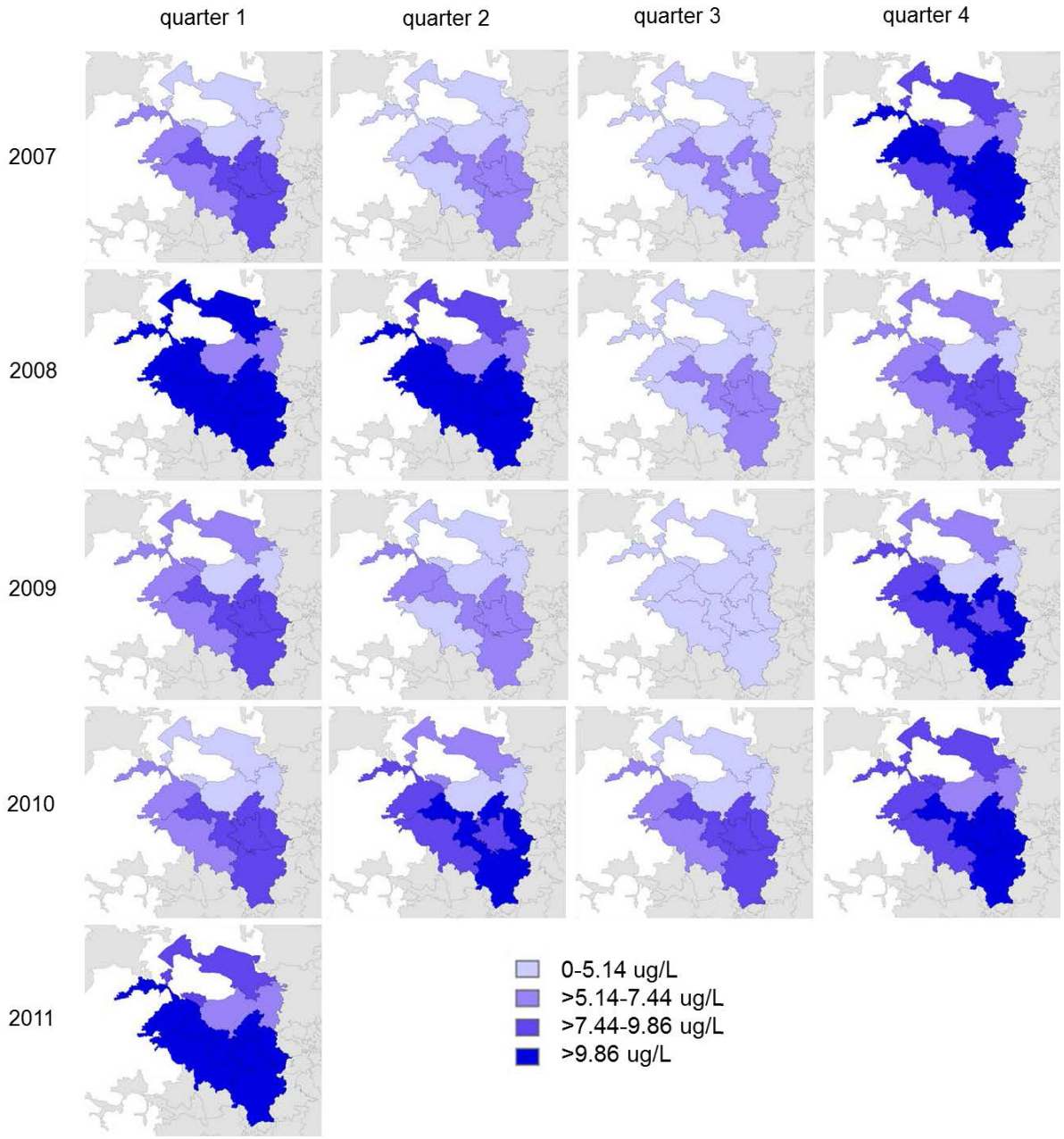


Figure 4.13: Modelled TCAA concentrations over time and space (by quartiles) for 8 water supply zones in Bradford
(Quarter 1 for years 2007 and 2011 as well as quarter 2 for year 2009 were extra/interpolated)

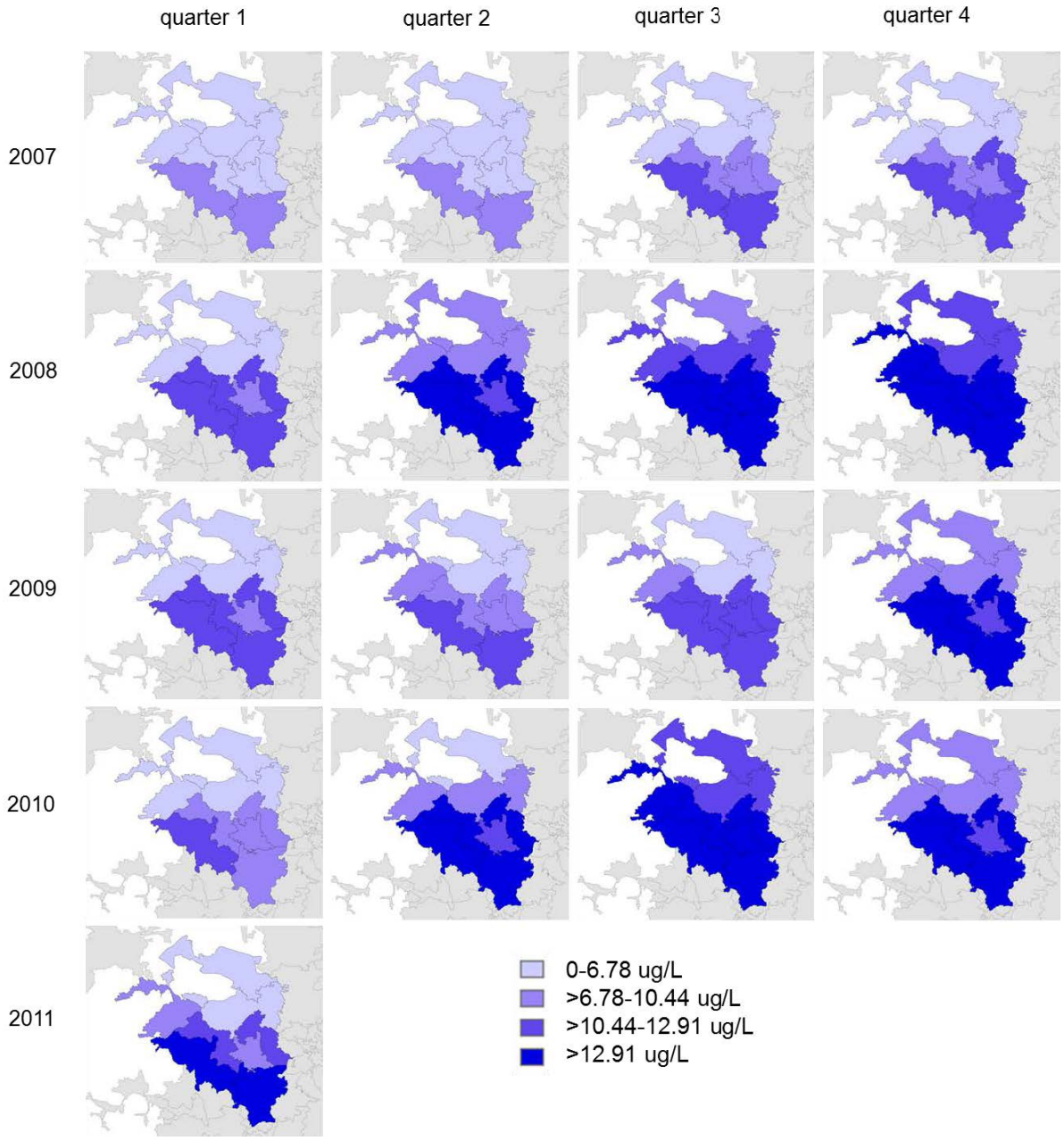


Figure 4.14: Modelled BDCAA concentrations over time and space (by quartiles) for 8 water supply zones in Bradford
(Quarter 1 for years 2007 and 2011 as well as quarter 2 for year 2009 were extra/interpolated)

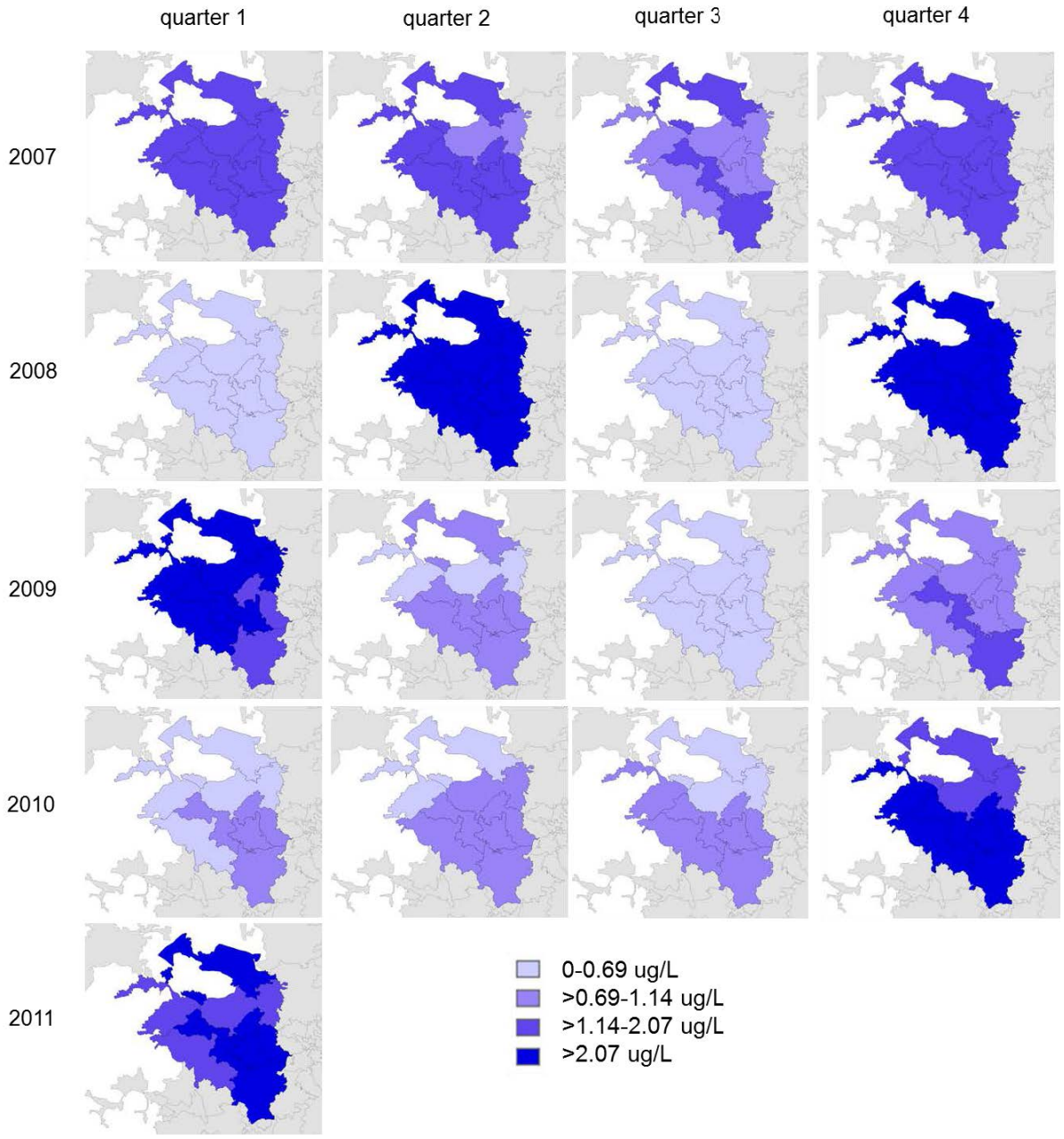
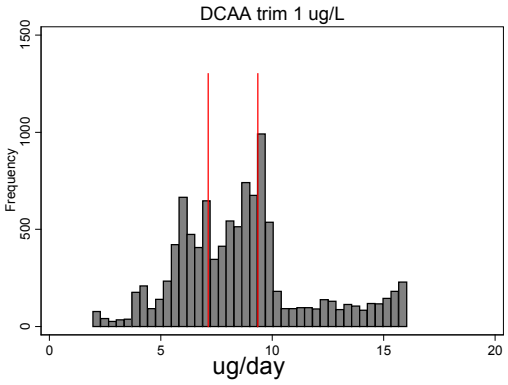
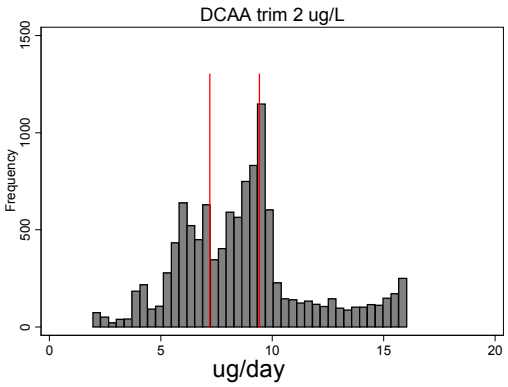


Figure 4.15: Frequency distributions of modelled trimester-specific HAA concentration by trimester

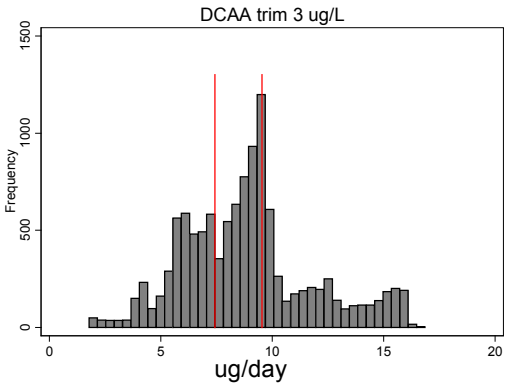
a) DCAA concentration in trimester 1 (cut-off tertile 1: 7.12 ug/L, cut-off tertile 2: 9.36 ug/L)



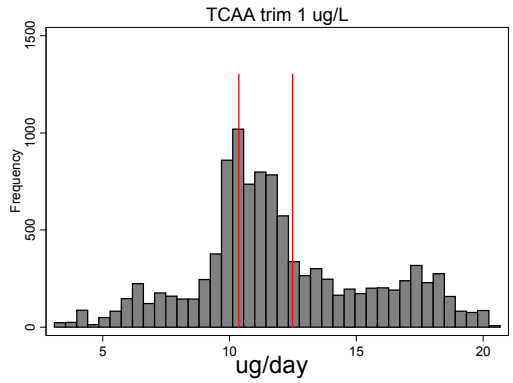
b) DCAA concentration in trimester 2 (cut-off tertile 1: 7.20 ug/L, cut-off tertile 2: 9.43 ug/L)



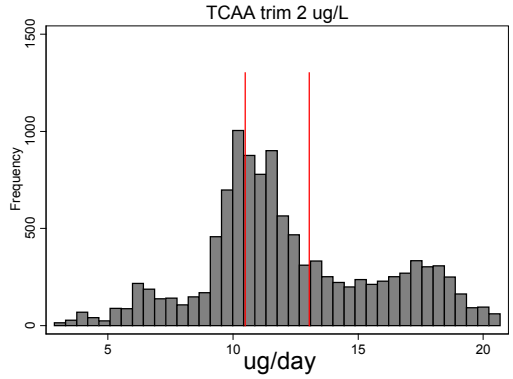
c) DCAA concentration in trimester 3 (cut-off tertile 1: 7.43 ug/L, cut-off tertile 2: 9.54 ug/L)



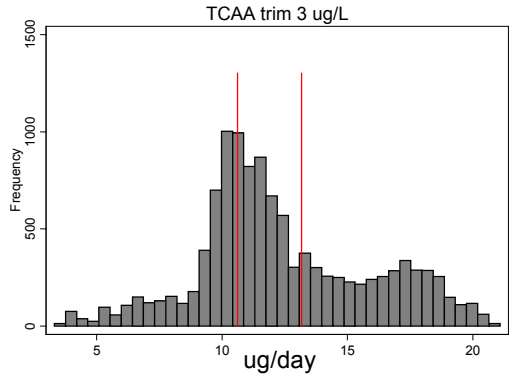
d) TCAA concentration in trimester 1 (cut-off tertile 1: 10.38 ug/L, cut-off tertile 2: 12.49 ug/L)



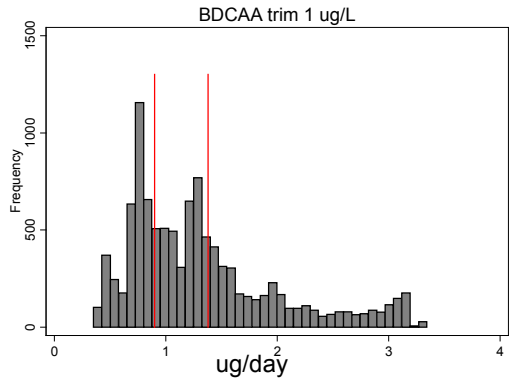
e) TCAA concentration in trimester 2 (cut-off tertile 1: 10.49 ug/L, cut-off tertile 2: 13.05 ug/L)



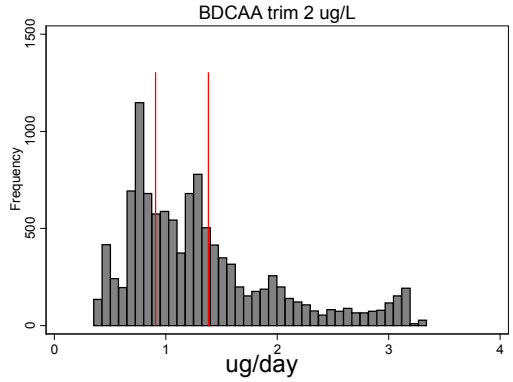
f) TCAA concentration in trimester 3 (cut-off tertile 1: 10.62 ug/L, cut-off tertile 2: 13.17 ug/L)



g) BDCAA concentration in trimester 1 (cut-off tertile 1: 0.90 ug/L, cut-off tertile 2: 1.38 ug/L)



h) BDCAA concentration in trimester 2 (cut-off tertile 1: 0.91 ug/L, cut-off tertile 2: 1.38 ug/L)



i) BDCAA concentration in trimester 3 (cut-off tertile 1: 0.91 ug/L, cut-off tertile 2: 1.39 ug/L)

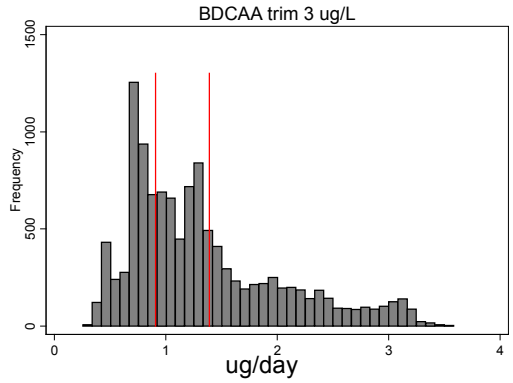
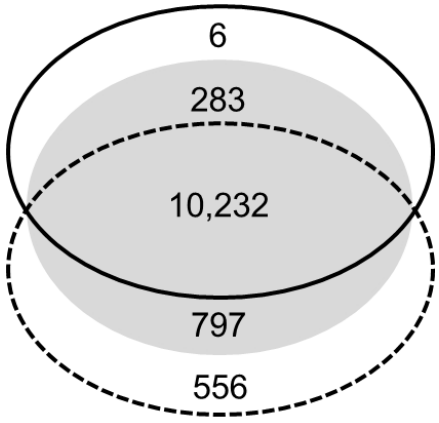


Figure 4.16: Numbers of women (out of 11,928) with a DCAA/TCAA/BDCAA concentration metric in each of their respective trimesters of pregnancy. Black line delineates trimester 1 (N=10,521), shaded area is trimester 2 (N=11,312), and dashed black line delineates trimester 3 (N=11,585)



CHAPTER 5 COMBINED METRIC

This chapter describes the method for combining total tap water consumption (Chapter 3) and area-level HAA concentrations for women's second trimester (Chapter 4) into one combined estimate of exposure to HAAs, completing Aim 1of this thesis (see Thesis Aims, Chapter 2). It also summarises the combined exposure estimates for the BiB cohort.

5.1 Background & Methods

As discussed in Chapter 1 (see section 1.1.1.5), failure to account for water consumption (Smith 2011) or for area-level concentrations (Savitz et al. 1995) can result in exposure misclassification. This is why individual total tap water consumption was combined with area-level DCAA, TCAA and BDCAA concentrations.

5.1.1 Boiling and filtering

Filtering and boiling of water may modify the concentrations of HAAs in the water, beverages or foods consumed, affecting the ingested dose of HAAs. These processes must therefore be taken into consideration.

5.1.1.1 Boiling

Though non-volatile, HAAs can be destroyed by chemical reactions at elevated temperatures.

Boiling tap water has a counter effect on the two main species of HAAs: increasing DCAA levels and decreasing TCAA levels (Chowdhury et al. 2010; Levesque et al. 2006; Ma 2008; Wu et al. 2001), such that average total concentrations of HAAs were found by several studies not to change (Chowdhury et al. 2010; Levesque et al. 2006).

These results might be explained by the two following phenomena occurring while boiling water: on one hand, the reaction of residual chlorine with DBP precursors favours the formation of DCAA (Krasner and Wright 2005); on the other hand, a reaction of decarboxylation of TCAA with chloroform[15] can occur favouring a decrease in levels of these substances (Zhang and Minear 2000).

Not all studies found this same result. Krasner and Wright (2005) observed an increase in DCAA levels after boiling water but did not observe a reduction of TCAA levels, and thus reported an overall

[15] $CCl_3COOH \rightarrow CHCl_3 + CO_2$

increase of total concentrations of HAAs. Kim (1997) observed no change in DCAA levels in distilled water following a 5-min boil but did observe some degradation of TCAA. Dojlido et al (1999) and Rahman et al (2011) both report decreases in all species of HAAs in tap water.

In general, HAA destruction reactions dominate at higher temperatures and for more highly halogenated (and more thermally labile) species, while HAA formation is more important at lower temperatures and for less halogenated species (Wu et al. 2001).

Averaging eight study results for DCAA and TCAA (Chowdhury et al. 2010; Dojlido et al. 1999; Kim 1997; Krasner and Wright 2005; Levesque et al. 2006; Ma 2008; Rahman et al. 2011; Wu et al. 2001), and two study results pertaining specifically to BDCAA (Krasner and Wright 2005; Ma 2008), DCAA was increased by 43.5%, and TCAA and BDCAA reduced by 36.9% and 56.5%, respectively (Table 5.1).

Type of boiling (kettle, microwave), duration of boiling, and amount of residual chlorine in the system are not taken into account.

Storage over time (whether in a refrigerator or not) has no effect on HAA concentrations (Chowdhury et al. 2010; Levesque et al. 2006; Rahman et al. 2011; Wu et al. 2001).

5.1.1.2 Filtering

Levels of HAAs treated with commercial point-of-use (POU) filtration devices such as activated carbon are significantly lower than in source tap water (Chowdhury et al. 2010; Levesque et al. 2006; Weinberg et al. 2006).

Egorov et al (2003)'s pilot study found that use of home filters (Brita™ [16], Aquaphor™ table-top pitcher filters (Russian-made), and a detachable faucet-tip Rodnik filter), all of which were mounted with cartridges which contained activated carbon, resulted in reduction of exposures to HAAs by a factor of 3.

The POU devices used in Weinberg et al's study (2006) were four American activated carbon filters (Brita Ultra model FF-100 (Oakland, CA, USA), PUR® Ultimate (Minneapolis, MN, USA), Teledyne Water Pik (Fort Collins, CO, USA), and Brita Pitcher (Oakland, CA, USA)). Among all HAAs, DCAA and TCAA had among the poorest average removal efficiencies. This is consistent with the higher solubility and polarity of these compounds compared to their bromine-containing counterparts.

[16] Brita filters also contain ion-exchange resins for remove of heavy metals, and thus a smaller amount of activated carbon (Egorov et al. 2003)

The latter had better removals presumably because their lower aqueous solubility made them more amenable to carbon adsorption (Weinberg et al. 2006).

A small study conducted at Imperial College London on tap water from the laboratory of the South Kensington campus also found that Brita Fjord filters (new and old) as well as PUR filters (new and old) effectively removed HAAs in water (Ma 2008). Similarly, water filtered using a domestic jug filter (fitted with ion exchange and activated carbon filtration) on chloraminated tap water in Sydney, Australia, resulted in large decreases (77-94%) in all species of HAAs in tap water (Rahman et al. 2011). Kim (1997) reported an average removal efficiency of 74% for DCAA and 71% TCAA for paired samples from six homes, but did not examine removal efficiency over the life of the filters (in Wright et al. (2006)).

The efficiency of home filters may vary depending on filter brand, age and characteristics of tap water (such as pH) (Egorov et al. 2003; Levesque et al. 2006). For example reductions of HAAs of 71% and 58% were observed using new vs. old Brita Classic Pitcher POU filters, respectively (Chowdhury et al. 2010). In Savitz's study (2005), filtration devices removed greater than 60% of HAA9 initially, but removal varied and declined about 20% over the 40 gallon pitcher POU capacity. HAA removal efficiency increases as the degree of bromination and halogenation increases.

Filter life, filtering frequency or specific filter types could not be taken into consideration here as no such detailed information was available. Instead, for DCAA and TCAA 17 results from seven filter studies were averaged (Chowdhury et al. 2010; Egorov et al. 2003; Kim 1997; Levesque et al. 2006; Ma 2008; Rahman et al. 2011; Savitz et al. 2005) using different types of both new and artificially aged filters, and for BDCAA, six results from two studies were used (Ma 2008; Savitz et al. 2005) to get reduction factors of 61.8% , 67.4% and 78.5% for DCAA, TCAA and BDCAA, respectively (Table 5.1).

5.1.2 Calculations

The following section explains how the combined metrics (in ug/day) for DCAA, TCAA, and BDCAA were calculated for each BiB woman.

Filtering of the cold water component of total tap water (an aggregate measure of cold and hot tap water) consumed at all locations (a combination of home, work, and other ("elsewhere") locations) was taken into account—mindful as previously that home and work are the only possible filtering locations for employed women, while home is the only possible filtering location for women out of

employment, for students, and for those missing employment information. Boiling of the hot component of total tap water was also uniformly taken into account.

For each HAA, the total tap water consumption amount (in L/day) was calculated by location (home, work, elsewhere), integrating filtering reduction factors for cold tap water ingestion (Table 5.1) when necessary, and boiling factors for total hot tap water intake. This calculation is described in Table 5.2.

To get the combined metric (ug/day), the consumption values adjusted for filtering and/or boiling factors were summed, and each one was then multiplied by the individual woman's HAA area-level concentration (ug/L), time-weighted over her second trimester of pregnancy (N=2,477). This was done differently for employed women (who filtered at home and work, only at home, only at work, nowhere) and unemployed women (as well as students and women with missing information). Table 5.3 is an example of the combined metric calculation for DCAA exposure (ug/day). Together, the two mutually exclusive groups (DCAA exposure among filterers ($ExpDCAA_f$,) and DCAA exposure among non-filterers ($ExpDCAA_u$)) make up total DCAA exposure (ExpDCAA). The same method applies to calculation of ExpTCAA and ExpBDCAA. These calculations exclude uncertain answers, as well as missing or inconsistent filtering information.

For the employed women, total filtered tap water (TTw_f) and total unfiltered tap water (TTw_u) consumed at work were either multiplied:

- by $[HAA]_{res}$, the HAA concentrations from women's residence WSZ concentrations only (N_f=1012, N_u=5211) (Method 1), or
- when available, by $[HAA]_{workres}$ which is the HAA concentration based on both work and residence WSZ concentrations (weighted by days reported spent at work) (N_f=395, N_u=441); for those whose work addresses were not geocoded or not within area, total tap water at work was multiplied by the HAA concentration from women's residence water supply zones only (N_f=617, N_u=4770) (Method 2) (see section 4.2.3.2).

  No HAA concentrations ($[HAA]_{workres}$) was taken into account for the women out of employment, students or women missing information, as these data were either not relevant or not available for them (see section 4.2.3.3).

5.2 Results

Combined metric levels in ug/day for 6223 women are presented in Table 5.5. The combined metrics were categorised here again by tertiles for the purposes of the epidemiologic analysis (Figure 5.1).

The combined metrics can be divided into 1012 filterers, and 5211 non-filterers. Compared to the sample size of filterers and nonfilterers reported in the water consumption chapter (Chapter 3), these figures represent a 10% and a 5% loss, respectively. Differences are due either to missing area-level concentration (due to lack of overlap with modelled exposure period) or missing water consumption information. If either the concentration or the water consumption value was missing, then the multiplication was not done (Figure 5.2).

Because reduction factors for boiling apply to all women's hot beverages, boiling is not as important a factor in modifying the combined metric levels as filtering is. Indeed filterers are exposed to 77% the DCAA exposures, 72% the TCAA exposures and 63% the BDCAA exposures on average compared to non-filterers (Table 5.6), decreases which are consistent with the magnitude of reduction factors: BDCAA>TCAA>DCAA reported in Table 5.1.

25% (251, or 250 for BDCAA, of 1012) of $ExpHAA_f$ and 5% (273 of 5211) of $ExpHAA_{nf}$ values differed depending on whether the calculation was based on WSZ concentrations at the residence only or at the combination of work and residence. This has to do with a combination of factors: few workplace WSZs were geocoded within Bradford to begin with, but a large proportion of women also work and live in the same WSZ resulting in no change to their exposure levels by this method (see Table A5 - 1 and Table A5 - 2).

Table 5.7 compares (using Spearman correlation coefficients) the number of people who fall into overlapping tertiles (low/medium or high) for exposure classification by ingestion, area-level concentrations and the combined metrics.

5.3 Discussion

In her PhD thesis, Smith (2011) developed the idea of the combined exposure metric which combines area-level exposure data and individual level water consumption data. When analysing the combined metric she created for THM exposures in the BiB cohort, she included interaction terms between Month, Year and WSZ in the ANOVA models, in order to separate all spatial and temporal components of variability from the residual. By isolating the variability in individual water use from other components of variability, she found that individual variability in water use is the most influential factor driving THM exposure, much more so than any temporal or spatial variation in THM concentrations at the tap (Smith 2011). This stark result was related to the fact that there is very little spatial variation in Bradford. The same is suspected for HAAs (see Chapter 4). Individual exposures to HAAs are therefore likely determined by total ingestion of tap water (non-boiled and boiled), hot drinks, and tap water-containing foods (Egorov et al. 2003), more so than by

concentration in the area of residence. Among volume of water intake, bottled and filtered water consumption, and effectiveness of point-of-use filtration in the home, Wright et al. (2006) also reported that volume of water intake was the most influential modifier of ingestion exposures. Information on tap-water containing foods was unfortunately not available from the BiB questionnaire.

Because ingestion is the major route of exposure to HAAs, accounting for uptake factors to transform exposures to ingested dose (taken up by the body) was deemed not necessary. Unless there is a big known difference between the uptake of DCAA, TCAA and BDCAA which current literature neither addresses nor shows, the added constant will not make any difference to the exposure estimates.

After comparing the number of people who fall into overlapping tertiles (low/medium or high) for exposure classification by ingestion, area-level concentrations and the combined metrics, total tap water and HAA concentrations do not have much agreement but total tap water and combined metrics are highly correlated (see the pink shade in Table 5.7). Area-level HAA concentrations for trimesters 1 and 2, and for trimesters 2 and 3 are correlated but not so much for trimesters 1 and 3 (though all of these correlations reach p-values<0.005) (see the brown shade in Table 5.7). All three combined metrics are highly correlated (blue shade in Table 5.7) while also being significantly correlated with all other exposure measures (except for one exception: DCAA combined and [TCAA] trim 1). Area-level HAA concentrations for trimester 2 and its respective combined HAA metric are highly correlated (as expected given that the combined metric was derived using trimester 2 area-level HAA concentration, see Table 1.4). The correlations are greater between the combined DCAA and TCAA metrics and total water than between the combined DCAA and TCAA metrics and area-level DCAA and TCAA concentrations, respectively; however the correlation coefficient is higher between the combined BDCAA and the area-level concentration for BDCAA, than between the combined DCAA and total tap water.

5.4 Tables

Table 5.1: Boiling and filtering factors used in the combined metric

	factors		% change	
	BOILING	**FILTERING**	**BOILING**	**FILTERING**
DCAA	1.44	0.38	43.51	-61.84
TCAA	0.63	0.33	-36.85	-67.43
BDCAA	0.44	0.22	-56.50	-78.50

Table 5.2: Filtering and boiling integrated into the water consumption variables

Definition	**Notation**	**Calculation**
Total filtered tap water consumed at home	TTh_f	$f(cold_h) + b(hot_h)$
Total filtered tap water consumed at work	TTw_f	$f(cold_w) + b(hot_w)$
Total unfiltered tap water consumed at home	TTh_u	$cold_h + b(hot_h)$
Total unfiltered tap water consumed at work	TTw_u	$cold_w + b(hot_w)$
Total unfiltered tap water consumed elsewhere	TTe_u	$cold_e + b(hot_e)$

Legend:
f filtering factor
b boiling factor

Table 5.3: Combined metric calculation for DCAA exposure (ug/day) using the residence water supply zone (WSZ) (Method 1)

		Calculation			Notation
	Filtering by location	**Total tap water ingestion, adjusted for filtering and/or boiling**	**x**	**Residence WSZ concentrations**	
Working women	HW	$(TTh_f + TTw_f + TTe_u)$	x	$[DCAA]_{res}$	$ExpDCAA_{1\ hw}$
	H~~W~~	$(TTh_f + TTw_u + TTe_u)$	x	$[DCAA]_{res}$	$ExpDCAA_{1\ h}$
	~~H~~W	$(TTh_u + TTw_f + TTe_u)$	x	$[DCAA]_{res}$	$ExpDCAA_{1\ w}$
	~~HW~~	$(TTh_u + TTw_u + TTe_u)$	x	$[DCAA]_{res}$	$ExpDCAA_{1\ nf}$
Not working, and students	H	$(TTh_f + TTe_u)$	x	$[DCAA]_{res}$	$ExpDCAA_{2\ f}$
	H	$(TTh_u + TTe_u)$	x	$[DCAA]_{res}$	$ExpDCAA_{2\ nf}$

Notation	**Calculation**
$ExpDCAA_f$	$ExpDCAA_{1\ hw} + ExpDCAA_{1\ h} + ExpDCAA_{1\ w} + ExpDCAA_{2\ f}$
$ExpDCAA_{nf}$	$ExpDCAA_{1\ nf} + ExpDCAA_{2\ nf}$

H: means the woman answered Yes to the filtering at home question; ~~H~~: means she answered no to the filtering at home question (and in some cases, don't know, NA, or missing). Ditto for W, workplace.

Shaded variables sum to DCAA exposure for filterers ($ExpDCAA_f$); unshaded variables sum to DCAA exposure for nonfilterers ($ExpDCAA_{nf}$)

(See Table 5.2 for definitions of TTh_f, TTh_u, TTw_f, TTw_u, TTe_f and TTe_u)

Table 5.4: Example of a combined metric calculation for DCAA exposure (ug/day) using a combination of work and residence water supply zone (WSZ) area-level concentrations (Method 2)

		Calculation						
	Filtering by location	**Total tap water ingestion at home and elsewhere***	**x**	**Residence WSZ concentrations**	**+**	**Total tap water ingestion at workplace***	**x**	**Work place WSZ concentrations**
Working women	HW	$(TTh_f + TTe_u)$	x	$[DCAA]_{res}$	+	TTw_f	x	$[DCAA]_{workres}$
	H~~W~~	$(TTh_f + TTe_u)$	x	$[DCAA]_{res}$	+	TTw_u	x	$[DCAA]_{workres}$
	~~H~~W	$(TTh_u + TTe_u)$	x	$[DCAA]_{res}$	+	TTw_f	x	$[DCAA]_{workres}$
	~~HW~~	$(TTh_u + TTe_u)$	x	$[DCAA]_{res}$	+	TTw_u	x	$[DCAA]_{workres}$
Not working /students	Same as previous table (Table 5.3)							

*adjusted for filtering and/or boiling
(See Table 5.2 for definitions of TTh_f, TTh_u, TTw_f, TTw_u, TTe_f and TTe_u)

Table 5.5: Summary of combined metric (total) based on residence concentrations only in ug/day

					Percentile Distribution				
n total = 11,928		mean	sd	min	25th %ile	Median	75th %ile	max	n
Exposure to DCAA	ExpDCAA	16.02	10.7	0.4	9.1	13.6	20.3	158.7	6223
Exposure to TCAA	ExpTCAA	18.15	11.9	0.4	10.3	15.5	23.1	131.7	6223
Exposure to BDCAA	ExpBDCAA	1.80	1.4	0.0	0.9	1.4	2.3	17.5	6223

Table 5.6: Summary of combined metric for filterers and non-filterers based on residence concentrations only in ug/day

					Percentile Distribution				
n total = 11,928		mean	sd	min	25th %ile	Median	75th %ile	max	n
Exposure to DCAA among filterers	$ExpDCAA_f$	12.85	10.4	0.4	6.4	10.1	16.3	158.7	1012
Exposure to DCAA among non-filterers	$ExpDCAA_u$	16.64	10.6	0.8	9.7	14.2	20.9	121.0	5211
Exposure to TCAA among filterers	$ExpTCAA_f$	13.70	11.5	0.4	6.5	10.8	17.1	131.7	1012
Exposure to TCAA among non-filterers	$ExpTCAA_u$	19.01	11.8	0.8	11.2	16.4	23.9	129.7	5211
Exposure to BDCAA among filterers	$ExpBDCAA_f$	1.20	1.2	0.0	0.4	0.9	1.5	17.5	1012
Exposure to BDCAA among non-filterers	$ExpBDCAA_u$	1.91	1.4	0.1	1.0	1.5	2.4	14.5	5211

Table 5.7: Pairwise Spearman correlation between all 13 exposure measures (rho, p-value and sample size), categorised by tertile

rho p-value N	Total Tap water (L/day)	[DCAA] in trimester 1 (ug/L)	[DCAA] in trimester 2 (ug/L)	[DCAA] in trimester 3 (ug/L)	[TCAA] in trimester 1 (ug/L)	[TCAA] in trimester 2 (ug/L)	[TCAA] in trimester 3 (ug/L)	[BDCAA] in trimester 1 (ug/L)	[BDCAA] in trimester 2 (ug/L)	[BDCAA] in trimester 3 (ug/L)	combined DCAA (ug/day)	combined TCAA (ug/day)	combined BDCAA (ug/day)
[DCAA] in trimester 1 (ug/L)	-0.01 0.501 8,487	1.00 10,521											
[DCAA] in trimester 2 (ug/L)	-0.01 0.277 9,185	0.43 <0.001 10,515	1.00 11,312										
[DCAA] in trimester 3 (ug/L)	-0.02 0.074 9,452	-0.02 0.022 10,232	0.44 <0.001 11,029	1.00 11,585									
[TCAA] in trimester 1 (ug/L)	0.00 0.789 8,487	0.29 <0.001 10,521	0.06 <0.001 10,515	-0.29 <0.001 10,232	1.00 10,521								
[TCAA] in trimester 2 (ug/L)	0.02 0.054 9,185	0.36 <0.001 10,515	0.26 <0.001 11,312	0.09 <0.001 11,029	**0.56*** <0.001 10,515	1.00 11,312							
[TCAA] in trimester 3 (ug/L)	0.01 0.201 9,452	**0.63*** <0.001 10,232	0.32 <0.001 11,029	0.25 <0.001 11,585	0.14 <0.001 10,232	**0.54*** <0.001 11,029	1.00 11,585						
[BDCAA] in trim 1 (ug/L)	-0.03 0.002 8,487	0.07 <0.001 10,521	-0.25 <0.001 10,515	-0.25 <0.001 10,232	0.18 <0.001 10,521	-0.20 <0.001 10,515	-0.21 <0.001 10,232	1.00 10,521					
[BDCAA] in trim 2 (ug/L)	0.00 0.935 9,185	0.12 <0.001 10,515	0.13 <0.001 11,312	-0.22 <0.001 11,029	0.50 <0.001 10,515	0.21 <0.001 11,312	-0.21 <0.001 11,029	0.25 <0.001 10,515	1.00 11,312				
[BDCAA] in trim 3 (ug/L)	0.02	0.34	0.11	0.20	0.22	**0.54***	0.26	0.03	0.24	1.00			

rho p-value N	Total Tap water (L/day)	[DCAA] in trimester 1 (ug/L)	[DCAA] in trimester 2 (ug/L)	[DCAA] in trimester 3 (ug/L)	[TCAA] in trimester 1 (ug/L)	[TCAA] in trimester 2 (ug/L)	[TCAA] in trimester 3 (ug/L)	[BDCAA] in trimester 1 (ug/L)	[BDCAA] in trimester 2 (ug/L)	[BDCAA] in trimester 3 (ug/L)	combined DCAA (ug/day)	combined TCAA (ug/day)	combined BDCAA (ug/day)
	0.023 9,452	<0.001 10,232	<0.001 11,029	<0.001 11,585	<0.001 10,232	<0.001 11,029	<0.001 11,585	0.010 10,232	<0.001 11,029	 11,585			
combined DCAA (ug/day)	**0.66***	0.21	0.44	0.23	0.02	0.16	0.21	-0.14	0.05	0.11	1.00		
	<0.001	<0.001	<0.001	<0.001	0.064	<0.001	<0.001	<0.001	<0.001	<0.001			
	6,223	5,734	6,223	6,061	5,734	6,223	6,061	5,734	6,223	6,061	6,223		
combined TCAA (ug/day)	**0.66***	0.17	0.17	0.12	0.26	0.43	0.31	-0.10	0.05	0.25	**0.68***	1.00	
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001		
	6,223	5,734	6,223	6,061	5,734	6,223	6,061	5,734	6,223	6,061	6,223	6,223	
combined BDCAA (ug/day)	**0.50***	0.06	0.15	-0.09	0.33	0.17	-0.10	0.14	**0.56***	0.18	**0.53***	**0.58***	1.00
	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	
	6,223	5,734	6,223	6,061	5,734	6,223	6,061	5,734	6,223	6,061	6,223	6,223	6,223

Bold with * means rho value>0.5; red indicates that a p-value < 0.005

Legend:

- total tap water and the combined metrics are highly correlated
- [HAA] trimesters 1 and 2, and trimesters 2 and 3 are correlated; trimesters 1 and 3 less so
- the 3 combined metrics are highly correlated
- trimester 2 of each [HAA] and its respective combined [HAA] are highly correlated (as expected)

5.5 Figures

Figure 5.1: Histograms with marked tertile cut points (N=6,223 for each of DCAA, TCAA, BDCAA)

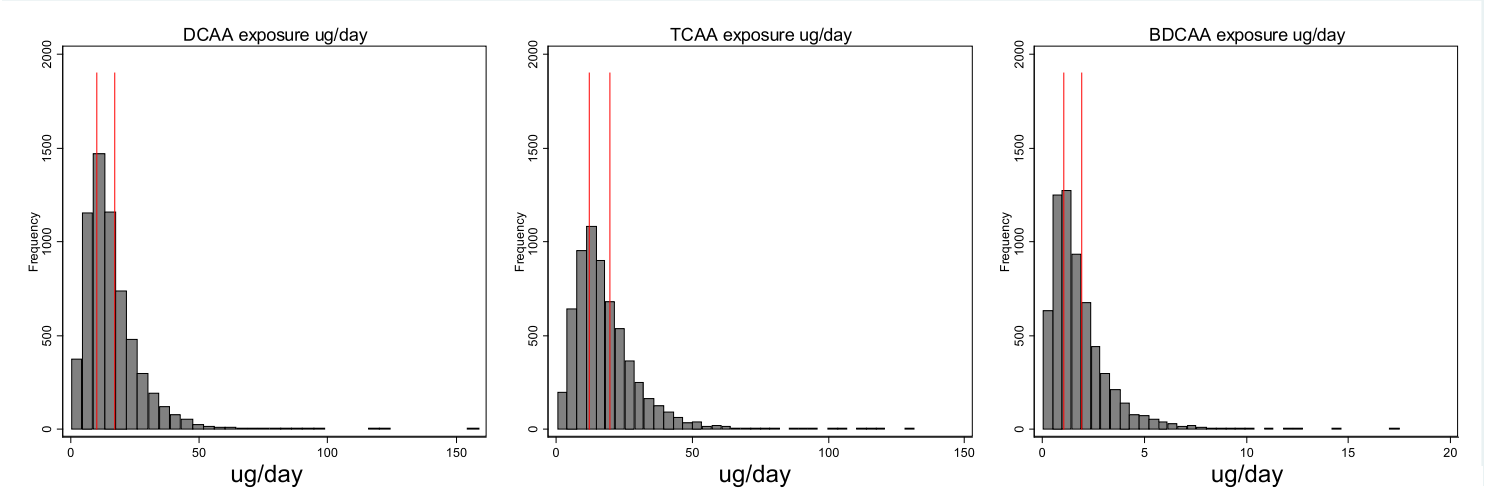
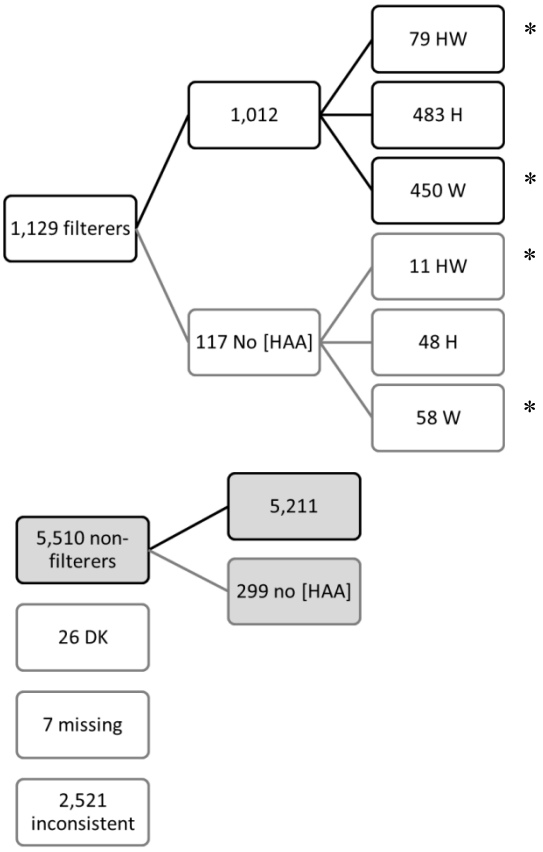


Figure 5.2: Filterers vs. non-filterers, and location of filtering if applicable (based on data presented in Table 3.5)



Legend:
HW: filtered at home and at work; H: filtered at home; W: filtered at work; DK: doesn't know filtering status; *to be eligible to be a work filterer, must be employed

CHAPTER 6 BIRTH OUTCOMES IN BORN IN BRADFORD

Setting the scene for Chapter 7, this chapter has multiple goals. It aims to a) describe the BiB birth outcomes, i.e. the distributions and rates of the three birth outcomes chosen for this thesis (birth weight, term low birth weight (term LBW), and small-for-gestational age (SGA)), as well as some obstetric outcomes, b) highlight differences in outcomes by ethnicity, and discuss whether this could be due to ethnicity itself or factors which correlate with ethnicity (e.g. environment), c) discuss the pros and cons of the three chosen outcomes in assessing fetal growth, and d) explore the representativeness of the BiB cohort with respect to Bradford, and England and Wales.

6.1 Background

This thesis examines the following common indicators of suboptimal growth during the fetal period: birth weight as a continuous measure, term low birth weight (term LBW), and small size for gestational age (SGA), all three of which are birth weight-based measures (see definitions in the Introduction (Chapter 1), section 1.1.2.1). Rates of preterm delivery (PTD), low birth weight (LBW) and very LBW (birth weight less than 1500g) are reported in this chapter for purposes of comparison with other studies.

The proportion of births delivered preterm (including stillbirths) in most developed countries is 5-9% (Goldenberg et al. 2008): 5% in the Nordic countries (Morgen et al. 2008), 7% in Hong Kong Chinese (Leung et al. 1998), and 12% in the United States with higher rates among Blacks compared to Whites (Martin et al. 2012). Moser et al (2008) reported an average 6.2% rate of preterm delivery among 621,793 live singletons in England and Wales in 2005.

In 2011, 7% of all live births in England & Wales were LBW according to the ONS (2012). While in 2005, Moser et al (2008) reported that 6.1% of 624,821 live singletons were < 2500g by weight at birth. According to the 2010 US National Vital Statistics, 8.1% of all US babies are LBW (Martin et al. 2012) (Table 6.1).

6.1.1 Ethnic differences in birth weight

6.1.1.1 Ethnic differences in fetal growth based on ultrasound studies

Many studies conducted in the UK have demonstrated shorter gestational ages and lower birth weights among ethnic minority groups in the UK, and in particular among individuals of South Asian

origin (Alvear and Brooke 1978; Chetcuti et al. 1985; Harding et al. 2004; Margetts et al. 2002; McFadyen et al. 1984; Patel et al. 2004; Wilcox et al. 1993). Differences in fetal growth patterns between ethnic groups have long been recognised (Cole et al. 1998; Rona and Chinn 1986).

Serial ultrasound measurements of biparietal diameter[17], head circumference and abdominal circumference *in utero* show that Black fetuses have significantly longer femur lengths than White fetuses (Davis et al. 1993; Shipp et al. 2001), and that Asian fetuses tend to have shorter femur lengths than White fetuses (Shipp et al. 2001). A study which compared humerus[18] lengths by ethnic group also found differences among African American and Asian (but not among Hispanic) fetuses, in comparison to White fetuses (Mastrobattista et al. 2004). These studies suggest that there may be differences in body length proportions that are important factors in understanding birth weight differences by ethnicity.

An ultrasound study of fetal growth patterns in women of Indian origin showed that the mean abdominal circumference measurements of the fetus throughout pregnancy were significantly smaller than measurements obtained in White European fetuses. The mean birth weight of the Indian babies was 340g less than that of the controls and Indian patients also exhibited a significantly shorter duration of gestation (Meire and Farrant 1981). Another study compared fetal growth curves for Bangladeshi mothers from serial ultrasonic estimates of both abdominal circumference and estimated fetal weight to the growth curve derived from a cohort of White Anglo-Saxon women (Gallivan et al. 1993). While average fetal weights at 28, 32, and 36 weeks as well as the birth weights were significantly different between the two groups, the study suggested that growth rates were similar such that even fetuses of uncomplicated Bangladeshi pregnancies were just "constitutively" smaller (Spencer et al. 1995). Either that and/or fetal maturity— i.e. the process of achieving full development or growth—occurs earlier in gestation in Black and South Asian compared to White European babies (Balchin and Steer 2007; Patel et al. 2004).

6.1.1.2 Nature versus Nurture

The debate over whether differences in birth weight—in particular between women of South Asian and White British origins—are due to particular environmental experiences such as diet and lifestyle, or whether they have a genetic underpinning is ongoing. The answer most likely involves a bit of both.

The effect of nutritional deficiencies on fetal size is complex. Naeye and Tafari (1985) studied the effect of maternal malnutrition, as assessed by skinfold thickness, on birth weight in an Ethiopian

[17] transverse diameter of the head
[18] the long bone in the arm running from shoulder to elbow

population. Women of comparable height with lower skinfold thickness had significantly lighter babies. Similarly, a 2010 study on 2394 Jamaican women found that poor maternal nutrition as indicated by low weight, height, and BMI is associated with smaller, shorter babies with smaller heads (Thame et al. 1997). Yet another study however concluded that this is only true for extreme malnutrition (Poppitt et al. 1994).

West (2011) reports that mean birth weights have increased and that the proportion of LBW babies in high income countries has fallen over the last century (Alberman 1991; Chike-Obi et al. 1996; Chowdhury et al. 2000). However, rates of LBW across low income countries remain high: around 20 million infants worldwide are born LBW (United Nations Children's Fund 2004), 95% of whom are born within low income countries. But there is still variation between low income regions, for example sub-Saharan African nations have reported a lower percentage of LBW births than South Asian nations despite their United Nations (UN) classification among the “least developed” countries. In fact, half of all LBW babies born in the world were born in South Asia (United Nations Children's Fund 2004). This variability, even amongst regions likely to experience equivalent nutritional challenges, suggests that LBW is probably not just a consequence of maternal nutrition and environment, as is generally thought (although it is in the extreme (Poppitt et al. 1994)).

This genetic underpinning for differences in birth weight is supported by the persistence of ethnic differences in migrant populations, as articulated by West (2011). For example, babies born to South Asian mothers in the US, Europe, New Zealand and UK weigh significantly less than the indigenous population. Within the US, a number of studies have also reported smaller mean birth weights and higher proportions of LBW among South Asians, particularly Asians of Indian origin (Fuentes-Afflick and Hessol 1997; Hayes et al. 2008; Madan et al. 2002). Interestingly, in the US, South Asian Indians are generally more affluent than other migrant groups but factors usually found to have a protective influence against LBW such as high levels of education and high social economic status, are reportedly not protective among US Asian Indians (Alexander et al. 2007; Gould et al. 2003). By contrast, other migrant groups in the US, for example Mexicans and Hispanics appear to have low rates of LBW despite seemingly unfavourable social and economic circumstances (Gould et al. 2003; Rosenberg et al. 2005). In these groups migrant status seems to confer some advantage in terms of birth weight whereas the continued lower birth weights of South Asians again suggests either an element of genetic predisposition or persistent environmental exposures after relocation to the host country. Similar trends have been observed in other high income countries including New Zealand (McCowan et al. 2004) and Singapore (Hughes et al. 1986). In Europe, a study of birth weight in Norway identified the rate of LBW to be higher among Pakistanis than any other ethnic group (Vangen et al. 2002).

In a large cohort in England and Wales, mean birth weights in 2005 were 3075g, 3082g and 3130g for Bangladeshis, Indians and Pakistanis, respectively. This compares to a mean birth weight of 3393g in the White British population (Moser et al. 2008). Furthermore, rates of LBW in Indians, Pakistanis and Bangladeshis in the UK (10.5%, 9.8% and 10.0% respectively) are almost double the rate for White British infants (5.6%). While these data do not exclude preterm births, ethnic differences persisted when analyses were restricted to births at 40 weeks gestation (Moser et al. 2008). Similar differences were reported for term babies in the UK Millenium Cohort Study (Kelly et al. 2009) which also identified significant differences in the prevalence of term LBW between South Asian and White British populations, particularly between Pakistanis compared to White British (Table 6.1 part b).

In the BiB cohort, marked differences in birth weight between Pakistani origin and White British origin infants have been found to persist even after adjustment for a wide range of potential masking and mediating characteristics, and important differences remain whether both parents were UK born, one was South Asian born or both were South Asian born (West 2011). This suggests that at least over two generations, environmental or lifestyle changes amongst parents who migrated to the UK and spent all of their life in the UK have not had a major impact on these differences.

Five other studies have compared differences in birth weight between first and second generation South Asian women in the UK compared to White babies. Four reported no increase in birth weights (Draper et al. 1995; Harding et al. 2004; Leon and Moser 2012; Margetts et al. 2002), and one study, the smallest with a sample of 331, found higher mean birth weights in second compared to first generation South Asian offspring (adjusted mean difference of 280g in birth weights) (Dhawan 1995). Despite this evidence in favour of a genetic underpinning to birth weight differences by ethnicity, West (2011) notes that these differences could also reflect persistent lifestyle and cultural behaviours that remain very similar in first and second generation mothers of close-knit South Asian communities. Differences in birth outcomes by ethnicity may also be affected by factors other than ethnicity and genes, including maternal exposures during pregnancy such as smoking and alcohol intake, maternal cardiovascular health and glycaemia, maternal size, and socio-economic factors (Kelly et al. 2009).

6.2 Methods

6.2.1 Data

All variables described in this chapter are defined in Chapter 2, sections 2.1.2 and 2.1.3.

The data described in this chapter (as in Chapter 7) were received in February 2013 from the BiB team and contained both the routine eClipse and the additional backfilled data.

Five babies whose gender could not be assigned at birth were excluded from the dataset over identifiability concerns. This procedure is in line with what was done in other cohort studies such as ALSPAC, the Avon Longitudinal Study of Parents and Children (Golding et al. 2001).

13,525 pregnancies had eClipse information. As in previous chapters, multiple births and multiple entries to BiB were excluded such that each woman was only counted once (see 3.2.3) resulting in a baseline sample size of 11874 eligible women with birth weight data.

6.2.2 Gestational age dating

Precise dating of a baby's age at birth is a prerequisite for any reliable measure of growth.

The ideal measure of gestational age would cover the period between day of conception, which takes place less than a day after ovulation (Wilcox 2010) and day of birth. Date of last menstrual period (LMP) and ultrasound dating are the two methods largely employed in practice, sometimes in combination, to estimate gestational age. Ultrasound dating is considered more accurate than menstrual dating (Gardosi et al. 1997), but both methods have their strengths and limitations.

Gestational age by LMP adds the number of days from the first day of the last normal menstrual period to delivery. By Naegele's rule, expected delivery is 280 days (40 weeks) from the first day of LMP, which assumes a 28 day cycle and ovulation on day 14.

Ultrasound dating measures the size of various parts of the embryo or fetus to estimate its weeks of gestational age (Gardosi et al. 1997). It is more accurate when done earlier in the pregnancy. The most accurate measurement for dating is the crown-rump length[19] of the fetus (which can be done between 7 and 13 weeks of gestation). After 13 weeks of gestation, fetal age may be estimated using the biparietal diameter, the head circumference, the length of the femur, the crown-heel length[20], and other fetal parameters.

6.2.3 SGA methods and derivation in BiB

Reference centile curves show the distribution of a measurement as it changes according to age.

[19] length from the top of the head (or crown) to the bottom of the buttocks (or rump)
[20] length from head to heel

Since 1995 the charts used in the UK have been based on the British 1990 (UK90) reference (Cole et al. 2011; Freeman et al. 1995) which pools birth data from five original studies collected between 1983 and 1993 and provides birth centiles for weight, length and head circumference from 23–44 weeks gestation.

When new UK–WHO charts replaced those based on the UK90 reference for children 0–4 years of age in May 2009 (CM Wright et al. 2010), the UK90 data had to be retained for preterm births as the WHO infants were all term births by design and the mean WHO birth weight was appreciably lower than in the UK (SACN/RCPCH Expert Group 2007), meaning that new birth centile charts for weight, length, and head circumference had to be constructed (Cole 2011).

Contrary to the original UK90-based birth centiles where birth data was amalgamated with postnatal data so that the centiles were smooth and uninterrupted from 23 weeks gestation through to 23 years (Cole et al. 1998), the construction of the new UK-WHO chart also necessitated reanalysis of the UK90 birth data without the postnatal data, as it became clear that birth and postnatal data need to be kept separate (Wright and Parkinson 2004). To do so, Cole et al. (2011) analysed the UK90 reference (by sex) using the LMS method. Briefly, the LMS method summarises the changing birth weight distribution by three curves representing the median (M), coefficient of variation (S) and skewness (L), the latter expressed as a Box-Cox power. Using penalised likelihood the method fits the three curves as cubic splines by non-linear regression; the extent of smoothing required can be expressed in terms of smoothing parameters or equivalent degrees of freedom (Cole and Green 1992). Centiles were generated for babies born 23–42 weeks of gestation, as the data were too sparse to generate birth centiles after 42 weeks (Cole et al. 2011).

The standard deviation scores (SDS) (or z-scores) of a child's birth weight measurement y can be calculated from the L, M and S curves, using values appropriate for the child's age and sex. I used the Microsoft Excel "LMSgrowth" add-in ("Measurement to/from SDS menu item") (Pan and Cole 2010) to define reference centiles for birth weight from 23–42 weeks (Cole et al. 2011). Serial ultrasound measurements of biparietal diameter, head circumference and abdominal circumference *in utero* confirm a slower rate of growth in the female fetus compared to that in the male fetus (Davis et al. 1993; Parker et al. 1984), which is why adjustment by sex was appropriate.

To calculate SGA, babies at or below the SDS cut-off corresponding to the tenth centile of a standard normal cumulative distribution, i.e. ≤ -1.282, were identified.

6.2.4 Statistical analysis

Pearson's chi-square tests, two-sample t-tests and one-way analysis of variance (ANOVA) tests were carried out in STATA 12.1.

6.3 Results

6.3.1 Frequency distribution of birth weight

Birth weight was recorded for 11,874 live singletons and was normally distributed. Mean birth weight was 3229.7g (95% CI: 3219.8, 3239.6) (Figure 6.1). The mean birth weight for babies born weighing 2500g or more, i.e. within a normal weight range, was 3322.3g (3313.9-3330.7); for LBW babies, it was 2092.8g (2065.7-2119.9).

After dividing birth weights into term and preterm births, the mean birth weight for term LBW babies was 2293.4g (2276.1-2310.8) (Figure 6.2, top figure). As expected, 50% (N=444) of 894 LBW babies were also preterm, but 33% (N=219) of the 663 preterm births were not LBW—as evidenced by the preterm births distributed across heavier birth weights (≥2500g) in Figure 6.2 (lower figure).

6.3.2 Preterm delivery (PTD), low birth weight (LBW), term LBW, and Small-for-Gestational Age (SGA)

7.5% of live singleton births in the BiB cohort were LBW, 0.7% very LBW, 5.6% preterm, and 4% term LBW (Table 6.2, Table 6.3, Table 6.4). These rates are the same whether considering all live births (N=13,199) irrespective of multiple births and whether women registered several births into the cohort (Figure 6.3 part a, Figure 6.4 part a, Figure 6.5 part a), or the set of 11,928 women with singleton live births (Figure 6.3 part b, Figure 6.4 part b, Figure 6.5 part b). Hereafter my analyses focus on the latter set of women as they are the focus of Chapter 7's epidemiologic analyses. Including both singletons and multiple live births, 8.5% were LBW and 0.9% was very LBW.

12.1% of live singleton births (born between 23-42 gestational weeks, N=11,863) are below the tenth centile of standardised birth weights, i.e. considered small-for-gestational age (SGA) (Table 6.5, Figure 6.5). SGA babies weighed on average 2547.4g (95% CI: 2526.5, 2568.4) compared to 3323.8g (3314.2-3333.3) for babies considered not SGA. Figure 6.6 shows the distribution of standard deviation scores (SDS) which has a mean a bit below zero at -0.17 (95% CI: -0.19, -0.15).

6.3.3 Obstetric outcomes

69% of births were spontaneous births, 10% had no labour, and 20% were induced (either medically, surgically, or both). Caesarean sections accounted for 23% of singleton births, a little under half of which are due to lack of spontaneous labour onset (Table 6.6).

Rates of pre-existing hypertension and hypertension with onset during labour only were low in the eligible set of mothers (1.0% and 0.7%, respectively). But 5.7% of mothers developed hypertension during the pregnancy (4.3% mild to moderate, 1.2% severe, and 0.1% unclassified) (Table 6.6). A greater proportion of these pregnancy-induced hypertensive women underwent a caesarean birth (43%) compared to non-hypertensive women (22%) ($p<0.001$ by chi^2 test). Rate of pre-eclampsia was 2.7% in this subset (Table 6.6). 2% (N=243) of BiB women had both pre-eclampsia and pregnancy-induced hypertension, and another 4% (N=448) had either one or the other.

While <1.0% of women reported suffering from diabetes prior to pregnancy, 8.1% of eligible women carrying singletons developed gestational diabetes during pregnancy making them more likely to have large-for-gestational age babies (Table 6.6).

6.3.4 Relationship between birth outcomes and demographic variables *6.3.4.1 Birth outcomes by ethnicity*

Eligible mothers are 37% Pakistani and 33% White British (Table 6.6).

The gestational age distribution of live singleton births varies by ethnic group (Table 6.2). The women in the Black group, though not numerous (N=204), had the highest rates of preterm births (9.8%), followed by the women in the Indian (7.7%), the White British (5.6%) and then the Pakistani group (4.9%) with one of the lowest preterm delivery rates ($p=0.035$ by chi^2 test). That Black babies are more likely born preterm is not a surprising finding, unlike the rate of preterm delivery among Pakistani babies (4.9%) which is below the BiB average (5.6%). Mean gestational age was highest in the White Other (39.5 weeks), Mixed (White and Black) (39.4 weeks) and White British groups (39.3 weeks) ($p<0.001$ by one-way ANOVA).

As expected, mean birth weights also varied significantly between the main ethnic groups (Figure 6.7, Table 6.3). The Bangladeshi, Indian and Pakistani distributions have the lowest birth weights, the White groups' distributions (including White British and White Other) have the heaviest birth weights, and the African and Mixed group distributions occupy an intermediate position. Pakistani babies are 223.2g (95% CI: 199.9g, 246.4) lighter than the White British group on average ($p<0.001$

by t-test); and Indian and Bangladeshi babies are lighter than the Pakistani babies (p<0. 001 by t-test) (mean difference of 82.9g (95% CI: 39.3, 126.5)).

As reported by Moser et al (2008), the rates of LBW babies among the Indian (11.3%), Pakistani (9.0%) and Black (7.8%) women are more than twice the rate among the White women (White British: 5.4%, White Other: 4.5%) (p<0.001 by chi^2 test) (Table 6.3). The rates of term LBW are also very different between Pakistani and White British (5.8% and 2.0%, respectively) while being more similar between Indian and Pakistani babies (5.3% and 5.8%) and between White British and Black babies (2.0% and 1.6%) (p<0.001 by chi^2 test) (Table 6.4).

Differences in birth weight between ethnic groups are partly accounted for by differences in gestational age, particularly for the Indian and Black babies. However, as seen in Figure 6.8, differences in birth weight between ethnic groups remain even when considering only those live singleton births delivered at 40 weeks. For example, where at 40 weeks mean birth weight was 3272.7g (95% CI: 3249.3, 3296.1) in the Pakistani group, it was 3484.2g (3458.8-3509.7) in the White British group (mean difference of 211.5g (95% CI: 177.0, 246.0) (p<0.001 by t-test)).

The South Asian groups also experienced greater rates of SGA: 20.6%, 16.5%, and 15.7% Bangladeshi, Indian and Pakistani babies are SGA, respectively, compared to 7.6% White British, 9.4% White other groups, and only 5.9% Black babies (p<0.001 by chi^2 test) (Table 6.5).

6.3.4.2 Birth outcomes by level of deprivation

These data confirms the Bradford Infant Commission's findings (BDIMC 2006) that the majority of the BiB cohort lives in the lowest quintiles of the Index of Multiple Deprivation 2010 (IMD 2010) (Table 6.7). With this uneven distribution in mind, term LBW rate among BiB babies born to the most deprived areas (quintile 1 of the IMD 2010) was 4.3% compared to 1.9% in the least deprived neighbourhoods (quintile 5) (p=0.145 by chi^2 test) and 1.8% in the second least deprived neighbourhoods (quintile 4) (p=0.045 by chi^2 test) (Table 6.7, Figure 6.9 part a). The same differences hold for LBW: 8.2% in quintile 1, compared to 4.3% in quintile 5 (p=0.071 by chi^2 test) and 3.9% in quintile 4 (p=0.009 by chi^2 test) (Table 6.7, Figure 6.9 part a).

6.3.4.3 Missing ethnicity and IMD 2010 data

The rates of LBW, term LBW, and SGA (8.7%, 4.5% and 13.3%, respectively) are highest among babies with missing deprivation information compared to all other deprivation groupings, and closest to quintile 1 (Figure 6.9 part a). Similarly, rates of LBW, term LBW, and SGA (8.7%, 4.6% and 13.3%, respectively) are the second highest among babies with missing ethnicity information after the

Pakistani group which they also most closely resemble (at least LBW) (Figure 6.9, part b). 17% of the sample has birth weight data but misses both ethnicity and IMD information (2030 of 11874).

6.3.5 Other demographics by ethnicity

In addition to intrinsic differences in expected growths between ethnic groups, ethnicity is also correlated with specific demographic, behavioural, and obstetric outcomes (Table 6.6). Pakistani origin mothers were slightly older at delivery than White British mothers (mean age 28.0 years and 27.0 years, respectively, mean difference 1.1 years, 95% CI 0.8, 1.3, $p<0.001$ by t-test) and were lighter at questionnaire completion (mean difference 6.5kg, 95% CI 5.8, 7.2, $p<0.001$ by t-test), shorter (mean difference 4.4cm, 95% CI 4.1, 4.6, $p<0.001$ by t-test) and had a lower BMI at questionnaire completion (mean difference 0.9, 95% CI 0.7, 1.2, $p<0.001$ by t-test). The proportion of Pakistani mother with gestational diabetes (10.9%) was double that of White British mothers (4.9%) ($p<0.001$ by chi^2 test), and the proportion with pregnancy-induced hypertension was also higher in Pakistani (6.4%) than in the White British group (4.6%) ($p<0.001$ by chi^2 test). Pre-eclampsia rate however was similar between ethnic groups ($p=0.922$ by chi^2 test). Differences in parity were also found (19% of White British mothers reported a parity of 2 or more compared with 40% of Pakistani origin mothers ($p<0.001$ by chi^2 test)), as well as marked differences in employment and marital status between the two ethnic groups: 32% of White British mothers were married and 63% currently employed and working which contrasts sharply with the Pakistani origin mothers who were almost all married (98%) but unlikely to be currently working (23%) (p's<0.001 by chi^2 test). Lastly, 29% of White British women were current smokers and 41% never smoked, while only 3% of Pakistani women were current smokers and 92% never smoked ($p<0.001$ by chi^2 test both for the comparison of current smokers rates between ethnic groups, and never smokers rates between ethnic groups); 34.6% of White British women reported drinking alcohol during pregnancy or in the 3 months preceding pregnancy compared to 0.3% of the Pakistani women (i.e. 11 women) ($p<0.001$ by chi^2 test) (Table 6.6).

6.4 Discussion

6.4.1 Preterm delivery (PTD), term LBW and SGA rates in BiB: comparisons with BDIMC

National birth weight data tend to be reported collectively for all gestations (premature and term). Variability in rates may also be due to regional differences in the criteria for registration of stillbirths and live births, differences in the extent of medical interventions (such as induction and caesarean section), and differences in the estimation of gestational age (Wilcox 2010) (see later section 6.4.3.2.1). Multiple births (twins, triplets etc.) are sometimes presented in the same statistic as well, without clear mention. I suspect that this may be the case for the results of the Bradford Infant

Mortality Commission which partly prompted the BiB birth cohort study. Indeed, 7.1% of live births were preterm, 9.7% babies LBW and 1.5% very LBW on average in the Bradford District between 1996 and 2003 (BDIMC 2006). These rates are quite a bit higher than those observes in BiB (Table 6.1). BiB rates are more in line with national trends reported by the Office of National Statistics and by Moser et al's large study (2008). However, within the average weight ranges (2500-4500g), BiB babies remain smaller on average than babies born in Yorkshire, or England and Wales as a whole (ONS 2012) (Figure 6.10). This is likely due to the high proportion of South Asian babies in this cohort, with lower average birth weights (Table 6.3, Table 6.4).

If the inclusion of multiple births does not explain the differences in rates noted in the BiB cohort compared to those reported by the 2006 Bradford District Infant Mortality Commission report, either a) Bradford birth outcome indicators greatly improved between the time that the 2006 Bradford District Infant Mortality Commission report came out and the time BiB recruitment began and was under way (2007-2010), or b) the BiB cohort is not fully representative of the Bradford community as a whole, or both (see later section 6.4.3.3).

When applied to the reference population, standardised birth weights (SDS) will by definition be distributed according to a normal distribution, centred at 0 with standard deviation of 1. A negative mean for the population distribution suggests that this population is more prone to being small-for-gestational age and sex than the reference population, with associated consequences. Figure 6.11 plots SGA rate by gestational age and sex. 51% of SGA babies are males. Likely in part due to relationship between prematurity and fetal growth restriction, a greater proportions of SGA babies are born at earlier gestational ages.

Rates of pregnancy-induced hypertension vary substantially in high-income countries. This is likely due to underascertainment and/or misclassification. Accurate pre-eclampsia statistics are also difficult to obtain because the condition ranges from extremely mild to severe, and mild cases—which may not have any effect on pregnancy—are not always included in official figures. In addition, as the majority of cases of pregnancy-induced hypertension and pre-eclampsia occur at term, the recent trend in Europe towards increasing rates of early elective delivery may reduce their frequency (Koopmans et al. 2009; MacDorman et al. 2010; Roberts et al. 2011). Despite these caveats, pregnancy-induced hypertension and pre-eclampsia rates observed in BiB are within the ranges (4% to 10%, and 2% to 5%, respectively) observed in other studies (Hernandez-Diaz et al. 2009; Klemmensen et al. 2007; Roberts et al. 2011; Ros et al. 1998). The same is true for the rate of gestational diabetes (8.1%) with 2-12 % of women developing gestational diabetes in the UK, more commonly if they are from ethnic minority groups (Department of Health 2012).

6.4.2 Relationship between birth outcomes and demographic variables

6.4.2.1 Birth outcomes by ethnicity

As per Moser's study (2008), I find many variations in rates of poor birth outcomes by ethnic groups. Moser et al (2008) reported that the Caribbean group had the highest percentage (9.7%) of live singletons born preterm followed by the African (7.0%), Indian (6.9%), Pakistani (6.8%) and then White British (6.1%) groups, and that mean gestational age was highest in the White groups (39.3 weeks for White British and 39.0 weeks for Pakistani). Compared to this study, the low rate of preterm delivery among Pakistani mothers in BiB (4.9%) is unexpected. Kelly et al (2009), albeit on a much smaller sample of Pakistani births from the Millenium Cohort Study (N=687 out of a total sample of 16,157 babies broken into six ethnic groups), also reported their lowest rate of preterm delivery among the Pakistani group (5.7%) though this rate is still greater than ours.

Moser et al (2008) reported the same relationship of birth weight by ethnicity as observed in this study: Bangladeshi, Indian and Pakistani babies are lightest, White groups (including White British and White Other) are heaviest, and the African and Mixed groups are in between. Pakistani babies were 223.2g lighter than the White British group on average. At 40 weeks gestation, the 212g difference observed between the White British and Pakistani groups remains significant and similar to the difference reported by Moser et al (2008) (approx. 210g per Figure 4 in that reference).

It is noteworthy that ethnicity information in BiB is self-reported. This could add noise to the variable as women's cultural vs. racial ethnic belongings may be different, and it is unclear which one is reported. Critically, ethnicity information is also collected for the mother, not for the baby. If a couple with different ethnic backgrounds parent a child, that baby's ethnicity may not be accurately reflected in the maternal ethnicity used in these analyses. This is likely to introduce additional random error in the data and effect estimates, as controlling for ethnicity may not be enough to capture the complexity of BiB babies' heritage.

6.4.2.1.1 Demographic breakdown by ethnicity: smoking during pregnancy

Ethnicity plays an important role either in itself as a genetic marker, or as a proxy marker for a host of singular maternal characteristics and behaviours (Table A7 - 2). If ethnicity itself as a genetic marker helps distinguish between pathologically and constitutively small babies at birth, differences may be underestimated if behaviours associated with ethnicity go unacknowledged.

For example, smoking during pregnancy is known to predict lower birth weight (see Chapter 7). Consistent with previous reports (Hawkins et al. 2008; Health Survey for England 2004), smoking and alcohol consumption were uncommon among Pakistani mothers. Given that smoking habits vary

between the two predominant ethnic groups (29% current smokers and 41% never smokers among the White British women vs. 3% current and 92% never smokers among the Pakistani group), babies' birth weights are naturally differentially affected in the two groups. As per above, babies born to Pakistani mothers weigh 223.2g (95% CI: 199.9, 246.4) less on average than the babies born to White British mothers. This difference increases to 284.2g (95% CI: 254.2, 314.2) for babies born to non-smoking Pakistani and non-smoking White British women. This increase in birth weight is due to an increase in average birth weights among the non-smoking White British mothers (3425.1g, 95% CI 3399.7, 3450.4, N=1,629) compared to the average for White British mothers (3359.9g, 95% CI 3342.5, 3377.2, N=3,953), 29% of whom smoked during pregnancy (rather than to a decrease in birth weights among non-smoking Pakistani mothers (3140.9g, 95% CI 3124.7, 3157.0, N=3,972) compared to average Pakistani mothers (3136.7g, 95% CI 3121.2, 3152.2, N=4,341)). At 40 weeks gestation, the difference in birth weights between White British and Pakistani babies born to mothers who never smoked is 271.4g (226.7, 316.2) ($p<0.001$ by t-test) (down from 284.2g). (There were too few Pakistani women who smoked (N=138) to compare the birth weights of babies born to smoking mothers.)

In sum, because White British mothers smoke substantially more than Pakistani mothers, their birth weights are more affected by smoking than Pakistani babies'. The same may be true for rates of alcohol and caffeine consumption and employment which are lower among the Pakistani women compared to the White British. Conversely, the rate of gestational diabetes, which is well established to increase birth weight (Dornhorst et al. 1992; Oldfield et al. 2007), is higher among the Pakistani women (Table 6.6). Further investigation of the effects of these factors on birth weight is outside the scope of this thesis.

The observation that prevalence of smoking among Pakistani and White British mothers is different (and that this difference carried over to birth weight) is not new. A similar result was found in a study in Nottingham where smoking in pregnancy (as recorded at the first visit) was prevalent in 28% of European mothers compared to only 2% of Asian mother, and the inter-ethnic differences in birth weight were most apparent for non-smokers, amounting to over 250g at 40 weeks' gestation (Wilcox et al. 1993).

6.4.2.2 Birth outcomes by level of deprivation

The LBW, term LBW and SGA rates by deprivation status reported for BiB (see Table 6.7) are consistent with (though again more moderate than) the Bradford District Infant Mortality Commission's report that 12.5% of the babies born to mothers living in the most deprived 20% of areas (quintile 1 of the IMD 2010) were LBW compared to 6.2% in the least deprived neighbourhoods (quintile 5) (BDIMC 2006)

I suspect that deprivation status per se is likely not the proximal cause of LBW. As such, a or several risk factors for LBW on a population basis must therefore be more common among lower socio-economic status women. A combination of lifestyle behaviours (including poor maternal nutrition during pregnancy), or maternal stress—whether chronic, psychological, social or physical—due to heightened disadvantage, perceived or real (Sapolsky 2005), could be the underlying explanation (see Chapter 7, section 7.4.8.3.2).

6.4.3 Pros and Cons of term LBW and SGA

6.3.3.1 Identifying pathologically small babies

Identifying pathologically small babies is a challenging undertaking. Several birth weight-based measures attempt to do so. I decided to study three of them, each with different advantages and limitations, in order to get the most well-rounded picture possible and hopefully a glimpse at the truth.

6.4.3.1.1 Term LBW

LBW is a recognised measure that is often used in the clinical setting, it has been linked with effects in adult life such as asthma, reduction in cognitive function, metabolic syndrome and heart disease (see Introduction (Chapter 1)). However, as LBW does not take age at birth into consideration, I elected to study term LBW, which has been linked to a number of environmental risk factors. Excluding preterm births means that growth restricted babies who are also born preterm may now be overlooked. Without a better understanding of whether fetal growth restriction causes prematurity, or whether fetal growth restriction and prematurity share the same aetiology, it is difficult to definitively determine which compromise (not accounting for gestational age at birth, or losing all preterm births) is best. Term LBW was chosen as the more restrictive but safer approximation of growth restriction.

6.4.3.1.2 SGA in a multi-ethnic cohort

The SGA measure is the most respected categorical measure of fetal growth based on birth weight data, as it identifies the smallest babies within each gestational age group. However, as a statistical construct, it is not clear that the smallest babies by birth weight are the only growth restricted babies in a cohort. In addition, the choice of appropriate reference by which to define SGA is critical and can vary: reference curves can be internal or external and must reflect the population at hand. On occasion they adjust for a host of possible factors such as height and weight of the parents, ethnicity and even smoking. They can also vary in age, and/or not be regionally representative.

The centiles used in the UK90 referent to generate the SGA analysed in this thesis were created based on birth data from five studies conducted between 1983 and 1993 and mainly based in East Anglia, UK (Cole et al. 2011). Nothing was known on the ethnic background of the women and babies included in the birth weight surveys which make up the referent, as the relevant information was not available. As such, this growth reference sample can be said to be representative of British ethnically Caucasian children (Cole et al. 1998), but was not necessarily developed for use on an a multi-ethnic population such as the BiB cohort. In light of this it does not come as a surprise that the SDS distribution in this cohort is centred below zero (section 6.3.2)—just as it was reported that the growth of infants in Bradford differs from the WHO standards (Wright et al. 2012).

Despite the above, producing ethnic group-specific references is not a satisfactory approach either, as the large and representative samples required are simply not available, the definition of ethnicity itself brings with it its own challenges and pitfalls, and ethnic group is only one of several non-pathological variables affecting weight—the others include parity, maternal height, weight at first visit, sex of the baby and social deprivation (Altman and Coles 1980; Gardosi et al. 1992). Adjustment of such ethnic factors may therefore not result in a true reflection of the growth potential of a fetus (Gardosi 2009). Cole et al (1998) suggest that a better answer would be to use a series of small-scale surveys to summarize the growth status of specific ethnic minority children in terms of their mean SDS on the British reference, which could be used to recalibrate the reference for use with such groups.

In addition to the difficulty of finding an appropriate reference for the population at hand, references are populated by data on babies who were born. In other words, the distribution of births at a given gestational age is not a random sample of all pregnancies that have attained that gestational age but restricted to those that *ended* at that point. The optimal approach would be to have longitudinal information on a large population of unselected pregnancies measured *in utero* by ultrasound to estimate fetal weight, allowing fetuses or births at a given point in gestation to be compared with the appropriate referent group of all fetuses of comparable gestational age (Hutcheon and Platt 2008; Savitz et al. 2002). This is what Gardosi's customised SGA measure does, which is arguably one of its greatest strengths (see Introduction (Chapter 1)).

The SGA measure can be derived for a greater number of babies than the term LBW, as it does not exclude 663 preterm births. However, for women for whom both SGA and tLBW were derived (N= 11,200, see Figure 6.5b), term LBW and SGA have 91% agreement by Kappa statistic for agreement between categorical variables.

6.4.3.1.3 Continuous birth weight

As well as term LBW and SGA, I have chosen to look at continuous birth weight, which makes no arbitrary judgement as to the appropriate cut-off. If in fact some condition or exposure shifts the entire birth weight distribution, the identification of that shift will be enhanced when examining mean birth weight as compared with the proportion below some cut point, such as 2500g. The rationale is that, although the growth restricted infants cannot be identified individually, shifting all birth weights downward will increase the number and proportion of births that are abnormal (Savitz et al. 2002).

If these inter-ethnic differences reported above remain unadjusted for, using cut-off limits to define high-risk populations can lead to significant error. For example, according to Gardosi et al (1994), ignoring a 200g+ "normal" difference and using the same tenth centile cut-off for all groups would mean that 75% of Asian SGA babies would in fact not be below this limit if their own norm was used.

For a summary of pros and cons of each measure investigated, see Table 6.8. All three of these birth outcome measures are birth weight-based. Birth weight alone cannot explain what contributes to differences in size. It reflects a number of components including bone, muscle, fat and fluids (Shields et al. 2006). Thus, a low birth weight does not indicate whether for example, an infant is universally small, has a large head and a small body, or is small but has a high percent body fat (West 2011).

6.4.3.2 Limitations of dating methods

Compounding the above mentioned issues, consideration must also be made for the challenges associated with accurate gestational age estimation. Both term LBW and SGA measurements rely on accurate estimation of the duration of gestation, errors in which will result in shifts in the percentile of weight for age (e.g. an infant of a given weight could represent a normal 36-week birth, a large 34-week birth, or a small 38-week birth), while continuous birth weight models adjust for gestational age (Savitz et al. 2002).

6.4.3.2.1 Dating based on last menstrual period (LMP)

There are two main issues with using LMP to estimate gestational age. The first is recall error, as most women do not keep menstrual diaries, particularly if not planning the pregnancy, or if the pregnancy resulted from contraceptive failure. Most women can recall their LMP within 1-2 days (Wegienka and Baird 2005), but for a few, the error is much greater. Late entry into prenatal care or becoming pregnant soon after a previous birth (such that there may not be a LMP to recall) add to the problem. Undetected miscarriages or delayed ovulation before the current pregnancy may cause erroneously long gestation, while early pregnancy spotting may be mistaken for LMP and cause erroneously short gestation. In addition, pregnant women who cannot recall any date for their LMP are likely to have less education and in other ways not be representative of the general population adding non-systematic bias to the equation (Buekens et al. 1984).

The second is the assumption that conception takes place immediately after ovulation, which in turn occurs at a variable time after LMP. The LMP approach assumes that all women have regular and/or 28-day cycles, while in fact, only 10% of women actually ovulate on day 14 (Baird et al. 1995; Lenton et al. 1984; Wilcox et al. 2000). A US study found that onset of LMP to ovulation was 17 days on average, but extended up to a maximum of 8 weeks (Baird et al. 1991).

Because the distribution of menstrual dating error is positively skewed, any birth weight at term can appear at later gestations than it actually should be, leading to an artificial flattening of the growth curve and apparent increase in post-term births (Gardosi et al. 1997). Unadjusted, these data do show this flattening tendency (Figure 6.12), while in reality growth *in utero* in a normal pregnancy continues without diminished velocity until birth (Williams et al. 1982)

6.4.3.2.2 Ultrasound dating

Ultrasound-based adjustment of LMP improves estimates of gestational age, and substantially reduces the percentage of babies born post-term (presumably by replacing many of the LMPs that had included long follicular phases[21]) (Wilcox 2010). But measurement error remains inevitable. A difference of just 2mm in head diameter of femur length can affect the estimated fetal growth by half a week (Gjessing et al. 2007).

Furthermore, and critically for this multi-ethnic population, ultrasound dating assumes that individual variation in growth is minimal, i.e. that the size of a fetus in the first half of pregnancy is a function of age alone, not of rate of growth, which is not strictly true. As such, any natural variations in the rate of fetal growth are automatically translated into variations in fetal age. For example, given two fetuses conceived on the same date, the one that grows more rapidly will be assigned an older age by ultrasound than the one growing more slowly. Natural variation in fetal growth in the first half of pregnancy adds error to the ultrasound estimate of fetal age (Henriksen et al. 1995). This may come to bear if an exposure damages the growth of fetuses but has no effect on length of pregnancy: such an exposure could, on the basis of ultrasound-assigned age, appear to increase the risk of preterm delivery (Wilcox 2010).

Based on the above, BiB's gestational age estimates—on whose origin I have little information other than that they summarised individual doctors' best estimates based on available LMP and ultrasound information (see Chapter 2)—are likely to contain some errors which could affect a) classification of

[21] i.e. phase of the menstrual cycle during which follicles in the ovary mature (days 0 to 14 of a 28 day cycle)

cases as term LBW, b) SGA derivations which are gestational age-specific and c) models on birth weight and term LBW which adjust for gestational age at birth.

6.4.3.3 Outcome misclassification

As summarised in Smith (2011), outcome misclassification in general can cause bias in health-risk estimates (in addition to loss of study power) (Armstrong 1998). If misclassification of fetal growth restriction is non-differential with regard to exposure status/level, effect estimates are simply biased towards the null (Armstrong 1998). Where a continuous outcome measure has been used, e.g. birth weight, random measurement error would cause loss of study power, but would not bias risk estimates (Armstrong 1998). However, if attrition from the cohort occurs preferentially in one subgroup (e.g. Pakistani) over the others, then the impact on estimates may be differential.

6.4.4 Cohort representativeness

The cohort profile published in 2012 (Wright et al. 2012) reports that >64% of all pregnant women in Bradford who registered for prenatal care at Bradford Royal Infirmary between 2007-2010 were in fact enrolled in BiB, suggesting good representativeness. It concludes that if anything, BiB may have recruited a lower proportion of younger mothers (age: 20-24 years) compared with Bradford mothers not in the cohort, and a higher proportion of South Asian mothers and nulliparous mothers (Wright et al. 2012).

In terms of internal validity, a particularly low rate of preterm delivery among BiB's Pakistani women was observed compared to previous studies. In addition, Pakistani women were on average significantly shorter White British women in BiB (see section 6.3.5), but their mean BMI (27.9 (27.8-28.1) (N=4,158)) was also lower than the White British women's (28.9 (28.7- 29.0) (N=3,832)) ($p<0.001$ by t-test). While the height pattern by ethnic group is similar to that reported in other studies (Kelly et al. 2009), Pakistani women enrolled in BiB have relatively low BMIs in comparison to national data (Health Survey for England). Weighing less for given heights, they may therefore represent a healthier subgroup of Pakistani women. People who agree to join research studies do tend to be healthier than the general population (also known as the "healthy volunteer effect" (Howe et al. 1988)).

Secondly, as recruitment to BiB occurred relatively late in pregnancy (26-28 weeks), women who give birth before 26-28 weeks' gestation—which, granted, will be a small proportion as these represent very preterm births—could not be captured (Wright et al. 2012).

Thirdly, attrition of enrolled women with difficult pregnancies or with poor birth outcomes is suspected. For instance, five babies whose gender could not be assigned had to be excluded from the dataset for identifiability reasons. Similarly, between the February 2012 and February 2013 data extracts received (see Chapter 2), there were a number of withdrawals from the study which precludes use of any of these participants' data (Table 6.9). When the number of term LBW babies is in the hundreds (450 to be specific, see Figure 6.4), losing 256 babies with possibly high term LBW prevalence could have an impact. It is not hard to imagine that the reason for many of these withdrawals could be related to a pregnancy complication or poor birth outcome which would differentially bias the sample. However the reason for withdrawal cannot be ascertained without the very supporting data which are not available to us.

Due to the potential non-random missingness pattern of ethnicity and deprivation status among those with birth weight information in this cohort, the differences in rates presented in this chapter stratified by ethnicity and deprivation status may underestimate the true differences in LBW, term LBW, and SGA rates between ethnicities and quintiles of deprivation. However, attrition, if present, remains low (< 10%).

Lastly, as described in section 6.3.4.3, a non-negligible portion of women who were recruited to BiB and eligible for this study did not report their ethnicity (17%), and could not be classified by level of deprivation as they lacked residence information (17%)—with substantial overlap between the two groups (97%). These women with missing information also had high rates of LBW, term LBW and SGA (see section 6.3.4.3). Their outcomes rates were most similar to those of the Pakistani and low IMD (quintile 1) women such that I suspect that demographic information on some of the most deprived women recruited to BiB may be missing.

The BiB cohort is unique in that it is based in a city which is unusually poor and deprived compared to the rest of the UK, and is made up of two large ethnic contingencies of women—namely White British and Pakistani women—to study separately and to contrast to one another. BiB constitutes a huge opportunity to study a population at high risk of poor birth outcomes in the UK.

6.5 Tables

Table 6.1: Comparison of rates with other studies a) overall and b) by ethnicity (the statistical construct SGA is not included as the prevalence is typically determined by definition) PTD=preterm delivery, LBW=low birth weight, vLBW=very low birth weight

a) overall

%	BiB	Bradford District Infant Mortality Commission (1996-2003)	England & Wales (2011) (ONS)	Moser et al (2008)	US National Vital Statistics for 2010 (Martin et al. 2012)
PTD	5.6	7.1		6.2	
LBW	7.5	9.7	7.0	6.1	8.1
vLBW	0.7	1.5			

b) By ethnicity

%	BiB		Moser et al (2008)		Kelly et al (2008)	
	WB	P	WB	P	WB	P
LBW	5.4	9.0	5.6	9.8	5.2	13.0
Term LBW	2.0	5.8			1.4*	6.0*

*these values were derived from regression models adjusting for gender, gestational age, parity, maternal age, maternal height, pre-pregnancy weight, any complications during pregnancy

Table 6.2: Gestational age at birth (and rate of preterm delivery) by ethnicity (11,928 eligible women with singletons, but only 11,875 had live births with recorded gestational age) (PTD=preterm delivery)

	Asian, Asian British			White		Black	Mixed		Other♯	Missing	**Total**
	Bangladeshi	Indian	Pakistani	White British	White Other^		White and Black*	White and South Asian**			
Gestational age, weeks (%)											
<28	NA	NA	0.1	0.2	NA	0.5	NA	NA	0.8	0.1	**0.2**
28-31	NA	0.3	0.7	0.7	0.4	0.5	1.0	NA	0.4	0.9	**0.7**
32-36	4.4	7.5	4.0	4.7	4.1	8.8	4.0	5.5	3.0	5.8	**4.7**
37-39	29.7	25.3	24.2	17.9	17.9	17.7	18.0	18.2	20.6	21.9	**21.4**
39-41	64.2	66.5	69.9	74.5	75.4	71.1	74.0	76.4	73.0	70.4	**71.6**
≥42	1.8	0.5	1.1	2.0	2.2	1.5	3.0	NA	2.3	0.9	**1.4**
Total	100	100	100	100	100	100	100	100	100	100	**100**
< 32 weeks (very preterm) (%)	0.0	0.3	0.9	0.9	0.4	1.0	1.0	0.0	1.1	1.1	**0.9**
< 37 weeks (preterm) (%)	4.4	7.7	4.9	5.6	4.5	9.8	5.0	5.5	4.1	6.8	**5.6**
Mean gestational age, weeks	39.1	38.9	39.1	39.3	39.5	38.9	39.4	39.2	39.2	39.0	**39.2**
Number of births (n)	229	388	4341	3954	268	204	100	55	267	2069	**11875**
Row percentage	1.9	3.3	36.6	33.3	2.3	1.7	0.8	0.5	2.2	17.4	**100.0**

^includes White Irish, and all other White self-reported groupings

*Mixed White and Black includes White and a) Black Caribbean, or b) Black African

** Mixed White and South Asian includes White and a) Indian, b) Pakistani, c) Bangladeshi, d) Indian Caribbean, e) African Indian

♯Other category includes Chinese, Japanese, Filipino and Vietnamese

NA means there were no data to calculate that gestational age %

Table 6.3: Birth weight (and LBW rate) by ethnicity (11,928 eligible women with singletons, but only 11,874 had live births with recorded birth weight, see Figure 6.3) (LBW=low birth weight)

	Asian, Asian British			White		Black	Mixed		Other	Missing	**Total**
	Bangladeshi	Indian	Pakistani	White British	White Other		White and Black	White and South Asian			
Birth weight, grams (%)											
< 1,500	NA	1.0	0.7	0.7	NA	0.5	1.0	NA	1.5	0.8	**0.7**
1,500-	0.9	2.3	1.7	1.2	1.5	2.9	2.0	NA	0.4	1.7	**1.5**
2,000-	6.1	8.0	6.7	3.5	3.0	4.4	3.0	5.5	3.0	6.1	**5.3**
2,500-	34.5	33.0	27.8	16.6	17.2	20.1	8.0	25.5	17.6	25.8	**23.2**
3,000-	42.8	38.7	39.8	38.0	34.3	35.8	46.0	41.8	40.8	37.8	**38.8**
3,500-	14.4	14.2	19.4	28.3	30.6	26.0	29.0	25.5	29.2	20.6	**23.0**
4,000-	1.3	2.6	3.5	10.2	10.5	9.8	10.0	1.8	5.6	6.6	**6.5**
4500-	NA	0.3	0.6	1.3	2.6	0.5	1.0	NA	1.5	0.6	**0.9**
≥ 5,000	NA	NA	0.0	0.1	0.4	NA	NA	NA	0.4	0.1	**0.1**
Total	100	100	100	100	100	100	100	100	100	100	**100**
< 1,500g (%)	0.0	1.0	0.7	0.7	0.0	0.5	1.0	0.0	1.5	0.8	**0.7**
< 2,500g (LBW) (%)	7.0	11.3	9.0	5.4	4.5	7.8	6.0	5.5	4.9	8.7	**7.5**
Mean birth weight, g	3068.2	3045.4	3136.7	3359.9	3412.8	3260.2	3352.2	3201.6	3308.3	3186.6	**3231.1**
(95% CI)	(3012.9, 3123.4)	(2994.2, 3096.6)	(3121.2, 3152.2)	(3342.5, 3377.2)	(3346.3, 3479.2)	(3181.4, 3339.0)	(3240.5, 3463.9)	(3074.0, 3329.3)	(3239.3, 3377.3)	(3162.7, 3210.4)	**(3219.8, 3239.6)**
Number of births (n)	229	388	4341	3953	268	204	100	55	267	2069	**11874**
Row percentage	1.9	3.3	36.6	33.3	2.3	1.7	0.8	0.5	2.2	17.4	**100.0**

Table 6.4: Birth weight at term (and term LBW rate) by ethnicity (11,928 eligible women with singletons, but only 11,211 had live births at term with recorded birth weight and gestational age, see Figure 6.4) (term LBW=low birth weight at term)

	Asian, Asian British			White		Black	Mixed		Other	Missing	**Total**
Birth weight at term, grams (%)	Bangladeshi	Indian	Pakistani	White British	White Other		White and Black	White and South Asian			
< 1,500 (at term)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	**NA**
1,500- (at term)	0.5	0.3	0.5	0.2	0.4	NA	NA	NA	NA	0.4	**0.3**
2,000- (at term)	4.1	5.0	5.3	1.8	2.0	1.6	3.2	3.9	2.0	4.2	**3.7**
2,500- (at term)	35.6	34.4	28.1	16.1	16.0	20.1	7.4	25.0	17.2	25.7	**23.2**
3,000- (at term)	43.8	41.9	41.6	39.7	35.9	38.6	47.4	42.3	42.6	40.0	**40.6**
3,500- (at term)	14.6	15.4	20.3	30.0	31.6	28.3	30.5	26.9	30.5	22.0	**24.3**
4,000- (at term)	1.4	2.8	3.7	10.8	10.9	10.9	10.5	1.9	5.9	7.0	**6.9**
4500- (at term)	NA	0.3	0.6	1.4	2.7	0.5	1.1	NA	1.6	0.7	**0.9**
≥ 5,000 (at term)	NA	NA	NA	0.1	0.4	NA	NA	NA	0.4	0.1	**0.1**
Total	100	100	100	100	100	100	100	100	100	100	**100**
< 2,500g at term (term LBW) (%)	4.6	5.3	5.8	2.0	2.3	1.6	3.2	3.9	2.0	4.6	**4.0**
Mean birth weight at term, g (95% CI)	3092.0 (3038.8, 3145.2)	3127.4 (3083.0, 3171.8)	3188.4 (3174.4, 3202.4)	3424.8 (3409.4, 3440.2)	3455.9 (3392.9, 3518.9)	3363.8 (3298.6, 3428.9)	3420.2 (3328.2, 3512.2)	3227.5 (3098.7, 3356.3)	3369.3 (3311.2, 3427.3)	3255.7 (3234.4, 3277.1)	**3290.1 (3281.2, 3299.0)**
Number of term births (n)	219	358	4130	3733	256	184	95	52	256	1928	**11211**
Row percentage	2.0	3.2	36.8	33.3	2.3	1.6	0.8	0.5	2.3	17.2	**100.0**

Table 6.5: Standard deviation scores (SDS) (and rate of SGA) by ethnicity (11,928 eligible women with singletons, but only11,863 had live births with recorded birth weight, sex of child, and gestational age, and born within 23-42 window of gestational weeks, see Figure 6.5) (SGA=small-for-gestational age)

	Asian, Asian British			White		Black	Mixed		Other	Missing	**Total**
	Bangladeshi	Indian	Pakistani	White British	White Other		White and Black	White and South Asian			
Standard deviation score (%)											
≤ -1.282 (10th percentile)	20.6	16.5	15.7	7.6	9.4	5.9	6.0	12.7	8.3	13.3	**12.1**
>-1.282 and ≤ -0.674 (10-25th p)	22.4	27.1	21.7	14.6	13.9	15.7	15.0	21.8	15.8	18.1	**18.4**
>-0.674 and ≤ 0 (25-50th p)	26.8	28.9	28.5	26.5	24.3	30.9	32.0	23.6	29.0	29.2	**27.9**
>0 and ≤ 0.674 (50-75th p)	18.4	16.0	20.6	26.1	25.8	26.5	26.0	29.1	27.4	20.8	**22.7**
>0.674 and ≤ 1.282 (75-90th p)	9.7	8.0	8.7	15.5	14.6	11.3	13.0	5.5	12.0	11.4	**11.7**
> 1.282 (90th percentile)	2.2	3.6	4.8	9.9	12.0	9.8	8.0	7.3	7.5	7.2	**7.1**
Total	100	100	100	100	100	100	100	100	100	100	**100**
SGA (i.e. SDS ≤ -1.282 (%))	20.6	16.5	15.7	7.6	9.4	5.9	6.0	12.7	8.3	13.3	**12.1**
Mean SDS	-0.466	-0.464	-0.351	0.055	0.070	0.019	-0.013	-0.267	-0.036	-0.213	**-0.172**
Number of births (n)	228	388	4339	3947	267	204	100	55	266	2069	**11863**
Row percentage	1.9	3.3	36.6	33.3	2.3	1.7	0.8	0.5	2.2	17.4	**100.0**

Table 6.6: Demographic, behavioural and obstetric variables by ethnicity: frequency (above) and percentage (below) within that ethnic group (11,928 eligible women with singletons, but only 11,875 had live births)

N total =11,875 N (% of ethnic groups represented)	White British 3954 (33.3%)	Pakistani 4341 (36.6%)	Other 1511 (12.7%)	Missing 2069 (17.4%)	Total 11,875	chi^2 test p-value (*one-way anova for continuous variables)
mother and baby demographic variables						
mother's age at time of birth						
<20	11.3	2.2	5.0	6.6	**6.3**	<0.001
20-24	27.1	25.7	20.4	27.2	**25.8**	
25-29	28.3	35.4	35.1	33.0	**32.6**	
30-34	20.1	24.2	25.9	21.3	**22.5**	
35-40	11.1	10.3	11.5	9.8	**10.6**	
≥40	2.2	2.2	2.1	2.2	**2.2**	
registerable parity						
no previous birth registration	48.4	31.7	47.6	33.0	**39.48**	<0.001
1 previous registerable birth	29.2	24.3	27.3	27.6	**26.88**	
2 or more previous registerable births	19.2	39.7	21.1	36.8	**30**	
missing	3.3	4.3	4.1	2.6	**3.64**	
marital status						
married or re-married	31.8	97.5	74.7	1.0	**55.9**	<0.001
single (never married)	64.6	1.0	23.7	0.7	**25.0**	
divorced, separated, or widowed	3.5	1.4	1.5	0.1	**1.9**	
missing	0.1	0.1	0.1	98.3	**17.2**	
cohabitation status						
living with a partner (father of the baby or not)	71.3	93.3	84.9	1.3	**68.9**	<0.001
not living with a partner	28.6	6.6	15.0	0.4	**13.9**	
missing	0.1	0.1	0.1	98.3	**17.2**	
maternal height (cm) (N)	164.1 (3891)	159.7 (4222)	160.8 (1481)	162.6 (38)	**161.7 (9632)**	<0.001*
maternal weight at questionnaire completion (kg) (N)	77.8 (3834)	71.3 (4161)	71.9 (1455)	74.0 (37)	**74.0 (9487)**	<0.001*
maternal weight at booking (kg) (N)	72.0 (3680)	65.2 (4043)	65.8 (1396)	65.8 (1860)	**67.6 (10,979)**	<0.001*
socio-economic variables						
maternal employment status						

N total =11,875 N (% of ethnic groups represented)	White British 3954 (33.3%)	Pakistani 4341 (36.6%)	Other 1511 (12.7%)	Missing 2069 (17.4%)	Total 11,875	chi^2 test p-value (*one-way anova for continuous variables)
employed	63.3	23.0	53.0	0.6	**36.3**	<0.001
not employed/student	36.7	76.9	47.0	1.1	**46.5**	
missing	0.0	0.1	0.0	98.4	**17.2**	
maternal education						
no formal education	19.9	25.8	12.9	0.5	**17.8**	<0.001
school	34.1	31.0	20.9	0.4	**25.4**	
further education	17.2	12.6	13.4	0.1	**12.1**	
higher education	19.2	25.9	40.3	0.4	**21.1**	
other, don't know, unknown foreign	9.6	4.5	12.1	0.2	**6.4**	
missing	0.1	0.2	0.4	98.3	**17.3**	
Index of Multiple Deprivation 2010 by quintiles						
quintile 1 (most deprived)	50.9	79.4	67.3	1.6	**54.8**	<0.001
quintile 2	21.5	14.0	20.0	0.2	**14.9**	
quintile 3	17.9	5.6	9.9	0.1	**9.3**	
quintile 4	6.0	0.5	1.5	0.0	**2.4**	
quintile 5 (least deprived)	3.5	0.2	1.1	0.0	**1.4**	
missing	0.2	0.2	0.3	98.1	**17.3**	
behavioural variables						
smoking during pregnancy						
Never a smoker	41.2	91.5	76.8	1.0	**57.1**	<0.001
Ever a smoker	30.2	5.1	15.3	0.3	**13.9**	
Currently a smoker	28.6	3.2	7.8	0.4	**11.7**	
missing	0.1	0.3	0.1	98.3	**17.3**	
consuming alcohol during (or 3 months before) pregnancy						
no	58.7	99.4	86.2	1.3	**67.1**	<0.001
yes	34.6	0.3	11.3	0.3	**13.1**	
missing	6.7	0.3	2.5	98.5	**19.8**	
caffeine intake during pregnancy						
no	61.2	84.4	77.6	1.1	**61.3**	<0.001
yes (>200 mg/day)	31.3	5.5	11.5	0.2	**13.9**	
missing	7.5	10.0	10.9	98.7	**24.8**	

N total =11,875 N (% of ethnic groups represented)	White British 3954 (33.3%)	Pakistani 4341 (36.6%)	Other 1511 (12.7%)	Missing 2069 (17.4%)	Total 11,875	chi^2 test p-value (*one-way anova for continuous variables)
obstetric outcomes						
gestational diabetes						
no	91.3	85.3	87.3	88.4	**88.1**	<0.001
yes	4.7	10.9	8.9	6.2	**7.8**	
missing	4.0	3.9	3.8	5.4	**4.2**	
pregnancy-induced hypertension						
no	89.5	91.0	90.6	89.8	**90.3**	0.015
yes (mild, moderate, and severe)	6.4	4.6	5.0	5.2	**5.3**	
missing	4.1	4.4	4.4	5.0	**4.4**	
pre-eclampsia						
no	93.2	93.0	93.1	91.9	**92.9**	0.713
yes	2.5	2.5	2.3	3.0	**2.6**	
missing	4.4	4.5	4.6	5.1	**4.6**	
route at birth						
vaginal	77.4	79.1	75.1	74.5	**77.2**	<0.001
caesarean	22.6	20.9	24.9	25.5	**22.8**	
spontaneous onset of labour						
spontaneous	68.3	71.0	69.4	67.1	**69.3**	<0.001
no labour	10.5	8.9	9.8	13.2	**10.3**	
induction (medical and surgical)	21.1	20.0	20.7	19.5	**20.4**	
missing	0.1	0.1	0.1	0.1	**0.1**	
presentation at birth						
cephalic	95.9	96.8	96.8	95.5	**96.3**	0.121
breech	3.9	3.0	3.0	4.2	**3.5**	
other and unknown	0.2	0.3	0.3	0.3	**0.3**	

Table 6.7: Rates of LBW, term LBW and SGA by IMD 2010 quintiles of deprivation (11,928 eligible mothers with singletons, but only 11,874 had live births with recorded birth weight, see Figure 6.3)

	Quintile	Population*		LBW		term LBW		SGA	
		Freq	%	Freq	%	Freq	%	Freq	%
Most deprived	1	6506	54.8	531	8.2	266	4.3	851	13.1
	2	1766	14.9	105	6.0	62	3.7	196	11.1
↕	3	1103	9.3	62	5.6	28	2.7	90	8.2
	4	284	2.4	11	3.9	5	1.8	17	6.0
Least deprived	5	164	1.4	7	4.3	3	1.9	14	8.5
	missing	2051	17.3	178	8.7	86	4.5	272	13.3
	total	11874	100	894		450		1440	

*among eligible set, with LBW values

Table 6.8: Pros and cons of major birth outcomes

	+	-
LBW	• cheap, precisely recorded, available in vast numbers • a powerful predictor of an individual baby's survival • on a population level, mean birth weight is associated with infant mortality • associated with health outcomes later in life (Wilcox 2001)	• mixes PT and term babies, such that it is not specific to prematurity or growth restriction aetiologies • causal role of birth weight in infant mortality controversial; if non-causal, may be an unimportant endpoint in itself, and inconsequential in the analysis of infant mortality or other outcomes (Wilcox 2001) • interventions to increase birth weight may therefore sometimes be wasted
term LBW	• better approximation of growth restriction than LBW	• does not take growth restricted babies among preterm births into account (approx. 50% of LBW are preterm (McKeown and Gibson 1951))
SGA	• captures almost all term LBWs and a few more	• mixes PT and term babies, like LBW • approx. 10% in each gestational age group by definition (even if based on a referent), though there should be more grow-restricted births among the preterm; would need SGA criteria based on all fetuses at a given gestational age (such as intrauterine fetal weight, currently mostly unavailable and/or unreliable) in order to do this (Wilcox 2010) • defines all small babies as grow-restricted, and no others

Table 6.9: Representativeness: difference in my sample size compared to sample size published in cohort profile paper (Wright et al. 2012)

Cohort Profile (Wright et al. 2012)	**My dataset**	**Difference**
13,818 babies	13,525 babies	293
13,455 singletons	13,199 singletons	256
177 twin sets	158 twin sets	19
3 triplet sets	3 triplet sets	
13,740 live births	13,453 live births	287
78 stillbirths	72 stillbirths	6
11,396 baseline questionnaires	11,391 non-duplicate questionnaires (of which 11,129 reported water consumption)	5

6.6 Figures

Figure 6.1: Birth weight distribution (live singleton births to eligible BiB women); solid line marks the 2500g cut point (any birth to the left of the solid line are considered LBW); dashed line marks the mean birth weight of 3229.7g (N=11,874)

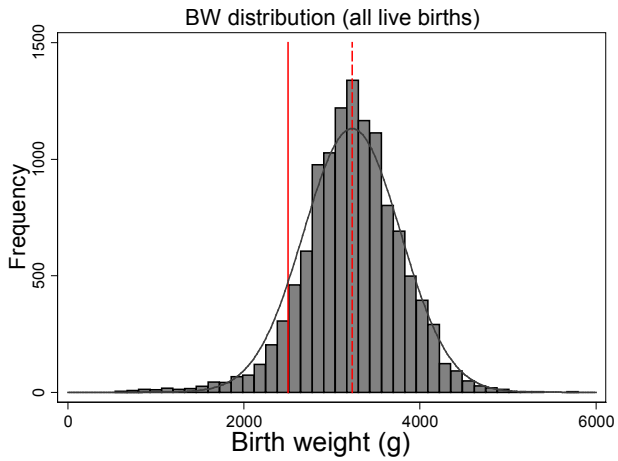


Figure 6.2: Birth weight distribution stratified by term (top plot, N=11,211) vs. preterm births (lower plot, N=663), live singleton births to eligible BiB mothers; solid line marks the 2500g cut point.
Any births to the left of that line on the top plot are considered term LBW
Any births to the left of that line on the lower plot are considered small and preterm (Wilcox 2010)

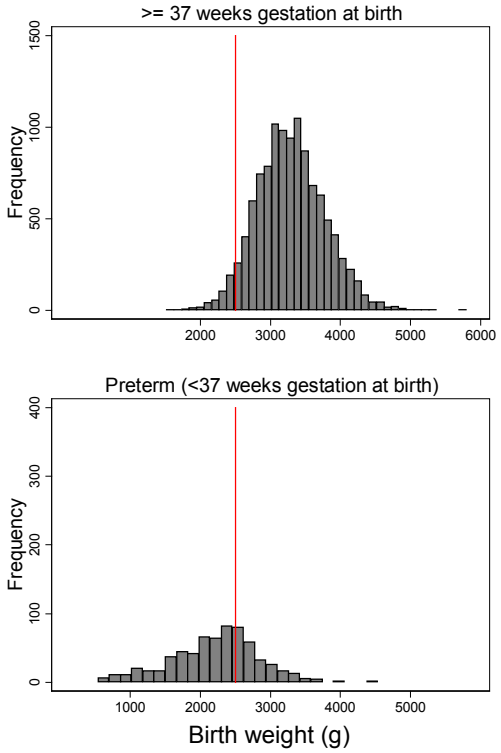
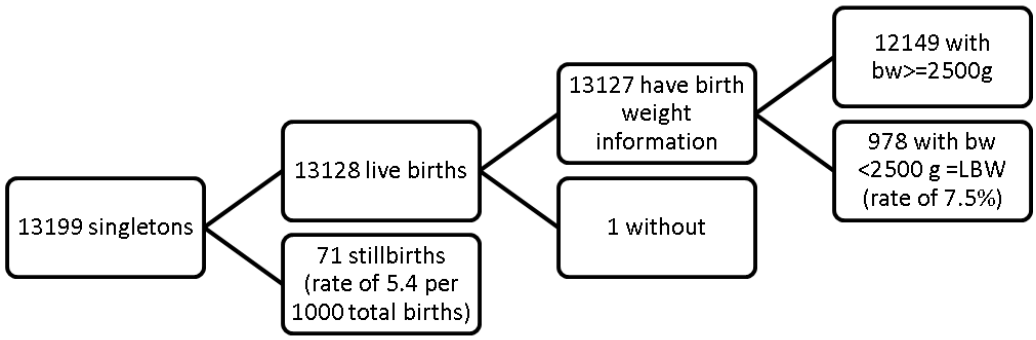


Figure 6.3: Flowchart of sample size loss for low birth weight (LBW)

a) Full cohort



b) Subset in this analysis

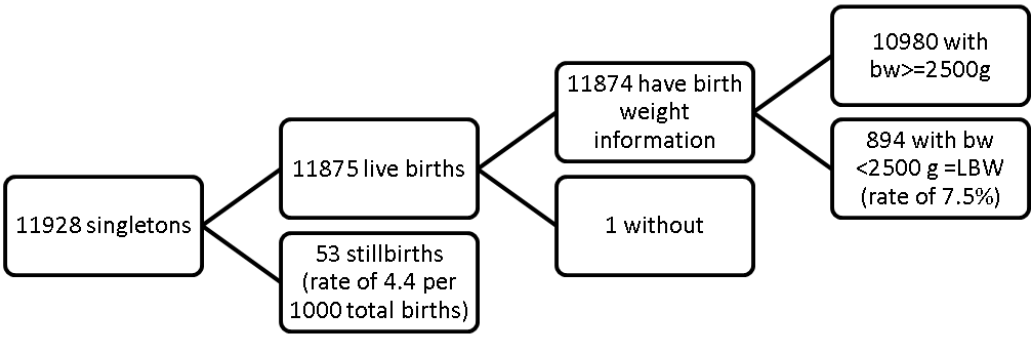
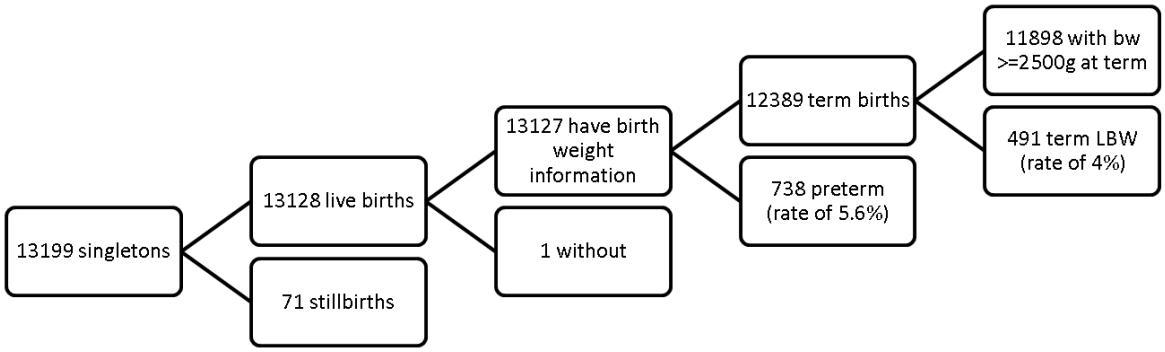


Figure 6.4: Flowchart of sample size loss for term low birth weight (term LBW)

a) Full cohort



b) Subset in this analysis

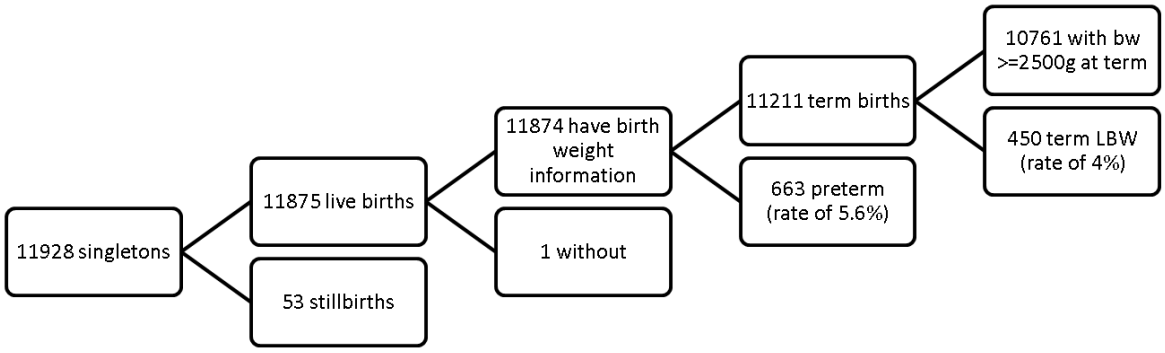
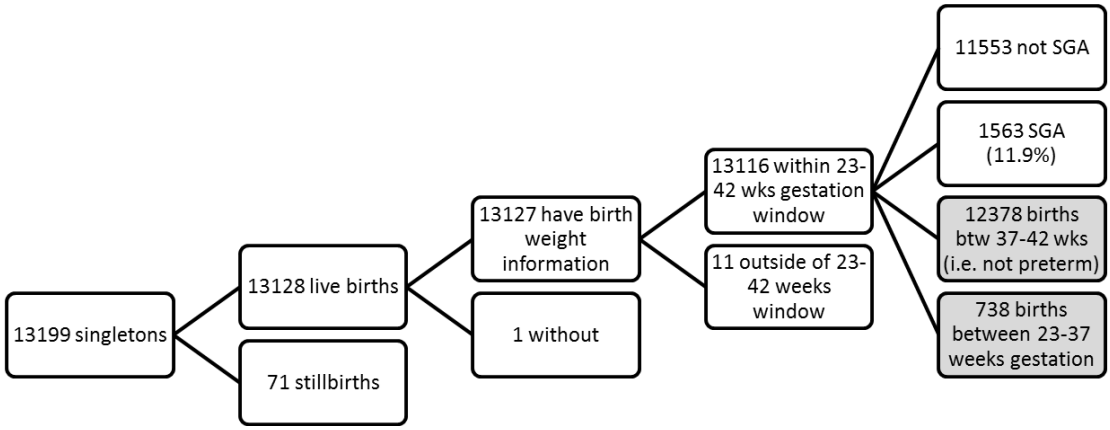


Figure 6.5: Flowchart of sample size loss for small-for-gestational age (SGA) (and shaded grey, for SGA and preterm)

a) Full cohort



b) Subset in this analysis

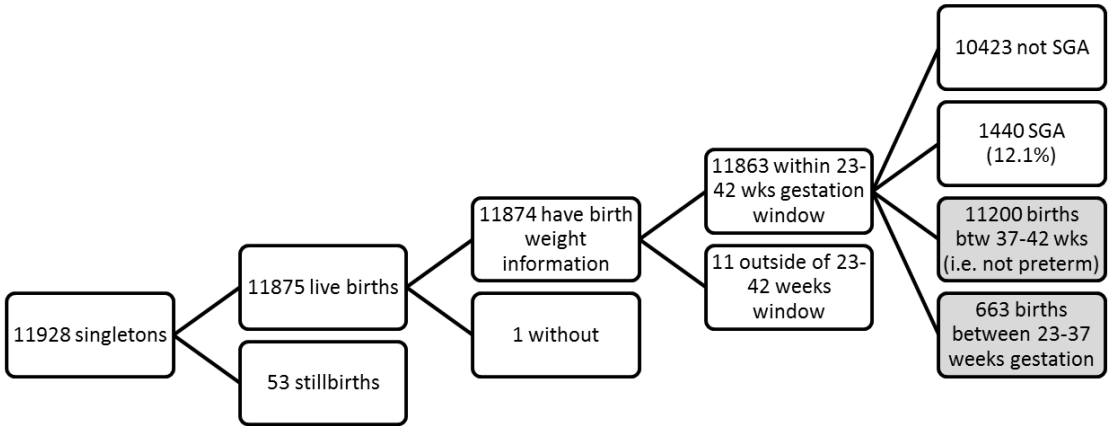


Figure 6.6: SDS (or z score) distribution; N=11,863 (11,928 eligible women with singletons, but only11,863 had live births with recorded birth weight, sex of child, and gestational age, and born within 23-42 window of gestational weeks, see Figure 6.5). Any births to the left of -1.282 are considered small-for-gestational age (SGA)

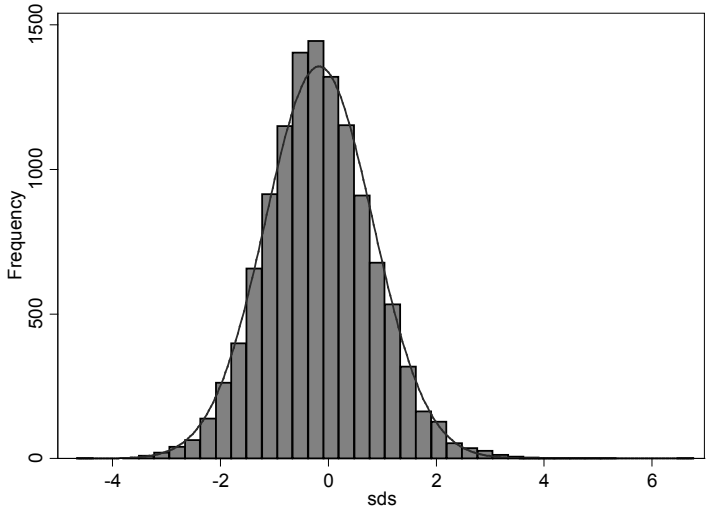


Figure 6.7: Birth weight distributions for selected ethnic groups: live singletons (P=Pakistani, WB=White British)

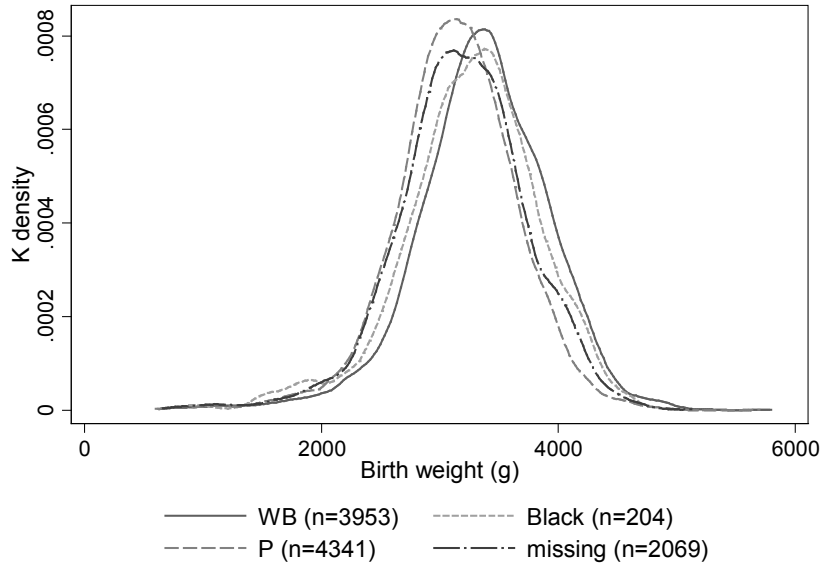


Figure 6.8: Mean birth weight (and 95% CI) of babies born at 40 weeks gestational age by ethnic group (sample sizes are in brackets)

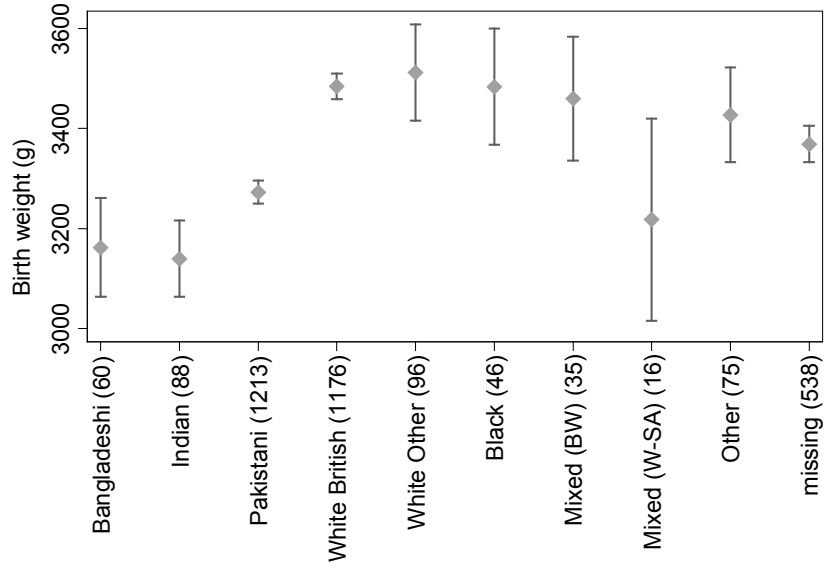
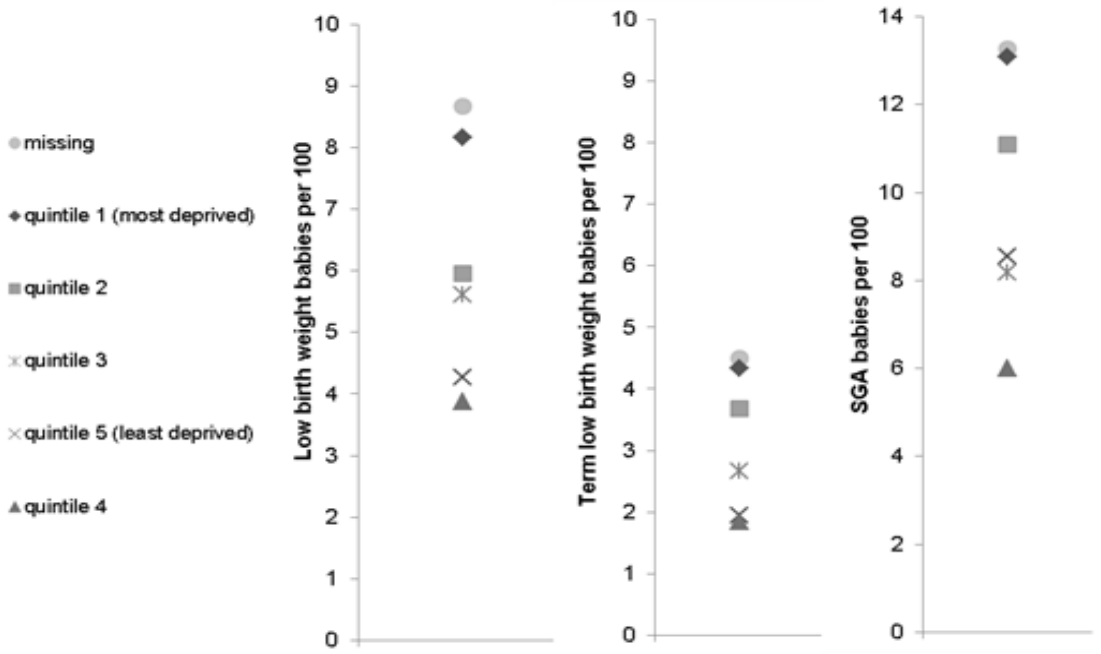


Figure 6.9: Low birth weight (LBW), term LBW and small-for-gestational age (SGA) rates by a) Index of Multiple Deprivation (IMD) quintiles of multiple deprivation 2010 and b) ethnicity.
Sample size for birth weight≥2500g=10980, for LBW=894 (see Figure 6.3);
Sample size for birth weight≥2500g at term=10761, for term LBW=450 (Figure 6.4);
Sample size for not SGA=10423, for SGA=1440 (Figure 6.5).

a) By IMD quintiles of multiple deprivation 2010



b) By ethnicity

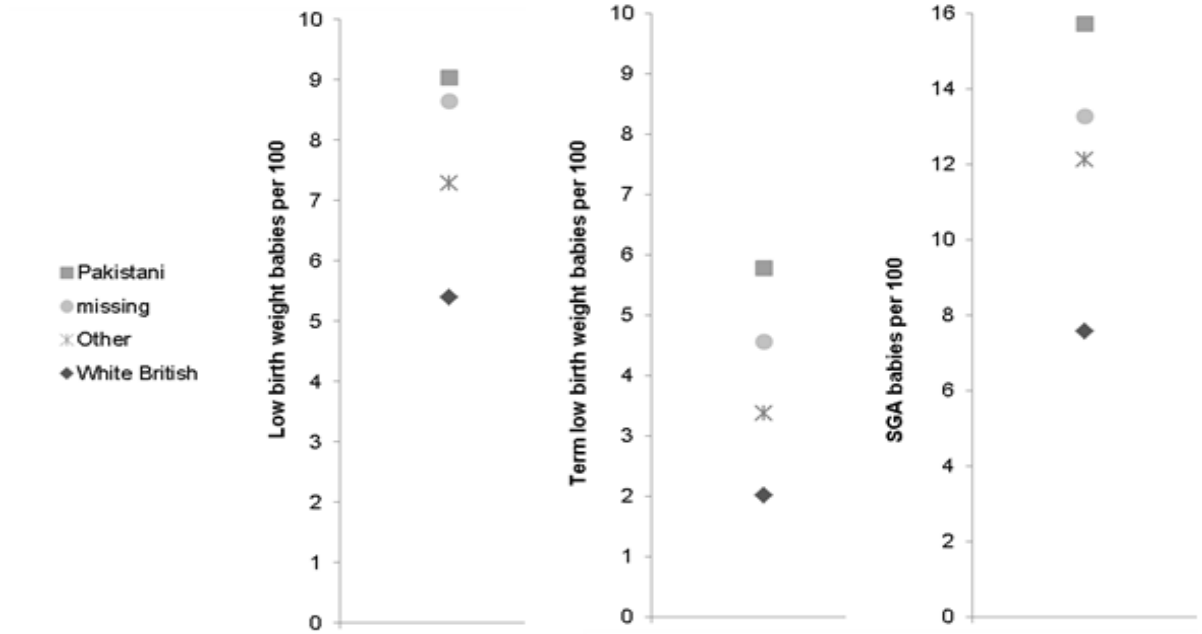
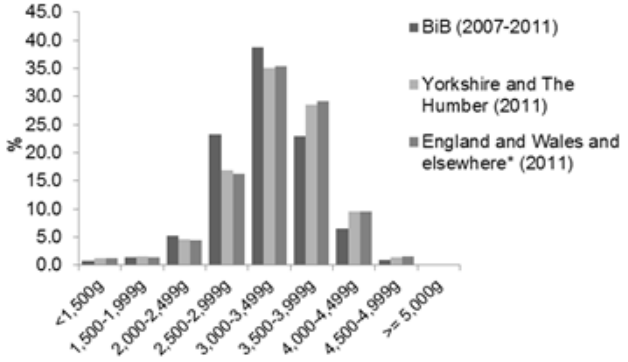
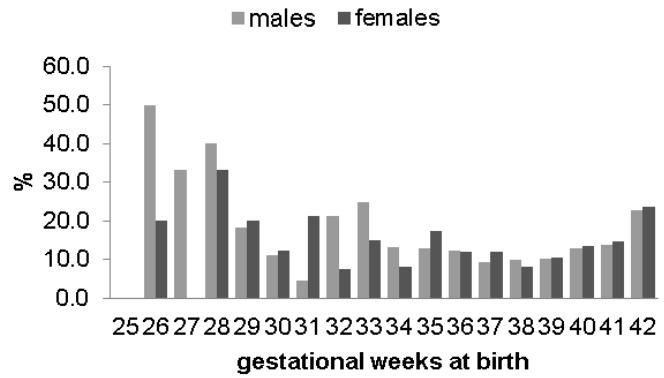


Figure 6.10: Categorical birth weight distributions for 11,874 live BiB babies, compared to birth weights in Yorkshire and The Humber, and England & Wales * (ONS 2011 figures) (11,928 eligible women with singletons, but only 11,874 had live births with recorded birth weight, see Figure 6.3)



*includes births to women whose usual residence is outside England and Wales

Figure 6.11: Rate of small-for-gestational age (SGA) by gestational age and sex of child (top: %, bottom: frequency); N=11,863 (11,928 eligible women with singletons, but only11,863 had live births with recorded birth weight, sex of child, and gestational age, and born within 23-42 window of gestational weeks, see Figure 6.5)



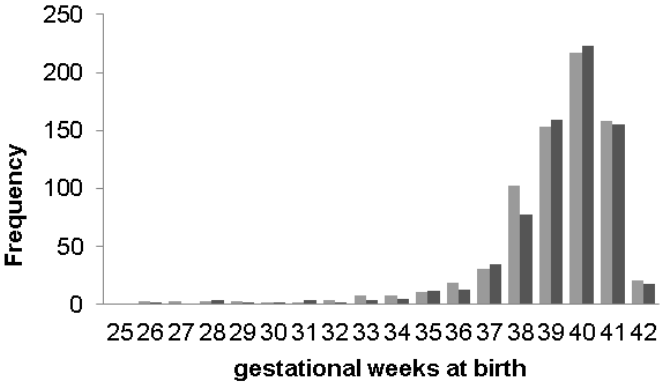
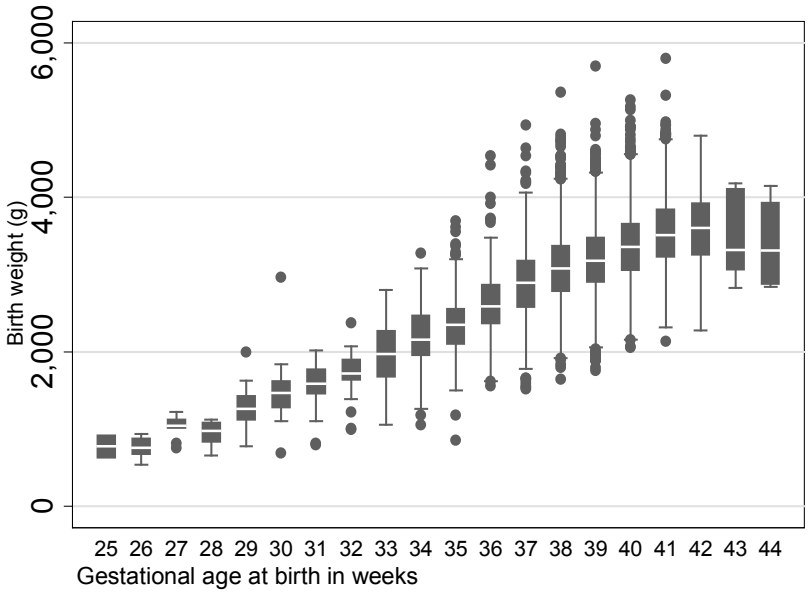


Figure 6.12: Birth weight (g) by gestational age (weeks) among 11,875 women (11,928 eligible women with singletons, but only 11,875 had live births with recorded gestational age)



CHAPTER 7 EPIDEMIOLOGIC ANALYSIS

This chapter presents the results of the epidemiologic analyses looking at the associations between exposure to haloacetic acids (HAAs) and risk of poor birth outcomes (Aim 2, Chapter 2). The three exposure measures of interest are mothers' water consumption, exposure to area-level DCAA, TCAA and BDCAA concentrations based on maternal water zone of residence, and overall exposure to DCAA, TCAA and BDCAA—which combines the previous two metrics. The outcomes of interest are continuous birth weight, and the dichotomised birth weight measures term low birth weight (LBW) and small-for-gestational age (SGA).

7.1 Background

There are three classes of HAAs exposures of interest here: water consumption (described in Chapter 3), trimester-specific area-level concentrations to DCAA, TCAA and BDCAA (Chapter 4), and their combined metrics (Chapter 5) (see Table 1.4). All are categorised by tertiles, which the variability in the exposure distributions can support. This also minimises any assumptions about the linearity of the relationship between exposure and outcome (see Figure 3.2 in Chapter 3 for distributions of water consumption with tertile cutpoints, see Figure 4.15 for distributions of area-level concentrations with tertile cutpoints, and see Figure 5.1 in Chapter 5 for histograms of combined metrics distributions with tertile cutpoints).

7.1.1 Possible confounders

In addition to ethnicity described in Chapter 6, the potential confounders discussed below are all known to be associated with birth weight (West 2011). Among them, smoking, caffeine intake, maternal education and BMI are also correlated with exposure (either with water consumption, or residence-based area-level DCAA, TCAA and/or BDCAA concentration, which make up the combined metric) and are therefore adjusted for accordingly (section 7.2.1).

7.1.1.1 Maternal/family characteristics

A number of non-behavioural maternal factors are relevant to birth weight and fetal growth.

7.1.1.1.1 Body Mass Index (BMI)

Short maternal stature has been associated with an increased risk of having a small-for-gestational age (SGA) baby and this association is reported to persist over at least two generations, i.e. birth weight is lower in infants whose grandmothers were of shorter stature (Klebanoff et al. 1997). A later study also

identified maternal height as a strong predictor of neonatal length and suggested that geographical differences in birth size can, in part, be explained by differences in maternal size (Leary et al. 2006).

Higher BMI is typically associated with an increased birth weight (Frederick et al. 2008; HAPO Study Cooperative Research Group 2010), and the associated reduced risk of delivering an SGA baby (Cnattingius et al. 1998). In the UK, South Asian women are slightly shorter than White British women (Kelly et al. 2009) but tend to have a higher BMI (Health Survey for England 2005). In BiB however, Pakistani women with weight information have lower BMI than the White British women which compounds the expectation that their babies will have lower birth weight on average (see Chapter 6).

7.1.1.1.2 Maternal age

An analysis of 36,056 New York City singleton babies found a significant progression of birth weight with advancing maternal age, with some evidence that young maternal age (<18 years) increases the risk of low birth weight (LBW) (MacLeod and Kiely 1988). The prevalence of SGA was later also found to be lowest in mothers aged 26-30 years old with similar increasing prevalence at younger and older ages outside this range (Lawlor et al. 2011). Socio-economic position and other characteristics shared by sisters in a large within-sister analysis appear to explain most of the association of young maternal age with adverse perinatal outcomes, but the association of older maternal age with preterm birth and SGA is not explained by this confounding and may even be masked by it (Lawlor et al. 2011).

7.1.1.1.3 Parity

Increasing parity has been associated with higher birth weights (Wilcox et al. 1996) although birth weight seemingly drops markedly with high parity (4 or more) (MacLeod and Kiely 1988). Joshi et al's study of 770 mothers and their babies from rural India found increasing parity (when mothers had 1 previous birth (i.e. primiparous), 2 previous births, or 3 or more previous births) was associated with increased birth weight and subscapular skinfold thickness (Joshi et al. 2005), but this study was not able to clarify whether weight and skinfold thickness reduced with 4 or more previous births as described by McCleod & Kielty (1988).

As recorded by West (2011), early studies in the UK suggested that there were differences in average maternal age and parity between South Asian mothers and White mothers (Asian women tending to start their families at younger ages and to have more children than the White women) (Lindley et al. 2004). While this may hold for first generation South Asian women, differences in parity have declined in second generation UK born South Asians (Dhawan 1995; Harding et al. 2004).

7.1.1.1.4 Socio-economic status (SES)

In high income countries, major disparities in birth weight have been evident across different social groups (Bambang et al. 2000). Lower social groups have lower birth weights (Spencer et al. 1999) and although mean birth weight has increased over the last century the social economic gradient in birth weight has remained unchanged (Spencer and Logan 2002). Social deprivation is also a risk factor for intrauterine growth restriction (Gardosi 2009).

On average, Pakistani and Bangladeshi communities in the UK are poor (Nazroo 2001; West 2011). As a consequence, social economic position may contribute to ethnic variations in birth size, and ethnicity may be a marker for social disadvantage in birth weight differences. Data from the Millenium Cohort Study (MCS) suggest that socio-economic factors accounted for 23% of the 305g difference in mean birth weight between Pakistani and White British babies (Kelly et al. 2009). Any effect of socio-economic position on birth weight is likely to be mediated by more proximal characteristics such as smoking. In fact, a systematic review concluded that maternal smoking was the strongest explanation for the association of socio-economic position with variation in birth weight (Kramer et al. 2000).

7.1.1.2 Behaviours

7.1.1.2.1 Smoking

A number of maternal health-related behaviours and the ensuing exposures of the fetus to maternal toxins can contribute to the development of a growth restricted fetus. Key amongst these is maternal smoking during pregnancy which is strongly and consistently associated with lower birth weight (Cliver et al. 1995; Horta et al. 1997; Kramer 1987; Nieuwenhuijsen et al. 2013; Roquer et al. 1995; Vik et al. 1996; Zaren et al. 1996) and an increased risk of having an SGA infant (Cnattingius et al. 1984; DiFranza and Lew 1995; Meis et al. 1997; O'Callaghan et al. 1997). Smoking has been shown to affect birth weight in a dose-response relationship causing a deficit of around 200g (Pringle et al. 2005) and up to 250g at term (Gardosi et al. 1995). Conversely, randomised controlled trial studies on smoking cessation during pregnancy have been found to mitigate the smoking associated deficits in birth weight (Haddow et al. 1991; Li et al. 1993; Lindley et al. 2000). Lastly, meta-analyses find that exposure of non-smoking pregnant women to second-hand tobacco smoke also reduces mean birth weight (by 33g or more (Leonardi-Bee et al. 2008)) and increases the risk of LBW (Leonardi-Bee et al. 2008; Salmasi et al. 2010).

In terms of mechanism, cigarette smoking is thought to reduce uterine blood flow, limiting fetal oxygenation and attenuating growth (Haworth et al. 1980). The quantity of cigarettes smoked per day positively correlated with the degree of intrauterine growth restriction (Andres and Day 2000; Cliver et al. 1995; Haworth et al. 1980; Meyer 1978).

As discussed in Chapter 6, historically, smoking is uncommon in South Asian women in the UK and therefore unlikely to be a major cause of lower birth weight in South Asian compared to other populations.

7.1.1.2.2 Alcohol

Alcohol is a teratogen that can cross the placenta during pregnancy to enter the baby's blood, potentially altering the development of the fetus. Frequent alcohol use early in pregnancy has been linked to congenital malformations of the heart, brain and kidney (RCOG 2006) and fetal death (Andersen et al. 2012). Heavy alcohol consumption throughout pregnancy causes fetal alcohol syndrome (Jones and Smith 1973) and has been associated with preterm birth and growth restriction (Patra et al. 2011). However, moderate consumption has not been shown to have any consistent effect (Oster 2013). A recent study in the BiB cohort found that binge drinking during the second trimester of pregnancy was associated with an increased risk of SGA birth, but found no association between any level of alcohol consumption and premature birth (Cooper et al. 2013).

Prolonged maternal ingestion of alcohol or other drugs (e.g. steroids, warfarin/Coumadin, hydantoin, cocaine, and heroin) are also implicated in the development of intrauterine growth restriction (Brodsky and Christou 2004; Lee et al. 2003).

7.1.1.2.3 Caffeine

Peck et al's review of the 16 human studies of caffeine and reproductive health published between 2000 and 2009 conclude that the studies of caffeine and fetal growth restriction are equivocal, with approximately half of the studies reporting weak associations with intrauterine growth restriction or reduced birth weight, and half observing no effects (Peck et al. 2010). The authors argue that the strength of the evidence for a potential effect of caffeine on fetal growth restriction is diminished by the inability to rule out alternative, credible explanations for the observed associations, namely confounding by pregnancy symptoms and aversions.

7.1.1.2.4 Maternal diet

Periods of famine in the Netherlands, Germany and the former Soviet Union have shown that severe maternal malnutrition can impair fetal growth (Barker 1994; Roseboom et al. 2001). The severity of the lack of food supply (Prentice and Cole 1994) and the length of malnutrition (Roseboom et al. 2001) correlated with the degree of growth delay. While the fetus is affected by chronic severe maternal malnutrition, it seems to be fairly resistant to acute malnutrition, particularly if it occurs late in gestation (Brodsky and Christou 2004; Schwartz 2000).

What's more, supplementation during pregnancy (such as iron supplementation) can lead to a significantly higher mean birth weight, a significantly lower incidence of low birth weight infants, and a significantly lower incidence of preterm low birth weight infants (Cogswell et al. 2003).

7.1.1.3 Pathophysiology

7.1.1.3.1 Pregnancy complications

Chronic maternal vascular disease due to hypertension, diabetes mellitus, renal disease, or collagen vascular disease is the most common cause of intrauterine growth restriction in developed countries (Lin and Santolaya-Forgas 1998). The most profound effects are observed if the hypertension is early onset, severe, or due to chronic hypertension with superimposed pre-eclampsia (Brodsky and Christou 2004; Resnik 2002) (see section 2.1.3.2).

7.1.1.3.2 Gestational diabetes

Maternal glucose intolerance and diabetes have been associated with greater birth weight, fetal adiposity and risk of macrosomia (Catalano et al. 2003; Jovanovic and Pettitt 2001). This association is continuous across the whole distribution of maternal glycaemia in pregnancy (Metzger et al. 2008).

As mentioned in Chapter 6, South Asian populations have a higher risk of gestational diabetes compared to other ethnic groups (Nanda et al. 2011). Thus, given the robust associations of greater maternal glycaemia in pregnancy with greater birth size and infant adiposity, were it for this variable alone, one would expect that South Asian babies to have higher average birth weights and greater adiposity than White babies (West 2011).

In addition to these factors, physical constraints such as large placental abnormalities, uterine masses, or multiple gestation (multiple births were excluded from analysis, see section 3.2.3) can lead to growth restricted fetuses (Brodsky and Christou 2004).

7.1.1.4 Fetal influences

7.1.1.4.1 Sex of child

Sex of child is known to be associated with birth weight. Boys have been reported to have a higher mean birth weight than girls and a lower rate of LBW (Halileh et al. 2008). This sex difference has been seen in most ethnic groups in the UK including South Asians (Margetts et al. 2002; West 2011).

7.1.1.4.2 Gestational age in weeks

Prematurity is known to be linked with growth restriction (Bukowski et al. 2001; Gardosi 2005; Jacobsson et al. 2008; Ott 1993; Tamura et al. 1984; Zeitlin et al. 2000) (Chapter 6). However, it is still unknown whether preterm and growth restricted babies share the same aetiology, or whether being growth restricted leads to preterm delivery (Savitz et al. 2002). Whether or not gestational age at

birth should be adjusted for in the models therefore remains debated. As it is strongly associated with birth weight, its effect is adjusted for and its possible mediation effect studied.

There is some evidence that gestational length varies by ethnic group (Kelly et al. 2009; Patel et al. 2004) although findings are inconsistent. Patel et al (2004) found a higher proportion of preterm deliveries in Asians compared to White Europeans *(*6.5% and 5.1 % respectively). However, data from the Millenium Cohort Study suggest that when South Asian populations are categorised according to country of origin, the proportion of preterm births varies between these different groups, with the proportion in Indian women (9.5%) being higher, and that in Pakistani women lower (5.7%) than the proportion in White women (6.7%) (Kelly et al. 2009). This was supported by findings in BiB (see section 6.3.4.1). Including preterm births in estimates of mean gestational length masks potential differences in term gestation and it has been suggested that what constitutes 'term' may vary by ethnic group with South Asian infants in particular reaching maturity at an earlier gestation than White populations (Balchin and Steer 2007; West 2011) (see section 6.1.1.1).

7.1.1.5 Access to antenatal care

Access to antenatal care has been shown to improve maternal, perinatal, and neonatal outcomes (Villar et al. 2001). However due to the nature of the free National Health Service (NHS) in the UK, and to the fact that this cohort was recruited exclusively from one hospital, the Bradford Royal Infirmary, access to healthcare is in theory the same for all BiB participants and is therefore not a primary cause for concern in this study and was therefore not included as a confounder in the model.

Nevertheless, Firdous & Bhopal (1989) identified low levels of literacy among South Asians in the UK for whom English was commonly a second language; both of these factors can hinder uptake of health services in terms of direct communication with health workers, but also in terms of a poor understanding of health services and how they operate (Abba, 2001).

7.1.1.6 Link with exposure

As per Smith et al. (2009), women's total tap water intake overall, at home and/or at work all differed by age, ethnicity, and income (i.e. a measure of socioeconomic status) categories in the BiB cohort. Caffeine intake (via coffee or tea consumption) is associated with tap water consumption by definition, and smoking was also found to be associated with increased tap water consumption—the two habits possibly occurring together.

Figure A7 - 1 in the appendix summarises the interrelationship between all variables involved including with the exposure.

7.2 Methods

7.2.1 Model selection process

Based on the evidence presented above for their potential to confound the relationship between exposure and outcome by being related (causally or not) to the exposure as well as causally related to the outcome, I have included the following ten covariates *a priori* in all the birth weight and term LBW models: sex of child, parity, ethnicity, maternal education, gestational age in weeks, BMI, maternal age, smoking during pregnancy, caffeine intake during pregnancy, and gestational diabetes. In the SGA models, all the above except for gestational age and sex of the child—information which is incorporated into the outcome variable—were included.

As described in section 6.3.3, less than 1% of births were to women with previous diabetes, pre-existing hypertension issues or who became hypertensive during labour, whilst prevalence of gestational diabetes, pregnancy-induced hypertension and pre-eclampsia were greater (8.1%, 5.6% and 2.7%, respectively). Pre-existing hypertension was associated with all other variables (all p-values<0.05), and all three hypertension variables were associated with pre-eclampsia, as expected given that pre-eclampsia is a condition of high blood pressure (see Table A7 - 1). Gestational diabetes was adjusted for but not pregnancy-induced hypertension or pre-eclampsia as these did not significantly improve the model fit (data not shown).

Initially, two other model selection methods were tested before settling on the *a priori* approach: a) forward stepwise regression (using STATA's "xi: sw, pe(0.05)[22] lockterm1 lr:" command), and b) inclusion of a set of seven *a priori* covariates[23] followed by any covariates from a list[24] that led to a significant change in the model deviance (p<0.05) or to a >5% change in the coefficient or log odds ratio (Greenland 1998; Toledano et al. 2005). These two alternative methods did not alter the results conclusion (and led to arguably over-parametrised models in the case of the latter method). As the effect of exposure in these models is small or inexistent, any covariate can easily lead to a 5% change in its coefficient/log OR such that all covariates are kept in the model by definition even when they do not have significant p-values. Using the uniform approach of including the same ten (or eight for SGA) covariates selected *a priori* in all models was deemed best to avoid confusion and added difficulties in interpretation due to too many slightly different models.

---

[22] pe= significance level required for addition to model
Sequential covariates selection: start with an empty model, find the most significant additional covariate (with an F-test) beyond exposure (which is forced into the model by lockterm1); if its p-value is less than the cut-off (0.05 in this case), add it to the model and re-fit the model with the new set of covariates; repeat two previous steps until no further covariates can be added.

[23] sex of child, ethnicity, gestational weeks, parity, maternal age, maternal BMI, and smoking during pregnancy

[24] Sequentially: maternal education, employment, IMD 2010 quintiles of deprivation, caffeine intake, alcohol consumption during pregnancy, exposure to second-hand smoke, gestational diabetes/pre-eclampsia/or pregnancy-induced hypertension, season of birth

7.2.2 Data and covariate derivation

The data described in this chapter were received in February 2013 from the BiB team and contained both the routine eClipse and the additional backfilled data. The derivation of ethnicity, BMI, maternal education as a proxy for SES and caffeine intake is described in Chapter 2.

7.2.2.1 Body Mass Index (BMI)

Body Mass Index (BMI) is defined as maternal weight in kilograms at the time of questionnaire divided by the square of the maternal height in metres (see Chapter 2, section 2.1.2). Because women will typically gain between 5kg and 18kg over the course of pregnancy (depending on pre-pregnancy weight) according to the Institute of Medicine (2009), usual BMI cut-off points for non-pregnant persons (<18.5=underweight, 18.5- <25=normal, 25- <30=overweight, >30=obese) are meaningless for women in their second trimester. BMI was therefore categorised into quartiles instead for analysis. This assumes that lighter and heavier women pre-pregnancy put on the same number of kilograms over the course of their pregnancy such that their categorisation by quartile will correctly estimate their BMI pre-pregnancy relative to one another.

7.2.3 Statistical analyses

Univariate and multiple linear and logistic regressions as well as multiple imputations with chained equations (*mi* command) for combined metrics and their water consumption and area-level concentration components on all three outcomes were all run in STATA 12.1. The lowest tertile of exposure (tertile 1) was used as the reference in all models.

The results of the interaction between combined metric and gestational age are presented for both birth weight and term LBW (not relevant to SGA, which already includes gestational age) to study the possible effect modification that gestational age may have on the results.

7.2.3.1 Sensitivity analyses

A number of sensitivity analyses for each of the three outcomes were performed to check the assumptions behind the derivation of the exposure measures.

7.2.3.2 Multiple imputation (combined metrics models only)

To avoid losing observations due to missing values, I ran multiple imputations by chained equations on my final models using combined metrics as the exposure. Multiple imputation enabled an increase in power, using all available information without restricting to values which have complete information for all covariates, and thus generating estimates with smaller confidence intervals and increasing the ability to reject the null hypotheses. In addition, if the reason a confounder variable was

missing was related to the outcome, a complete case analysis would have introduced bias in the model's estimates (Rubin 1996). Joint imputation of all variables removes the potential bias that incorporating them consecutively, in a given order, might impose.

Only one birth (out of all singleton live births eligible for the analysis, i.e. N=11,875) was missing a birth weight value, was excluded from the associated analysis and not imputed. The maximum imputed sample size was therefore 11,874 for the continuous birth weight models, and 11,211 for the term LBW models (excluding 663 preterm births). Because the SGA referent used is specified to be only valid for the 23-42 week range of gestational ages, a further 11 babies were excluded, resulting in a maximum imputed sample size of 11,863 babies for the SGA models (see flowcharts in Figure 6.3, Figure 6.4 and Figure 6.5 for details).

There were no missing values for sex of child, gestational age (in completed weeks), and maternal age. All other variables (including the exposure) that had missing data were imputed. The binomial variables caffeine intake and gestational diabetes were imputed by logistic regression; missing values in the smoking, maternal education, and ethnicity variables were imputed by multinomial logistic regression; and the exposure (combined DCAA, TCAA, or BDCAA metric, categorised into tertiles), parity and BMI (quartiles) were imputed using ordinal logistic regression.

The missingness pattern was assumed to be missing at random. That is, the probability of a particular value being missing depends only on the observed data, and the complete cases are not a random sample. If the missing data are missing at random as assumed, complete cases analysis gives biased results but a correctly specified multiple imputation does not.

Ten different datasets were imputed to fit the epidemiologic regression models, in order to account for uncertainty about the imputed values (Rubin 1987). For the linear models on the continuous outcome birth weight, the imputation models were regressed on birth weight. For the logistic models on binary outcomes, the imputation models were regressed on the binary outcome (either term LBW or SGA); there was no material difference when imputing on birth weight or standardised birth weight (not presented).

7.3 Results

7.3.1 Relationship between combined metrics and birth weight or standardised birth weight

There is no apparent descriptive relationship between exposure (combined metrics) and outcome (Figure 7.1, Figure 7.2). These scatter plots do not signal an increasing or decreasing trend between

exposure and outcome, and the distributions of birth weight or standardised birth weight by high (third) vs. low (first) tertiles of exposure appear to overlap.

For example, mean birth weight of singletons exposed to high (> 17.29 ug/day) vs. low (0 - 10.41 ug/day) DCAA tertiles, weigh 3223.9g (95% CI: 3200.0, 3247.8) vs. 3200.7g (3176.8, 3224.6), respectively, corresponding to a 23.2g mean difference, or <1% change in birth weight, in the contrary direction to the hypothesis. High (> 19.79 ug/day) vs. low (0 - 12.02 ug/day) TCAA tertiles, and high (> 1.93 ug/day) vs. low (0 - 1.04 ug/day) BDCAA tertiles of exposure are also associated with a <1% change in mean birth weight in the contrary direction to hypothesis.

Differences for term LBW and SGA rates by exposure group are also in the opposite direction to the hypotheses that higher exposure leads to a decrease in birth weight and an increase in term LBW and SGA rates. Term LBW prevalences among the high DCAA, TCAA and BDCAA tertiles of exposure are 3.7%, 3.7% and 3.3% vs. 4.7%, 4.8% and 4.7% in the low exposure tertiles respectively; SGA prevalences among the high DCAA, TCAA and BDCAA tertiles are 12.0%, 12.7% and 12.2% vs. 13.4%, 13.6%, and 12.7% in the low exposure tertiles, respectively (Figure A7 - 2, Figure A7 - 3, Figure A7 - 4 and Figure A7 - 5). For full descriptive details on birth weight, LBW, term LBW, and SGA prevalences by tertiles of exposure to each of the 16 exposure measures of interest in this analysis (4 water consumption only, 9 area-level concentrations, and 3 combined metrics), see Table 7.1.

These descriptive plots show that differences in mean birth weights by exposure groups and in rates of term LBW and SGA by exposure group appear to be small and in the contrary direction to hypothesis. It can therefore be anticipated that a strong association between combined metrics and outcome is unlikely to be detected.

7.3.2 Continuous birth weight models

7.2.2.1 Main results

Table 7.2 presents the results of the crude and multiple linear regression models of combined DCAA, TCAA and BDCAA metrics on birth weight, with no evidence of an association. After multiple imputation, still no results are significant, and the trend becomes more systematically contrary to hypothesis, i.e. greater exposures predicting higher birth weight (Table 7.2).

Breaking down the combined metric, Table 7.3 presents results for water consumption during pregnancy, with evidence of an association between cold tap water consumption and increasing birth weight in the adjusted models (mean birth weight change between tertile 3 and tertile 1 of exposure is

24.0g (95% CI: 1.4, 46.6)). No other adjusted models are significant. The trend appears generally contrary to the hypothesis that higher exposures predict lower birth weight.

Table 7.4 on trimester-specific area-level concentrations presents only one significant effect for third tertile of exposure to TCAA in trimester 1 compared to reference after adjustment: -28.3 (95% CI: -52.0, -4.6). This may be a chance finding given that no other exposures or exposure tertiles come out significant in the nine models (18 mean birth weight differences) presented. In these area-level concentration models, it is noteworthy that the trend is mostly in the hypothesized direction. None of the Wald F tests for joint significance of the exposure measures in any of these adjusted models are significant (data not shown), though they are borderline significant for TCAA trimester 1 (p=0.065) and BDCAA trimester 2 (p=0.074).

Comparing crude to adjusted models, nearly no changes in effect direction were found. However, significance in p-values for the third tertiles of total tap water and bottled water, and for both second and third tertiles of total water and area-level exposure to DCAA for all three trimesters is lost. In addition, most effect estimates are closer to the null in the adjusted compared to crude models, suggesting that the association is indeed confounded by the covariates adjusted for.

Figure 7.3, Figure 7.4 and Figure 7.5 depict the mean change in birth weight compared to reference for these adjusted models.

Spearman correlations between the ten covariates included *a priori* in these models are presented in Table A7 - 2 and show that smoking, ethnicity, and caffeine co-vary (rho between ethnicity and smoking is -0.42 ($p<0.001$); rho between ethnicity and caffeine intake is -0.28 ($p<0.001$); rho between smoking and caffeine intake 0.35 ($p<0.001$)), and parity is significantly correlated to all three of these as well—though with smaller correlation coefficients (rho between parity and smoking is -0.11 ($p<0.001$); rho between parity and ethnicity is 0.10 ($p<0.001$); rho between parity and caffeine intake is 0.02 ($p=0.038$)).

The association between the combined exposures of interest and birth weight or growth restriction could be modified by gestational age at birth (which is highly significant in the models, and thought to either be a mediator of adverse birth outcomes or a cause in and of itself). However, no interaction terms were significant (Table A7 - 3).

7.3.2.2 Additional analyses

For the continuous birth weight models, the combined metrics models to the White British and Pakistani women were stratified to compare effects in the two predominant ethnic groups. In light of

previous literature regarding the difference in accuracy of reporting of water ingestion by employment status (reporting being more accurate in the women who stay at home (Smith et al. 2013)), analyses both of the combined metric and the water consumption variables were also restricted to the unemployed women to see if coefficients would change.

7.3.2.2.1 Stratification by ethnicity

The ethnicity covariate is strongly associated with birth weight in all models. Being of Pakistani origin predicts decrease in birth weight of approximately 250g compared to being White British, while being of an ethnicity other than White British or Pakistani predicts a decrease in birth weight of approximately 160g compared to being White British (Table A7 - 9).

Because of an interest in the possibility that HAA exposure affects different ethnic groups in different ways, adjusted models were stratified to women of White British vs. Pakistani origin, excluding the Other grouping. Though no interaction terms were significant, there were some differences in the effect of exposure on birth weight by ethnic group, with more significant associations appearing amongst the Pakistani women (Table A7 - 10).

In the stratified analyses, babies born to Pakistani women were significantly heavier at birth with increasing maternal water consumption (contrary to hypothesis), and lighter with greater area-level maternal HAA concentration exposure (in line with hypothesis). For the combined metric, though not significant, it seems that effects on White British and Pakistani babies are operating in opposite directions to one another. HAA exposures were similar among White British and Pakistani women.

7.3.2.2.2 Restricting to unemployed women only

Previous work on water consumption in this cohort concluded that unemployed women have less error in their exposure estimates. However no differences in exposure coefficients were found when restricting the analyses to unemployed women only (data not shown).

7.3.2.3 Sensitivity analyses

For each of the three outcomes alternative exposure measures than the ones originally modelled were considered. The results did not change.

The results of the continuous birth weight models do not change when excluding outliers, i.e. any summary consumption measure greater than 10L/day which is an extreme water consumption average per day (data not shown).

I then aimed to isolate women who had potentially no HAA exposure during pregnancy (i.e. those who consumed only bottled water) and compare their outcomes to filterers (who have reduced HAA exposure for a given daily tap water consumption), and women who reported not filtering their tap water at all. Unfortunately, there were very few cases with zero exposure (N=21); however the (non-significant) effect was in the hypothesized direction: mean birth weight change was -116.3g (95% CI: -294.2, 61.5) for filterers compared to exclusive bottled water drinkers and -119.4g (95%: -295.7, 57.0) for non-filterers compared to exclusive bottled water drinkers. (It must be said that women were categorised as being only filterers or only non-filterers because of the way the questionnaire was formulated, realising that this is unlikely to accurately reflect the reality) Data quality did not allow us to look at this.

To challenge the accuracy of the area-level exposure metric, for those women who reported being employed and who listed a place of work which was located within Bradford's eight water supply zones (WSZ) (22% of the cohort, i.e. 2,628 women of 11,928 eligible women), the DCAA, TCAA and BDCAA concentrations of the work WSZ were identified and these concentrations' contribution were weighted based on number of days of work the woman reported per week; women were considered to be exposure to their residence WSZ concentrations the rest of the time. For the remaining 78% of eligible women (i.e. the unemployed, or those with insufficient or inadequate work location information to enable linkage to the appropriate WSZ corresponding to their work place), the same residence only measure was used as in the original metric (see the methods section of Chapter 4 and the Appendix to Chapter 4 for further details). On average, mean exposure levels were similar between residence only, and measures combining residence and workplace concentration or residence only (see Appendix to Chapter 4, Table A4 - 12), with TCAA concentrations perhaps the only exception, being slightly higher in the latter metric. The results of the continuous birth weight models on this newly defined cohort were very similar to the original area-level concentrations assigned to residence WSZs only, with TCAA average concentration in trimester 1 significant both by individual tertiles as well as jointly by Wald F test (p=0.017). Additionally, BDCAA average concentration in trimester 2 becomes significant for tertile 2 of exposure (mean birth weight change=-35.8g (95% CI: -60.5, -11.1)) as well as overall (Wald F test p=0.017) (Table A7 - 4). These effects are significant in the hypothesized direction.

The same concept is applied to the combined metric: instead of area-level concentrations based on the residence only, the combined metric now multiplies water consumption by area-level concentrations based on both work and residence water supply zones if available. On an aggregate level, there are nearly no differences between the exposure metrics for combined DCAA, TCAA, and BDCAA based on residence only vs. based on a combination of residence and work (if available) and residence only. There is no evidence of an association between combined metrics of exposure and birth weight (Table

A7 - 5). The coefficients for the combined metrics of exposure are slightly reduced, i.e. closer to the null, in the multiple imputation models compared to the complete case analyses. The confidence intervals cover approximately the same range.

7.3.3 Term LBW

Approx. 4% of births were categorised as term LBW.

7.3.3.1 Main results

The multiple linear regression models of combined DCAA, TCAA and BDCAA on term LBW predicted a significant decreased risk of term LBW when comparing those in the highest tertile of exposure (tertile 3) to those in the lowest (tertile 1): odds ratios (OR) for third tertile of exposure to DCAA (combined metric) vs. first tertile: 0.68 (95% CI: 0.46, 0.99); OR for third tertile of exposure to TCAA (combined metric) vs. first tertile: 0.67 (95% CI: 0.46, 0.97); and OR for third tertile of exposure to BDCAA (combined metric) vs. first tertile: 0.62 (95% CI: 0.42, 0.90) (Table 7.5). However the Wald chi^2 test for joint significance is only significant for the BDCAA combined exposure metric in the adjusted models (p=0.044). These significant results are lost after imputing missing covariates, though the trends remain (Table 7.5). The ORs for the association between combined metrics of exposure and the imputed term LBW models are very similar to those in the complete case analyses (Figure 7.6).

As per water consumption models for continuous birth weight, the trend in the term LBW models ORs is contrary to hypothesis, with increasing exposure to cold tap water, total tap water and total water all predicting a significant decrease in risk of term LBW in the adjusted models (see Table 7.6 and Figure 7.7). The Wald chi^2 tests for joint significance of the cold tap water and total tap water measures are significant (p=0.017, and p=0.046, respectively)) in the adjusted models (borderline significant for total water, p=0.081). Although not significant and involving small numbers, increasing tertiles of bottled water consumption appear to predict increasing risk of term LBW. No models of trimester-specific area-level concentrations on term LBW are significant, either by individual tertiles (Table 7.7, Figure 7.8) or overall (data not shown), with ORs mostly equal to 1.

There is an attenuation of odds ratios (ORs) towards 1 (i.e. no effect) in the adjusted models compared to crude models, though some significant associations in the crude models do remain after adjustment: e.g. cold tap water, total water, and combined BDCAA exposure (Table 7.6, Table 7.7).

7.3.3.2 Interaction between exposure and gestational age (combined metric models only)

To identify a possible role for interaction among term babies only, an interaction term was included between the combined metric and continuous gestational age to the final fully adjusted logistic regression models of term LBW. Here again no interaction terms were significant, neither in the complete case analysis nor in the multiple imputation models (Table A7 - 6).

7.3.3.3 Sensitivity analyses

The results did not change after carrying out a number of sensitivity analyses to challenge the exposure assessment assumptions. The results of the term LBW models do not change when excluding any summary consumption measure greater than 10L/day (data not shown). The results of models on term LBW using HAA concentrations combining modelled averages from women's residence water supply zone and workplace water supply zone if available are no different to the null results of the original concentration measure (Table A7 - 7). The third tertiles of combined DCAA, TCAA and BDCAA exposure remains significant compared to reference when including work location information into the combined metric of exposure, again only reaching overall significance in BDCAA (p=0.042) (Table A7 - 8).

7.3.4 SGA

Approx. 13% of births were categorised as SGA in the final complete case analyses of combined exposures.

7.3.4.1 Main results

11 births outside the 23-42 weeks of gestation range (range for which the SGA variable is valid) were excluded from these models. In addition, sex of the child and gestational age were not covariates in the SGA models because they are redundant with the outcome.

Sex of child was incorporated to verify its effect and it was never significant in these models (complete case analysis) and so excluded (while it is invariably significantly associated with continuous birth weight and term LBW in those respective models (see detailed output in Table A7 - 9 and Table A7 - 11)).

Table 7.8 presents the results of the complete case and imputed analyses, respectively, of the crude and multiple linear regression models of combined DCAA, TCAA and BDCAA metrics on birth weight, with no evidence of an association (see also Figure 7.9). The Wald chi^2 tests for joint significance of the combined exposure measures in any of these adjusted models are not significant either (data not shown). The confidence intervals for ORs of the SGA models for which covariates were imputed are narrower than for the complete case analysis.

The results of the models of water consumption exposure on SGA are contrary to the hypothesis, as observed for continuous birth weight and term LBW models, with no significant adjusted effects (Table 7.9 and Figure 7.10). The Wald tests for joint significance are not significant in the adjusted models either for cold tap water, total tap water or total water (p=0.060, p=0.208, p=0.167 respectively). No models of trimester-specific area-level concentration on SGA are significant, either by individual tertiles (Table 7.10, Figure 7.11) or overall (data not shown), with ORs mostly equal to 1.

As with continuous birth weight models, the effect estimates in the adjusted models compared to crude models are mostly in the same direction. Significant p-values for the third tertiles of total tap water, bottled water, and area-level exposure to DCAA in the first trimester, and for both second and third tertiles of total water and area-level exposure to DCAA for second and third trimesters are lost.

7.3.4.2 Sensitivity analyses

The results did not change after carrying out the sensitivity analyses to challenge the exposure assessment assumptions. The results of the SGA models do not change when excluding any summary consumption measure greater than 10L/day (data not shown). The results of models on SGA using HAA concentrations combining modelled averages from women's residence water supply zone and workplace water supply zone if available are no different to the null results of the original concentration measure, with just the third tertile of exposure to TCAA concentration in trimester 1 becoming significant (OR: 1.26, 95% CI: 1.03, 1.53) (Table A7 - 12). Including work location information on a fraction of the cohort makes no difference either to the results of combined DCAA, TCAA or BDCAA on risk of SGA (Table A7 - 13).

7.3.5 Effect sizes of covariates in combined models

In the combined metrics models on all three outcomes, the continuous variable gestational age is significant, predicting an average birth weight increase of approximately 177g for every additional week of gestation completed at birth. As suggested by the literature, all covariates in the combined models were significant predictors of birth weight (all $p < 0.05$ by Wald test) (see Table A7 - 9, Table A7 - 11 and Table A7 - 14).

The direction of effects is also as predicted in the literature: having a baby girl, being Pakistani or of Pakistani origin, smoking during pregnancy, or ingesting >200mg of caffeine per day each individually predict lower birth weight, when all other variables are held constant. Conversely, being a second- or third-time mother, having a relatively high BMI, being educated at university-level, or

developing gestational diabetes during the pregnancy each predict heavier birth weight. The only unusual direction is maternal age: the reference group (25 to 29 year olds) have the lightest babies, while younger (< 25 years) and older (>=35 years) mothers have on average heavier babies contrary to expectations. All of these findings hold for combined DCAA, TCAA, and BDCAA results (Table A7 - 9, Table A7 - 11 and Table A7 - 14).

7.3.6 Multiple imputation

7.3.6.1 Model results after multiple imputation (only on combined metrics models)

When restricting to complete case analysis, the prevalence of term LBW and SGA was 4.1% and 12.9% respectively. After including all available data (nb. the outcome did not need to be imputed), the prevalences were 4.0% and 12.1% respectively. If anything, this difference means that conducting the complete case analysis excludes a small set of healthy babies, as the prevalence decreases when all observations are included.

No significant differences are found after comparing crude and adjusted complete case models to crude and adjusted models after multiple imputation.

7.3.6.2 Missing data

Of the eligible set of BiB mothers who gave birth to live singleton babies whose birth weight was recorded (for the continuous birth weight models N=11,874, for the term LBW models N=11,211, for the SGA models N=11,863), the combined metric of exposure was the biggest limiting factor (48% missing). In addition to the exposure, the following covariates had missing values to be estimated by multiple imputation: caffeine intake, maternal BMI, smoking status, maternal education, ethnicity, gestational diabetes, and parity (ranging from 25% to 4% missing). Sex of child, maternal age at delivery and weeks of gestational age at birth were all complete variables. For details of the percentage of each covariate with missing values for each of the three outcomes, see Table 7.11, Table 7.12 and Table 7.13.

7.3.6.3 Assessing imputations

7.3.6.3.1 Comparing proportions in complete case, raw and imputed datasets

Multiple imputation brings the sample sizes from 5,040 to 11,874 for continuous birth weight model, from 4,782 to 11,211 for term LBW models, and from 5,034 to 11,863 for SGA models. As the gaps filled are non-negligible, I compared the proportions in each categorical variable, pre- and post-imputation, to assess the reliability of these imputations on an aggregate level. The proportions are all are similar (see Table A7 - 15 for a case study of the multiple imputations generated in the continuous birth weight model), but with proportions from the imputed dataset being closer to the proportions in

the raw data (i.e. without restrictions to exposure and all other covariates) than to the proportions in the complete case dataset.

Compared to the proportions in the raw data, the complete case analysis tends to underestimate the proportion of White British women, women who drink coffee, who are ex-smokers, who have further or higher education, with no previous children, and whilst overestimating the proportion of Pakistani women, less than 25 years old, who have never smoked, with 2 or more previous babies, and little if any formal education. The proportions imputed for the term LBW and SGA models are very similar to those of the continuous birth weight model (data not shown). Interpreting the results of the complete case analysis alone would give a skewed view of the reality in the full dataset, therefore the results of the multiple imputation models were interpreted.

7.3.6.3.2 Imputations among women with missing ethnicity information

Among 2069 eligible women who did not have ethnicity information in the original dataset, 38% were imputed to be White British, 48% Pakistani and 14% Other, compared to 40%, 44%, and 15% in the original raw data respectively (Table A7 - 16).

The output of the multiple imputation analysis confirms my findings in Chapter 6, that the women potentially at greatest risk of low birth weight are the same women who have missing (ethnicity and IMD) data (see Figure 6.9 b in Chapter 6). For example, 4.6% are term LBW among N=1,928 term births with missing ethnicity (compared to 4.1% in the complete case analysis), and 13.3% are SGA among N=2,069 with missing ethnicity data (compared to 12.9% in the complete case analysis). Because women with highest risk of poor birth outcomes are likely to be Pakistani in the original dataset (and to have gestational diabetes, lower formal education attainment, and 2 or more previous babies), when missing, Pakistani ethnicity is imputed at higher rates that the average in the original dataset. Mean birth weight among the 987 (of 2069) women imputed to be Pakistani is 3091g (vs. 3317g for White British and 3157g for Other); term LBW rate among these Pakistani women is 6.7% (White British 1.9%, Other 4.8%), and their SGA rate is 17.3% (WB 8.3%, Other 13.7%).

7.4 Discussion

In this chapter, the association between three metrics of exposure to HAAs and three measures of growth restriction were investigated. Though results are non-significant for the most part, there are a few trends. Few metrics show significant effects in these models, however many confounders did.

Increasing tertiles of tap water consumption predict a heavier weight at birth, and a higher probability of being >2500g at term after adjusting for potential confounders. Increasing tertiles of area-level

HAA concentrations (particularly TCAA during the first trimester exposure window) predict a lighter weight at birth, and possibly even a higher risk of being SGA. None of the adjusted models using combined metrics were significant, but the third tertile of exposure to the combined BDCAA metric predicts a higher probability of being of normal weight (>2500g) at term. Based on the output of the combined metric models with interaction terms for combined metrics and gestational age, the lowest tertile of exposure may be at slightly (non-significant) higher risk of lighter birth weight and of term LBW in the earliest respective gestational ages. The power in the term LBW models is very low, as the prevalence of low birth weight at term is only about 4%.

A number of sensitivity analyses were conducted to question the assumptions of these models, none of which substantially altered the conclusions. In addition, with the possible exception of a few differences for the continuous BW models, the crude and adjusted results are very similar.

7.4.1 Water consumption

Smith et al (2013) showed that individual water intake is influential in determining TCAA exposure variability in this cohort, which is why we thought it interesting to investigate the sole effect of water consumption on birth outcomes.

7.4.1.1 Mediation hypothesis

The assumption I made is that water consumption reflects an individual mother's (and by extension, baby's) exposure to HAAs. However, not only can water consumption contain and thus reflect exposure to a number of other water contaminants (e.g. other DBPs such as THMs, HANs, nitrosamines, yet unknown DBPs, or nitrates, arsenic), but it also constitutes an exposure in itself, to water, consumption of which is vital for survival. The finding that increasing cold tap water consumption predicts a heavier weight at birth and a higher probability of being >2500g at term is in line with previous studies which show that increased water consumption is associated with better birth outcomes (Aggazzotti et al. 2004; Savitz et al. 1995; JM Wright et al. 2010). Greater water consumption itself appears to have a protective effect on birth outcomes.

It is possible that the positive effects of tap water consumption on birth weight suggested by these results are compounded by other positive determinants of birth weight (e.g. maternal education which also predict increased water consumption) either directly (e.g. making water consumption a possible confounder of the relationship in this example between a maternal education and birth weight), indirectly via a mediation pathway (e.g. education is positively associated with tap water consumption which in turn leads to heavier weight at birth), or perhaps both.

Greater maternal education (as a proxy for higher socio-economic status), quintiles of IMD 2010 and maternal physical exercise (for which I selected to look at a categorical variable for number of hours spent doing physical exercise in the week preceding the baseline questionnaire[25]) all predict greater water consumption (JM Wright et al. 2010)[26]. All three of these are also strongly positively associated of birth weight in crude associations, exhibiting a dose-response association (see Table A7 - 17, step 1).

The mediation hypothesis was tested using Baron and Kenny's four step approach (1986), choosing continuous birth weight as the outcome and cold tap water as the water consumption "mediator" variable (as it was borderline significant in final adjusted models on continuous birth weight). In step 2, physical activity predicts a step change in cold tap water consumption, with women who reported exercising for ≥3 hours/week drinking on average 0.4L/day more than those who reported no exercise. Greater maternal education predicted a very slight increase in cold tap water consumption, while the pattern associated with IMD 2010 quintiles was unclear and non-significant, ruling the latter out as a possible mediator in this case (Table A7 - 17, step 2). In step 3, cold tap water consumption was significantly positively associated with birth weight, as per the main analyses (mean change in birth weight was 16.48g (95% CI: 2.14, 30.83)) (Table A7 - 17, step 3). Based on the output of step 4, I conclude that water consumption is a mediator of the association between maternal education and birth weight because all three first steps are significant and water consumption remains significant after controlling for maternal education (water consumption is in fact a partial mediator here because maternal education also remains significant after adjustment). The effect of physical activity on birth weight could not be established to be mediated by water consumption in this dataset (water consumption is not significant in step 4), however physical activity is consistently correlated with water consumption variables. Spearman's correlations between water consumption variables (cold tap water, total tap water, bottled water and total fluid), and maternal education, IMD 2010 quintiles, and physical activity are presented in the appendix (Table A7 - 18). The example of water consumption mediating the relationship between maternal education and birth weight opens up the possibility that it and other predictors of both water consumption and birth weight could account for the water consumption results.

[25] 74% of women did not do any physical activity in the week preceding enrolment to BiB, 14% did less than 1 hour per week, 9% between 1-3 hours, and only 2.4% report exercising for 3 or more hours a week (3539 of 11874 women do not have a reported physical activity information)

[26] Castano-Vinyals et al (2011) recently noted the opposite trend; however this study may not be comparable as it is based in Spain which has a very different climate and water source to the UK. (Castano-Vinyals G, Cantor KP, Villanueva CM, Tardon A, Garcia-Closas R, Serra C, et al. 2011. Socioeconomic status and exposure to disinfection by-products in drinking water in spain. Environ Health 10:18.)

I also find that the physical activity variable, which is strongly associated with birth weight on its own (see Table A7 - 17), is never a significant predictor of birth weight when added to the full final models (data not shown). No significant improvements to the final models are made by adding physical activity as an additional covariate to the models either. This suggests that water consumption may be a confounder of the association between physical activity and birth weight, though this hypothesis would need further investigation.

7.4.1.2 Unmeasured confounding

It cannot be excluded that levels of water consumption are associated with unmeasured confounding; for instance, per Wright (2010), a study found that participants reporting no water intake (nb. all those reporting zero total water intake were excluded from these analyses, but there remain individuals with as little as 0.20 L (i.e. 1 cup) reported total daily water intake) also reported increased soft drink consumption and less fruit, vegetable and low- and median-fat dairy product intake, all of which are indicative of generally unhealthier lifestyles (Popkin et al. 2005). This means that lower water consumption is associated with higher unhealthy lifestyle choices (in this case, poor diet), which could be the factors involved in lowering birth weight, not the water consumption itself.

In favour of the argument that water consumption (e.g. hydration) itself predicts better birth outcomes, I find a graduated response to water consumption: the more you drink during pregnancy, the heavier your baby at birth on average. If there were a greater positive effect on birth weight from consumption of bottled water (which is thought not to contain any HAAs, i.e. represents zero exposure) compared to the tap water types, the positive effects of greater water consumption could be verified to mask the smaller but nonetheless real negative effect of the HAAs contained in tap water. Unfortunately, there were not enough women with zero exposure (based on bottled water) in order to assess this effect difference. In addition, selecting for bottled water consumption itself may isolate a particular (perhaps richer) group of women, biasing the results.

Given these results, I conclude that it is possible that the effect of water consumption itself—or possibly of variables with strong positive effects on birth weight which also predict water consumption—drown out any hypothesized negative effects of the HAAs assumed to be contained in tap water. Other than HAA-containing tap water, no other sources of exposure from food have been taken into account.

7.4.2 Area-level concentrations

The area-level concentrations results indicate that increasing area-level HAA concentrations predict lighter weights at birth supporting the hypothesis.

In addition to being an area-level exposure metric which does not account for individual behaviour, the issue with area-level exposure metrics is that their derivation depends entirely on when and where a woman spends her time while pregnant. In reality, women clearly will not consume tap water only in the water supply zone of their residence. In addition, their place of residence may well change over the course of pregnancy.

To address this, the sophistication of the spatial component of this metric was increased by including both the work and residence information for those women for whom it was available. It is not that surprising that there are few differences in the results after adding work location information to the exposure measures (see Table A7 - 4, Table A7 - 7, and Table A7 - 12), given that the concentrations in the sensitivity analysis and in the original models scarcely differed. In fact, approx. 78% of BiB women had the same area-level HAA concentrations values in these two different analyses, either because they didn't report being employed outside of the home, they were employed but didn't report a valid workplace address that could be geocoded, or they reported a work address which fell outside of the catchment area of Bradford's eight water supply zones, meaning they could not be matched to modelled HAA concentrations.

As it is important to characterize exposure variability to assess the potential for exposure misclassification (Nieuwenhuijsen et al. 2009b), the modelling effort described in Chapter 4 attempted to model area-level concentrations both spatially and temporally. However, there is very little spatial variability in HAA exposures in Bradford. In addition, and as reported in previous studies, seasonality of HAAs is not well defined (Parvez et al. 2011); this was particularly true because it was only possible to model HAAs on a quarterly basis given available data. This means that when considering area-level concentrations alone (without a component for individual behaviour), variability in exposure is quite limited, such that there is not a lot of contrast between high and low exposed individuals, and that risk estimation is more difficult because of a higher probability of exposure misclassification (Nieuwenhuijsen et al. 2009b) (see distributions in Chapter 4, Figure 4.15). Studies from Scandinavia (Cedergren et al. 2002; Hwang et al. 2002; Kallen and Robert 2000; Magnus et al. 1999) and Taiwan (Hwang et al. 2008) have also shown low levels of DBPs with a similar small range (Nieuwenhuijsen et al. 2009b).

My Bayesian model developed, I could be including the uncertainty calculated in the derivation of time-weighted area-level estimates in the epidemiologic study. However, for reasons of time constraints with respect to the probable minimum added value of such an effort, I do not carry that uncertainty forward.

7.4.3 Combined exposure

In the models using the combined metric, either no effect was found, or the positive effect on birth weight suggested by individual tap water consumption and the negative trend of area-level concentrations on birth weight seem to cancel each other out. This is true for all models except perhaps for BDCAA on term LBW (brominated species have been reported to be more toxic than chlorinated ones (Plewa et al. 2010)), though the reported effect is in the opposite direction to the prediction.

The combined metric presented in this work combines the best available data sources and knowledge available to date to produce the most accurate exposure assessment possible, short perhaps of using biomarker data. However this combined metric also has some limitations. It likely integrates different types of misclassification error. As in Iszatt et al (2011), the exposure estimates are subject to random error from both classical error (i.e. error in sampling measurements of HAAs and quantity of water consumed by individual participants) and Berkson error (i.e. assigning modelled area-level HAA concentration to individual participants). Therefore, exposure misclassification may not simply attenuate risk (Armstrong 1998) but also increase the uncertainty in the estimates, as reflected by wide confidence intervals (Iszatt et al. 2011).

The reason only area-level concentrations for these three candidates (DCAA, TCAA and BDCAA) were modelled was the scarcity of HAA concentration data points collected. As a result, I was unable to assign individual women an exposure to the mixture of HAA5 (sum of MCAA, DCAA, TCAA, MBAA and DBAA) or HAA9 (sum of all 9 HAAs), or to the relative proportion of HAA5 to brominated HAAs (Parvez et al. 2011). However DCAA and TCAA are highly correlated with total HAAs (both DCAA and TCAA were correlated with $r=0.8$ ($p<0.001$) with HAA9, see Table A4 - 9), and when summed they are assumed to be good proxies for total HAAs in Bradford as they make up the great majority by concentration of the nine HAAs (see Chapter 4). I did run the combined models summing all three combined metrics derived in lieu of the exposure, but found no effect of this sum on birth outcomes (data not show). The assumption that all women are continuously exposed to the same level exposure throughout the time period at hand (whether considering a specific trimester of pregnancy in the case of area-level concentrations, or assuming that second trimester exposure is representative of the whole pregnancy in the case of the combined metric) may not be met. In addition, not all women were recruited at the exact same time in their pregnancy; rather women were recruited during a range of weeks of gestation (median difference between the date of questionnaire completion and the estimated date of conception is 184 days (26.3 weeks), with a 6 day interquartile range (N=9,775)), meaning that they reported behaviours in the baseline questionnaire at slightly different times of pregnancy. However controlling for days of gestation at the time of questionnaire completion did not alter the results (data not shown).

One of the biggest limitations of the combined metric exposure measures is the amount of filtering information which was not available. As described in Chapter 3 (Table A3 - 1), 9,071 women answered the question about whether or not they filtered their tap water at home (10% of which reported that they did), but only 1,466 employed women answered the question about whether or not they filtered their tap water at work (43% of which reported that they did). This corresponds to 76% of 11,874 eligible women for the filtering at home question, and only 34% of the 4,315 who were working for the filtering at work question. In all, 56% (or N=6,609) of the 11,874 total number of women who gave birth to live singleton babies with a birth weight record answered the filtering question for either (or both) of the two locations of interest. Knowing whether or not a woman filtered her tap water was essential to assigning her a combined DCAA, TCAA or BDCAA metric of exposure. If that filtering question was not answered, derivation of the combined metric of exposure could not be performed. As a result, 5,265 women were dropped from the combined metric calculation due to missing filtering information.

In addition, a greater proportion of Pakistani origin women over White British women answered the filtering questions, such that the complete cases analyses may be somewhat skewed by the higher proportion of their answers. A greater proportion of the women who answered the filtering question were of Pakistani origin (54%) proportionally to their share of the cohort (44%), and a smaller proportion of White British (32%) compared to their cohort share (40%) (The "Other" group was well represented in filtering question answer (14% and 15%, respectively)). That being said, amongst all women with ethnicity information who did respond to a filtering question (at either or both locations, N=6,619), the Pakistani women less frequently reported filtering their tap water (11% said yes) compared to the White British or Other groups (24% and 23%, respectively).

Counterbalancing the exclusion of a higher proportion of White British women from the combined metric models based on missing filtering information, proportionally more Pakistani women will have been excluded from the adjusted analyses than White British women, because of missing covariate information. This is one of the points made in Chapter 6 regarding representativeness of BiB and missing data among the most vulnerable.

Considering the results of the models after multiple imputation addressed successfully these imbalances. Indeed, as part of the imputation process, the partial information of those women excluded from the complete case analysis is used. Using multiple imputation to impute the missing confounders preserves the sample size. In this case though, the conclusions after imputation regarding the relationship between HAA exposure (in ug/day) and birth outcomes do not change.

The choice of filtering and boiling factors, based in this study solely on average factors sourced from the literature depended on literature bias, and water quality and type of filters used in those specific studies. In the future, further sensitivity analyses of the combined metric analysis with different (or without) filtering and boiling factors, and with imputation of missing filtering information may be warranted.

7.4.4 Effect of HAAs when comparing the results of the cold tap water and combined metric models on continuous birth weight

The effects of cold tap water consumption on continuous birth weight are significant in the adjusted linear regression models: women in tertile 3 of exposure (drinking >1.4 L/day) had significantly heavier babies (24.0grams (CI: 1.4, 46.6)) compared to women in the reference group.

As above, this result suggests that increased maternal cold tap water consumption—or the variables positively associated with water consumption such as socio-economic status, education, level of physical exercise, ethnicity or smoking and caffeine habits—is associated with heavier babies at birth, and as such could constitute a healthy or even protective behaviour in terms of birth weight.

However, after combining the volumes of cold and hot tap water reported to be consumed by each cohort woman (L/day) with the modelled HAA concentrations (ug/L) into a combined metric for exposure to DCAA, TCAA and BDCAA (ug/day) and taking boiling and filtering into account, the effect of HAAs on continuous birth weight in adjusted models disappeared for all three HAAs (seeTable 7.2).

The difference between the two exposure measures (cold tap water consumption and the combined metrics) being the HAAs themselves, this result could mean that accounting for HAAs—the very chemicals of interest in this thesis—leads to a reduction to the null, or cancelling out, of any positive effect that water consumption has on birth weight, irrespective of quantity of water consumed.

While there is no intake of HAAs without water consumption—the exposures, and thus their effects, are inextricably paired—, one could imagine that higher concentrations of HAAs may have a more drastic effect counter the positive effect of water consumption (especially after boiling is taken into consideration for DCAA, as boiling increases DCAA's concentrations).

However, contrary to the significant positive effect of cold tap water on birth weight, there are no significant effects of total tap water (or bottled water, or total water) consumption on continuous birth weight (see Table 7.3), which is the consumption measure included in the combined metric, with the

appropriate boiling and filtering reduction/increase factors applied to it. In addition one must question the clinical significance of a 24 gram increase in birth weight, when the mean birth weight in this population is 3,231.1grams (95% CI: 3219.8, 3239.6) (N=11,874), i.e. 0.7% change in average birth weight.

In the term LBW models, the protective significant effects observed for cold tap water, total water and total water does carry over to the results of the combined DCAA, TCAA and BDCAA metrics or even grow stronger (see Table 7.5 and Table 7.6). This supports the theory that it is total tap water, not cold tap water, that is the primary consumption measure of interest. This result, though on an outcome which exclude all preterm births many of which also be growth restricted, highlights that the reversal of effects observed for the continuous birth weight models may be a chance finding or indeed due to other factors, perhaps related to preterm status.

7.4.5 Stratification by ethnicity: White British compared to Pakistani women

Though not significant, upon stratification of the models using the combined measure as exposure metrics, a trend towards a reduction in birth weight is observed in the White British women for DCAA and TCAA models, and a trend in the opposite direction for Pakistani women for DCAA, TCAA and BDCAA models (Table A7 - 10). If the direction of the effect of HAAs is different between the two major ethnic groups represented in BiB, then this may explain why non-stratified analyses will not reveal any effects of HAA on birth outcomes.

These results are similar—though less accentuated—to similar models also set in BiB but looking at air pollutants (fine particulate matter of diameter 2.5um or less, i.e. PM2.5, in particular). This work was conducted by Anna Schembari and Mark Nieuwenhuijsen, at CREAL in Barcelona, Spain (unpublished to date, personal communication). They report a significant interaction between exposure and ethnicity such that increasing PM2.5 during the third trimester of pregnancy has a statistically significant negative effect on birth weight in White British origin babies, but not in the Pakistani origin babies. If these results were reproduced in another population, it may suggest that the two predominant ethnic groups react differently to the same environment. This may be explained by a) genetic factors, as highlighted in Chapter 6, such that babies of Pakistani and South Asian origin in general tend to have lower birth weights than White babies (in this sample, 6% of Pakistani are term LBW compared to 2% among White British and 3% among Other groups, and 16% of Pakistani are SGA compared to 8% and 12% among White British and Other, respectively), and/or b) behavioural factors such as the time spent at home, cooking habits and diet, in addition to the smoking, caffeine intake and deprivation factors adjusted for (see Chapter 6 for details, as well as Table A7 - 9, Table A7 - 11 and Table A7 - 14). In fact, some of these characteristics (such as smoking or caffeine intake,

which predict lower birth weight but only apply to the White British women who already tend to have heavier babies on average) may mask even larger differences between ethnicities than the differences reported in this study.

7.4.5.1 Using HAAs and birth weight as proxies for other variables

As discussed above, the individual DCAA, TCAA (and BDCAA) exposures modelled, derived and studied in this thesis are good proxies for exposure to total HAAs in Bradford (Table A4 - 9). The lack of correlation between these HAAs and other DBP species in Bradford as highlighted in Chapter 4 make it quite unlikely however that HAAs are a good proxy for other DBPs. Beyond this, it is reasonable to question how reliable birth weight alone is as a marker of ill health (see Introduction (Chapter 1)). Indeed a study in the BiB cohort aimed at describing the growth pattern from birth to age two years in a subset of 1,434 UK-born White British and Pakistani infants found that Pakistani infants were lighter and had shorter predicted mean length at birth than White British infants, but gained weight and length quicker in infancy. By age 2, both ethnic groups had similar weight, but Pakistani infants were on average taller than White British infants. In other words, weight at birth is just one endpoint; upon follow-up of a subset of this very same cohort, differences in growth trajectories appear by ethnicity (Fairley et al. 2013).

West's study finds that while markedly lighter overall, Pakistani infants in the BiB cohort had similar skinfold thicknesses and greater total fat mass (as indicated by cord leptin) for a given birth weight than White British infants, indicating that Pakistani babies were more adipose than White British babies. If the birth weight endpoint is considered alone and out of context, this study's authors conclude that any efforts to reduce ethnic inequalities in birth weight need to consider differences in adiposity and the possibility that increasing birth weight in South Asian infants might inadvertently worsen health by increasing relative adiposity (West et al. 2013).

7.4.6 Stratification to unemployed women only

Restricting the analyses to unemployed women (N=5,519) could increase the precision of the risk estimates; however, there were no difference in the results or overall conclusion after stratification. Though the estimates may be more accurate, the confidence intervals are wider in the stratified group, as the sample size is split in half.

There are two reasons why unemployed women may have less error in their exposure estimates than employed women: a) using the urinary TCAA biomarker as a gold standard of exposure, TCAA ingestion in the home as recorded by the BiB questionnaire was found to be a valid proxy for TCAA exposure for unemployed women, but less so for employed women in the BiB cohort (Smith et al. 2013) and b) even though I tried to account for the work water supply zone location of the employed

women in sensitivity analyses presented earlier, the data to do this were scarce meaning that the area-level HAA estimates among the unemployed women (for whom there is only one water supply zone to account for) are likely to contain less error. This is based on the underlying assumption that if a woman is employed she spends her work time at work only and her free time at home only, while if a woman is unemployed she spends all of her time at home. It is not straightforward to predict how these errors will affect risk estimates given the two types of error (classical and Berkson) at play, but I expect estimates among the unemployed to represent better estimates of risk as there is generally less exposure measurement error. This finding also highlights the existence of recall error associated with self-reported information via questionnaire, and future work would benefit from validating this.

7.4.7 Gestational age

Excluding gestational age form the final models does not alter the results, though it strongly affects the model fit. It is to be noted that fetal growth is not linear and further work exponentiating or categorising the continuous gestational age variable may improve the fit. It is unclear however that any less rigidly linear inclusion of gestational age in the models would have altered the conclusions of the association of interest. In addition, if an environmental exposure were to affect growth in early pregnancy, gestational age (as measured by ultrasound) would be underestimated. This means for example that a baby thought to be born at 38 weeks gestation at a healthy weight of 3100g may in fact be underweight if he was really born at 40 weeks of gestational age (mean birth weight at 40 weeks gestation is 3343g in the BiB cohort). Ultimately, along with possible non-linear growth during pregnancy, and the possible errors in measurement of gestational age itself (further discussed in Chapter 6), it is difficult to assess whether gestational age is a mediator of the association between exposure and outcome because that exposure on outcome relationship is overall non-significant.

7.4.8 Model selection process

After trying several different methods to select best fit models, an *a priori* selection of variables which are both important predictors of the exposure and/or birth outcomes and of relatively good quality, were used.

7.4.8.1 Discussion of coefficients

All covariates coefficients but maternal age were in the expected direction (Table A7 - 9, Table A7 - 11 and Table A7 - 14): when holding all other variables constant in the models (whether the continuous birth weight, term LBW or SGA models), female babies were lighter than males at birth; parity and quartiles of BMI had graduated positive effects the higher their values; the higher the mother's educational attainment the heavier her baby at birth; report of > 200mg/day caffeine intake

predicted lower birth weights while report of gestational diabetes predicted higher birth weight, as expected. Current smokers had lower weight babies.

The lightest babies were to 25 to 29 year old mothers, with increases in birth weight (or decreases in risk of term LBW or SGA) for both the younger and older groups of mothers. This is slightly unusual, however maternal age was only significant in the continuous birth weight models not in the term LBW or SGA models; it was not included *a priori* for its contribution to the model, but because it is customary to do so.

Models with most significant covariate effects were the continuous birth weight models, most likely because they are better powered models. Nevertheless, all covariates remain intercorrelated (Table A7 - 2), making differentiation between the variables which drive the associations difficult.

The robustness of the caffeine report as associated with lower birth or increased risk of term LBW or SGA was quite surprising. However, such an association may reflect known as the Stein-Susser epiphenomenon: if nausea is a marker of good implantation as has been suggested (perhaps reflecting a favourable balance of hormones produced by a healthy placenta) (Leviton and Cowan 2002) and women with prominent nausea tend to reduce caffeine consumption, then those who continue drinking coffee may be the subgroup of women at higher risk of pregnancy complication for reasons unrelated to coffee consumption. This may explain the strong association observed here, even if it is not a true effect of the caffeine itself.

There is a limitation in the methods relating to the calculation of caffeine consumption. Indeed, this variable is derived from the food section of the baseline questionnaire. Subsequently in the water section, women are asked about their daily coffee consumption during pregnancy, the answer to which is added to the "hot tap water" consumption, then to "total tap water" consumption and in turn to "total water" consumption summary categories, thus potentially including caffeine consumption both in the exposure as well as adjusting for it in the models. I don't suspect any great impact on the association between exposure to HAAs and birth weight, but this means that interpreting the caffeine intake covariate must be done with caution. Note that the effect of hot tap water on birth outcomes alone was not examined, in part for this reason.

7.4.8.2 Adjusting for a large number of covariates

Choosing which variables to include as confounders in a regression model is a delicate issue, and one for which there is no simple or universally established set of "rules". The main statistical trade-off is that failing to include important confounders may lead to biased exposure estimates, whereas

including large numbers of covariates in a regression model can lead to loss of efficiency (inflated variances). However, including 8 to 10 variables is still only a modest number of covariates.

7.4.8.3 Covariates not included

Other than the variables which were considered for inclusion in the models, but then decided not to for reasons discussed previously[27], the following are possible confounders reported in the literature to have links with birth outcomes and which could not be included in the models because they were unavailable: maternal diet and folate/iron supplementation prior to pregnancy; paan chewing; nicotine replacement; intake of illegal drugs; pregnancy weight gain; HIV/AIDS status and other infections acquired during pregnancy; heart, kidney or lung problems; certain medications taken during pregnancy (e.g. acne medications, or medications to delay onset of labour); birth spacing; paternal factors (such as height, weight, and ethnicity). Assisted reproductive technology data which has been shown to contribute to an increase in preterm deliveries was not available either (Wilcox 2010).

In addition, the effects of other environmental exposures (such as trihalomethanes (THMs), or air pollutants) for which data are already available in this cohort were not included. I suggest that the joint effect of exposure to the HAAs presented in this thesis, the trihalomethanes (THMs) and the air pollutant data which are now available in BiB be examined at a later stage, as this would constitute a very interesting study.

7.4.8.3.1 Alcohol

I considered including alcohol in the models, however the results did not materially change after its inclusion. In addition, the quality of the questionnaire-derived variable was quite poor, heavily subject to recall bias, with approximately 20% missing data, and while I was in fact mostly interested in the period of pregnancy itself, it includes alcohol consumption during pregnancy as well as during the three months preceding pregnancy. In this cohort, the alcohol variable is also highly correlated with smoking ($r=0.33$, $p<0.001$, $N=9,517$), ethnicity ($r=-0.37$, $p<0.001$, $N=9,489$) and caffeine intake ($r=0.21$, $p<0.001$, $N=8,682$), reducing the added benefit of incorporating it and leading to potentially highly collinear models (Table A7 - 19). In addition, the evidence regarding alcohol consumption during pregnancy, particularly the effects of moderate quantities of alcohol consumption at various stages of pregnancy, is not well established (Oster 2013). In the BiB cohort itself only binge drinking was found to affect SGA with no association found between any level of alcohol consumption and premature birth (Cooper et al. 2013).

[27] Hypertension, pre-existing diabetes, employment status, physical activity of the mother, alcohol consumption, having multiple births, and level of deprivation/income

7.4.8.3.2 Stress

Stress is a very important factor which has not been considered in this analysis, and would likely play an important role given BiB's study population. Maternal stress (of various forms and definitions) has been linked with preterm delivery. For instance, in a prospective cohort study of 1,962 pregnant women in central North Carolina between 1996 and 2000, in which 12% delivered preterm, Dole et al. (2003) reported an increased risk of preterm birth among women with high counts of pregnancy-related anxiety, with life events to which the respondent assigned a negative impact weight and with a perception of racial discrimination. The mechanism of action of maternal and fetal stress could involve activation of cells in the placenta, decidua, and fetal membranes to produce corticotropin-releasing hormone, which in turn enhances prostaglandin production in these tissues to promote parturition (Lockwood 1999).

Dejin-Karlsson et al (2000) find that a lack of psychosocial resources, such as social stability, social participation, emotional and instrumental support, all increased the likelihood of delivering an infant that was SGA. Simultaneous exposure to what was defined as a poor total network index, as well as a poor total support index showed a significantly increased odds ratio for having a small-for-gestational age baby. In addition, the authors reported an interaction between immigrant status and poor total network or poor total support, in a synergistic direction, in their sample (Dejin-Karlsson et al. 2000).

Last but not least, work by Traviss et al (2012) in the BiB cohort reveals that maternal mental health in pregnancy has an independent influence on infant growth up to six months and is associated with ethnicity which is itself associated with deprivation. Taking example from Traviss et al (2012) recent study, further exploration of the complex relationship in the BiB population between symptoms of maternal distress, ethnicity, deprivation, health behaviours, environmental exposures and early infant growth, would make a very interesting area for future research and may answer some of the remaining questions pertaining to the ethnic disparities observed in BiB. Stressors are extremely hard to quantify. Their perception may vary highly between individuals, while biomarkers of stress (such as glucocorticoid hormones in saliva) have very short half-lives and are highly variable depending on the timing and circumstances of measurement. However sources of stress themselves can be quantified.

7.4.9 Conclusion

Mean birth weights in each tertile of exposure are very similar, as are the rates of tLBW or SGA by exposure tertile (for term LBW see Figure A7 - 2 and Figure A7 - 3; for SGA see Figure A7 - 4 and Figure A7 - 5) (Table 7.1). Also there is a high level of variability in birth weight (SD=545.0g). The study has low power to detect the differences observed in the birth weight (and probabilities of tLBW or SGA) between the tertile of exposure to combined DCAA, TCAA or BDCAA.

I find nearly no significant results in this chapter. If there are any, they are opposite effects to the hypothesis for water consumption models suggesting that either there really is no effect of the HAAs contained in tap water on birth outcomes, or that the benefits of tap water consumption itself and the healthy behaviours associated with water consumption (e.g. higher socio-economic status, greater physical health via physical exercise, healthier lifestyle choices in general) outweigh any possible negative effects of the HAAs it contains, rendering these undetectable. The trend for area-level concentration results fits with the hypothesis that greater exposure to HAAs predicts poorer birth outcomes, particularly for TCAA, though the results are mostly not significant, and the exposure range is quite limited. As for the models using the metric combining individual water consumption and area-level concentrations which combines the best available data sources that can be derived to date (short perhaps of biomarker data, which come with their own unique set of challenges), the results are not significant. While the study power is quite low given the very small expected effects, these results may be not significant because of the opposite direction of the effect of its components (water consumption, and area-level HAA concentrations), or because HAAs do not affect the birth outcomes under consideration in this cohort.

7.5 Tables

Table 7.1: Descriptive data of birth weight (in grams), and prevalence of LBW, term LBW, and SGA (UK 1990), by tertiles of exposure for each of 16 exposure measures of interest (6 water consumption variables, 9 trimester-specific area-level concentration variables, and 3 combined metrics) (live births only) CI: confidence interval

	Exposure mean (95% CI)	N.	**birth weight (g)** mean (95% CI)	**LBW (rate)** prevalence (%) (95% CI)	N.	**term LBW (rate)** prevalence (%) (95% CI)	N.	**SGA (rate)** prevalence (%) (95% CI)	N.
water consumption (L/day)									
cold tap water									
0 - 0.8	0.57 (0.56, 0.57)	3506	3227.8 (3209.6, 3246.0)	7.9 (7.1, 8.9)	3506	4.4 (3.7, 5.1)	3307	12.0 (10.9, 13.1)	3503
> 0.8 - 1.4	1.14 (1.14, 1.15)	2826	3229.9 (3209.8, 3250.1)	6.9 (6.0, 7.9)	2826	3.7 (3.0, 4.5)	2682	13.1 (11.9, 14.4)	2823
> 1.4	2.06 (2.03, 2.09)	2861	3251.7 (3231.4, 3271.9)	6.8 (5.9, 7.8)	2860	3.5 (2.8, 4.3)	2712	10.8 (9.7, 12.0)	2857
total tap water									
0 - 1.2	0.88 (0.87, 0.89)	3811	3223.7 (3206.3, 3241.1)	8.0 (7.2, 8.9)	3811	4.5 (3.8, 5.2)	3593	12.4 (11.4, 13.5)	3806
> 1.2 - 1.8	1.58 (1.57, 1.59)	2905	3232.5 (3212.8, 3252.2)	6.7 (5.8, 7.6)	2905	3.5 (2.9, 4.3)	2767	12.4 (11.2, 13.7)	2903
> 1.8	2.69 (2.66, 2.73)	2981	3259.2 (3239.1, 3279.4)	7.1 (6.2, 8.1)	2980	3.6 (3.0, 4.4)	2816	10.9 (9.8, 12.1)	2977
bottled water									
0 - 0.4	0.25 (0.24, 0.26)	1608	3252.5 (3225.3, 3279.7)	7.5 (6.2, 8.9)	1608	3.7 (2.8, 4.8)	1521	12.1 (10.6, 13.8)	1607
> 0.4 - 0.8	0.70 (0.69, 0.71)	993	3288.1 (3254.2, 3322.1)	6.1 (4.7, 7.8)	992	3.6 (2.5, 5.0)	942	10.3 (8.5, 12.4)	990
> 0.8	1.52 (1.48, 1.55)	1105	3316.9 (3283.5, 3350.3)	6.4 (5.1, 8.0)	1105	3.3 (2.3, 4.6)	1048	9.1 (7.5, 11.0)	1104
total fluid									
0 - 1.4	1.07 (1.06, 1.08)	3743	3204.1 (3186.9, 3221.2)	8.1 (7.2, 9.0)	3743	4.7 (4.0, 5.4)	3547	13.4 (12.3, 14.5)	3737
> 1.4 - 2.0	1.86 (1.86, 1.87)	3405	3245.3 (3226.9, 3263.7)	6.8 (6.0, 7.7)	3405	3.5 (2.9, 4.2)	3218	11.4 (10.4, 12.5)	3403
> 2.0	3.19 (3.15, 3.23)	2638	3277.7 (3255.9, 3299.5)	6.9 (5.9, 7.9)	2637	3.4 (2.7, 4.2)	2497	10.4 (9.3, 11.6)	2634
area-level concentrations (ug/L)									
Average [DCAA]									
trimester 1									
0-7.12	5.69 (5.65, 5.73)	3498	3241.2 (3222.9, 3259.4)	7.4 (6.6, 8.4)	3497	3.8 (3.2, 4.5)	3278	11.4 (10.4, 12.6)	3494
>7.12 - 9.36	8.40 (8.38, 8.42)	3485	3223.0 (3204.8, 3241.2)	7.6 (6.7, 8.5)	3485	4.1 (3.5, 4.9)	3299	12.3 (11.2, 13.4)	3481
>9.36	11.82 (11.74, 11.89)	3494	3221.3 (3203.1, 3239.5)	7.4 (6.6, 8.4)	3494	4.0 (3.4, 4.7)	3310	12.6 (11.5, 13.7)	3492

	Exposure mean (95% CI)	N.	**birth weight (g)** mean (95% CI)	**LBW (rate)** prevalence (%) (95% CI)	N.	**term LBW (rate)** prevalence (%) (95% CI)	N.	**SGA (rate)** prevalence (%) (95% CI)	N.
trimester 2									
0-7.20	5.76 (5.72, 5.79)	3758	3252.7 (3234.8, 3270.5)	7.5 (6.7, 8.4)	3758	3.9 (3.3, 4.6)	3538	11.4 (10.4, 12.4)	3757
>7.20 - 9.43	8.53 (8.51, 8.55)	3749	3211.7 (3194.3, 3229.1)	7.6 (6.8, 8.5)	3748	4.2 (3.6, 5.0)	3541	12.8 (11.8, 14.0)	3744
>9.43	11.76 (11.69, 11.83)	3755	3224.0 (3206.7, 3241.3)	7.4 (6.6, 8.3)	3755	4.0 (3.4, 4.7)	3557	12.5 (11.5, 13.6)	3750
trimester 3									
0-7.43	5.87 (5.83, 5.91)	3843	3261.4 (3243.8, 3278.9)	7.1 (6.3, 8.0)	3843	3.4 (2.9, 4.1)	3618	10.6 (9.6, 11.6)	3841
>7.43 - 9.54	8.70 (8.69, 8.72)	3847	3218.4 (3201.1, 3235.6)	7.7 (6.9, 8.6)	3847	4.4 (3.7, 5.1)	3650	12.9 (11.9, 14.0)	3843
>9.54	12.01 (11.94, 12.08)	3844	3205.8 (3188.7, 3223.0)	7.9 (7.1, 8.8)	3843	4.4 (3.8, 5.1)	3617	12.9 (11.9, 14.1)	3838
Average [TCAA]									
trimester 1									
0-10.38	8.59 (8.54, 8.65)	3503	3233.1 (3214.6, 3251.5)	7.6 (6.8, 8.6)	3503	4.1 (3.5, 4.9)	3303	12.0 (11.0, 13.2)	3500
>10.38 - 12.49	11.38 (11.36, 11.40)	3485	3223.5 (3205.2, 3241.8)	7.7 (6.9, 8.7)	3485	4.1 (3.5, 4.9)	3282	11.7 (10.7, 12.8)	3481
>12.49	15.97 (15.90, 16.04)	3489	3228.9 (3211.0, 3246.9)	7.1 (6.3, 8.0)	3488	3.7 (3.1, 4.4)	3302	12.6 (11.5, 13.7)	3486
trimester 2									
0-10.49	8.73 (8.68, 8.79)	3758	3222.4 (3204.3, 3240.4)	8.4 (7.5, 9.3)	3758	4.4 (3.7, 5.1)	3531	12.7 (11.7, 13.8)	3752
>10.49 - 13.05	11.58 (11.56, 11.60)	3748	3242.2 (3224.5, 3259.8)	7.3 (6.4, 8.1)	3748	4.0 (3.3, 4.7)	3540	11.8 (10.8, 12.9)	3747
>13.05	16.48 (16.42, 16.54)	3756	3223.9 (3207.0, 3240.9)	6.8 (6.1, 7.7)	3755	3.8 (3.2, 4.5)	3565	12.2 (11.2, 13.3)	3752
trimester 3									
0-10.62	9.04 (8.98, 9.09)	3841	3214.8 (3197.2, 3232.4)	8.6 (7.7, 9.5)	3841	4.6 (3.9, 5.3)	3618	12.6 (11.6, 13.7)	3836
>10.62 - 13.17	11.72 (11.69, 11.74)	3847	3234.0 (3216.4, 3251.5)	7.3 (6.5, 8.1)	3847	3.7 (3.1, 4.4)	3611	12.4 (11.4, 13.5)	3844
>13.17	16.48 (16.42, 16.54)	3846	3236.8 (3220.0, 3253.6)	6.9 (6.1, 7.7)	3845	3.9 (3.3, 4.6)	3656	11.4 (10.4, 12.4)	3842
Average [BDCAA]									
trimester 1									
0-0.90	0.71 (0.70, 0.71)	3493	3236.4 (3218.3, 3254.5)	7.3 (6.5, 8.2)	3492	4.0 (3.3, 4.7)	3293	11.6 (10.6, 12.7)	3490
>0.90 - 1.38	1.16 (1.15, 1.16)	3492	3229.4 (3211.2, 3247.5)	7.2 (6.4, 8.2)	3492	3.8 (3.2, 4.6)	3299	12.1 (11.0, 13.2)	3488
>1.38	2.09 (2.07, 2.11)	3492	3219.7 (3201.2, 3238.2)	7.9 (7.0, 8.8)	3492	4.1 (3.5, 4.9)	3295	12.6 (11.5, 13.8)	3489
trimester 2									
0-0.91	0.71 (0.70, 0.71)	3757	3244.0 (3226.2, 3261.7)	7.4 (6.6, 8.3)	3757	4.0 (3.4, 4.7)	3545	11.4 (10.4, 12.4)	3754
>0.91 - 1.38	1.16 (1.15, 1.16)	3759	3217.0 (3199.6, 3234.4)	7.6 (6.7, 8.4)	3759	3.9 (3.3, 4.6)	3536	12.7 (11.7, 13.8)	3754
>1.38	2.09 (2.07, 2.11)	3746	3227.4 (3210.0, 3244.8)	7.6 (6.7, 8.5)	3745	4.2 (3.6, 4.9)	3555	12.6 (11.6, 13.7)	3743
trimester 3									
0-0.91	0.71 (0.70, 0.71)	3843	3224.6 (3207.0, 3242.2)	8.1 (7.2, 9.0)	3843	4.0 (3.4, 4.7)	3597	11.9 (10.9, 12.9)	3839

	Exposure mean (95% CI)	N.	**birth weight (g)** mean (95% CI)	**LBW (rate)** prevalence (%) (95% CI)	N.	**term LBW (rate)** prevalence (%) (95% CI)	N.	**SGA (rate)** prevalence (%) (95% CI)	N.
>0.91 - 1.39	1.16 (1.15, 1.16)	3849	3241.7 (3224.7, 3258.8)	7.0 (6.2, 7.8)	3849	3.9 (3.3, 4.6)	3659	12.0 (11.0, 13.1)	3845
>1.39	2.13 (2.11, 2.15)	3842	3219.2 (3201.8, 3236.5)	7.7 (6.9, 8.6)	3841	4.3 (3.7, 5.0)	3629	12.5 (11.5, 13.6)	3838
combined metric (ug/day)									
DCAA exposure									
0 - 10.41	7.07 (6.96, 7.17)	2067	3200.7 (3176.8, 3224.6)	8.6 (7.4, 9.9)	2067	4.7 (3.8, 5.8)	1948	13.4 (12.0, 15.0)	2065
> 10.41 - 17.29	13.71 (13.63, 13.80)	2063	3223.9 (3201.2, 3246.6)	6.6 (5.6, 7.8)	2063	3.6 (2.8, 4.5)	1963	13.1 (11.7, 14.7)	2063
> 17.29	27.32 (26.85, 27.79)	2065	3223.9 (3200.0, 3247.8)	7.1 (6.0, 8.3)	2065	3.7 (2.9, 4.6)	1958	12.0 (10.6, 13.5)	2061
TCAA exposure									
0 - 12.02	8.01 (7.89, 8.13)	2066	3204.8 (3181.1, 3228.6)	8.6 (7.4, 9.9)	2066	4.8 (3.9, 5.9)	1950	13.6 (12.1, 15.1)	2064
> 12.02 - 19.79	15.69 (15.59, 15.78)	2062	3219.7 (3196.5, 3242.8)	6.8 (5.7, 8.0)	2062	3.5 (2.8, 4.5)	1953	12.3 (10.9, 13.8)	2062
> 19.79	30.77 (30.25, 31.29)	2067	3224.0 (3200.3, 3247.7)	7.0 (5.9, 8.2)	2067	3.7 (2.9, 4.6)	1966	12.7 (11.2, 14.2)	2063
BDCAA exposure									
0 - 1.04	0.65 (0.63, 0.66)	2066	3203.0 (3179.5, 3226.4)	8.3 (7.1, 9.5)	2066	4.7 (3.8, 5.7)	1952	12.7 (11.3, 14.2)	2066
> 1.04 - 1.93	1.45 (1.44, 1.46)	2066	3211.4 (3187.6, 3235.1)	7.7 (6.6, 8.9)	2066	4.1 (3.3, 5.1)	1952	13.6 (12.2, 15.2)	2063
> 1.93	3.30 (3.23, 3.36)	2063	3234.2 (3210.8, 3257.5)	6.3 (5.3, 7.5)	2063	3.3 (2.5, 4.1)	1965	12.2 (10.8, 13.7)	2060

Table 7.2: Complete case analysis (CC) and Results after Multiple Imputation (MI): Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (in ug/day) and continuous birth weight (in grams) by linear regression (Complete case analysis, N=5,040; After Multiple Imputation using Chained Equations, N=11,874)
Sample restricted to all live singleton births to eligible mothers.

Combined exposure (ug/day)	**CC**			**MI**		
	n	**Crude mean change in BW (g)[#]**	**Adjusted* mean change in BW (g)**	**n[1]**	**Crude mean change in BW (g)[#]**	**Adjusted* mean change in BW (g)**
DCAA exposure						
0 - 10.41	1563	0.0	0.0	3851	0.0	0.0
> 10.41 - 17.29	1678	17.9 (-19.3, 55.2)	-0.2 (-28.6, 28.2)	3905	24.9 (-9.0, 58.8)	5.1 (-19.1, 29.3)
> 17.29	1799	20.0 (-16.6, 56.7)	7.7 (-21.2, 36.5)	4118	31.4 (-5.9, 68.6)	5.7 (-20.7, 32.1)
TCAA exposure						
0 - 12.02	1590	0.0	0.0	3849	0.0	0.0
> 12.02 - 19.79	1719	10.0 (-26.8, 46.9)	3.3 (-24.8, 31.5)	3925	21.8 (-12.4, 56.1)	4.3 (-20.5, 29.1)
> 19.79	1731	12.9 (-23.9, 49.7)	-0.7 (-29.1, 27.8)	4100	29.9 (-9.1, 68.8)	-1.8 (-28.2, 24.6)
BDCAA exposure						
0 - 1.04	1737	0.0	0.0	3914	0.0	0.0
> 1.04 - 1.93	1652	4.0 (-32.4, 40.4)	-4.8 (-32.5, 23.0)	3936	18.1 (-17.9, 54.0)	3.1 (-21.3, 27.5)
> 1.93	1651	33.8 (-2.6, 70.2)	11.4 (-16.5, 39.2)	4024	37.9 (-0.5, 76.2)	8.8 (-17.7, 35.4)

[#]mean change in birth weight (95% confidence intervals)
[1] based on average proportion in each category over 10 imputations
*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table 7.3: Complete case analysis: Crude and adjusted association between water consumption (four different water types, in L/day) and continuous birth weight (in grams) by linear regression
Sample restricted to all live singleton births to eligible mothers

Water consumption (L/day)		n	Crude mean change in BW (g)[♯]	Adjusted* mean change in BW (g)
cold tap water				
	0 - 0.8	2858	0.0	0.0
	> 0.8 - 1.4	2260	2.2 (-27.8, 32.3)	19.9 (-3.0, 42.8)
	> 1.4	2308	23.9 (-6.0, 53.9)	24.0 (1.4, 46.6)
total tap water				
	0 - 1.2	2993	0.0	0.0
	> 1.2 - 1.8	2387	-0.4 (-29.8, 29.0)	1.0 (-21.2, 23.2)
	> 1.8	2491	34.3 (5.3, 63.4)	16.3 (-6.4, 39.1)
bottled water				
	0 - 0.4	1195	0.0	0.0
	> 0.4 - 0.8	798	7.7 (-41.8, 57.1)	-2.3 (-40.0, 35.4)
	> 0.8	872	55.1 (6.9, 103.2)	16.9 (-20.7, 54.5)
total water				
	0 - 1.4	2947	0.0	0.0
	> 1.4 - 2.0	2819	31.1 (2.9, 59.3)	10.9 (-10.6, 32.4)
	> 2.0	2169	72.6 (42.3, 102.9)	15.8 (-8.4, 40.0)

[♯]mean change in birth weight (95% confidence intervals)
*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes
(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table 7.4: Complete case analysis: Crude and adjusted association by linear regression between average modelled area-level concentrations of DCAA, TCAA and BDCAA (based on residence water supply zone) (in ug/L) and continuous birth weight (in grams), by trimester of pregnancy
Sample restricted to all live singleton births to eligible mothers

Area-level concentrations (in ug/L)	**n**	**Crude mean change in BW (g)[♯]**	**Adjusted* mean change in BW (g)**
Average [DCAA]			
trimester 1			
0-7.12	2389	0.0	0.0
>7.12 - 9.36	2427	-40.9 (-71.6, -10.1)	5.8 (-17.5, 29.2)
>9.36	2338	-49.1 (-80.1, -18.1)	-0.4 (-24.0, 23.2)
trimester 2			
0-7.20	2426	0.0	0.0
>7.20 - 9.43	2547	-48.6 (-78.9, -18.4)	-22.0 (-45.0, 1.0)
>9.43	2612	-46.0 (-76.1, -15.9)	-14.4 (-37.3, 8.5)
trimester 3			
0-7.43	2519	0.0	0.0
>7.43 - 9.54	2579	-52.2 (-82.1, -22.2)	-16.5 (-39.2, 6.1)
>9.54	2613	-54.8 (-84.7, -25.0)	-17.4 (-40.0, 5.2)
Average [TCAA]			
trimester 1			
0-10.38	2281	0.0	0.0
>10.38 - 12.49	2534	-15.6 (-46.4, 15.2)	-15.3 (-38.6, 8.0)
>12.49	2339	-21.1 (-52.5, 10.3)	-28.3 (-52.0, -4.6)
trimester 2			
0-10.49	2359	0.0	0.0
>10.49 - 13.05	2686	16.5 (-13.6, 46.6)	2.6 (-20.1, 25.3)
>13.05	2540	-6.0 (-36.5, 24.5)	-6.7 (-29.7, 16.3)
trimester 3			
0-10.62	2376	0.0	0.0
>10.62 - 13.17	2678	13.6 (-16.6, 43.7)	10.1 (-12.6, 32.8)
>13.17	2657	8.0 (-22.2, 38.2)	10.5 (-12.2, 33.2)
Average [BDCAA]			
trimester 1			
0-0.90	2625	0.0	0.0
>0.90 - 1.38	2404	1.1 (-29.0, 31.2)	-0.5 (-23.3, 22.3)
>1.38	2125	-18.2 (-49.3, 12.9)	-10.5 (-34.1, 13.1)
trimester 2			
0-0.91	2836	0.0	0.0
>0.91 - 1.38	2364	-21.3 (-51.0, 8.4)	-22.3 (-44.8, 0.2)
>1.38	2385	-22.8 (-52.4, 6.9)	-22.3 (-44.8, 0.1)
trimester 3			
0-0.91	2775	0.0	0.0
>0.91 - 1.39	2407	16.8 (-13.0, 46.6)	-4.9 (-27.3, 17.5)
>1.39	2529	-13.2 (-42.6, 16.2)	-5.8 (-27.9, 16.3)

[♯]mean change in birth weight (95% confidence intervals)
*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes
(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table 7.5: Complete case analysis (CC) and Results after Multiple Imputation (MI): Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (in ug/day) and risk of term LBW by logistic regression (Complete case analysis, Ncases=195, Nnon-cases=4,587, N=4,782, 4.1% prevalence of term LBW; Multiple Imputation using Chained Equations (with LBW in imputation algorithm), Ncases=450, Nnon-cases=10761, Ntotal=11,211, 4.0% prevalence of term LBW)

Combined exposure (ug/day)	**CC**				**MI**			
	cases (n)	**non-cases (n)**	**Crude OR (95% CI)**	**Adjusted* OR (95% CI)**	**cases (n)[1]**	**non-cases (n)[1]**	**Crude OR (95% CI)**	**Adjusted* OR (95% CI)**
DCAA exposure								
0 - 10.41	71	1,404	1.00	1.00	172	3,444	1.00	1.00
> 10.41 - 17.29	61	1,539	0.78 (0.55, 1.11)	0.80 (0.55, 1.15)	139	3,551	0.79 (0.59, 1.04)	0.81 (0.59, 1.11)
> 17.29	63	1,644	0.76 (0.54, 1.07)	0.68 (0.46, 0.99)	139	3,766	0.74 (0.55, 1.00)	0.72 (0.52, 1.00)
TCAA exposure								
0 - 12.02	74	1,431	1.00	1.00	170	3,447	1.00	1.00
> 12.02 - 19.79	60	1,570	0.74 (0.52, 1.05)	0.71 (0.49, 1.02)	140	3,545	0.80 (0.62, 1.04)	0.81 (0.62, 1.07)
> 19.79	61	1,586	0.74 (0.53, 1.05)	0.67 (0.46, 0.97)	140	3,769	0.75 (0.54, 1.05)	0.73 (0.50, 1.05)
BDCAA exposure								
0 - 1.04	81	1,562	1.00	1.00	173	3,507	1.00	1.00
> 1.04 - 1.93	63	1,497	0.81 (0.58, 1.14)	0.82 (0.58, 1.17)	149	3,547	0.85 (0.67, 1.08)	0.89 (0.69, 1.14)
> 1.93	51	1,528	0.64 (0.45, 0.92)	0.62 (0.42, 0.90)	128	3,707	0.70 (0.49, 0.99)	0.70 (0.48, 1.01)

[1] average sample size over 10 imputations

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table 7.6: Complete case analysis: Crude and adjusted association between water consumption (in L/day) and risk of term LBW by logistic regression (all live singletons to eligible mother) (OR: Odds Ratio)

Water consumption (L/day)		cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
cold tap water					
	0 - 0.8	123	2574	1.00	1.00
	> 0.8 - 1.4	79	2069	0.80 (0.60, 1.07)	0.70 (0.52, 0.95)
	> 1.4	75	2122	0.74 (0.55, 0.99)	0.67 (0.49, 0.92)
total tap water					
	0 - 1.2	130	2700	1.00	1.00
	> 1.2 - 1.8	82	2188	0.78 (0.59, 1.03)	0.72 (0.54, 0.98)
	> 1.8	84	2278	0.77 (0.58, 1.01)	0.73 (0.53, 0.99)
bottled water					
	0 - 0.4	36	1099	1.00	1.00
	> 0.4 - 0.8	28	726	1.18 (0.71, 1.95)	1.19 (0.69, 2.04)
	> 0.8	27	804	1.03 (0.62, 1.70)	1.15 (0.66, 1.99)
total water					
	0 - 1.4	136	2663	1.00	1.00
	> 1.4 - 2.0	92	2572	0.70 (0.53, 0.92)	0.73 (0.55, 0.98)
	> 2.0	69	1993	0.68 (0.50, 0.91)	0.77 (0.55, 1.07)

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes
(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table 7.7: Complete case analysis: Crude and adjusted association between average modelled area-level concentrations for DCAA, TCAA and BDCAA by trimester of pregnancy (in ug/L) (based on residence water supply zone) and risk of term LBW by logistic regression

Area-level concentrations (ug/L)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
Average [DCAA]				
trimester 1				
0-7.12	84	2171	1.00	1.00
>7.12 - 9.36	89	2220	1.04 (0.76, 1.40)	0.80 (0.58, 1.10)
>9.36	95	2125	1.16 (0.86, 1.56)	0.93 (0.68, 1.28)
trimester 2				
0-7.20	79	2215	1.00	1.00
>7.20 - 9.43	106	2313	1.28 (0.95, 1.73)	1.17 (0.86, 1.61)
>9.43	102	2379	1.20 (0.89, 1.62)	1.08 (0.79, 1.48)
trimester 3				
0-7.43	79	2305	1.00	1.00
>7.43 - 9.54	103	2355	1.28 (0.95, 1.72)	1.09 (0.80, 1.50)
>9.54	111	2354	1.38 (1.03, 1.85)	1.26 (0.92, 1.72)
Average [TCAA]				
trimester 1				
0-10.38	79	2080	1.00	1.00
>10.38 - 12.49	104	2301	1.19 (0.88, 1.60)	1.16 (0.85, 1.59)
>12.49	85	2135	1.05 (0.77, 1.43)	1.06 (0.76, 1.47)
trimester 2				
0-10.49	92	2132	1.00	1.00
>10.49 - 13.05	102	2446	0.97 (0.72, 1.29)	1.00 (0.74, 1.35)
>13.05	93	2329	0.93 (0.69, 1.24)	0.87 (0.64, 1.19)
trimester 3				
0-10.62	95	2157	1.00	1.00
>10.62 - 13.17	95	2422	0.89 (0.67, 1.19)	0.92 (0.68, 1.25)
>13.17	103	2435	0.96 (0.72, 1.28)	0.92 (0.68, 1.24)
Average [BDCAA]				
trimester 1				
0-0.90	97	2388	1.00	1.00
>0.90 - 1.38	85	2198	0.95 (0.71, 1.28)	1.00 (0.73, 1.37)
>1.38	86	1930	1.10 (0.82, 1.48)	1.13 (0.83, 1.54)
trimester 2				
0-0.91	109	2572	1.00	1.00
>0.91 - 1.38	85	2150	0.93 (0.70, 1.25)	0.94 (0.69, 1.28)
>1.38	93	2185	1.00 (0.76, 1.33)	0.98 (0.73, 1.33)
trimester 3				
0-0.91	103	2512	1.00	1.00
>0.91 - 1.39	83	2216	0.91 (0.68, 1.23)	0.99 (0.72, 1.34)
>1.39	107	2286	1.14 (0.87, 1.51)	1.11 (0.83, 1.49)

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table 7.8: Complete case analysis (CC) and Results after Multiple Imputation (MI): Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (in ug/day) and risk of being SGA by logistic regression (Complete case analysis: Ncases=649, Nnon-cases=4,388, N=5,037, 12.9% prevalence of SGA; Multiple Imputation using Chained Equations (with SGA in imputation algorithm), Ncases=1,440, Nnon-cases=10,423, Ntotal=11,863, 12.1% prevalence of SGA)

Combined exposure (ug/day)	**CC**				**MI**			
	cases (n)	**non-cases (n)**	**Crude OR (95% CI)**	**Adjusted* OR (95% CI)**	**cases (n)[1]**	**non-cases (n)[1]**	**Crude OR (95% CI)**	**Adjusted* OR (95% CI)**
DCAA exposure								
0 - 10.41	213	1,349	1.00	1.00	496	3,376	1.00	1.00
> 10.41 - 17.29	224	1,454	0.98 (0.80, 1.19)	0.96 (0.78, 1.18)	475	3,395	0.95 (0.82, 1.11)	0.97 (0.82, 1.14)
> 17.29	212	1,585	0.85 (0.69, 1.04)	0.87 (0.70, 1.07)	469	3,652	0.87 (0.74, 1.03)	0.91 (0.74, 1.12)
TCAA exposure								
0 - 12.02	218	1,371	1.00	1.00	496	3,364	1.00	1.00
> 12.02 - 19.79	213	1,506	0.89 (0.73, 1.09)	0.89 (0.73, 1.10)	465	3,439	0.92 (0.79, 1.06)	0.94 (0.80, 1.09)
> 19.79	218	1,511	0.91 (0.74, 1.11)	0.95 (0.77, 1.17)	479	3,620	0.90 (0.75, 1.08)	0.95 (0.77, 1.18)
BDCAA exposure								
0 - 1.04	226	1,511	1.00	1.00	485	3,440	1.00	1.00
> 1.04 - 1.93	223	1,428	1.04 (0.86, 1.27)	1.07 (0.88, 1.31)	492	3,438	1.01 (0.88, 1.17)	1.04 (0.89, 1.22)
> 1.93	200	1,449	0.92 (0.75, 1.13)	0.95 (0.77, 1.17)	464	3,546	0.93 (0.78, 1.11)	0.97 (0.79, 1.18)

[1] average sample size over 10 imputations

*adjusted for 8 variables: ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table 7.9: Complete case analysis: Crude and adjusted association between water consumption (in L/day) and risk of being Small-for-Gestational Age by logistic regression (all live singletons to eligible mother) (OR: Odds Ratio)

Water consumption L/day	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
cold tap water				
0 - 0.8	350	2,507	1.00	1.00
> 0.8 - 1.4	304	1,955	1.11 (0.94, 1.31)	1.04 (0.88, 1.23)
> 1.4	244	2,062	0.85 (0.71, 1.01)	0.84 (0.71, 1.01)
total tap water				
0 - 1.2	384	2,607	1.00	1.00
> 1.2 - 1.8	298	2,088	0.97 (0.82, 1.14)	0.96 (0.82, 1.14)
> 1.8	268	2,221	0.82 (0.69, 0.97)	0.86 (0.72, 1.02)
bottled water				
0 - 0.4	146	1,048	1.00	1.00
> 0.4 - 0.8	86	710	0.87 (0.66, 1.15)	0.89 (0.66, 1.19)
> 0.8	79	793	0.72 (0.54, 0.95)	0.79 (0.58, 1.07)
total water				
0 - 1.4	402	2,541	1.00	1.00
> 1.4 - 2.0	332	2,487	0.84 (0.72, 0.99)	0.92 (0.79, 1.08)
> 2.0	221	1,946	0.72 (0.60, 0.85)	0.84 (0.69, 1.01)

*adjusted for 8 variables: ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table 7.10: Complete case analysis: Crude and adjusted association between average modelled area-level concentrations for DCAA, TCAA and BDCAA by trimester of pregnancy (in ug/L) (based on residence water supply zone) and risk of being SGA by logistic regression

Area-level concentrations (ug/L)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
Average [DCAA]				
trimester 1				
0-7.12	264	2,124	1.00	1.00
>7.12 - 9.36	289	2,136	1.09 (0.91, 1.30)	0.95 (0.79, 1.14)
>9.36	303	2,033	1.20 (1.01, 1.43)	1.04 (0.87, 1.25)
trimester 2				
0-7.20	260	2,165	1.00	1.00
>7.20 - 9.43	328	2,218	1.23 (1.04, 1.46)	1.14 (0.95, 1.36)
>9.43	332	2,276	1.21 (1.02, 1.44)	1.07 (0.90, 1.28)
trimester 3				
0-7.43	264	2,255	1.00	1.00
>7.43 - 9.54	332	2,245	1.26 (1.06, 1.50)	1.13 (0.95, 1.35)
>9.54	332	2,277	1.25 (1.05, 1.48)	1.11 (0.93, 1.32)
Average [TCAA]				
trimester 1				
0-10.38	258	2,021	1.00	1.00
>10.38 - 12.49	302	2,229	1.06 (0.89, 1.27)	1.09 (0.91, 1.30)
>12.49	296	2,043	1.13 (0.95, 1.36)	1.16 (0.96, 1.39)
trimester 2				
0-10.49	295	2,059	1.00	1.00
>10.49 - 13.05	314	2,372	0.92 (0.78, 1.09)	0.96 (0.81, 1.14)
>13.05	311	2,228	0.97 (0.82, 1.16)	0.98 (0.82, 1.16)
trimester 3				
0-10.62	295	2,079	1.00	1.00
>10.62 - 13.17	328	2,349	0.98 (0.83, 1.16)	1.03 (0.86, 1.22)
>13.17	305	2,349	0.92 (0.77, 1.09)	0.93 (0.78, 1.10)
Average [BDCAA]				
trimester 1				
0-0.90	297	2,327	1.00	1.00
>0.90 - 1.38	287	2,113	1.06 (0.90, 1.26)	1.02 (0.86, 1.22)
>1.38	272	1,853	1.15 (0.97, 1.37)	1.10 (0.92, 1.32)
trimester 2				
0-0.91	317	2,516	1.00	1.00
>0.91 - 1.38	296	2,065	1.14 (0.96, 1.35)	1.17 (0.98, 1.39)
>1.38	307	2,078	1.17 (0.99, 1.39)	1.13 (0.95, 1.34)
trimester 3				
0-0.91	323	2,449	1.00	1.00
>0.91 - 1.39	288	2,117	1.03 (0.87, 1.22)	1.01 (0.85, 1.20)
>1.39	317	2,211	1.09 (0.92, 1.28)	1.05 (0.88, 1.24)

*adjusted for 8 variables: ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table 7.11: Proportions missing in continuous birth weight models

Variable	Complete	Incomplete	Imputed	Total	% missing
Combined metric (DCAA, TCAA or BDCAA)	6195	5679	5679	11874	48%
Caffeine intake	8935	2939	2939	11874	25%
Maternal BMI (quartiles)	9478	2396	2396	11874	20%
Smoking status	9825	2049	2049	11874	17%
Maternal education	9821	2053	2053	11874	17%
Ethnicity	9805	2069	2069	11874	17%
Gestational diabetes	11381	493	493	11874	4%
Parity	11442	432	432	11874	4%
Maternal age	11874	0	0	11874	0%
Gestational age (in weeks)	11874	0	0	11874	0%
Sex of child	11874	0	0	11874	0%

Table 7.12: Proportions missing in term LBW models

Variable	Complete	Incomplete	Imputed	Total	% missing
Combined metric (DCAA, TCAA or BDCAA)	5869	5342	5342	11211	48%
Caffeine intake	8470	2741	2741	11211	24%
Maternal BMI (quartiles)	8974	2237	2237	11211	20%
Smoking status	9301	1910	1910	11211	17%
Maternal education	9297	1914	1914	11211	17%
Ethnicity	9283	1928	1928	11211	17%
Gestational diabetes	10761	450	450	11211	4%
Parity	10794	417	417	11211	4%
Maternal age	11211	0	0	11211	0%
Gestational age (in weeks)	11211	0	0	11211	0%
Sex of child	11211	0	0	11211	0%

Table 7.13: Proportions missing in SGA models

Variable	Complete	Incomplete	Imputed	Total	% missing
Combined metric (DCAA, TCAA or BDCAA)	6189	5674	5674	11863	48%
Caffeine intake	8926	2937	2937	11863	25%
Maternal BMI (quartiles)	9468	2395	2395	11863	20%
Smoking status	9814	2049	2049	11863	17%
Maternal education	9810	2053	2053	11863	17%
Ethnicity	9794	2069	2069	11863	17%
Gestational diabetes	11371	492	492	11863	4%
Parity	11432	431	431	11863	4%
Maternal age	11863	0	0	11863	0%

7.6 Figures

Figure 7.1: Scatter plots (and LOWESS smoother) of the DCAA, TCAA and BDCAA combined metrics of exposure (ug/day) by birth weight (live births, Nmax=11,928)

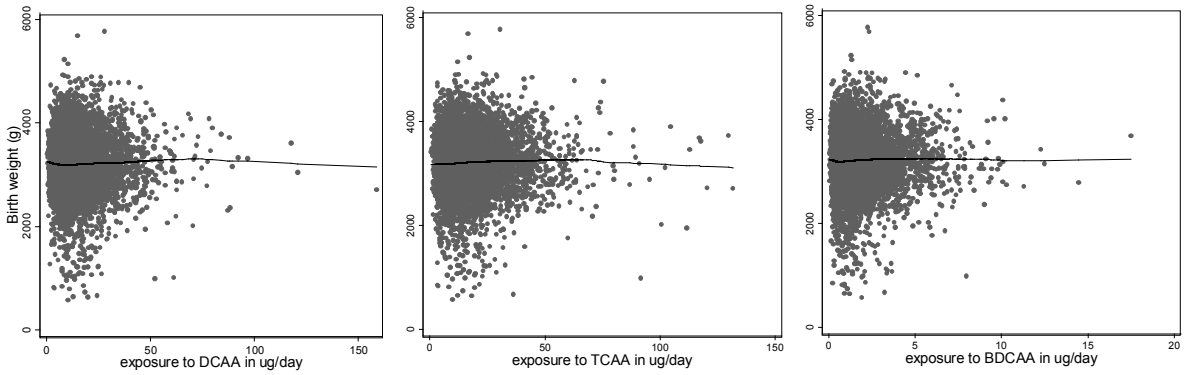


Figure 7.2: Scatter plots (and LOWESS smoother) of DCAA, TCAA and BDCAA combined metrics of exposure (ug/day) by standardised (or relative) birth weight (live births, Nmax=11,928)

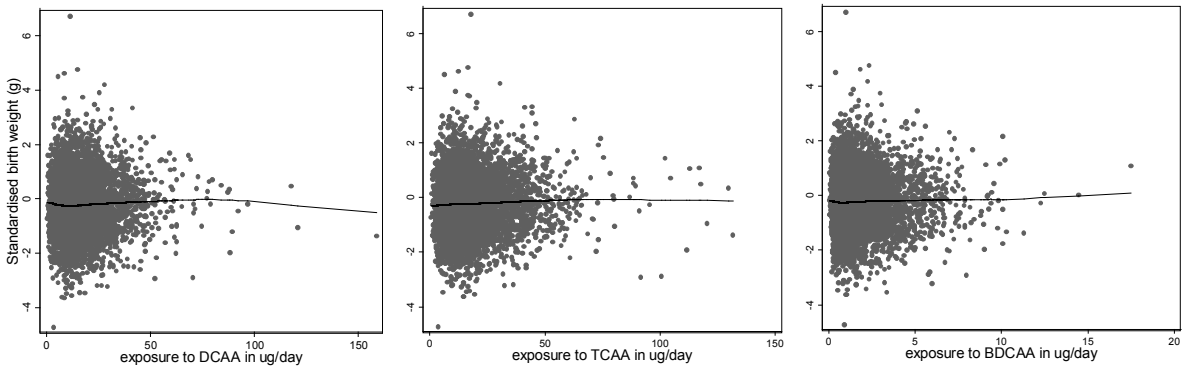


Figure 7.3: Complete Case Analysis (left) and results after Multiple Imputation (right): Adjusted coefficients (and 95% CI) for combined metric exposure on continuous birth weight (in grams)

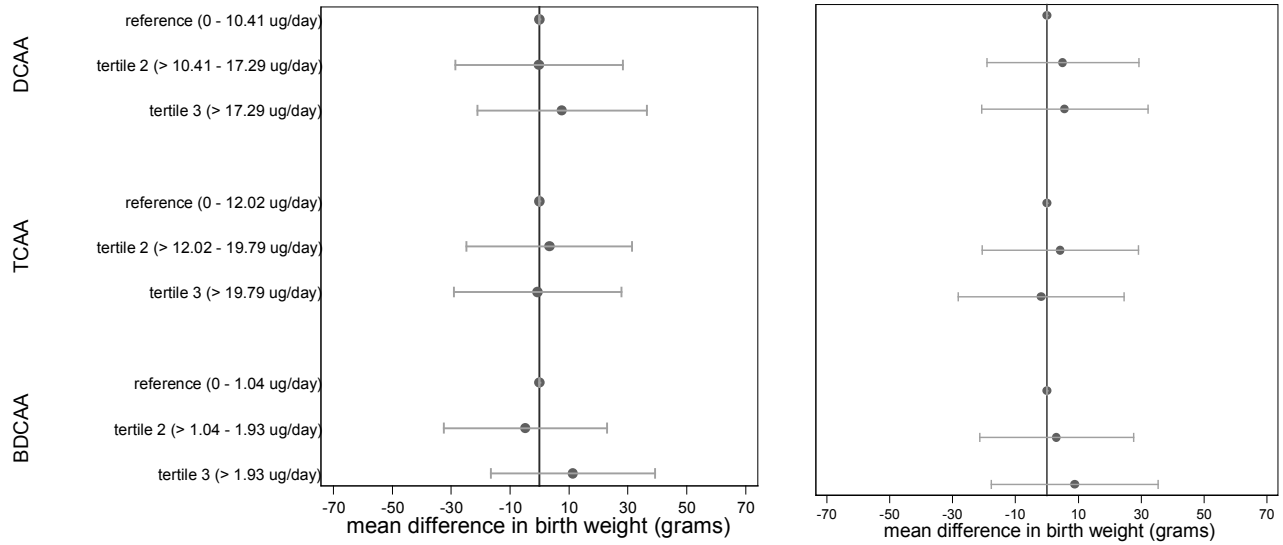


Figure 7.4: Complete Case Analysis: Adjusted coefficients (and 95% CI) for water consumption on continuous birth weight (in grams); Ctw: cold tap water; Ttw: total tap water; Bw: bottled water; Tw: total water

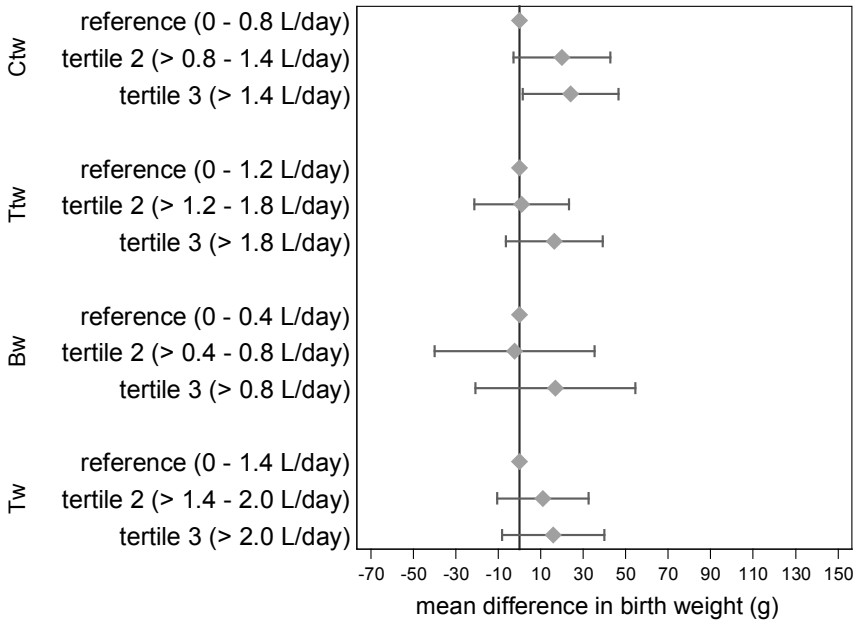


Figure 7.5: Complete Case Analysis: Adjusted coefficients (and 95% CI) for exposure to area-level concentration to DCAA, TCAA and BDCAA on continuous birth weight (in grams)

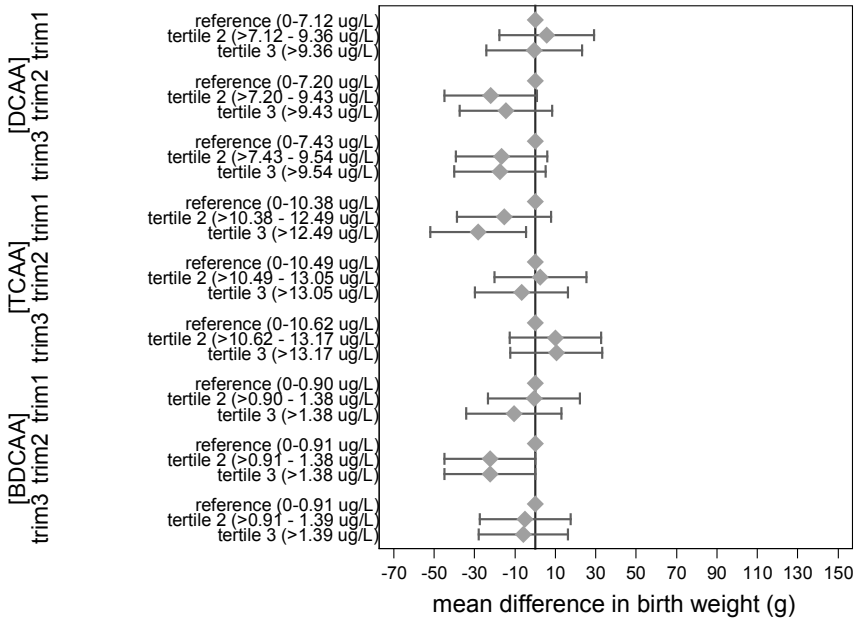


Figure 7.6: Complete Case Analysis (left) and results after Multiple Imputation (right): Adjusted Odds Ratios (and 95% CI) for combined metric exposure on term LBW

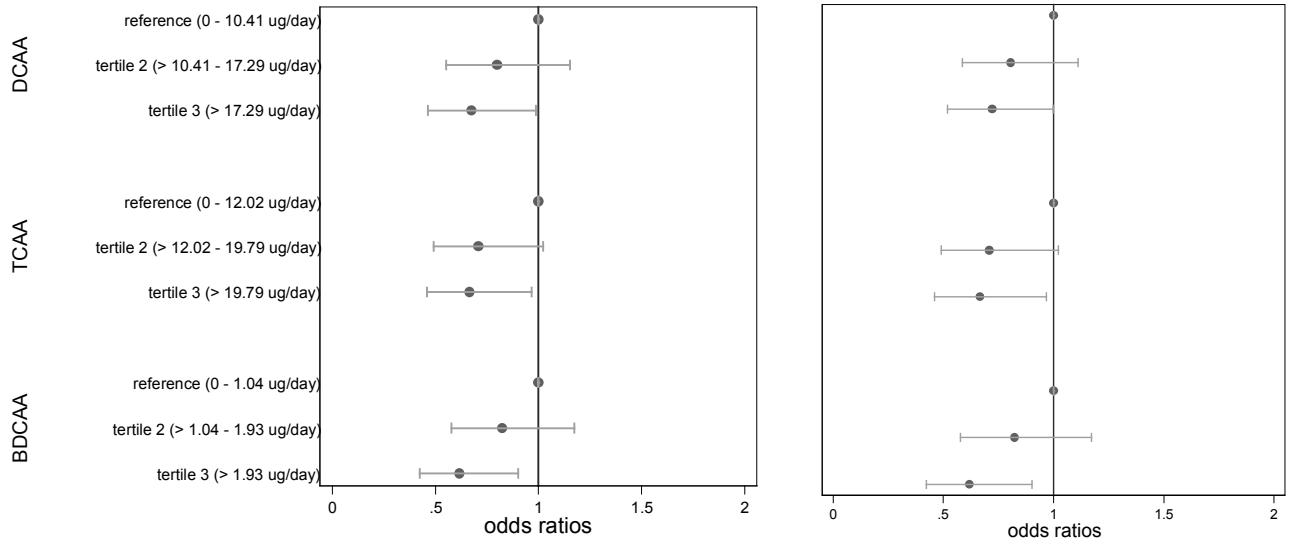


Figure 7.7: Complete Case Analysis: Adjusted Odds Ratios (and 95% CI) for water consumption on term LBW. Ctw: cold tap water; Ctw: cold tap water; Ttw: total tap water; Bw: bottled water; Tw: total water

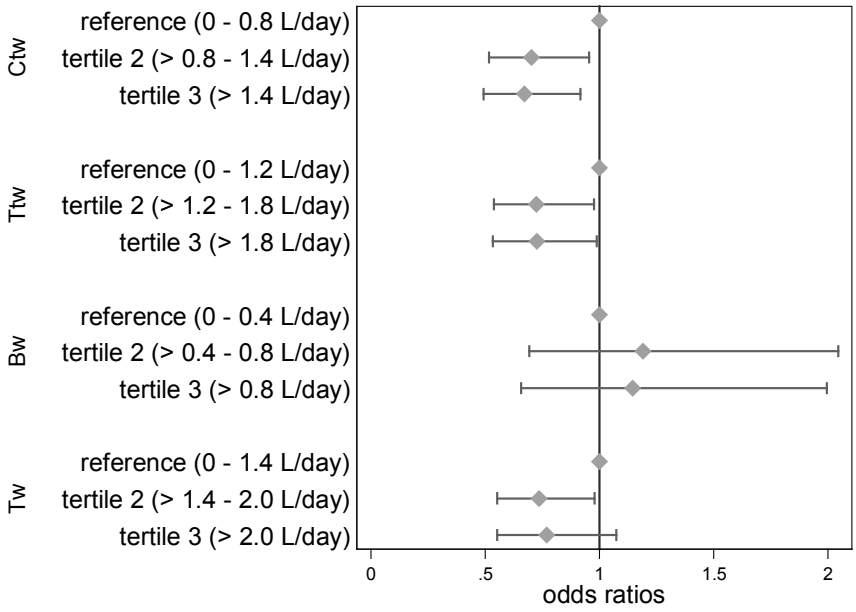


Figure 7.8: Complete Case Analysis: Adjusted Odds Ratios (and 95% CI) for exposure to area-level concentration to DCAA, TCAA and BDCAA on term LBW

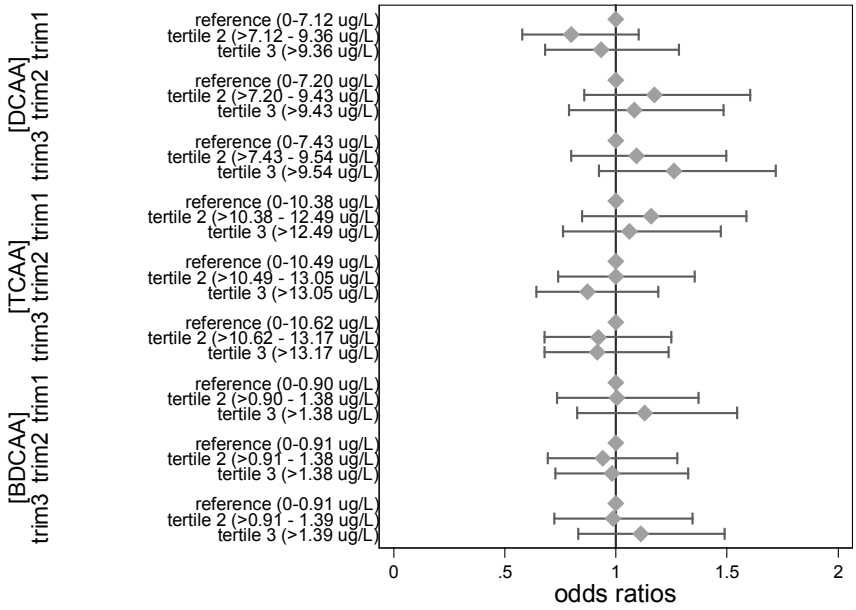


Figure 7.9: Complete Case Analysis (left) and results after Multiple Imputation (right): Adjusted Odds Ratios (and 95% CI) for combined metric exposure on SGA

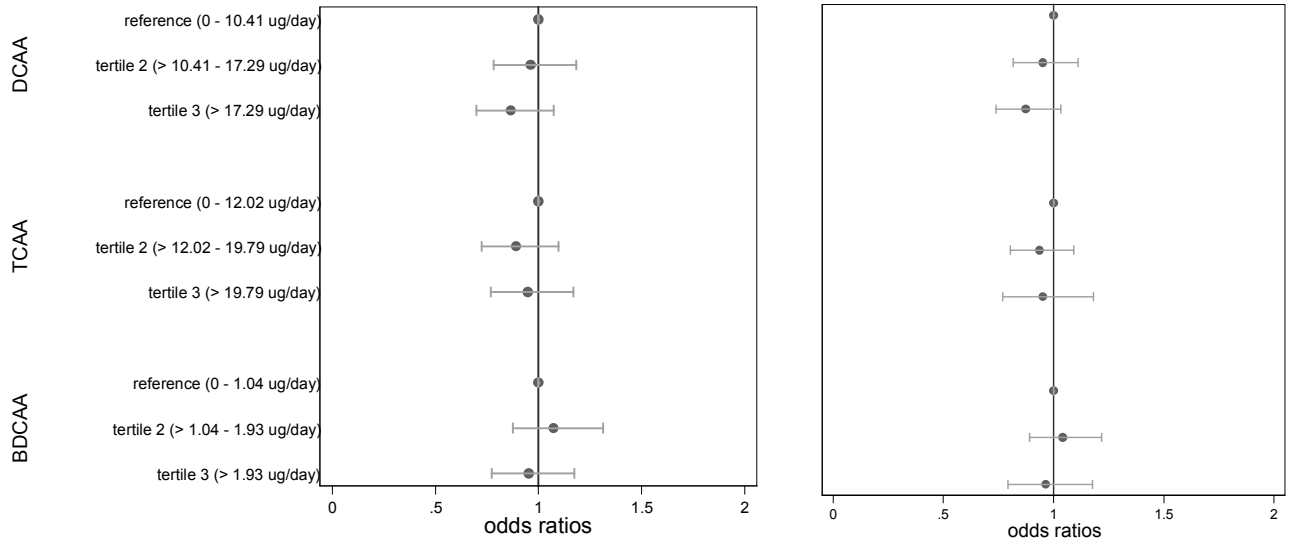


Figure 7.10: Complete Case Analysis: Adjusted Odds Ratios (and 95% CI) for water consumption exposures on SGA (excluding births born before 23 weeks of gestation or after 42 weeks of gestation); Ctw: cold tap water; Ttw: total tap water; Bw: bottled water; Tw: total water

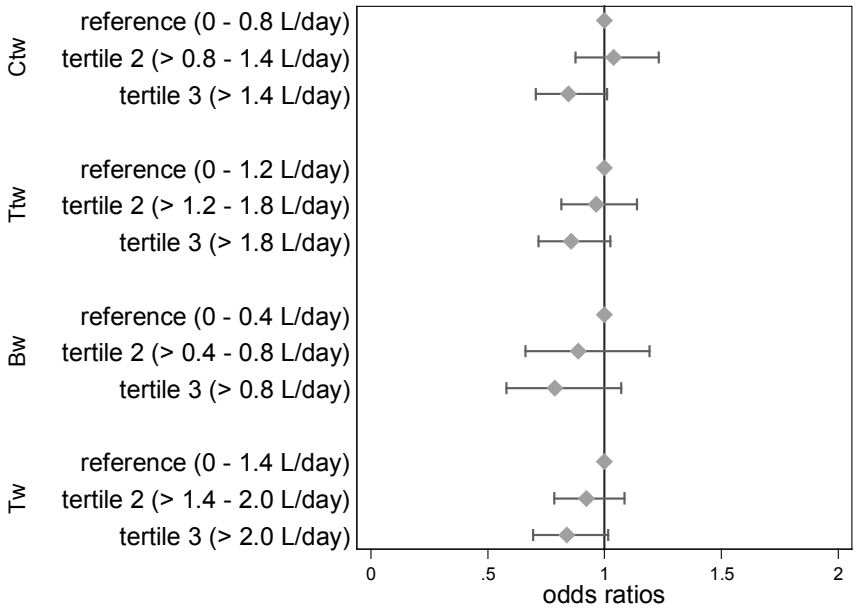
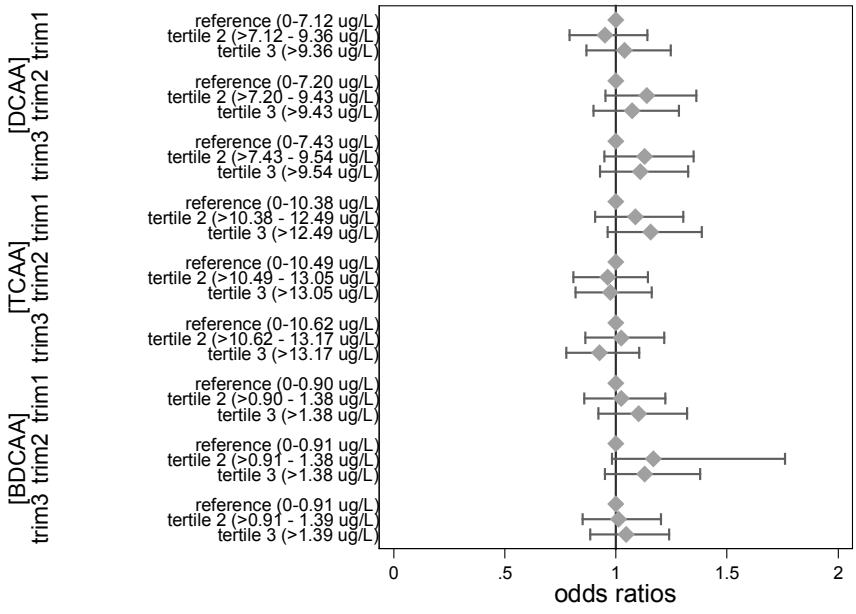


Figure 7.11: Complete Case Analysis: Adjusted Odds Ratios (and 95% CI) for exposure to area-level concentration to DCAA, TCAA and BDCAA on SGA



CHAPTER 8 REPEAT QUESTIONNAIRE STUDY

This chapter presents the results of a repeat questionnaire study conducted on a subset of BiB participants in late 2010 to assess their water consumption behaviours over the course of late pregnancy (Aim 3, Chapter 2). It is arguably an extension to Chapter 3 on women's water consumption, but is placed here at the end of the thesis as it was not included in the final epidemiologic analyses that Chapters 3 through 6 lead up to and used some different methodology.

8.1 Background

Water consumption patterns can vary between individuals and groups and are influenced by source water quality perception and various socio-demographic factors (Wright et al. 2006). Forssen et al. (2009) and others suggest that the self-reported intake and use of tap water varies considerably during the course of pregnancy.

Combining information on individual water use with area-level concentrations (as done in Chapter 5) improves on using either exposure assessment on its own. However, the individual water use information should be evaluated for measurement error because within-subject variability in questionnaire data may be substantial (Forssen et al. 2009) and attenuate risk estimates overall (Nieuwenhuijsen et al. 2009a). But irrespective of whether the reported increases and decreases in water use are intentional (perhaps even related to participation in the study) or reflective of random variability, studies that rely on only one reporting period to reflect water-related exposure throughout pregnancy could be subject to considerable exposure misclassification (Forssen et al. 2009).

Smith's work on exposure assessment and analysis of exposure metrics in this very cohort (Smith 2011) confirms that, particularly when spatial variability in DBPs is limited across the study area, failure to incorporate individual water use and its variability into exposure assessment can result in exposure misclassification, which may lead to loss of study power and bias of risk estimates (Armstrong 1998; King et al. 2004).

Several studies have actually tried to quantify the amount of misclassification that is likely to result from ignoring variation in individual behaviour (King et al. 2004; Waller et al. 2001; Whitaker et al. 2003b; Wright and Bateson 2005; Wright et al. 2006; Zender et al. 2001). Reif et al. (2000) for instance estimated that 20% nondifferential misclassification of subjects with low to intermediate exposures into the high exposure groups could result in substantial attenuation of the observed effect

estimates for the high exposure group. This may be of particular concern for smaller studies, since even minor exposure misclassification can result in substantial bias and reduced statistical power to detect subtle increases in health risk (Wright et al. 2006).

8.1.1 Aim of study

The aim of this study is to investigate the variability in volumes of water consumed daily by BiB women in their (possibly critical) third trimester of pregnancy, in order to assess the consistency of the baseline measurements available on the full cohort. To do so repeat water questionnaire data were collected on a subset of women enrolled in the main BiB cohort to assess their water consumption trends over the course of late pregnancy. This should minimise exposure misclassification (Wright et al. 2006) by enabling us to estimate within- and between-subject variability, and correct for attenuation of effect estimates due to exposure measurement error in the final epidemiologic models (Bateson and Wright 2010).

In addition, the third trimester of pregnancy is suspected to be a critical exposure period for fetal growth as the rate of fetal growth and weight gain increases dramatically and reaches its peak at about week 33 (Owen et al. 1996; Williams et al. 1982). If that is correct, relying on second trimester estimates in the full epidemiologic investigation of the effect of prenatal HAAs on fetal growth without considering the possibility of behaviour change during the third trimester period could lead to misclassification and biased inference.

8.1.2 Water consumption during pregnancy using repeat questionnaire

Previous studies all conclude that questionnaires are a valid measure of exposure during pregnancy, but tend to slightly overestimate tap water intakes (Barbone et al. 2002; Kaur et al. 2004; Maskiell et al. 2006; Shimokura et al. 1998; Smith et al. 2012).

The few previous published studies to assess the variability of tap water consumption over the course of women's pregnancies using repeat questionnaire or interviews found that the self-reported ingestion of tap water can vary during the course of pregnancy. Windham et al (1992) examined changes in tap water consumption between the time before pregnancy and the time of an interview performed after delivery (or pregnancy loss) in 1,926 US women. This study showed that women were more likely to report increased rather than decreased consumption of tap water during pregnancy. More recently, Forssen et al (2009) found considerable variation in behaviour between two questionnaires administered in early and mid-pregnancy to a population of 1,990 US women (for example, 33% of women changed their ingestion of cold tap water by $\geq$ 1 L/day). However, a

validation study comparing repeated 7-day diaries carried out in 2008 within the BiB cohort itself reported little change in overall tap water consumption over the third trimester of pregnancy (Smith et al. 2012).

8.1.3 Assessment of water consumption at baseline in BiB

In the water section of the BiB baseline questionnaire (administered between week 26-28 of pregnancy), women were asked about the number of glasses/cups they consumed per day of tap water, bottled water, tea, coffee and squash at home, work or place of study, and elsewhere; about their use and type of water filter; and about the frequency and duration of their showering, bathing and swimming habits (see Chapter 3). A detailed description of the water consumption and water use patterns at baseline of BiB women is in preparation (Smith et al. in preparation). This work shows that women consume on average 1.92 (95% CI: 1.90, 1.93) litres of total water per day (N=11,334), the greatest proportion of which is tap water (85% of total water consumed) consumed in the home (79% of total water consumed). Maternal age, ethnicity, employment status, smoking status, cohabitation status, and amount of weekly physical exercise reported are consistent predictors of total water consumption during the second trimester of pregnancy.

8.2 Methods

8.2.1 RQS recruitment

During the last months of recruitment to BiB (between September and December 2010), I mailed out 619 (mailing 1) and 616 (mailing 2) repeat water questionnaires in hard-copy to the homes of each eligible BiB participant. (Three women's Q1 responses were invalid, which is why total number of questionnaires sent out for Q2 dropped to 616). Eligibility to the study included being proficient in English (as determined by the language of administration of the baseline questionnaire), having completed the water section of the baseline questionnaire upon recruitment and being within the appropriate period of pregnancy during the RQS period. The mailings were timed to correspond with two windows of pregnancy for each woman: 30 to 33 weeks of pregnancy (to capture an early third trimester window of pregnancy, hereafter called "Q1"), and 36 to 39 weeks of pregnancy (late third trimester window, hereafter "Q2") (Figure A8 - 1).

The packet received by each eligible woman included a personalised invitation letter and information sheet briefly describing the aims of the study, a recruitment form to confirm contact details and any changes in address or work status, the water use questionnaire, and a Freepost return envelope (see Appendix C). The recruitment form containing all identifiable confidential information was kept

separate from the questionnaire upon completion. Consent was assumed when women filled out and returned the questionnaires to us.

In order to be credible repeats, the questions in the repeat questionnaire were identical to the water section of BiB's baseline questionnaire, with two notable differences: a) questions about water consumption in the repeat questionnaires were asked about the week directly preceding the date of questionnaire completion ("on a typical day in the past week of your pregnancy, how much of the following did you drink?") compared to a vaguer "on a typical day, how much of the following do you drink?" in the baseline questionnaire, and b) the repeat questionnaires asked about perception of tap water drinking and use habits: "do you think that [your tap water drinking habits] have changed since you completed the last questionnaire?" (yes/no).

The repeat questionnaires also asked whether women were employed and/or on maternity leave or sick leave, or whether they were full-time students at questionnaire completion. The variable for employment status therefore changes over time, and contains some missing data.

This RQS was approved by the Bradford Research Ethics Committee in September 2010.

8.2.2 Data preparation

As described in Chapter 3 (section 3.2.2.1), and consistent with previous work in this area (Iszatt et al. 2011), from the five water types queried in the repeat questionnaires, the following four summary water consumption variables were derived per time point: cold tap water (tap water+squash), hot tap water (coffee+tea), total tap water (cold+hot tap water), and total water (total tap water+bottled water). Only the latter two are analysed in this chapter, as total tap water is pertinent to assessing HAA exposure, and total water is the most complete summary variable available on water consumption. These measures were also sorted by location: home, outside of the home (which combines the work/study, and elsewhere categories), and in total. All questionnaire entries were converted from glasses/cups per day to litres per day (the relation "1 cup or glass = 200ml" was explicitly stated in the questionnaires). Boiling and filtering reduce DBP concentrations in tap water, and bottled water is assumed to contain no HAAs (Kim 1997; Krasner and Wright 2005; Levesque et al. 2006; Savitz et al. 2005).

8.2.3 Statistics

Water consumption values at 2 or more of the 3 possible time points were necessary for inclusion into the analysis (N=254).

Linear mixed regression models (with random intercepts and random slopes) were run on total tap water and total water consumption (L/day) run in WinBUGS 1.4.1 (Lunn et al. 2009). A number of interactions were studied. All significant covariates at $p \leq 0.05$ in univariate analyses on total water consumption were included in the final model. I then used a Bayesian approach in order to be able to impute missing outcome data (total tap water, or total water values) and occasional missing covariate data (e.g. employment status) and to handle censored observations. In a Bayesian framework, missing data are included in the model as an extra parameter, and estimated at every iteration during the Markov Chain Monte Carlo (MCMC) sampling process. The parameters' posterior distribution propagates the posterior uncertainty in the imputed missing data (Daniels and Hogan 2008; Gelman et al. 2004). All results presented are posterior means and 95% credible intervals.

Table A8 - 1 lists the equations for the linear mixed models (all with random intercepts and random slopes) and interactions presented in this chapter.

The variance partition coefficient (VPC) was also reported for each model, which is the proportion of total residual variability due to between-subject variability (Goldstein et al. 2002). This will help us interpret the possible source of this residual variability.

All linear mixed models for total water and total tap water consumption were bounded at zero to only allow the imputation of positive values. Non informative priors were given to all parameters.

Other statistics tests such as chi^2 tests, t-tests, and ANOVAs were performed in STATA 11.1 (StataCorp 2009).

Of note, the sample sizes in this chapter were not restricted to women with birth outcomes, and did not exclude multiple births or even multiple entries (as was done in the rest of this thesis, as explained in section 3.2.3) because most of this information was not available to me at the time that this study was conducted and analysed.

8.2.4 Treatment of missing data

Of the 254 respondents, 127 completed all 3 questionnaires and 127 completed the baseline and only one of Q1 or Q2. In addition, employment status was missing for a total of 145 women (56 at Q1 and 89 at Q2), of which a total of 18 women (11 at Q1 and 7 at Q2) women responded to the water consumption questions (therefore will contribute new information after their employment status is imputed).

I assumed that no answer in a given outcome category of the repeat water questionnaire (e.g. a box in the questionnaire was left blank) meant that the participant did not consume any water in that particular category. Such missing values were assigned a value of zero in the summary tables (Table A8 - 2). This assumption that no answer means no consumption was tested in a sensitivity analysis by treating zeros as censored observations in order to check influence on the effect estimates.

8.2.5 Sensitivity analysis

I ran sensitivity analyses comparing each model (on total tap water, and on total water consumption), with and without covariates, to

a) models which excluded potential outliers (per standardized residuals criteria) (data not shown),

b) models making a different distributional assumption for the observed data likelihood (specifically, assuming a Student's t-distribution with 4 degrees of freedom which is more robust to outliers than the normal distribution), and

c) models censoring i) any summary value containing missing entries in an otherwise completed questionnaire. The sampling space is assigned the range: X-6 L/day, where X is the sum of the components of the summary measure which were actually specified. X was then assumed to represent the minimum volume ingested by that individual, but may be incomplete and underestimate true consumption and ii) total water consumption values above 6 L/day, which are considered excessive and questionable. The sampling space is then assigned the range: 0 - X L/day so as to incorporate the information provided but censor any total water consumption values greater than 6 L/day.

A total of 248 outcomes were censored in Model 1c on total tap water, and 252 in Model 1c on total water.

8.3 Results

8.3.1 Descriptive results

1008 women registered to BiB within the RQS recruitment period, of which 882 (88%) completed the water section of the baseline questionnaire. Of these, women who were not English speaker (140), and those not in the relevant window of pregnancy (123) were excluded, leaving us with 619 (61%) women eligible for a study invitation (Figure A8 - 2).

A total of 381 (31% of 619+616 sent out) repeat questionnaires were returned to us: 209 women (34% of the 619 possible returns) responded to Q1, 172 (28%) responded to Q2, and 127 (21%) responded

to both Q1 and Q2. This means that 82 (13%) women responded only to Q1, and 45 (7%) women responded only to Q2.

Table 8.1 presents the demographic and behavioural characteristics at baseline of the RQS subset compared to the rest of the cohort. Proportionally more women in the subset were White British and less women were Pakistani compared to the rest of the cohort. The RQS subset was a bit older, and more likely to have achieved a higher educational degree. Households are more frequently smaller (in the 0-2 member range rather than in the ≥5 member range) among the women in the subset. RQS women are more likely to be having their first baby, have less gestational diabetes on average and are more physically active in their leisure time, but they are also more likely to have smoked in the past (slightly less likely to currently smoke). They are also substantially less likely to drink any caffeinated drinks during pregnancy compared to the rest of the cohort (97% in RQS vs. 30% in the rest of the cohort).

In Figure 8.1 and Figure 8.2, summaries of water consumption are presented by water type (e.g. cold tap water, hot tap water, etc.) and location (in the home, or out of the home), and graphically compare medians at baseline, Q1 and Q2 in the RQS subset of women. RQS women's mean weeks of gestation was 26.5 ± 0.7 weeks (range: 24.4, 29.7) at baseline, 31.9 ± 1.7 weeks (range: 23.3, 40.9) at Q1, and 37.2 ± 1.2 weeks (range: 33.1, 43.3) at Q2, which fall within the targets. As described in Chapter 3, and as similarly found for the whole BiB cohort (Smith et al. in preparation), the greatest proportion of total water drank is cold tap water (Figure 8.1), and the most important location of water consumption is the home (Figure 8.2). These findings also mirror those of another UK study (Kaur et al. 2004). Because of the small contributions of different tap water types at different locations and the many missing/zeros in some of the categories (Table A8 - 2), I focus on total tap water (the sum of cold and hot tap water) and total water (the sum of total tap water and bottled water) trends at all locations in this analysis.

Table 8.2 (Model 1) presents the results of univariate analyses. There is a significant increase in total tap water consumption (time pattern: 0.13 L/day (0.04, 0.23)) and total water consumption (0.14 L/day (0.03, 0.26)) per time point, corresponding to an increase of about ¾ of a 200ml glass or cup of water per time point, i.e. from baseline to Q1 and from Q1 to Q2.

8.3.2 Variable selection for water consumption

In univariate linear mixed models with random intercepts and slopes ran in STATA (frequentist models), the key predictors of total water consumption trend from baseline to Q2 are women's ethnicity (p=0.012), smoking status at baseline (p=0.025), employment status (p=0.025), physical

activity level in the past week (p=0.002), physical activity in paid work (p=0.050) and total number of household members (p=0.033) (data not shown). The direction of these effects are: the White British drink significantly more than all the other ethnic groups, current smokers drink significantly more than ever smokers who drink more than never smokers (i.e. a gradient effect), unemployed women drink significantly less than women who were employed, working women who report exercising weekly drink significantly more than inactive women (both at home and work), and finally, having 5 or more household members predicts significantly less total water consumption than having a 0-2 member household. These observations again are consistent with previous descriptive work on water consumption in this cohort (Smith et al. in preparation).

Maternal age, season of the baby's birth, report of exposure to second-hand smoke or of consumption of alcoholic or caffeinated drinks during pregnancy, maternal/paternal highest education level achieved, or the Index of Multiple Deprivation (IMD) score for 2010 were not significant in these univariate analyses on total water consumption and thus excluded from the final model.

8.3.3 Adjusting for covariates – Model 2

Model 2 adjusts for a number of covariates identified above. There is a significant increase in total tap water consumption (time pattern: 0.19 L/day (0.08, 0.30)) and total water consumption (0.22 L/day (0.09, 0.35)) per time point, after adjusting for ethnicity, employment status, smoking status, hours spent exercising weekly (in leisure and in paid work), and total number of household members (Table 8.3, Model 2). This corresponds to an increase of approximately 1 glass or cup of water per time point.

The effects of ethnicity, employment status and physical exercise are significant in the final model with six covariates for total water; in addition to these, smoking status and maternal physical exercise exerted in paid employment were significant predictors of total tap water consumption.

Physical activity (for leisure and in paid work) and number of members of household are correlated variables (Spearman rank correlation coefficients between 0.15 and 0.27, $p<0.05$). Removing the 'physical activity at work' and 'total number of household members' variables from the model to avoid possible multicollinearity did not change the model fit, nor the direction of the estimates or their significance compared to the univariate analyses (data not shown).

8.3.4 Effect modification with time

8.3.4.1 Ethnicity (allowing for separate random slopes by ethnic group) – Model 3

I studied the interaction between maternal ethnicity and time in the adjusted models (Model 2) by allowing for separate random slopes and random intercepts for women in each of the three ethnic group (Table 8.4, Model 3). Pakistani women increased their total water (and total tap water) consumption at a steeper rate over time than any other ethnicity from baseline to Q2, suggesting that the relationship between water consumption and ethnicity changes significantly over time in this group of women.

Whereas Pakistani women drink more over time (total water mean slope = 0.49 L/day (0.19, 0.80)), the variability around the random slopes for Pakistani women is also greater (total water random slopes SD=0.99 L/day (0.70, 1.30)) than for the other two groups, White British (0.21 L/day (0.02, 0.43) and 'Other ethnic' 0.14 L/day (0.01, 0.35) groups, respectively. Though less marked, the same pattern is true for total tap water. This suggests that there is a lot of heterogeneity in trends over time for the Pakistani women.

8.3.4.2 Employment (simple interaction) – Model 4

Model 4 is the adjusted Model 2 but with a "simple" interaction between employment and time. Employment was also found to be a significant modifier of total water consumption over the pregnancy period (Table 8.5, Model 4). The same holds true for total tap water. Of note, women tend to remain more and more at home as pregnancy progresses: indeed the proportion of employed women who work at the time of questionnaire completion decreases from 56% at baseline, to 47% at Q1, to 13% at Q2. The lack of significant results in both the "employed and on maternity leave" and the "full-time students" categories is likely due to the insufficient number of subjects available in these two groups for interaction analysis. The employed and on maternity leave group is the only group which decreases its trend in water consumption over time (slope of -0.20 (-0.64, 0.23)), though this trend is not significant. This could be explained by the fact that employed women taking maternity leave join the unemployed group, who drink less on average than the employed and working, hence their decreasing water consumption trend over time.

8.3.4.3 Smoking (simple interaction) – Model 5

Model 5 is the adjusted Model 2 but with a "simple" interaction between smoking status and time. Never smokers (65% of RQS women) and ever smokers (25%) significantly increase their water consumption over time by approximately 1 glass or cup a day (slopes of 0.26 (0.11, 0.41) and 0.25 (0.01, 0.49), respectively), while current smokers (10%) do not (-0.08 (-0.46, 0.30)) (Table 8.6, Model 5). Smoking status was asked at baseline; there is no data on women's putative change in smoking behaviour over late pregnancy.

Ethnicity remains a predictor of total water (and total tap water) consumption in all models, as does the first category of maternal physical exercise for leisure at baseline (<1 hour per day) relative to no physical exercise in both Models 4 and 5. For total tap water, RQS women who at baseline report mostly standing or walking at work (21%) are significant predictors of consumption in both Models 4 and 5, drinking less than the reference unemployed group. Women who live in households with 4 members of household at baseline (18%) (but not 5 or more) are also significant predictors of consumption in Model 5, as in Model 3.

Figure 8.3 depicts the slopes of the different groups included in interaction models (Models 3, 4, and 5).

Table 8.7 summarises the differences in total water consumption in different strata of the effect modifiers compared to reference (White British, or employed & working, or never smokers), at baseline, at Q1 and at Q2, as predicted by Models 3, 4 and 5. It shows for instance that ignoring differences between women due to ethnicity would overestimate Pakistani women's total water consumption by 0.34 L/day compared to the White British women's consumption at Q2, and underestimating unemployed women's consumption at Q1 by 0.45 L/day compared to the employed and working.

The difference between Models 4 and 5 which include a simple interaction and Model 3 which allows for random slopes by ethnic group is that the random slopes for the former two are based on the overall mean time effect, while they are based on ethnicity-specific means for the latter. The reason for the two types of interaction rests on the observation that rate of change in water consumption over time was different by ethnic group. As such, I ran the interaction allowing for random slopes and intercepts by ethnic group (i.e. allowing for differences in variability), while for the employment and smoking interaction models, I just wanted to check for differences in slopes by group, hence the "simple" interaction.

8.3.5 Self-perception of change in behaviour

Lastly, women's self-perception of change in total tap water drinking habits (yes/no) was compared to the water consumption they actually reported drinking in the questionnaires (Table 8.8). (This question was only asked with respect to total tap water, which is most relevant to HAA exposure.) I hoped to verify whether the change observed in the water volumes women reported consuming was also reflected in their perception of their tap water drinking habits, as a means to distinguish between measurement error (i.e. an imprecise questionnaire) and actual behaviour change over late pregnancy.

(The underlying assumption of this approach being that women's own "perception" is the truth, i.e. gold standard.) There were no statistically significant differences between total tap water volumes consumed by the women who said they did change their tap water drinking habits, and those who said they did not. Table 8.8 presents t-test comparisons.

8.3.6 Variance Partition Coefficient (VPC)

On average, a greater proportion of within- as opposed to between-subject variability is explained by the models (VPCs<0.50). In the interaction model of ethnicity and time (Model 3), the VPCs for Pakistani and White British women were in the 0.42 - 0.52 range, but only 0.10 - 0.18 in the "Other" ethnic category, suggesting that the random effects (capturing between-women variation) explain very little of the variability in this group.

8.3.7 Sensitivity analysis: Models 1b and 1c

I ran sensitivity analyses to test the assumptions made, and found similar effect estimates to those reported (data not shown). As an example, the results of the unadjusted model with a t-distribution (4 degrees of freedom) (Model 1b), and the results after censoring its extreme and unsure outcome values (Model 1c) are presented in Table A8 - 3. Both Models 1b and 1c's VPCs increase for the total tap water models (and for total water models, though less markedly).

Both for total tap water and total water consumption overall, censoring the dependent variables I was unsure of as well as extreme values (water consumption>6 L/day) increased the fixed time pattern estimate to 0.64 L/day (0.53, 0.75) from 0.13 L/day (0.04, 0.23) for total tap water, and 0.66 L/day (0.55, 0.77) from 0.14 L/day (0.03, 0.26) for total water (Table 8.2). These increases constitute a difference in consumption of approx. 2 glasses/cups per time point but may be overestimating the actual increase in water consumption over time because of the constrained sampling space in the likelihood of these censored models.

8.4 Discussion

This study focused on the trends in total tap water (sum of cold tap water, squash, hot tea and coffee) and total water (sum of total tap water and bottled water) consumption of a subset of 254 BiB women over their third trimester of pregnancy (between weeks 27 and 39 of pregnancy). It provides new information on a small but significant increase in water consumption over the third trimester of pregnancy. Overall, unemployed women reported drinking significantly less total water than working women (Model 2), while current smokers and the White British women drank on average the most. Over time, being Pakistani, being employed and working (as opposed to on maternity leave), being

unemployed, or never having a history of cigarette smoking all predict an increase both in total tap water and total water consumption (Models 3, 4, 5). These results should be considered in any future modelling of third trimester exposure estimates in order to help minimise exposure misclassification.

Little is known about women's water consumption behaviour in late pregnancy, a window thought to be critical to DBP exposure as the rate of fetal growth and weight gain dramatically accelerates. This increase in consumption over the course of pregnancy is consistent with Windham et al's results (1992). The considerable variation in women's responses whether over time, by strata, or over time by strata also agrees with Forssen et al's study (2009). However these results are not consistent with Smith et al's study (2012) which found no change in absolute volume of tap water intake between two diaries administered approximately 9 weeks apart (mean difference= -0.01 L/day (-0.30, 0.27)) within a very small sample of the BiB cohort (N=14).

If the finding of an increase in water consumption over the second half of pregnancy is a true increase, this result suggests that exposure estimates based solely on volumes of tap water (the summary measure most relevant to HAA exposure) reported to be consumed at baseline (i.e. at the end of their second trimester of pregnancy) may underestimate exposure during the critical period by up to 2 glasses/day at the end of the third trimester (Model 2). This underestimation may be even greater in different strata of women: based on these models, Pakistani women would drink 0.98 L/day more at Q2 on average than they do at baseline while their White British counterparts drink only 0.14 L/day more at Q2 than they do at baseline (Figure 8.3). Similarly, both the employed and working women, and the unemployed women will drink 0.52 L/day more at Q2 on average than they do at baseline, while the employed women who are on maternity leave may drink less at Q2 than they do at baseline. Never smokers drink 0.52 L/day more at Q2 than they do at baseline, while current smokers drink approximately 0.16 L/day less at Q2 than they do at baseline. Because these models suggest that current smokers drink on average 0.58 L/day more than never smokers at baseline (Table 8.7), differences in slope directions mean that this difference in consumption at baseline may in fact become milder over the course of the late third trimester.

Previous studies have shown a difference in reliability of questionnaire responses by employment status. A validation study in BiB which compared questionnaire responses to simultaneous diary entries concluded that agreement between the two instruments was consistently much stronger for responses given by unemployed women compared to employed women (Smith et al. 2012). The authors' explanation for this observation is that employed women lead busier lives and may find it more difficult to interpret what constitutes a "typical" day as patterns of behaviour are likely to be different across working and non-working days (Smith 2011). In contrast, the unemployed women are less pressed for time and therefore take more care in their responses (Smith et al. 2012). Shimokura et

al's study (1998) which assessed variation in water use based on a 3-day diary among pregnant women in North Carolina, USA, also concluded that employed women had more heterogeneous consumption patterns over time compared to women working part-time or less, suggesting that their responses may be less reliable. Even though both of these studies were based on very small sample sizes, they reinforce the conclusion that employment status has a significant impact on lifestyle behaviours as well as the self-reporting of these behaviours. The third trimester is possibly not only more biologically relevant a period of exposure to DBPs, but also a more stable period in which to measure exposure as more pregnant women are home-based later on in pregnancy.

In all, stratification by ethnicity, employment and smoking status is essential. Failing to take such factors into account in exposure assessments may result in differential misclassification error. Exposure based on imprecise individual water consumption estimates may be especially biased in the BiB cohort in which individual behaviour is likely to drive differences in exposure to HAAs as it does exposure to TCAA (Smith et al. 2012) and THMs (Smith 2011), because of the low variability in tap water DBP concentrations measured across the Bradford area.

This study has a number of limitations. Whilst employment status was updated at each questionnaire in this study, factors such as smoking, alcohol consumption, and physical exercise were only measured at baseline. Any changes in those behaviours over the course of late pregnancy could therefore not be accounted for. Also, due to the timing of the release of the BiB data extracts containing complete data on certain variables, in 2012 when this work was done, maternal weight, height, BMI and in particular gestational diabetes could not be included in these analyses.

Moreover, the subset of women who volunteered to participate in the RQS may not be fully representative of BiB as a whole. Firstly, there are some unavoidable differences between groups due to the design and eligibility criteria of the RQS: all women enrolled in the RQS were English-speakers, responded to the third version (of three) of the baseline questionnaire and had pregnancies in a similar period (80% in winter) compared to the rest of BiB mothers whose babies were born randomly over the whole BiB recruitment period 2007 to 2010. This explains why season was not a strong predictor of water consumption in this study, as per Forssen et al's findings (2007).

Secondly, a higher proportion of the subset was White British and a smaller proportion Pakistani than in the rest of the cohort (Table 8.1). The underrepresented Pakistani group exhibits higher rates of gestational diabetes and lower probabilities of smoking and consuming alcohol, factors which are potential confounders in the full epidemiologic study on fetal growth. The Pakistani group are also less likely to be in employed work which is known to increase the reliability of reported consumption patterns. The "Other" ethnic category closely tracks the White British group in terms of

demographics. Water consumption at baseline also differs significantly between this subset and the rest of the cohort. Indeed, this subset of women drank systematically more at baseline (0.63 L/day more on average total water) than the rest of the BiB cohort (Table A8 - 4). However, water consumption trends by ethnicity, employment and smoking status in this subset were representative of the rest of the cohort (Table A8 - 5, Table A8 - 6, Table A8 - 7). In all, there appears to be a non-random selection bias in the type of women who chose to volunteer to the RQS. This is manageable if information from the RQS is applied in a stratified way to the full cohort.

The variance partition coefficient (VPC), which represents the proportion of total variability due to between-subject variability ($\frac{\sigma_B^2}{\sigma_B^2+\sigma_W^2}$), can be used as measure of reliability of a given repeated measurement. There was a greater proportion of within-subject vs. between-subject variability in the subset for total tap and total water consumption. In contrast, both Barbone et al (2002) who compared changes in water use in Italian women late in their pregnancy and after delivery, and Shimokura et al (1998) found that the between-subject variation was larger than within-subject variation for total water intake (ICC=0.81 (ICC=0.42 for full-time employed women), and ICCs≥0.77, respectively). Forssen et al (2009) found that almost 60% of the variation in total intake was due to between-subject variation. Smith et al (2011) also found low within-subject variability for tap water intakes across combined locations and at home, but high within-subject variability for tap water intake specifically outside the home.

This being said, the average interval between questionnaires in this analysis was 37.6 (36.0, 39.2) days from baseline to Q1, and 39.6 (37.4, 41.7) days from Q1 to Q2. This means that at baseline women were asked to estimate their water drinking habits since pregnancy began, 185 days or so ago (26.5 ± 0.7 weeks), while at Q1 and Q2, recall was only expected over the past week of pregnancy, which may or may not have been representative of their current trimester of pregnancy. The recall time and the number of days between repeats in this study are different to previous studies (Kaur et al. 2004; Shimokura et al. 1998). This could explain why such considerable within-subject variability in behaviours was observed.

Taken together, these data suggest that women do increase their water consumption as pregnancy progresses. This could be explained by a number of reasons relating to perceived health benefits, or physiologic changes such as increased thirst (JM Wright et al. 2010), discomfort during late pregnancy, inactivity/boredom, increased food consumption, increased time spent at home—which is the location of most water consumption, even though the unemployed women typically drink less water than the employed. (On this point, I don't know whether it is the characteristics of the women who tend to be unemployed that explains their lower drinking consumption, or whether it is the fact of

remaining at home itself which is less conducive to drinking). However ultimately, with these study data, true behaviour change cannot be differentiated from measurement error. Asking women about their perception of change in drinking habits did not help us answer this question: women who said they did change habits over time, do not report drinking more or less than the ones who did not. Moreover, actual water change over time does not predict whether women report changing behaviour (Table 8.8).

The observed differences in water consumption over pregnancy, though varying by strata, represent consistent and marked trends, such that measurement error alone is unlikely to explain them. This observation ties in with the argument made in previous studies (Forssen et al. 2009; Kaur et al. 2004; Shimokura et al. 1998) that any large variation is more likely to reflect an actual change rather than inaccurate reporting. In addition, the unemployed women, who are thought to be more accurate in their reporting, were once more found to drink less water. Given that women stay at home more as pregnancy progresses, reporting should become increasingly accurate suggesting that any changes in water consumption reported are more likely to reflect real change.

On the other hand, water intake over a set period of time is always difficult to measure and may be subject to non-differential error (JM Wright et al. 2010). Faulty memory (a question such as "how much water did you drink, on average, a month ago?" is hard to answer accurately), lack of knowledge (the type of tap water consumed at a friend's house, or at a restaurant may not be known), and varying perceptions (e.g. what constitutes 200ml, i.e. a cup/glass's worth) may explain the noise inherent in questionnaire responses (Savitz 2012). In addition, none of the water questionnaires asked about other possible uses of drinking water such as in the making of soups or meals, or via secondary pathways of exposure (pharmaceuticals, occupation, etc.) (Arbuckle et al. 2002), all of which might increase the potential for inaccuracies in women's responses.

In addition, the true comparability of the water consumption values derived from the baseline questionnaire, which was administered in an interview setting, compared to the repeat questionnaires, which were mailed out and completed in each participant's home, is not guaranteed. Indeed, questionnaires already tend to slightly overestimate tap water intakes (Barbone et al. 2002; Kaur et al. 2004; Maskiell et al. 2006; Shimokura et al. 1998; Smith et al. 2012) by up to 0.41 L/day (95% CI: 0.13, 0.69) according to Smith et al (2012). As such, answering a repeat questionnaire focused exclusively on water consumption (in contrast to the baseline questionnaire which included many questions on other lifestyle habits etc.), could have led RQS volunteers to report their water consumption in the repeat questionnaires differently to the baseline but whether this would have led to more overestimation is impossible to know for certain. There is no reason to suspect a systematic trend of increasing overestimation with each additional questionnaire.

In conclusion, relying on second trimester estimates to calculate HAA exposure metrics will likely underestimate critical third trimester exposures leading to attenuation of risk estimates. In order to cover both the behaviour change and the measurement error explanations for this observed increase in water consumption over time, findings from the RQS should be incorporated into epidemiologic study to improve the accuracy of risk estimates. This can be achieved by stratified imputation of consumption in the third trimester and running an epidemiologic model including a measurement error component (Richardson and Gilks 1993). I also recommend prospective collection of information on maternal lifestyle factors during more than one trimester of pregnancy. This will better capture likely true changes in maternal behaviour and improve our ability to understand their impacts on the fetus at critical periods of development.

8.5 Tables

Table 8.1: Demographic and behavioural characteristics of the RQS subset and the rest of the cohort, as reported at enrolment to BiB

		RQS			Rest of the cohort			Pearson chi^2/Fisher exact test
variables	**categories**	**no.**	**%**	**n**	**no.**	**%**	**n**	
Demographic variables								
maternal age	<20 years	14	5.5	254	616	5.5	11,121	0.003
	20-24 years	57	22.4		2,791	25.1		
	25-29 years	59	23.2		3,636	32.7		
	30-34 years	82	32.3		2,595	23.3		
	35-39 years	34	13.4		1,212	10.9		
	≥40 years	8	3.1		271	2.4		
	missing				2,398			
self-reported ethnicity	White British	123	48.4	254	4,365	39.3	11,093	<0.001
	Pakistani	74	29.1		5,053	45.6		
	Other	57	22.4		1,675	15.1		
	missing				2,426			
current employment status	employed	142	55.9	254	4,309	38.7	11,129	<0.001
	employed but on maternity/sick leave	13	5.1		432	3.9		
	not working	90	35.4		6,012	54.0		
	full-time student	9	3.5		376	3.4		
	missing				2,390			
marital status	married or re-married	165	65.0	254	7,598	68.3	11,132	0.536
	single	83	32.7		3,289	29.5		
	separated, divorced or widowed	6	2.4		245	2.2		
	missing				2,387			
cohabitation status	living with baby's father or another partner	219	86.2	254	9,286	83.5	11,122	0.246
	not living with a partner	35	13.8		1,836	16.5		
	missing				2,397			
parity	no previous children	122	50.6	241	4,979	39.4	12,634	0.001
	1 previous child	66	27.4		3,662	29.0		
	2 or more previous children	53	22.0		3,993	31.6		
	missing	13			885			

variables	categories	RQS no.	RQS %	RQS n	Rest of the cohort no.	Rest of the cohort %	Rest of the cohort n	Pearson chi^2/Fisher exact test
gestational diabetes (developed	No	109	88.6	123	3,758	79.9	4,706	0.016
during this pregnancy)	Yes	14	11.4		948	20.1		
	missing	131			8,813			
maternal physical exercise #	None	132	52.0	254	6,990	74.8	9,348	<0.001
	< 1 hour	71	28.0		1,233	13.2		
	≥ 1 hour but < 3 hours	42	16.5		901	9.6		
	≥ 3 hours	9	3.5		224	2.4		
	missing				4,171			
physical activity involved in mother's	not in paid employment	103	40.6	254	5,285	56.5	9,348	<0.001
paid work	sitting most of the time	72	28.3		1,916	20.5		
	standing or walking most of the time	53	20.9		1,599	17.1		
	(vigorous) physical activity involved	26	10.2		548	5.9		
	missing				4,171			
season at birth of child	winter (jan-mar)	198	79.5	249	3,150	24.0	13,113	<0.001
	spring (apr-jun)	7	2.8		2,875	21.9		
	summer (jul-sep)	0	0.0		3,616	27.6		
	autumn (oct-dec)	44	17.7		3,472	26.5		
	missing	5			406			
Behavioural variables								
maternal smoking status	Currently a smoker	25	9.8	254	1,550	13.9	11,121	<0.001
	Ever a smoker	64	25.2		1,824	16.4		
	Never a smoker	165	65.0		7,747	69.7		
	missing				2,398			
maternal exposure to second hand	No	168	66.1	254	7,546	68.1	11,076	<0.001
smoke during pregnancy	Yes, >1 hour a day	18	7.1		1,518	13.7		
	Yes, <1 hour a day	68	26.8		2,012	18.2		
	missing				2,443			
maternal alcohol consumption during	No	214	84.9	252	9,046	84.0	10,769	0.018*
pregnancy	Yes	38	15.1		1,723	16.0		
	missing	2			6			
caffeinated drinks consumption by	No (0 cups)	247	97.2	254	4,061	30.0	13,519	<0.001
mother during pregnancy	Yes (1 or more cups per day)	7	2.8		9,458	70.0		
	missing				2,744			
maternal drug usage during	No	253	99.6	254	9,368	98.7	9,494	0.285*

variables	categories	RQS no.	RQS %	RQS n	Rest of the cohort no.	Rest of the cohort %	Rest of the cohort n	Pearson chi^2/Fisher exact test
pregnancy	Yes	1	0.4		125	1.3		
	Don't remember	0	0.0		1	0.0		
	missing				4,025			
Socio-economic variables								
total members of household	0, 1, 2 members	93	36.6	254	2,805	25.2	11,134	<0.001
	3 members	62	24.4		2,708	24.3		
	4 members	46	18.1		1,957	17.6		
	5 or more members	53	20.9		3,664	32.9		
	missing				2,385			
maternal education level^	None	31	12.2	254	2,422	21.8	11,112	<0.001
	School	55	21.7		3,433	30.9		
	Further	50	19.7		1,594	14.3		
	Higher)	97	38.2		2,815	25.3		
	Other	21	8.3		848	7.6		
	missing				2,407			
paternal education level^	None	36	14.2	254	1,704	15.3	11,104	0.493
	School	63	24.8		2,663	24.0		
	Further	23	9.1		1,143	10.3		
	Higher	75	29.5		2,786	25.1		
	Other	57	22.4		2,808	25.3		
	missing				2,415			
IMD 2010 score (not derived from questionnaire)	1st quintile (least deprived)	143	56.3	254	7,411	66.6	11,121	0.001*
	2nd quintile	64	25.2		1,986	17.9		
	3rd quintile	27	10.6		1,220	11.0		
	4th quintile	15	5.9		319	2.9		
	5th quintile (most deprived)	5	2.0		185	1.7		
	missing				2,398			

*Fisher's exact test
^ see Table 2.2
# hours in the past week spent swimming, jogging, doing aerobics, playing tennis, exercising at the gym, etc.
Red signal a p-value<0.05

Table 8.2: Univariate linear mixed models of RQS women's a) total tap water ("TAP") and b) total water ("Total") consumption (in L/day) over the 3 time points of interest: baseline, Q1 (30-33 weeks of pregnancy) and Q2 (36-39 weeks of pregnancy) (Model 1) (number of observations over 3 time points=762, number of missing observations=127)

a. TAP	Model 1	
Fixed effect		
Main intercept	1.86	(1.71, 2.01)
Time pattern	0.13	(0.04, 0.23)
Random effect		
Random intercepts SD	0.76	(0.65, 0.89)
Random slopes SD	0.09	(0.00, 0.25)
Measurement error SD	1.01	(0.94, 1.09)
Variance Partition Coefficient	0.37	(0.28, 0.46)

b. Total	Model 1	
Fixed effect		
Main intercept	2.22	(2.07, 2.37)
Time pattern	0.14	(0.03, 0.26)
Random effect		
Random intercepts SD	0.73	(0.58, 0.87)
Random slopes SD	0.29	(0.07, 0.47)
Measurement error SD	1.08	(1.00, 1.16)
Variance Partition Coefficient	0.35	(0.26, 0.44)

Red means significant (i.e. 95% credible interval does not cross zero).

Table 8.3: Adjusted linear mixed models of RQS women's a) total tap water ("TAP") and b) total water ("Total") consumption trend (in L/day) over the 3 time points of interest: baseline, Q1 and Q2 (Model 2) (number of observations over 3 time points=762, number of missing observations=127)

Model 2		a. TAP		b. Total	
Fixed effect					
Main intercept		1.94	(1.50, 2.37)	2.40	(1.95, 2.85)
Time pattern		0.19	(0.08, 0.30)	0.22	(0.09, 0.35)
Ethnicity	White British	ref		ref	
	Pakistani	-0.15	(-0.50, 0.20)	-0.22	(-0.58, 0.14)
	Other	-0.38	(-0.70, -0.06)	-0.35	(-0.67, -0.02)
Smoking status at baseline	Never	ref		ref	
	Currently	0.47	(0.02, 0.92)	0.34	(-0.13, 0.80)
	Ever	0.21	(-0.12, 0.53)	-0.01	(-0.34, 0.32)
Employment status (over time)**	Employed and currently working	ref		ref	
	Employed and on maternity leave	-0.09	(-0.46, 0.29)	-0.08	(-0.49, 0.34)
	Not employed	-0.26	(-0.56, 0.03)	-0.42	(-0.75, -0.09)
	Full-time student	-0.06	(-0.75, 0.63)	-0.23	(-0.96, 0.49)
Maternal physical exercise for leisure	None	ref		ref	
	Some but <1 hour	0.32	(0.03, 0.60)	0.43	(0.14, 0.72)
	≥1 hour but <3 hours	0.25	(-0.11, 0.62)	0.29	(-0.08, 0.67)
	≥3 hours	0.19	(-0.55, 0.93)	0.51	(-0.24, 1.28)
Maternal physical exercise in paid work	Not in paid work	ref		ref	
	Mostly sitting at work	-0.10	(-0.50, 0.29)	0.04	(-0.38, 0.46)
	Mostly standing/walking at work	-0.41	(-0.82, -0.01)	-0.30	(-0.73, 0.12)
	(Vigorous) physical activity	-0.13	(-0.63, 0.38)	-0.29	(-0.81, 0.23)
Total household members	0, 1, 2 members in the household	ref		ref	
	3 members	-0.06	(-0.38, 0.25)	-0.20	(-0.53, 0.13)
	4 members	0.34	(-0.02, 0.70)	0.26	(-0.11, 0.63)
	≥5 members	-0.04	(-0.40, 0.33)	-0.09	(-0.47, 0.29)
Random effect					
Random intercepts SD		0.72	(0.59, 0.84)	0.67	(0.52, 0.82)
Random slopes SD		0.10	(0.00, 0.26)	0.32	(0.03, 0.50)
Measurement error SD		1.01	(0.94, 1.08)	1.06	(0.98, 1.15)
Variance Partition Coefficient		0.34	(0.25, 0.43)	0.33	(0.24, 0.43)

**N=145 missing (18 of which have outcomes, 127 of which do not such that their imputation does not add anything to the results), missing data (assumed to be missing at random) were imputed based on a probability estimated from the observed data

Red means significant (i.e. 95% credible interval does not cross zero).

Table 8.4: Linear mixed models of RQS women's a) total tap water ("TAP") and b) total water ("Total") consumption trend (in L/day) over the 3 time points of interest: baseline, Q1 and Q2 allowing for random slopes (intercepts) by ethnic group (Model 3) (number of observations over 3 time points=762, number of missing observations=127)

Model 3		a. TAP		b. Total	
Fixed effect					
Main intercept		1.95	(1.52, 2.38)	2.39	(1.95, 2.82)
Time trend by ethnicity	White British	0.08	(-0.07, 0.23)	0.07	(-0.09, 0.24)
	Pakistani	0.37	(0.14, 0.62)	0.49	(0.19, 0.80)
	Other	0.17	(-0.04, 0.40)	0.20	(-0.02, 0.43)
Ethnicity	White British	ref		ref	
	Pakistani	-0.37	(-0.76, 0.02)	-0.50	(-0.87, -0.12)
	Other	-0.44	(-0.80, -0.08)	-0.42	(-0.78, -0.07)
Smoking status at baseline	Never	ref		ref	
	Currently	0.43	(-0.01, 0.86)	0.28	(-0.15, 0.72)
	Ever	0.21	(-0.11, 0.53)	0.08	(-0.25, 0.41)
Employment status (over time)**	Employed and currently working	ref		ref	
	Employed and on maternity leave	-0.02	(-0.39, 0.35)	0.03	(-0.36, 0.42)
	Not employed	-0.21	(-0.51, 0.09)	-0.36	(-0.68, -0.05)
	Full-time student	0.01	(-0.66, 0.67)	-0.16	(-0.80, 0.49)
Maternal physical exercise for leisure	None	ref		ref	
	Some but <1 hour	0.25	(-0.02, 0.53)	0.32	(0.05, 0.59)
	≥1 hour but <3 hours	0.24	(-0.12, 0.59)	0.24	(-0.10, 0.59)
	≥3 hours	0.02	(-0.65, 0.70)	0.24	(-0.42, 0.91)
Maternal physical exercise in paid work	Not in paid work	ref		ref	
	Mostly sitting at work	-0.01	(-0.39, 0.38)	0.16	(-0.23, 0.55)
	Mostly standing/walking at work	-0.40	(-0.78, -0.02)	-0.22	(-0.61, 0.16)
	(Vigorous) physical activity	-0.01	(-0.50, 0.49)	-0.14	(-0.63, 0.36)
Total members in household	0, 1, 2 members in the household	ref		ref	
	3 members	-0.03	(-0.34, 0.27)	-0.18	(-0.48, 0.13)
	4 members	0.39	(0.05, 0.73)	0.34	(0.00, 0.68)
	≥5 members	0.01	(-0.34, 0.36)	0.01	(-0.33, 0.34)
Random effect					
Random intercepts SD	White British	0.82	(0.65, 1.00)	0.85	(0.67, 1.04)
	Pakistani	0.55	(0.11, 0.89)	0.28	(0.01, 0.65)
	Other	0.42	(0.07, 0.69)	0.27	(0.01, 0.57)
Random slopes SD	White British	0.11	(0.01, 0.29)	0.21	(0.02, 0.43)
	Pakistani	0.58	(0.07, 1.00)	0.99	(0.70, 1.30)
	Other	0.13	(0.00, 0.35)	0.14	(0.01, 0.35)
Measurement error SD		0.98	(0.90, 1.06)	1.01	(0.93, 1.09)
Variance Partition Coefficient	White British	0.42	(0.31, 0.53)	0.43	(0.32, 0.54)
	Pakistani	0.43	(0.27, 0.58)	0.52	(0.37, 0.65)
	Other	0.18	(0.03, 0.35)	0.10	(0.00, 0.27)

Red means significant (i.e. 95% credible interval does not cross zero).

Table 8.5: Linear mixed models of RQS women's total tap water ("TAP") (**a**) and total water ("Total") (**b**) consumption trend (in L/day) over the 3 time points of interest: baseline, Q1 and Q2 with interaction between employment and time (Model 4) (number of observations over 3 time points=762, number of missing observations=127)

Model 4		a. TAP		b. Total	
Fixed effect					
Main intercept		1.94	(1.48, 2.39)	2.39	(1.92, 2.87)
Employment status x Time Interaction	Employed and currently working	0.21	(0.00, 0.41)	0.26	(0.03, 0.49)
	Employed and on maternity leave	-0.06	(-0.46, 0.34)	-0.20	(-0.64, 0.23)
	Not employed	0.22	(0.06, 0.38)	0.26	(0.08, 0.44)
	Full-time student	0.16	(-0.47, 0.82)	0.41	(-0.32, 1.14)
Ethnicity	White British	ref		ref	
	Pakistani	-0.16	(-0.50, 0.19)	-0.23	(-0.59, 0.13)
	Other	-0.39	(-0.71, -0.07)	-0.37	(-0.70, -0.05)
Smoking status at baseline	Never	ref		ref	
	Currently	0.47	(0.02, 0.91)	0.34	(-0.12, 0.79)
	Ever	0.20	(-0.12, 0.52)	-0.01	(-0.34, 0.32)
Employment status (over time)* *	Employed and currently working	ref		ref	
	Employed and on maternity leave	0.22	(-0.38, 0.82)	0.42	(-0.22, 1.06)
	Not employed	-0.29	(-0.67, 0.09)	-0.44	(-0.86, -0.03)
	Full-time student	-0.03	(-0.84, 0.77)	-0.30	(-1.13, 0.53)
Maternal physical exercise for leisure	None	ref		ref	
	Some but <1 hour	0.32	(0.03, 0.60)	0.43	(0.14, 0.71)
	≥1 hour but <3 hours	0.24	(-0.12, 0.61)	0.28	(-0.09, 0.65)
	≥3 hours	0.18	(-0.55, 0.92)	0.51	(-0.24, 1.26)
Maternal physical exercise in paid work	Not in paid work	ref		ref	
	Mostly sitting at work	-0.11	(-0.51, 0.28)	0.03	(-0.40, 0.45)
	Mostly standing/walking at work	-0.42	(-0.83, -0.01)	-0.31	(-0.75, 0.12)
	(Vigorous) physical activity	-0.15	(-0.65, 0.35)	-0.32	(-0.85, 0.21)
Total members in household	0, 1, 2 members in the household	ref		ref	
	3 members	-0.06	(-0.38, 0.26)	-0.19	(-0.51, 0.13)
	4 members	0.34	(-0.01, 0.70)	0.26	(-0.10, 0.63)
	≥5 members	-0.03	(-0.40, 0.34)	-0.08	(-0.46, 0.30)
Random effect					
Random intercepts SD		0.71	(0.59, 0.84)	0.66	(0.49, 0.81)
Random slopes SD		0.10	(0.00, 0.26)	0.34	(0.11, 0.50)
Measurement error SD		1.01	(0.94, 1.09)	1.06	(0.98, 1.15)
Variance Partition Coefficient		0.34	(0.25, 0.43)	0.33	(0.24, 0.42)

Red means significant (i.e. 95% credible interval does not cross zero).

Table 8.6: Linear mixed models of RQS women's total tap water ("TAP") (**a**) and total water ("Total") (**b**) consumption trend (in L/day) over the 3 time points of interest: baseline, Q1 and Q2 with interaction between smoking status and time (Model 5) (number of observations over 3 time points=762, number of missing observations=127)

Model 5		a. TAP		b. Total	
Fixed effect					
Main intercept		1.92	(1.48, 2.36)	2.39	(1.93, 2.84)
Smoking status x Time Interaction	Never	0.22	(0.09, 0.36)	0.26	(0.11, 0.41)
	Currently	0.00	(-0.35, 0.34)	-0.08	(-0.46, 0.30)
	Ever	0.17	(-0.04, 0.39)	0.25	(0.01, 0.49)
Ethnicity	White British	ref		ref	
	Pakistani	-0.14	(-0.50, 0.21)	-0.22	(-0.58, 0.15)
	Other	-0.37	(-0.70, -0.05)	-0.35	(-0.68, -0.03)
Smoking status at baseline	Never	ref		ref	
	Currently	0.66	(0.13, 1.19)	0.58	(0.04, 1.11)
	Ever	0.25	(-0.13, 0.63)	0.00	(-0.38, 0.38)
Employment status (over time)**	Employed and currently working	ref		ref	
	Employed and on maternity leave	-0.10	(-0.49, 0.28)	-0.09	(-0.51, 0.32)
	Not employed	-0.27	(-0.58, 0.03)	-0.45	(-0.77, -0.11)
	Full-time student	-0.08	(-0.76, 0.61)	-0.25	(-0.98, 0.47)
Maternal physical exercise for leisure	None	ref		ref	
	Some but <1 hour	0.32	(0.04, 0.60)	0.43	(0.14, 0.72)
	≥1 hour but <3 hours	0.25	(-0.12, 0.61)	0.29	(-0.09, 0.67)
	≥3 hours	0.18	(-0.56, 0.91)	0.50	(-0.25, 1.25)
Maternal physical exercise in paid work	Not in paid work	ref		ref	
	Mostly sitting at work	-0.11	(-0.51, 0.28)	0.03	(-0.39, 0.44)
	Mostly standing/walking at work	-0.41	(-0.82, -0.01)	-0.31	(-0.73, 0.11)
	(Vigorous) physical activity	-0.13	(-0.63, 0.37)	-0.29	(-0.81, 0.22)
Total members in household	0, 1, 2 members in the household	ref		ref	
	3 members	-0.06	(-0.39, 0.26)	-0.20	(-0.52, 0.13)
	4 members	0.34	(-0.01, 0.70)	0.27	(-0.10, 0.63)
	≥5 members	-0.04	(-0.41, 0.33)	-0.08	(-0.46, 0.29)
Random effect					
Random intercepts SD		0.72	(0.60, 0.84)	0.67	(0.51, 0.82)
Random slopes SD		0.10	(0.00, 0.26)	0.33	(0.11, 0.50)
Measurement error SD		1.01	(0.94, 1.08)	1.06	(0.98, 1.15)
Variance Partition Coefficient		0.34	(0.25, 0.43)	0.34	(0.24, 0.43)

Red means significant (i.e. 95% credible interval does not cross zero)

Table 8.7: Differences in total water consumption between each (interaction) category and the reference category, at baseline, at Q1 and at Q2.
All models are adjusted for ethnicity, smoking status, employment status, maternal physical exercise (leisure and in paid work), total members of household.
1 cup = 1 glass = 200ml
A positive difference means the reference group consumes more.

		at baseline		at Q1		at Q2	
		Mean	95% CrI	Mean	95% CrI	Mean	95% CrI
Model 3	White British	ref		ref		ref	
	Pakistani	0.50	(0.12, 0.87)	0.08	(-0.32, 0.48)	-0.34	(-0.98, 0.30)
	Other	0.42	(0.07, 0.78)	0.30	(0.01, 0.59)	0.18	(-0.23, 0.59)
Model 4	Employed and working	ref		ref		ref	
	Employed and on maternity leave	-0.42	(-1.06, 0.22)	0.05	(-0.38, 0.48)	0.52	(-0.14, 1.19)
	Not employed	0.44	(0.03, 0.86)	0.45	(0.11, 0.78)	0.45	(-0.03, 0.92)
	Full-time student	0.30	(-0.53, 1.13)	0.16	(-0.66, 0.97)	0.02	(-1.34, 1.35)
Model 5	Never smoker	ref		ref		ref	
	Current smoker	-0.58	(-1.11, -0.04)	-0.24	(-0.71, 0.25)	0.11	(-0.60, 0.81)
	Ever smoker	0.00	(-0.38, 0.38)	0.01	(-0.33, 0.35)	0.02	(-0.46, 0.49)

Red means significant (i.e. 95% credible interval does not cross zero).

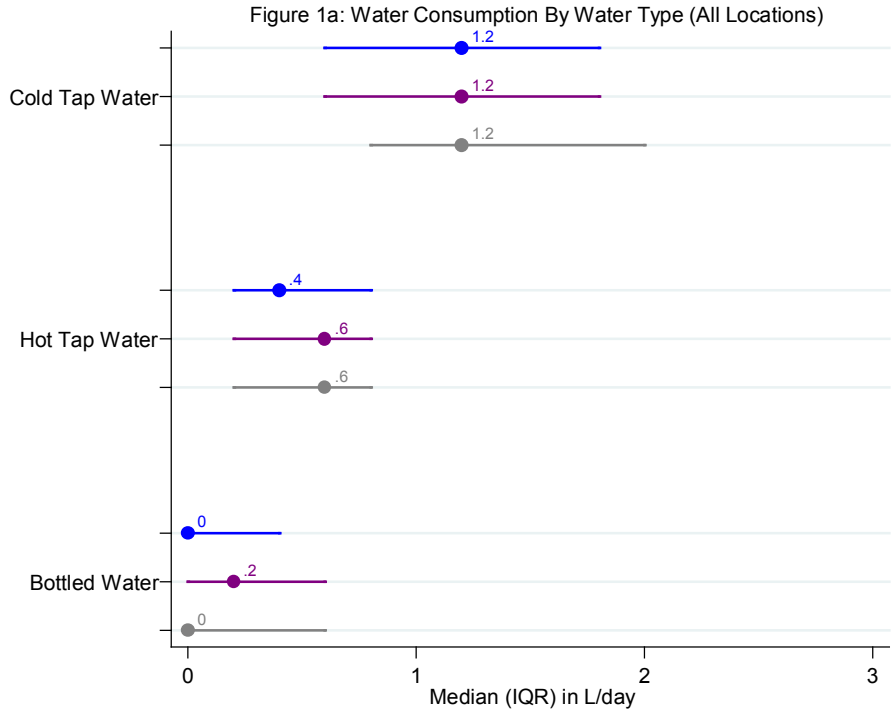
Table 8.8: Comparison between women's self-perception of change in total tap water drinking habits, vs. quantitative water consumption actually reported

Thinks her tap water drinking habits have changed since previous questionnaire	Difference between baseline and Q1				Difference between Q2 and Q1			
	mean	(95% CI)	p-value*	n	mean	(95% CI)	p-value*	n
yes (changed)	0.24	(-0.22, 0.71)	0.293	28	-0.22	(-0.85, 0.41)	0.478	31
no (did not change)	0.18	(-0.08, 0.45)	0.170	171	0.04	(-0.20, 0.28)	0.742	92
missing	0.24			10	1.60			4
total	0.19	(-0.03, 0.42)	0.090	209	0.03	(-0.22, 0.27)	0.839	127
difference between yes and no categories**	-0.06	(-0.74, 0.62)	0.865		0.26	(-0.28, 0.81)	0.342	

*t test: Is the difference in total tap water consumption different from 0? (stratified by yes/no self-perceived a change in behaviour)
**2 sample t-test

8.6 Figures

Figure 8.1 a and b: Line plots summarising and comparing water consumption medians (by water type) at baseline (26-28 weeks of pregnancy), Q1 (30-33 weeks of pregnancy) and Q2 (36-39 weeks of pregnancy) of the women enrolled in the Repeat Questionnaire Study (RQS) (N=254)



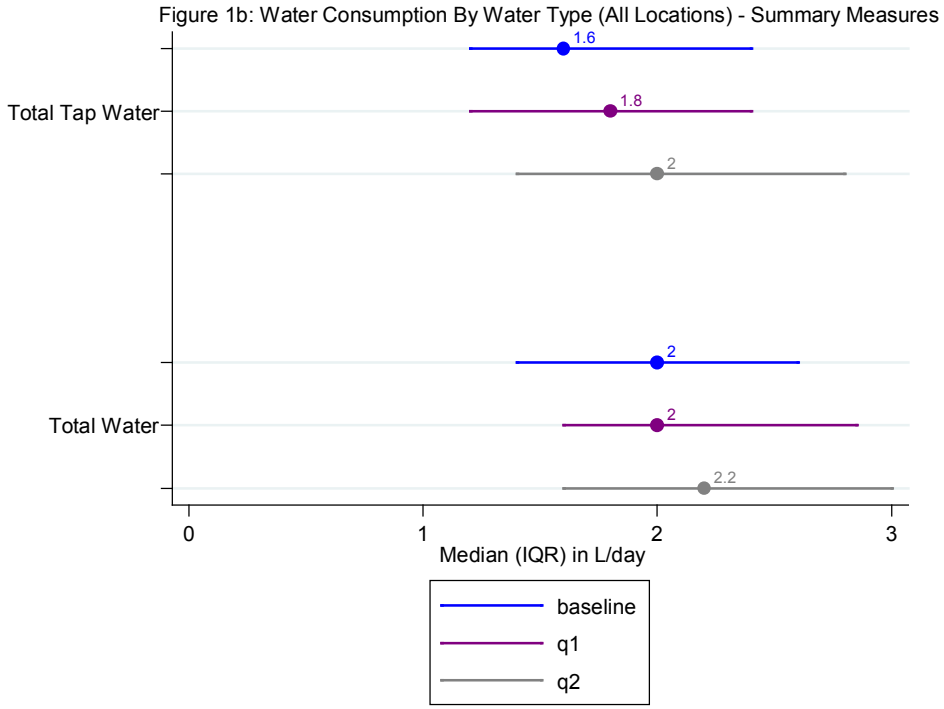
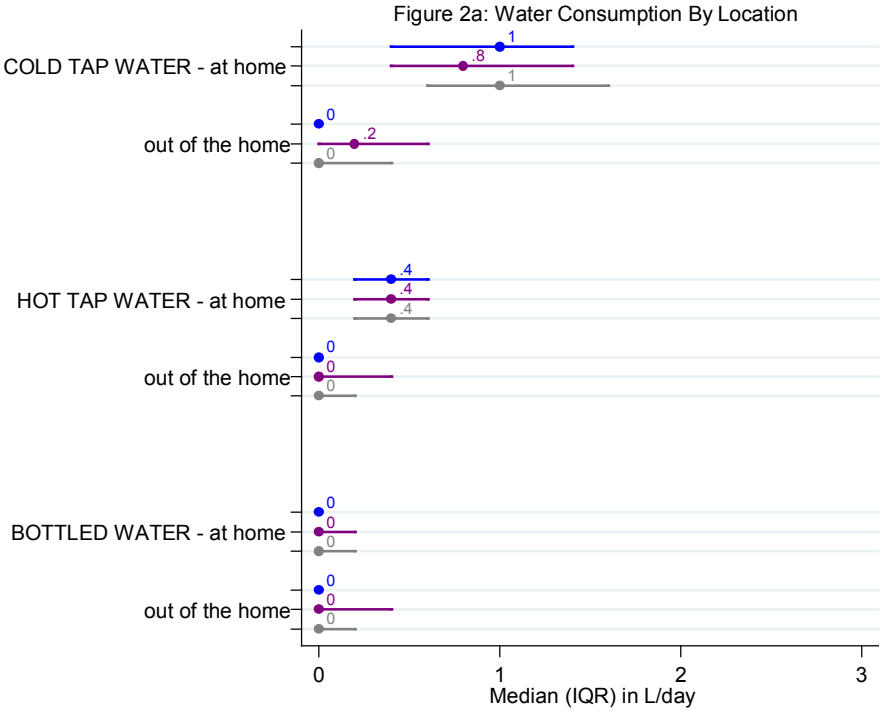


Figure 8.2 a and b: Line plots summarising and comparing water consumption medians (by location) at baseline (26-28 weeks of pregnancy), Q1 (30-33 weeks of pregnancy) and Q2 (36-39 weeks of pregnancy) of the women enrolled in the Repeat Questionnaire Study (RQS) (N=254)



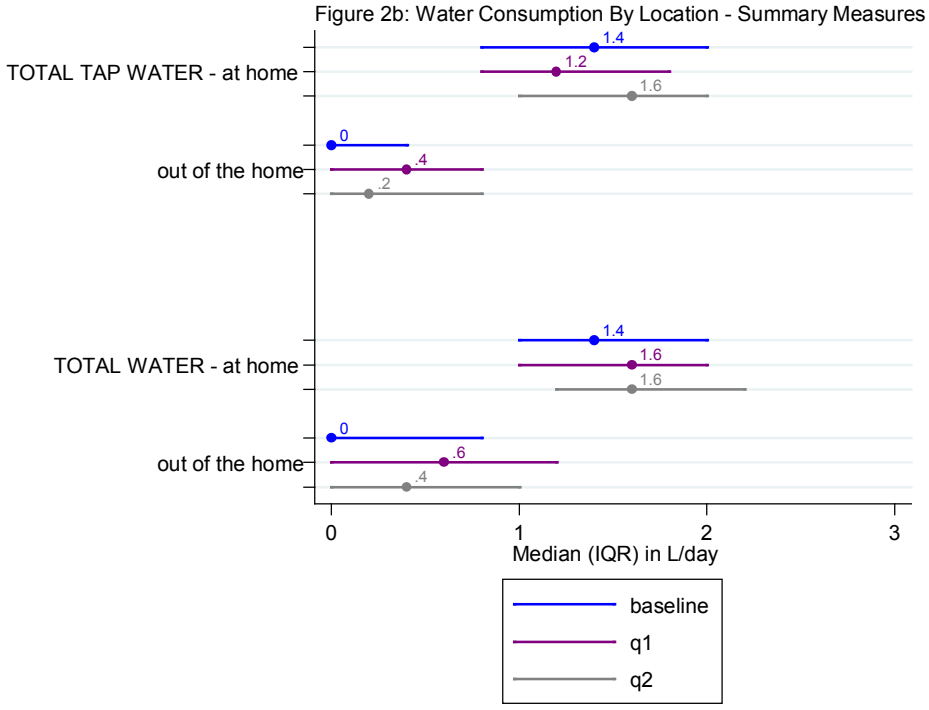
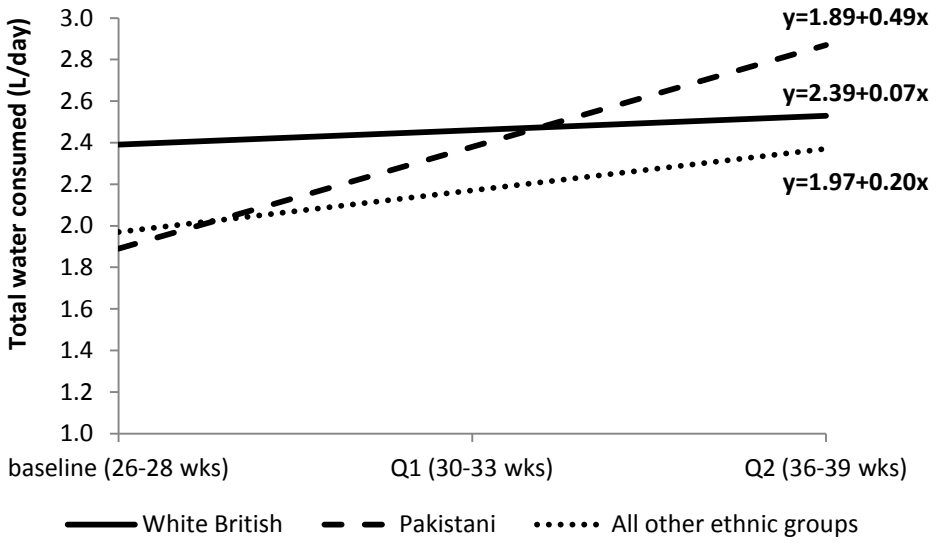
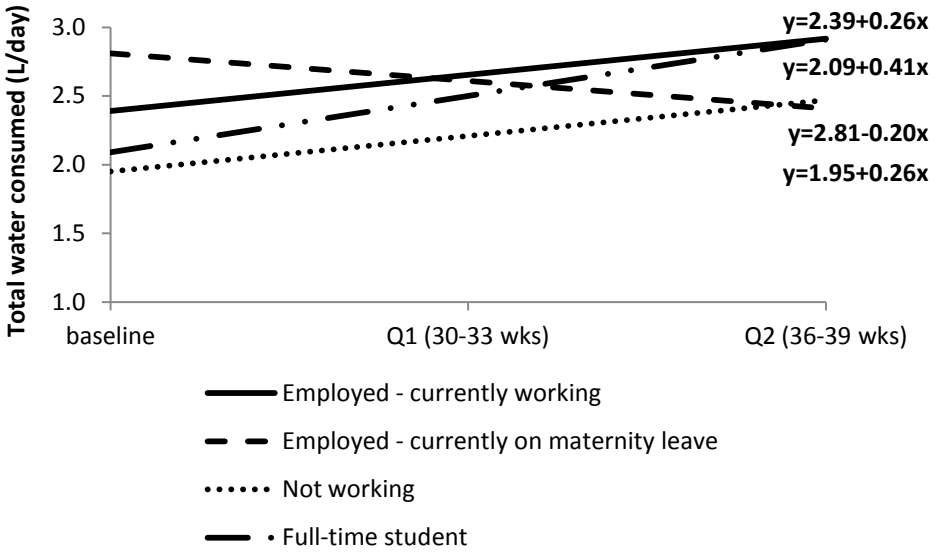


Figure 8.3: Interaction plots by category for total water consumption

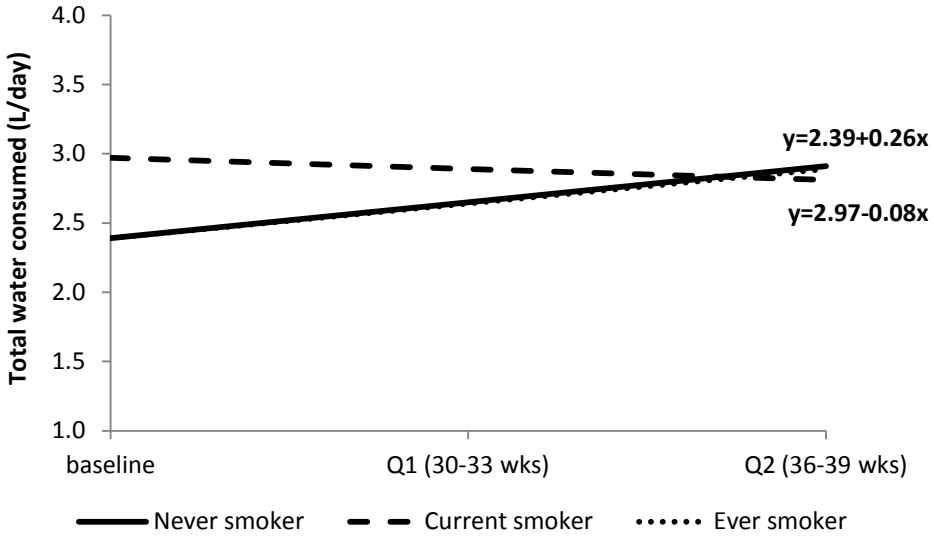
a. Model 3: Interaction between ethnicity and time, allowing for random slopes (and intercepts) by ethnic group



b. Model 4: Interaction between employment status over time and time



c. Model 5: Interaction between smoking status at baseline and time



CHAPTER 9 DISCUSSION

This chapter concludes the thesis by summarising briefly the main results by aim (Thesis Aims, Chapter 2), discussing the interpretation of these results in a wider context, and offering some recommendations for future research.

9.1 Aim 1

9.1.1 Summary

The first aim of this thesis was to generate exposure estimates to ingested dichloroacetic acid (DCAA), trichloroacetic acid (TCAA) and bromodichloroacetic acid (BDCAA) for the pregnancy period of each BiB participant. To do so, each individual woman's total tap water consumption during pregnancy was combined in turn with modelled area-level DCAA, TCAA and BDCAA concentrations, and weighted to each woman's specific trimester based on postcode of residence. Boiling and filtering factors were taken into account in this calculation (Table 5.1).

Women's average tap water consumption was 1.7 ± 0.9 L/day, the majority of which is cold tap water consumed in the home (Chapter 3). 158 DCAA and TCAA valid data points and 143 valid BDCAA data points were collected from eight water supply zones in Bradford; their medians were 9.90ug/L (IQR=9.05), 10.70ug/L (IQR=7.05) and 1.20ug/L (IQR=1.43) respectively, with DCAA and TCAA together representing the largest proportion of all HAAs by mass (Table 4.5). Though not total HAAs, the sum of these three HAAs is below the US's maximum contaminant level of 60ug/L for HAA5 and below the 80ug/L standard for HAA9 proposed by the European Union's Drinking Water Directive (DHI 2008). Bayesian hierarchical models were fit to predict mean DCAA, TCAA, and BDCAA concentration levels by water supply zone and time. Averaging and weighting daily HAA predictions to correspond to each woman's first, second or third trimester of pregnancy resulted in plausible values which were similar for women's three trimesters of pregnancy (Table 4.8). Water consumption and HAA occurrence reported in Bradford compared well to (albeit quite variable) data from previous studies (Table 3.1 and Table 4.6).

Correcting total tap water consumption for reduction and/or boiling factors as appropriate, combined DCAA, TCAA and BDCAA metrics were 16.02 ± 10.7 ug/day, 18.15 ± 11.9 ug/day, and 1.80 ± 1.4 ug/day, respectively (Table 5.5). The three combined metric calculations included the same total tap water consumption component for each woman. They were calculated using the average HAA concentrations corresponding to each woman's second trimester of pregnancy, to be as faithful to the

time of baseline questionnaire administration as possible. As such, these levels are assumed to represent second trimester exposures. It is the differences in filtering and in particular boiling factors by HAA that lead to proportionally larger overall combined exposures to DCAA than to TCAA when compared to DCAA and TCAA mean observed and modelled concentrations.This is the first study of this magnitude to produce such a combined metric for HAAs.

9.1.2 Discussion and recommendations

With the focus of this thesis on DCAA, TCAA, and BDCAA, I ask whether DCAA, TCAA, and BDCAA are good proxies for total HAAs (HAA9 or HAA5) in a given sample. Given that DCAA and TCAA dominate by mass, and BDCAA was the most prevalent brominated species in this sample, I think they are. As mentioned in Chapter 7, the combined metrics for DCAA, TCAA and BDCAA were summed in a final attempt to detect an effect on birth outcomes, but this made no difference to the results, i.e. no association was found between the sum of DCAA, TCAA and BDCAA on either birth weight, term LBW or SGA.

I then turn to the question of whether total HAAs are good surrogates for all the other DBPs present in the sample. This represents an increasing challenge of exposure assessment in epidemiologic studies of DBPs given that over 600 DBPs have been identified (Richardson et al. 2007). Given evolving water treatment strategies with alternative disinfectants to chlorine, there is also a growing interest in emerging (unregulated) DBPs (e.g. nitrosamines, the most toxic emerging DBP). Many of these DBPs are not examined for toxicity nor typically measured in drinking water supplies. At this stage, I don't believe that HAAs have been established to be good proxies for other DBPs present in the sample.

This is relevant to exposure assessment of epidemiologic studies because finding an association between brominated HAA species for example (thought to be the most toxic of all HAA species (Plewa et al. 2010)) and a particular adverse health effect may not mean that brominated HAAs are the cause. This effect may in fact be due to the presence of other brominated DBPs of high health concern (e.g. haloacetonitriles, halonitromethanes, haloacetaldehydes) which are often not measured but which correlate with brominated HAAs. Predicting which DBPs will be present in any given water sample is a tremendous challenge because there are so many determinants of DBPs (NOM content which is affected by water source, treatment methods, treatment distribution, and season all affect DBP formation), and their concentrations once determined may be unique to each location and its own unique set of conditions. The relationship between specific DBPs (as well as their interactions in a specific context) must be very well studied before being able to assess proxy status. In BiB, the relationship between the THMs modelled by Drs Rachel Smith and James Bennett, and the HAAs

described in this thesis could be further studied in the short term. But indeed the nature of that relationship would likely only apply specifically to Bradford.

The few studies available to date suggest that mixture effects may be complex and unpredictable (e.g. inhibitory, additive, synergistic) (Komulainen 2004). Going forward, a detailed assessment of specific DBP mixtures will be required to provide a better understanding of any observed results (Nieuwenhuijsen et al. 2009b). If this is not done, one could erroneously be concluding that none of the DBPs that HAAs could be surrogates for in this study for example have any effect on birth outcomes either.

Future work could also incorporate the uncertainty from area-level modelling into the combined metric and its subsequent epidemiologic modelling.

9.2 Aim 2

9.2.1 Summary

This thesis's second aim was to examine the epidemiologic association between prenatal exposure to DCAA, TCAA and BDCAA as estimated under Aim 1 and birth weight, term low birth weight (LBW), and small-for-gestational age (SGA) as measures of fetal growth, adjusting the analyses for potential confounders.

Chapter 6 describes the outcome variables in the BiB cohort, and delves into some of the differences in birth weight to be expected by ethnicity: Pakistani babies were expected to be and are lighter at birth than White British babies. The chapter also justifies why continuous birth weight, term LBW and SGA based on Cole et al's standardised birth weights method, each with its own limitations, are the best birth weight-based indicators available to try and isolate effects on fetal growth. Finally, with regard to BiB's representativeness, Chapter 6 exposes that a) late second trimester recruitment to BiB meant that no combined metric exposure estimates for early pregnancy time periods could be measured, and no early deliveries could be recorded and included in BiB; and b) women with very poor birth outcomes may have been artificially excluded from the cohort (as they may not have wished to continue participating, or because disclosure of specific information may have compromised participants' anonymity). However, these threats to validity were deemed minimal as the study sample is otherwise so large.

Chapter 7 works through in detail the association between the combined DCAA, TCAA and BDCAA metrics of exposure (as well as its components: maternal water consumption, and area-level DCAA,

TCAA and BDCAA concentrations by place of residence) and risk of adverse birth outcomes in the BiB cohort. No clear association between the combined metric of exposure and either continuous birth weight, term LBW or SGA was found. The significant associations reported for the adjusted comparison of the highest versus lowest tertile of exposure to combined DCAA, TCAA and BDCAA and term LBW (complete case analyses) were contrary to hypothesis in terms of direction, and lost after multiple imputation. As a result, they were likely chance results given the large number of analyses conducted. Several stratifications and sensitivity analyses were carried out but the results of the combined metric remained non-significant.

This study did confirm previous literature's findings that higher levels of water consumption are protective against adverse birth outcomes. There appears to be a trade-off between drinking less tap water to avoid the possible risk of exposure to chemicals, or drinking more tap water to benefit from hydration. Based on these results, the benefits of hydration appear to outweigh the risks of chemical exposure with respect to in terms of birth outcomes. Of course, women could choose to consume other beverages to avoid the conundrum altogether, and every other option, at high enough dose, may also be associated with its own set of risks. As per the Paracelsus adage: it is the dose that makes the poison.

Whether or not a woman answered the filtering question in the water questionnaire affected her inclusion in Chapter 7's complete case analyses. A large proportion of women were therefore excluded (although they are included in the multiple imputations). In fact, the Pakistani women answered the filtering questions more frequently than the White British women, but answered *not* filtering their tap water more frequently when they did, such that proportionally more White British women may have been excluded due to this criterion. As an alternative the assumption could have been made that if a woman did not respond to either of the filtering questions, her answer was negative. Indeed, people tend not to fill out questions when their answer is 'no' (as per the assumptions made for zero water consumption in Chapter 8). The anticipated effect on results is minimised however because the combined metric analyses were replicated in the full cohort after carefully imputing any missing data.

9.2.2 Discussion and recommendations

The Bradford District Infant Mortality Commission reported that a third of the Bradford District babies that died in infancy were born at full term weighing 2500 grams or more (BDIMC 2006). Further still, babies born to Bradford District residents *in all birth weight ranges* had a significantly higher risk of post-neonatal (i.e. within the first seven days of life) and infant death than babies of a similar birth weight across the whole of England and Wales (BDIMC 2006). This means that the high

rate of post-neonatal mortality in the Bradford District was not explained solely by a greater proportion of babies born preterm or low birth weight.

One of the premises behind studying birth outcomes is that they are surrogate markers for infant mortality and future adult health (see Introduction (Chapter 1)). But many other lifestyle factors can also influence post-neonatal mortality. Prior to using birth weight as a means of predicting adult health, I believe it prudent to think of birth weight above all as an indicator of the fetal experience. Based on the results of this study, prenatal exposure to HAAs does not appear to negatively impact the health of fetuses *in utero*.

9.2.2.1 "Low birth weight paradox"

It has been acknowledged for many years that small babies from high-risk populations can have lower mortality than small babies from a lower-risk population. This is known as the "low birth weight paradox". Considering the example of altitude: the shift (towards the left, i.e. towards lower birth weights) of birth weight in populations living at high altitude might be expected to produce higher mortality because more babies are subjected to the higher risks at lower weights. In fact, total mortality is the same in the population living at higher vs. at lower altitudes because weight-specific mortality rates are not the same in the two populations. It is as if specific mortality rates at high altitude have shifted with the birth weight distribution, leading to no net change in mortality, such that at any given weight below 2,500, babies born at high altitude have slightly better survival than other babies.

This is because altitude affects birth weight but not mortality and the association between birth weight and mortality is an indirect one. If another factor is present that affects both birth weight and mortality, birth weight becomes a collider and stratifying on birth weight biases the relationship between altitude and mortality within those birth weight strata. In MacMahon's words, "the marked relationship between weight and mortality is the result of factors…that affect them both", and it is a mistake to analyse birth weight-specific mortality at all (p.221 in Wilcox(2010)).

9.2.2.2 Continuous birth weight

Wilcox also discusses the concepts of dominant vs. residual birth weight distributions in his latest book (Wilcox 2010). The dominant birth weight distribution is the Gaussian part of the distribution, and typically contains 95-98% of births. The residual birth weight distribution is the excess portion in the lower tail, lying outside the main Gaussian distribution. Although residual birth weights are few, they contribute a major portion of infant deaths.

If an environmental insult shifts a population's entire birth weight distribution, such a shift in mean birth weight will be driven by shifts within the dominant distribution and will not be sensitive to changes in the tail of the birth weight distribution. As changes in the overall mean of the dominant distribution, without changes in the residual, do not seem to be of importance for infant mortality or other forms of morbidity (Savitz et al. 2002), detecting such shifts in the overall mean birth weight may not be relevant to the study of the long-term effects of HAA exposures.

9.2.2.3 Transient effects

What's more, the trajectories of fetal growth *in utero* are known to differ by race, gender, and plurality (Brenner et al. 1976; Williams et al. 1982) and may be affected by exogenous factors. But such effects cannot be detected from size or weight at birth. Only longitudinal information on fetal size through the course of pregnancy would enable the study of attained weight at specified points in pregnancy, as well as of growth during specified intervals (Savitz et al. 2002). If environmental insults have a transient effect on growth, for example slowing it for a time and allowing the fetus to catch up later, it will not be captured by birth weight-based measures of growth restriction.

9.2.2.4 Outcomes going forward

Given the above, it is necessary to continue validating that birth outcomes are predictive of survival and future health, by following-up the BiB cohort to test the delayed consequences of prenatal or early childhood exposure to DBPs (growth, delayed puberty, obesity, neurodevelopmental deficits). While I believe we focused on the best available outcomes available to us in this thesis, I would encourage future studies to continue developing better birth outcome measures in order to improve the classification of fetal growth restriction

According to Brodsky and Christou (2004) , there are two main patterns of fetal growth restriction. If fetal growth is impaired during the first or second trimester, the infant will have symmetric growth restriction. This proportional lack of growth is caused by reduced fetal cellular proliferation of all organs and occurs in approximately 20% to 30% of growth restricted infants (Spencer et al. 1999). In contrast, asymmetric growth, in which an infant has a smaller abdominal size compared to head size and which is the most common form of growth restriction (~70-80%) (Lin et al. 1991), will occur if the decrease in growth velocity happens in the last trimester, and is attributed to the ability of the fetus to adapt, redistributing its cardiac output to the spleen, adrenal, coronary, and cerebral circulations. Although some overlap can occur, the timing of the growth delay is more important than the aetiology in determining the pattern of growth restriction.

Requiring birth length to be recorded at birth would help go beyond outcome measures based on birth weight only, in order for instance to examine symmetric vs. asymmetric growth restriction, using such

measures as the ponderal index (=birth weight/(birth length)^3 x100), an indicator for thinness of the newborn. I hope future studies will be able to pair this more nuanced approach to the definition of fetal growth restriction with improvements in fetal growth standards towards taking both intrauterine-based and birth weight-based standards into account (Hutcheon and Platt 2008).

9.2.2.5 Confounding

Most of the known or suspected risk factors for birth weight adjusted for in this study are correlated which means it is difficult to disentangle their respective effects on birth weight (Table A7 - 2).

This is relevant to the discussion over whether to use ethnic-specific references to generate measures of deviation from subgroup norms such as small-for-gestational age (SGA). Choice of reference curve for SGA derivation is indeed critical, and I decided not to adjust SGA thresholds for ethnicity using previous estimates a) because of non-availability of such ethnic-specific estimates and b) because of the risk involved in making the strong assumption that adjusting for ethnicity (but not deprivation status, nor parity, nor any other demographic or behavioural factor) is sufficient. In a population of women who not only had different ethnic origin but were also smaller maternal size (Spencer et al. 1995), an ethnic standard derived from the smaller mothers will apply to ethnic mothers who are smaller than the general population. But it will not apply to mothers who are larger than average for their ethnic group, and whose babies should be expected to grow to a heavier weight. The fact that a baby is SGA may be overlooked in this case (Gardosi 1995).

In terms of uncontrolled confounding, a major element of birth weight relevant to this cohort and which w taken into account is consanguinity. Consanguinity is a major risk factor for congenital anomaly (Sheridan et al. 2013). The majority of Pakistani-origin mothers in the BiB cohort were related to the father of their child, and this was a more common practices than for their parents (64% vs. 55%) (Wright et al. 2012). As questionnaire information is available on participants' relatedness by blood, accounting for the effects of consanguinity would be an interesting avenue to explore further.

9.3 Aim 3

9.3.1 Summary

The third and last aim of the thesis, described in full in Chapter 8, was to investigate water use patterns in the third trimester of pregnancy by evaluating the agreement of individual water use values reported in BiB questionnaires at baseline (at approx. 26-28 weeks of pregnancy) and at two later time points in pregnancy (30-33 weeks of pregnancy, and 36-39 weeks of pregnancy) in a subset of BiB women.

Overall, Chapter 8 reports that women tend to slightly increase their consumption over the third trimester of pregnancy by approx. one 200ml glass per time point, i.e. every 5 to 6 weeks, or approx. 2 glasses from the end of the second trimester of pregnancy to the end of the third. Unemployed women reported drinking significantly less total water than working women (Model 2), while current smokers and the White British women drank on average the most. Over time, being Pakistani, being employed and working (as opposed to on maternity leave), being unemployed, or never having smoked cigarettes all predicted an increase both in total tap water and total water consumption (Models 3, 4, 5).

9.3.2 Discussion and recommendations

Definitive differentiation between behaviour change (e.g. as in women drinking more tap water as the pregnancy progresses, perhaps because they are more uncomfortable or more idle/at home) and variability in questionnaire responses due to measurement error (i.e. questionnaire are imprecise) is not possible from the repeat questionnaire study. However I infer from these results that both behaviour change and measurement error likely contributed to this change over time such that assessing women's late pregnancy water consumption and validating questionnaires by other methods (such as biomarker studies or further questionnaire assessment by daily water diary) remain critical to developing the most accurate late trimester exposure estimates.

Because the contribution of individual water consumption is so important to women's overall exposure to HAAs in a location like Bradford where area-level variation in HAA concentrations is relatively limited, these results should be considered in any future modelling of third trimester HAA exposure estimates in order to help minimise exposure misclassification.

Future epidemiologic analyses could also consider the change in behaviour over the course of pregnancy/high within-subject variability found in Chapter 8, by including these repeat data in a Bayesian hierarchical model with a measurement error component to account for the imprecision of exposure estimates. Another interesting avenue would be to include Dr Rachel Smith's work on THMs and/or the air pollutant information derived for the Bradford area by Dr Kees De Hoogh and CREAL (Barcelona, Spain) with the HAA exposures presented in this thesis, in order study multiple exposures (see Chapter 7).

9.4 Final thoughts

This thesis studies the best possible exposure metric given current knowledge, improving the accuracy of exposure estimates from the methods used to date in the field of the environmental epidemiology of the understudied HAAs. In addition, it considers outcome measures carefully selected to identify pathologically small babies. And it does so at an unprecedented scale in a population of 10,000+ singleton babies which is relatively disadvantaged compared to the rest of the UK meaning that there is a real need to understand underlying causes of increased morbidity and ill health.

Despite this, this study still does not find an epidemiologic association between prenatal exposure to HAAs and adverse birth outcomes. As such, it certainly contributes to the small existing body of evidence on the subject: of the seven epidemiologic studies published to date to my knowledge, four find that HAAs may be associated with adverse birth weight-based outcomes, and three others report only non-significant (and even inverse) effects. While there will always be a number of limitations associated with human studies—they are inherently "messy"—, these results must be accepted at face value.

Studies in the field of DBP research are difficult to conduct in general because of the limited predictability and ever increasing breadth of the chemicals of interest. I believe that we (as a society) should address the risks associated with any technology (even incredibly beneficial ones to public health like water disinfection) if most of the population is exposed to it and there is reasonable doubt as to its potential to harm health. But given that research resources will always be limited, we should consider studying DBPs differently going forward. Perhaps with a more focused purpose (e.g. are we only interested in HAAs, or are we interested in HAAs as proxies for all other DBPs), in a more highly exposed population (we worked with a vulnerable population in Bradford, but one which was not particularly highly exposed to HAAs), and after putting a renewed emphasis on the need for a better understanding of the biological mechanisms of action of DBPs in general and HAAs in particular. Given that a 1% increased risk of SGA, as was reported in the meta-analysis by Grellier et al. (2010) for the association between THMs and SGA, is not detectable in an epidemiologic study, but only in a meta-analysis (S. Cordier, Gordon conference, 2012), perhaps the next step is first of all to conduct such a meta-analysis for HAAs.

But whatever it is, as Savitz suggested in an editorial in the Epidemiology and Society section of the journal Epidemiology (Savitz 2010), the marginal costs vs. marginal benefits (as far as they can be determined) of any new study must be weighed. 'Because we can' (i.e. the funding is in place, the cohort recruited, data routinely collected, and the manpower ready and willing) cannot be the only reason to embark on a new study. Instead any new DBP study must have a chance of providing an

incremental benefit in light of what has already been done. In the meantime, this thesis has led to the rather reassuring conclusion that there is no discernable issue with the specific class of DBPs which are HAAs at the levels observed in Bradford. Perhaps this is an excellent result for the people of Bradford.

REFERENCES

Aggazzotti G, Righi E, Fantuzzi G, Biasotti B, Ravera G, Kanitz S, et al. 2004. Chlorination by-products (cbps) in drinking water and adverse pregnancy outcomes in italy. J Water Health 2:233-247.

Al Salmi I, Hoy WE, Kondalsamy-Chennakesavan S, Wang Z, Gobe GC, Barr EL, et al. 2008. Disorders of glucose regulation in adults and birth weight: Results from the australian diabetes, obesity and lifestyle (ausdiab) study. Diabetes care 31:159-164.

Alberman E. 1991. Are our babies becoming bigger? Journal of the Royal Society of Medicine 84:257-260.

Alexander GR, Himes JH, Kaufman RB, Mor J, Kogan M. 1996. A united states national reference for fetal growth. Obstet Gynecol 87:163-168.

Alexander GR, Kogan MD, Himes JH. 1999. 1994-1996 u.S. Singleton birth weight percentiles for gestational age by race, hispanic origin, and gender. Matern Child Health J 3:225-231.

Alexander GR, Wingate MS, Mor J, Boulet S. 2007. Birth outcomes of asian-indian-americans. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics 97:215-220.

Alston TA. 1991. Inhibition of vitamin b12-dependent methionine biosynthesis by chloroform and carbon tetrachloride. Biochem Pharmacol 42:R25-28.

Altman DG, Coles EC. 1980. Nomograms for precise determination of birth weight for dates. Br J Obstet Gynaecol 87:81-86.

Alvear J, Brooke OG. 1978. Fetal growth in different racial groups. Archives of disease in childhood 53:27-32.

Amy GL, Siddiqui M, Zhai W, Debroux J. 1994. Survey on bromide in drinking water and impacts on dbp formation. Denver, CO, US:American Water Works Association Research Foundation

Amy GL, Siddiqui M, Ozekin K, Zhu HW, Wang C. 1998. Empirically based models for predicting chlorination and ozonation by-products: Trihalomethanes haloacetic acids chloral hydrate and bromate. Washington, D.C., US:US Environmental Protection Agency, Office of Water.

Andersen AM, Andersen PK, Olsen J, Gronbaek M, Strandberg-Larsen K. 2012. Moderate alcohol intake during pregnancy and risk of fetal death. Int J Epidemiol 41:405-413.

Andres RL, Day MC. 2000. Perinatal complications associated with maternal tobacco use. Seminars in neonatology : SN 5:231-241.

APHO AoPHO. 2008. Health profiles. Bradford, UK.

Arbuckle TE, Hrudey SE, Krasner SW, Nuckols JR, Richardson SD, Singer P, et al. 2002. Assessing exposure in epidemiologic studies to disinfection by-products in drinking water: Report from an international workshop. Environ Health Perspect 110 Suppl 1:53-60.

Armstrong BG. 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup Environ Med 55:651-656.

Ashworth A. 1998. Effects of intrauterine growth retardation on mortality and morbidity in infants and young children. Eur J Clin Nutr 52 Suppl 1:S34-41; discussion S41-32.

Ates N, Kaplan SS, Sahinkaya E, Kitis M, Dilek FB, Yetis U. 2007. Occurrence of disinfection by-products in low doc surface waters in turkey. J Hazard Mater 142:526-534.

Attene-Ramos MS, Wagner ED, Plewa MJ. 2010. Comparative human cell toxicogenomic analysis of monohaloacetic acid drinking water disinfection byproducts. Environ Sci Technol 44:7206-7212.

Austin EW, Okita JR, Okita RT, Larson JL, Bull RJ. 1995. Modification of lipoperoxidative effects of dichloroacetate and trichloroacetate is associated with peroxisome proliferation. Toxicology 97:59-69.

Baird DD, Weinberg CR, Wilcox AJ, McConnaughey DR, Musey PI. 1991. Using the ratio of urinary oestrogen and progesterone metabolites to estimate day of ovulation. Stat Med 10:255-266.

Baird DD, McConnaughey DR, Weinberg CR, Musey PI, Collins DC, Kesner JS, et al. 1995. Application of a method for estimating day of ovulation using urinary estrogen and progesterone metabolites. Epidemiology 6:547-550.

Balchin I, Steer PJ. 2007. Race, prematurity and immaturity. Early Hum Dev 83:749-754.

Bambang S, Spencer NJ, Logan S, Gill L. 2000. Cause-specific perinatal death rates, birth weight and deprivation in the west midlands, 1991-93. Child: care, health and development 26:73-82.

Barbone F, Valent F, Brussi V, Tomasella L, Triassi M, Di Lieto A, et al. 2002. Assessing the exposure of pregnant women to drinking water disinfection byproducts. Epidemiology 13:540-544.

Barker DJ. 1997. The long-term outcome of retarded fetal growth. Clin Obstet Gynecol 40:853-863.

Barker DJ. 2002. Fetal programming of coronary heart disease. Trends Endocrinol Metab 13:364-368.

Barker DJP. 1994. Mothers, babies and disease in later life. London, UK:B.M.J. Publishing.

Baron RM, Kenny DA. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of personality and social psychology 51:1173-1182.

Bateson P, Barker D, Clutton-Brock T, Deb D, D'Udine B, Foley RA, et al. 2004. Developmental plasticity and human health. Nature 430:419-421.

Bateson TF, Wright JM. 2010. Regression calibration for classical exposure measurement error in environmental epidemiology studies using multiple local surrogate exposures. Am J Epidemiol 172:344-352.

BDIMC BDIMC. 2006. Bradford district infant mortality commission summary report.

Bougeard C. 2009. Haloacetic acids and other disinfection by-products in uk treated waters: Occurrence, formation and precursor investigation. Cranfield, UK:Cranfield University.

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

Bradford Teaching Hospitals NHS Foundation Trust BTHNHST. 2005. Guidelines for obstetric ultrasound.

Branum AM, Schoendorf KC. 2002. Changing patterns of low birthweight and preterm birth in the united states, 1981-98. Paediatr Perinat Epidemiol 16:8-15.

Brenner WE, Edelman DA, Hendricks CH. 1976. A standard of fetal growth for the united states of america. Am J Obstet Gynecol 126:555-564.

Brodsky D, Christou H. 2004. Current concepts in intrauterine growth restriction. Journal of intensive care medicine 19:307-319.

Buekens P, Delvoye P, Wollast E, Robyn C. 1984. Epidemiology of pregnancies with unknown last menstrual period. Journal of epidemiology and community health 38:79-80.

Bukowski R, Gahn D, Denning J, Saade G. 2001. Impairment of growth in fetuses destined to deliver preterm. Am J Obstet Gynecol 185:463-467.

Cancho B, Ventura F, Galceran MT. 1999. Behavior of halogenated disinfection by-products in the water treatment plant of barcelona, spain. Bulletin of environmental contamination and toxicology 63:610-617.

Carlson M, Hardy D. 1998. Controlling dbps with monocholoramine. Denver, CO, US:American Water Works Association.

Carstairs V, Morris R. 1991. Deprivation and health in scotland. . Aberdeen, Scotland:Aberdeen University Press.

Castano-Vinyals G, Cantor KP, Villanueva CM, Tardon A, Garcia-Closas R, Serra C, et al. 2011. Socioeconomic status and exposure to disinfection by-products in drinking water in spain. Environ Health 10:18.

Catalano PM, Thomas A, Huston-Presley L, Amini SB. 2003. Increased fetal adiposity: A very sensitive marker of abnormal in utero development. Am J Obstet Gynecol 189:1698-1704.

Cedergren MI, Selbing AJ, Lofman O, Kallen BA. 2002. Chlorination byproducts and nitrate in drinking water and risk for congenital cardiac defects. Environ Res 89:124-130.

Cemeli E, Wagner ED, Anderson D, Richardson SD, Plewa MJ. 2006. Modulation of the cytotoxicity and genotoxicity of the drinking water disinfection byproduct lodoacetic acid by suppressors of oxidative stress. Environ Sci Technol 40:1878-1883.

Chen W, Weisel C. 1998. Halogenated dbp concentrations in a distribution system. Journal of the American Water Works Association 90:151-163.

Chetcuti P, Sinha SH, Levene MI. 1985. Birth size in indian ethnic subgroups born in britain. Archives of disease in childhood 60:868-870.

Chike-Obi U, David RJ, Coutinho R, Wu SY. 1996. Birth weight has increased over a generation. Am J Epidemiol 144:563-569.

Choi H, Perera F, Pac A, Wang L, Flak E, Mroz E, et al. 2008. Estimating individual-level exposure to airborne polycyclic aromatic hydrocarbons throughout the gestational period based on personal, indoor, and outdoor monitoring. Environ Health Perspect 116:1509-1518.

Chowdhury S, Ammari F, Burden AC, Gregory R. 2000. Secular trend in birth weight in native white and immigrant south asian populations in leicester, uk: Possible implications for incidence of type 2 diabetes in the future. . Practical Diabetes International 17:104-108.

Chowdhury S, Rodriguez MJ, Serodes J. 2010. Model development for predicting changes in dbp exposure concentrations during indoor handling of tap water. Sci Total Environ 408:4733-4743.

Clausson B, Gardosi J, Francis A, Cnattingius S. 2001. Perinatal outcome in sga births defined by customised versus population-based birthweight standards. BJOG : an international journal of obstetrics and gynaecology 108:830-834.

Cliver SP, Goldenberg RL, Cutter GR, Hoffman HJ, Davis RO, Nelson KG. 1995. The effect of cigarette smoking on neonatal anthropometric measurements. Obstet Gynecol 85:625-630.

Cnattingius S, Axelsson O, Eklund G, Lindmark G, Meirik O. 1984. Factors influencing birthweight for gestational age, with special respect to risk factors for intrauterine growth retardation. Early Hum Dev 10:45-55.

Cnattingius S, Bergstrom R, Lipworth L, Kramer MS. 1998. Prepregnancy weight and the risk of adverse pregnancy outcomes. N Engl J Med 338:147-152.

Cogswell ME, Parvanta I, Ickes L, Yip R, Brittenham GM. 2003. Iron supplementation during pregnancy, anemia, and birth weight: A randomized controlled trial. Am J Clin Nutr 78:773-781.

Cole TJ, Green PJ. 1992. Smoothing reference centile curves: The lms method and penalized likelihood. Stat Med 11:1305-1319.

Cole TJ, Freeman JV, Preece MA. 1998. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. Stat Med 17:407-429.

Cole TJ, Williams AF, Wright CM. 2011. Revised birth centiles for weight, length and head circumference in the uk-who growth charts. Annals of human biology 38:7-11.

Cooper DL, Petherick ES, Wright J. 2013. The association between binge drinking and birth outcomes: Results from the born in bradford cohort study. Journal of epidemiology and community health 67:821-828.

Costet N, Garlantezec R, Monfort C, Rouget F, Gagniere B, Chevrier C, et al. 2012. Environmental and urinary markers of prenatal exposure to drinking water disinfection by-products, fetal growth, and duration of gestation in the pelagie birth cohort (brittany, france, 2002-2006). Am J Epidemiol 175:263-275.

Cowman GA, Singer PC. 1996. Effect of bromide ion on haloacetic acid speciation resulting from chlorination and chloramination of aquatic humic substances. Environmental Science and Technology 30:16-24.

Daniels MJ, Hogan JW. 2008. Missing data in longitudinal studies: Strategies for bayesian modeling and sensitivity analysis. 1st ed. New York, NY, USA:Chapman & Hall/CRC.

Davis RO, Cutter GR, Goldenberg RL, Hoffman HJ, Cliver SP, Brumfield CG. 1993. Fetal biparietal diameter, head circumference, abdominal circumference and femur length. A comparison by race and sex. The Journal of reproductive medicine 38:201-206.

de Souza-Pinto NC, Eide L, Hogue BA, Thybo T, Stevnsner T, Seeberg E, et al. 2001. Repair of 8-oxodeoxyguanosine lesions in mitochondrial dna depends on the oxoguanine dna glycosylase (ogg1) gene and 8-oxoguanine accumulates in the mitochondrial dna of ogg1-defective mice. Cancer research 61:5378-5381.

Dejin-Karlsson E, Hanson BS, Ostergren PO, Lindgren A, Sjoberg NO, Marsal K. 2000. Association of a lack of psychosocial resources and the risk of giving birth to small for gestational age infants: A stress hypothesis. BJOG : an international journal of obstetrics and gynaecology 107:89-100.

Department of Health DOH. 2012. Diabetes and pregnancy. Available: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/browsable/DH_4915754 [accessed 04/05/2012 2012].

Dhawan S. 1995. Birth weights of infants of first generation asian women in britain compared with second generation asian women. BMJ 311:86-88.

DHI. 2008. Final report on establishment of a list of chemical parameters for the revision of the drinking water directive.

Diehl AC, Speitel Jr GE, Symons JM, Krasner SW, Hwang CJ, Barrett SE. 2000. Dbp formation during chloramination. Journal AWWA, 92:76-90.

DiFranza JR, Lew RA. 1995. Effect of maternal cigarette smoking on pregnancy complications and sudden infant death syndrome. The Journal of family practice 40:385-394.

Dion-Fortier A, Rodriguez MJ, Serodes J, Proulx F. 2009. Impact of water stagnation in residential cold and hot water plumbing on concentrations of trihalomethanes and haloacetic acids. Water Res 43:3057-3066.

Dodds L, King W, Woolcott C, Pole J. 1999. Trihalomethanes in public water supplies and adverse birth outcomes. Epidemiology 10:233-237.

Dojlido J, Zbiec E, Swietlik R. 1999. Formation of the haloacetic acids during ozonation and chlorination of water in warsaw waterworks (poland). . Water research 33:3111-3118.

Dole N, Savitz DA, Hertz-Picciotto I, Siega-Riz AM, McMahon MJ, Buekens P. 2003. Maternal stress and preterm birth. Am J Epidemiol 157:14-24.

Dornhorst A, Paterson CM, Nicholls JS, Wadsworth J, Chiu DC, Elkeles RS, et al. 1992. High prevalence of gestational diabetes in women from ethnic minority groups. Diabetic medicine : a journal of the British Diabetic Association 9:820-825.

Dow JL, Green T. 2000. Trichloroethylene induced vitamin b(12) and folate deficiency leads to increased formic acid excretion in the rat. Toxicology 146:123-136.

Draper ES, Abrams KR, Clarke M. 1995. Fall in birth weight of third generation asian infants. BMJ 311:876.

DWI DWI. 2008. National tap water consumption survey for england and wales DWI 70/2/217 Phase Two.

DWI DWI. 2010. The water supply regulations 2010

Egorov AI, Tereschenko AA, Altshul LM, Vartiainen T, Samsonov D, LaBrecque B, et al. 2003. Exposures to drinking water chlorination by-products in a russian city. Int J Hyg Environ Health 206:539-551.

EPA UEPA. 2003. Method 552.3: Determination of haloacetic acids and dalapon in drinking water by liquid-liquid microextraction, derivatization and gas chromatography with electron capture detection. Revision 1.0. Technical Support Center USEPA Office of Ground Water and Drinking Water.

EPA UEPA. 2006. Stage 2 disinfectants and disinfection byproducts rule.

EPA USEPA. 1979. National interim primary drinking water regulations. Fed Registr 44.

EPA USEPA. 1998. National primary drinking water regulations: Disinfectants and disinfection by-products; final rule. Fed Registr 63.

EPA USEPA. 2000. Estimated per capita water ingestion in the us. Washington, D.C., US:Office of Science and Technology, Office of Water.

Ershow AG, Brown LM, Cantor KP. 1991. Intake of tapwater and total water by pregnant and lactating women. Am J Public Health 81:328-334.

Fairley L, Petherick ES, Howe LD, Tilling K, Cameron N, Lawlor DA, et al. 2013. Describing differences in weight and length growth trajectories between white and pakistani infants in the uk: Analysis of the born in bradford birth cohort study using multilevel linear spline models. Archives of disease in childhood 98:274-279.

Firdous R, Bhopal RS. 1989. Reproductive health of asian women: A comparative study with hospital and community perspectives. Public health 103:307-315.

Fleischacker SJ, Randtke SJ. 1983. Formation of organic chlorine in public water supplies. . Journal of the American Water Works Association 75:132-138.

Forssen UM, Herring AH, Savitz DA, Nieuwenhuijsen MJ, Murphy PA, Singer PC, et al. 2007. Predictors of use and consumption of public drinking water among pregnant women. J Expo Sci Environ Epidemiol 17:159-169.

Forssen UM, Wright JM, Herring AH, Savitz DA, Nieuwenhuijsen MJ, Murphy PA. 2009. Variability and predictors of changes in water use during pregnancy. J Expo Sci Environ Epidemiol 19:593-602.

Frederick IO, Williams MA, Sales AE, Martin DP, Killien M. 2008. Pre-pregnancy body mass index, gestational weight gain, and other maternal characteristics in relation to infant birth weight. Matern Child Health J 12:557-567.

Freeman JV, Cole TJ, Chinn S, Jones PR, White EM, Preece MA. 1995. Cross sectional stature and weight reference curves for the uk, 1990. Archives of disease in childhood 73:17-24.

Fuentes-Afflick E, Hessol NA. 1997. Impact of asian ethnicity and national origin on infant birth weight. Am J Epidemiol 145:148-155.

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

Gallard H, von GU. 2002. Chlorination of natural organic matter: Kinetics of chlorination and of thm formation. Water Res 36:65-74.

Gallivan S, Robson SC, Chang TC, Vaughan J, Spencer JA. 1993. An investigation of fetal growth using serial ultrasound data. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 3:109-114.

Gang D, Clevenger TE, Banerji SK. 2003. Relationship of chlorine decay and thms formation to nom size. J Hazard Mater 96:1-12.

Gang DD, Segar RLJ, Clevenger TE, Shankha KB. 2002. Using chlorine demand to predict thm and haa9 formation. . Journal of the American Water Works Association 94:76-86.

Garcia-Villanova RJ, Garcia C, Gomez JA, Garcia MP, Ardanuy R. 1997. Formation, evolution and modeling of trihalomethanes in the drinkingwater of a town: I. At the municipal treatment utilities. Water Res 31:1299-1308.

Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM. 1992. Customised antenatal growth charts. Lancet 339:283-287.

Gardosi J, Mongelli M, Wilcox M, Chang A. 1994. Screening and assessment of fetal growth. In: A critical appraisal of fetal surveillance, (Van Geijn HP, Copray F, eds). Amsterdam:Elsevier, 437-450.

Gardosi J. 1995. Ethnic differences in fetal growth. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 6:73-74.

Gardosi J, Mongelli M, Wilcox M, Chang A. 1995. An adjustable fetal weight standard. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 6:168-174.

Gardosi J, Vanner T, Francis A. 1997. Gestational age and induction of labour for prolonged pregnancy. Br J Obstet Gynaecol 104:792-797.

Gardosi J. 1998. The application of individualised fetal growth curves. J Perinat Med 26:137-142.

Gardosi J. 2006. New definition of small for gestational age based on fetal growth potential. Horm Res 65 Suppl 3:15-18.

Gardosi J, Clausson B, Francis A. 2007. The use of customised versus population-based birthweight standards in predicting perinatal mortality. BJOG : an international journal of obstetrics and gynaecology 114:1301-1302; author reply 1303.

Gardosi J. 2009. Intrauterine growth restriction: New standards for assessing adverse outcome. Best practice & research Clinical obstetrics & gynaecology 23:741-749.

Gardosi JO. 2005. Prematurity and fetal growth restriction. Early Hum Dev 81:43-49.

Gelbaya TA, Nardo LG. 2005. Customised fetal growth chart: A systematic review. J Obstet Gynaecol 25:445-450.

Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. Stat Sci 7:457-472.

Gelman A, Carlin JB, Stern HS, al e. 2004. Texts in statistical science In: Bayesian data analysis 2nd ed Boca Raton, FL: 2004.:CRC Press, LLC.

Gelman A. 2006. Prior distributions for variance parameters in hierarchical models(comment on an article by browne and draper). Bayesian Analysis 1:515-533.

Gelman A, Hill J. 2007. Data analysis using regression and multilevel/hierarchical models New York, NY:Cambridge University Press.

Gjessing HK, Grottum P, Eik-Nes SH. 2007. A direct method for ultrasound prediction of day of delivery: A new, population-based approach. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 30:19-27.

Glaser JA, Foerst DL, McKee GD, Quave SA, Budde WL. 1981. Trace analyses for wastewaters. Environ Sci Technol 15:1426-1435.

Gluckman PD, Hanson MA. 2004. The developmental origins of the metabolic syndrome. Trends Endocrinol Metab 15:183-187.

Goldenberg RL, Culhane JF, Iams JD, Romero R. 2008. Epidemiology and causes of preterm birth. Lancet 371:75-84.

Golding J, Pembrey M, Jones R. 2001. Alspac--the avon longitudinal study of parents and children. I. Study methodology. Paediatr Perinat Epidemiol 15:74-87.

Goldstein H, Browne WJ, Rasbash J. 2002. Partitioning variation in multilevel models. . Understanding Statistics 1:223-232.

Golfinopoulos SK. 2000. The occurrence of trihalomethanes in the drinking water in greece. Chemosphere 41:1761-1767.

Golfinopoulos SK, Nikolaou AD. 2005. Survey of disinfection by-products in drinking water in athens, greece. . Desalination 176:13-24.

Goslan EH, Krasner SW, Villanueva CM, Toledano MB, Kogevinas M, Stephanou EG, et al. in prep. Disinfection byproduct occurrence in european waters.

Gould JB, Madan A, Qin C, Chavez G. 2003. Perinatal outcomes in two dissimilar immigrant populations in the united states: A dual epidemiologic paradox. Pediatrics 111:e676-682.

Graves CG, Matanoski GM, Tardiff RG. 2001. Weight of evidence for an association between adverse reproductive and developmental effects and exposure to disinfection by-products: A critical review. Regul Toxicol Pharmacol 34:103-124.

Greenland SR, K.J.;. 1998. Modern epidemiology. 2nd edition ed. Philadelphia:Lippincott Williams & Wilkins.

Grellier J, Bennett J, Patelarou E, Smith RB, Toledano MB, Rushton L, et al. 2010. Exposure to disinfection by-products, fetal growth, and prematurity: A systematic review and meta-analysis. Epidemiology 21:300-313.

Gulmezoglu M, de Onis M, Villar J. 1997. Effectiveness of interventions to prevent or treat impaired fetal growth. Obstet Gynecol Surv 52:139-149.

Haddow JE, Knight GJ, Kloza EM, Palomaki GE, Wald NJ. 1991. Cotinine-assisted intervention in pregnancy to reduce smoking and low birthweight delivery. Br J Obstet Gynaecol 98:859-865.

Halileh S, Abu-Rmeileh N, Watt G, Spencer N, Gordon N. 2008. Determinants of birthweight: Gender based analysis. Matern Child Health J 12:606-612.

HAPO Study Cooperative Research Group. 2010. Hyperglycaemia and adverse pregnancy outcome (hapo) study: Associations with maternal body mass index. BJOG : an international journal of obstetrics and gynaecology 117:575-584.

Harding S, Rosato MG, Cruickshank JK. 2004. Lack of change in birthweights of infants by generational status among indian, pakistani, bangladeshi, black caribbean, and black african mothers in a british cohort study. Int J Epidemiol 33:1279-1285.

Hashimoto S, Azuma T, Otsuki A. 1998. Distribution, sources and stability of haloacetic acids in tokyo bay, japan. Environ Toxicol Chem 17:798-805.

Hassoun EA, Ray S. 2003. The induction of oxidative stress and cellular death by the drinking water disinfection by-products, dichloroacetate and trichloroacetate in j774.A1 cells. Comp Biochem Physiol C Toxicol Pharmacol 135:119-128.

Hattersley AT, Tooke JE. 1999. The fetal insulin hypothesis: An alternative explanation of the association of low birthweight with diabetes and vascular disease. Lancet 353:1789-1792.

Hawkins SS, Lamb K, Cole TJ, Law C. 2008. Influence of moving to the uk on maternal health behaviours: Prospective cohort study. BMJ 336:1052-1055.

Haworth JC, Ellestad-Sayed JJ, King J, Dilling LA. 1980. Fetal growth retardation in cigarette-smoking mothers is not due to decreased maternal food intake. Am J Obstet Gynecol 137:719-723.

Hayes DK, Lukacs SL, Schoendorf KC. 2008. Heterogeneity within asian subgroups: A comparison of birthweight between infants of us and non-us born asian indian and chinese mothers. Matern Child Health J 12:549-556.

Health Survey for England. 2004. Risk factors for cardiovascular disease. London:Joint Health Surveys Unit.

Health Survey for England. 2005. The health of ethnic minority groups. London, UK:Health and Social Care Information Centre.

Hediger ML, Overpeck MD, Ruan WJ, Troendle JF. 2002. Birthweight and gestational age effects on motor and social development. Paediatr Perinat Epidemiol 16:33-46.

Heller-Grossman L, Manka J, Limoni-Relis B, Rebhun M. 1993. Formation and distribution of haloacetic acids, thm and tox in chlorination of bromide-rich lake water. . Water research 27:1323-1331.

Henriksen TB, Wilcox AJ, Hedegaard M, Secher NJ. 1995. Bias in studies of preterm and postterm delivery due to ultrasound assessment of gestational age. Epidemiology 6:533-537.

Hernandez-Diaz S, Toh S, Cnattingius S. 2009. Risk of pre-eclampsia in first and subsequent pregnancies: Prospective cohort study. BMJ 338:b2255.

Hinckley AF, Bachand AM, Reif JS. 2005. Late pregnancy exposures to disinfection by-products and growth-related birth outcomes. Environ Health Perspect 113:1808-1813.

Hoffman CS, Mendola P, Savitz DA, Herring AH, Loomis D, Hartmann KE, et al. 2008a. Drinking water disinfection by-product exposure and duration of gestation. Epidemiology 19:738-746.

Hoffman CS, Mendola P, Savitz DA, Herring AH, Loomis D, Hartmann KE, et al. 2008b. Drinking water disinfection by-product exposure and fetal growth. Epidemiology 19:729-737.

Horta BL, Victora CG, Menezes AM, Halpern R, Barros FC. 1997. Low birthweight, preterm births and intrauterine growth retardation in relation to maternal smoking. Paediatr Perinat Epidemiol 11:140-151.

Horton BJ, Luben TJ, Herring AH, Savitz DA, Singer PC, Weinberg HS, et al. 2011. The effect of water disinfection by-products on pregnancy outcomes in two southeastern us communities. J Occup Environ Med 53:1172-1178.

Howe GR, Chiarelli AM, Lindsay JP. 1988. Components and modifiers of the healthy worker effect: Evidence from three occupational cohorts and implications for industrial compensation. Am J Epidemiol 128:1364-1375.

Hughes K, Tan NR, Lun KC. 1986. Ethnic group differences in low birthweight of live singletons in singapore, 1981-3. Journal of epidemiology and community health 40:262-266.

Hunter ES, 3rd, Rogers EH, Schmid JE, Richard A. 1996. Comparative effects of haloacetic acids in whole embryo culture. Teratology 54:57-64.

Hunter ES, 3rd, Rogers E, Blanton M, Richard A, Chernoff N. 2006. Bromochloro-haloacetic acids: Effects on mouse embryos in vitro and qsar considerations. Reprod Toxicol 21:260-266.

Hutcheon JA, Platt RW. 2008. The missing data problem in birth weight percentiles and thresholds for "small-for-gestational-age". Am J Epidemiol 167:786-792.

Hutcheon JA, Zhang X, Cnattingius S, Kramer MS, Platt RW. 2008. Customised birthweight percentiles: Does adjusting for maternal characteristics matter? BJOG : an international journal of obstetrics and gynaecology 115:1397-1404.

Hutcheon JA, Zhang X, Platt RW, Cnattingius S, Kramer MS. 2011. The case against customised birthweight standards. Paediatr Perinat Epidemiol 25:11-16.

Huxley R, Neil A, Collins R. 2002. Unravelling the fetal origins hypothesis: Is there really an inverse association between birthweight and subsequent blood pressure? Lancet 360:659-665.

Huxley R, Owen CG, Whincup PH, Cook DG, Rich-Edwards J, Smith GD, et al. 2007. Is birth weight a risk factor for ischemic heart disease in later life? Am J Clin Nutr 85:1244-1250.

Hwang BF, Magnus P, Jaakkola JJ. 2002. Risk of specific birth defects in relation to chlorination and the amount of natural organic matter in the water supply. Am J Epidemiol 156:374-382.

Hwang BF, Jaakkola JJ, Guo HR. 2008. Water disinfection by-products and the risk of specific birth defects: A population-based cross-sectional study in taiwan. Environ Health 7:23.

Hwang CJ, Amy GL, Bruchet A, Croué J-P, Krasner SW, Leenheer JA. 2001. Polar nom: Characterization, dbps, treatment. . Denver, CO, US:American Water Works Association Research Foundation

IARC IAfRoC. 2004. Some drinking-water disinfectants and contaminants, including arsenic IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 84:512 pages.

IHDP IHaDP. 1990. Enhancing the outcomes of low-birth-weight, premature infants. A multisite, randomized trial. . JAMA 263:3035-3042.

Infante-Rivard C. 2004. Drinking water contaminants, gene polymorphisms, and fetal growth. Environ Health Perspect 112:1213-1216.

Institute of Medicine IOM. 2009. Weight gain during pregnancy: Reexamining the guidelines. Available: http://www.iom.edu/~/media/Files/Report%20Files/2009/Weight-Gain-During-Pregnancy-Reexamining-the-Guidelines/Resource%20Page%20-%20Weight%20Gain%20During%20Pregnancy.pdf.

IPCS IPoCS. 2000. Environmental health criteria 216: Disinfectants and disinfectant by-products (report). (Commision of the European Communities). Geneva, Switzerland World Health Organization (WHO).

Iszatt N, Nieuwenhuijsen MJ, Nelson P, Elliott P, Toledano MB. 2011. Water consumption and use, trihalomethane exposure, and the risk of hypospadias. Pediatrics 127:e389-397.

Itoh S, Echigo S. 2008. Disinfection by-products in water. Tokyo, Japan (in Japanese).

Jacobsson B, Ahlin K, Francis A, Hagberg G, Hagberg H, Gardosi J. 2008. Cerebral palsy and restricted growth status at birth: Population-based case-control study. BJOG : an international journal of obstetrics and gynaecology 115:1250-1255.

Jeong CH, Wagner ED, V.R. S, S. A, Richardson SD, Daiber EJ, et al. 2012. Occurrence and toxicity of disinfection byproducts in european drinking waters in relation with the hiwate epidemiology study. Environ Sci Technol 46:12120-12128.

Johnson W. 2009. The growth of bradford infants. :University of Lougborough.

Johnson W, Cameron N, Dickson P, Emsley S, Raynor P, Seymour C, et al. 2009. The reliability of routine anthropometric data collected by health workers: A cross-sectional study. International journal of nursing studies 46:310-316.

Jones KL, Smith DW. 1973. Recognition of the fetal alcohol syndrome in early infancy. Lancet 302:999-1001.

Joshi NP, Kulkarni SR, Yajnik CS, Joglekar CV, Rao S, Coyaji KJ, et al. 2005. Increasing maternal parity predicts neonatal adiposity: Pune maternal nutrition study. Am J Obstet Gynecol 193:783-789.

Jovanovic L, Pettitt DJ. 2001. Gestational diabetes mellitus. JAMA 286:2516-2518.

Kallen BA, Robert E. 2000. Drinking water chlorination and delivery outcome-a registry-based study in sweden. Reprod Toxicol 14:303-309.

Kargalioglu Y, McMillan BJ, Minear RA, Plewa MJ. 2002. Analysis of the cytotoxicity and mutagenicity of drinking water disinfection by-products in salmonella typhimurium. Teratog Carcinog Mutagen 22:113-128.

Karowicz-Bilinska A, Suzin J, Sieroszewski P. 2002. Evaluation of oxidative stress indices during treatment in pregnant women with intrauterine growth retardation. Med Sci Monit 8:CR211-216.

Kaur S, Nieuwenhuijsen MJ, Ferrier H, Steer P. 2004. Exposure of pregnant women to tap water related activities. Occup Environ Med 61:454-460.

Kelly Y, Panico L, Bartley M, Marmot M, Nazroo J, Sacker A. 2009. Why does birthweight vary among ethnic groups in the uk? Findings from the millennium cohort study. Journal of public health (Oxford, England) 31:131-137.

Kim H. 1997. Human exposure to dichloroacetic acid and trichloroacetic acid from chlorinated water during household use and swimming. New Brunswick, NJ, US:Rutgers, The State University of New Jersey.

Kim H, Weisel CP. 1998. Dermal absorption of dichloro- and trichloroacetic acids from chlorinated water. J Expo Anal Env Epid 8:555-575.

Kim J, Chung Y, Shin D, Kim M, Lee Y, Lim Y, et al. 2002. Chlorination by-products in surface water treatment process. . Desalination 151:1-9.

Kim YJ, Hong YC, Lee KH, Park HJ, Park EA, Moon HS, et al. 2005. Oxidative stress in pregnant women and birth weight reduction. Reprod Toxicol 19:487-492.

King WD, Dodds L, Armson BA, Allen AC, Fell DB, Nimrod C. 2004. Exposure assessment in epidemiologic studies of adverse pregnancy outcomes and disinfection byproducts. J Expo Anal Environ Epidemiol 14:466-472.

King WD, Dodds L, Allen AC, Armson BA, Fell D, Nimrod C. 2005. Haloacetic acids in drinking water and risk for stillbirth. Occup Environ Med 62:124-127.

Klebanoff MA, Schulsinger C, Mednick BR, Secher NJ. 1997. Preterm and small-for-gestational-age birth across generations. Am J Obstet Gynecol 176:521-526.

Klemmensen AK, Olsen SF, Osterdal ML, Tabor A. 2007. Validity of preeclampsia-related diagnoses recorded in a national hospital registry and in a postpartum interview of the women. Am J Epidemiol 166:117-124.

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

Komulainen H. 2004. Experimental cancer studies of chlorinated by-products. Toxicology 198:239-248.

Konje JC, Cade JE, Study Group C. 2008. Maternal caffeine intake during pregnancy and risk of fetal growth restriction: A large prospective observational study. BMJ 337:a2332.

Koopmans CM, Bijlenga D, Groen H, Vijgen SM, Aarnoudse JG, Bekedam DJ, et al. 2009. Induction of labour versus expectant monitoring for gestational hypertension or mild pre-eclampsia after 36 weeks' gestation (hypitat): A multicentre, open-label randomised controlled trial. Lancet 374:979-988.

Kramer MD, Lynch CF, Isacson P, Hanson JW. 1992. The association of waterborne chloroform with intrauterine growth retardation. Epidemiology 3:407-413.

Kramer MS. 1987. Determinants of low birth weight: Methodological assessment and meta-analysis. Bulletin of the World Health Organization 65:663-737.

Kramer MS, Seguin L, Lydon J, Goulet L. 2000. Socio-economic disparities in pregnancy outcome: Why do the poor fare so poorly? Paediatr Perinat Epidemiol 14:194-210.

Krasner SW, McGuire MJ, Jacangelo JG, Patania NL, Reagan KM, Aieta EM. 1989. The occurrence of disinfection by-products in us drinking water. . Journal of the American Water Works Association 81:41-53.

Krasner SW. 1999. Chemistry of disinfection by-product formation. In: Formation and control of disinfection by-products in drinking wate, (Singer PC, ed). Denver, CO, US:American Water Works Association, 27-52.

Krasner SW, Wright JM. 2005. The effect of boiling water on disinfection by-product exposure. Water Res 39:855-864.

Krasner SW, Weinberg HS, Richardson SD, Pastor SJ, Chinn R, Sclimenti MJ, et al. 2006. Occurrence of a new generation of disinfection byproducts. Environ Sci Technol 40:7175-7185.

Krasner SW. 2009. The formation and control of emerging disinfection by-products of health concern. Philos Transact A Math Phys Eng Sci 367:4077-4095.

Lad M. 2011. The english indices of deprivation 2010 Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6871/1871208.pdf [accessed 18/10/2013 2013].

Larson JL, Bull RJ. 1992. Metabolism and lipoperoxidative activity of trichloroacetate and dichloroacetate in rats and mice. Toxicol Appl Pharmacol 115:268-277.

Lawlor DA, Davey Smith G, Ebrahim S. 2003. Life course influences on insulin resistance: Findings from the british women's heart and health study. Diabetes care 26:97-103.

Lawlor DA, Owen CG, Davies AA, Whincup PH, Ebrahim S, Cook DG, et al. 2006. Sex differences in the association between birth weight and total cholesterol. A meta-analysis. Ann Epidemiol 16:19-25.

Lawlor DA, Leary S, Davey Smith G. 2009. Theoretical underpinning for the use of intergenerational studies in life course epidemiology. . In: Family matters: Designing, analysing and understanding family-based studies in life course epidemiology Oxford, UK.:Oxford University Press.

Lawlor DA, Mortensen L, Andersen AM. 2011. Mechanisms underlying the associations of maternal age with adverse perinatal outcomes: A sibling study of 264 695 danish women and their firstborn offspring. Int J Epidemiol 40:1205-1214.

Leary S, Fall C, Osmond C, Lovel H, Campbell D, Eriksson J, et al. 2006. Geographical variation in relationships between parental body size and offspring phenotype at birth. Acta obstetricia et gynecologica Scandinavica 85:1066-1079.

Lee PA, Chernausek SD, Hokken-Koelega AC, Czernichow P. 2003. International small for gestational age advisory board consensus development conference statement: Management of short children born small for gestational age, april 24-october 1, 2001. Pediatrics 111:1253-1261.

Lenton EA, Landgren BM, Sexton L. 1984. Normal variation in the length of the luteal phase of the menstrual cycle: Identification of the short luteal phase. Br J Obstet Gynaecol 91:685-689.

Leon DA. 2004. Biological theories, evidence, and epidemiology. Int J Epidemiol 33:1167-1171.

Leon DA, Moser KA. 2012. Low birth weight persists in south asian babies born in england and wales regardless of maternal country of birth. Slow pace of acculturation, physiological constraint or both? Analysis of routine data. Journal of epidemiology and community health 66:544-551.

Leonardi-Bee J, Smyth A, Britton J, Coleman T. 2008. Environmental tobacco smoke and fetal health: Systematic review and meta-analysis. Archives of disease in childhood Fetal and neonatal edition 93:F351-361.

Leung TN, Roach VJ, Lau TK. 1998. Incidence of preterm delivery in hong kong chinese. The Australian & New Zealand journal of obstetrics & gynaecology 38:138-141.

Levallois P, Gingras S, Marcoux S, Legay C, Catto C, Rodriguez M, et al. 2012. Maternal exposure to drinking-water chlorination by-products and small-for-gestational-age neonates. Epidemiology 23:267-276.

Levesque S, Rodriguez MJ, Serodes J, Beaulieu C, Proulx F. 2006. Effects of indoor drinking water handling on trihalomethanes and haloacetic acids. Water Res 40:2921-2930.

Leviton A, Cowan L. 2002. A review of the literature relating caffeine consumption by women to their risk of reproductive hazards. Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association 40:1271-1310.

Lewis C, Suffet IH, Ritz B. 2006. Estimated effects of disinfection by-products on birth weight in a population served by a single water utility. Am J Epidemiol 163:38-47.

Li CC, Maloney CA, Cropley JE, Suter CM. 2010. Epigenetic programming by maternal nutrition: Shaping future generations. Epigenomics 2:539-549.

Li CQ, Windsor RA, Perkins L, Goldenberg RL, Lowe JB. 1993. The impact on infant birth weight and gestational age of cotinine-validated smoking reduction during pregnancy. JAMA 269:1519-1524.

Liang L, Singer PC. 2003. Factors influencing the formation and relative distribution of haloacetic acids and trihalomethanes in drinking water. Environ Sci Technol 37:2920-2928.

Lin CC, Su SJ, River LP. 1991. Comparison of associated high-risk factors and perinatal outcome between symmetric and asymmetric fetal intrauterine growth retardation. Am J Obstet Gynecol 164:1535-1541; discussion 1541-1532.

Lin CC, Santolaya-Forgas J. 1998. Current concepts of fetal growth restriction: Part i. Causes, classification, and pathophysiology. Obstet Gynecol 92:1044-1055.

Lindley AA, Becker S, Gray RH, Herman AA. 2000. Effect of continuing or stopping smoking during pregnancy on infant birth weight, crown-heel length, head circumference, ponderal index, and brain:Body weight ratio. Am J Epidemiol 152:219-225.

Lindley J, Dale A, Dex S. 2004. Ethnic differences in women's demographic, family characteristics and economic activity profiles, 1992 to 2002. Labour Market trends 112:153-165.

Lira PI, Ashworth A, Morris SS. 1996. Low birth weight and morbidity from diarrhea and respiratory infection in northeast brazil. The Journal of pediatrics 128:497-504.

Liviac D, Creus A, Marcos R. 2010. Genotoxicity testing of three monohaloacetic acids in tk6 cells using the cytokinesis-block micronucleus assay. Mutagenesis 25:505-509.

Lockwood CJ. 1999. Stress-associated preterm delivery: The role of corticotropin-releasing hormone. Am J Obstet Gynecol 180:S264-266.

Luben TJ, Olshan AF, Herring AH, Jeffay S, Strader L, Buus RM, et al. 2007. The healthy men study: An evaluation of exposure to disinfection by-products in tap water and sperm quality. Environ Health Perspect 115:1169-1176.

Lunn D, Spiegelhalter D, Thomas A, Best N. 2009. The bugs project: Evolution, critique and future directions. Stat Med 28:3049-3067.

Lunn DJ, Thomas A, Best N, Spiegelhalter D. 2000. Winbugs -- a bayesian modelling framework: Concepts, structure, and extensibility. . Statistics and Computing 10:325-337.

Ma A. 2008. Investigation into different tap water household handling treatments to determine the effects on disinfectant by-products. London, UK:Imperial College London.

MacDorman MF, Declercq E, Zhang J. 2010. Obstetrical intervention and the singleton preterm birth rate in the united states from 1991-2006. Am J Public Health 100:2241-2247.

MaCH MaCHC. 2001. Confidential enquiry of stillbirths and deaths in infancy.CESDI

MacLeod S, Kiely JL. 1988. The effects of maternal age and parity on birthweight: A population-based study in new york city. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics 26:11-19.

Madan A, Holland S, Humbert JE, Benitz WE. 2002. Racial differences in birth weight of term infants in a northern california population. Journal of perinatology : official journal of the California Perinatal Association 22:230-235.

Magnus P, Jaakkola JJ, Skrondal A, Alexander J, Becher G, Krogh T, et al. 1999. Water chlorination and birth defects. Epidemiology 10:513-517.

Malliarou E, Collins C, Graham N, Nieuwenhuijsen MJ. 2005. Haloacetic acids in drinking water in the united kingdom. Water Res 39:2722-2730.

Mamelle N, Cochet V, Claris O. 2001. Definition of fetal growth restriction according to constitutional growth potential. Biol Neonate 80:277-285.

Margetts BM, Mohd Yusof S, Al Dallal Z, Jackson AA. 2002. Persistence of lower birth weight in second generation south asian babies born in the united kingdom. Journal of epidemiology and community health 56:684-687.

Martin JA, Hamilton BE, Ventura SJ, Osterman MJK, Wilson EC, Mathews TJ. 2012. Births: Final data for 2010. Natl Vital Stat Rep 61.

Maskiell KE, Heyworth JS, McCaul KA. 2006. Validation of a water consumption questionnaire for a study of the adverse health outcomes associated with disinfection by-products. Int J Environ Health Res 16:145-153.

Mastrobattista JM, Pschirrer ER, Hamrick MA, Glaser AM, Schumacher V, Shirkey BA, et al. 2004. Humerus length evaluation in different ethnic groups. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine 23:227-231.

Matsubasa T, Uchino T, Karashima S, Kondo Y, Maruyama K, Tanimura M, et al. 2002. Oxidative stress in very low birth weight infants as measured by urinary 8-ohdg. Free Radic Res 36:189-193.

McCormick MC. 1985. The contribution of low birth weight to infant mortality and childhood morbidity. N Engl J Med 312:82-90.

McCowan L, Stewart AW, Francis A, Gardosi J. 2004. A customised birthweight centile calculator developed for a new zealand population. The Australian & New Zealand journal of obstetrics & gynaecology 44:428-431.

McFadyen IR, Campbell-Brown M, Abraham R, North WR, Haines AP. 1984. Factors affecting birthweights in hindus, moslems and europeans. Br J Obstet Gynaecol 91:968-972.

McGuire MJ, McLain JL, Obolensky A. 2002. Information collection rule data analysis. Denver, CO, US.:American Water Works Association Research Foundation

McKeown T, Gibson JR. 1951. Observations of all births (23,970) in birmingham, 1947. Iv. Premature birth. British medical journal 2:513-517.

McRae BM, LaPara TM, Hozalski RM. 2004. Biodegradation of haloacetic acids by bacterial enrichment cultures. Chemosphere 55:915-925.

Meire HB, Farrant P. 1981. Ultrasound demonstration of an unusual fetal growth pattern in indians. Br J Obstet Gynaecol 88:260-263.

Meis PJ, Michielutte R, Peters TJ, Wells HB, Sands RE, Coles EC, et al. 1997. Factors associated with term low birthweight in cardiff, wales. Paediatr Perinat Epidemiol 11:287-297.

Metzger BE, Coustan DR. 1998. Summary and recommendations of the fourth international workshop-conference on gestational diabetes mellitus. The organizing committee. Diabetes care 21 Suppl 2:B161-167.

Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, Coustan DR, et al. 2008. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 358:1991-2002.

Meyer MB. 1978. How does maternal smoking affect birth weight and maternal weight gain? Evidence from the ontario perinatal mortality study. Am J Obstet Gynecol 131:888-893.

Morgen CS, Bjork C, Andersen PK, Mortensen LH, Nybo Andersen AM. 2008. Socioeconomic position and the risk of preterm birth--a study within the danish national birth cohort. Int J Epidemiol 37:1109-1120.

Moser K, Stanfield KM, Leon DA. 2008. Birthweight and gestational age by ethnic group, england and wales 2005: Introducing new data on births. Health statistics quarterly / Office for National Statistics:22-31, 34-55.

Muthayya S, Kurpad AV, Duggan CP, Bosch RJ, Dwarkanath P, Mhaskar A, et al. 2006. Low maternal vitamin b12 status is associated with intrauterine growth retardation in urban south indians. Eur J Clin Nutr 60:791-801.

Myatt L, Cui X. 2004. Oxidative stress in the placenta. Histochem Cell Biol 122:369-382.

Naeye RL, Tafari N. 1985. Biologic bases for international fetal growth curves. Acta paediatrica Scandinavica Supplement 319:164-169.

Nanda S, Savvidou M, Syngelaki A, Akolekar R, Nicolaides KH. 2011. Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. Prenatal diagnosis 31:135-141.

Narotsky MG, Best DS, McDonald A, Godin EA, Hunter ES, 3rd, Simmons JE. 2011. Pregnancy loss and eye malformations in offspring of f344 rats following gestational exposure to mixtures of regulated trihalomethanes and haloacetic acids. Reprod Toxicol 31:59-65.

Nazroo J. 2001. Ethnicity, class and health. London, UK.

NCI NCI. 1976. Report on the carcinogenesis bioassay of chloroform (Carcinogenesis Technical Report Series).Carcinogenesis Program, Division of Cancer Cause and Prevention.

Nieuwenhuijsen MJ, Toledano MB, Eaton NE, Fawell J, Elliott P. 2000a. Chlorination disinfection byproducts in water and their association with adverse reproductive outcomes: A review. Occup Environ Med 57:73-85.

Nieuwenhuijsen MJ, Toledano MB, Elliott P. 2000b. Uptake of chlorination disinfection by-products; a review and a discussion of its implications for exposure assessment in epidemiological studies. J Expo Anal Environ Epidemiol 10:586-599.

Nieuwenhuijsen MJ, Toledano MB, Bennett J, Best N, Hambly P, de Hoogh C, et al. 2008. Chlorination disinfection by-products and risk of congenital anomalies in england and wales. Environ Health Perspect 116:216-222.

Nieuwenhuijsen MJ, Grellier J, Smith R, Iszatt N, Bennett J, Best N, et al. 2009a. The epidemiology and possible mechanisms of disinfection by-products in drinking water. Philos Transact A Math Phys Eng Sci 367:4043-4076.

Nieuwenhuijsen MJ, Martinez D, Grellier J, Bennett J, Best N, Iszatt N, et al. 2009b. Chlorination disinfection by-products in drinking water and congenital anomalies: Review and meta-analyses. Environ Health Perspect 117:1486-1493.

Nieuwenhuijsen MJ, Smith R, Golfinopoulos S, Best N, Bennett J, Aggazzotti G, et al. 2009c. Health impacts of long-term exposure to disinfection by-products in drinking water in europe: Hiwate. J Water Health 7:185-207.

Nieuwenhuijsen MJ, Dadvand P, Grellier J, Martinez D, Vrijheid M. 2013. Environmental risk factors of pregnancy outcomes: A summary of recent meta-analyses of epidemiological studies. Environ Health 12:6.

Nikolaou AD, Lekkas TD, Golfinopoulos SK. 2004. Kinetics of the formation and decomposition of chlorination by-products in surface waters. . Chemical Engineering Journal 100:139-148.

Nissinen TK, Miettinen IT, Martikainen PJ, Vartiainen T. 2002. Disinfection by-products in finnish drinking waters. Chemosphere 48:9-20.

O'Callaghan MJ, Harvey JM, Tudehope DI, Gray PH. 1997. Aetiology and classification of small for gestational age infants. Journal of paediatrics and child health 33:213-218.

Oldfield MD, Donley P, Walwyn L, Scudamore I, Gregory R. 2007. Long term prognosis of women with gestational diabetes in a multiethnic population. Postgraduate medical journal 83:426-430.

ONS OoNS. 2012. Release: Characteristics of birth 1, england and wales, 2011.Vital Statistics Output Branch, Office of National Statistics.

Oster E. 2013. Expecting better: Why the conventional pregnancy wisdom is wrong-and what you really need to know. New York, New York, USA:The Penguin Press HC.

Ott WJ. 1993. Intrauterine growth retardation and preterm delivery. Am J Obstet Gynecol 168:1710-1715; discussion 1715-1717.

Owen CG, Whincup PH, Odoki K, Gilg JA, Cook DG. 2003. Birth weight and blood cholesterol level: A study in adolescents and systematic review. Pediatrics 111:1081-1089.

Owen P, Donnet ML, Ogston SA, Christie AD, Howie PW, Patel NB. 1996. Standards for ultrasound fetal growth velocity. Br J Obstet Gynaecol 103:60-69.

Pals JA, Ang JK, Wagner ED, Plewa MJ. 2011. Biological mechanism for the toxicity of haloacetic acid drinking water disinfection byproducts. Environ Sci Technol 45:5791-5797.

Pan H, Cole TJ. 2010. Lmsgrowth, a microsoft excel add-in to access growth references based on the lms method. http://wwwhealthforallchildrencouk Version 2.69.

Parker AJ, Davies P, Mayho AM, Newton JR. 1984. The ultrasound estimation of sex-related variations of intrauterine growth. Am J Obstet Gynecol 149:665-669.

Parrish JM, Austin EW, Stevens DK, Kinder DH, Bull RJ. 1996. Haloacetate-induced oxidative damage to DNA in the liver of male b6c3f1 mice. Toxicology 110:103-111.

Parvez S, Rivera-Nunez Z, Meyer A, Wright JM. 2011. Temporal variability in trihalomethane and haloacetic acid concentrations in massachusetts public drinking water systems. Environ Res 111:499-509.

Patel RR, Steer P, Doyle P, Little MP, Elliott P. 2004. Does gestation vary by ethnic group? A london-based study of over 122,000 pregnancies with spontaneous onset of labour. Int J Epidemiol 33:107-113.

Patelarou E, Kargaki S, Stephanou EG, Nieuwenhuijsen M, Sourtzi P, Gracia E, et al. 2011. Exposure to brominated trihalomethanes in drinking water and reproductive outcomes. Occup Environ Med 68:438-445.

Patra J, Bakker R, Irving H, Jaddoe VW, Malini S, Rehm J. 2011. Dose-response relationship between alcohol consumption before and during pregnancy and the risks of low birthweight, preterm birth and small for gestational age (sga)-a systematic review and meta-analyses. BJOG : an international journal of obstetrics and gynaecology 118:1411-1421.

Paz I, Gale R, Laor A, Danon YL, Stevenson DK, Seidman DS. 1995. The cognitive outcome of full-term small for gestational age infants at late adolescence. Obstet Gynecol 85:452-456.

Peck JD, Leviton A, Cowan LD. 2010. A review of the epidemiologic evidence concerning the reproductive health effects of caffeine consumption: A 2000-2009 update. Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association 48:2549-2576.

Peters RJB, Erkelens C, De Leer EWB, De Galan L. 1991. The analysis of halogenated acetic acids in dutch drinking water. Water research 25: 473-477.

Plewa MJ, Kargalioglu Y, Vankerk D, Minear RA, Wagner ED. 2002. Mammalian cell cytotoxicity and genotoxicity analysis of drinking water disinfection by-products. Environ Mol Mutagen 40:134-142.

Plewa MJ, Wagner ED, Richardson SD, Thruston AD, Jr., Woo YT, McKague AB. 2004. Chemical and biological characterization of newly discovered iodoacid drinking water disinfection byproducts. Environ Sci Technol 38:4713-4722.

Plewa MJ, Simmons JE, Richardson SD, Wagner ED. 2010. Mammalian cell cytotoxicity and genotoxicity of the haloacetic acids, a major class of drinking water disinfection by-products. Environ Mol Mutagen 51:871-878.

Popkin BM, Barclay DV, Nielsen SJ. 2005. Water and food consumption patterns of u.S. Adults from 1999 to 2001. Obesity research 13:2146-2152.

Poppitt SD, Prentice AM, Goldberg GR, Whitehead RG. 1994. Energy-sparing strategies to protect human fetal growth. Am J Obstet Gynecol 171:118-125.

Porter CK, Putnam SD, Hunting KL, Riddle MR. 2005. The effect of trihalomethane and haloacetic acid exposure on fetal growth in a maryland county. Am J Epidemiol 162:334-344.

Prentice AM, Cole TJ. 1994. Seasonal changes in growth and energy status in the third world. The Proceedings of the Nutrition Society 53:509-519.

Pringle PJ, Geary MP, Rodeck CH, Kingdom JC, Kayamba-Kay's S, Hindmarsh PC. 2005. The influence of cigarette smoking on antenatal growth, birth size, and the insulin-like growth factor axis. J Clin Endocrinol Metab 90:2556-2562.

Rahman MD, Driscoll T, Clements M, Armstrong BK, Cowie CT. 2011. Effects of tap water processing on the concentration of disinfection by-products. J Water Health 9:507-514.

Ray JG, Laskin CA. 1999. Folic acid and homocyst(e)ine metabolic defects and the risk of placental abruption, pre-eclampsia and spontaneous pregnancy loss: A systematic review. Placenta 20:519-529.

Raymer JH, Pellizzari E, Childs B, Briggs K, Shoemaker JA. 2000. Analytical methods for water disinfection byproducts in foods and beverages. J Expo Anal Environ Epidemiol 10:808-815.

Raynor P. 2008. Born in bradford, a cohort study of babies born in bradford, and their parents: Protocol for the recruitment phase. BMC Public Health 8:327.

RCOG RCoOaG. 2002. The investigation and management of the small-for-gestational age fetus. London.

RCOG RCoOaG. 2006. Alcohol consumption and outcomes of pregnancy. . RCOG Statement No 5.

Reckhow DA, Singer PC. 1984. The removal of organic halide precursors by preozonation and alum coagulation. Journal of the American Water Works Association 76:151-157.

Reckhow DA, Singer PC. 1985. Mechanisms of organic halide formation during fulvic acid chlorination and implications with respect to preozonation. In: Water chlorination: Chemistry, environmental impact and health effects, Vol. 5, (Jolley RL, ed). Chelsea, MI, US, 1229-1257.

Reckhow DA, Singer PC, Malcolm RL. 1990. Chlorination of humic materials: Byproduct formation and chemical interpretations. Environmental Science and Technology 24 1655-1664.

Reif JS, Bachand A, Andersen M. 2000. Reproductive and developmental effects of disinfection by-products. (Final Report). Ottawa, Canada:Bureau of Reproductive and Child Health, Health Canada.

Resnik R. 2002. Intrauterine growth restriction. Obstet Gynecol 99:490-496.

Rice D, Barone S, Jr. 2000. Critical periods of vulnerability for the developing nervous system: Evidence from humans and animal models. Environ Health Perspect 108 Suppl 3:511-533.

Richards M, Hardy R, Kuh D, Wadsworth ME. 2002. Birthweight, postnatal growth and cognitive function in a national uk birth cohort. Int J Epidemiol 31:342-348.

Richardson S, Gilks WR. 1993. Conditional independence models for epidemiological studies with covariate measurement error. Stat Med 12:1703-1722.

Richardson SD, Plewa MJ, Wagner ED, Schoeny R, Demarini DM. 2007. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: A review and roadmap for research. Mutat Res 636:178-242.

Roberts CL, Ford JB, Algert CS, Antonsen S, Chalmers J, Cnattingius S, et al. 2011. Population-based trends in pregnancy hypertension and pre-eclampsia: An international comparative study. BMJ open 1:e000101.

Rodriguez M, Serodes J, Levesque S. 2004. Behavior of trihalomethanes and haloacetic acids in a drinking water distribution system. . Water research 38:4367-4382.

Rona RJ, Chinn S. 1986. National study of health and growth: Social and biological factors associated with height of children from ethnic groups living in england. Annals of human biology 13:453-471.

Rook JJ. 1974. Formation of haloforms during chlorination of natural waters. . Water Treatment and Examination 23:234-243.

Roquer JM, Figueras J, Botet F, Jimenez R. 1995. Influence on fetal growth of exposure to tobacco smoke during pregnancy. Acta paediatrica (Oslo, Norway : 1992) 84:118-121.

Ros HS, Cnattingius S, Lipworth L. 1998. Comparison of risk factors for preeclampsia and gestational hypertension in a population-based cohort study. Am J Epidemiol 147:1062-1070.

Roseboom TJ, van der Meulen JH, Ravelli AC, Osmond C, Barker DJ, Bleker OP. 2001. Effects of prenatal exposure to the dutch famine on adult disease in later life: An overview. Twin research : the official journal of the International Society for Twin Studies 4:293-298.

Rosenberg TJ, Raggio TP, Chiasson MA. 2005. A further examination of the "epidemiologic paradox": Birth outcomes among latinas. Journal of the National Medical Association 97:550-556.

Rubin DB. 1987. Multiple imputation for nonresponse in surveys. New York, USA.

Rubin DB. 1996. Multiple imputation after 18+ years. Journal of the American Statistical Society 91:473–489.

SACN/RCPCH Expert Group. 2007. Application of who growth standards in the uk. http://www.sacn.gov.uk/pdfs/report_growth_standards_2007_08_10.pdf.

Sadiq R, Rodriguez MJ. 2004. Disinfection by-products (dbps) in drinking water and predictive models for their occurrence: A review. Sci Total Environ 321:21-46.

Salmasi G, Grady R, Jones J, McDonald SD. 2010. Environmental tobacco smoke exposure and perinatal outcomes: A systematic review and meta-analyses. Acta obstetricia et gynecologica Scandinavica 89:423-441.

Sapolsky RM. 2005. The influence of social hierarchy on primate health. Science 308:648-652.

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

Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. 2002. Epidemiologic measures of the course and outcome of pregnancy. Epidemiologic reviews 24:91-101.

Savitz DA, Singer PC, Hartmann KE, Herring AH, Weinberg HS, Makarushka C, et al. 2005. Drinking water disinfection by-products and pregnancy outcome. Denver, Colorado:American Water Works Association (AWWA) Research Foundation.

Savitz DA, Singer PC, Herring AH, Hartmann KE, Weinberg HS, Makarushka C. 2006. Exposure to drinking water disinfection by-products and pregnancy loss. Am J Epidemiol 164:1043-1051.

Savitz DA. 2010. The etiology of epidemiologic perseveration: When enough is enough. Epidemiology 21:281-283.

Savitz DA. 2012. Invited commentary: Biomarkers of exposure to drinking water disinfection by-products--are we ready yet? Am J Epidemiol 175:276-278.

Scholl TO, Johnson WG. 2000. Folic acid: Influence on the outcome of pregnancy. Am J Clin Nutr 71:1295S-1303S.

Scholl TO, Stein TP. 2001. Oxidant damage to DNA and pregnancy outcome. J Matern Fetal Med 10:182-185.

Schwartz ID. 2000. Failure to thrive: An old nemesis in the new millennium. Pediatrics in review / American Academy of Pediatrics 21:257-264; quiz 264.

Serodes JB, Rodriguez MJ, Li H, Bouchard C. 2003. Occurrence of thms and haas in experimental chlorinated waters of the quebec city area (canada). Chemosphere 51:253-263.

Sheridan E, Wright J, Small N, Corry PC, Oddie S, Whibley C, et al. 2013. Risk factors for congenital anomaly in a multiethnic birth cohort: An analysis of the born in bradford study. Lancet.

Shields BM, Knight BA, Powell RJ, Hattersley AT, Wright DE. 2006. Assessing newborn body composition using principal components analysis: Differences in the determinants of fat and skeletal size. BMC pediatrics 6:24.

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

Shipp TD, Bromley B, Mascola M, Benacerraf B. 2001. Variation in fetal femur length with respect to maternal race. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine 20:141-144.

Singer PC. 1999. Formation and control of disinfection by-products in drinking water. Denver, CO, US:American Water Works Association,.

Singer PC, Weinberg HS, Brophy K, Liang L, Roberts M, Grisstede I, et al. 2002. Relative dominance of haas and thms in treated drinking water (report). Denver, Colorado, US:American Water Works Association Research Foundation.

Sizonenko SV, Borradori-Tolsa C, Bauthay DM, Lodygensky G, Lazeyras F, Huppi P. 2006. Impact of intrauterine growth restriction and glucocorticoids on brain development: Insights using advanced magnetic resonance imaging. Mol Cell Endocrinol 254-255:163-171.

Small N. 2012. Infant mortality and migrant health in babies of pakistani origin born in bradford, uk. J Intercult Stud 33:549-564.

Smith GC. 2004. First trimester origins of fetal growth impairment. Semin Perinatol 28:41-50.

Smith MK, Randall JL, Read EJ, Stober JA. 1989. Teratogenic activity of trichloroacetic acid in the rat. Teratology 40:445-451.

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

Smith RB, Toledano MB, Wright J, Raynor P, Nieuwenhuijsen MJ. 2009. Tap water use amongst pregnant women in a multi-ethnic cohort. Environ Health 8 Suppl 1:S7.

Smith RB. 2011. Assessment and validation of exposure to disinfection byproducts during pregnancy, in an epidemiological study examining associated risk of adverse fetal growth outcomes. . London, UK.:Imperial College London.

Smith RB, Nieuwenhuijsen MJ, Wright J, Raynor P, Cocker J, Jones K, et al. 2012. Biomarker validation of trichloroacetic acid exposure via drinking water during pregnancy. Submitted (Environmental Research).

Smith RB, Nieuwenhuijsen MJ, Wright J, Raynor P, Cocker J, Jones K, et al. 2013. Validation of trichloroacetic acid exposure via drinking water during pregnancy using a urinary tcaa biomarker. Environ Res.

Smith RB, Edwards SC, Wright J, Nieuwenhuijsen MJ, Toledano MB. in preparation. Patterns of tap water use during pregnancy in the born in bradford cohort. Pediatrics.

Smith SC, Baker PN, Symonds EM. 1997. Increased placental apoptosis in intrauterine growth restriction. Am J Obstet Gynecol 177:1395-1401.

Sohn J, Amy G, Cho J, Lee Y, Yoon Y. 2004. Disinfectant decay and disinfection by-products formation model development: Chlorination and ozonation by-products. Water Res 38:2461-2478.

Spencer JA, Chang TC, Robson SC, Gallivan S. 1995. Fetal size and growth in bangladeshi pregnancies. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 5:313-317.

Spencer N, Logan S. 2002. Social influences on birth weight. Journal of epidemiology and community health 56:326-327.

Spencer NJ, Logan S, Gill L. 1999. Trends and social patterning of birthweight in sheffield, 1985-94. Archives of disease in childhood Fetal and neonatal edition 81:F138-140.

StataCorp. 2009. Stata statistical software. Part Release 11. College Station, TX:StataCorp LP.

Sung W, Reilley-Matthews B, O'Day DK, Horrigan K. 2000. Modeling dbp formation. Journal of the American Water Works Association 92:53-63.

Tamura RK, Sabbagha RE, Depp R, Vaisrub N, Dooley SL, Socol ML. 1984. Diminished growth in fetuses born preterm after spontaneous labor or rupture of membranes. Am J Obstet Gynecol 148:1105-1110.

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

Taylor DJ, Howie PW. 1989. Fetal growth achievement and neurodevelopmental disability. Br J Obstet Gynaecol 96:789-794.

Thame M, Wilks RJ, McFarlane-Anderson N, Bennett FI, Forrester TE. 1997. Relationship between maternal nutritional status and infant's weight and body proportions at birth. Eur J Clin Nutr 51:134-138.

Thomas P, Peabody J, Turnier V, Clark RH. 2000. A new look at intrauterine growth and the impact of race, altitude, and gender. Pediatrics 106:E21.

Toledano MB, Nieuwenhuijsen MJ, Best N, Whitaker H, Hambly P, de Hoogh C, et al. 2005. Relation of trihalomethane concentrations in public water supplies to stillbirth and birth weight in three water regions in england. Environ Health Perspect 113:225-232.

Traviss GD, West RM, House AO. 2012. Maternal mental health and its association with infant growth at 6 months in ethnic groups: Results from the born-in-bradford birth cohort study. PloS one 7:e30707.

Tu YK, West R, Ellison GT, Gilthorpe MS. 2005. Why evidence for the fetal origins of adult disease might be a statistical artifact: The "reversal paradox" for the relation between birth weight and blood pressure in later life. Am J Epidemiol 161:27-32.

Tung H-H, Unz RF, Xie YF. 2006. Haa removal by gac adsorption. Journal of the American Water Works Association 98:107-112.

United Nations Children's Fund WHO, .;. 2004. Low birthweight: Country, regional and global estimates. New York, Geneva:UNICEF, WHO.

van Wassenaer A. 2005. Neurodevelopmental consequences of being born sga. Pediatr Endocrinol Rev 2:372-377.

Vangen S, Stoltenberg C, Skjaerven R, Magnus P, Harris JR, Stray-Pedersen B. 2002. The heavier the better? Birthweight and perinatal mortality in different ethnic groups. Int J Epidemiol 31:654-660.

Vik T, Jacobsen G, Vatten L, Bakketeig LS. 1996. Pre- and post-natal growth in children of women who smoked in pregnancy. Early Hum Dev 45:245-255.

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

Villar J, Ba'aqeel H, Piaggio G, Lumbiganon P, Miguel Belizan J, Farnot U, et al. 2001. Who antenatal care randomised trial for the evaluation of a new model of routine antenatal care. Lancet 357:1551-1564.

Waller K, Swan SH, Windham GC, Fenster L. 2001. Influence of exposure assessment methods on risk estimates in an epidemiologic study of total trihalomethane exposure and spontaneous abortion. J Expo Anal Environ Epidemiol 11:522-531.

Wang W, Ye B, Yang L, Li Y, Wang Y. 2007. Risk assessment on disinfection by-products of drinking water of different water sources and disinfection processes. Environ Int 33:219-225.

Waterland RA, Michels KB. 2007. Epigenetic epidemiology of the developmental origins hypothesis. Annual review of nutrition 27:363-388.

Watson W. 1993. Mathematical modelling of the formation of thms and haas in chlorinated natural waters. Denver, CO, US.:American Water Works Association Research Foundation.

Wegienka G, Baird DD. 2005. A comparison of recalled date of last menstrual period with prospectively recorded dates. Journal of women's health (2002) 14:248-252.

Weinberg HS, Krasner SW, Richardson SD, Thruston ADJ. 2002. The occurrence of disinfection by-products (dbps) of health concern in drinking water: Results of a nationwide dbp occurrence study. Fed Reg EPA/600/R-02/068.

Weinberg HS, Pereira VR, Singer PC, Savitz DA. 2006. Considerations for improving the accuracy of exposure to disinfection by-products by ingestion in epidemiologic studies. Sci Total Environ 354:35-42.

Weisel CP, Chen WJ. 1994. Exposure to chlorination by-products from hot water uses. Risk analysis : an official publication of the Society for Risk Analysis 14:101-106.

Wen SW, Fung KF, Huang L, Demissie K, Joseph KS, Allen AC, et al. 2005. Fetal and neonatal mortality among twin gestations in a canadian population: The effect of intrapair birthweight discordance. American journal of perinatology 22:279-286.

West J. 2011. Explaining differences in birth size and adiposity between pakistani and white british babies. Leeds, UK:The University of Leeds.

West J, Manchester B, Wright J, Lawlor DA, Waiblinger D. 2011. Reliability of routine clinical measurements of neonatal circumferences and research measurements of neonatal skinfold thicknesses: Findings from the born in bradford study. Paediatr Perinat Epidemiol 25:164-171.

West J, Lawlor DA, Fairley L, Bhopal R, Cameron N, McKinney PA, et al. 2013. Uk-born pakistani-origin infants are relatively more adipose than white british infants: Findings from 8704 mother-offspring pairs in the born-in-bradford prospective birth cohort. Journal of epidemiology and community health 67:544-551.

Westerhoff P, Chao P, Mash H. 2004. Reactivity of natural organic matter with aqueous chlorine and bromine. Water Res 38:1502-1513.

Whincup PH, Kaye SJ, Owen CG, Huxley R, Cook DG, Anazawa S, et al. 2008. Birth weight and risk of type 2 diabetes: A systematic review. JAMA 300:2886-2897.

Whitaker HJ, Nieuwenhuijsen MJ, Best N, Fawell J, Gowers A, Elliott P. 2003a. Description of trihalomethane levels in three uk water suppliers. Journal Exposure Analysis and Environmental Epidemiology 13:17-23.

Whitaker HJ, Nieuwenhuijsen MJ, Best NG. 2003b. The relationship between water concentrations and individual uptake of chloroform: A simulation study. Environ Health Perspect 111:688-694.

WHO WHO. 2008. Guidelines for drinking water quality. 3rd:451-455.

Wigle DT, Arbuckle TE, Turner MC, Berube A, Yang Q, Liu S, et al. 2008. Epidemiologic evidence of relationships between reproductive and child health outcomes and environmental chemical contaminants. Journal of toxicology and environmental health Part B, Critical reviews 11:373-517.

Wilcox AJ, Dunson D, Baird DD. 2000. The timing of the "fertile window" in the menstrual cycle: Day specific estimates from a prospective study. BMJ 321:1259-1262.

Wilcox AJ. 2001. On the importance--and the unimportance--of birthweight. Int J Epidemiol 30:1233-1241.

Wilcox AJ. 2010. Fertility and pregnancy: An epidemiologic perspective:OUP USA.

Wilcox M, Gardosi J, Mongelli M, Ray C, Johnson I. 1993. Birth weight from pregnancies dated by ultrasonography in a multicultural british population. BMJ 307:588-591.

Wilcox MA, Chang AM, Johnson IR. 1996. The effects of parity on birthweight using successive pregnancies. Acta obstetricia et gynecologica Scandinavica 75:459-453.

Williams DT, LeBel GL, Benoit FM. 1997. Disinfection by-products in canadian drinking water. . Chemosphere 34:299-316.

Williams RL, Creasy RK, Cunningham GC, Hawes WE, Norris FD, Tashiro M. 1982. Fetal growth and perinatal viability in california. Obstet Gynecol 59:624-632.

Windham GC, Swan SH, Fenster L, Neutra RR. 1992. Tap or bottled water consumption and spontaneous abortion: A 1986 case-control study in california. Epidemiology 3:113-119.

Wood SN. 2003. Thin plate regression splines. J R Statist Soc B 65:95–114.

Wright CM, Parkinson KN. 2004. Postnatal weight loss in term infants: What is normal and do growth charts allow for it? Archives of disease in childhood Fetal and neonatal edition 89:F254-257.

Wright CM, Williams AF, Elliman D, Bedford H, Birks E, Butler G, et al. 2010. Using the new uk-who growth charts. BMJ 340:c1140.

Wright J, Small N, Raynor P, Tuffnell D, Bhopal R, Cameron N, et al. 2012. Cohort profile: The born in bradford multi-ethnic family cohort study International Journal of Epidemiology 1-14.

Wright JM, Schwartz J, Vartiainen T, Maki-Paakkanen J, Altshul L, Harrington JJ, et al. 2002. 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5h)-furanone (mx) and mutagenic activity in massachusetts drinking water. Environ Health Perspect 110:157-164.

Wright JM, Schwartz J, Dockery DW. 2003. Effect of trihalomethane exposure on fetal development. Occup Environ Med 60:173-180.

Wright JM, Schwartz J, Dockery DW. 2004. The effect of disinfection by-products and mutagenic activity on birth weight and gestational duration. Environ Health Perspect 112:920-925.

Wright JM, Bateson TF. 2005. A sensitivity analysis of bias in relative risk estimates due to disinfection by-product exposure misclassification. J Expo Anal Environ Epidemiol 15:212-216.

Wright JM, Murphy PA, Nieuwenhuijsen MJ, Savitz DA. 2006. The impact of water consumption, point-of-use filtration and exposure categorization on exposure misclassification of ingested drinking water contaminants. Sci Total Environ 366:65-73.

Wright JM, Hoffman CS, Savitz DA. 2010. The relationship between water intake and foetal growth and preterm delivery in a prospective cohort study. BMC Pregnancy Childbirth 10:48.

Wu WW, Benjamin MM, Korshin GV. 2001. Effects of thermal treatment on halogenated disinfection by-products in drinking water. Water Res 35:3545-3550.

Xie Y. 2003. Disinfection by-products in drinking water: Formation, analysis and control. Chelsea, MI, US.

Xu X, Mariano TM, Laskin JD, Weisel CP. 2002. Percutaneous absorption of trihalomethanes, haloacetic acids, and haloketones. Toxicol Appl Pharmacol 184:19-26.

Xu X, Weisel CP. 2003. Inhalation exposure to haloacetic acids and haloketones during showering. Environ Sci Technol 37:569-576.

Xu X, Weisel CP. 2005. Human respiratory uptake of chloroform and haloketones during showering. J Expo Anal Environ Epidemiol 15:6-16.

Yasmin S, Osrin D, Paul E, Costello A. 2001. Neonatal mortality of low-birth-weight infants in bangladesh. Bulletin of the World Health Organization 79:608-614.

Zaren B, Lindmark G, Gebre-Medhin M. 1996. Maternal smoking and body composition of the newborn. Acta paediatrica (Oslo, Norway : 1992) 85:213-219.

Zeitlin J, Ancel PY, Saurel-Cubizolles MJ, Papiernik E. 2000. The relationship between intrauterine growth restriction and preterm delivery: An empirical approach using data from a european case-control study. BJOG : an international journal of obstetrics and gynaecology 107:750-758.

Zender R, Bachand AM, Reif JS. 2001. Exposure to tap water during pregnancy. J Expo Anal Environ Epidemiol 11:224-230.

Zhang J, Sun K. 2013. Invited commentary: The incremental value of customization in defining abnormal fetal growth status. Am J Epidemiol 178:1309-1312.

Zhang SH, Miao DY, Liu AL, Zhang L, Wei W, Xie H, et al. 2010. Assessment of the cytotoxicity and genotoxicity of haloacetic acids using microplate-based cytotoxicity test and cho/hgprt gene mutation assay. Mutat Res 703:174-179.

Zhang X, Minear RA. 2000. Decomposition of trihaloacetic acids and formation of the corresponding trihalomethanes in water. . 219th ACS National Meeting Preprints of Extended Abstracts 40:34-36.

Zhang Y, Collins C, Graham N, M.R. T, J. H, Nieuwenhuijsen M. 2010. Speciation and variation in the occurrence of haloacetic acids in three water supply systems in england. Water and Environment Journal 24:237-245.

Zhou WS, Xu L, Xie SH, Li YL, Li L, Zeng Q, et al. 2012. Decreased birth weight in relation to maternal urinary trichloroacetic acid levels. Sci Total Environ 416:105-110.

APPENDICES TO CHAPTERS 2, 3, 4, 5, 7 AND 8

For each chapter, tables are presented first, then figures.

Chapter 2

Table A2 - 1: Codebook of variables under consideration in this thesis including source of data (data was either self-reported by participants in the baseline questionnaire administered at recruitment to BiB, or drawn from the NHS maternal records system eClipse)

Name	Source	Description	Coding
ethgrp3	questionnaire	Ethnic group	1= White British (reference) 2= Pakistani origin 3= Other
eclmumageUP_cat2	eClipse	Mother's age at time of delivery	1= 25-29 years (reference) 2= <25 years 3 = 30-34 years 4= ≥ 35 years
eclregpartUP_cat	eClipse	Parity	0= no previous registerable births (reference) 1= one previous registerable birth 2= two or more previous registerable births
mumeducation_new2	questionnaire	Maternal education	1= School (reference) 2= No formal education 3= Further education 4= Higher education 5= Other, Don't know, and Foreign education (of unknown equivalency)
smokingX0_new	questionnaire	Maternal active smoking during pregnancy	1= never a smoker (reference) 2= ever a smoker 3= current smoker
caffdrink_cat2	questionnaire	A binary variable derived from the conversion of L/day to mg/day of caffeine consumption. It is split at a cut-point of 200mg/day which is the current recommended maximum for caffeine intake during pregnancy	0 = 0-200 mg/day (reference) 1 = >200 mg/day
drvgesdiabUP	eClipse	Gestational diabetes (derived from the GTT and backfill notes)	0= no 1= yes
eclbabysexUP	eClipse	Sex of child	0= male 1= female
eclgestwksUP	eClipse	Number of completed weeks of gestation	range: 25-44 weeks
BMIquest_quart	questionnaire	Maternal BMI (in quartiles) based on maternal weight at questionnaire completion (see Chapter 7)	1= quartile 1 (reference) 2= quartile 2 3= quartile 3 4= quartile 4

Name	Source	Description	Coding
workstatX0_new	questionnaire	Employment status; derived from 4 other variables: a) whether the mother ever worked (job0evrwrk), b) is currently employed (job0curemp), c) is a full time student (job0studen) or d) is currently is on maternity/sick leave (job0matscl)	1 = employed 2 = not employed/student
imd_2010_quintile_nat	(*)	IMD quintiles of deprivation 2010	1= quintile 1 (most deprived) (reference) 2= quintile 2 3= quintile 3 4= quintile 4 5= quintile 5 (least deprived)
alcohol2bis	questionnaire	Alcohol pregnancy during pregnancy and/or 3 months before	1= no 2= yes
bkfdiabeteUP_2	eClipse	Pre-existing diabetes	0=no 1= yes (includes Type 1, Type 2 and unknown type)
bkfhyperpiUP_new 2	eClipse	Pregnancy-induced hypertension	0= no 1= yes (includes mild to moderate, severe and not classified)
bkfhyperexUP	eClipse	Pre-existing hypertension	0= no 1= yes
bkfhyperlbUP	eClipse	Hypertension during labour only	0= no 1= yes
bkfpreeclmUP	eClipse	Pre-eclampsia	0=no 1= yes

(*) derived by the Department for Communities and Local Government (Lad 2011)

Chapter 3

Table A3 - 1: Proportion who reported consumption at Home

n total = 11,928	valid n	%	possible total for each water type	Filterers	% Filterers	Non Filterers	% Non Filterers	Filterers and Non Filterers	% Filterers and Non Filterers
Cold tap water at Home (L/day)									
Non zero, valid values	9,112	76.4	11,928						
zero values	44	0.4							
missing	2,772	23.2							
Filtered cold tap water at Home	909	10.0		909	10.0			9,071	99.6
Did not filter cold tap water at Home	8,162	89.6				8,162	89.6		
Doesn't know her filtering at Home	31	0.3							
Information on filtering at Home missing	10	0.1							
Hot tap water at Home (L/day)									
Non zero, valid values	7,576	63.5	11,928						
zero values	118	1.0							
missing	4,234	35.5							
Total tap water at Home (L/day)									
Non zero, valid values	9,679	81.1	11,928						
zero values	9	0.1							
missing	2,240	18.8							
Bottled water at Home (L/day)									
Non zero, valid values	1,425	11.9	11,928						
zero values	420	3.5							
missing	10,083	84.5							

n total = 11,928	valid n	%	possible total for each water type	Filterers	% Filterers	Non Filterers	% Non Filterers	Filterers and Non Filterers	% Filterers and Non Filterers
Total water at Home (L/day)									
Non zero, valid values	9,788	82.1	11,928						
zero values	1	0.0							
missing	2,139	17.9							

Table A3 - 2: Proportion who reported consumption Outside the Home

n total = 11,928	valid n	%	possible total	Filterers	% Filterers	Non Filterers	% Non Filterers	Filterers and Non Filterers	% Filterers and Non Filterers
Cold tap water at Outside the Home (L/day)									
Non zero, valid values	1,749	14.7	11,928						
zero values	79	0.7							
missing	10,100	84.7							
Filtered cold tap water at Work [8]	633	36.2		633	36.2			1,466	83.8
Did not filter cold tap water at Work [9]	833	47.6				833	47.6		
Doesn't know her filtering at Work [10]	43	2.5							
Information on Filtering at Work missing [11]	0	0.0							
Inconsistent [12]	53	3.0							
Not Applicable [13]	187	10.7							
Hot tap water at Outside the Home (L/day)									
Non zero, valid values	1,661	13.9	11,928						
zero values	74	0.6							
missing	10,193	85.5							
Total tap water at Outside the Home (L/day)									
Non zero, valid values	2,715	22.8	11,928						
zero values	51	0.4							
missing	9,162	85.5							
Bottled water at Outside the Home (L/day)									
Non zero, valid values	2,093	17.5	11,928						
zero values	346	2.9							
missing	9,489	79.6							
Total water at Outside the Home (L/day)									

n total = 11,928	valid n	%	possible total	Filterers	% Filterers	Non Filterers	% Non Filterers	Filterers and Non Filterers	% Filterers and Non Filterers
Non zero, valid values	3,946	33.1	11,928						
zero values	23	0.2							
missing	7,959	66.7							

Legend for both Table A3 - 1 and Table A3 - 2

[1] Employed, and Filtered cold tap water at Home and Work (Home is Yes; Work is Yes)

[2] Employed, and Filtered cold tap water at Home only (Home is Yes; Work is No, Don't know, NA or missing); Out of employment, and Filtered cold tap water at Home (Home=1)

[3] Employed, and Filtered cold tap water at Work only (Home is No, Don't know or missing; Work is Yes)

[4] Employed, and Did not filter cold tap water at all (Home is No; Work is No); Employed, and Did not filter cold tap water at Home (Home is No; Work is Don't know, NA or missing); Employed, and Did not filter cold tap water at Work (Home is Don't know or missing; Work is No); Out of employment, and Did not filter cold tap water at Home (Home=2)

[5] Employed, and Doesn't know her filtering either at Home or Work (Home is Don't know, Work is Don't know); Employed, and Doesn't know her filtering at Home (Home is Don't know; Work is NA or missing); Employed, and Doesn't know her filtering at Work (Home is missing; Work is Don't know); Out of employment, and Doesn't know her filtering at Home (Home=3)

[6] Employed, and Information on filtering at Home missing (Home is missing; Work is NA or missing); Out of employment, and information on filtering at Home missing

[7] Employed, but didn't report drinking cold tap water at Home (=0); Employed, but didn't report drinking cold tap water at Home (=missing); Employed, but didn't report drinking cold tap water at Work (=0); Employed, but didn't report drinking cold tap water at Work (=missing); Out of employment, and didn't report drinking cold water at home (=0); Out of employment, and didn't report drinking cold water at home (=missing)

And

[8] Employed, and Filtered tap water at Work

[9] Employed, and Did not filter cold tap water at Work

[10] Employed, and Doesn't know her filtering at Work; Employed, and Information on filtering at Work N/A

[11] Employed, and Information on filtering at Work missing

[12] Employed, but didn't report drinking cold tap water at Work (=0); Employed, but didn't report drinking cold tap water at Work (=missing)

[13] Not employed, full-time students or missing employment status

Chapter 4

Table A4 - 1: Relationships between HAA levels and HAA determinants

Relationship		references
[HAA] increases as	pH decreases	Krasner 1999; Singer 1999
	temperature increases*	Krasner et al. 1989; Williams et al. 1997 ; Mallariou et al. 2005
	bromide decreases	Diehl et al. 2000
	chlorine dose increases	Carlson and Hardy 1998
	residence time decreases	Chen and Weisel 1998

*typically, but other factors to take into consideration too, such as changes in the raw water quality and the nature of natural organic matter (NOM)

Table A4 - 2: HAA predictive models (adapted from a table in Bougeard's PhD thesis (2009))

Reference	Species	Predictive model for HAA	R^2	Water source	Comment
Watson (1993)	MCAA	$1.634\ (TOC)^{0.753}\ (Br^- + 0.01)^{-0.085}\ (pH)^{-1.124}\ (Cl_2)^{0.509}\ (t)^{0.300}$	0.82	not real water	water quality and chlorination conditions do not represent situation encountered in real water utilities
	DCAA	$0.605\ (TOC)^{0.291}\ (UV)^{0.726}\ (Br^- + 0.01)^{-0.568}\ (Cl_2)^{0.48}\ (t)^{0.239}\ (T)^{0.665}$	0.97		
	TCAA	$87.182\ (TOC)^{0.355}\ (UV)^{0.901}\ (Br^- + 0.01)^{0.679}\ (pH)^{1.732}\ (Cl_2)^{0.881}\ (t)^{0.264}$	0.98		
	MBAA	$0.176\ (TOC)^{1.664}\ (UV)^{-0.624}\ (Br^-)^{0.795}\ (pH)^{-0.927}\ (t)^{0.145}\ (T)^{0.45}$	0.8		
	DBAA	$84.945\ (TOC)^{-0.62}\ (UV)^{0.651}\ (Br^-)^{1.073}\ (Cl_2)\text{-}0.2\ (t)^{0.12}\ (T)^{0.657}$	0.95		
Amy et al. (1998)	HAA6	**DOC-based model** $HAA6 = 9.98\ (DOC)^{0.935}\ (Cl_2)^{0.443}\ (Br^-)^{-0.031}\ (T)^{0.387}\ (pH)^{-0.655}\ (t)^{0.178}$	0.87	raw water	specific to water quality and operating conditions
Amy et al. (1998)	HAA6	DOC-based model $HAA6 = 5.22\ (DOC)^{0.585}\ (Cl_2)^{0.565}\ (Br^-)^{-0.031}\ (t)^{0.153}$	0.92	coagulated water (alum or iron)	does not take pH and temperature into account
Sohn et al. (2004)	HAA6	UV-based model $HAA6 = 171.4\ (UV)^{0.584}\ (Cl_2)^{0.398}\ (Br^-)^{-0.091}\ (T)^{0.396}\ (pH)^{-0.645}\ (t)^{0.178}$	0.80	raw water	
Sohn et al. (2004)	HAA6	DOC*UV-based model $HAA6 = 101.2\ (DOC*UV)^{0.452}\ (Cl_2)^{0.194}\ (Br^-)^{-0.0698}\ (T)^{0.346}\ (pH)^{-0.623}\ (t)^{0.180}$	0.85	raw water	
Sohn et al. (2004)	HAA6	UV-based models (developed from EPA 1998 database) $HAA6 = 63.7\ (UV)^{0.419}\ (Cl_2)^{0.640}\ (Br^-)^{-0.066}\ (t)^{0.161}$	0.92	coagulated water (alum or iron)	does not take pH and temperature into account (use of below equation necessary for correction)
Sohn et al. (2004)	HAA6	DOC*UV-based models (developed from EPA 1998 database) $HAA6 = 30.7\ (DOC*UV)^{0.302}\ (Cl_2)^{0.541}\ (Br^-)^{-0.012}\ (t)^{0.161}$	0.94	coagulated water (alum or iron)	does not take pH and temperature into account (use of below equation necessary for correction)
Sohn et al. (2004)	HAA6	pH and temperature correction $HAA6 = (HAA6_{@pH=\ 7.5,\ T=\ 20°C})*(0.932)^{(pH-7.5)}\ (1.021)^{(T-20)}$	0.85	coagulated water (alum or iron)	this equation modifies the coagulated water HAA models, so that they are applicable under different pH and temperature
Sung et al. (2000)	HAA5	$4.8 * 10^4\ [OH^-]^{0.35}\ (C_0(1\text{-}\exp(-kt))^{0.43})(UV254)^{0.34}$	0.74	raw water	$[OH^-]$ calculated from the raw water pH and temperature

Reference	Species	Predictive model for HAA	R^2	Water source	Comment
Gang et al. (2002)	HAA9	Model based on **chlorine demand** $HAA9 = \beta C_0 [\ 1\text{-}fe^{-kR*t} - (1\text{-}f)e^{-kS*t}\ ]$	0.98	from raw to treated water	model can be applied accurately from raw to alum treated water, but may not perform well outside the typical conditions (pH=8±0.2, temperature=25°C, and chlorine residual=1.0±0.5 mg/l)
Villanueva et al. (2003)	HAAs	Linear **regression** in function of various THM species	0.57-0.97	NR	models do not consider chlorine dose, temperature etc.
Serodes et al. (2003)	HAAs	Single linear and non-linear regression models for water of single-utility	0.56-0.92	NR	variation according to the DBP and utility at stake

T= temperature (in °C)
t = time in hours
Cl_2= chlorine dose (mg/l)
β = ratio of the concentration of HAA9 formed (ug/l) to the concentration of chlorine consumed (mg/l)
C_0= initial chlorine concentration (mg/l)
f= fraction of the chlorine demand attributed to rapid reactions
k_R, k_S = first order rate constants for rapid and slow reactions, respectively
NR: not reported

Table A4 - 3: DCAA and possible predictors (r, p-value, N) (bold with * means rho>0.5; red means p-value <0.005)

	sqrt(DCAA)	temperature	UV254	TOC	colour	bromide	free chlorine	total chlorine	conductivity	pH	turbidity
sqrt(DCAA)	1 158										
temperature	-0.35 <0.001 134	1 148									
UV254	-0.32 <0.001 158	0.33 <0.001 148	1 176								
TOC	-0.18 0.023 158	0.30 <0.001 148	**0.89*** <0.001 176	1 176							
colour	-0.29 0.001 139	0.37 <0.001 127	**0.76*** <0.001 154	**0.74*** <0.001 154	1 154						
bromide	-0.15 0.060 156	0.09 0.282 146	0.18 0.017 174	0.21 0.006 174	0.20 0.014 152	1 174					
free chlorine	0.39 <0.001 140	-0.21 0.019 128	-0.20 0.012 156	-0.10 0.228 156	-0.22 0.007 151	-0.08 0.353 154	1 156				
total chlorine	0.38 <0.001 140	-0.23 0.010 128	-0.22 0.005 156	-0.11 0.158 156	-0.24 0.003 151	-0.09 0.268 154	**0.97*** <0.001 156	1 156			
conductivity	-0.22 0.010 139	0.10 0.254 127	-0.29 <0.001 154	-0.33 <0.001 154	-0.23 0.004 154	0.08 0.357 152	-0.09 0.294 151	-0.07 0.374 151	1 154		
pH	-0.26 0.002 139	0.17 0.064 127	-0.05 0.545 154	-0.10 0.223 154	0.04 0.603 154	0.08 0.316 152	-0.24 0.003 151	-0.27 0.001 151	0.23 0.004 154	1 154	
turbidity	-0.07 0.441 139	-0.25 0.005 128	0.10 0.233 155	0.03 0.757 155	0.20 0.014 154	0.12 0.154 153	-0.06 0.428 152	-0.05 0.521 152	0.08 0.308 154	0.06 0.425 154	1 155

Table A4 - 4: TCAA and possible predictors (r, p-value, N) (bold with * means rho>0.5; red means p-value <0.005)

	sqrt(TCAA)	temperature	UV254	TOC	colour	bromide	free chlorine	total chlorine	conductivity	pH	turbidity
sqrt(TCAA)	1 158										
temperature	0.31 <0.001 134	1 148									
UV254	0.34 <0.001 158	0.33 <0.001 148	1 176								
TOC	0.33 <0.001 158	0.30 <0.001 148	**0.89*** <0.001 176	1 176							
colour	0.11 0.184 139	0.37 <0.001 127	**0.76*** <0.001 154	**0.74*** <0.001 154	1 154						
bromide	-0.26 0.001 156	0.09 0.282 146	0.18 0.017 174	0.21 0.006 174	0.20 0.014 152	1 174					
free chlorine	0.05 0.562 140	-0.21 0.019 128	-0.20 0.012 156	-0.10 0.228 156	-0.22 0.007 151	-0.08 0.353 154	1 156				
total chlorine	0.01 0.867 140	-0.23 0.010 128	-0.22 0.005 156	-0.11 0.158 156	-0.24 0.003 151	-0.09 0.268 154	**0.97*** <0.001 156	1 156			
conductivity	-0.25 0.003 139	0.10 0.254 127	-0.29 <0.001 154	-0.33 <0.001 154	-0.23 0.004 154	0.08 0.357 152	-0.09 0.294 151	-0.07 0.374 151	1 154		
pH	-0.33 <0.001 139	0.17 0.064 127	-0.05 0.545 154	-0.10 0.223 154	0.04 0.603 154	0.08 0.316 152	-0.24 0.003 151	-0.27 0.001 151	0.23 0.004 154	1 154	
turbidity	-0.30 <0.001 139	-0.25 0.005 128	0.10 0.233 155	0.03 0.757 155	0.20 0.014 154	0.12 0.154 153	-0.06 0.428 152	-0.05 0.521 152	0.08 0.308 154	0.06 0.425 154	1 155

Table A4 - 5: BDCAA and possible predictors (r, p-value, N) (bold with * means rho>0.5; red means p-value <0.005)

	ln(BDCAA)	temperature	UV254	TOC	colour	bromide	free chlorine	total chlorine	conductivity	pH	turbidity
ln(BDCAA)	1 143										
temperature	-0.01 0.934 121	1 148									
UV254	-0.17 0.045 143	0.33 <0.001 148	1 176								
TOC	-0.16 0.062 143	0.30 <0.001 148	**0.89*** <0.001 176	1 176							
colour	-0.28 0.002 126	0.37 <0.001 127	**0.76*** <0.001 154	**0.74*** <0.001 154	1 154						
bromide	-0.09 0.313 141	0.09 0.282 146	0.18 0.017 174	0.21 0.006 174	0.20 0.014 152	1 174					
free chlorine	0.02 0.849 127	-0.21 0.019 128	-0.20 0.012 156	-0.10 0.228 156	-0.22 0.007 151	-0.08 0.353 154	1 156				
total chlorine	0.01 0.891 127	-0.23 0.010 128	-0.22 0.005 156	-0.11 0.158 156	-0.24 0.003 151	-0.09 0.268 154	**0.97*** <0.001 156	1 156			
conductivity	0.07 0.457 126	0.10 0.254 127	-0.29 <0.001 154	-0.33 <0.001 154	-0.23 0.004 154	0.08 0.357 152	-0.09 0.294 151	-0.07 0.374 151	1 154		
pH	-0.04 0.685 126	0.17 0.064 127	-0.05 0.545 154	-0.10 0.223 154	0.04 0.603 154	0.08 0.316 152	-0.24 0.003 151	-0.27 0.001 151	0.23 0.004 154	1 154	
turbidity	0.09 0.342 126	-0.25 0.005 128	0.10 0.233 155	0.03 0.757 155	0.20 0.014 154	0.12 0.154 153	-0.06 0.428 152	-0.05 0.521 152	0.08 0.308 154	0.06 0.425 154	1 155

Table A4 - 6: Days of work reported

No of days of work/week	Employed women No (%)	Women with geocoded Bradford work addresses No (%)
1	16 (0.5)	14 (0.5)
2	113 (3.6)	95 (3.6)
3	413 (13.0)	333 (12.7)
4	368 (11.6)	306 (11.6)
5	2,157 (68.0)	1,787 (68.0)
6	76 (2.4)	67 (2.6)
7	27 (0.9)	25 (1.0)
missing	1 (0.0)	1 (0.0)
total	3,171	2,628

Table A4 - 7: Days assigned as work days

No of work days /week	Assigned work days
1	Mondays only
2	Mondays & Tuesdays
3	Mondays through Wednesdays
4	Mondays through Thursdays
5	Mondays through Fridays
6	Mondays through Saturdays
7	Every day

Table A4 - 8: Summary HAA determinants (including 2009q2)

	temperature	**UV254**	**TOC**	**colour**	**bromide**	**free chlorine**	**total chlorine**	**conductivity**	**pH**	**turbidity**
units	deg C	abs/m	mg/L-C	mg/L	mg/L	mg/L	mg/L	uS/cm		FTU
mean	12.2		1.7						7.8	
sd	3.9		0.4						0.4	
p50	11.0	2.7	1.6	1.0	0.010	0.15	0.20	183	7.7	0.1
iqr	6.2	1.4	0.6	0.9	0.002	0.25	0.25	40	0.4	0.1
min	4.7	0.1	0.6	0.3	0.006	0.03	0.05	119	7.3	0.1
max	20.1	5.7	2.8	3.7	0.156	1.00	1.00	615	9.2	0.6
No. <LOD^	0	1	0	89*	168*	20	0	0	0	95*
Ntotal	156	184	184	162	182	164	164	162	162	163
missing	28	0	0	22	2	20	20	22	22	21

^ and replaced by 2/3LOD

***** these variables had different LOD's over time: colour: <0.5 to <1.5; bromide: <0.009 to <0.027; turbidity: <0.01 to <0.16

Table A4 - 9: Pairwise Pearson correlations (r, p-value, N)

a) between individual HAAs and HAA9 (sum of nine known HAAs) and HAA5 (sum of MCAA, DCAA, TCAA, MBAA and DBAA)
Of note, there are many missing values, such that HAA9 is an actual sum of 9 HAAs in only 33% of cases (ditto for HAA5). In particular, TBAA had 24 values<MRL, 32 values<LOD and no detectable values; and MBAA had 16<MRL, 25<LOD and 15 detectable values (see Table 4.1). However, DCAA and TCAA dominate the sums, and they are both present in 97% of HAA9 and HAA5 sums.

	HAA9	HAA5	MCAA	MBAA	DCAA	DBAA	BCAA	TCAA	TBAA	BDCAA	DBCAA
HAA5	**0.98***	1									
	<0.001										
	172	172									
MCAA	0.01	0.06	1								
	0.885	0.561									
	98	98	98								
MBAA	0.01	0.04	-0.19	1							
	0.952	0.762	0.154								
	56	56	56	56							
DCAA	**0.79***	**0.81***	0.04	-0.20	1						
	<0.001	<0.001	0.675	0.147							
	166	166	92	56	166						
DBAA	-0.09	-0.11	-0.33	0.21	-0.03	1					
	0.415	0.293	0.002	0.116	0.765						
	89	89	84	56	89	89					
BCAA	0.44	0.32	-0.18	-0.24	0.37	0.23	1				
	<0.001	<0.001	0.114	0.079	<0.001	0.047					
	140	140	79	56	140	78	140				
TCAA	**0.81***	**0.82***	0.01	0.11	0.39	-0.17	0.17	1			
	<0.001	<0.001	0.953	0.437	<0.001	0.106	0.047				
	166	166	92	56	166	89	140	166			
TBAA	-0.46	-0.44	**1.00***	-0.19	**-0.52***	-0.49	-0.48	-0.30	1		
	<0.001	0.001	<0.001	0.154	<0.001	<0.001	<0.001	0.025			
	56	56	56	56	56	56	56	56	56		
BDCAA	0.44	0.31	0.49	-0.25	0.17	-0.08	0.31	0.39	-0.31	1	
	<0.001	<0.001	<0.001	0.067	0.040	0.460	<0.001	<0.001	0.021		
	151	151	86	56	151	89	129	151	56	151	
DBCAA	0.32	0.26	**0.59***	0.40	0.36	-0.40	-0.07	0.32	**0.67***	**0.68***	1
	0.002	0.012	<0.001	0.002	<0.001	0.001	0.534	0.002	<0.001	<0.001	
	95	95	66	56	95	67	94	95	56	95	95

Bold with * means rho>0.5; red means p-value <0.005

b) between three transformed HAAs of interest (excluding 2009 q2)

	sqrt(DCAA)	sqrt(TCAA)	ln(BDCAA)
sqrt(DCAA)	1 158		
sqrt(TCAA)	0.34 <0.001 158	1 158	
ln(BDCAA)	0.08 0.318 143	0.17 0.045 143	1 143

Bold with * means rho>0.5; red means p-value <0.005

Table A4 - 10: Parameters for DCAA, TCAA, and BDCAA models – Imputation models

parameters	categories	DCAA model Mean (95% Cred. Int.)	TCAA model Mean (95% Cred. Int.)	BDCAA model Mean (95% Cred. Int.)
TEMPERATURE				
Intercept α_1		10.24 (9.17, 11.41)	10.40 (9.34, 11.50)	
WSZ factor γ1	zone 1	REF	REF	
	zone 2	-0.63 (-1.97, 0.63)	-0.86 (-2.14, 0.38)	
	zone 3	-1.39 (-2.83, 0.00)	-1.56 (-3.01, -0.12)	
	zone 4	-1.49 (-2.95, -0.13)	-1.65 (-2.98, -0.32)	
	zone 5	-1.26 (-2.68, 0.19)	-1.42 (-2.83, -0.01)	
	zone 6	-0.90 (-2.31, 0.45)	-1.13 (-2.50, 0.19)	
	zone 7	-1.60 (-3.13, -0.04)	-1.75 (-3.27, -0.32)	
	zone 8	-0.99 (-2.31, 0.28)	-1.24 (-2.60, -0.02)	
quarter φ1	quarter 1 (jan-mar)	REF	REF	
	quarter 2 (apr-jun)	4.40 (3.54, 5.23)	4.51 (3.61, 5.39)	
	quarter 3 (jul-sep)	8.52 (7.62, 9.44)	8.52 (7.55, 9.48)	
	quarter 4 (oct-dec)	0.45 (-0.52, 1.35)	0.46 (-0.45, 1.33)	
CONDUCTIVITY				
Intercept α2		189.80 (184.20, 195.20)	192.60 (187.30, 197.70)	185.90 (179.90, 191.50)
WSZ factor γ2	zone 1	REF	REF	REF
	zone 2	0.70 (-7.91, 9.74)	-2.15 (-10.49, 6.19)	-1.16 (-10.43, 8.30)
	zone 3	-3.94 (-12.53, 5.18)	-3.41 (-11.79, 4.89)	1.22 (-7.76, 11.43)
	zone 4	0.22 (-8.31, 8.54)	-2.05 (-9.89, 6.50)	2.31 (-7.14, 11.60)
	zone 5	27.87 (18.90, 36.98)	26.01 (17.92, 33.87)	31.61 (22.28, 40.59)
	zone 6	5.06 (-3.24, 13.96)	-4.85 (-13.26, 3.23)	3.66 (-5.33, 12.84)
	zone 7	-32.61 (-41.26, -23.84)	-33.26 (-41.34, -24.96)	-29.24 (-38.12, -19.73)
	zone 8	-6.31 (-14.77, 2.48)	-6.00 (-14.24, 2.60)	-1.09 (-10.40, 8.44)
TOC				
Intercept α3			1.41 (1.26, 1.55)	
WSZ factor γ3	zone 1		REF	
	zone 2		0.05 (-0.12, 0.22)	
	zone 3		0.06 (-0.12, 0.25)	
	zone 4		0.02 (-0.16, 0.20)	
	zone 5		-0.15 (-0.34, 0.05)	
	zone 6		0.04 (-0.12, 0.22)	
	zone 7		-0.11 (-0.30, 0.09)	
	zone 8		0.02 (-0.15, 0.20)	
quarter φ2	quarter 1		REF	
	quarter 2		0.15 (0.04, 0.25)	
	quarter 3		0.63 (0.51, 0.76)	
	quarter 4		0.55 (0.43, 0.68)	
TOTAL CHLORINE				
Intercept α4				0.27 (0.24, 0.31)
WSZ factor γ_4	zone 1			REF
	zone 2			0.13 (0.08, 0.18)
	zone 3			0.08 (0.04, 0.13)
	zone 4			-0.01 (-0.05, 0.03)
	zone 5			0.04 (-0.01, 0.08)
	zone 6			0.10 (0.06, 0.15)
	zone 7			0.01 (-0.03, 0.06)
	zone 8			0.09 (0.05, 0.14)
quarter φ_3	quarter 1			REF
	quarter 2			-0.08 (-0.11, -0.05)
	quarter 3			-0.15 (-0.18, -0.12)
	quarter 4			-0.06 (-0.09, -0.03)

zone 1=Airedale (ADL), zone 2=Bradford Central (BCE), zone 3=Bradford South East (BSE), zone 4=Bradford South West (BSW), zone 5=Graincliffe (GCF), zone 6=Idle/Pudsey (IPY), zone 7=Keighley (KLY), zone 8=Shipley/Bingley (SPY)

Table A4 - 11: Summary statistics of trimester-weighted HAA concentrations (in ug/L) for each woman according to Method 2 (including only women whose given trimester completely overlaps with the modelled period)

In ug/L		Mean	SD	Min	Max	N	% of N= 11,928
Average [DCAA]	trimester 1	8.72	2.94	1.95	16.04	9,024	75.7
	trimester 2	8.75	2.85	1.95	16.04	9,702	81.3
	trimester 3	8.92	2.87	1.78	16.73	9,970	83.6
Average [TCAA]	trimester 1	11.85	3.39	3.10	20.67	9,024	75.7
	trimester 2	12.14	3.49	2.86	20.67	9,702	81.3
	trimester 3	12.34	3.39	3.32	20.94	9,970	83.6
Average [BDCAA]	trimester 1	1.33	0.68	0.35	3.34	9,024	75.7
	trimester 2	1.32	0.67	0.34	3.34	9,702	81.3
	trimester 3	1.33	0.68	0.26	3.55	9,970	83.6

Table A4 - 12: Comparison of weighted HAA concentrations (in ug/L) for each woman, by Method 1 and Method 2. Red means p-value <0.005

in ug/L			Method 2		Method 1		Difference		t	p-value*
		N	Mean	SD	Mean	SD	Mean	SD		
Average [DCAA]	trimester 1	9024	8.72	2.94	8.71	2.96	0.01	0.61	1.1351	0.256
	trimester 2	9702	8.75	2.85	8.74	2.88	0.01	0.62	1.7342	0.083
	trimester 3	9970	8.92	2.87	8.91	2.90	0.01	0.63	2.0664	0.039
Average [TCAA]	trimester 1	9024	11.85	3.39	11.94	3.42	-0.09	0.70	-12.802	<0.001
	trimester 2	9702	12.14	3.49	12.24	3.53	-0.09	0.72	-12.685	<0.001
	trimester 3	9970	12.34	3.39	12.43	3.44	-0.09	0.75	-12.381	<0.001
Average [BDCAA]	trimester 1	9024	1.33	0.68	1.33	0.68	0.00	0.03	-5.1941	<0.001
	trimester 2	9702	1.32	0.67	1.32	0.67	0.00	0.03	-5.4834	<0.001
	trimester 3	9970	1.33	0.68	1.33	0.68	0.00	0.04	-6.4648	<0.001

*by paired t-test

Table A4 - 13: THM-HAA correlations reported in the literature (in decreasing order of correlation) (chronological order)

Reference	Location	Correlated species	Correlation	Correlation coefficient
Nissinen et al. (2002)	Finland	HAA6-TTHM	Pearson	r=0.90 (r^2=0.81)
Wright et al. (2002)	US	HAA5-TTHM		r=0.35 r^2=0.12
Villanueva et al. (2003)	Spain	HAA9-TTHM	Pearson	r=0.815 (r^2=0.66)
Serodes et al. (2003)	Canada	HAA9-TTHM		combined: r^2 = 0.63
King et al. (2004)	Canada	HAA3 (DCAA+TCAA +BCAA) -TTHM	Pearson	r=0.74 in Nova Scotia and r=0.52 in Ontario r^2=0.55
Malliarou et al. (2005)	UK	HAA6-TTHM		r=0.2–0.3 (r^2=0.04-0.09) (by region: r= 0.85, 0.87, 0.10) r^2=0.72, 0.76, 0.01)
Ates et al. (2007)	Turkey		Linear regression	r =0.92 (r^2=0.85)
Bougeard (2009) (PhD thesis)	UK			r^2 = 0.81
Y Zhang et al. (2010)	UK	HAA9-TTHM		r^2=0.88 (by regions: r=poor †)

† data not provided

Figure A4 - 1: Bradford (blue), rest of Yorkshire (grey), BiB women's workplaces (black dots)

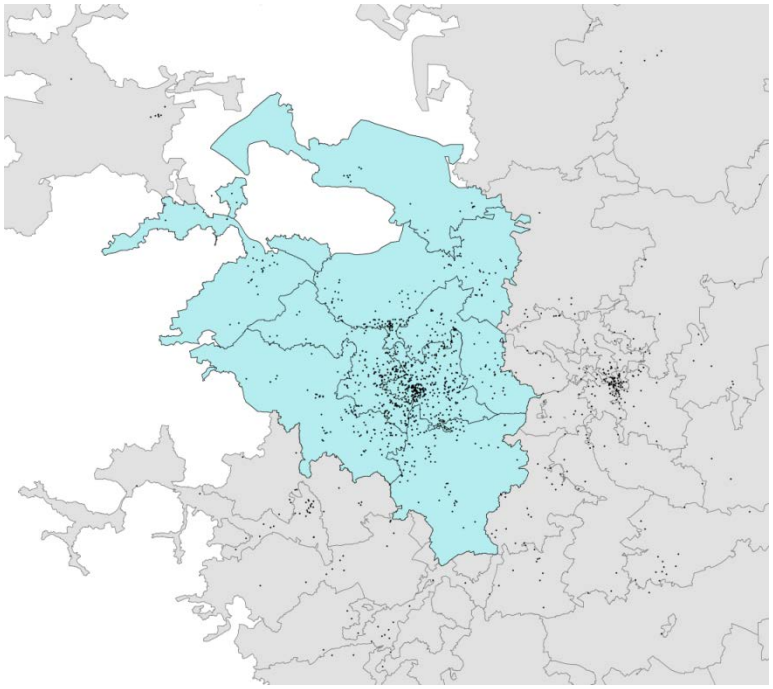
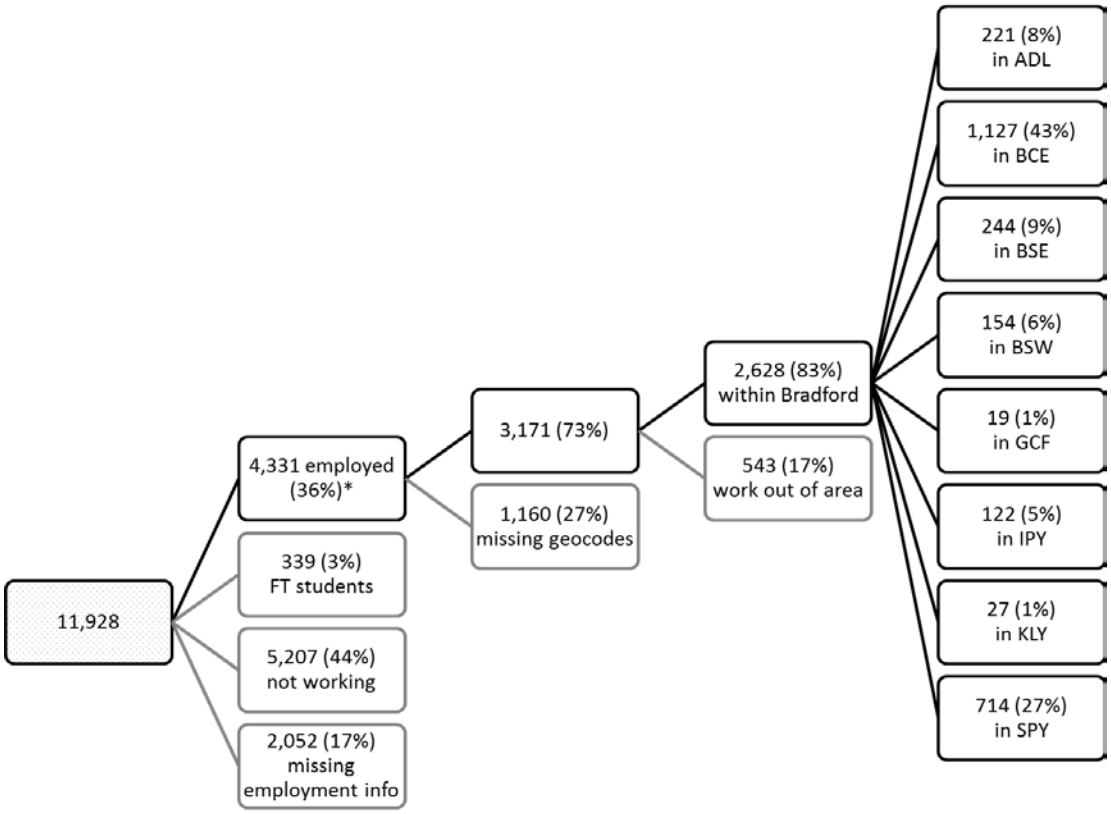


Figure A4 - 2: Flowchart of women eligible for Method 2 time-weighted area-level HAA concentrations



% are calculated from the total in the previous cell
*employed and either currently working or currently on maternity leave
ADL: Airedale
BCE: Bradford Central
BSE: Bradford South East
BSW: Bradford South West
GCF: Graincliffe
IPY: Idle/Pudsey
KLY: Keighley
SPY: Shipley and Bingley

Figure A4 - 3: Correlation matrix between three HAAs of interest

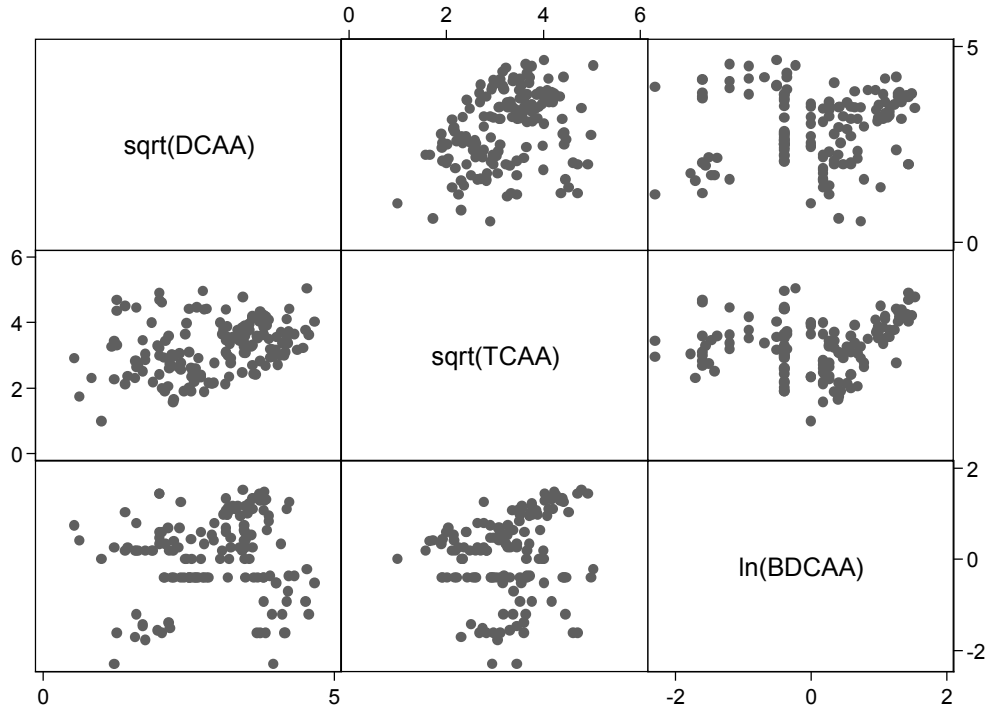
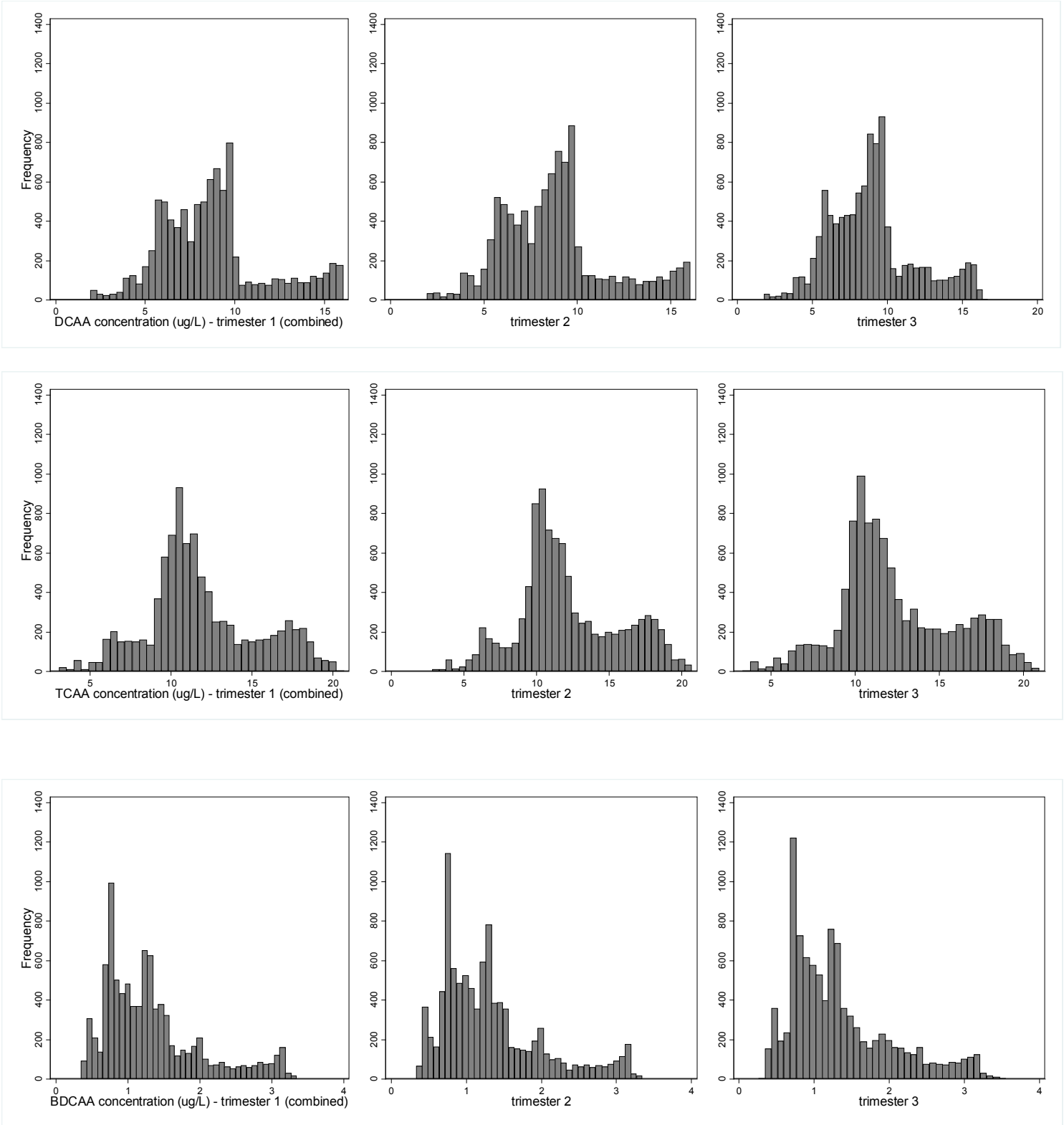


Figure A4 - 4: Frequency distributions of modelled trimester-specific HAA exposure for women in the BiB cohort; DCAA (row 1), TCAA (row 2) and BDCAA (row 3) for trimesters 1 (column 1), 2 (column 2) and 3 (column 3) as derived by Method 2
Ntrim1=9024, Ntrim2=9702, Ntrim3=9970 for all three HAAs



Chapter 5

Table A5 - 1: Summary of combined metric for filterers and non-filterers based on combination of residence/work concentrations if available, residence only if not (in ug/day)

					Percentile Distribution				
n total = 11,928		mean	sd	min	25th %ile	Median	75th %ile	max	n
$ExpDCAA_f$	filterers	12.87	10.4	0.4	6.4	10.1	16.3	158.7	1012
$ExpDCAA_u$	non-filterers	16.64	10.6	0.8	9.7	14.2	20.9	117.8	5211
$ExpTCAA_f$	filterers	13.64	11.4	0.4	6.6	10.8	17.0	131.7	1012
$ExpTCAA_u$	non-filterers	18.98	11.8	0.8	11.2	16.4	23.9	129.7	5211
$ExpBDCAA_f$	filterers	1.20	1.2	0.0	0.4	0.9	1.5	17.5	1012
$ExpBDCAA_f$	non-filterers	1.91	1.4	0.1	1.0	1.5	2.4	14.5	5211

Table A5 - 2: Summary of combined metric (total) based on residence/work concentrations if available, residence only if not (in ug/day)

n total = 11,928		mean	sd	min	25th %ile	Median	75th %ile	max	n
Exposure to DCAA	ExpDCAA	16.03	10.6	0.4	9.1	13.6	20.3	158.7	6223
Exposure to TCAA	ExpTCAA	18.11	11.9	0.4	10.3	15.5	23.0	131.7	6223
Exposure to BDCAA	ExpBDCAA	1.80	1.4	0.0	0.9	1.4	2.3	17.5	6223

Chapter 7

Table A7 - 1: Correlation between pregnancy complication variables (eClipse data) (Spearman's rho (same as Cramer's V[28]), p-value, N) (Nmax=11,928, the red colour signals a p-value <0.05)

rho p-value N	gestational diabetes	pre-existing hypertension	pregnancy-induced hypertension	hypertension during labour only	pre-eclampsia
previous diabetes	0.0156 0.0976 11,312	0.0904 <0.001 11,347	0.0080 0.3927 11,345	-0.0055 0.5644 11,018	0.0169 0.0722 11,326
gestational diabetes	NA	0.0448 <0.001 11,302	0.0093 0.3205 11,300	0.0035 0.7140 10,974	0.0024 0.7964 11,281
pre-existing hypertension		NA	0.2026 <0.001 11,343	0.0289 0.0024 11,018	0.0630 <0.001 11,320
pregnancy-induced hypertension			NA	0.2328 <0.001 11,024	0.5389 <0.001 11,318
hypertension during labour only				NA	0.1214 <0.001 10,992

[28] a measure of association between two nominal variables. In 2x2 tables, the range is -1 to 1

Table A7 - 2: Spearman (pairwise) correlations between 10 covariates in the birth weight models: Spearman's rho, p-value and sample size.

Nmax= 11875	sex of child	ethnicity	parity	maternal age	BMI	smoking	maternal education	caffeine intake	gestational diabetes
gestational age	0.01	-0.07	-0.10	-0.03	0.02	0.04	0.03	0.01	-0.25
	0.198	<0.001	<0.001	<0.001	0.018	<0.001	0.011	0.484	<0.001
	11875	9806	11443	11875	9479	9826	9822	8936	11382
sex of child	NA	0.00	0.00	0.00	-0.02	-0.01	-0.02	0.01	0.00
		0.774	0.631	0.636	0.040	0.176	0.025	0.620	0.808
		9806	11443	11875	9479	9826	9822	8936	11382
ethnicity		NA	0.10	-0.02	-0.07	-0.42	0.08	-0.28	0.09
			<0.001	0.116	<0.001	<0.001	<0.001	<0.001	<0.001
			9428	9806	9442	9790	9787	8910	9424
parity			NA	0.19	0.18	-0.11	-0.15	0.02	0.08
				<0.001	<0.001	<0.001	<0.001	0.038	<0.001
				11443	9105	9443	9441	8614	11003
maternal age				NA	0.11	-0.04	0.03	0.01	0.11
					<0.001	<0.001	0.013	0.296	<0.001
					9479	9826	9822	8936	11382
BMI (quartiles)					NA	0.01	0.00	-0.01	0.12
						0.483	0.980	0.613	<0.001
						9467	9456	8612	9113
smoking						NA	-0.13	0.35	-0.08
							<0.001	<0.001	<0.001
							9806	8933	9443
maternal education							NA	-0.08	-0.01
								<0.001	0.469
								8919	9439
caffeine intake								NA	-0.05
									<0.001
									8579

Red means p-value <0.005

Table A7 - 3: Interaction terms (combined metric x GA) from linear models on continuous birth weight

	Complete case analysis		Analysis with multiple imputation	
Combined metric for:	F	Prob>F	F	Prob>F
DCAA	2.02	0.1324	0.60	0.5533
TCAA	1.13	0.3243	0.28	0.7551
BDCAA	0.11	0.8991	0.12	0.8867

Table A7 - 4: Sensitivity Analysis: Crude and adjusted association between average modelled area-level concentrations of DCAA, TCAA and BDCAA by trimester of pregnancy (based on residence and work water supply zones if available, residence water supply zone only if not) (in ug/L) and continuous birth weight (in grams) by linear regression

Area-level concentrations (ug/L)	**n**	**Crude mean change in BW (g)[#]**	**Adjusted* mean change in BW (g)[#]**
Average [DCAA]			
trimester 1			
0-7.12	1925	0.0	0.0
>7.12 - 9.36	2033	-33.0 (-66.4, 0.4)	-1.7 (-27.3, 23.9)
>9.36	1940	-58.6 (-92.4, -24.9)	-12.2 (-38.1, 13.7)
trimester 2			
0-7.20	1956	0.0	0.0
>7.20 - 9.43	2167	-36.2 (-69.0, -3.4)	-30.3 (-55.4, -5.2)
>9.43	2135	-34.0 (-66.9, -1.1)	-16.0 (-41.3, 9.3)
trimester 3			
0-7.43	2054	0.0	0.0
>7.43 - 9.54	2192	-29.2 (-61.6, 3.2)	-13.5 (-38.2, 11.1)
>9.54	2144	-53.4 (-86.0, -20.9)	-30.6 (-55.5, -5.8)
Average [TCAA]			
trimester 1			
0-10.38	1893	0.0	0.0
>10.38 - 12.49	2063	-40.8 (-74.2, -7.4)	-17.8 (-43.3, 7.7)
>12.49	1942	-41.1 (-75.0, -7.2)	-37.6 (-63.5, -11.8)
trimester 2			
0-10.49	1956	0.0	0.0
>10.49 - 13.05	2211	-16.7 (-49.3, 16.0)	-12.9 (-37.8, 12.0)
>13.05	2091	-35.8 (-68.9, -2.7)	-20.5 (-45.7, 4.8)
trimester 3			
0-10.62	1975	0.0	0.0
>10.62 - 13.17	2210	-20.0 (-52.6, 12.7)	-17.4 (-42.2, 7.4)
>13.17	2205	-15.4 (-48.1, 17.3)	-4.6 (-29.4, 20.3)
Average [BDCAA]			
trimester 1			
0-0.90	2174	0.0	0.0
>0.90 - 1.38	1965	8.5 (-24.2, 41.2)	3.2 (-21.8, 28.2)
>1.38	1759	-20.4 (-54.1, 13.3)	-12.8 (-38.5, 13.0)
trimester 2			
0-0.91	2334	0.0	0.0
>0.91 - 1.38	1936	-29.5 (-61.9, 2.8)	-35.8 (-60.5, -11.1)
>1.38	1988	-19.1 (-51.2, 13.1)	-20.0 (-44.6, 4.5)
trimester 3			
0-0.91	2293	0.0	0.0
>0.91 - 1.39	1980	24.4 (-8.0, 56.8)	-0.7 (-25.3, 23.9)
>1.39	2117	-14.6 (-46.5, 17.2)	-11.7 (-35.8, 12.4)

[#]mean change in birth weight (95% confidence interval)

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table A7 - 5: Sensitivity Analysis: Complete case analysis: Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (derived from area-level concentrations based on residence and work water supply zones if available, residence water supply zone only if not) (in ug/day) and continuous birth weight (in grams) by linear regression (N=5,040)

Combined exposure (ug/day) (N=5,040)	**n**	**Crude mean change in BW (g)[#]**	**Adjusted* mean change in BW (g)[#]**
DCAA exposure			
0 - 10.41	1563	0.0	0.0
> 10.41 - 17.29	1680	22.8 (-14.4, 60.1)	6.0 (-22.5, 34.4)
> 17.29	1797	23.0 (-13.7, 59.6)	11.0 (-17.9, 39.8)
TCAA exposure			
0 - 12.02	1589	0.0	0.0
> 12.02 - 19.79	1719	6.9 (-30.0, 43.8)	0.6 (-27.5, 28.8)
> 19.79	1732	12.9 (-23.9, 49.7)	0.0 (-28.5, 28.4)
BDCAA exposure			
0 - 1.04	1734	0.0	0.0
> 1.04 - 1.93	1656	2.0 (-34.4, 38.4)	-6.4 (-34.2, 21.3)
> 1.93	1650	31.8 (-4.6, 68.3)	9.9 (-18.0, 37.8)

[#]mean change in birth weight (95% confidence interval)

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table A7 - 6: Interaction terms (combined metric x GA) from logistic models on term LBW (i.e. born ≥ 37 weeks gestation)

	Complete case analysis		Analysis after multiple imputation	
Combined metric for:	chi^2	Prob>chi^2	F	Prob>F
DCAA	1.55	0.4608	0.48	0.6224
TCAA	1.35	0.5092	0.09	0.9164
BDCAA	0.59	0.7460	0.08	0.9258

Table A7 - 7: Sensitivity Analysis: Crude and adjusted association between average modelled area-level concentrations of DCAA, TCAA and BDCAA by trimester of pregnancy (based on residence and work water supply zones if available, residence water supply zone only if not) (in ug/L) and risk of term LBW by logistic regression
(OR: Odds Ratio)

Area-level concentrations (ug/L)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
Average [DCAA]				
trimester 1				
0-7.12	65	1759	1.00	1.00
>7.12 - 9.36	77	1857	1.12 (0.80, 1.57)	0.92 (0.64, 1.31)
>9.36	91	1753	1.40 (1.01, 1.94)	1.15 (0.81, 1.61)
trimester 2				
0-7.20	66	1781	1.00	1.00
>7.20 - 9.43	98	1963	1.35 (0.98, 1.85)	1.30 (0.93, 1.81)
>9.43	88	1948	1.22 (0.88, 1.69)	1.15 (0.82, 1.63)
trimester 3				
0-7.43	69	1875	1.00	1.00
>7.43 - 9.54	88	2003	1.19 (0.87, 1.65)	1.12 (0.80, 1.57)
>9.54	100	1930	1.41 (1.03, 1.93)	1.37 (0.98, 1.90)
Average [TCAA]				
trimester 1				
0-10.38	67	1724	1.00	1.00
>10.38 - 12.49	89	1874	1.22 (0.88, 1.69)	1.07 (0.76, 1.51)
>12.49	77	1771	1.12 (0.80, 1.56)	1.11 (0.78, 1.58)
trimester 2				
0-10.49	73	1777	1.00	1.00
>10.49 - 13.05	99	2002	1.20 (0.88, 1.64)	1.17 (0.84, 1.61)
>13.05	80	1913	1.02 (0.74, 1.41)	0.95 (0.68, 1.33)
trimester 3				
0-10.62	77	1796	1.00	1.00
>10.62 - 13.17	85	2002	0.99 (0.72, 1.36)	1.00 (0.72, 1.39)
>13.17	95	2010	1.10 (0.81, 1.50)	1.04 (0.76, 1.44)
Average [BDCAA]				
trimester 1				
0-0.90	87	1980	1.00	1.00
>0.90 - 1.38	73	1797	0.92 (0.67, 1.27)	1.01 (0.72, 1.41)
>1.38	73	1592	1.04 (0.76, 1.43)	1.12 (0.80, 1.56)
trimester 2				
0-0.91	95	2114	1.00	1.00
>0.91 - 1.38	79	1758	1.00 (0.74, 1.36)	1.06 (0.77, 1.47)
>1.38	78	1820	0.95 (0.70, 1.30)	0.98 (0.71, 1.35)
trimester 3				
0-0.91	91	2075	1.00	1.00
>0.91 - 1.39	70	1825	0.87 (0.64, 1.20)	0.97 (0.70, 1.36)
>1.39	96	1908	1.15 (0.86, 1.54)	1.16 (0.85, 1.58)

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes
(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table A7 - 8: Sensitivity Analysis: Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (derived from area-level concentrations based on residence and work water supply zones if available, residence water supply zone only if not) (in ug/day) and risk of term LBW by logistic regression; 4.1% prevalence of term LBW (Ncases=195, Nnon-cases=4,587, N=4,782)
(OR: Odds Ratio)

Combined exposure (ug/day) (N=4,782)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
DCAA exposure				
0 - 10.41	71	1,404	1.00	1.00
> 10.41 - 17.29	61	1,540	0.78 (0.55, 1.11)	0.80 (0.55, 1.15)
> 17.29	63	1,643	0.76 (0.54, 1.07)	0.68 (0.46, 0.99)
TCAA exposure				
0 - 12.02	74	1,430	1.00	1.00
> 12.02 - 19.79	61	1,567	0.75 (0.53, 1.06)	0.72 (0.50, 1.04)
> 19.79	60	1,590	0.73 (0.51, 1.03)	0.65 (0.45, 0.94)
BDCAA exposure				
0 - 1.04	81	1,560	1.00	1.00
> 1.04 - 1.93	63	1,500	0.81 (0.58, 1.13)	0.82 (0.57, 1.16)
> 1.93	51	1,527	0.64 (0.45, 0.92)	0.62 (0.42, 0.90)

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes
(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table A7 - 9: Complete case analysis: Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (in ug/day) and continuous birth weight (in grams) by linear regression including all covariate coefficients (N=5,040)
DF: degrees of freedom

Combined exposure ug/day (N=5,040)	n	Crude mean change in BW (g)[#]	Adjusted* mean change in BW (g)	F(DF, 5017)	Prob>F
DCAA exposure					
0 - 10.41	1563	0.0	0.0	0.19	0.8257
> 10.41 - 17.29	1678	17.9 (-19.3, 55.2)	-0.2 (-28.6, 28.2)		
> 17.29	1799	20.0 (-16.6, 56.7)	7.7 (-21.2, 36.5)		
constant		3200.2 (3173.4, 3227.0)	-3706.5 (-3973.8, -3439.2)		
Gestational Age (in weeks)	5040	NA	176.8 (170.1, 183.5)		
Sex of child					
male	2578	NA	0.0	112.14	<0.0001
female	2462	NA	-122.8 (-145.6, -100.1)		
Ethnicity					
White British	1651	NA	0.0	116.80	<0.0001
Pakistani	2719	NA	-252.3 (-284.7, -220.0)		
Other	670	NA	-160.6 (-200.6, -120.6)		
Parity					
0 previous births	1923	NA	0.0	49.84	<0.0001
1 previous birth	1353	NA	117.5 (87.7, 147.3)		
2 or more previous births	1764	NA	158.8 (125.4, 192.1)		
Maternal age					
< 25 years old	1724	NA	49.4 (18.2, 80.6)	3.44	0.0161
25-29 years old	1604	NA	0.0		
30-34 years old	1120	NA	28.7 (-3.4, 60.8)		
≥ 35 years old	592	NA	14.9 (-25.6, 55.3)		
BMI					
quartile 1	1333	NA	0.0	67.68	<0.0001
quartile 2	1245	NA	71.2 (39.0, 103.3)		
quartile 3	1265	NA	154.3 (121.9, 186.6)		
quartile 4	1197	NA	230.5 (196.8, 264.1)		
Smoking					
never a smoker	3696	NA	0.0	36.37	<0.0001
ever a smoker	669	NA	22.4 (-15.5, 60.3)		
current smoker	675	NA	-158.0 (-198.9, -117.0)		
Maternal education					
No education	1326	NA	-17.3 (-47.4, 12.7)	3.55	0.0067
School	1638	NA	0.0		
Further education	614	NA	18.4 (-20.0, 56.9)		
Higher education	1145	NA	47.6 (14.6, 80.7)		
Other, Don't know, Unknown foreign	317	NA	-1.3 (-51.4, 48.7)		
Caffeinated drinks					
0 mg/day	4279	NA	0.0	9.15	0.0025
0-200 mg/day	761	NA	-55.5 (-91.4, -19.5)		
Gestational diabetes					
no	4554	NA	0.0	11.03	0.0009
yes	486	NA	68.7 (28.1, 109.2)		
TCAA exposure					
0 - 12.02	1590	0.0	0.0	0.05	0.9556

Combined exposure ug/day (N=5,040)	n	Crude mean change in BW (g)[#]	Adjusted* mean change in BW (g)	F(DF, 5017)	Prob>F
> 12.02 - 19.79	1719	10.0 (-26.8, 46.9)	3.3 (-24.8, 31.5)		
> 19.79	1731	12.9 (-23.9, 49.7)	-0.7 (-29.1, 27.8)		
constant		3205.5 (3178.9, 3232.1)	-3703.3 (-3970.6, -3436.0)		
Gestational Age (in weeks)	5040	NA	176.7 (170.0, 183.4)		
Sex of child					
male	2578	NA	0.0	112.03	<0.0001
female	2462	NA	-122.8 (-145.5, -100.0)		
Ethnicity					
White British	1651	NA	0.0	117.52	<0.0001
Pakistani	2719	NA	-252.9 (-285.3, -220.6)		
Other	670	NA	-160.9 (-200.9, -120.9)		
Parity					
0 previous births	1923	NA	0.0	49.62	<0.0001
1 previous birth	1353	NA	117.2 (87.3, 147.0)		
2 or more previous births	1764	NA	158.6 (125.2, 192.0)		
Maternal age					
< 25 years old	1724	NA	49.1 (17.9, 80.3)	3.40	0.0170
25-29 years old	1604	NA	0.0		
30-34 years old	1120	NA	28.6 (-3.5, 60.7)		
≥ 35 years old	592	NA	14.8 (-25.6, 55.2)		
BMI					
quartile 1	1333	NA	0.0	68.00	<0.0001
quartile 2	1245	NA	71.5 (39.4, 103.6)		
quartile 3	1265	NA	154.6 (122.3, 187.0)		
quartile 4	1197	NA	231.1 (197.5, 264.8)		
Smoking					
never a smoker	3696	NA	0.0	36.25	<0.0001
ever a smoker	669	NA	22.4 (-15.5, 60.4)		
current smoker	675	NA	-157.7 (-198.7, -116.7)		
Maternal education					
No education	1326	NA	-17.2 (-47.3, 12.8)	3.54	0.0068
School	1638	NA	0.0		
Further education	614	NA	18.4 (-20.1, 56.8)		
Higher education	1145	NA	47.7 (14.6, 80.7)		
Other, Don't know, Unknown foreign	317	NA	-1.6 (-51.7, 48.4)		
Caffeinated drinks					
0 mg/day	4279	NA	0.0	8.61	0.0034
0-200 mg/day	761	NA	-52.9 (-88.2, -17.6)		
Gestational diabetes					
no	4554	NA	0.0	11.07	0.0009
yes	486	NA	68.8 (28.3, 109.4)		
BDCAA exposure					
0 - 1.04	1737	0.0	0.0	0.67	0.5122
> 1.04 - 1.93	1652	4.0 (-32.4, 40.4)	-4.8 (-32.5, 23.0)		
> 1.93	1651	33.8 (-2.6, 70.2)	11.4 (-16.5, 39.2)		
constant		3201.0 (3175.6, 3226.4)	-3703.1 (-3970.2, -3435.9)		
Gestational Age (in weeks)	5040	NA	176.7 (170.0, 183.4)		
Sex of child					
male	2578	NA	0.0	112.39	<0.0001
female	2462	NA	-123.0 (-145.7, -100.2)		

Combined exposure ug/day (N=5,040)	n	Crude mean change in BW (g)[♯]	Adjusted* mean change in BW (g)	F(DF, 5017)	Prob>F
Ethnicity					
White British	1651	NA	0.0	117.10	<0.0001
Pakistani	2719	NA	-252.4 (-284.7, -220.0)		
Other	670	NA	-160.4 (-200.4, -120.5)		
Parity					
0 previous births	1923	NA	0.0	49.80	<0.0001
1 previous birth	1353	NA	117.4 (87.6, 147.2)		
2 or more previous births	1764	NA	158.7 (125.4, 192.0)		
Maternal age					
< 25 years old	1724	NA	49.4 (18.2, 80.6)	3.44	0.0161
25-29 years old	1604	NA	0.0		
30-34 years old	1120	NA	28.9 (-3.2, 61.0)		
≥ 35 years old	592	NA	15.3 (-25.2, 55.7)		
BMI					
quartile 1	1333	NA	0.0	67.86	<0.0001
quartile 2	1245	NA	71.0 (38.9, 103.1)		
quartile 3	1265	NA	154.2 (121.8, 186.6)		
quartile 4	1197	NA	230.7 (197.1, 264.3)		
Smoking					
never a smoker	3696	NA	0.0	36.16	<0.0001
ever a smoker	669	NA	22.7 (-15.3, 60.6)		
current smoker	675	NA	-157.3 (-198.3, -116.4)		
Maternal education					
No education	1326	NA	-17.2 (-47.3, 12.8)	3.57	0.0065
School	1638	NA	0.0		
Further education	614	NA	18.6 (-19.9, 57.0)		
Higher education	1145	NA	47.8 (14.7, 80.9)		
Other, Don't know, Unknown foreign	317	NA	-1.8 (-51.8, 48.2)		
Caffeinated drinks					
0 mg/day	4279	NA	0.0	9.15	0.0025
0-200 mg/day	761	NA	-54.3 (-89.4, -19.1)		
Gestational diabetes					
no	4554	NA	0.0	11.18	0.0008
yes	486	NA	69.2 (28.6, 109.7)		

[♯]mean change in birth weight (95% confidence intervals)

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table A7 - 10: Comparing the exposure coefficients in the original combined DCAA, TCAA and BDCAA (ug/day) models on continuous birth weight to the same coefficients stratified to White British women only, and to Pakistani women only (The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

	White British			Pakistani			Total (un-stratified, see Table 7.2)		
(n=5040)	*(total n)* n	Crude mean change in BW (g)	Adjusted* mean change in BW (g)	*(total n)* n	Crude mean change in BW (g)	Adjusted* mean change in BW (g)	*(total n)* n	Crude mean change in BW (g)	Adjusted* mean change in BW (g)
DCAA exposure	*(1651)*			*(2719)*			*(5040)*		
0 - 10.41	461	0.0	0.0	884	0.0	0.0	1563	0.0	0.0
> 10.41 - 17.29	456	-19.8 (-92.9, 53.2)	-30.0 (-83.7, 23.6)	1001	41.3 (-5.8, 88.3)	17.8 (-19.6, 55.2)	1678	17.9 (-19.3, 55.2)	-0.2 (-28.6, 28.2)
> 17.29	734	-79.2 (-145.0, -13.5)	-23.9 (-74.6, 26.8)	834	28.8 (-20.4, 78.0)	24.6 (-15.0, 64.2)	1799	20.0 (-16.6, 56.7)	7.7 (-21.2, 36.5)
TCAA exposure	*(1651)*			*(2719)*			*(5040)*		
0 - 12.02	463	0.0	0.0	900	0.0	0.0	1590	0.0	0.0
> 12.02 - 19.79	511	-55.4 (-126.4, 15.6)	-45.2 (-97.5, 7.1)	986	40.4 (-6.5, 87.4)	39.7 (2.2, 77.1)	1719	10.0 (-26.8, 46.9)	3.3 (-24.8, 31.5)
> 19.79	677	-72.4 (-139.1, -5.7)	-15.9 (-65.8, 34.0)	833	21.3 (-27.7, 70.2)	13.1 (-26.2, 52.4)	1731	12.9 (-23.9, 49.7)	-0.7 (-29.1, 27.8)
BDCAA exposure	*(1651)*			*(2719)*			*(5040)*		
0 - 1.04	542	0.0	0.0	976	0.0	0.0	1737	0.0	0.0
> 1.04 - 1.93	523	-18.8 (-86.7, 49.2)	0.7 (-49.1, 50.4)	894	-1.9 (-49.1, 45.2)	-6.0 (-43.6, 31.5)	1652	4.0 (-32.4, 40.4)	-4.8 (-32.5, 23.0)
> 1.93	586	-6.0 (-72.0, 60.0)	9.8 (-38.7, 58.4)	849	39.3 (-8.5, 87.1)	23.2 (-15.0, 61.3)	1651	33.8 (-2.6, 70.2)	11.4 (-16.5, 39.2)

*adjusted for 9 variables: gestational age at birth (in completed weeks), sex of child, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table A7 - 11: Complete case analysis: Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (in ug/day) and risk of term LBW by logistic regression including all covariate coefficients; 4.1% prevalence of term LBW (Ncases=195, Nnon-cases=4,587, N=4,782) OR: Odds Ratio; DF: degrees of freedom

Combined exposure ug/day (N=4,782)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)	chi^2 (DF)	Prob > chi^2
DCAA exposure						
0 - 10.41	71	1,404	1.00	1.00	4.18	0.1235
> 10.41 - 17.29	61	1,539	0.78 (0.55, 1.11)	0.80 (0.55, 1.15)		
> 17.29	63	1,644	0.76 (0.54, 1.07)	0.68 (0.46, 0.99)		
Gestational Age (in weeks)			NA	0.38 (0.32, 0.44)		
Sex of child						
male	87	2,345	NA	1.00	4.98	0.0257
female	108	2,242	NA	1.41 (1.04, 1.92)		
Ethnicity						
White British	32	1,521	NA	1.00	39.77	<0.0001
Pakistani	146	2,451	NA	4.84 (2.89, 8.10)		
Other	17	615	NA	1.86 (0.97, 3.58)		
Parity						
0 previous births	77	1,727	NA	1.00	10.01	0.0067
1 previous birth	44	1,239	NA	0.60 (0.39, 0.90)		
2 or more previous births	74	1,621	NA	0.53 (0.34, 0.81)		
Maternal age						
< 25 years old	57	1,577	NA	0.65 (0.43, 0.97)	7.00	0.0720
25-29 years old	71	1,449	NA	1.00		
30-34 years old	38	1,034	NA	0.84 (0.54, 1.29)		
≥ 35 years old	29	527	NA	1.28 (0.78, 2.11)		
BMI						
quartile 1	66	1,182	NA	1.00	8.04	0.0452
quartile 2	48	1,138	NA	0.77 (0.51, 1.15)		
quartile 3	37	1,179	NA	0.53 (0.34, 0.83)		
quartile 4	44	1,088	NA	0.69 (0.45, 1.07)		
Smoking						
never a smoker	151	3,375	NA	1.00	5.00	0.0823
ever a smoker	19	616	NA	1.60 (0.91, 2.82)		
current smoker	25	596	NA	1.72 (0.99, 2.98)		
Maternal education						
No education	55	1,199	NA	0.81 (0.56, 1.18)	4.76	0.3124
School	77	1,491	NA	1.00		
Further education	19	561	NA	0.74 (0.43, 1.26)		
Higher education	37	1,051	NA	0.66 (0.43, 1.02)		
Other, Don't know, Unknown foreign	7	285	NA	0.58 (0.25, 1.32)		
Caffeinated drinks						
0 mg/day	161	3,915	NA	1.00	11.08	0.0009
0-200 mg/day	34	672	NA	2.21 (1.39, 3.54)		
Gestational diabetes						
no	164	4,159	NA	1.00	1.61	0.2048
yes	31	428	NA	0.75 (0.49, 1.17)		
TCAA exposure						
0 - 12.02	74	1,431	1.00	1.00	5.50	0.0638
> 12.02 - 19.79	60	1,570	0.74 (0.52, 1.05)	0.71 (0.49, 1.02)		

Combined exposure ug/day (N=4,782)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)	chi^2 (DF)	Prob > chi^2
> 19.79	61	1,586	0.74 (0.53, 1.05)	0.67 (0.46, 0.97)		
Gestational Age (in weeks)			NA	0.38 (0.32, 0.44)		
Sex of child						
male	87	2,345	NA	1.00	5.11	0.0239
female	108	2,242	NA	1.42 (1.05, 1.93)		
Ethnicity						
White British	32	1,521	NA	1.00	40.77	<0.0001
Pakistani	146	2,451	NA	4.90 (2.93, 8.19)		
Other	17	615	NA	1.85 (0.97, 3.56)		
Parity						
0 previous births	77	1,727	NA	1.00	9.91	0.0070
1 previous birth	44	1,239	NA	0.60 (0.40, 0.91)		
2 or more previous births	74	1,621	NA	0.52 (0.34, 0.81)		
Maternal age						
< 25 years old	57	1,577	NA	0.65 (0.43, 0.98)	7.06	0.0700
25-29 years old	71	1,449	NA	1.00		
30-34 years old	38	1,034	NA	0.85 (0.55, 1.30)		
≥ 35 years old	29	527	NA	1.30 (0.79, 2.15)		
BMI						
quartile 1	66	1,182	NA	1.00	8.04	0.0452
quartile 2	48	1,138	NA	0.78 (0.52, 1.16)		
quartile 3	37	1,179	NA	0.53 (0.34, 0.83)		
quartile 4	44	1,088	NA	0.70 (0.45, 1.08)		
Smoking						
never a smoker	151	3,375	NA	1.00	5.23	0.0732
ever a smoker	19	616	NA	1.61 (0.92, 2.83)		
current smoker	25	596	NA	1.75 (1.01, 3.02)		
Maternal education						
No education	55	1,199	NA	0.80 (0.55, 1.17)	4.66	0.3242
School	77	1,491	NA	1.00		
Further education	19	561	NA	0.74 (0.44, 1.27)		
Higher education	37	1,051	NA	0.67 (0.43, 1.04)		
Other, Don't know, Unknown foreign	7	285	NA	0.57 (0.25, 1.31)		
Caffeinated drinks						
0 mg/day	161	3,915	NA	1.00	10.38	0.0013
0-200 mg/day	34	672	NA	2.13 (1.34, 3.37)		
Gestational diabetes						
no	164	4,159	NA	1.00	1.47	0.2247
yes	31	428	NA	0.76 (0.49, 1.18)		
BDCAA exposure						
0 - 1.04	81	1,562	1.00	1.00	6.24	0.0441
> 1.04 - 1.93	63	1,497	0.81 (0.58, 1.14)	0.82 (0.58, 1.17)		
> 1.93	51	1,528	0.64 (0.45, 0.92)	0.62 (0.42, 0.90)		
Gestational Age (in weeks)			NA	0.38 (0.33, 0.44)		
Sex of child						
male	87	2,345	NA	1.00	4.98	0.0257
female	108	2,242	NA	1.41 (1.04, 1.92)		
Ethnicity						
White British	32	1,521	NA	1.00	39.76	<0.0001
Pakistani	146	2,451	NA	4.84 (2.89, 8.11)		
Other	17	615	NA	1.89 (0.98, 3.63)		
Parity						

Combined exposure ug/day (N=4,782)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)	chi^2 (DF)	Prob > chi^2
0 previous births	77	1,727	NA	1.00	9.95	0.0069
1 previous birth	44	1,239	NA	0.60 (0.40, 0.90)		
2 or more previous births	74	1,621	NA	0.53 (0.34, 0.81)		
Maternal age						
< 25 years old	57	1,577	NA	0.64 (0.43, 0.97)	6.84	0.0772
25-29 years old	71	1,449	NA	1.00		
30-34 years old	38	1,034	NA	0.83 (0.54, 1.28)		
≥ 35 years old	29	527	NA	1.26 (0.77, 2.08)		
BMI						
quartile 1	66	1,182	NA	1.00	8.05	0.0450
quartile 2	48	1,138	NA	0.77 (0.52, 1.16)		
quartile 3	37	1,179	NA	0.53 (0.34, 0.83)		
quartile 4	44	1,088	NA	0.69 (0.45, 1.08)		
Smoking						
never a smoker	151	3,375	NA	1.00	4.80	0.0909
ever a smoker	19	616	NA	1.59 (0.90, 2.79)		
current smoker	25	596	NA	1.70 (0.98, 2.95)		
Maternal education						
No education	55	1,199	NA	0.82 (0.56, 1.19)	4.80	0.3081
School	77	1,491	NA	1.00		
Further education	19	561	NA	0.73 (0.43, 1.25)		
Higher education	37	1,051	NA	0.66 (0.43, 1.03)		
Other, Don't know, Unknown foreign	7	285	NA	0.57 (0.25, 1.30)		
Caffeinated drinks						
0 mg/day	161	3,915	NA	1.00	10.09	0.0015
0-200 mg/day	34	672	NA	2.10 (1.33, 3.33)		
Gestational diabetes						
no	164	4,159	NA	1.00	1.62	0.2030
yes	31	428	NA	0.75 (0.48, 1.17)		

*adjusted for 10 variables: gestational age at birth (in completed weeks), sex of child, ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table A7 - 12: Sensitivity Analysis: Crude and adjusted association between average modelled area-level concentrations of DCAA, TCAA and BDCAA by trimester of pregnancy (based on residence and work water supply zones if available, residence water supply zone only if not) (in ug/L) and risk of being SGA by logistic regression (OR: Odds Ratio)

Area-level concentrations (ug/L)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
Average [DCAA]				
trimester 1				
0-7.12	228	1,696	1.00	1.00
>7.12 - 9.36	247	1,786	1.03 (0.85, 1.25)	0.94 (0.77, 1.14)
>9.36	267	1,671	1.19 (0.98, 1.44)	1.05 (0.87, 1.28)
trimester 2				
0-7.20	221	1,734	1.00	1.00
>7.20 - 9.43	292	1,874	1.22 (1.01, 1.47)	1.18 (0.98, 1.43)
>9.43	284	1,849	1.21 (1.00, 1.45)	1.10 (0.90, 1.33)
trimester 3				
0-7.43	233	1,821	1	1.00
>7.43 - 9.54	286	1,905	1.17 (0.98, 1.41)	1.10 (0.91, 1.33)
>9.54	288	1,853	1.21 (1.01, 1.46)	1.11 (0.92, 1.34)
Average [TCAA]				
trimester 1				
0-10.38	211	1,681	1	1.00
>10.38 - 12.49	261	1,800	1.16 (0.95, 1.40)	1.10 (0.90, 1.34)
>12.49	270	1,672	1.29 (1.06, 1.56)	1.26 (1.03, 1.53)
trimester 2				
0-10.49	233	1,721	1	1.00
>10.49 - 13.05	288	1,922	1.11 (0.92, 1.33)	1.11 (0.92, 1.34)
>13.05	276	1,814	1.12 (0.93, 1.35)	1.08 (0.90, 1.31)
trimester 3				
0-10.62	240	1,734	1	1.00
>10.62 - 13.17	296	1,913	1.12 (0.93, 1.34)	1.11 (0.92, 1.33)
>13.17	271	1,932	1.01 (0.84, 1.22)	1.00 (0.83, 1.21)
Average [BDCAA]				
trimester 1				
0-0.90	257	1,916	1	1.00
>0.90 - 1.38	246	1,717	1.07 (0.89, 1.29)	1.04 (0.86, 1.26)
>1.38	239	1,520	1.17 (0.97, 1.42)	1.14 (0.94, 1.38)
trimester 2				
0-0.91	271	2,061	1	1.00
>0.91 - 1.38	258	1,676	1.17 (0.98, 1.40)	1.22 (1.01, 1.47)
>1.38	268	1,720	1.18 (0.99, 1.42)	1.15 (0.95, 1.38)
trimester 3				
0-0.91	281	2,010	1	1.00
>0.91 - 1.39	242	1,737	1.00 (0.83, 1.20)	0.98 (0.81, 1.18)
>1.39	284	1,832	1.11 (0.93, 1.32)	1.08 (0.90, 1.29)

*adjusted for 8 variables: ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table A7 - 13: Sensitivity Analysis: Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (in ug/day) (derived from area-level concentrations based on residence and work water supply zones if available, residence water supply zone only if not) and risk of being SGA by logistic regression; 12.9% prevalence of term LBW (Ncases=649, Nnon-cases=4,388, N=5,037) (OR: Odds Ratio)

Combined exposure (ug/day) (N=5,037)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)
DCAA exposure				
0 - 10.41	213	1,349	1.00	1.00
> 10.41 - 17.29	225	1,455	0.98 (0.80, 1.20)	0.97 (0.79, 1.19)
> 17.29	211	1,584	0.84 (0.69, 1.03)	0.86 (0.70, 1.07)
TCAA exposure				
0 - 12.02	218	1,370	1.00	1.00
> 12.02 - 19.79	215	1,504	0.90 (0.73, 1.10)	0.90 (0.73, 1.11)
> 19.79	216	1,514	0.90 (0.73, 1.10)	0.94 (0.76, 1.15)
BDCAA exposure				
0 - 1.04	225	1,509	1.00	1.00
> 1.04 - 1.93	224	1,431	1.05 (0.86, 1.28)	1.08 (0.88, 1.32)
> 1.93	200	1,448	0.93 (0.76, 1.14)	0.96 (0.78, 1.18)

*adjusted for 8 variables: ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

Table A7 - 14: Complete case analysis: Crude and adjusted association between combined metric for DCAA, TCAA and BDCAA (in ug/day) and risk of being SGA by logistic regression including all covariate coefficients; 12.9% prevalence of term LBW (Ncases=649, Nnon-cases=4,388, N=5,037) OR: Odds Ratio; DF: degrees of freedom

Combined exposure ug/day (N=5,037)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)	chi^2 (DF)	Prob > chi^2
DCAA exposure						
0 - 10.41	213	1,349	1.00	1.00	1.84	0.3988
> 10.41 - 17.29	224	1,454	0.98 (0.80, 1.19)	0.96 (0.78, 1.18)		
> 17.29	212	1,585	0.85 (0.69, 1.04)	0.87 (0.70, 1.07)		
Ethnicity						
White British	130	1,519	NA	1.00	88.41	<0.0001
Pakistani	438	2,281	NA	3.76 (2.85, 4.96)		
Other	81	588	NA	2.31 (1.67, 3.21)		
Parity						
0 previous births	305	1,616	NA	1.00	33.96	<0.0001
1 previous birth	143	1,210	NA	0.57 (0.46, 0.72)		
2 or more previous births	201	1,562	NA	0.53 (0.41, 0.68)		
Maternal age						
< 25 years old	245	1,479	NA	0.82 (0.65, 1.03)	3.83	0.2799
25-29 years old	220	1,383	NA	1.00		
30-34 years old	118	1,000	NA	0.90 (0.70, 1.15)		
≥ 35 years old	66	526	NA	1.05 (0.77, 1.44)		
BMI						
quartile 1	232	1,100	NA	1.00	25.70	<0.0001
quartile 2	175	1,070	NA	0.83 (0.67, 1.04)		
quartile 3	131	1,132	NA	0.59 (0.47, 0.75)		
quartile 4	111	1,086	NA	0.60 (0.46, 0.77)		
Smoking						
never a smoker	492	3,203	NA	1.00	24.01	<0.0001
ever a smoker	54	613	NA	0.99 (0.71, 1.37)		
current smoker	103	572	NA	2.03 (1.50, 2.75)		
Maternal education						
No education	189	1,136	NA	1.05 (0.85, 1.30)	6.05	0.1957
School	224	1,414	NA	1.00		
Further education	65	549	NA	0.81 (0.60, 1.09)		
Higher education	142	1,001	NA	0.82 (0.64, 1.04)		
Other, Don't know, Unknown foreign	29	288	NA	0.80 (0.53, 1.22)		
Caffeinated drinks						
0 mg/day	549	3,727	NA	1.00	7.67	0.0056
0-200 mg/day	100	661	NA	1.47 (1.12, 1.93)		
Gestational diabetes						
no	602	3,949	NA	1.00	4.75	0.0293
yes	47	439	NA	0.70 (0.50, 0.96)		
TCAA exposure						
0 - 12.02	218	1,371	1.00	1.00	1.16	0.5602
> 12.02 - 19.79	213	1,506	0.89 (0.73, 1.09)	0.89 (0.73, 1.10)		
> 19.79	218	1,511	0.91 (0.74, 1.11)	0.95 (0.77, 1.17)		
Ethnicity						
White British	130	1,519	NA	1.00	89.34	<0.0001
Pakistani	438	2,281	NA	3.78 (2.86, 4.99)		
Other	81	588	NA	2.31 (1.67, 3.21)		

Combined exposure ug/day (N=5,037)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)	chi^2 (DF)	Prob > chi^2
Parity						
0 previous births	305	1,616	NA	1.00	33.74	<0.0001
1 previous birth	143	1,210	NA	0.58 (0.46, 0.72)		
2 or more previous births	201	1,562	NA	0.53 (0.41, 0.68)		
Maternal age						
< 25 years old	245	1,479	NA	0.82 (0.66, 1.03)	3.79	0.2853
25-29 years old	220	1,383	NA	1.00		
30-34 years old	118	1,000	NA	0.90 (0.70, 1.15)		
≥ 35 years old	66	526	NA	1.06 (0.77, 1.44)		
BMI						
quartile 1	232	1,100	NA	1.00	26.10	<0.0001
quartile 2	175	1,070	NA	0.83 (0.67, 1.03)		
quartile 3	131	1,132	NA	0.59 (0.46, 0.75)		
quartile 4	111	1,086	NA	0.59 (0.46, 0.77)		
Smoking						
never a smoker	492	3,203	NA	1.00	23.79	<0.0001
ever a smoker	54	613	NA	0.99 (0.71, 1.37)		
current smoker	103	572	NA	2.03 (1.50, 2.74)		
Maternal education						
No education	189	1,136	NA	1.05 (0.84, 1.30)	5.94	0.2040
School	224	1,414	NA	1.00		
Further education	65	549	NA	0.81 (0.60, 1.10)		
Higher education	142	1,001	NA	0.82 (0.64, 1.04)		
Other, Don't know, Unknown foreign	29	288	NA	0.80 (0.53, 1.22)		
Caffeinated drinks						
0 mg/day	549	3,727	NA	1.00	6.56	0.0104
0-200 mg/day	100	661	NA	1.42 (1.09, 1.85)		
Gestational diabetes						
no	602	3,949	NA	1.00	4.71	0.0299
yes	47	439	NA	0.70 (0.50, 0.97)		
BDCAA exposure						
0 - 1.04	226	1,511	1.00	1.00	1.27	0.5309
> 1.04 - 1.93	223	1,428	1.04 (0.86, 1.27)	1.07 (0.88, 1.31)		
> 1.93	200	1,449	0.92 (0.75, 1.13)	0.95 (0.77, 1.17)		
Ethnicity						
White British	130	1,519	NA	1.00	88.89	<0.0001
Pakistani	438	2,281	NA	3.78 (2.86, 4.98)		
Other	81	588	NA	2.32 (1.67, 3.22)		
Parity						
0 previous births	305	1,616	NA	1.00	33.79	<0.0001
1 previous birth	143	1,210	NA	0.57 (0.46, 0.72)		
2 or more previous births	201	1,562	NA	0.53 (0.41, 0.68)		
Maternal age						
< 25 years old	245	1,479	NA	0.82 (0.66, 1.03)	3.72	0.2933
25-29 years old	220	1,383	NA	1.00		
30-34 years old	118	1,000	NA	0.90 (0.70, 1.15)		
≥ 35 years old	66	526	NA	1.05 (0.77, 1.44)		
BMI						
quartile 1	232	1,100	NA	1.00		
quartile 2	175	1,070	NA	0.83 (0.67, 1.03)	26.30	<0.0001
quartile 3	131	1,132	NA	0.59 (0.46, 0.75)		
quartile 4	111	1,086	NA	0.59 (0.46, 0.76)		

Combined exposure ug/day (N=5,037)	cases (n)	non-cases (n)	Crude OR (95% CI)	Adjusted* OR (95% CI)	chi^2 (DF)	Prob > chi^2
Smoking						
never a smoker	492	3,203	NA	1.00	23.80	<0.0001
ever a smoker	54	613	NA	0.99 (0.72, 1.38)		
current smoker	103	572	NA	2.03 (1.50, 2.74)		
Maternal education						
No education	189	1,136	NA	1.05 (0.85, 1.30)	5.99	0.2002
School	224	1,414	NA	1.00		
Further education	65	549	NA	0.81 (0.60, 1.09)		
Higher education	142	1,001	NA	0.82 (0.64, 1.04)		
Other, Don't know, Unknown foreign	29	288	NA	0.81 (0.53, 1.23)		
Caffeinated drinks						
0 mg/day	549	3,727	NA	1.00	6.55	0.0105
0-200 mg/day	100	661	NA	1.42 (1.08, 1.85)		
Gestational diabetes						
no	602	3,949	NA	1.00	4.81	0.0283
yes	47	439	NA	0.70 (0.50, 0.96)		

*adjusted for 8 variables: ethnicity, parity, maternal age at delivery, mother's BMI at questionnaire completion (quartiles), smoking status during pregnancy, maternal education, caffeine intake, and gestational diabetes

(The red colour signals a comparison with the reference group that is below a critical p-value of 0.05.)

Table A7 - 15: For each of the three combined metrics (DCAA, TCAA and BDCAA in ug/day) in the continuous birth weight model, proportion of women by category, comparing
-the complete case (CC) dataset of N=5,040
-the raw analysis ("raw") analysis, i.e. not accounting for availability of any other covariates or exposure (different N depending on availability of data, see (N) in Table)
-the imputed (MI) dataset of N=11,874

		CC	raw	MI
		5040	(N)	11874
DCAA, TCAA or BDCAA exposure			**(6195)**	
	tertile 1	**0.31**	**0.33**	**0.32**
	tertile 2	**0.33**	**0.33**	**0.33**
	tertile 3	**0.36**	**0.33**	**0.35**
Caffeinated drinks			(8935)	
	0 mg/day	0.85	0.81	0.82
	0-200 mg/day	0.15	0.19	0.18
BMI			(9478)	
	quartile 1	0.26	0.25	0.25
	quartile 2	0.25	0.25	0.25
	quartile 3	0.25	0.25	0.25
	quartile 4	0.24	0.25	0.25
Smoking			(9825)	
	never a smoker	0.73	0.69	0.69
	ever a smoker	0.13	0.17	0.17
	current smoker	0.13	0.14	0.14
Maternal education			(9821)	
	School	0.33	0.31	0.31
	No education	0.26	0.21	0.22
	Further education	0.12	0.15	0.14
	Higher education	0.23	0.25	0.25
	Other, Don't know, Unknown foreign	0.06	0.08	0.08
Ethnicity			(9805)	
	White British	0.33	0.40	0.40
	Pakistani	0.54	0.44	0.45
	Other	0.13	0.15	0.15
Gestational diabetes			(11381)	
	no	0.90	0.92	0.92
	yes	0.10	0.08	0.08
Parity			(11442)	
	0 previous births	0.38	0.41	0.41
	1 previous birth	0.27	0.28	0.28
	2 or more previous births	0.35	0.31	0.31
Sex of child			(11874)	
	male	0.51	0.52	0.52
	female	0.49	0.48	0.48
Maternal age			(11874)	
	<25	0.34	0.32	0.32
	25-29	0.32	0.33	0.33
	30-34	0.22	0.23	0.23
	>=35	0.12	0.13	0.13
Gestational Age (weeks)			(11874)	
	25	0.00	0.00	0.00
	26	0.00	0.00	0.00
	27	0.00	0.00	0.00

	CC 5040	raw (N)	MI 11874
28	0.00	0.00	0.00
29	0.00	0.00	0.00
30	0.00	0.00	0.00
31	0.00	0.00	0.00
32	0.00	0.00	0.00
33	0.00	0.00	0.00
34	0.01	0.01	0.01
35	0.01	0.01	0.01
36	0.02	0.02	0.02
37	0.05	0.05	0.05
38	0.17	0.16	0.16
39	0.24	0.25	0.25
40	0.29	0.28	0.28
41	0.18	0.18	0.18
42	0.01	0.01	0.01
43	0.00	0.00	0.00
44	0.00	0.00	0.00

Table A7 - 16: Breakdown of missing values by categories compared to original raw data (example of DCAA combined metric model on continuous birth weight)

	Original data	% of 11874 total (including missing)	% of total (excluding missing)	average of 10 imputed datasets, among women with missing ethnicity	% among missing
ethnicity					
White British	3,953	33.3	40.3	**788**	**38.1**
Pakistani	4,341	36.6	44.3	**987**	**47.7**
Other	1,511	12.7	15.4	**294**	**14.2**
missing	**2,069**	**17.4**			
smoking					
Never	6,782	57.1	69.0	**1,424**	**69.5**
Ever	1,650	13.9	16.8	**324**	**15.8**
Currently	1,393	11.7	14.2	**301**	**14.7**
missing	**2,049**	**17.3**			
caffeine consumption					
No	7,280	61.3	81.5	**2,446**	**83.2**
Yes	1,655	13.9	18.5	**493**	**16.8**
missing	**2,939**	**24.8**			
maternal education					
School	3,015	25.4	30.7	**647**	**31.5**
No education	2,109	17.8	21.5	**491**	**23.9**
Further education	1,432	12.1	14.6	**288**	**14.0**
Higher education	2,502	21.1	25.5	**466**	**22.7**
Other, Don't know, Unknown foreign	763	6.4	7.8	**161**	**7.8**
missing	**2,053**	**17.3**			
DCAA combined metric					
tertile 1	2,067	17.4	33.4	**1,784**	**31.4**
tertile 2	2,063	17.4	33.3	**1,842**	**32.4**
tertile 3	2,065	17.4	33.3	**2,054**	**36.2**
missing	**5,679**	**47.8**			
parity					
no previous registerable birth	4,687	39.5	41.0	**148**	**34.2**
1 previous registerable birth	3,192	26.9	27.9	**121**	**27.9**
more previous registerable births	3,563	30.0	31.1	**164**	**37.9**
missing	**432**	**3.6**			
Gestational Diabetes					
No	10,459	88.1	91.9	**449**	**91.0**
Yes	922	7.8	8.1	**44**	**9.0**
missing	**493**	**4.2**			
BMI (quartiles)					
1	2,368	19.9	25.0	**601**	**25.1**
2	2,370	20.0	25.0	**605**	**25.2**
3	2,372	20.0	25.0	**594**	**24.8**
4	2,368	19.9	25.0	**597**	**24.9**
missing	**2,396**	**20.2**			

Table A7 - 17: Water consumption as mediator of positive predictors of birth weight (X=exposure=either maternal education, quintiles of IMD 2010 or physical activity in the past week; M=mediator=cold tap water (L/day); O=outcome=birth weight (g))
DK: don't know

	Step 1 X→O		Step 2 X→O		Step 3 M→O	Step 4 X+M→O	
maternal education							
no education	reference		reference		/	reference	
school	46.88 (16.35, 77.42)		0.04 (-0.01, 0.08)		/	48.26 (16.74, 79.77)	
further education	88.37 (51.54, 125.20)	10.5 (<0.001)	0.06 (0.00, 0.11)	2.8 (0.024)	/	92.07 (53.88, 130.27)	8.5 (<0.001)
higher education	87.38 (55.58, 119.18)		0.08 (0.03, 0.12)		/	74.71 (41.94, 107.48)	
other, DK, unknown foreign	105.74 (60.30, 151.18)		0.02 (-0.05, 0.09)		/	98.79 (51.37, 146.20)	
cold tap water (L/day)	/		/		**16.48 (2.14, 30.83)**	**15.66 (1.33, 30.00)**	
constant	3180.84 (3157.42, 3204.27)		1.17 (1.13, 1.20)		3215.92 (3195.25, 3236.59)	3162.30 (3132.91, 3191.68)	
IMD 2010 quintiles							
quintile 1 (most deprived)	reference		reference		/	reference	
quintile 2	62.60 (33.87, 91.33)		0.05 (0.01, 0.09)		/	61.55 (31.76, 91.34)	
quintile 3	128.91 (94.05, 163.78)	31.4 (<0.001)	-0.02 (-0.07, 0.03)	2.0 (0.095)	/	129.18 (93.02, 165.35)	28.8 (<0.001)
quintile 4	237.84 (172.93, 302.74)		-0.03 (-0.13, 0.06)		/	230.89 (164.09, 297.68)	
quintile 5 (least deprived)	238.36 (153.71, 323.02)		0.03 (-0.10, 0.15)		/	235.43 (149.20, 321.67)	
cold tap water (L/day)	/		/		**16.48 (2.14, 30.83)**	**16.70 (2.42, 30.98)**	
constant	3202.07 (3188.80, 3215.35)		1.20 (1.18, 1.22)		3215.92 (3195.25, 3236.59)	3179.72 (3157.73, 3201.71)	
physical activity in past week							
none	reference		reference		/	reference	
< 1 hour/week	44.01 (9.56, 78.45)		0.06 (0.01, 0.11)		/	41.43 (5.77, 77.09)	
1-3 hours per week	130.07 (90.04, 170.10)	18.1 (<0.001)	0.17 (0.11, 0.23)	32.0 (<0.001)	/	129.49 (87.72, 171.25)	16.0 (<0.001)
≥ 3 hours per week	148.05 (71.50, 224.60)		0.45 (0.34, 0.55)		/	141.90 (61.97, 221.83)	
cold tap water (L/day)	/		/		**16.48 (2.14, 30.83)**	**11.59 (-4.78, 27.96)**	
constant	3216.77 (3203.15, 3230.39)		1.18 (1.16, 1.20)		3215.92 (3195.25, 3236.59)	3200.75 (3176.87, 3224.63)	

(The red colour signals a p-value < 0.05.)

Table A7 - 18: Spearman correlations (Spearman's rho, p-value and sample size)

water consumption:	physical activity	maternal education	quintiles of IMD 2010 (1->5 most to least deprived)
cold tap water	0.0708	0.0217	-0.005
	<0.001	0.0373	0.6332
	7756	9177	9175
total tap water	0.0751	-0.0035	0.0424
	<0.001	0.7286	<0.001
	8211	9679	9678
bottled water	0.0859	0.1068	0.0896
	<0.001	<0.001	<0.001
	2847	3695	3700
total water	0.1629	0.0593	0.1229
	<0.001	<0.001	<0.001
	8290	9767	9766

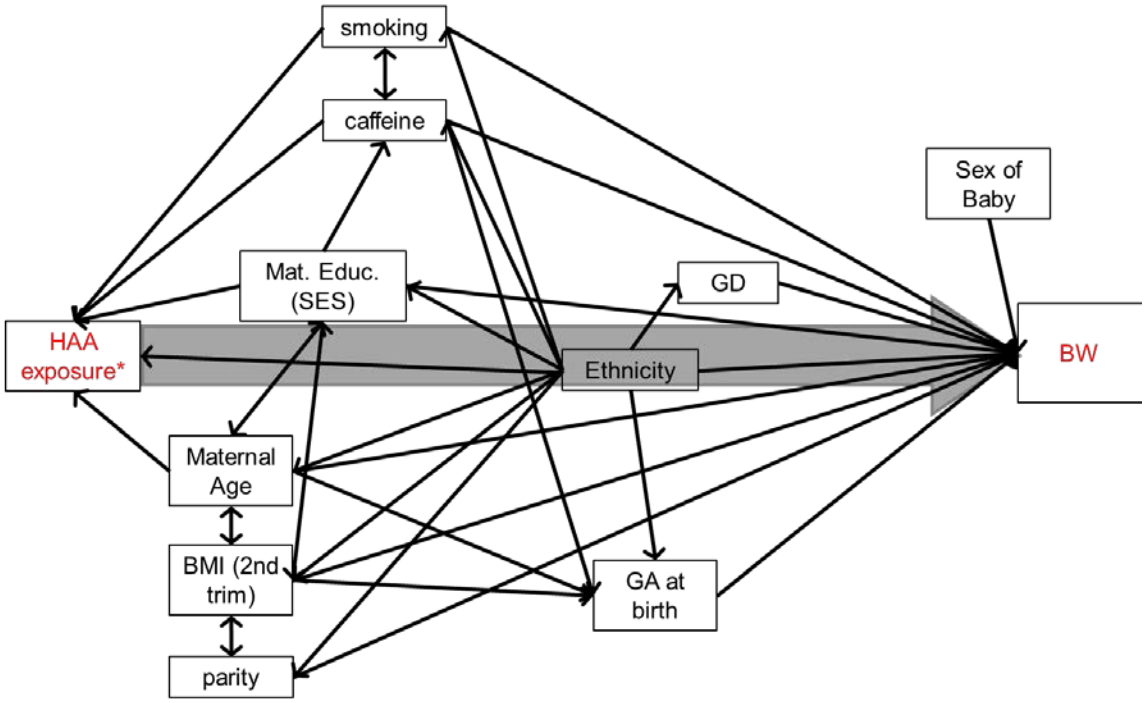
(The red colour signals a p-value < 0.05)

Table A7 - 19: Correlation between alcohol consumption during pregnancy and the three months preceding pregnancy and smoking, ethnicity, caffeine intake, parity and gestational diabetes (Spearman's rho (same as Cramer's V), p-value, N) (Nmax=11,928)

rho p-value N	alcohol consumption
smoking	0.33 <0.001 9517
ethnicity	-0.37 <0.001 9489
caffeine intake	0.21 <0.001 8682
parity	-0.13 <0.001 9169
gestational diabetes	-0.06 <0.001 9147

(The red colour signals a p-value <0.05).

Figure A7 - 1: Diagram of the relationships between the exposure and outcome (in red), and 10 covariates included in the continuous birth weight model (GD: gestational diabetes; GA: gestational age; BMI: body mass index; BW: birth weight)



*the exposure combines individual water consumption and area-level HAA concentrations from each mother's area of residence, either of which a covariate may be associated with.

Figure A7 - 2: Prevalence of term LBW by DCAA, TCAA, BDCAA exposure tertile (ug/day) (N total=5,869)

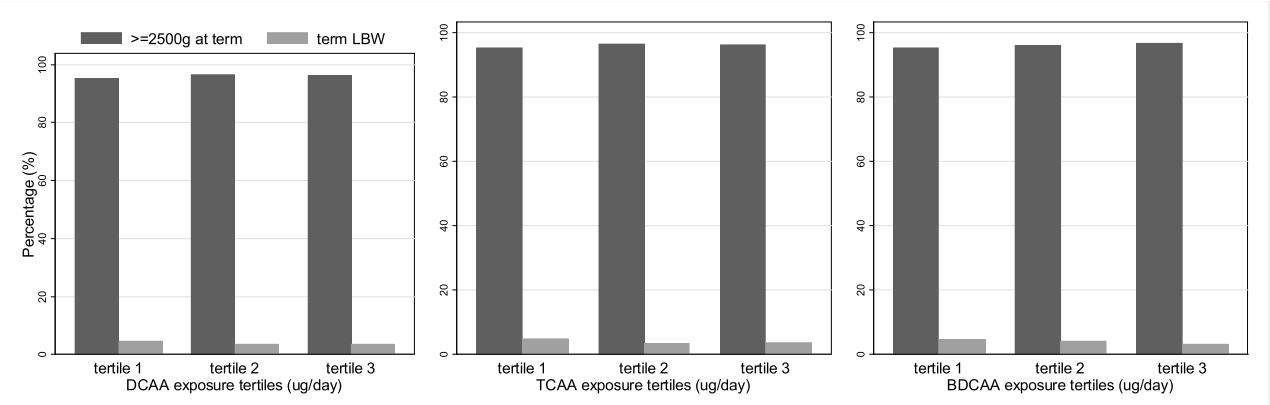


Figure A7 - 3: Boxplots of DCAA, TCAA and BDCAA exposure (ug/day) by term LBW (N total=5,869)

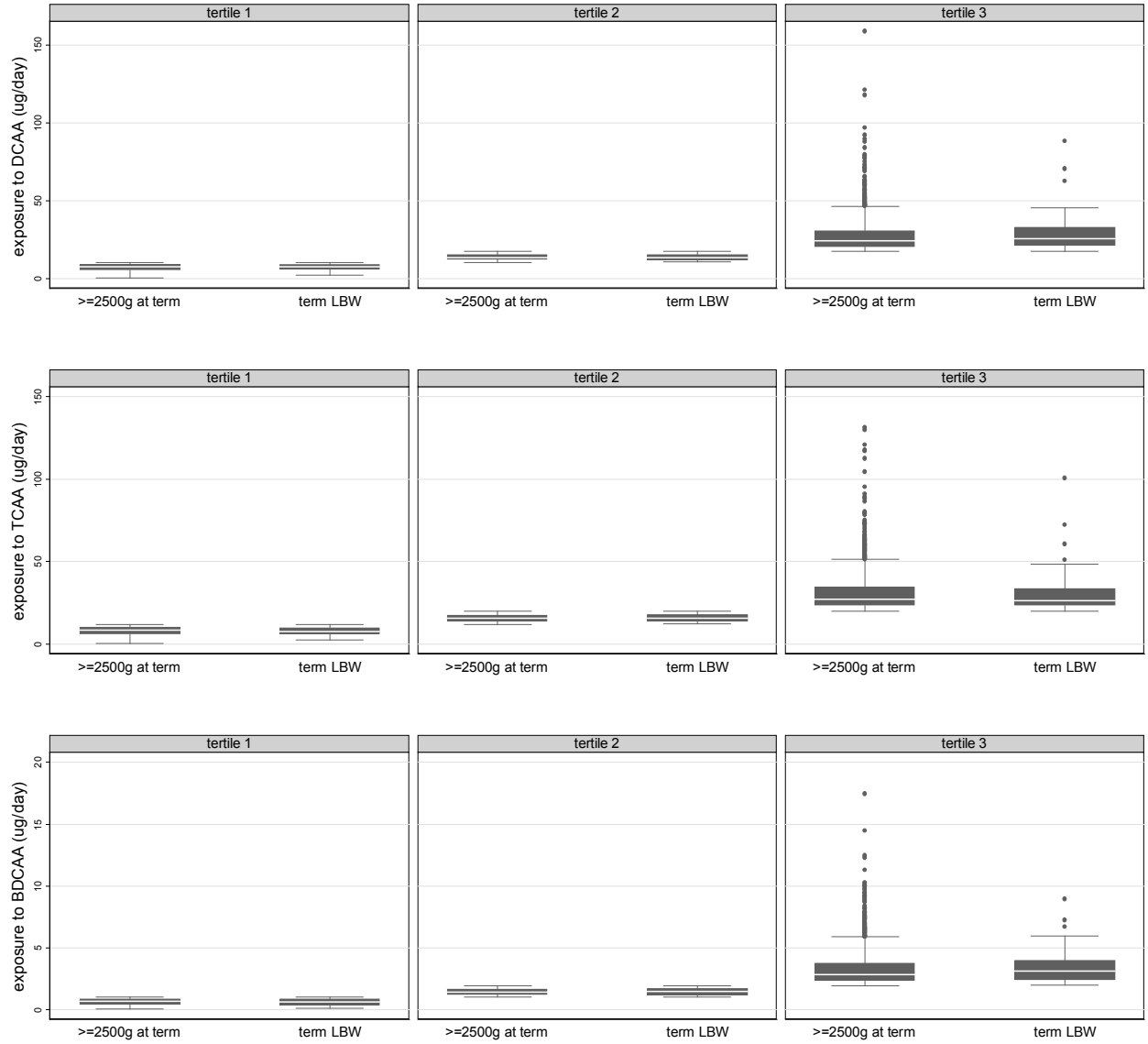


Figure A7 - 4: Prevalence of SGA (UK 1990) by DCAA, TCAA, BDCAA exposure tertile (ug/day) (N total=6,189) AGA=average for gestational age

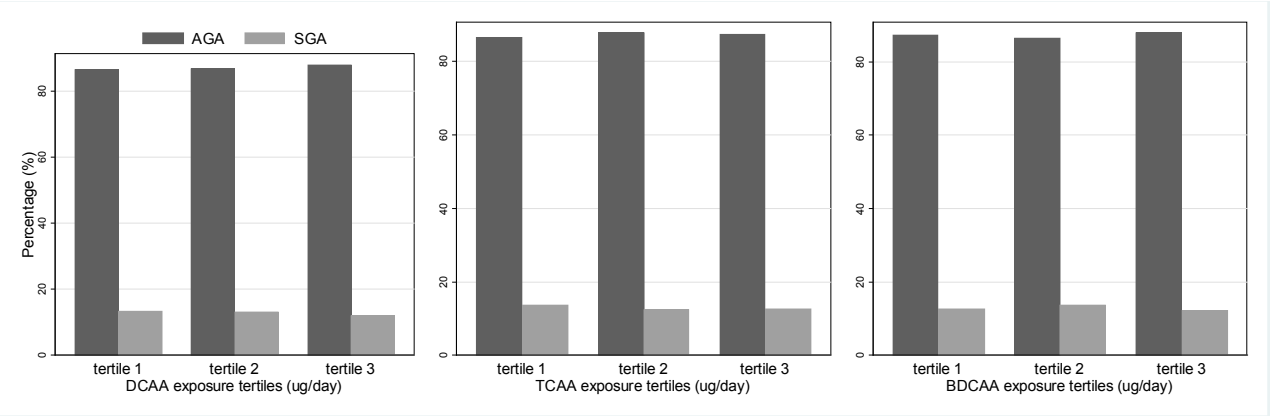
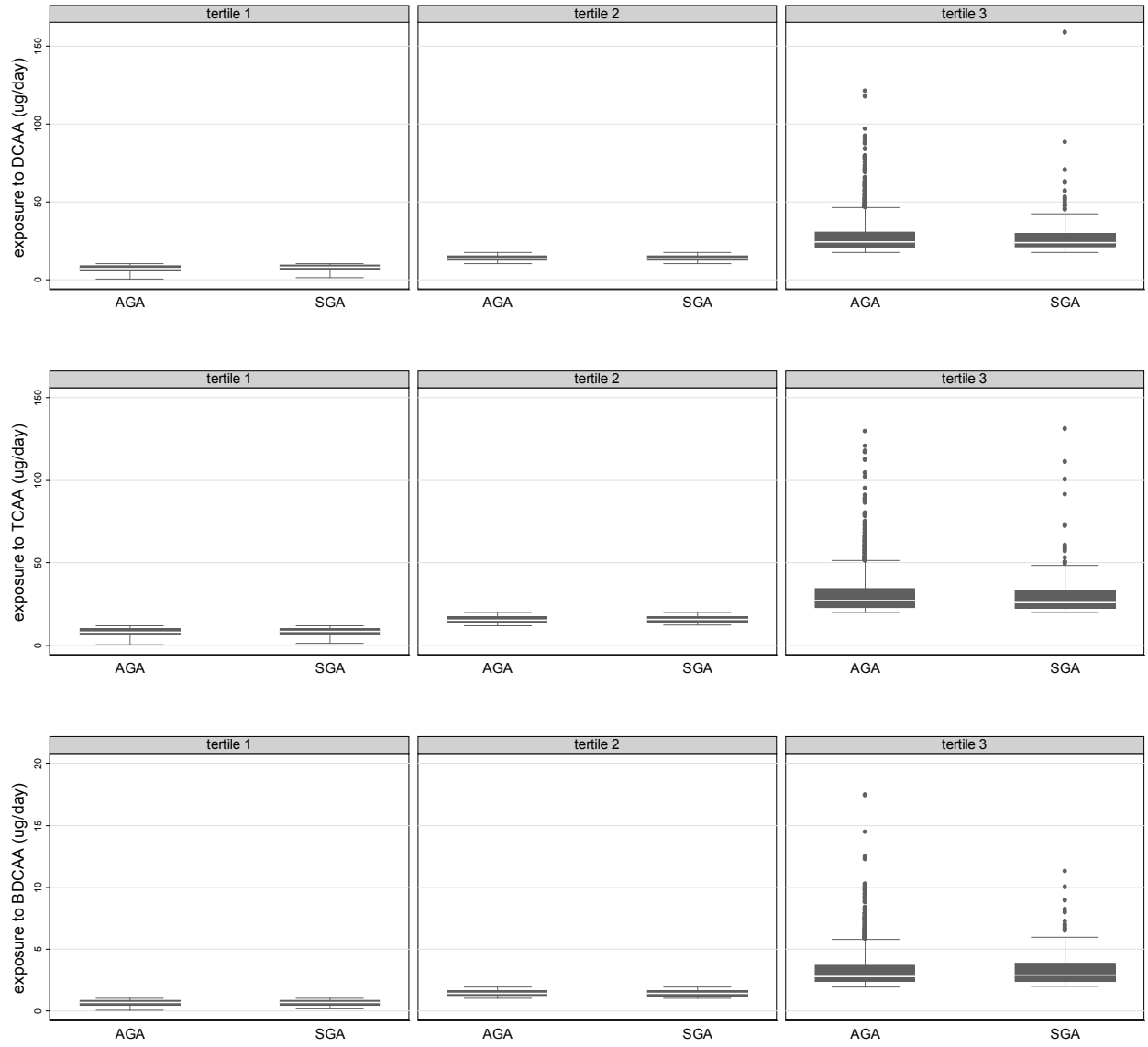


Figure A7 - 5: Boxplots of DCAA, TCAA and BDCAA exposure (ug/day) by SGA (UK 1990) (N total = 6,189)



Chapter 8

Table A8 - 1: Equations for the linear mixed models (all with random intercepts and random slopes) presented in Chapter 8

Model 1:	$y_{ij} = u_{0i} + u_{1i}\, time_{ij} + \varepsilon_{ij}$
	$u_{0i} \sim \mathrm{N}\,(\beta_0\,,\ \sigma^2_{u0})$
	$u_{1i} \sim \mathrm{N}\,(\beta_1\,,\ \sigma^2_{u1})$
Model 2:	$y_{ij} = u_{0i} + u_{1i}\, time_{ij} + \beta_2 E_{ij} + \beta_3 W_{ij} + \beta_4 S_{ij} + \sum_{x=1}^{p} \alpha_x\, C_{xij} + \varepsilon_{ij}$
	$u_{0i} \sim \mathrm{N}\,(\beta_0\,,\ \sigma^2_{u0})$
	$u_{1i} \sim \mathrm{N}\,(\beta_1\,,\ \sigma^2_{u1})$
Model 3:	$y_{ij} = u_{0i} + u_{1i}\, time_{ij} + \sum_g \beta_{2g}\, E_{gij} + \beta_3 W_{ij} + \beta_4 S_{ij} + \sum_{x=1}^{p} \alpha_x\, C_{xij} + \varepsilon_{ij}$
	$u_{0i} \sim \mathrm{N}\,(\beta_0\,,\ \sigma^2_{u0g})$
	$u_{1i} \sim \mathrm{N}\,(\beta_{1\mathrm{g}}\,,\ \sigma^2_{u1g})$
Model 4:	$y_{ij} = u_{0i} + u_{1i}\, time_{ij} + \beta_2 E_{ij} + \sum_g \gamma_g(W_{gij})\, time_{ij} + \sum_g \beta_{3\mathrm{h}}\, W_{gij} + \beta_4 S_{ij} + \sum_{x=1}^{p} \alpha_x\, C_{xij} + \varepsilon_{ij}$
	$u_{0i} \sim \mathrm{N}\,(\beta_0\,,\ \sigma^2_{u0})$
	$u_{1i} \sim \mathrm{N}\,(\beta_1\,,\ \sigma^2_{u1})$
Model 5:	$y_{ij} = u_{0i} + u_{1i}\, time_{ij} + \beta_2 E_{ij} + \beta_3 W_{ij} + \sum_g \gamma_g(S_{gi})\, time_{ij} + \sum_g \beta_{4\mathrm{k}}\, S_{gi} + \sum_{x=1}^{p} \alpha_x\, C_{xij} + \varepsilon_{ij}$
	$u_{0i} \sim \mathrm{N}\,(\beta_0\,,\ \sigma^2_{u0})$
	$u_{1i} \sim \mathrm{N}\,(\beta_1\,,\ \sigma^2_{u1})$

Legend: where i is subject, j is time point, g is ethnicity category, h is employment category, k is smoking category, p is total number of covariates, and:

y_{ij} = outcome (either total tap water, or total water) in L/day

u_{1i} = random slopes

u_{0i} = random intercepts

γ_g = interaction term

β_2, β_3, β_4 = main effect of ethnicity, employment, and smoking, respectively

β_{1g} = main time effect for every ethnic group (Model 3)

β_{2g} = main effect of ethnicity by ethnicity category (Model 3)

β_{3h} = main effect of work by work category (Model 4)

β_{4k} = main effect of smoking by smoking category (Model 5)

α_x = main effect of other covariates (C)

ε_{ij} = residual error

β_0 = main intercept

β_1 = main time effect

σ^2_{u0} = random intercepts variance

σ^2_{u1} = random slopes variance

σ^2_{u0g} = random intercepts variance by ethnic group (Model 3)

σ^2_{u1g} = random slopes variance by ethnic group (Model 3)

E=ethnicity (3 categories, where the reference category is White British)

W=employment status (4 categories, where the reference category is Employed and currently working)

S=smoking (3 categories, where the reference category is never smoker)

C=other covariates (maternal physical exercise for leisure, maternal physical exercise at work, total number of household members)

Table A8 - 2: Number of entries (and percentage of total number of entries) in each water consumption category (by water type, and by location of consumption) that were left blank in the original repeat questionnaires, and in the summary variables derived for this analysis. These missing entries were subsequently replaced by 0's in the main analysis. There were no such missing entries in the baseline data.

original variables in the questionnaire						**summary variables used in the manuscript**					
		Q1 (N=209)	**% blank**	**Q2 (N=172)**	**% blank**			**Q1 (N=209)**	**% blank**	**Q2 (N=172)**	**% blank**
tap water	home	19	9%	16	9%						
	work/study	88	42%	97	56%						
	elsewhere	94	45%	83	48%						
squash	home	33	16%	29	17%						
	work/study	95	45%	104	60%						
	elsewhere	99	47%	89	52%						
						cold tap water (tap water+squash)	home	42	20%	35	20%
							out of the home	124	59%	113	66%
							total	124	59%	113	66%
tea	home	26	12%	18	10%						
	work/study	90	43%	99	58%						
	elsewhere	96	46%	83	48%						
coffee	home	57	27%	56	33%						
	work/study	95	45%	104	60%						
	elsewhere	111	53%	92	53%						
						hot tap water (tea+coffee)	home	64	31%	59	34%
							out of the home	117	56%	112	65%
							total	117	56%	112	65%
bottled water	home	65	31%	56	33%	bottled water	home	65	31%	56	33%
	work/study	90	43%	94	55%						
	elsewhere	104	50%	84	49%						
							out of the home	116	56%	105	61%
							total	117	56%	107	62%
						total tap water (cold+hot)	home	71	34%	66	38%
							out of the home	128	61%	116	67%
							total	128	61%	116	67%
						total water (total tap+bottled water)	home	82	39%	74	43%
							out of the home	130	62%	116	67%
							total	130	62%	116	67%

Table A8 - 3: Univariate linear mixed models of RQS women's a) total tap water ("TAP") and b) total water ("Total") consumption (in L/day) over the 3 time points of interest: baseline, Q1 (30-33 weeks of pregnancy) and Q2 (36-39 weeks of pregnancy) (Model 1).
Example of sensitivity analyses results for the unadjusted model with a t-distribution (4 degrees of freedom) (Model 1b), and after censoring extreme and unsure outcome values (Model 1c).

a. TAP	Model 1		Model 1b		Model 1c	
Fixed effect						
Main intercept	1.86	(1.71, 2.01)	1.73	(1.61, 1.85)	1.85	(1.72, 1.97)
Time pattern	0.13	(0.04, 0.23)	0.13	(0.06, 0.20)	0.64	(0.53, 0.75)
Random effect						
Random intercepts SD	0.76	(0.65, 0.89)	0.70	(0.61, 0.80)	0.70	(0.58, 0.82)
Random slopes SD	0.09	(0.00, 0.25)	0.20	(0.06, 0.32)	0.31	(0.14, 0.45)
Measurement error SD	1.01	(0.94, 1.09)	0.76	(0.68, 0.84)	0.78	(0.70, 0.87)
Variance Partition Coefficient	0.37	(0.28, 0.46)	0.48	(0.39, 0.58)	0.49	(0.38, 0.60)

b. Total	Model 1		Model 1b		Model 1c	
Fixed effect						
Main intercept	2.22	(2.07, 2.37)	2.08	(1.97, 2.20)	2.19	(2.07, 2.31)
Time pattern	0.14	(0.03, 0.26)	0.12	(0.03, 0.20)	0.66	(0.55, 0.77)
Random effect						
Random intercepts SD	0.73	(0.58, 0.87)	0.62	(0.52, 0.72)	0.60	(0.47, 0.72)
Random slopes SD	0.29	(0.07, 0.47)	0.24	(0.07, 0.38)	0.34	(0.18, 0.47)
Measurement error SD	1.08	(1.00, 1.16)	0.85	(0.76, 0.95)	0.81	(0.72, 0.90)
Variance Partition Coefficient	0.35	(0.26, 0.44)	0.38	(0.28, 0.49)	0.43	(0.30, 0.54)

Red means significant (i.e. 95% credible interval does not cross zero).

Table A8 - 4: Water consumption values at baseline in the RQS and the rest of the cohort (means (95% confidence intervals))

	RQS			Rest			Difference		*p** (2-way)
	mean	(95% CI)	n	mean	(95% CI)	n	mean	(95% CI)	
Total tap water L/day	1.84	(1.70, 1.98)	254	1.33	(1.31, 1.35)	13519	-0.51	(-0.64, -0.38)	<0.001
Total water L/day	2.19	(2.06, 2.33)	254	1.56	(1.54, 1.58)	13519	-0.63	(-0.77, -0.48)	<0.001

*two-sample t-test, p-value red if significant at <0.05

Table A8 - 5: Water consumption values at baseline in the RQS and the rest of the cohort (means (95% confidence intervals)), stratified by ethnicity

	RQS			Rest of cohort			Difference		*p***
	mean	(95% CI)	n	mean	(95% CI)	n	mean	(95% CI)	
Total tap water L/day									
White British	2.11	(1.88, 2.33)	123	1.81	(1.77, 1.84)	4365	-0.30	(-0.50, -0.10)	0.235
Pakistani	1.61	(1.42, 1.80)	74	1.49	(1.47, 1.51)	5053	-0.11	(-0.28, 0.05)	
Other	1.56	(1.32, 1.81)	57	1.48	(1.44, 1.53)	1675	-0.08	(-0.33, 0.17)	
Total water L/day									
White British	2.47	(2.26, 2.69)	123	2.20	(2.17, 2.24)	4365	-0.27	(-0.49, -0.06)	0.738
Pakistani	1.84	(1.65, 2.04)	74	1.65	(1.62, 1.67)	5053	-0.20	(-0.38, -0.02)	
Other	2.03	(1.80, 2.27)	57	1.88	(1.83, 1.93)	1675	-0.15	(-0.42, 0.11)	

** Two-factor ANOVA, p for interaction term between two factors, ethnicity and inclusion in RQS

Table A8 - 6: Water consumption values at baseline in the RQS and the rest of the cohort (means (95% confidence intervals)), stratified by employment status

	RQS			Rest			Diff		*p***
	mean	(95% CI)	n	mean	(95% CI)	n	mean	(95% CI)	
Total tap water L/day									
Employed & currently working	1.83	(1.65, 2.00)	142	1.67	(1.64, 1.70)	4309	-0.16	(-0.33, 0.02)	0.662
Employed & currently on mat leave	2.15	(1.41, 2.90)	13	1.68	(1.59, 1.77)	432	-0.47	(-1.01, 0.07)	
Not working	1.82	(1.59, 2.06)	90	1.57	(1.55, 1.59)	6012	-0.25	(-0.43, -0.07)	
Full-time student	1.73	(0.58, 2.89)	9	1.58	(1.48, 1.68)	376	-0.15	(-0.80, 0.50)	
Total water L/day									
Employed & currently working	2.28	(2.11, 2.44)	142	2.17	(2.14, 2.21)	4309	-0.10	(-0.29, 0.08)	0.089
Employed & currently on mat leave	2.82	(2.20, 3.43)	13	2.06	(1.96, 2.16)	432	-0.75	(-1.35, -0.16)	
Not working	1.99	(1.74, 2.24)	90	1.69	(1.67, 1.71)	6012	-0.30	(-0.49, -0.12)	
Full-time student	1.93	(0.90, 2.97)	9	1.94	(1.83, 2.04)	376	0.00	(-0.70, 0.70)	

** Two-factor ANOVA, p for interaction term between two factors, employment status and inclusion in RQS

Table A8 - 7: Water consumption values at baseline in the RQS and the rest of the cohort (means (95% confidence intervals)), stratified by smoking status

	RQS			Rest			Diff		p
	mean	(95% CI)	n	mean	(95% CI)	n	mean	(95% CI)	
Total tap water L/day									
Currently	2.47	(1.86, 3.08)	25	1.92	(1.86, 1.98)	1550	-0.55	(-1.02, -0.08)	0.033
Ever	2.08	(1.79, 2.36)	64	1.70	(1.65, 1.74)	1824	-0.38	(-0.63, -0.12)	
Never	1.65	(1.50, 1.80)	165	1.54	(1.52, 1.56)	7747	-0.12	(-0.25, 0.02)	
Total water L/day									
Currently	2.80	(2.15, 3.45)	25	2.21	(2.14, 2.27)	1550	-0.59	(-1.09, -0.10)	0.219
Ever	2.34	(2.06, 2.62)	64	2.16	(2.11, 2.21)	1824	-0.18	(-0.46, 0.10)	
Never	2.04	(1.89, 2.19)	165	1.78	(1.76, 1.80)	7747	-0.26	(-0.40, -0.12)	

** Two-factor ANOVA, p for interaction term between two factors, smoking status and inclusion in RQS

Figure A8 - 1: Timeline of recruitment to BiB (26-28 weeks of pregnancy) in grey, and to RQS (Q1 at 30-33 weeks, Q2 at 36-39 weeks of pregnancy) in pink

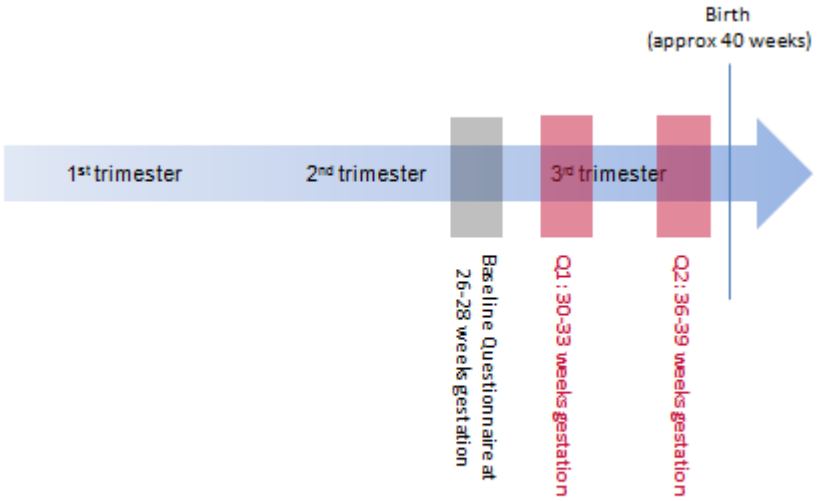
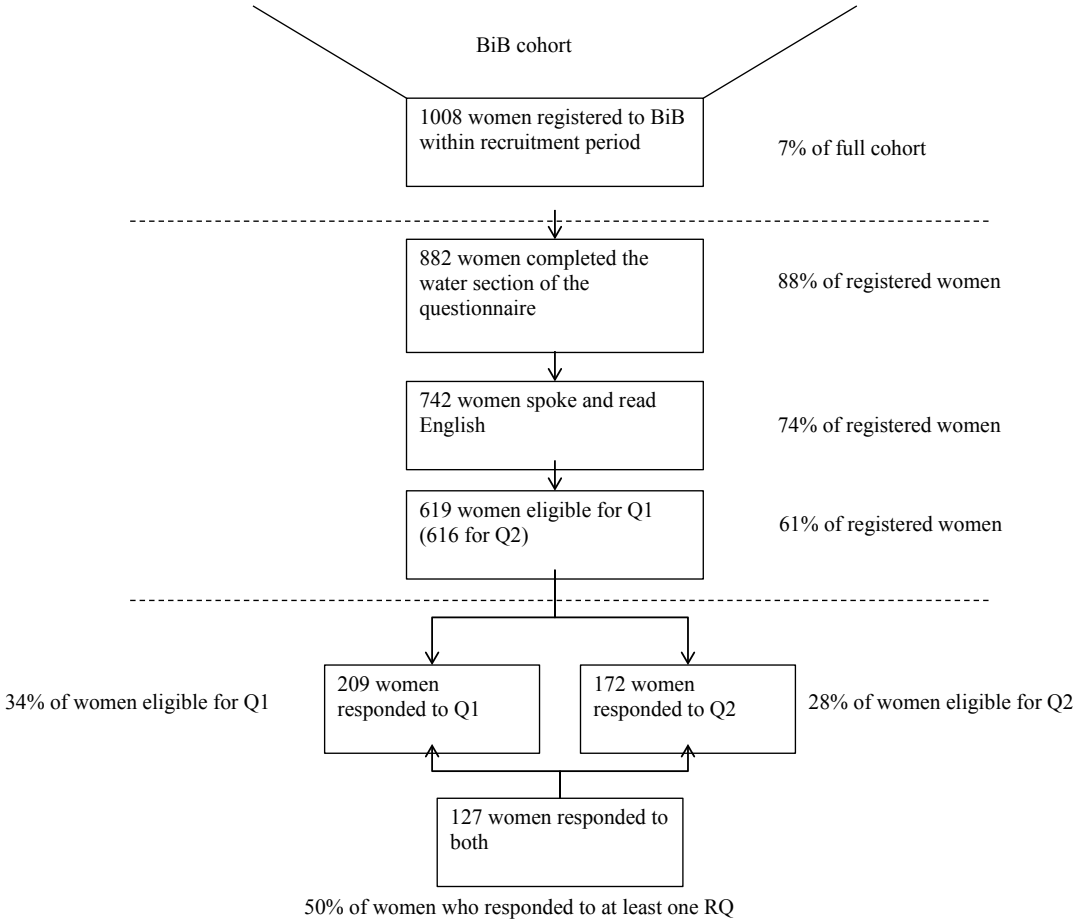


Figure A8 - 2: Diagram of eligibility



APPENDIX A

BiB baseline questionnaire – water section

Section I Water Consumption

***I1. On a typical day how much of the following do you drink?**

		At home	At work/study	Elsewhere
a) Tap water	Glasses per day:			
b) Bottled water (Includes water cooler)	Glasses per day:			
c) Tea (any sort)	Cups per day:			
d) Coffee	Cups per day:			
e) Squash (Including any other drinks made with tap water)	Glasses per day:			

I2. Do you filter the water you drink at home? (Cross ONE box ONLY)

☐ Yes ☐ No ☐ Don't Know

I3. Do you filter the water you drink at work? (Cross ONE box ONLY)

☐ Yes ☐ No ☐ Don't Know ☐ N/A

I4. In a typical week while you have been pregnant how often and for how long do you undertake the following?

(if you do not do any then fill in 0)

	Times per week	Minutes each time
Shower		
Bath		
Swim		

APPENDIX B

Manuscript contribution

Jeong CH, Wagner ED, Siebert VR, Anduri S, Richardson SD, Daiber EJ, McKague AB, Kogevinas M, Villanueva CM, Goslan EH, Luo W, Isabelle LM, Pankow JF, Grazuleviciene R, Cordier S, **Edwards SC**, Righi E, Nieuwenhuijsen MJ, Plewa MJ, *Occurrence and Toxicity of Disinfection Byproducts in European Drinking Waters in Relation with the HIWATE Epidemiology Study*, dx.doi.org/10.1021/es3024226 Environ. Sci. Technol. 2012, 46, 12120−12128

Environmental Science & Technology

Article

pubs.acs.org/est

Occurrence and Toxicity of Disinfection Byproducts in European Drinking Waters in Relation with the HIWATE Epidemiology Study

Clara H. Jeong,[†] Elizabeth D. Wagner,[†] Vincent R. Siebert,[†] Sridevi Anduri,[‡,$] Susan D. Richardson,[‡] Eric J. Daiber,[‡,$] A. Bruce McKague,[§] Manolis Kogevinas,[∥] Cristina M. Villanueva,[∥] Emma H. Goslan,[⊥] Wentai Luo,[#] Lorne M. Isabelle,[#] James F. Pankow,[#] Regina Grazuleviciene,[▽] Sylvaine Cordier,[○] Susan C. Edwards,[◆] Elena Righi,[¶] Mark J. Nieuwenhuijsen,[∥] and Michael J. Plewa[†,*]

[†]College of Agricultural, Consumer and Environmental Sciences, University of Illinois at Urbana−Champaign, Urbana, IL, United States

[‡]National Exposure Research Laboratory, U.S. Environmental Protection Agency, Athens, Georgia, United States

[§]CanSyn Chem. Corp., Toronto, Ontario, Canada

[∥]Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; Hospital del Mar Research Institute (IMIM); CIBER Epidemiologia y Salud Publica (CIBERESP)

[⊥]Cranfield Water Science Institute, Cranfield University, Cranfield, Bedford, United Kingdom

[#]Portland State University, Portland, Oregon, United States

[▽]Department of Environmental Sciences, Vytautas Magnus University, Kaunas, Lithuania

[○]National Institute for Health and Medical Research (INSERM), U1085-IRSET; University of Rennes I, Rennes, France

[◆]MRC-HPA Centre for Environment & Health, Imperial College, London, United Kingdom

[¶]Department of Public Health Sciences, University of Modena and Reggio Emilia, Modena, Italy

Supporting Information

ABSTRACT: The HIWATE (**H**ealth **I**mpacts of long-term exposure to disinfection byproducts in drinking **WATE**r) project was a systematic analysis that combined the epidemiology on adverse pregnancy outcomes and other health effects with long-term exposure to low levels of drinking water disinfection byproducts (DBPs) in the European Union. The present study focused on the relationship of the occurrence and concentration of DBPs with in vitro mammalian cell toxicity. Eleven drinking water samples were collected from five European countries. Each sampling location corresponded with an epidemiological study for the HIWATE program. Over 90 DBPs were identified; the range in the number of DBPs and their levels reflected the diverse collection sites, different disinfection processes, and the different characteristics of the source waters. For each sampling site, chronic mammalian cell cytotoxicity correlated highly with the numbers of DBPs identified and the levels of DBP chemical classes. Although there was a clear difference in the genotoxic responses among the drinking waters, these data did not correlate as well with the chemical analyses. Thus, the agents responsible for the genomic DNA damage observed in the HIWATE samples may be due to unresolved associations of combinations of identified DBPs, unknown emerging DBPs that were not identified, or other toxic water contaminants. This study represents the first to integrate quantitative in vitro toxicological data with analytical chemistry and human epidemiologic outcomes for drinking water DBPs.

INTRODUCTION

The introduction of water disinfection greatly reduced the incidence of waterborne infectious diseases.[1] Although chlorine is the most common disinfectant, alternatives include ozone, chloramines, chlorine dioxide, and UV radiation.[2−4] An unintended consequence of disinfection is the formation of drinking water disinfection byproducts (DBPs) from the reaction between organic and inorganic materials in the water and disinfectants. Chemical classes of DBPs include halomethanes, haloacetic acids (HAAs) and nitrogen-containing DBPs (N-DBPs); to date, more than 600 DBPs have been identified in drinking water.[5,6] The spectrum of DBP generation depends on the source water, pH, temperature, disinfection type, and processes.[5,7−9] Less than 20 DBPs are currently regulated in the United States and in other countries.[6,10]

Received: June 15, 2012
Revised: September 5, 2012
Accepted: September 7, 2012
Published: September 7, 2012

Previous epidemiological studies reported associations between DBPs in chlorinated water and increased cancer risk[11−15] as well as DBPs and adverse pregnancy outcomes including spontaneous abortion, low birth weight (LBW), small-for-gestational-age (SGA), still birth, and preterm delivery.[16−19] HAAs were teratogenic in mice embryos;[20] mixtures of trihalomethanes (THMs) and HAAs were teratogenic in rats.[21]

In 2006, the European Union (EU) established the project HIWATE (**H**ealth **I**mpacts of long-term exposure to disinfection byproducts in drinking **WATE**r) to investigate potential human health risks associated with long-term exposure to DBPs.[22] Pregnancy cohorts (N ∼23 000) were included from France, Lithuania, Spain, Italy, and the United Kingdom (Table 1). These locations encompassed a variety of disinfectants and treatments including chlorine, ozone, chlorine dioxide, and desalination with reverse osmosis. Metrics for adverse pregnancy outcomes were LBW, SGA, preterm delivery, fetal growth restriction (FGR), and parameters derived from ultrasound medical diagnosis.

Table 1. HIWATE Water Sampling Locations and Applied Disinfection Methods[a]

sample number	sampling location (site)	disinfection method
HIWATE 1	Barcelona, Spain (Badalona)	Cl_2−Cl_2
HIWATE 2	Barcelona, Spain (Hospitalet del Llobregat)	Blend of Cl_2−Cl_2, Cl_2−O_3-Cl_2, Desal-RO-ClO_2
HIWATE 3	Barcelona, Spain (Sabadell)	Blend of (ClO_2/Cl_2)-Cl_2, Cl_2−Cl_2
HIWATE 4	Kaunas, Lithuania (Petruniusai)	Cl_2
HIWATE 5	Modena, Italy	ClO_2
HIWATE 6	Kaunas, Lithuania (Viciunai)	Cl_2
HIWATE 7	Valencia, Spain	Cl_2−Cl_2
HIWATE 8	Rennes, France	O_3−Cl_2
HIWATE 9	Asturias, Spain	Cl_2
HIWATE 10	Bradford, U.K. (Shipley)	Cl_2
HIWATE 11	Bradford, U.K. (Airedale)	Cl_2

[a]Cl_2 = chlorination, O_3 = ozonation, ClO_2 = chlorine dioxide, Desal-RO = desalination with reverse osmosis.

This project represents the first systematic analysis combining DBP analytical chemistry and in vitro mammalian cell toxicology with adverse pregnancy outcomes. Our objectives were to (i) obtain disinfected drinking water from HIWATE cities, extract and concentrate the organic fraction and chemically analyze for DBPs, (ii) determine the relative chronic cytotoxicity and acute genotoxicity in mammalian cells for each HIWATE sample, and (iii) analyze for correlations between the toxicity data and the occurrence and concentrations of DBPs.

EXPERIMENTAL SECTION

Chemicals and Reagents. General reagents were purchased from Sigma-Aldrich Co. (St. Louis, MO) and Fisher Scientific Co. (Itasca, IL). Media and fetal bovine serum (FBS) were purchased from Fisher Scientific Co. (Itasca, IL). Chemical standards were purchased from Sigma-Aldrich, ChemService (West Chester, PA), Orchid Cellmark (Westminster, BC, Canada), and TCI America (Waltham, MA) at the highest level of purity.

Sample Preparation. Drinking water samples (20 L) were collected from 11 different distribution systems from 7 cities within 5 European countries, where an epidemiologic study of reproductive outcomes was being conducted. Samples were collected from March-June 2010 using 2 L Teflon bottles (headspace-free) and were commercially shipped in coolers with icepacks to the U.S. Environmental Protection Agency (U.S. EPA) laboratory in Athens, GA. Water samples were analyzed immediately upon arrival using XAD resins.[23] The final extract (2 mL in ethyl acetate) was equally divided for GC/MS analysis and genotoxicity/cytotoxicity analysis. For toxicity analyses the solvent ethyl acetate was evaporated with a stream of dry N_2 and exchanged to dimethylsulfoxide (DMSO) resulting in a 10^5 × concentration. These samples were stored in glass Supelco 1-mL Micro Reaction Vessels (no. 27036) at −20 °C.

Broad-Screen GC/MS Analyses. Half of the extract was derivatized with diazomethane[24] to identify halo-acids (through their corresponding methyl esters) while the other half was analyzed directly for other DBPs. Comprehensive gas chromatography/mass spectrometry (GC/MS) analyses were performed on a high-resolution magnetic sector mass spectrometer (Autospec, Waters, Inc.) in electron ionization mode, equipped with an Agilent model 6890 gas chromatograph and operated at an accelerating voltage of 8 kV and source temperature of 200 °C, in both low-resolution (1000) and high-resolution (10 000) modes. Injections of 1 μL of the extracts were introduced via a split/splitless injector (in splitless mode) onto a GC column (ZB-5, 30 m × 0.25 mm ID, 0.25 μm film thickness, Phenomenex (Torrance, CA). The GC temperature program consisted of an initial temperature of 35 °C (4 min) followed by an increase at 9 °C/min to 285 °C (held for 30 min). Transfer lines were held at 280 °C and the injection port at 250 °C. To prevent decomposition of trihalonitromethanes, separate analyses were made with an injection port temperature of 180 °C.[25] For analysis of data by the Massworks expert system,[26] extracts were analyzed in the continuum mode at 1000 resolution.

Mass spectra of unknown compounds in the drinking water extracts were subjected to library database searching (National Institute of Standards and Technology and Wiley databases). For DBPs not present in either database, high-resolution-MS and Massworks software (Cerno Bioscience, Norwalk, CT) were used to provide empirical formulas for molecular ions and fragments. Mass spectra were also interpreted extensively to provide tentative structural identifications. When possible, pure standards were obtained to confirm identifications through a match of GC retention times and mass spectra.

GC × GCTOFMS Measurements. GC × GC-time-of-flight-MS (GC × GC-TOF-MS) measurements were conducted using a Leco Pegasus 4D GC × GC-TOF mass spectrometer (Leco Corp., St. Joseph, Michigan). One μL of the extracts was introduced via a split/splitless injector (in splitless mode). A DB-VRX (45 m, 0.25 mm i.d., 1.4 um film thickness, Agilent, Santa Clara CA) served as the primary column and a Stabilwax (1.5 m, 0.25 mm i.d., 0.25 um film thickness, Restek, Bellefonte, PA) as the secondary column. The primary GC oven program consisted of an initial temperature of 45 °C (3 min), an increase at 10 °C/min to 145 °C (3 min), an increase at 5 °C/min to 240 °C, and final hold of 20 min. The secondary GC oven was 13 °C above the

Table 2. Levels of DBPs by Chemical Classes and Correlation with CHO Cell Cytotoxic Potency Index and Genotoxic Potency Index

HIWATE sample number	21 DBPs[c] (μg/L)	4 THMs (μg/L)	9 HAAs (μg/L)	4 HANs (μg/L)	2 HKs (μg/L)	CH (μg/L)	CP (μg/L)	U.S.-regulated DBPs (μg/L)	unregulated DBPs (μg/L)
1	115	70.9	36.0	6.47	0.21	1.27	<LOD[d]	94.7	20.1
2	91.1	66.8	19.5	4.70	<LOD	<LOD	<LOD	77.1	13.9
3	202	139	51.5	8.88	1.11	1.80	<LOD	168	33.8
4	3.24	3.24	<LOD	<LOD	<LOD	<LOD	<LOD	3.24	<LOD
5	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
6	1.11	1.11	<LOD	<LOD	<LOD	<LOD	<LOD	1.11	<LOD
7	118	83.9	26.7	4.80	<LOD	2.46	<LOD	103	14.6
8	27.8	14.7	13.3	<LOD	<LOD	<LOD	<LOD	26.5	1.25
9	92.9	55.0	<LOD	4.02	6.86	23.2	3.78	55.0	37.9
10	40.6	22.6	13.3	<LOD	3.28	1.43	<LOD	29.0	11.6
11	45.2	29.3	11.6	<LOD	3.08	1.23	<LOD	36.8	8.37
cytotoxic potency index[a]	$r = 0.77$	$r = 0.77$	$r = 0.75$	$r = 0.73$	$r = 0.04$	$r = 0.04$	$r = -0.02$	$r = 0.76$	$r = 0.60$
	$P \leq 0.006$	$P \leq 0.006$	$P \leq 0.009$	$P = 0.011$	$P = 0.913$	$P = 0.905$	$P = 0.947$	$P \leq 0.006$	$P = 0.051$
genotoxic potency index[b]	$r = 0.36$	$r = 0.37$	$r = 0.40$	$r = 0.36$	$r = -0.08$	$r = -0.12$	$r = -0.18$	$r = 0.38$	$r = 0.22$
	$P = 0.273$	$P = 0.260$	$P = 0.221$	$P = 0.281$	$P = 0.827$	$P = 0.720$	$P = 0.600$	$P = 0.248$	$P = 0.521$

[a]The CHO cell cytotoxic potency index value corresponds to ($LC_{50}^{-1} \times 10^3$) for each HIWATE sample. [b]The CHO cell genotoxic potency index value is the reciprocal HIWATE sample concentration factor that was calculated to induce a 50% SCGE tail DNA value $\times 10^4$. [c]These 21 quantitatively measured DBPs are listed in the text. [d]LOD = limit of detection.

primary GC oven. The modulator offset was 20 °C above to the primary GC oven. The modulation period was 7 s with 1.5 s hot pulse. The transfer line and source temperature were maintained at 248 and 200 °C, respectively. The MS data were acquired from m/z 35 to 500 at rate of 150 spectra/s in electron ionization mode.

Quantitative Chemical Analyses. THMs (chloroform, bromodichloromethane, dibromochloromethane, and bromoform), haloacetonitriles (dichloroacetonitrile, bromochloroacetonitrile, dibromoacetonitrile, and trichloroacetonitrile), haloketones (1,1-dichloro- and 1,1,1-trichloropropanone), trichloroacetaldehyde (chloral hydrate), and trichloronitromethane (chloropicrin) were extracted using a modified form of U.S. EPA Method 551.1.[27] HAAs (chloro-, bromo-, dichloro-, trichloro-, bromochloro-, dibromo-, bromodichloro-, dibromochloro-, and tribromoacetic acid) were analyzed using a modified form of U.S. EPA Method 552.3.[28] The limit of detection for each DBP was 1 μg/L, with the exception of chloroacetic acid (detection limit was 2 μg/L).

Chinese Hamster Ovary Cells. Chinese hamster ovary (CHO) cell line AS52, clone 11−4−8 was used for the biological assays.[29−31] CHO cells were maintained on glass culture plates in Ham's F12 medium containing 5% fetal bovine serum (FBS), 1% antibiotics (100 U/mL sodium penicillin G, 100 μg/mL streptomycin sulfate, 0.25 μg/mL amphotericin B in 0.85% saline), and 1% glutamine at 37 °C in a humidified atmosphere of 5% CO_2.

CHO Cell Chronic Cytotoxicity Assay. This assay measures the reduction in cell density on flat-bottom 96-well microplates as a function of the concentration of the test sample over a period of approximately 72 h (~3 cell cycles).[32,33] Microliters of the sample in DMSO were diluted with F12 +FBS medium to analyze a range of concentration factors. This assay was calibrated; the detailed procedure was published[32,33] and is presented in the Supporting Information (SI). For each HIWATE sample concentration factor, four replicate wells were analyzed. The experiments were repeated 2−3 times. A concentration−response curve was generated for each sample. A regression analysis was conducted with each curve. The LC_{50} ($\%C^1/_2$) values were calculated from the regression analysis and represents the sample concentration factor that induced a 50% reduction in cell density as compared to the concurrent negative controls.

CHO Cell Single Cell Gel Electrophoresis (SCGE) Assay. Single cell gel electrophoresis (SCGE, or Comet) assay quantitatively measures genomic DNA damage in individual nuclei induced by a test agent.[34−36] We employed microplate methodology;[35] the detailed procedure is presented in the SI. The SCGE metric for genomic DNA damage induced by the HIWATE samples was the %Tail DNA value which is the amount of DNA that migrated from the nucleus into the microgel.[37] Within each concentration factor range with >70% cell viability, a concentration−response curve was generated and regression analysis was used to fit the curve. The concentration factor inducing a 50% Tail DNA value was calculated from each concentration−response curve.

Statistical Analyses. For the cytotoxicity assay, a one-way analysis of variance (ANOVA) test was conducted to determine if the HIWATE sample induced a statistically significant level of cell death at a specific concentration factor. If a significant F value ($P \leq 0.05$) was obtained, a Holm-Sidak multiple comparison versus the control group analysis was performed to identify the lowest cytotoxic concentration factor. The power of the test statistic ($1-\beta$) was maintained as ≥0.8 at $\alpha = 0.05$.

For the SCGE assay, the %Tail DNA values are not normally distributed which limits the use of parametric statistics.[38] The mean %Tail DNA value for each microgel was calculated and these values were averaged among all of the microgels for each HIWATE sample concentration factor. Averaged mean values express a normal distribution according to the central limit theorem.[38] A one-way ANOVA test was conducted on these averaged %Tail DNA values.[39] If a significant F value of $P \leq 0.05$ was obtained, a Holm-Sidak multiple comparison versus the control group analysis was conducted (power ≥0.8; $\alpha = 0.05$).

The mammalian cell cytotoxicity and genotoxicity analyses were compared with the following analytical chemical metrics: (i) the numbers of DBPs identified in each HIWATE sample,

Table 3. CHO Cell Chronic Cytotoxicity Analyses of the HIWATE Samples

sample number	concentration factor range	lowest cytotoxic concentration factor[a]	LC_{50} value[b] (conc. factor ± SE)	r^{2c}	ANOVA test statistic[d]
HIWATE 1	0−150	60	102.7 ± 4.2	0.95	$F_{10, 37} = 58.4$; $P \leq 0.001$
HIWATE 2	0−150	70	107.8 ± 3.8	0.97	$F_{10, 37} = 59.2$; $P \leq 0.001$
HIWATE 3	0−300	50	79.1 ± 4.1	0.96	$F_{19, 76} = 130$; $P \leq 0.001$
HIWATE 4	0−300	22.5	107.5 ± 3.7	0.97	$F_{19, 76} = 134$; $P \leq 0.001$
HIWATE 5	0−1000	300	605.8 ± 4.3	0.98	$F_{9, 33} = 69.2$; $P \leq 0.001$
HIWATE 6	0−800	300	366.9 ± 4.1	0.99	$F_{10, 37} = 113$; $P \leq 0.001$
HIWATE 7	0−350	50	122.1 ± 3.4	0.99	$F_{11, 44} = 212$; $P \leq 0.001$
HIWATE 8	0−300	70	162.5 ± 4.8	0.98	$F_{10, 37} = 71.7$; $P \leq 0.001$
HIWATE 9	0−300	80	140.0 ± 4.9	0.98	$F_{10, 37} = 90.8$; $P \leq 0.001$
HIWATE 10	0−300	80	128.9 ± 4.6	0.97	$F_{11, 40} = 77.9$; $P \leq 0.001$
HIWATE 11	0−600	100	164.4 ± 5.2	0.98	$F_{10, 37} = 78.0$; $P \leq 0.001$

[a]Lowest cytotoxic concentration was the lowest concentration factor of the HIWATE sample in the concentration−response curve that induced a statistically significant reduction in cell density as compared to the concurrent negative controls. [b]The LC_{50} value is the fold concentration factor of the HIWATE sample, determined from a regression analysis of the data, that induced a cell density of 50% as compared to the concurrent negative controls. The LC_{50} error term was calculated as $\Sigma\bar{X}_{SE}$. [c]r^2 is the coefficient of determination for the regression analysis upon which the LC_{50} value was calculated. [d]The degrees of freedom for the between-groups and residual associated with the calculated *F*-test result and the resulting probability value.

(ii) the chromatographic peak areas for all DBPs in the entire sample, (iii) peak areas for specific classes of DBPs, (iv) the total concentration of 21 selected DBPs and, (v) concentrations of specific DBP classes within the group of 21 DBPs. A Pearson's Product Moment correlation test was conducted.

RESULTS AND DISCUSSION

Chemical Analyses. Over 90 DBPs were identified in the samples, including several haloacids (including 3- and 4-carbon acids and diacids), halophenols, haloamides, halonitromethanes, haloketones, haloaldehydes, and haloalkenes (Table S1, SI). Approximately 300 chromatographic peaks were observed in the original GC/MS chromatograms (Figure S1, SI; including DBPs and other compounds present in the raw waters prior to disinfection). With GC × GC-TOF-MS analyses, these peaks were resolved into >1000 peaks (Figure S2, SI). Several DBPs identified were not in mass spectral library databases and these identifications were made through the methods outlined previously and utilizing Massworks software. Several new DBPs were identified, including *cis*- and *trans*-2,3-dibromo-3-chloropropenoic acid, 3,3-dibromo-2-chloropropenoic acid, and several halophenols and haloalkenes. Twenty-one target DBPs, including four U.S-regulated THMs, nine HAAs, four haloacetonitriles (HANs), two haloketones (HKs), trichloroacetaldehyde (chloral hydrate), and trichloronitromethane (chloropicrin) were quantified (Table 2).

Substantial differences were observed in the DBPs from the different locations. As expected, drinking waters from coastal Spain (Barcelona and Valencia) had relatively high DBP levels with many brominated (and some iodinated) species due to higher levels of total organic carbon (TOC), bromide and iodide in their source waters (surface water), as well as the use of chlorine as a disinfectant. Drinking waters from coastal Spain averaged 90 and 33 μg/L for THM4 and HAA9, respectively (Table 2). In contrast, drinking water from Modena, Italy had fewer DBPs at much lower levels; these were primarily chlorine-containing species. The source water for Modena is a low-TOC groundwater that is treated with low chlorine dioxide doses (0.1 mg/L), which forms fewer DBPs as compared to other disinfectants.[3,40−42] None of the 21 target DBPs were detected in the drinking water from Modena, but a few were detected in the broad screen analyses due to lower detection limits. Drinking water from other locations (samples 4, 6, 8−11, Table 1) expressed intermediate DBP levels with a mix of chloro-bromo species probably due to lower levels of bromide and TOC in their source waters as compared to waters from coastal Spain (Table S1, SI and Table 2).

Of the N-DBPs,[43] haloacetonitriles, haloamides, and halonitromethanes were prevalent in drinking waters from coastal cities in Spain (samples 1−3, 7), which involved treatment with chlorine, alone or in combination with ozone or chlorine dioxide. Previous research demonstrated that ozonation increased the formation of halonitromethanes when used prior to chlorination or chloramination.[44−46] While chloramination increases the formation of some N-DBPs,[5] none of the cities in this study employed chloramines.

CHO Cell Chronic Cytotoxicity. CHO cell chronic cytotoxicity analyses of each HIWATE sample are summarized in Table 3. The concentration factor is the fold concentration of the isolated organic material as compared to the original water. The lowest concentration factor of each sample which induced a statistically significant reduction in cell density as compared to its concurrent negative control was determined by an ANOVA test statistic. The data from replicated experiments were averaged and plotted (Figure 1A, Figures S3−S13, SI); regression analyses were used to calculate the LC_{50} ($\%C^1/_2$) value for each sample. Based on the LC_{50} values, the descending rank order of chronic cytotoxicity was, sample 3 > sample 1 > sample 2 ≈ sample 4 > sample 7 > sample 10 > sample 9 > sample 8 ≈ sample 11 > sample 6 > sample 5. Samples from Barcelona, Spain were ranked as the 3 most cytotoxic. We calculated the cytotoxicity index value ($LC_{50}^{-1} \times 1000$) for each HIWATE sample (Figure 1B, Table S2, SI).

CHO Cell Acute Genotoxicity. CHO cell acute genotoxicity analyses of each HIWATE sample are summarized in Table 4. The lowest genotoxic concentration factor was that which induced a statistically significant amount of genomic DNA damage as compared to the concurrent negative control. Figure 2A (Figures S14−S24, SI) illustrates the concentration−response curves. Based on 50% Tail DNA values, the descending rank order of genotoxicity was, sample 10 > sample 4 > sample 7 > sample 1 ≈ sample 2 > sample 3 > sample 9 > sample 11 > sample 6 > sample 8 ≫ sample 5. We calculated

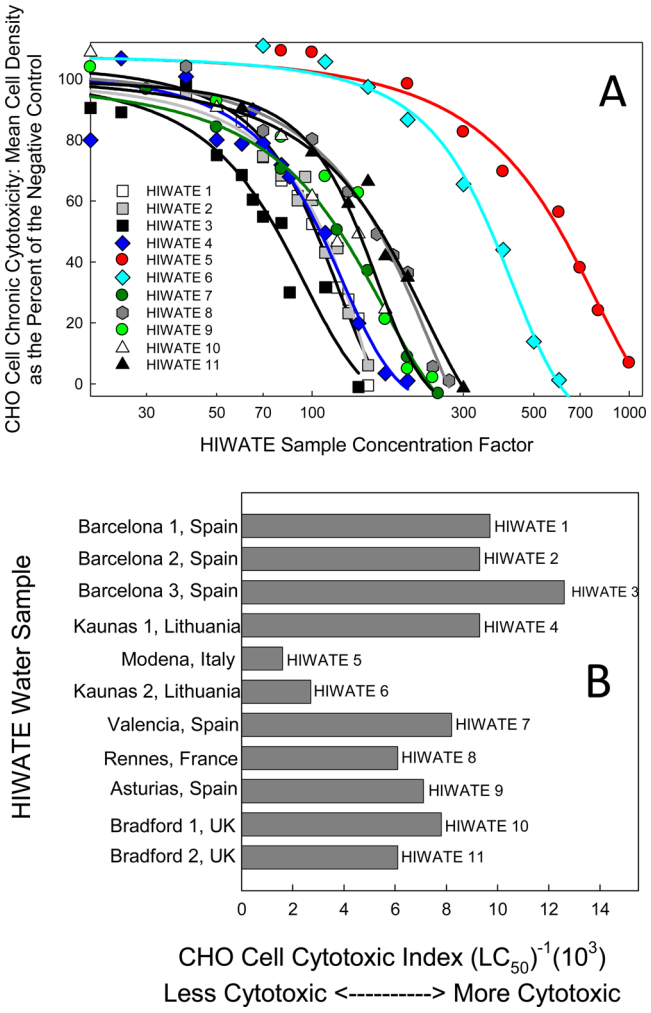


Figure 1. (A) Log-linear plot of the concentration–response curves of 11 HIWATE samples illustrating CHO cell chronic (72 h) cytotoxicity. (B) The distributions of the CHO cell cytotoxic index values for each HIWATE sample.

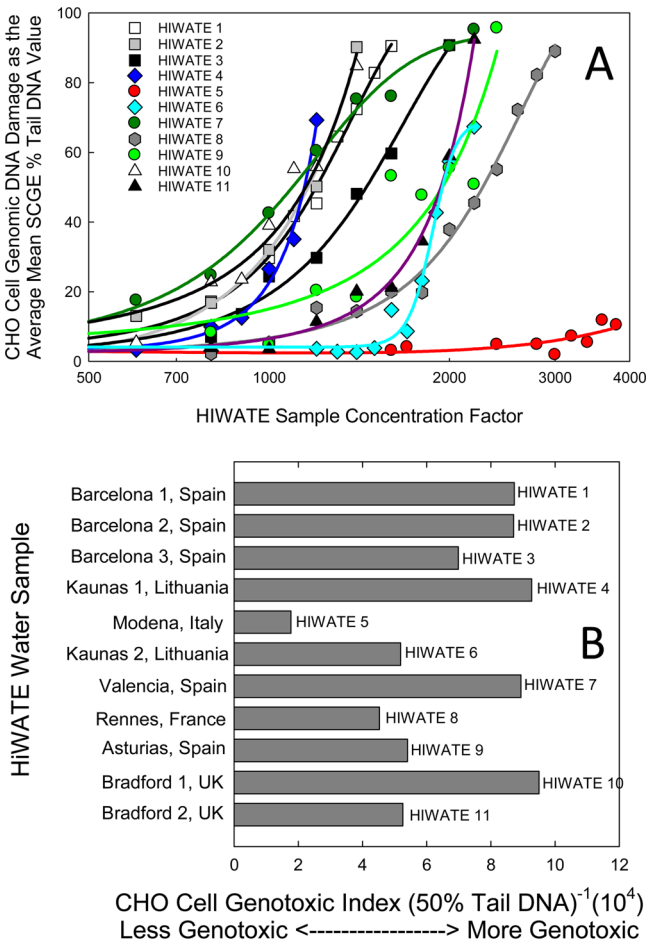


Figure 2. (A) Log–linear plot of the concentration–response curves of 11 HIWATE samples illustrating CHO cell acute (4 h) genotoxicity. (B) The distributions of the CHO cell genotoxic index values for each HIWATE sample.

Table 4. CHO Cell SCGE Genotoxicity Analyses of the HIWATE Samples

sample number	concentration factor range	lowest genotoxic concentration factor[a]	50% tail DNA Value[b] (Conc. Factor ± SE)	r^{2c}	ANOVA test statistic[d]
HIWATE 1	0–1700	1000	1146 ± 9.1	0.99	$F_{9,33} = 10.5$; $P \leq 0.001$
HIWATE 2	0–2000	1000	1148 ± 3.8	0.99	$F_{10,37} = 43.6$; $P \leq 0.001$
HIWATE 3	0–2000	900	1430 ± 2.0	0.99	$F_{8,37} = 133$; $P \leq 0.001$
HIWATE 4	0–1600	1000	1079 ± 3.3	0.99	$F_{10,37} = 78.0$; $P \leq 0.001$
HIWATE 5	0–3800	3600	5659 ± 1.4	0.74	$F_{10,38} = 3.14$; $P \leq 0.005$
HIWATE 6	0–2200	1600	1925 ± 2.0	0.90	$F_{11,39} = 20.0$; $P \leq 0.001$
HIWATE 7	0–2200	600	1119 ± 2.9	0.95	$F_{9,26} = 76.3$; $P \leq 0.001$
HIWATE 8	0–3000	2000	2206 ± 5.6	0.98	$F_{13,50} = 9.29$; $P \leq 0.001$
HIWATE 9	0–2400	1600	1847 ± 6.9	0.89	$F_{10,42} = 6.87$; $P \leq 0.001$
HIWATE 10	0–1600	400	1052 ± 6.2	0.98	$F_{8,50} = 7.18$; $P \leq 0.001$
HIWATE 11	0–2400	1600	1901 ± 3.7	0.99	$F_{10,33} = 15.9$; $P \leq 0.001$

[a]The lowest genotoxic concentration was the lowest concentration factor of the HIWATE sample in the concentration–response curve that induced a statistically significant amount of genomic DNA damage as compared to the negative control. [b]The SCGE 50% Tail DNA value is the HIWATE sample concentration factor determined from a regression analysis of the data that was calculated to induce a 50% SCGE Tail DNA value. The 50% SCGE Tail DNA value error term was calculated as $\sum \bar{X}_{SE}$. [c]r^2 is the coefficient of determination for the regression analysis upon which the SCGE % Tail DNA value was calculated. [d]The degrees of freedom for the between-groups and residual associated with the calculated *F*-test result and the resulting probability value.

the genotoxic index value as 50% Tail DNA^{-1} × 10^4 for each sample (Figure 2B; Table S2, SI).

Correlation of Toxicology, Chemistry, and Epidemiology. To investigate correlations between DBP occurrence and DBP classes with mammalian cell toxicity, we applied a Pearson's Product Moment statistical test.[38] The cytotoxic potency index values statistically significantly correlated with the number of identified DBPs (r = 0.76; $P \leq 0.005$, Table 5) and the level of 21 target DBPs (r = 0.77; $P \leq 0.006$, Table 2). The genotoxic potency index values were not correlated with either of these metrics or with any DBP chemical class (Table 2 and Table 5). Interestingly, the cytotoxicity and genotoxicity indices indicated a good correlation (r = 0.74; $P \leq 0.009$). The cytotoxic potency index showed a good correlation with the U.S-regulated DBPs (r = 0.78; $P \leq 0.006$) and unregulated DBPs (r = 0.60; $P \leq 0.05$; Table 2).

Cytotoxicity was significantly correlated with the relative concentrations of the following DBP classes: THMs (r = 0.74; $P \leq 0.01$), haloacids (r = 0.75; $P \leq 0.008$), other monoacids (r = 0.68; $P \leq 0.021$), halodiacids (r = 0.80; $P \leq 0.003$), haloamides (r = 0.68; $P \leq 0.021$), haloaromatics (r = 0.64; $P \leq 0.035$), brominated (r = 0.68; $P \leq 0.022$), chlorinated (r = 0.78; $P \leq 0.005$), and iodinated (r = 0.82; $P \leq 0.002$) DBPs (Table 6). There were no statistically significant correlations with genotoxicity and the above DBP classes, although there were associations or trends in relationships between genotoxicity and the relative concentrations of haloacids (r = 0.54; P = 0.088), haloaromatics (r = 0.52; P = 0.103), chlorinated (r = 0.56; P = 0.073) and iodinated (r = 0.53; P = 0.093) DBPs (Table 6). It should be noted that some highly polar components might have been missed by GC/MS and this may explain, in part, the reduced correlation seen with the genotoxicity data and analytical chemistry of the water samples. Recently several papers have been published on novel methods to detect polar iodinated/brominated DBPs.[47−49]

Epidemiology results on water DBPs and birth outcomes from Lithuania, Spain and France were recently published[50−52] and the present analysis included water samples from the cities covered in those studies. An expanded discussion of the associations among the epidemiology studies and this work is presented in the Supporting Information (Table S3). It should be noted, however, that the drinking water samples for the epidemiologic analyses and the current analytical chemical and toxicological evaluations were not collected at the same time. Existing epidemiological studies on birth outcomes including those in the HIWATE project, have evaluated a limited number of DBPs (usually only THMs) through environmental analyses of drinking water or, in the case of the French study,[52] through an evaluation of biomarkers of haloacetic acid metabolites in urine. The analyses of water toxicity presented in this paper were limited in number due to their complexity, but they provide an overall evaluation of differences of toxicity in different geographic areas. It is the first time that this evaluation was done to specifically correspond with areas examined in epidemiological studies. Expanding the chemical and toxicological characterization of water samples may enhance the resolving power of epidemiological investigations and the evaluation of dose−response relationships. In addition, the relationship between the analytical chemistry, quantitative in vitro toxicology, and the epidemiology may provide additional mechanistic evidence on potential health effects of water DBPs.

This paper focused on the relationship of the occurrence and concentration of DBPs with mammalian cell toxicity. The range of the number of DBPs identified and their levels reflect the diverse collection sites, different disinfection processes, and the

Table 5. Description of Each HIWATE Sample, DBPs Identified and Gross Correlation with the Rank Order of CHO Cell Cytotoxicity and Genotoxicity[a]

sample number	number of identified DBPs	rank order of number of identified DBPs	rank order of cytotoxic potency index[b]	rank order of genotoxic potency index[c]
HIWATE 1	86	1	2	4
HIWATE 2	76	5	3	5
HIWATE 3	85	2	1	6
HIWATE 4	41	7	3	2
HIWATE 5	13	11	11	11
HIWATE 6	18	10	10	9
HIWATE 7	83	3	5	3
HIWATE 8	77	4	8	10
HIWATE 9	45	6	7	7
HIWATE 10	41	7	6	1
HIWATE 11	40	9	8	8

[a]Correlation with the rank order of CHO cell cytotoxicity: r = 0.78 ($P \leq 0.005$). Correlation with the rank order of CHO cell genotoxicity: r = 0.52 (P = 0.105). Rank order where 1 is the highest response and 11 is the lowest response. [b]The CHO cell cytotoxic potency index value is in arbitrary units and the value corresponds to ($LC_{50}^{-1} \times 10^3$) for each HIWATE sample. [c]The CHO cell genotoxic potency index value is the reciprocal HIWATE sample concentration factor that was calculated to induce a 50% SCGE tail DNA value ×10^4 and is presented in arbitrary units.

Table 6. Pearson Product Moment Correlation Analyses of the Relative Concentrations of Each DBP Group Versus CHO Cell Chronic Cytotoxicity or Acute Genotoxicity

relative concentration of DBP class[a]	cytotoxic potency index value[b] ($LC_{50}^{-1} \times 10^3$)	genotoxic potency index value[c] (50% tail DNA^{-1} × 10^4)
THMs	r = 0.74	r = 0.45
	$P \leq 0.010$	P = 0.164
haloacids	r = 0.75	r = 0.54
	$P \leq 0.008$	P = 0.088
other monoacids	r = 0.68	r = 0.42
	$P \leq 0.021$	P = 0.201
halodiacids	r = 0.80	r = 0.40
	$P \leq 0.003$	P = 0.221
haloamides	r = 0.68	r = 0.45
	$P \leq 0.021$	P = 0.170
haloaromatics	r = 0.64	r = 0.52
	$P \leq 0.035$	P = 0.103
brominated DBPs	r = 0.68	r = 0.46
	$P \leq 0.022$	P = 0.154
chlorinated DBPs	r = 0.78	r = 0.56
	$P \leq 0.005$	P = 0.073
iodinated DBPs	r = 0.82	r = 0.53
	$P \leq 0.002$	P = 0.093

[a]Relative concentration is defined as the integrated area for each chromatographic peak summed for each DBP chemical class. [b]The CHO cell cytotoxic potency index value corresponds to ($LC_{50}^{-1} \times 10^3$) for each HIWATE sample. [c]The CHO cell genotoxic potency index value is the reciprocal HIWATE sample concentration factor that was calculated to induce a 50% SCGE tail DNA value ×10^4.

different characteristics of the source waters. CHO cytotoxicity was well correlated with the numbers of DBPs identified and the levels of DBP chemical classes. Although there was a clear difference in genotoxic responses, these data did not correlate well with chemical analyses of the HIWATE samples. Thus, the agents responsible for the genomic DNA damage observed in the HIWATE samples may be due to unresolved associations of combinations of identified DBPs, unknown emerging DBPs that were not identified, or other toxic water contaminants.

We are continuing to compare the epidemiology with the in vitro toxicity and analytical chemistry analyses. Future study will investigate the possible association between chronic cytotoxicity, acute genotoxicity, multivariate comparisons of identified DBPs and epidemiology across the entire HIWATE program. We plan to compare other in vitro and molecular toxicity metrics and rates of adverse pregnancy measurements. Finally, we propose to determine the contribution of source water, and disinfection chemistry to the observed toxicity and epidemiology results and develop solutions to protect the public health and the environment.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional information is available in the Supporting Information on experimental methods, including additional figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 217-333-3614; e-mail: mplewa@illinois.edu.

Notes

The authors declare no competing financial interest.
[$]Student Services Authority.

■ ACKNOWLEDGMENTS

The work was supported by HIWATE (www.hiwate.org) project, funded under the EU (Contract No. Food-CT-2006-036224). We appreciate support by the Center of Advanced Materials for the Purification of Water with Systems (WaterCAMPWS), a National Science Foundation Science and Technology Center, under Award CTS-0120978. C.J. was supported by a NIEHS Predoctoral Fellowship under Grant No. T32 ES007326. We thank Adonina Tardon and Ferran Ballester for their assistance with obtaining the Spanish water samples. We thank Jessie Merriam and Matthew Bloodgood for assistance with analyzing GC/MS data, Mike Riley and Leco Corporation for their loan of the Pegasus software, Yongdong Wang and Ming Gu of Cerno Bioscience for assistance with Massworks, and Sean Crook of Waters Corporation for assistance with high resolution analyses. This paper has been reviewed in accordance with the U.S. EPA's peer and administrative review policies and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. EPA.

■ REFERENCES

(1) Cutler, D.; Miller, G. The role of public health improvements in health advances: The twentieth-century United States. *Demography* **2005**, *42* (1), 1−22.

(2) Reckhow, D. A.; Linden, K. G.; Kim, J.; Shemer, H.; Makdissy, G. Effect of UV treatment on DBP formation. *J. Am.Water Works Assoc.* **2010**, *102* (6), 100−113.

(3) Zhang, X.; Echigo, S.; Minear, R. A.; Plewa, M. J., Characterization and comparison of disinfection by-products of four major disinfectants. In *Natural Organic Matter and Disinfection By-Products: Characterization and Control in Drinking Water*; Barrett, S. E.; Krasner, S. W.; Amy, G. L., Eds.; American Chemical Society: Washington, DC, 2000; pp 299−314.

(4) Hua, G. H.; Reckhow, D. A. Comparison of disinfection byproduct formation from chlorine and alternative disinfectants. *Water Res.* **2007**, *41* (8), 1667−1678.

(5) Richardson, S. D.; Plewa, M. J.; Wagner, E. D.; Schoeny, R.; DeMarini, D. M. Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: A review and roadmap for research. *Mutat. Res.* **2007**, *636*, 178−242.

(6) Richardson, S. D. Disinfection by-products: Formation and occurrence in drinking water. In *Encyclopedia of Environmental Health*; Nriagu, J. O., Ed.; Elsevier: Burlington, 2011; Vol. *1*, pp 110−136.

(7) Krasner, S. W. The formation and control of emerging disinfection by-products of health concern. *Philos. Trans. R. Soc., A* **2009**, *367* (1904), 4077−4095.

(8) Ye, B.; Wang, W.; Yang, L.; Wei, J.; Xueli, E. Formation and modeling of disinfection by-products in drinking water of six cities in China. *J. Environ. Monit.* **2011**, *13* (5), 1271−1275.

(9) Ueno, H.; Moto, T.; Sayato, Y.; Nakamuro, K. Disinfection by-products in the chlorination of organic nitrogen compounds: By-products from kynurenine. *Chemosphere* **1996**, *33* (8), 1425−1433.

(10) U. S. Environmental Protection Agency, National primary drinking water regulations: Stage 2 disinfectants and disinfection byproducts rule. *Fed. Regist.* **2006**, *71*, (2), 387-493.

(11) Villanueva, C. M.; Cantor, K. P.; Cordier, S.; Jaakkola, J. J.; King, W. D.; Lynch, C. F.; Porru, S.; Kogevinas, M. Disinfection byproducts and bladder cancer: A pooled analysis. *Epidemiology* **2004**, *15* (3), 357−367.

(12) Villanueva, C. M.; Cantor, K. P.; Grimalt, J. O.; Malats, N.; Silverman, D.; Tardon, A.; Garcia-Closas, R.; Serra, C.; Carrato, A.; Castano-Vinyals, G.; Marcos, R.; Rothman, N.; Real, F. X.; Dosemeci, M.; Kogevinas, M. Bladder cancer and exposure to water disinfection by-products through ingestion, bathing, showering, and swimming in pools. *Am. J. Epidemiol.* **2007**, *165* (2), 148−156.

(13) Costet, N.; Villanueva, C. M.; Jaakkola, J. J.; Kogevinas, M.; Cantor, K. P.; King, W. D.; Lynch, C. F.; Nieuwenhuijsen, M. J.; Cordier, S. Water disinfection by-products and bladder cancer: Is there a European specificity? A pooled and meta-analysis of European case-control studies. *Occup. Environ. Med.* **2011**, *68* (5), 379−385.

(14) Bull, R. J.; Reckhow, D. A.; Li, X. F.; Humpage, A. R.; Joll, C.; Hrudey, S. E. Potential carcinogenic hazards of non-regulated disinfection by-products: Haloquinones, halo-cyclopentene and cyclohexene derivatives, N-halamines, halonitriles, and heterocyclic amines. *Toxicology* **2011**, *286* (1−3), 1−19.

(15) Rahman, M. B.; Driscoll, T.; Cowie, C.; Armstrong, B. K. Disinfection by-products in drinking water and colorectal cancer: A meta-analysis. *Int. J. Epidemiol.* **2010**, *39* (3), 733−745.

(16) Yang, C. Y.; Xiao, Z. P.; Ho, S. C.; Wu, T. N.; Tsai, S. S. Association between trihalomethane concentrations in drinking water and adverse pregnancy outcome in Taiwan. *Environ. Res.* **2007**, *104* (3), 390−395.

(17) Wright, J. M.; Schwartz, J.; Dockery, D. W. Effect of trihalomethane exposure on fetal development. *Occup. Environ. Med.* **2003**, *60* (3), 173−180.

(18) Grellier, J.; Bennett, J.; Patelarou, E.; Smith, R. B.; Toledano, M. B.; Rushton, L.; Briggs, D. J.; Nieuwenhuijsen, M. J. Exposure to disinfection by-products, fetal growth, and prematurity: A systematic review and meta-analysis. *Epidemiology* **2010**, *21* (3), 300−313.

(19) Hinckley, A. F.; Bachand, A. M.; Reif, J. S. Late pregnancy exposures to disinfection by-products and growth-related birth outcomes. *Environ. Health Perspect.* **2005**, *113* (12), 1808−1813.

(20) Hunter, E. S., III; Rogers, E. H.; Schmid, J. E.; Richard, A. Comparative effects of haloacetic acids in whole embryo culture. *Teratology* **1996**, *54* (2), 57−64.
(21) Narotsky, M. G.; Best, D. S.; McDonald, A.; Godin, E. A.; Hunter, E. S., 3rd; Simmons, J. E. Pregnancy loss and eye malformations in offspring of F344 rats following gestational exposure to mixtures of regulated trihalomethanes and haloacetic acids. *Reprod. Toxicol.* **2011**, *31* (1), 59−65.
(22) Nieuwenhuijsen, M. J.; Smith, R.; Golfinopoulos, S.; Best, N.; Bennett, J.; Aggazzotti, G.; Righi, E.; Fantuzzi, G.; Bucchini, L.; Cordier, S.; Villanueva, C. M.; Moreno, V.; La Vecchia, C.; Bosetti, C.; Vartiainen, T.; Rautiu, R.; Toledano, M.; Iszatt, N.; Grazuleviciene, R.; Kogevinas, M. Health impacts of long-term exposure to disinfection by-products in drinking water in Europe: HIWATE. *J. Water Health* **2009**, *7* (2), 185−207.
(23) Richardson, S. D. *XAD Resin Extraction of Disinfection By-Products Drinking Water: SOP - RSB-003.1- Revision No. 1*; Environmental Protection Agency: Athens, GA, 2011.
(24) Pressman, J. G.; Richardson, S. D.; Speth, T. F.; Miltner, R. J.; Narotsky, M. G.; Hunter, E. S., 3rd; Rice, G. E.; Teuschler, L. K.; McDonald, A.; Parvez, S.; Krasner, S. W.; Weinberg, H. S.; McKague, A. B.; Parrett, C. J.; Bodin, N.; Chinn, R.; Lee, C. F.; Simmons, J. E. Concentration, chlorination, and chemical analysis of drinking water for disinfection byproduct mixtures health effects research: U.S. EPA's Four Lab Study. *Environ. Sci. Technol.* **2010**, *44* (19), 7184−7192.
(25) Chen, P. H.; Richardson, S. D.; Krasner, S. W.; Majetich, G.; Glish, G. L. Hydrogen abstraction and decomposition of bromopicrin and other trihalogenated disinfection byproducts by GC/MS. *Environ. Sci. Technol.* **2002**, *36* (15), 3362−3371.
(26) Wang, Y.; Liao, X. H.; Gu, Y.; Chen, R.; Zeng, J. Preliminary study on oxygen content monitoring for port wine stains during PDT using diffuse reflection spectra. *Guang pu xue yu guang pu fen xi = Guang pu* **2010**, *30* (12), 3363−3366.
(27) EPA Method 551.1. *Determination of Chlorination Disinfection Byproducts, Chlorinated Solvents, And Halogenated Pesticides/Herbicides in Drinking Water by Liquid-Liquid Extraction and Gas Chromatography with Electron-Capture Detection, Revision 1.0*; National Exposure Research Laboratory, USEPA Office of Research and Development: Cincinnati, OH., 1995.
(28) EPA Method 552.3. *Determination of Haloacetic Acids and Dalapon in Drinking Water by Liquid-Liquid Microextraction, Derivatization and Gas Chromatography with Electron Capture Detection, Revision 1.0*; U.S. EPA Office of Research and Development: Cincinnati OH: 2003.
(29) Wagner, E. D.; Rayburn, A. L.; Anderson, D.; Plewa, M. J. Analysis of mutagens with single cell gel electrophoresis, flow cytometry, and forward mutation assays in an isolated clone of Chinese hamster ovary cells. *Environ. Mol. Mutagen.* **1998**, *32* (4), 360−368.
(30) Tindall, K. R.; Stankowski, L. F., Jr.; Machanoff, R.; Hsie, A. W. Detection of deletion mutations in pSV2gpt-transformed cells. *Mol. Cell. Biol.* **1984**, *4* (7), 1411−1415.
(31) Hsie, A. W.; Brimer, P. A.; Mitchell, T. J.; Gosslee, D. G. The dose-response relationship for ultraviolet-light-induced mutations at the hypoxanthine-guanine phosphoribosyltransferase locus in Chinese hamster ovary cells. *Somatic Cell Genet.* **1975**, *1* (4), 383−389.
(32) Plewa, M. J.; Kargalioglu, Y.; Vankerk, D.; Minear, R. A.; Wagner, E. D. Mammalian cell cytotoxicity and genotoxicity analysis of drinking water disinfection by-products. *Environ. Mol. Mutagen.* **2002**, *40* (2), 134−142.
(33) Plewa, M. J.; Wagner, E. D., *Mammalian Cell Cytotoxicity and Genotoxicity of Disinfection By-Products*; Water Research Foundation: Denver, CO, 2009; p 134.
(34) Rundell, M. S.; Wagner, E. D.; Plewa, M. J. The comet assay: Genotoxic damage or nuclear fragmentation? *Environ. Mol. Mutagen.* **2003**, *42* (2), 61−67.
(35) Wagner, E. D.; Plewa, M. J., Microplate-based comet assay. In *The Comet Assay in Toxicology*; Dhawan, A.; Anderson, D., Eds.; Royal Society of Chemistry: London, 2009; pp 79−97.
(36) Tice, R. R.; Agurell, E.; Anderson, D.; Burlinson, B.; Hartmann, A.; Kobayashi, H.; Miyamae, Y.; Rojas, E.; Ryu, J. C.; Sasaki, Y. F. Single cell gel/comet assay: Guidelines for in vitro and in vivo genetic toxicology testing. *Environ. Mol. Mutagen.* **2000**, *35* (3), 206−221.
(37) Kumaravel, T. S.; Jha, A. N. Reliable Comet assay measurements for detecting DNA damage induced by ionising radiation and chemicals. *Mutat. Res.* **2006**, *605* (1−2), 7−16.
(38) Box, G. E. P.; Hunter, W. G.; Hunter, J. S. *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*; Wiley & Sons Inc.: New York, NY, 1978.
(39) Lovell, D. P.; Omori, T. Statistical issues in the use of the comet assay. *Mutagenesis* **2008**, *23* (3), 171−182.
(40) Richardson, S. D.; Thruston, A. D., Jr.; Rav-Acha, C.; Groisman, L.; Popilevsky, I.; Juraev, O.; Glezer, V.; McKague, A. B.; Plewa, M. J.; Wagner, E. D. Tribromopyrrole, brominated acids, and other disinfection byproducts produced by disinfection of drinking water rich in bromide. *Environ. Sci. Technol.* **2003**, *37* (17), 3782−3793.
(41) Weinberg, H. S.; Krasner, S. W.; Richardson, S. D.; Thruston, A. D., Jr. *The Occurrence of Disinfection By-Products (DBPs) of Health Concern in Drinking Water: Results of a Nationwide DBP Occurrence Study*, EPA/600/R02/068; U.S. Environmental Protection Agency National Exposure Research Laboratory: Athens, GA, 2002; pp 1−460.
(42) Richardson, S. D.; Rav-Acha, C.; Simpson, G. C., Chlorine dioxide chemistry, reactions, and disinfection by-products. In *State of the Science of Chlorine Dioxide in Drinking Water*; Gates, D.; Ziglio, G.; Ozekin, K., Eds.; Water Research Foundation and Fondazione AMGA: Denver, CO., 2009; pp 21−49.
(43) Plewa, M. J.; Wagner, E. D.; Muellner, M. G.; Hsu, K. M.; Richardson, S. D., Comparative mammalian cell toxicity of N-DBPs and C-DBPs. In *Occurrence, Formation, Health Effects and Control of Disinfection By-Products in Drinking Water*; Karanfil, T., Krasner, S. W., Westerhoff, P., Xie, Y., Eds.; American Chemical Society: Washington, DC, 2008; Vol. *995*, pp 36−50.
(44) Choi, J.; Richardson, S. D. Formation of halonitromethanes in drinking water. In *Proceedings of the Water Quality Technology Conference*; American Water Works Association: Denver, CO, 2004.
(45) Hu, J.; Song, H.; Addison, J. W.; Karanfil, T. Halonitromethane formation potentials in drinking waters. *Water Res.* **2010**, *44* (1), 105−114.
(46) Krasner, S. W.; Weinberg, H. S.; Richardson, S. D.; Pastor, S. J.; Chinn, R.; Sclimenti, M. J.; Onstad, G. D.; Thruston, A. D., Jr. The occurrence of a new generation of disinfection by-products. *Environ. Sci. Technol.* **2006**, *40* (23), 7175−7185.
(47) Zhai, H.; Zhang, X. Formation and decomposition of new and unknown polar brominated disinfection byproducts during chlorination. *Environ. Sci. Technol.* **2011**, *45* (6), 2194−2201.
(48) Ding, G.; Zhang, X. A picture of polar iodinated disinfection byproducts in drinking water by (UPLC/)ESI-tqMS. *Environ. Sci. Technol.* **2009**, *43* (24), 9287−9293.
(49) Zhang, X.; Talley, J. W.; Boggess, B.; Ding, G.; Birdsell, D. Fast selective detection of polar brominated disinfection byproducts in drinking water using precursor ion scans. *Environ. Sci. Technol.* **2008**, *42* (17), 6598−6603.
(50) Grazuleviciene, R.; Nieuwenhuijsen, M. J.; Vencloviene, J.; Kostopoulou-Karadanelli, M.; Krasner, S. W.; Danileviciute, A.; Balcius, G.; Kapustinskiene, V. Individual exposures to drinking water trihalomethanes, low birth weight and small for gestational age risk: A prospective Kaunas cohort study. *Environ. Health* **2011**, *10*, 32.
(51) Villanueva, C. M.; Gracia-Lavedan, E.; Ibarluzea, J.; Santa Marina, L.; Ballester, F.; Llop, S.; Tardon, A.; Fernandez, M. F.; Freire, C.; Goni, F.; Basagana, X.; Kogevinas, M.; Grimalt, J. O.; Sunyer, J.; Project, I. Exposure to trihalomethanes through different water uses and birth weight, small for gestational age, and preterm delivery in Spain. *Environ. Health Perspect.* **2011**, *119* (12), 1824−1830.
(52) Costet, N.; Garlantezec, R.; Monfort, C.; Rouget, F.; Gagniere, B.; Chevrier, C.; Cordier, S. Environmental and urinary markers of prenatal exposure to drinking water disinfection by-products, fetal growth, and duration of gestation in the PELAGIE birth cohort

(Brittany, France, 2002−2006). *Am. J. Epidemiol.* **2012**, *175* (4), 263−275.

APPENDIX C

Repeat questionnaire study invitation package (participant information sheet, invitation letter, recruitment form, water questionnaire)

Imperial College London

BORNINBRADFORD For a Healthy Future

Participants' Information Sheet:
A study about tap water and how babies grow and develop before birth

This information sheet gives details of a study which has been set up within the Born in Bradford project and which you are invited to join. Please take some time to read the following information carefully.

- **What is this study about?**

This study is about the disinfection of our tap water with chlorine, and whether by-products created by this disinfection process might affect babies' growth and development before they are born. Disinfection of our tap water is essential to kill germs and prevent us from getting sick from waterborne diseases, but our work hopes to learn more about possible health effects of disinfection by-products.

In the BiB questionnaire you completed a few weeks ago, you may recall that you were asked questions about how much water you drink, and how often you shower and bathe. Those questions were part of our research into this subject.

This study looks at your water use over the third trimester of pregnancy (i.e. 28th week of pregnancy until delivery), in order to investigate the possible effects of disinfection by-products in tap water on babies' growth during this important stage of your pregnancy.

- **Why is it important to do this research?**

We are all are exposed to tap water, through drinking, eating and washing, therefore we may all be exposed to the disinfection by-products that exist in water. With so many people potentially exposed to these by-products, it is very important that we investigate any possible health effects that they may have.

It is also really important to investigate babies' growth and development, because we know that low birth-weight is associated with poorer health later in life. If we can learn what factors contribute to low birth-weight then we can improve a baby's chance of a healthy start to life.

- **What advantages are there to taking part?**

You get to be part of a very select group of mothers in the BiB study who will be studied in greater detail. The main advantage of taking part is that you will be helping us to understand more about how the water we drink and use might affect babies during pregnancy. You may also find that you enjoy taking part in research that could help mothers and babies in the future.

- **What will I have to do if I decide to take part?**

It should take you about 10 minutes, in your own time, to fill out the questionnaire and recruitment form, seal them in the prepaid and pre-labelled return envelope we have sent you, and put it in the post to us at your nearest post box or post office. We ask you to do this once between your 30th and 33rd week of pregnancy, and again once between your 36th and 39th week of pregnancy.

PLEASE TURN OVER ➡

Frequently Asked Questions

- **Why do you need to know if I am currently working, and where I work?**

We need to account for which part of town the drinking water you use at work comes from. If you are working in a different part of town during the day from where you live, we need to know in order to correctly estimate the tap water you are drinking.

- **Why do I need to fill out two questionnaires?**

We are looking at how your water use changes over your last trimester of pregnancy. The BiB questionnaire you have completed collected information about your water use during the first 2 trimesters of pregnancy. We would now like to repeat this questionnaire both earlier (between weeks 30-33 of pregnancy) and later on (between weeks 36-39) in your third trimester of pregnancy.

- **Should I stop drinking tap water – is it unsafe for my baby?**

You should carry on doing things as you usually do. We do not yet know if the disinfection by-products in water affect babies' growth *in utero*; we are investigating many potential factors that may contribute to low birth-weight, disinfection by-products in tap water are just one of those factors.

- **What if I change my mind and want to withdraw from the study?**

You are free to change your mind and withdraw from the study at any time. If you decide to withdraw, you should let Susan Edwards know either by telephone or e-mail. Whatever you decide your medical care will not be affected in any way.

- **What will happen to the information you collect?**

Information will be stored for use by researchers from the Born in Bradford project and from Imperial College London. All the information which is collected about you during the course of the research will be kept strictly confidential. Any information about you will have your name and address removed so that you cannot be recognised from it. All the results of the study will be presented on a group basis, no individuals will be identified.

- **How do you ensure confidentiality?**

All the information we collect about you will be stored in strict confidence, as is required by law in the Data Protection and Human Tissue Acts. A personal ID number will be the only way the information can be linked to you.

- **How can I get to know the findings of the project?**

The general findings of the project will be published in scientific journals.

- **Who is organising and funding the research?**

The research is organised by the Born in Bradford study and Imperial College London. It is funded by the MRC (Medical Research Council) and the HiWATE project (Health Impacts of Long-Term Exposure to Disinfection By-Products in Drinking Water), which is funded by the 6th Framework Programme of the European Union.

- **What if I have any questions or problems?**

You can contact the study researcher, Susan Edwards, if you want to ask any questions about the study. If you have any questions about your general health or pregnancy, you should contact your doctor, midwife or health visitor.

Susan Edwards
Imperial College London, Department of Epidemiology and Biostatistics
St Mary's Campus, Norfolk Place, Paddington, London, W2 1PG, UK
☎ Tel: 020 7594 3285 ☎ Mob: 0783 326 1680 ✉ E-mail: s.edwards10@imperial.ac.uk

Thank you for reading this supplementary information!

Imperial College London

BORNINBRADFORD For a Healthy Future

Name
Address

Bradford, Date

Dear,

The Born in Bradford team, in collaboration with researchers at Imperial College in London, invite you to join a small study which has been set up within the Born in Bradford project.

Participation should take less than 10 minutes of your time, and can be achieved within the comfort of your home. Your participation is completely voluntary.

We are studying whether the water we drink in Bradford affects the health of our babies at birth. We are particularly interested in learning more about **how much water** women drink, how long they spend showering/bathing/swimming when they are pregnant, and if this **changes** over the course of pregnancy.

Please help us in this important research by completing and returning to us the following two items **within the next few days**:

1. the enclosed **yellow questionnaire** on your water consumption and usage between your 30th and 33rd week of pregnancy, and
2. the enclosed **green recruitment** form.

TURN OVER PLEASE ➡

Included in this invitation pack is a Freepost return envelope (already addressed with postage already paid for so you do NOT need a stamp). After completing the yellow and green forms, please seal them in the envelope provided and put it in the post to us at your nearest post box or post office. We guarantee that your information will be kept absolutely confidential.

In a few weeks time, we will send you another identical questionnaire asking you to complete the exact same exercise. This will provide us with information on your water use later on in your pregnancy, between your 36th and 39th week of pregnancy. Please contact us if you do not wish to receive this repeat questionnaire.

We have enclosed a ***Participants' Information Sheet***, and list of ***Frequently Asked Questions***, to explain the study in more detail and answer your questions. Please do not hesitate to contact us with any further questions you may have.

We hope that you find this information interesting, and that you will agree to help us conduct this important research about the environmental health of our families.

We thank you tremendously for your time in helping us better understand water use patterns during pregnancy.

Yours Sincerely,

Susan Edwards & Research Team

Imperial College London
Department of Epidemiology and Biostatistics
St Mary's Campus, Norfolk Place, Paddington, London, W2 1PG, UK
☎ Tel: 020 7594 3285
☎ Mob: 0783 326 1680
✉E-mail: s.edwards10@imperial.ac.uk

Born in Bradford
Bradford Institute for Health Research
Bradford Royal Infirmary
Duckworth Lane, Bradford, BD9 6RJ, UK

Imperial College London

BORNINBRADFORD For a Healthy Future

Recruitment Form (weeks 30-33)

Title of Project: BiB Water Use Study

Name of Researcher: Susan Edwards

BiB Study ID:

Repeat Questionnaire Study ID:

Participant's information	HALEY SMITH 8 ASPREY DRIVE ALLERTON BRADFORD BD15 7TX
Today's Date	*Day* *Month* *Year* ☐☐ ☐☐ ☐☐☐☐
Week of pregnancy you are in	_____ weeks
Date of Birth	*Day* *Month* *Year* ☐☐ ☐☐ ☐☐☐☐
Have you moved address since you enrolled in BiB? *If yes, what is your new address?* Flat/House No.: Street: City/County: Postcode: *What was the date of your move?*	Yes / No *Day* *Month* *Year* ☐☐ ☐☐ ☐☐☐☐

FORM CONTINUES ON THE NEXT PAGE
PLEASE TURN OVER ➡

Are you currently a full-time student?	Yes / No
Are you currently working? *If yes, how many days in a typical recent week do you go to work?* *(enter 0 if you work mainly <u>at</u> or <u>from</u> home)*	Yes / No No. of days: ☐☐
If you are working, what is the address of your main place of work? Building Name or No.: Street: City/County: Postcode:	
Are you currently on maternity leave? *If yes, since what date have you been on maternity leave?*	Yes / No *Day* *Month* *Year* ☐☐ ☐☐ ☐☐☐☐
Are you currently on sick leave? *If yes, when did you start sick leave?*	Yes / No *Day* *Month* *Year* ☐☐ ☐☐ ☐☐☐☐

After completion, please seal this form (along with the yellow questionnaire) in the envelope provided and put it in the post to us at your nearest post box or post office: *Born in Bradford, Bradford Institute for Health Research, Bradford Royal Infirmary, Duckworth Lane, Bradford, BD9 6RJ, UK.*

Thank you!

Imperial College London

BORNINBRADFORD For a Healthy Future

Title of Project: BiB Water Use Study

BiB Study ID:

Name of Researcher: Susan Edwards

Repeat Questionnaire Study ID: (week 30-33)

TODAY's DATE: _____

W1. **On a typical day in the past week of your pregnancy how much of the following did you drink?**

(fill in no. of glasses/cups per day; 1 glass is 200ml, 1 cup is 200ml, 1 mug = 2 cups)

	at home	at work/study	elsewhere
1a. Tap water	Glasses per day: ☐☐	Glasses per day: ☐☐	Glasses per day: ☐☐
1b. Bottled water (includes water cooler)	Glasses per day: ☐☐	Glasses per day: ☐☐	Glasses per day: ☐☐
1c. Tea (any sort)	Cups per day: ☐☐	Cups per day: ☐☐	Cups per day: ☐☐
1d. Coffee	Cups per day: ☐☐	Cups per day: ☐☐	Cups per day: ☐☐
1e. Squash (including any other drinks made with tap water)	Glasses per day: ☐☐	Glasses per day: ☐☐	Glasses per day: ☐☐

W2. **Did you filter the water you drink at home during the past week?** *(cross ONE box only)*

☐ Yes ☐ No ☐ I don't know

W3. **Did you filter the water you drink at work during the past week?** *(cross ONE box only)*

☐ Yes ☐ No ☐ I don't know ☐ I was not working

QUESTIONNAIRE CONTINUES ON THE NEXT PAGE
PLEASE FLIP OVER ➡

W4. **In the past week of your pregnancy, how often and for how long did you undertake the following?** *(if you do not do any, then fill in 0)*

	Times per week	Minutes each time
4a. **Shower**	☐☐	☐☐☐
4b. **Bath**	☐☐	☐☐☐
4c. **Swim**	☐☐	☐☐☐

4d. **Do you have a shower installed at home?**
☐ Yes ☐ No

4e. **Do you have a bath installed at home?**
☐ Yes ☐ No

W5. **Do you think that your overall water use habits have changed since you completed the last questionnaire, upon enrolment to BiB?** *(cross ONE box only)*
☐ Yes ☐ No

W6. **Do you think that the following specific water use habits have changed since you completed the last questionnaire?**

6a. **Your tap water drinking habits?** *(cross ONE box only)*
☐ Yes ☐ No

6b. **Your showering habits?**
☐ Yes ☐ No

6c. **Your bathing habits?**
☐ Yes ☐ No

6d. **Your swimming habits?**
☐ Yes ☐ No

After completion, please seal this form (along with the green recruitment form) in the envelope provided and put it in the post to us at your nearest post box or post office: *Born in Bradford, Bradford Institute for Health Research, Bradford Royal Infirmary, Duckworth Lane, Bradford, BD9 6RJ, UK.*

Thank you for your time! We couldn't do this research without your help!