Maternal health exposures and pregnancy outcome: Examining symptoms of nausea and vomiting in early pregnancy, maternal caffeine consumption, and spontaneous abortion utilizing regression and propensity score methodologies

Ronna L. Chan

A dissertation submitted to the faculty of The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Maternal and Child Health (reproductive epidemiology), School of Public Health

Chapel Hill
2008

Approved by:

Sandra L. Martin, Ph.D. (Chair, Academic Advisor)

Andrew F. Olshan, Ph.D. (Co-Research Advisor)

David. A Savitz, Ph.D. (Co-Research Advisor)

Amy H. Herring, Sc.D.

Julie L. Daniels, Ph.D.

Herbert B. Peterson, M.D.

ABSTRACT

Ronna L. Chan

Maternal health exposures and pregnancy outcome: Examining symptoms of nausea and vomiting in early pregnancy, maternal caffeine consumption, and spontaneous abortion utilizing regression and propensity score methodologies

(Under the direction of Andrew F. Olshan, Ph.D., and David A. Savitz, Ph.D.)

This dissertation addressed two important issues concerning pregnant women: caffeine consumption and nausea and vomiting in early pregnancy (NVP). Factors influencing NVP have not been well established and methodological limitations in previous research merit closer evaluation on the caffeine consumption-spontaneous abortion (SAB) and NVP-SAB relationships. This study examined potential risk factors or markers for NVP, and the associations between NVP, caffeine exposures, and SAB using refined classifications of NVP and traditional and novel analytic methodologies.

Our analyses included 2,430 newly pregnant women, who participated in a prospective cohort epidemiologic study from 2000 to 2004. Detailed NVP and caffeine consumption data and other health and pregnancy outcome information were collected through interviews, ultrasound assessments, and medical and vital records. Modified Poisson regression with robust error variance models generated risk ratios for potential risk factors to NVP characteristics. Discrete-time continuation ratio logistic survival models were used to estimate week-specific pregnancy loss associated with NVP and caffeine exposures.

Additionally, we introduced propensity score methods to evaluate the effects of caffeine exposures on SAB.

Most maternal characteristics were not associated with having NVP, except for plurality. Increased risk for delayed onset was found with increasing maternal age, among non-Hispanic Blacks, Hispanics, or age at menarche between 12 and 13 years old. Additionally, older women and non-Hispanic Blacks were less likely to experience symptoms beyond first trimester. Absence of NVP is associated with an increased risk for SAB, compared to having any symptoms. Furthermore, age was found to modify the association between symptom severity and duration and SAB. Reduced risks for SAB was found across all age groups for longer duration, but the effects were much pronounced in the oldest age groups. There was little indication of potential harmful effects of caffeine on SAB; our results showed no overall meaningful differences in estimates from traditional covariate adjustment versus propensity score models.

Using improved exposure assessments and analytic methodologies, we identified characteristics that are associated with subclasses of NVP, in addition to providing better understanding in NVP-SAB and caffeine-SAB relationships. Finally, our findings suggest that propensity score methodology is an important addition in studying pregnancy health.

ACKNOWLEDGEMENTS

This dissertation could not have been completed and my time spent as a graduate student at UNC would not have been fruitful if it were not for the dedication of the many professors at UNC. First, my sincere gratitude goes out to my dissertation committee, Drs. Sandra Martin, Andrew Olshan, David Savitz, Amy Herring, Julie Daniels, and Herbert Peterson for their time, continuous support and guidance through my dissertation.

I am especially grateful for the opportunities that Drs. David Savitz and Andrew Olshan have provided for me as graduate research assistant. My involvement with the Right from the Start (RFTS) study and the Healthy Men Study in the past five years has strengthen me as a collaborator and a leader, and the skills that I learned have helped me develop into a better researcher. Not only did I learn a lot from them and the other investigators and staff on the projects, we had our share of fun too!

The use of RFTS data for this dissertation would not be possible without the many women who participated in the study. Thank you for your interests in learning more about pregnancy health and for allowing us to peak inside your lives for a short period of time! Additionally, this study would not be successful without the other Principle Investigators (Drs. Katherine Hartmann, Philip Singer, and Howard Weinberg) and study staff (Christina Makarushka, Yanfang Jiang, David Kleckner, Kerri Kruse, and Andrea Lindsay). I also would like to thank the participating medical practices and water companies in three cities. Another big thanks goes to the American Water Works Association Research Foundation

(AwwaRF) and the U.S. Environmental Protection Agency (EPA) for funding RFTS, under Cooperative Agreement Nos. CR825625-01, CR827268-01, and CR82826-01.

I thank the Maternal and Child Health Bureau for funding my training in the doctoral program for the first two years under the Maternal and Child Health Epidemiology Doctoral Training Grant.

I also cannot say thank you enough to Linda Southern (retired) and Nancy Colvin at the Departments of Maternal and Child Health and Epidemiology for never getting tired of answering my questions and for helping me navigate this system known as graduate school. Nancy, thanks for letting me just pop into your office for brief and sometimes lengthy visits. I have also enjoyed our many chats about life, our dogs, careers, etc., and I know we will continue to do so for many years to come.

Thank you to my many friends for their support. Whether I've known them for just a few years or since high school and college, each of them adds a little brightness to my life. I have especially enjoyed our hikes at the local trails and state parks, kayaking trips, impromptu dinners at the Village Tavern, meeting up for margaritas, tea times at Guglhupf, meeting up at the "half-way mark" just to catch up, watching men's college basketball games, and shopping excursions in New York City, Chicago, and Boston.

Finally, I thank my family for their unwavering love and support and for being there every step of the way. Mom and Dad, you made the right moves back in 1988. Thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CATI	Computer Assisted Telephone Interviewing
DAG	Directed Acyclic Graph
DBP	Disinfection by-product
EGA	Estimated gestational age
hCG	Human Chorionic Gonadotropin
LMP	Last menstrual period
MAR	Missing at random
MCAR	Missing completely at random
mg	Milligram
mg/day	Milligram per day
mg/oz	Milligram per ounce
mL	Milliliter
NVP	Nausea and vomiting in early pregnancy
OR	Odds ratio
oz	Ounce

RFTS	Right from the Start
RR	Risk ratio
SAB	Spontaneous abortion
UNC	The University of North Carolina at Chapel Hill

CHAPTER 1: INTRODUCTION AND SPECIFIC AIMS

1.1 Introduction

Nausea and vomiting in early pregnancy (NVP) affects 50%-90% of pregnant women some time before 20 weeks' completed gestation.[1, 2] The exact cause for NVP is not known, but the biochemical changes during pregnancy may be responsible for this phenomenon. Several hormones may contribute to the symptoms of NVP. Most notably, human chorionic gonadotropin (hCG) has been proposed as the most likely contributor to NVP due to the rapid increase in its levels in the first trimester.[3] Although some studies have reported that the presence of NVP is associated with decreased adverse pregnancy outcomes, particularly spontaneous abortion (SAB),[3-8] the role of NVP has not been studied extensively. Earlier studies had methodological limitations which may potentially underestimate the association between NVP and SAB.

Spontaneous abortions or miscarriages are pregnancies that are lost prior to 20 completed weeks' gestation from the first day of the last menstrual period (LMP). Studies of SABs are often challenging, primarily due to the difficulty in identifying women early in pregnancy. In most cases, recognition of pregnancy can be detected with standard assays for hCG prior to a missed period or via home urine pregnancy test kits by the time of a missed menstrual period. Nonetheless, challenges in identifying women early in their pregnancy persist because many women may not be aware that they were pregnant due to

unintentional and/or unplanned pregnancies. Because most pregnancy losses are concentrated in the earlier weeks of gestation, early enrollment of women into the research study is critical in the early identification of pregnancies and the identification of pregnancy losses. Current knowledge on the etiology of SAB is limited but several risk factors have been shown to be influential. Advanced maternal age and history of previous SABs have shown to be the strongest predictors,[9, 10] while other risk factors such as cigarette smoking,[11] caffeine consumption,[12-14] occupational chemical exposures,[15] and drinking water disinfection by-products[16] are also possible contributors. Earlier epidemiologic studies have repeatedly demonstrated that absence of nausea in the early gestation period may be indicative of a pregnancy already lost or as a predictor of an impending pregnancy loss. Additionally, timing of onset and duration of NVP may also potentially be indicators of adverse pregnancy outcomes such as SAB, though no prior research has examined these relationships.

Finally, observational data are often used in reproductive and perinatal research to make inferences on causal effects. Because it is often unjust and unethical to randomize pregnant women to specific exposures, observational research is the best type of study in studying exposure-outcome relationships. However, the causal interpretation from observational data is complicated because the exposure groups may differ systematically with respect to relevant observed covariates, and therefore, not be directly comparable and these differences may lead to biased effect estimates.[17-20] One notable research question in reproductive epidemiology that often rely on the use of observational data is the association between caffeine exposure in pregnancy and the risk of SAB. Results from earlier epidemiologic studies have generated much controversy and public health

interest since the 1980's; however, conclusive evidence has not been established.[21] The concern for the impact of caffeine consumption is caffeine's ability to affect pregnant women differently than non-pregnant women. Methodological limitations in earlier studies, including differences in measuring and quantifying exposure and control of confounding, pose a challenge to the interpretation of results from previous studies.

1.2 Specific Aims

Symptoms of nausea and vomiting in early pregnancy (NVP) are common phenomenon that affects the majority of pregnancy women some time during the course of their pregnancy. Spontaneous abortion (SAB) is defined as pregnancy loss at less than 20 completed weeks' gestation and occurs in approximately 10% of all clinically-recognized pregnancies.[22] In this dissertation, I addressed the following specific aims:

1.2.1 Specific aim 1

Specific aim 1 of this study was to determine various maternal factors that influence symptoms of NVP. Potential risk factors or markers for the characteristics of NVP symptoms involving severity, time of onset, and duration were examined in specific aim 1. Factors that were examined include sociodemographic variables, maternal health behaviors, and women's reproductive and medical histories. We hypothesized that women are more likely to experience NVP if they were younger, African-American, non-smokers, non-drinkers, 'younger' age at menarche, multiparous, and had prior SABs. The associations between NVP and its possible predictors are stronger for women with NVP with delayed symptom onset and with longer symptom duration.

1.2.2 Specific aim 2

Specific aim 2 of this study was to evaluate the relationship between NVP and pregnancy loss. Moreover, we examined symptoms of NVP with respect to symptom severity and duration. Covariates of interest included sociodemographic variables, maternal health behaviors, reproductive history, and medical history. We tested the hypothesis that the absence of NVP in early pregnancy is associated with a higher risk of spontaneous abortion, with higher risk associated with women with no symptoms or short symptom duration.

1.2.3 Specific aim 3

Specific aim 3 of this study was to apply propensity score methods in studying maternal health behaviors and exposures in pregnancy. The goal of many reproductive and perinatal epidemiologic studies is to estimate the causal effects of exposures on health outcomes. The use of randomized studies is not always justified or ethical; consequently, observational data are often used as an alternative to interpret causal effects. Traditional covariate adjustments in multivariable analysis may not be sufficient to correct the bias in estimation that potentially arises from the systematic differences in observed covariates across exposure status. However, propensity score was developed by Rosenbaum and Rubin to reduce bias in observational studies.[23]

Because pregnancy is a time in which many women modify their behavior for the health of the fetus or under the advice of clinicians; therefore, investigators are often interested in studying the effects of maternal exposures and behaviors on pregnancy health. In specific aim #3, we illustrated the application of propensity score methods in

study reproductive health in an extension of an earlier study that evaluated the association between caffeine exposure at three time periods and SAB.[24] Caffeine provides a useful example as it's consumption around the time of pregnancy is influenced by a number of factors and the use of propensity score methods allowed us to compare findings from traditional multivariable models with covariate adjustments with propensity score models.

CHAPTER 2: BACKGROUND AND SIGNIFICANCE

2.1 Nausea and vomiting in early pregnancy

2.1.1 NVP characteristics

Little is known about the natural history of early pregnancy, and pregnant women may not know what to expect with regard to symptoms.[25] Nausea and vomiting in early pregnancy (NVP) is a common phenomenon that affects 50%-90% of pregnant women.[4, 5, 26] Symptoms may differ in the timing of onset, duration, and severity, both among individual women and among individual pregnancies of the same woman, and some women experience symptoms of nausea during pregnancy with or without episodes of vomiting. Although the presence of NVP is unpleasant and can be somewhat debilitating, they may protect women from ingesting substances that could potentially damage the embryo during the crucial stages of embryogenesis or organogenesis.[27] Due to its common occurrence, the onset of NVP symptoms often is the first indicator to a woman that she is pregnant. Any delay in the onset of symptoms may delay a woman's actions to make necessary behavioral changes to protect her embryo.[25] In spite of the many theories that have been generated to explain early pregnancy symptoms, the etiology of NVP remains unclear. The secretion of human chorionic gonadotropin (hCG) and the rapid rise in its levels in the first trimester have been proposed as the most likely contributors to NVP.[2, 28] It was demonstrated that presence of symptoms during early

pregnancy has a protective effect on the health of the growing fetus, where various mechanisms have been proposed to explain the feto-protective effect of NVP symptoms; this included the reduction of fetal exposure to potential teratogens present in the mother's diet, a decrease in energy intake that elicits hormonal changes to stimulate early placental and fetal growth, and other nutritional and non-nutritional pathways.[29]

In a comprehensive review by Gadsby et al.[30] on the timing of onset, duration, and severity of NVP, the authors found that symptoms of nausea and vomiting episodes in pregnancy began between the fourth and seventh week after the last menstrual period (LMP) in 70% of women; however, a small proportion (7%) noted NVP symptoms prior to their missed menstrual period. Women in this study population reported symptom onset before nine weeks' gestation, and decrease in symptoms was noted in 30% by the 10^{th} week, another 30% by the 12^{th} week, and the last 30% by 16^{th} week. An analysis of the 24-hour variation in NVP in this study demonstrated that symptoms occurred commonly between 6 am and noon, but symptoms were experienced by many women throughout the day. While most studies did not examine the natural 24-hour variation in NVP symptoms, these studies provided further confirmation that onset of NVP is early in the gestational period; however, the duration and severity of the symptoms varied between the studies.[8, 31, 32]

While most women's experiences with NVP symptoms are restricted to early pregnancy; approximately 20% of pregnant women continue to experience the symptoms beyond 20 weeks gestation.[1] Data from an earlier study indicated NVP in late pregnancy resulted in lower prenatal weight gains and lower infant birth weight, although this study focused only on pregnant teenagers.[33] In a small study of Midwestern women receiving

care from prenatal care clinics in their first trimester, 70% of women reported experiencing symptoms of nausea and vomiting during the early gestational period; however, 32% continued to have NVP symptoms past the first 20 weeks. Women with late NVP tended to be older, had higher parity, and gained less weight in the pregnancy than women who did not experience late NVP.[34]

The most severe form of NVP, hyperemesis gravidarum, affects 0.5%-2% of pregnancies, and it is one of the most common reason for hospitalization in pregnancy, second only to preterm labor.[35] Hyperemesis gravidarum is a diagnosis of exclusion based on a typical presentation in the absence of other disorders that could explain the findings. Despite a lack of uniform definition of hyperemesis gravidarum, persistent vomiting, dehydration, ketonuria, electrolyte disturbance, and weight loss greater than 5% are the usual criteria.[36]

2.1.2 Risk factors or markers for NVP

Symptoms of NVP begin typically at two to four weeks gestation;[30, 37] generally peak in intensity between 8 and 12 weeks, and subside by 20 weeks' gestation.[37-39] To date, no confirmed predictors of NVP has been established; however, a few previous studies found that the presence of NVP is associated with older maternal age, occupation such as manual or service workers, cigarette smoking, and infant gender.[3, 5, 39] However, other studies yielded conflicting results on these factors.[31, 40] Furthermore, chronic illnesses such as hypertension, liver and renal diseases,[40] maternal hormones,[41]vitamin use,[32] gravidity, plurality, increasing number of prior miscarriages,[42] and stress[43] have also shown to be predictors for NVP.

2.2 Nausea and vomiting in early pregnancy & spontaneous abortion

2.2.1 Spontaneous abortion

Spontaneous abortion (SAB) or miscarriages are pregnancies that are lost prior to 20 completed weeks' gestation from the first day of the LMP. Although between 10%-12% of pregnancies that are identified resulted in SAB, the proportion of losses that take place prior to six weeks' gestation may be two to three times higher.[22] Pregnancy loss can be stressful and traumatizing for some women and women who experience a pregnancy loss may exhibit symptoms of distress such as depression, anxiety, and prolonged grief.[44-46] Advanced maternal age and history of previous SABs are the strongest predictors for SAB;[9, 10] other potential risk factors include the presence of uterine fibroids,[47, 48] incompetent cervix,[49] occupational chemicals,[15] cigarette smoking,[11] caffeine consumption,[12-14] drinking water disinfection by-products[16]

2.2.2 The association between NVP and SAB

The role of NVP has been examined in a few previous research studies in relation to pregnancy outcomes. Several studies have found that symptoms of NVP are strongly associated with a decreased risk of SAB.[3-5] Tierson et al.[4] examined the patterns of NVP on pregnancy outcome and reported that while 90% of the study population experienced some form of NVP symptoms during pregnancy, the 10% of women who had no symptoms of NVP had a higher incidence of SAB. A meta-analysis was performed on 11 previous studies to examine the effects of NVP on the risk of SAB.[8] The authors found that the presence of NVP were significantly associated with a decreased risk of miscarriage.

Although much previous research found consistent association between positive pregnancy outcome and the presence of NVP symptoms during pregnancy, one study found a difference in pregnancy outcome dependent the severity of NVP. In this retrospective cohort study of women in a large city on the West Coast, Weigel and Weigel[7] did not find any association of NVP with pregnancy outcomes such as perinatal death, fetal anomalies, preterm delivery, and low birthweight. However, symptoms of nausea and accompanying vomiting episodes were associated with a decreased risk of SAB (adjusted OR=0.18, 95%CI: 0.06, 0.53), but women with nausea symptoms without any vomiting episodes had a miscarriage risk equal to that in the overall study population (adjusted OR=0.32, 95%CI: 0.15, 0.68). Results from a prospective cohort study on Ecuadorian prenatal patients confirmed previous findings that women who experienced nausea only (adjusted OR= 0.45, 95%CI: 0.22, 0.94) and those who experienced nausea with vomiting had a decreased risk of spontaneous abortion (adjusted OR=0.66, 95%CI: 0.46, 0.99).[50] Selected studies on the risk factors of NVP and its association with SAB are detailed in Tables 1 and 2. Figure 1 is a Directed Acyclic Graph (DAG) which provides a visual diagram on the causal relationship between NVP and SAB, along with covariates of interests;[51, 52] each linked covariates by arrows represents direct causal effects, either protective or causative, of one variable on another. Maternal demographics such as race/ethnicity, education, marital status, and age have been known to influence maternal health behaviors during pregnancy. Maternal age at the time of pregnancy and parity may impact a woman's symptoms of NVP. Women with certain maternal demographic characteristics may be more prone to experience prior pregnancy losses. Prior SABs,

woman's age at menarche, parity, maternal health behaviors may cause SAB in a woman's "current" pregnancy.

2.3 An application of propensity score methods in studying maternal health exposures and behaviors in pregnancy health

2.3.1 The use of observational study data and propensity score methods

The goal of many epidemiologic studies is to estimate the causal effect of some specific exposures on a health outcome,[53] and reproductive and perinatal health studies are no exception. Investigators often rely on the use of observational research data for estimating causal effects; causal interpretation from observational data is complicated because exposure groups may differ systematically with respect to relevant observed covariates and, therefore, may not be directly comparable and such differences may lead to biased effect estimates. [17-20] Traditional covariate adjustments may not be sufficient to correct this bias in estimation; propensity score methods have been named as another approach to control confounding and reduce bias in observational data.

The concept of propensity scores has been used in pharmacoepidemiology and other disciplines, but reproductive and perinatal epidemiologic studies have been slow to adopt this method. Propensity score was developed by Rosenbaum and Rubin[23] to reduce bias in observational studies. In theory, the effect of an exposure can be measured among individuals with the same probability of being exposed, thus allowing control for confounding; individuals with the same propensity score have the same chance of being exposed even though in truth some were and some were not. Therefore, this method is often conceptualized "virtual randomization" and a useful tool to evaluate maternal

exposures and outcome relationships where randomization to specific exposures is not ethical.

2.3.2 Propensity score

Under the assumption of strongly ignorable treatment or exposure assignment, a propensity score is defined as the conditional probability of exposure to a potential risk factor given an individual's set of observed covariates.[23] Each individual has a vector of observed covariates measured at baseline, X_i and an indicator of exposure, Z_i, where Z_i =1 if exposed and $Z_i = 0$ if unexposed. The propensity score, $e(X_i)$, is the probability of exposure for a person with covariates X_i, that is,

$$e(X_i) = P(Z_i = 1 \mid X_i = x_i),$$

where we assume that Z_i , i =1, and the number of study subjects are independent conditional on X. It is a function of the observed covariates X such that the conditional distribution of X given the propensity score $e(X_i)$ is the same for subjects with $Z_i = 1$ and Z_i =0. The propensity score is estimated from observed data using logistic regression.[17] Individuals with the same estimated propensity score are thought to have the same probability of being exposed, although they may have different sets of X's.

2.3.3 An example: Caffeine and the risk of SAB

While most sources of caffeine comes from coffee and that coffee drinking in many countries and societies is common, it is considered a harmless habit in many countries. However, caffeine is a central nervous system stimulant known to cross

placental membranes and the blood-brain barrier during pregnancy,[54] slower in metabolism later in gestational age during pregnancy,[55, 56] potential to influence cellular development by increasing cAMP,[57] and the interference with utero-placental circulation with increased catecholamines.[58] Because of the properties of caffeine, its use by pregnant women has generated much controversy in the past 30 years. Some earlier epidemiologic studies have implicated caffeine as a risk factor for adverse pregnancy outcomes such as preterm birth,[59] birth defects,[26] fetal death (>20 weeks' or >140 days' gestation),[60, 61] and small-for-gestational-age.[62] Despite many studies in recent decades that evaluated the association between caffeine exposure and SAB,[12-14, 60, 63-68] these studies and others showed conflicting results due to methodological limitations.[21] Consequently, this poses a major challenge to the interpretation of results from previous studies. For example, in a prospective cohort study in New Haven, the researchers found that the odds for SAB was 2.6 times (95%CI: 1.3, 5.3) higher for women who consumed >300 mg/day than those who were non-consumers.[63] Bech[60] and Tolstrup [66] found a positive association between caffeine exposure and SAB in their respective studies, but the magnitude of association was not quite as high that found in the New Haven study. In the Danish National Birth Cohort study,[60] women who consumed ⩾8 cups of coffee had 1.5 times (95%CI: 1.0, 2.2) and 2.3 times (95%CI: 1.3, 3.9) the average hazard for SAB in <140 days gestation and between 140-195 days gestation respectively. A recent published study showed very little to no evidence overall that caffeine consumption increases the risk of SAB, despite the study's use of prospective cohort data and evaluated caffeine exposure at three time points.[24] However, a different study that was

published in the same year yielded conflicting results.[69] A summary of epidemiologic studies that evaluated the association between caffeine and SAB is shown in Table 3.

2.4 Limitations of prior studies

2.4.1 Limitations in studying SAB

Early stages of development are crucial to the survival of an embryo developing into a healthy fetus; many pregnancies that result in losses typically occur prior to 20 completed weeks' gestation with many losses occurring right before a woman is aware of her pregnancy. Because of the difficulties in recognizing pregnancies and pregnancy losses, many epidemiologic studies that are designed to recruit participants at a later gestational age would have missed ascertaining relevant information from women who already experienced pregnancy losses. This consequently leads to a major methodological issue known as left truncation, which resulted because of a selection process in the study design. Left truncation is of special concern when we have subjects who had the event of interest (i.e. pregnancy loss) but are not in the study.[70, 71] SAB data are often left truncated due to the high prevalence of unplanned pregnancies, high proportion of loss experienced by the source population prior to enrollment into study, and study subjects for SAB studies are commonly recruited after conception; therefore, the true cohort of pregnant women cannot be known and many women who experience pregnancy loss generally would have lost their pregnancies before they are even aware that they were pregnant, or that they first become aware of their pregnancy at the time of the loss. For this reason, many women who otherwise would have been eligible in the study are left out in the analysis data set because they have already experienced the event.

While many methodological limitations are unique to the studies on the effect of NVP symptoms on the risk of SAB and the effect of caffeine exposure on SAB, another possible limitation to studying SAB is the potential inaccuracy in gestational age dating. Estimated gestational age is determined based on the first day of a woman's LMP or with the use of ultrasound. Errors in gestational age dating can occur as a result of recall error, where a woman is unable to recall the exact date of her last menstrual period and she makes up a date, or she is asked to recall in which week of the month she had her last menstrual period. Self-reported or assigned LMP can lead to digit preference.[72, 73] Error in estimating gestational age can also be influenced by the time in which ultrasound is used for pregnancy dating. Because assessment of gestational age using ultrasound is shown to be more accurate for a younger pregnancy relative to an older pregnancy, the timing in which ultrasound is used for pregnancy dating can influence the assignment of gestational age relative to a pregnancy's true age. Regardless of the sources of errors, inaccuracy in the dating of pregnancy age can lead to potential inaccuracy in dating exposure or potentially misclassifying an early pregnancy loss that takes place close to the cutpoint, by definition of SAB, as a fetal demise.

2.4.2 Limitations of previous NVP studies

Although findings from some previous research on the effects of NVP on SAB showed fairly consistent results, methodological limitations could have potentially biased the estimates. Often, these previous research were conducted using retrospective cohort and case-control study designs,[35-38] with the exception of the Early Pregnancy Study,[38] the Diana Project,[74] and the Ecuadorian study,[50] where the reproductive outcomes were

examined prospectively. Retrospective cohort and case-control studies are subjected to exposure misclassification, especially when studies recruit women in later gestation which would bias the effect estimates. Exposures could also be non-differentially or differentially misclassified depending whether women with pregnancy losses recall past exposures more accurately in comparison to women with no history of pregnancy losses. The variability in the classification of exposure groups in previous studies may introduce non-differential misclassification which may generate either an underestimation or an overestimation on the effect of NVP symptoms in relation to SAB.

Analytic techniques used in earlier studies posed additional challenges in studying the association between NVP and SAB. Most investigators utilized logistic regression to model the odds of NVP and SAB, even in prospective cohort studies;[50, 74] while this method was suitable for some study designs, the use of logistic regression modeling did not take into consideration that women are observed to be at risk for SAB for varying lengths of time because they entered the studies at different gestational ages. Because logistic regression does not account for left truncation and assumes that follow-up time for all subjects are the same, it is especially problematic when the left truncation is differential with respect to exposure status; thus using logistic regression to analyze SAB data would create a biased estimate of effect. In addition to the limitations from study designs and analytic methods, the comparability of study results in earlier studies is further compromised, because the operational definitions of NVP varied across studies. Some studies utilized a dichotomous approach in categorization of NVP symptoms, while other studies categorized NVP ordinally. Those that coded NVP data as “any versus

none" did not take into account symptom severity, duration, or timing of onset of the symptoms.

2.4.3 Limitations of previous caffeine exposure studies

Methodological limitations in previous studies on the association between caffeine exposure and SAB have hampered the interpretation of earlier findings. Few major methodological issues included exposure assessment, quantifying caffeine exposure, and identification of pregnancy symptoms.

Exposure assessment

Caffeine consumption generally changes drastically when a woman becomes pregnant. The decrease in intake may be due to the advice of clinicians or in response to an increase in olfactory sensitivity or the presence of pregnancy symptoms such as NVP. Consequently, ascertaining information on caffeine intake during pregnancy is inherently more difficult. Because many pregnant women modify their behavior during the course of the pregnancy, it is not enough to elicit a single measurement to capture true caffeine intake; nevertheless, many earlier studies have done so.[68, 75-77] Only a handful of studies attempted to incorporate temporal changes in caffeine intake. Despite this approach, a number of studies ascertained information on change in consumption, if applicable, and then computed an average consumption value.[14, 64, 76, 78] While this form of exposure assessment is a slight improvement over single exposure measurement over the course of the pregnancy; nevertheless, this approach is not robust enough to capture multiple changes in consumption over the course of the pregnancy.

Quantifying caffeine exposure

The amount of caffeine present in a beverage depends on the method of preparation and the beverage cup size. Only a small number of prior studies took caffeinated beverage preparation method[65, 76] into consideration and a few studies collected information on beverage cup size.[65, 76, 79] The range of caffeine assigned to each caffeinated beverage has been inconsistent across studies; depending on a study's methods in quantifying caffeine exposure, the absolute caffeine intake of an individual based on the reported amount in one study may be different in another. This large variability in quantifying caffeine exposure often leads to measurement error and make comparing results across studies difficult.

Accounting for pregnancy symptoms

NVP, a common type of pregnancy symptoms, often causes pregnant women to develop aversions toward specific foods, drinks, tastes, or smells, leading to potentially drastic lifestyle and behavioral modification during the course of viable pregnancy, while women with non-viable pregnancies may carry on with pre-pregnancy behavior because such behavior is not deterred by pregnancy symptoms. Caffeine exposure assessments, in most studies, are likely to take place after changes have already been made under the influence of NVP; women who retained their pre-pregnancy behavior (i.e. caffeine intake) in light of the absence of NVP symptoms could be interpreted as causally linked to SAB. Therefore, it is unclear to what extent is caffeine exposure truly affecting the pregnancy outcome or if the exposure-outcome relationship is an artifact of NVP. One

review[21] uncovered studies on caffeine exposure and SAB often did not take NVP into consideration because data 1) were not collected [63, 66, 75, 80, 81], 2) were collected but were incomplete[60], or 3) were not incorporated into the analysis.[64, 79] While many studies that examined caffeine exposures and SAB often failed to take into consideration NVP symptoms, those that attempted only had done so simplistically. Because of methodological limitations in earlier studies, particularly in failing to address confounding by NVP symptoms, the interpretation of results from previous studies on the effects of caffeine and SAB is further limited.

2.5 Contribution of the proposed research

The proposed study addressed some of the limitations in prior research by using data from the Right from the Start (RFTS) study to 1) identify potential risk factors for NVP, 2) examine the association between NVP and spontaneous abortion, and 3) examine the effect of caffeine exposure at three time points and SAB. Because most pregnancy losses are concentrated in the earlier weeks of gestation, studies of SAB are often challenging, primarily due to the difficulty in identifying women early in pregnancy. The protocol of RFTS called for the identification of women early in their pregnancy (<12 completed weeks gestation) or women who were trying to conceive; therefore, we were able to successfully track the course of pregnancy events and outcomes for all study participants. With a prospective study design, we were able to collect detailed information on maternal behaviors and exposures in early weeks of gestation. We were able to follow-up and confirm pregnancy losses through participant self-report and medical record abstraction, and for other pregnancy outcomes via

participant self-report, medical record abstraction, and linkages of vital fetal death records.

Nausea and vomiting in early pregnancy was classified based on symptom severity, time of onset, and duration. Caffeine exposure was classified as caffeine from coffee only and caffeine from caffeinated coffee, caffeinated iced and hot tea, and caffeinated sodas. Pregnancy outcomes, including SABs, were ascertained in the parent study through three mechanisms which included participant self-report, medical record abstraction, and matching of vital records data. Estimated gestational age was assigned based on the women's last menstrual period or assessed with the use of endovaginal ultrasound dating.

Modified Poisson regression with robust variance was used to model the log risk of maternal age, maternal race/ethnicity, marital status, education, smoking, alcohol use, pregnancy loss history, plurality, parity, and age at menarche as potential risk factors for NVP. Discrete hazard models were fitted to examine the association between NVP and SAB, with maternal age, maternal race/ethnicity, education, marital status, education, smoking, alcohol use, age at menarche, parity, plurality, and SAB history as potential effect measure modifiers and/or potential confounders. We used the example of caffeine intake and SAB to illustrate the application of propensity score methods in observational study data. These improved study features and the application of more sophisticated and appropriate analytic methods allowed us to produce a more valid magnitude of association and for us to account for the time-varying outcome that was not considered in the analyses of earlier research studies.

Table 1. Selected literature on studies of risk factors or markers influencing nausea and vomiting in early pregnancy

Reference (Year)	**Study design, study population, and data collection**	**Risk factors or markers examined for NVP**	**Assessment of NVP symptoms**	**Main findings**
O'Brien (1995)	**Design**: Cross-sectional **Population**: pregnant women who sought prenatal care <16 weeks in three Midwestern clinics **Study subjects:** 147	Maternal age, Education, Parity, Occupation, Cigarette smoking, Other demographic characteristics (not reported)	Measurement collected via questionnaire on NVP: Occurrence, Duration, Frequency, Amount,	Maternal age, Occupation (working outside home), Parity, Cigarette smoking, Infant gender
Zhang (1999)	**Design**: Case-control **Population**: Sample of 29 hospitals in Shanghai, China. Cases-perinatal deaths; controls-live births born around time of the case **Study subjects:** 1,875 cases, 1,875 controls	Chronic hypertension, Chronic liver disease, Chronic renal illness, Passing smoking, Maternal age, Gravidity	Measurement of vomiting only; ascertained from prenatal care card	Chronic hypertension, Chronic liver disease, Chronic renal illness, Passing smoking
Zhou (1999)	**Design**: Secondary data analysis from a RCT **Population:** Volunteers recruited through newspaper advertisement for RCT to evaluate the efficacy of P6 acupressure on NVP **Study subjects**: 103	Maternal age, Previous SAB, Parity, Employment status, Infant gender, Infant birthweight	Measurement on NVP severity from mail questionnaire	Infant birthweight

Table 1 (cont). Selected literature on studies of risk factors or markers influencing nausea and vomiting in early pregnancy

Reference (Year)	Study design, study population, and data collection	Risk factors or markers examined for NVP	Assessment of NVP symptoms	Main findings
Källén (2003)	**Design**: Prospective cohort **Population:** Women seeking care in prenatal care clinic in Sweden **Study subjects**: 3,675 women	Maternal age, Parity, Cigarette smoking, Occupation, Vitamin use	Measured via questionnaire for NVP severity (none, nausea only, both), number of times per day, and symptom duration	Maternal age, Parity, Occupation (worked outside home), Vitamin use
Lagiou (2003)	**Design**: Prospective cohort **Population**: Boston women who participated in study on breast cancer among White and Asian women **Study subjects:** 402	Prolactin, Estradiol, E3, Sex hormone-binding globulin, Progesterone	Measured via self-report interview on NVP severity (none, nausea only, both)	Prolactin, Estradiol
Louik (2006)	**Design:** Case-control **Population:** Birth defects surveillance of The Slone Epidemiology Center at Boston University. Cases-women with infants with congenital malformations; controls-women with live birth in greater metropolitan areas of Boston, Philadelphia, Toronto, and San Diego **Study subjects:** 17,158 cases, 5,329 controls	Plurality, Gravidity, Planned pregnancy, Education, Maternal age, Maternal race/ethnicity, Income ($/year), Cigarette smoking, Pre-pregnancy weight, Obstetrics characteristics	Measured via interview where 1) NVP asked as single event (1988-1992) 2) Nausea and vomiting asked as separate entity (>1992) in terms of occurrence (yes/no), duration, and time of onset	Maternal age, Maternal race/ethnicity, Gravidity, Pre-pregnancy weight

Table 2. Summary of selected epidemiologic studies examining the association between nausea and vomiting in early pregnancy and spontaneous abortion.

Reference (Year)	**Study design, study population, and data collection**	**Assessment of NVP symptoms**	**Covariates adjusted for in multivariable model(s)**	**Main findings Adjusted ORs or RRs (95%CI)**
Klebanoff (1985)	**Design**: Prospective cohort **Population:** 56,000 pregnancies from Collaborative Perinatal Project **Study subjects**: 9,098 primiparous women enrolled in CPP within 14 weeks' gestation	Measured occurrence of vomiting since LMP	Maternal age, Maternal race, Education, Gravidity, Cigarette smoking, Pre-pregnancy weight	OR=0.7, p=0.002 (no 95% CI reported)
Tierson (1986)	**Design**: Prospective cohort **Population:** Women with singleton pregnancies in Albany, NY **Study subjects**: 414	Measured occurrence of NVP, onset and stop dates, and time of day in occurrence during interviews that took place at 12, 16, 20, 30, and 38 weeks	None reported	Study reported that women with no symptoms of nausea and/or vomiting experienced significantly greater proportion of SABs ($p<0.0015$). No adjusted estimates reported.
Weigel (1985)	**Design**: Meta-analysis **Population:** N/A **Study subjects**: 10 previous studies	N/A	N/A	OR=0.4 (0.3, 0.4)

Table 2 (cont). Summary of selected epidemiologic studies examining the association between nausea and vomiting in early pregnancy and spontaneous abortion.

Reference (Year)	**Study design, study population, and data collection**	**Assessment of NVP symptoms**	**Covariates adjusted for in multivariable model(s)**	**Main findings Adjusted ORs or RRs (95%CI)**
Weigel (1989)	**Design**: Retrospective cohort **Population:** Women with singleton births and delivered at UCLA **Study subjects**: 873	Measured occurrence of NVP since LMP at intake interview, and NVP since last visit at subsequent interviews	Maternal age, Maternal race/ethnicity, Housewife status, Anti-emetic medication	*Nausea only:* OR=0.32 (0.2, 0.7) *Nausea with vomiting:* OR=0.2 (0.1, 0.5)
Weigel (2006)	**Design**: Prospective cohort **Population:** Women in 1st trimester of pregnancy and planned to have prenatal care and delivery at the Hospital Patronato Municipal San Jose in Quito, Ecuador **Study subjects**: 849	Measurement occurrence, timing, intensity, peak and duration for nausea and vomiting at two time intervals (~11 weeks and weeks 24 to 28)	Maternal age, Neighborhood altitude, Cigarette smoking, Alcohol consumption, Antiemetic drugs, Vitamin-mineral supplements	*Nausea only:* OR=0.5 (0.2, 0.9) *Nausea with vomiting:* OR=0.7 (0.5, 1.0) Did not examine NVP with respect to its timing, intensity, peak, and duration

Figure 1. Directed Acyclic Graph showing causal relationship between nausea and vomiting in early pregnancy and spontaneous abortion, Right from the Start (2000-2004)

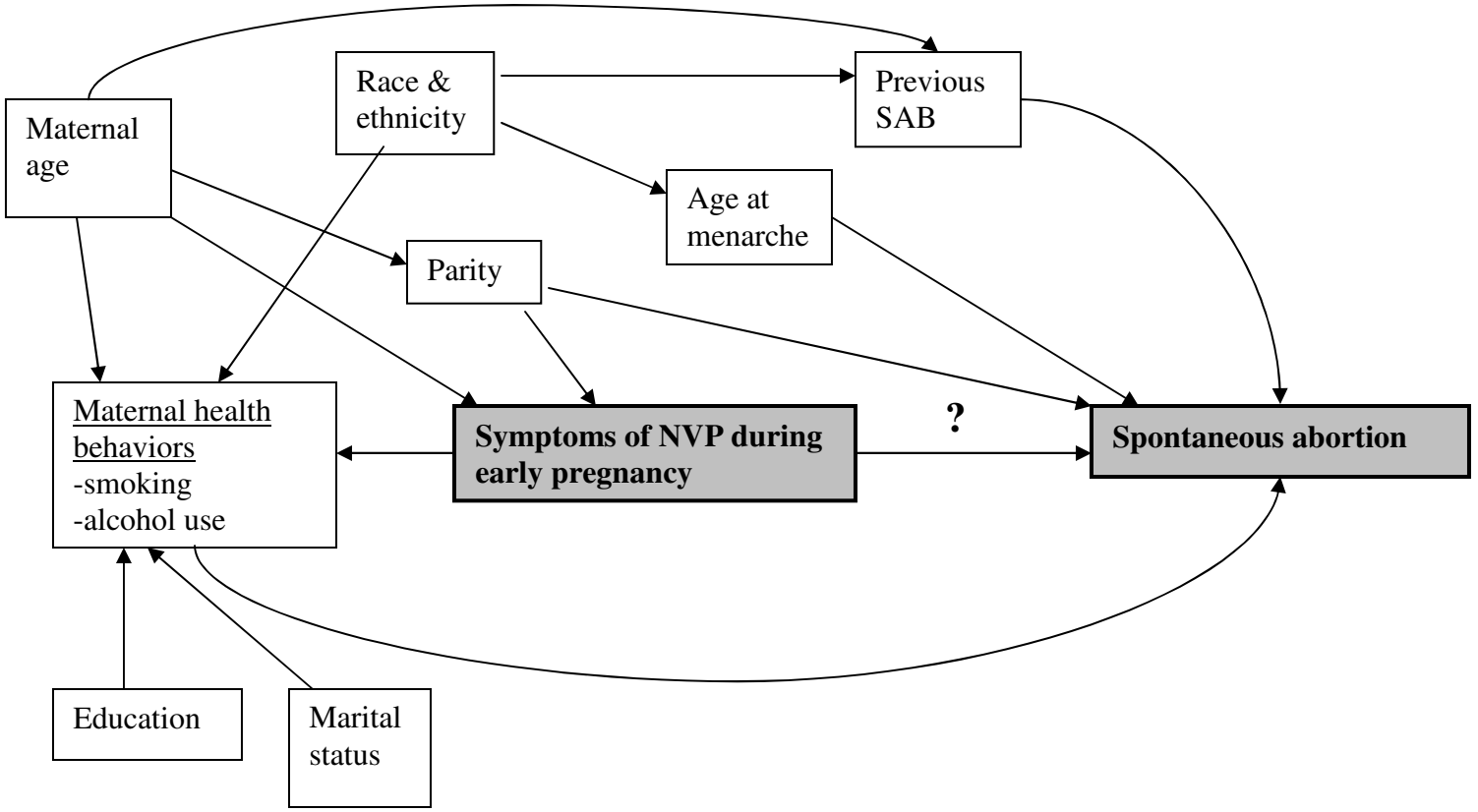


Table 3. Summary of selected epidemiologic studies examining the association between caffeine consumption and spontaneous abortion.

Reference (Year)	Study design, study population, and data collection	Assessment of caffeine consumption	Assessment of NVP symptoms	Covariates adjusted for in multivariable model(s)	Main findings Adjusted ORs or RRs (95%CI)
Axelsson (1989)	**Design:** Retrospective cohort **Population:** Women employed at Swedish hospital between 1980 and 1984 **Study subjects:** 769	Coffee consumption ascertained in cups for before and during 1st trimester of pregnancy	None collected		*Coffee before pregnancy:* >3 cups/day OR=1.3 (1.0, 1.8) *Coffee during 1st trimester:* >3 cups/day OR=1.6 (1.1, 2.3)
Fenster (1991)	**Design**: Case-control **Population:** Residents of Santa Clara, CA with SAB by 20 weeks' gestation **Study subjects**: 607 cases, 1,284 controls	Cases asked of events during entire pregnancy; controls for the first 20 weeks Caffeine consumption (coffee, tea, and soda)	Measured as "any" versus "none" Did not incorporate into analysis	Maternal age, Maternal race, Marital status, Insurance coverage, Cigarette smoking, Alcohol consumption, Previous SAB, Previous therapeutic abortion	*Total caffeine:* >300 mg/day OR = 1.2 (0.8, 1.9)
Kline (1991)	**Design**: Case-control **Population:** Cases-identified in both public and private facilities of NYC hospital; controls-women registered for prenatal care in same facilities **Study subjects**: 899 cases, 1,462 controls	Information collected during interview on caffeine consumption (coffee, tea, cocoa, and sodas) for "during pregnancy" and "peri-fertilization" (2 mos before and 1 mos after LMP).	Collected but did not specify symptom classification Did not incorporate into analysis	Maternal age, Maternal ethnicity, Cigarette smoking, Alcohol consumption	*Total caffeine:* >225mg/day during preg (normal karyotype) OR=1.9 (1.3, 2.6) >225mg/day during preg (abnormal karyotype) OR=1.6 (1.1, 2.3)

Table 3 (cont). Summary of selected epidemiologic studies examining the association between caffeine consumption and spontaneous abortion.

Reference (Year)	**Study design, study population, and data collection**	**Assessment of caffeine consumption**	**Assessment of NVP symptoms**	**Covariates adjusted for in multivariable model(s)**	**Main findings Adjusted ORs or RRs (95%CI)**
Parazzini (1991)	**Design:** Case-control **Population:** Cases with ≥2 SABs and no term pregnancy; controls were women admitted for normal delivery **Study subjects:** 94 cases, 176 controls	Single assessment of coffee intake during pregnancy	None collected		*Coffee during pregnancy:* Any coffee OR=1.4 (0.7, 2.6)
Infante-Rivard (1993)	**Design**: Case-control **Population:** Cases-fetal loss and controls with normal pregnancy in obstetric & pediatric hospital in Montréal, Quebec **Study subjects**: 331 cases, 993 controls	Information on caffeine consumption (coffee, tea, and cola) before pregnancy and during pregnancy collected via in-person interview of both cases and controls	None collected	Period of pregnancy, Maternal age, Education, Cigarette smoking, Alcohol consumption, Uterine abnormality, Work schedule	*Total caffeine (before preg):* 48 to 162 mg/day OR = 1.3 (0.9, 2.0) 163 to 321 mg/day OR=1.4 (0.9, 2.0) >321 mg/day OR =1.9 (1.2, 2.9) *(during preg)*: 48 to 162 mg/day OR = 1.2 (0.8, 1.6) 163 to 321 mg/day OR=2.0 (1.3, 2.9) >321 mg/day OR =2.6 (1.4, 5.0)

Table 3 (cont). Summary of selected epidemiologic studies examining the association between caffeine consumption and spontaneous abortion.

Reference (Year)	Study design, study population, and data collection	Assessment of caffeine consumption	Assessment of NVP symptoms	Covariates adjusted for in multivariable model(s)	Main findings Adjusted ORs or RRs (95%CI)
Dlugosz (1996)	**Design**: Prospective cohort **Population:** Women planning to deliver at Yale-New Haven Hospital **Study subjects**: 2,849	Caffeine consumption (coffee, tea, and soda) collected via interview before 16 weeks' gestation Total caffeine derived by: -107mg/cup (coffee) -34 mg/cup (tea) -47 mg/can (soda)	None collected	Maternal age, Maternal race/ethnicity, Education, Gravidity, Parity, Cigarette smoking, Alcohol consumption, Prior SAB, Prior TAB, Pre-pregnancy weight, Age at menarche	*Total caffeine:* 1-150 mg/day OR=0.81 (0.5,1.2) 151-300 mg/day OR=0.89 (0.5,1.6) >300 mg/day OR=1.8 (0.9, 3.5) *Coffee:* ⩾3 cups/day: OR = 2.6 (1.3, 5.3)
Fenster (1997)	**Design**: Prospective cohort **Population:** Women receiving health care at 3 Kaiser Permanente facilities in CA **Study subjects**: 5,144 pregnant women	1st trimester interview on caffeine consumption (coffee, tea, and soda, cups or cans/day) before and during pregnancy Total caffeine derived by: -107mg/cup (coffee) -34 mg/cup (tea) -47 mg/can (soda)	Measured as "any" versus "none" Did not incorporate into analysis	Maternal age, Maternal race, Pregnancy history, Cigarette smoking, Alcohol consumption, Employment, Gestational age at interview, Marital status, Socioeconomic status	*Total caffeine:* >300mg/day during preg OR=1.3 (0.8, 2.1) >300mg/day before preg. OR=1.3 (0.9, 1.7) *Coffee:* 2 cups/day before preg OR=1.1 (0.8, 1.5) ⩾3 cups/day before preg OR=1.1 (0.8, 1.6) 2 cups/day after preg OR=1.5 (1.0, 2.2) ⩾3 cups/day after preg OR=0.8 (0.4, 1.7)

Table 3 (cont). Summary of selected epidemiologic studies examining the association between caffeine consumption and spontaneous abortion.

Reference (Year)	Study design, study population, and data collection	Assessment of caffeine consumption	Assessment of NVP symptoms	Covariates adjusted for in multivariable model(s)	Main findings Adjusted ORs or RRs (95%CI)
Parazzini (1998)	**Design**: Case-control **Population:** Cases-admitted for SAB in Milan; controls-women who gave birth at term in same hospitals where cases were identified **Study subjects**: 782 cases, 1,542 controls	Caffeine consumption (coffee, tea, and cola) was collected during interview	Measured in terms of occurrence and intensity (none, low, moderate, or serious); only ~50% of study population had NVP data because collection started midway through study period	Maternal age, Education, Previous live births, Previous SAB, Alcohol consumption, Cigarette smoking, Nausea intensity in 1st trimester	*Coffee:* 2-3 cups/day in 1st trimester OR=1.8 (1.2, 2.3) ⩾4 cups/day in 1st trimester OR=4.0 (2.6, 6.2)
Cnattingius (2000)	**Design**: Case-control **Population:** Cases-identified from Dept. OB/GYN at Uppsala University, Sweden; controls—antenatal clinics in Uppsala County **Study subjects**: 562 cases, 953 controls	Information on weekly caffeine consumption (coffee, tea, cocoa, chocolate, sodas, and medications), starting 4 weeks before LMP until most completed week gestation	Nausea (never, sometimes but not daily, daily but not all day, daily all day) Vomiting (never, sometimes but not daily, daily) Scores were assigned based on responses; multivariable analysis adjusted for presence/absence of NVP symptoms	Maternal age, Cigarette smoking, Alcohol consumption, Gravidity History of SAB, NVP	No overall main results *Total caffeine among non-smokers*: (normal karyotype) 300-499 mg/day OR = 1.8 (0.8, 3.8) ⩾500 mg/day OR= 2.2 (0.8, 6.4) (abnormal karyotype) 300-499 mg/day OR = 0.9 (0.5, 1.7) ⩾500 mg/day OR= 1.8 (0.8, 3.9)

Table 3 (cont). Summary of selected epidemiologic studies examining the association between caffeine consumption and spontaneous abortion.

Reference (Year)	Study design, study population, and data collection	Assessment of caffeine consumption	Assessment of NVP symptoms	Covariates adjusted for in multivariable model(s)	Main findings Adjusted ORs or RRs (95%CI)
Wen (2001)	**Design**: Prospective cohort **Population:** Women enrolled in HMO in Twin Cities, MN **Study subjects**: 650	Caffeine consumption (coffee, tea, soda, hot chocolate, other foods with chocolate) was collected via mailed-in monthly food frequency questionnaire -139.2 mg/cup (coffee) -64.0 mg/cup (tea) -46.0 mg/can (cola)	Measured as nausea status "yes" vs. "no" and its duration (days) No measurement for vomiting episodes	Found no differences in association between unadjusted and adjusted analysis, only unadjusted RRs were presented	*Total caffeine* *With or without nausea:* 100-299 mg/day RR = 2.0 (1.0, 4.1) ⩾300 mg/day RR =2.5 (1.0, 6.4) *With nausea:* 20-99 mg/day RR = 2.5 (0.5, 11.3) ⩾100 mg/day RR =0.6 (0.1, 3.9) *Without nausea:* 20-99 mg/day RR = 1.1 (0.4, 2.8) ⩾100 mg/day RR =1.6 (0.6, 4.2)
Giannelli (2003)	**Design**: Case-control **Population:** Cases-nulliparous women with confirmed SAB in UK; controls-nulliparous women attended prenatal care **Study subjects**: 160 cases, 314 controls	Cases, on average, were interviewed 3 weeks after controls Measured caffeine consumption (coffee, tea, and cola) before and during pregnancy	Measured occurrence and classified by severity: mild, moderate, or severe	Maternal age, Nausea severity, Gestational age	*Total caffeine (during pregnancy):* 301-500 mg/day OR=1.9 (1.0, 3.6) >500 mg/day OR = 2.2 (1.1, 4.4) *Coffee (during pregnancy):* 107-321 mg/day OR=2.2 (1.3, 3.6) >321 mg/day OR = 2.3 (1.2, 4.4)

Table 3 (cont). Summary of selected epidemiologic studies examining the association between caffeine consumption and spontaneous abortion.

Reference (Year)	Study design, study population, and data collection	Assessment of caffeine consumption	Assessment of NVP symptoms	Covariates adjusted for in multivariable model(s)	Main findings Adjusted ORs or RRs (95%CI)
Tolstrup (2003)	**Design**: Nested case-control **Population:** Random sample from general population of Copenhagen, Denmark **Study subjects**: 303 cases, 1,381 controls	Caffeine consumption (coffee and tea) before pregnancy was collected at enrollment interview	None collected	Maternal age, Marital status, Cigarette smoking, Pregnancy history	*Total caffeine (before pregnancy):* 75-300 mg/day OR=1.3 (0.8, 2.1) 300-500 mg/day OR= 1.5 (0.9, 2.4) 501-900 mg/day OR= 1.4 (0.9, 2.4) >900 mg/day OR= 1.7 (1.0, 3.0)
Bech (2005)	**Design**: Prospective cohort **Population:** Danish National Birth Cohort **Study subjects**: 88,482	Coffee consumption (number of cups drank/day) was collected during interview	None collected	Maternal age, Parity, Pre-pregnancy BMI, Cigarette smoking, Alcohol consumption, Socio-occupational status	*Coffee (during pregnancy):* 4-7 cups/day: HR=1.2 (0.9, 1.6) ⩾8 cups/day: HR =1.5 (1.0, 2.2)

Table 3 (cont). Summary of selected epidemiologic studies examining the association between caffeine consumption and spontaneous abortion.

Reference (Year)	Study design, study population, and data collection	Assessment of caffeine consumption	Assessment of NVP symptoms	Covariates adjusted for in multivariable model(s)	Main findings Adjusted ORs or RRs (95%CI)
Savitz (2008)	**Design:** Prospective cohort **Population:** Women recruited from three US cities who were early in pregnancy or trying to conceive a pregnancy **Study subjects:** 2,407 women with singleton pregnancies	Caffeinated coffee, caffeinated iced tea, caffeinated hot tea, and caffeinated soda were collected at time of interview (number of cups, cup size, brewing methods, timing of change in consumption, consumption amount after change) Assessment at three time points: pre-pregnancy, four weeks post-LMP, and around time of telephone interview	Symptoms of nausea and vomiting were collected separately for severity (NVP status), time of onset and stop dates	Maternal age Maternal education Maternal race/ethnicity Marital status Alcohol use Vitamin use NVP	*Coffee* *(at time of interview)* >0 to <200.8mg/day OR=1.2 (0.7, 1.9) ⩾200.8 mg/day OR=1.0 (0.6, 1.6) >372.9 mg/day OR=0.7 (0.3, 1.8) *Total caffeine* *(at time of interview)* >0 to <144.3 mg/day OR=1.0 (0.7, 1.4) ⩾144.3 mg/day OR=1.2 (0.9, 1.7) >273.2 mg/day OR=1.3 (0.9, 1.9)
Weng (2008)	**Design:** Population-based prospective cohort **Population:** Kaiser Permanente Medical Care Program members who resided in the San Francisco areas and who had a positive pregnancy test in a Kaiser facility **Study subjects:** 1.063	Caffeinated coffee, caffeinated tea, caffeinated soda at the time of the in-person interview (types of drinks, timing of initial drink, frequency and amount of intake, whether changed consumption patterns since pregnancy, intake amount after change)	NVP (yes/no)	Maternal age Maternal race Maternal education Household income Marital status Smoking Alcohol use Jacuzzi use EMF NVP	*Caffeine* <200mg/day aHR=1.4 (0.9, 2.2) ⩾200mg/day aHR=2.2 (1.3, 3.7) *Coffee only* <200mg/day aHR=1.2 (0.6, 2.2) ⩾200mg/day aHR =2.5 (1.2, 5.1)

CHAPTER 3: METHODS

3.1 Overview of Right from the Start study

The Right from the Start (RFTS) study was a multi-site, prospective epidemiologic cohort study that recruited pregnant women and women, who were planning a pregnancy between 2000 and 2004 from three distinct metropolitan areas in three states, in prenatal care clinics and directly from the communities. The primary study aim of RFTS was to examine the effects of drinking-water disinfection by-products on SAB.[82, 83] In the parent study, endovaginal ultrasound was used to collect information on the health of the fetus and telephone interviews were conducted at two time intervals to ascertain behavioral and relevant pregnancy-related information. Fibroid sub-study participants received additional endovaginal ultrasounds and participated in an additional telephone interview; whereas women in the blood trihalomethane biomarker study each completed a 24-hour water activity diary and provided the study with 12 mL of tap water from their residence and 10 mL of blood during a home visit.

This study included 2,430 women whose pregnancies started between 2000 and 2004 based on the LMP. Women included in this analysis were those who completed at least the baseline interview and had a valid last menstrual period date.

Data on NVP symptoms were collected in the telephone interviews, before 16 weeks' and between 22-26 weeks' gestation. All study participants were asked by telephone about their experiences with nausea symptoms. Information on vomitin

episodes, if any, was ascertained only from women with symptoms of nausea. Symptom duration was recorded independently for nausea and vomiting and severity was asked in regard to vomiting episodes. Participant self-reports, medical records abstraction, and linkages with vital records were use to determine and to confirm pregnancy outcomes.

Also, at the time of the telephone interview, information on daily consumption for "a typical week" was collected on coffee, iced and hot tea, and sodas. Questions were asked about all sources, including the number of cups consumed daily, beverage sizes, and brewing methods for coffee. We also collected information regarding changes in intake on each caffeinated beverage, if applicable, with respect to the timing of change and consumption amount prior to the change.

3.2 Overview of proposed research using Right from the Start data

With the data from RFTS study, we examined the risk factors or markers for the occurrence of NVP and evaluate the association between NVP and SAB. Additionally, we used this data to illustrate the application of propensity score models in pregnancy health research by applying the data and methods in studying the effects of maternal caffeine consumption and SAB. Maternal age, maternal race/ethnicity, marital status, education, cigarette smoking, alcohol use, parity, plurality, pregnancy loss history, and age at menarche were assessed as potential risk factors or markers for NVP. These predictors were chosen because they were shown in the literature as some of the known risk factors for NVP. Modified Poisson regression with robust error variance was applied to examine potential risk factors for NVP. Discrete-time hazard models were employed to describe the association between NVP and SAB, with covariates maternal age, maternal

race/ethnicity, marital status, education, cigarette smoking, alcohol use, age at menarche, parity, and previous pregnancy loss considered as potential effect measure modifiers or potential confounders. Finally, in the propensity score analysis, a subset of the covariates listed above was used to estimate the probability of caffeine use at three time points: pre-pregnancy, four weeks post-LMP, and around the time of telephone interview. We then evaluated the association between caffeine exposure at three time points and the risk for SAB by using discrete-time hazard models.

3.3 Study sites selection and participant recruitment

3.3.1 Selection of study sites and study eligibility criterion

Study sites for RFTS were chosen because of their differences in drinking water disinfection by-products (DBPs) characteristics. Women from three U.S. metropolitan cities (Raleigh, North Carolina; Memphis Tennessee; and Galveston, Texas) who were eligible to participate in this research study must have been 18 years of age or older and less than 12 completed weeks pregnant, or women between the ages of 18 to 45 and who were trying to become pregnant for less than six months. All participants lived within the city limits of the study sites and did not plan to move out of the study areas prior to delivery. For participants who were pregnant, they must have had a positive pregnancy test and must not have taken any medications prescribed to become pregnant. If she did not have a positive pregnancy test, she was not considered pregnant by the study until she notified the study staff with a positive pregnancy result. Additionally, participants must be able to speak, read, and write in English. If a participant met all eligibility criteria but had unknown date of LMP (last menstrual period), an ultrasound EGA (estimated

gestational age) was used to determine her eligibility. Summary of participant recruitment, enrollment, and women who are included in the final analysis are shown in Figures 2 to 4.[83]

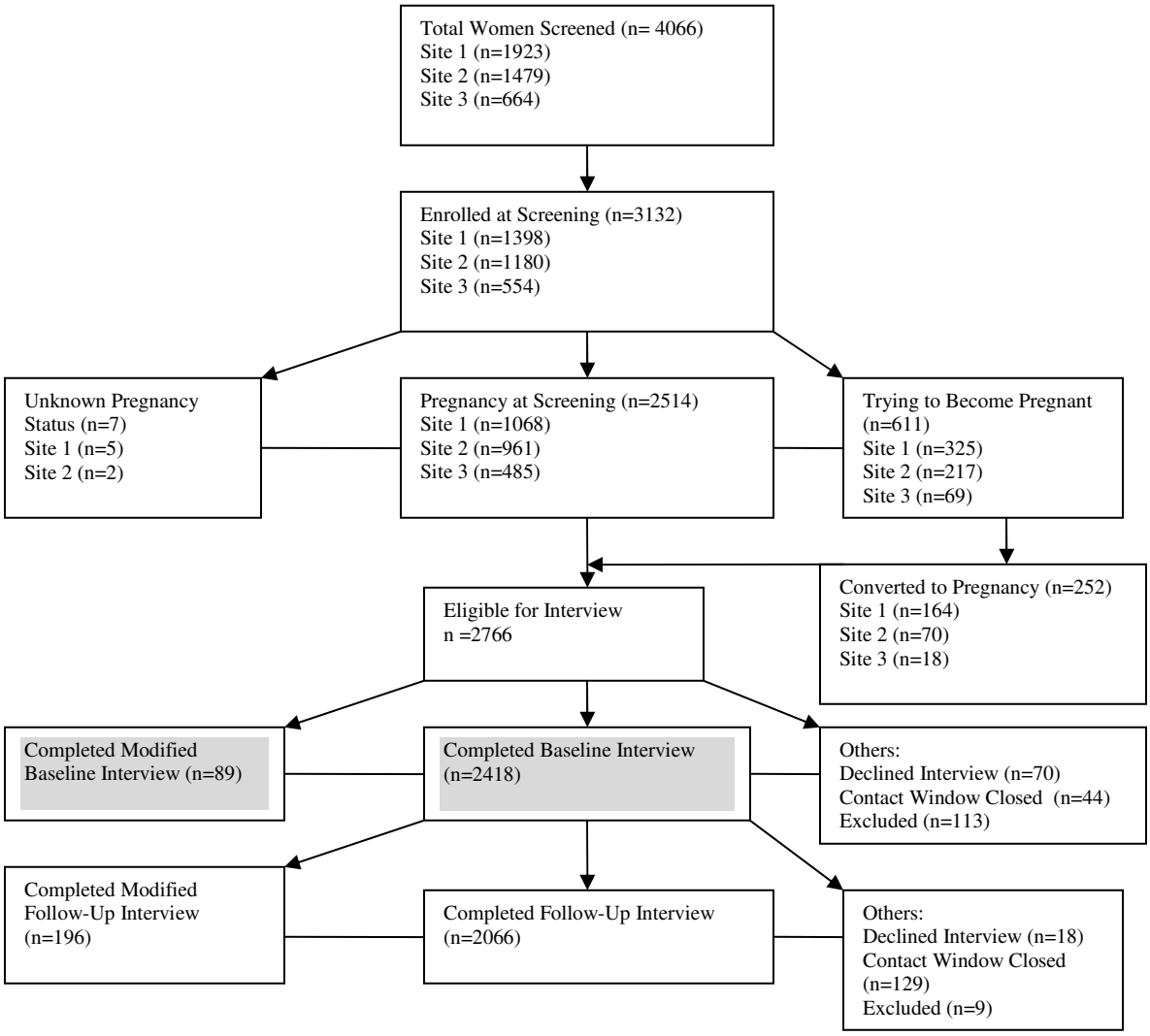


Figure 2. Recruitment and enrollment summary, Right from the Start (2000-2004)
Adapted from AwwaRF report (Savitz et al., 2005)

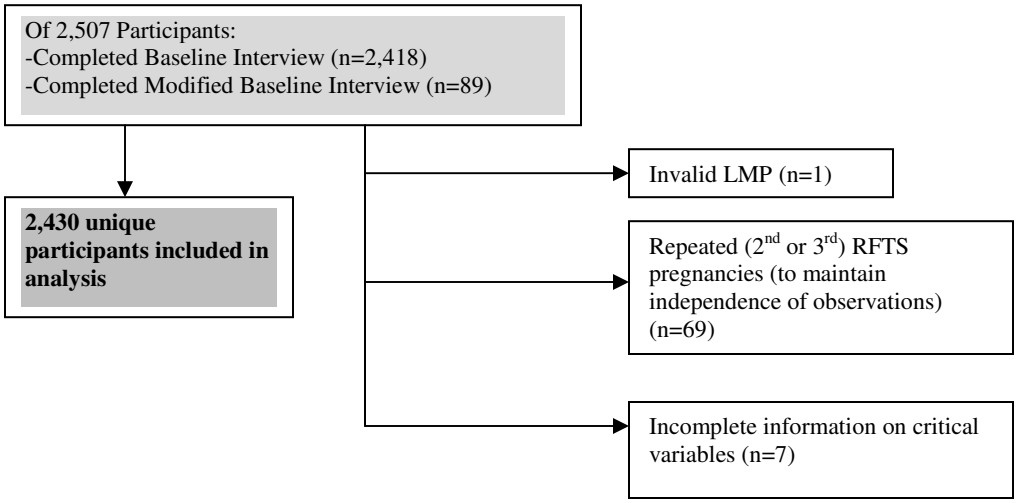


Figure 3. Construction of analysis data set for aim 1 and aim 2 analyses: Right from the Start (2000-2004). Adapted from AwwaRF report (Savitz et al., 2005)

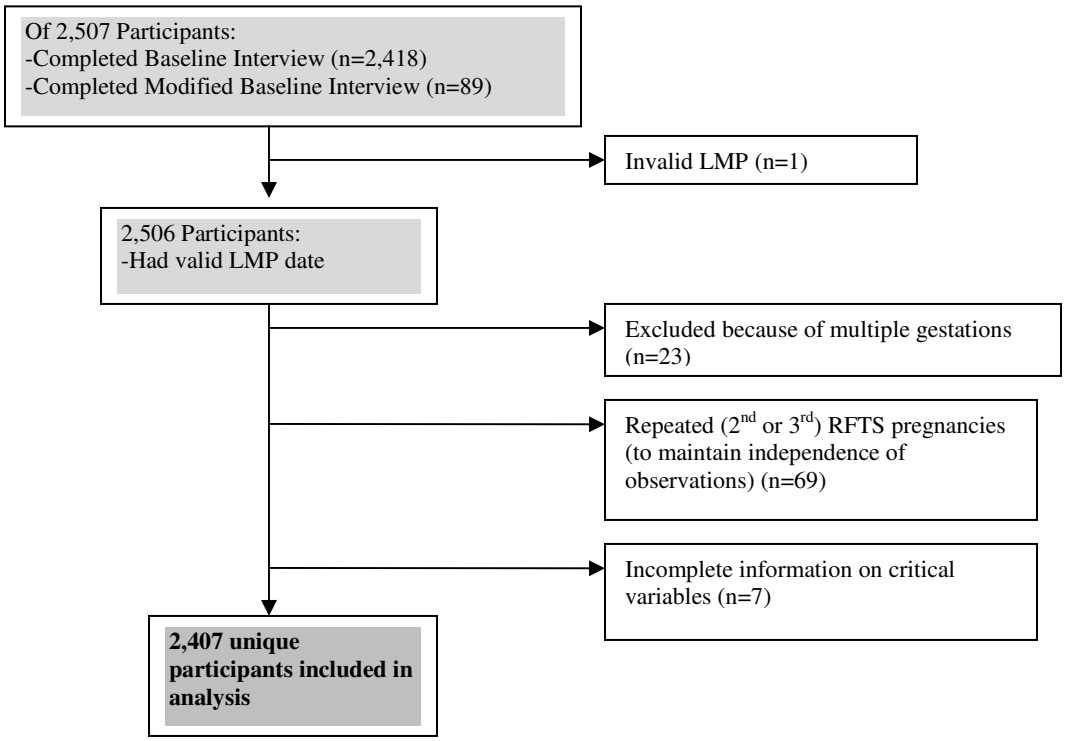


Figure 4. Construction of analysis data set for aim 3 analysis: Right from the Start (2000-2004). Adapted from AwwaRF report (Savitz et al., 2005)

3.3.2 Recruitment of potential study participants

Several strategies were developed to facilitate the identification of potential participants and recruitment of participants—recruitment through medical practices and community-based recruitment. Cumulative result of RFTS recruitment strategies is shown in Figure 5[83]

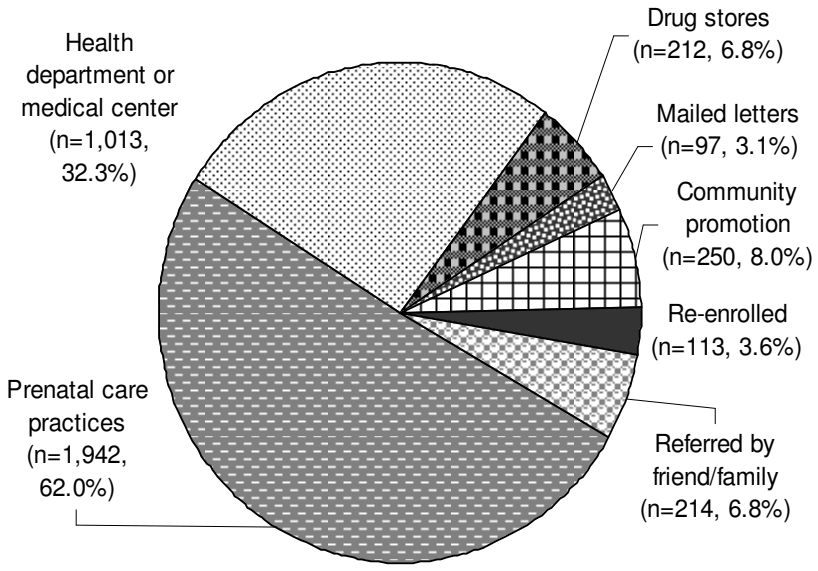


Figure 5. Recruitment sources for all study sites*, Right from the Start (2000-2004)
*Participants can be recruited from multiple avenues. Adapted from AwwaRF report (Savitz et al., 2005)

Recruitment through medical practices

Through collaborative relationships with 50 private obstetrics and gynecology clinics, public prenatal care sites including county health department clinics and university medical centers, RFTS study staff was able to identify and recruit women who sought preconception counseling, prenatal care services, or pregnancy testing, into the study. Study coordinators

worked with each medical practice to develop individualized plans for recruitment and collaboration to best accommodate the differences in clinic flow and patient populations amongst the practices. Brochures and fliers with information about the study were posted in the waiting rooms and examination rooms, and contact information for interested women was collected and then transferred to The University of North Carolina at Chapel Hill (UNC). In addition to displaying promotional materials, many clinics discussed RFTS with their patients; clinic staff collected contact information for all interested women to be faxed to the study office or forwarded their patient calls directly to the study office if a woman were interested in learning more about the study. Regardless of the mode of contact, study staff further explained the study and invited interested women to complete the screening interview.

Community-based recruitment

RFTS utilized multiple avenues in the communities to generate awareness about the study throughout the three study sites. Fliers with tear-off tables were regularly posted at local childcare facilities, grocery stores, public libraries, churches, drug stores, book stores, retail stores, beauty salons. Brochures and information cards were displayed at the waiting and examination rooms of pediatric and family practices, and promotional materials were also distributed by study coordinators at local health fairs and community events. Rebate coupons for home pregnancy test kits were available at local drug stores and eligible women who contacted RFTS and enrolled into the study redeemed a $5 coupon. Letters describing the study and pregnancy test rebate were mailed to two target groups: new home owners in the study areas and women who delivered a child within the

past three years. All study-related promotional and information materials included the study's toll-free telephone number and the study's web site address for potential participants to contact the study staff. When women contacted RFTS study office, study staff provided information about the study, explained the study protocol and invited them to complete a screening interview to determine eligibility.

3.4 Data collection

3.4.1 Screening interview

A short telephone screening interview was conducted with potential study participants to gather information (e.g. women's date of birth, residential address, LMP, pregnancy status, pregnancy-related medication use) that was necessary to establish eligibility for RFTS study. If a woman was deemed eligible to participate in the study and she agreed to participate in RFTS, additional information was collected and an endovaginal ultrasound was scheduled.

3.4.2 Participant consent

The RFTS research team employed a vigorous consent process to ensure that all study protocols was explained thoroughly and that all potential study participants understood all parts of the study activities. All study participants provided a signed consent prior to their first trimester ultrasound or participating in any study activities.

3.4.3 Monetary and non-monetary participant incentives

Study participants were eligible to receive numerous incentives from the study. The amount of incentives varies depending on the type of study activity completed by each participant in the main study and/or in one of the sub-studies. Incentives included six free pregnancy test kits for women who enrolled into the study while trying to conceive, one free endovaginal ultrasound to all newly pregnant women, and monetary incentives in varying amount for the completion of one and both telephone interviews ($10 for baseline or $20 for modified baseline interview and $10 for follow-up or modified follow-up interview). An additional $10 was given to women who completed all study activities in the main study.

3.4.4 First trimester endovaginal ultrasound

First trimester endovaginal ultrasounds were conducted around eight weeks' gestation (median gestational age at ultrasound=9.3 weeks) by ARDMS® certified clinical sonographers, using state-of-the-art ultrasound technology. All clinical sonographers were required to attend study-specific training and a trained study staff reviewed all data collection forms and still images prior to data entry. We collected measurements on the gestational sac, yolk sac, fetal pole, and fetal heart rate for use in clinical pregnancy dating.

3.4.5 Study telephone interviews

Some sociodemographic information (e.g. maternal age, mother's highest educational attainment, maternal race) were obtained during the screening interview. As part of the study activities, participants were asked to complete two telephone interviews

that took place at two different time intervals in the course of the study. The baseline interview took place around one to two weeks after the screening interview and no later than 16 completed weeks' gestation. This interview ascertained information on a range of topics such as:

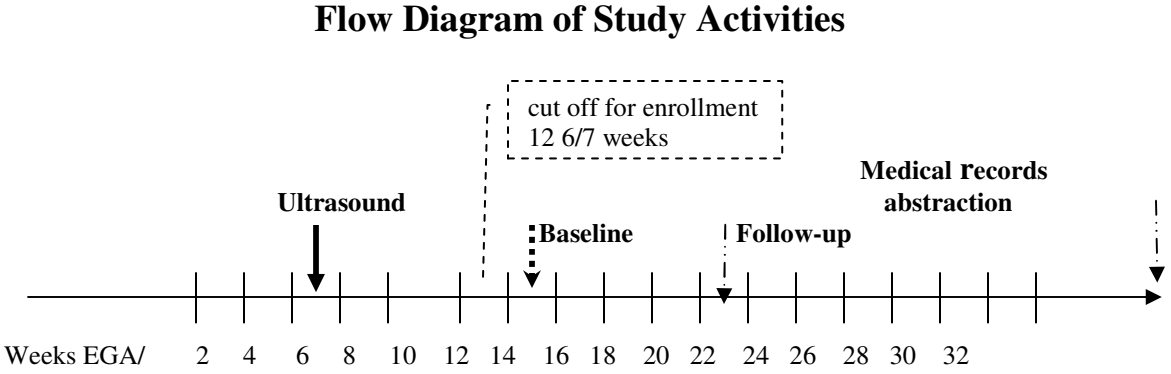
- Employment history
- Health behaviors (water, alcohol and caffeine consumption, tobacco and illicit drug use, physical activity, and time in pool, hot tubs, and Jacuzzi)
- Water exposure characteristics
- Current and previous reproductive histories (menstrual history, previous pregnancy history, time to conception, NVP, vaginal bleeding, infections, diabetes, medications taken to prevent pregnancy complications, and prenatal care)
- Physical and sexual violence
- Vitamin and mineral supplement use
- Other social, household, and income information.

Participants were contacted between 22 to 26 weeks of their pregnancy for the follow-up telephone interview. This interview gathered additional information on the following topics:

- Changes in health behaviors and water exposures

- Current and past pregnancy and medical histories (diseases and injuries, surgical procedures, fever, and other health problems)
- Pregnancy-related symptoms and information regarding prenatal care and delivery choices, and partner information.
- Paternal information (age, race/ethnicity, education)

Study participants who had been identified as having pregnancy losses during either phase of the telephone interviews were asked to complete the modified version of the baseline or follow-up interview. The modified interviews ascertained identical information as those in the "regular" interviews, but the language of the modified version took into account the sensitivity of the recent pregnancy loss. Figure 6 shows a flow diagram of RFTS study activities.[83]



IDEAL TIME FRAMES AND DEADLINES FOR ULTRASOUNDS AND CATI

All Participants

First Trimester Ultrasound ideal time: 6 2/7 – 7 5/7 weeks; no later than 14 0/7 weeks
Baseline CATI: preferably within 2 weeks of enrolling and no later than 16 0/7 weeks
Follow-up CATI: 20 0/7 – 25 0/7 weeks

Figure 6. Diagram of study activities, Right from the Start (2000-2004)
*Adapted from AwwaRF report (Savitz et al., 2005)

3.4.6 Medical record reviews and vital record linkages

Participants' medical records were reviewed and relevant information was abstracted at the end of pregnancy, whether it resulted in a loss or in a live birth. Information such as first day of the LMP, gestational age at delivery, pregnancy outcome, and other maternal demographic status from the medical chart abstraction was used to clarify discordance and to supplement the data collected during the telephone interviews. Additionally, vital and fetal death records were linked with existing participant data to confirm and/or to provide supplemental information on pregnancy outcomes. Infant's birthweight, date of birth, and gender were key pieces of information abstracted from the vital records and similar information on the infant was obtained from the fetal death records.

Because of the comprehensive nature of RFTS, multiple sources including participant self-report, vital records, and medical records, were used to gather participant and pregnancy-related information; information collected such as pregnancy outcome and pregnancy outcome date, abstracted data from medical records took precedence over data from vital records, which took precedence over self-reported data.

3.5 Collection of NVP information

3.5.1 NVP assessment

Information on NVP was collected via self-report from all eligible participants at two time intervals, the baseline (Appendix A) and follow-up (Appendix B) telephone interviews. For this study, symptoms of nausea was defined as 'nausea or feeling sick to

their stomach at any time' during the index pregnancy and a vomiting episode was defined as a woman who 'had nausea so bad that she vomited'.

Information regarding the date of onset and stop date for nausea during pregnancy was collected on all women at the baseline and follow-up telephone interviews. Similarly, information on the date of onset and stop date on vomiting was also gathered during the same two time intervals; however, questions pertaining to vomiting episodes were ascertained only from women who experienced symptoms of nausea.

Onset and stop dates for NVP symptoms were collected down to the specific month, day, and year for each woman, and separately for symptoms of nausea and for vomiting episodes. If the women were unsure of their "day" of event, they were asked one additional question to describe the timing of occurrence with respect to the "week in the month". For women who experienced vomiting along with nausea symptoms, two more questions were used to collect information on the number of times per day and per week she vomited when her "vomiting was the worst" to provide further indication on the intensity of NVP.

Additionally, women who experienced symptoms of NVP were asked about changes in their health behaviors during their pregnancy such as coffee, soft drinks, and water consumption, physical activity level, change in intake of prenatal vitamin or supplements, cigarette smoking, etc. as a direct result of NVP symptoms at the baseline telephone interview. These questions on health behavioral changes were tailored specifically to each woman based on her reported activities and/or consumption of beverages reported in the earlier sections of the interview. For example, if a woman who experienced NVP symptoms and reported consuming coffee during her pregnancy, her

health behavior modification questions included one on whether her coffee intake increased or decreased as a direct result of her NVP symptoms. Similarly, a non-consumer with NVP symptoms was not asked about her changes in coffee consumption.

3.5.2 Variable creation and coding for NVP

Information collected on symptoms of NVP from the telephone interviews was used for the coding and creation of variables to indicate NVP characteristics in this study population.

Symptom severity

Severity of women's experience with NVP symptoms was categorized into three levels of "no symptoms", "symptoms of nausea only", and "symptoms of nausea and vomiting episodes". Women were considered to have no symptoms if they did not experience any symptoms of nausea and any vomiting during the pregnancy. They were classified into the second group, "having symptoms of nausea only", if they reported experiencing symptoms of nausea, without any vomiting episodes as a result of the nausea symptoms. Lastly, women were classified as having "symptoms of nausea and vomiting episodes" if they reported experiencing both symptoms of nausea and vomiting during their pregnancy.

Assigning timing of onset and stop date

As noted in the earlier section of this proposal, information pertaining to NVP symptoms was ascertained at both telephone interviews; nevertheless, these questions asked women to make references back to the symptoms that they experienced during the

early gestational period. That is, women were asked to report the exact dates of symptoms onset and the exact stop dates for nausea and/or for vomiting episodes that occurred during early pregnancy at Time 1 (baseline interview) and they were subsequently asked to best recall the same information in regard to their NVP symptoms for the early pregnancy period at Time 2 (follow-up interview).

Because the primary focus were symptoms of nausea and vomiting during the period of early pregnancy (approximately four weeks after the LMP), we made the assumption that dates of onset and stop dates given at Time 1 were reported more accurately than those given for the same events at Time 2. Therefore, dates of onset and stop dates for each symptom reported at the baseline interview took precedence over the dates reported in the follow-up interview, in order to minimize misclassification and recall bias. An exception was made to use the date(s) reported in the follow-up interview if a woman reported experiencing either nausea or vomiting episodes, but was unable to provide the corresponding dates of onset and/or stop dates at the baseline interview but one was provided during the follow-up interview.

The onset and stop dates for nausea symptoms and vomiting episodes were collected not only in calendar year and the month of occurrence but also for the "day" of the month, so we were able to almost pinpoint the date of onset and stop dates for these symptoms with respect to the gestational period, down to the number of days from the LMP. To further reduce misclassification in the assignment for timing of NVP, women were asked to report the week in the month in which they experienced either nausea symptoms or vomiting episodes if they were not able to provide the exact "day" of symptom occurrence. In scenarios where the "day" of event was unknown but the "week

in month" estimate was reported, we were able to impute the "day" of the month according to the week (1^{st}, 2^{nd}, 3^{rd}, 4^{th}, and 5^{th}) in the month. This allowed us to approximate the onset and stop dates of NVP in relation to the LMP and to classify the dates relative to the gestational period. Using an algorithm, the "day" of the month corresponding to the "week in the month" is imputed as follow:

First week = Day 4

Second week = Day 11

Third week = Day 18

Fourth week = Day 25

Fifth week = Day 30

Unknown week= Day 15

Time of symptom onset and symptom duration.

Time of onset was subdivided into "typical onset", defined into out study as symptom onset that took place prior to the start of the second trimester ($\leqslant$13 weeks gestation) and "delayed onset", which was defined as symptom that commenced after the first trimester (>13 trimester gestation). Symptom duration was calculated using dates of symptom onset and stop dates for nausea and vomiting episodes. Using the available information, duration was categorized as "symptoms restricted within first trimester" and "symptoms lasting beyond first trimester" in Manuscript 1 or "short", "moderate", and "severe" in Manuscript 2. We considered gestational weeks 0 to 12 as first trimester, weeks 13 to 26 as the second trimester, and week 27 and greater was assigned as the third trimester.

3.6 Collection of caffeine intake information

Information on the daily consumption of caffeine from coffee and from non-coffee beverages for a typical week (study participant not on vacation or experience unusual event(s) or circumstance(s) that would affect caffeine consumption) was ascertained during the telephone interview (Appendix C). Detailed questions were asked of caffeinated coffee, with respect to the number of cups consumed daily on a typical week, beverage sizes, and brewing method (brewed versus instant). We considered cup sizes for caffeinate coffee and caffeinated hot tea to be small (7oz), medium (13oz), and large (20oz). Iced tea was categorized as small (7 oz), medium (16 oz), and large (28 oz), and sodas as small (10 oz), medium (18 oz), and large (29 oz). Daily caffeine exposure (mg) was calculated based on amount consumed (according to the number of cups and cup sizes) and brewing methods. Bunker's caffeine contents were then applied to the amount consumed to determine the total daily exposure value. We applied Bunker's parameter of caffeine content [84] to calculate the total exposure value (mg/day). Caffeine contents (mg/oz) for each type of caffeinate beverage were assigned as follow:

Brewed caffeinated coffee =28.6869 mg/oz

Instant caffeinated coffee =13.0108 mg/oz

Half brewed and instant caffeinate coffee = 20.8489 mg/oz

Unknown caffeinated coffee = 26.7688 mg/oz

Hot and iced caffeinated tea = 9.7581 mg/oz

Caffeinated sodas = 3.4881 mg/oz

In addition to collecting data on women's "current consumption" during the interviews, we also considered changes in consumption, if any, since their LMP on each caffeinated beverage. If changes in consumption occurred, timing of changes in month, date, and year were determined via women's self-report. Because we collected information on changes in consumption not only in calendar year and the month of occurrence but also for the "day" of the month, we were able to almost pinpoint the date of change with respect to the gestational period, down to the number of days from the LMP. In scenarios where the "day" of event was unknown but the "week in month" estimate was reported, we imputed the "day" of the month accordingly and unknown week was assigned the mid-month value, using the same algorithm that was used to assign the "day" of the month for the date of onset and stop date for NVP.

3.7 Collection of SAB information

SAB is typically defined as pregnancy loss prior to 20 completed weeks' gestation. In RFTS, SAB is classified as loss of pregnancy if the end of pregnancy occurred on or before 140 days gestation as calculated by the last menstrual period. The RFTS cohort (n=2,430) had 88.8% (n=2,158) live births, 0.5 % fetal deaths (n=12), and 10.7% (n=260) resulted in SAB. Regardless of the pregnancy outcome, the outcome information was obtained through participant self-report, review of medical records abstraction from prenatal care provider offices, university hospitals, emergency rooms, and other health care settings, and through data linkages to vital records to document live birth or fetal death.

3.7.1 Calculation of gestational age

Whether for a loss or a delivery, the first step in assigning outcomes was to estimate gestational age at pregnancy outcome. In the RFTS cohort, we found self-reported LMP to be highly reliable in dating pregnancy onset; therefore, gestational age in this study population was primarily determined based on women's self-reported LMP and secondarily with the use of ultrasound dating. If both self-reported LMP and LMP from ultrasound dating were available and the two dates agreed within seven days, the self-reported LMP was used to calculate gestational age; however, information from ultrasound dating was used if they differed for more than seven days or if a woman's LMP were unknown or were deemed unreliable.

3.8 Measurement of covariates of interest

Maternal age, maternal race/ethnicity, marital status, education, cigarette smoking, alcohol use, age at menarche, parity, plurality, and pregnancy loss history were examined as potential risk factors or markers for NVP in the analysis for Manuscript 1. Manuscript 2 examined the association between NVP and spontaneous abortion; many covariates were evaluated as effect measure modifiers or potential confounders and the DAG[51] (Figure 1) was drawn to provide a visual assessment on the covariates of interest. Manuscript 3 focused on the association between caffeine exposure and SAB. A list of covariates, similar to those used in Manuscripts 1 and 2, was used in Manuscript 3 as potential confounders in the traditional covariate adjustment analysis and to estimate individual propensity score for the propensity score analysis. Covariates of interest that

were used in the analysis of Manuscripts 1, 2, and 3 and the respective variable coding schemes are shown in Tables 4, 5, and 6.

Table 4. Variables included in the analysis for Manuscript 1: Risk factors or markers for nausea and vomiting in early pregnancy

Variable Name	Description	Variable Coding
Study Outcome		
Nausea and vomiting during early pregnancy	Self-reported experiences with NVP during the baseline and follow-up telephone interviews	
	Severity:	0= Any symptoms 1= No symptoms
	Time of symptom onset:	0=Typical onset 1=Delayed onset
	Duration:	0=Symptoms restricted to first trimester 1=Symptoms lasting past first trimester
Predictors		
Age	Maternal age at study enrollment	1= 18-24 years old 2= 25-29 years old 3= 30-34 years old 4= ≥ 35 years old
Race/ethnicity	Mother's race/ethnicity	0= Non-Hispanic White 1= Non-Hispanic Black 2= Hispanic 3= Asian/Other
Education	Maternal education	0= <12 years 1= 12 to <16 years 2= ⩾16 years
Smoking	Cigarette smoking	0= Non-smoker 1= Smoked any
Alcohol	Alcohol use during pregnancy	0= No alcohol use 1= Any alcohol use
Marital status	Marital status	0= Married 1= Other

Table 4 (cont). Variables included in the analysis for Manuscript 1: Risk factors or markers for nausea and vomiting in early pregnancy

Variable Name	Description	Variable Coding
Predictors		
Parity	Parity	0= 0 (nulliparous) 1= 1 2= 2+
SAB history	Pregnancy loss history	0= No prior SAB 1= ⩾1 prior pregnancy with no prior SAB 2= ⩾1 prior pregnancy with ⩾1 prior SAB
Age at menarche	Mother's age at menarche	0= <11 years old 1= 12 to 13 years old 2= ⩾14 years old
Plurality	Plurality for index pregnancy	0= Singleton 1= Multiple

Table 5. Variables included in the analysis for Manuscript 2: Symptoms of nausea and vomiting in early pregnancy and spontaneous abortion

Variable Name	Description	Variable Coding
Study Outcome		
Spontaneous abortion	Pregnancy loss at <20 completed weeks gestation	0= No SAB 1= SAB
Main Exposure		
Nausea and vomiting during early pregnancy	Self-reported experiences with NVP during the baseline and follow-up telephone interviews	
	Severity:	0= Both nausea and vomiting 1= Nausea symptoms only 2= No symptoms
	Duration	0= Short 1= Moderate 2= Long
Covariates		
Age	Maternal age at study enrollment	1= 18-24 years old 2= 25-29 years old 3= 30-34 years old 4= ⩾ 35 years old
Race/ethnicity	Mother's race/ethnicity	0= Non-Hispanic White 1= Non-Hispanic Black 2= Hispanic 3= Asian/Other
Marital status	Marital status	0= Married 1= Other
Education	Maternal education	0= <12 years 1= 12 to <16 years 2= ⩾16 years

Table 5 (cont). Variables included in the analysis for Manuscript 2: Symptoms of nausea and vomiting during early pregnancy and spontaneous abortion

Variable Name	Description	Variable Coding
Covariates		
Smoking	Cigarette smoking	0= Non-smoker 1= Smoked any
Alcohol	Alcohol use during pregnancy	0= No alcohol use 1= Any alcohol use
Age at menarche	Mother's age at menarche	0= <11 years old 1= 12 to 13 years old 2= ⩾14 years old
Parity	Parity	0= 0 (nulliparous) 1= 1 2= 2+
SAB history	Pregnancy loss history	0= No prior SAB 1= ⩾1 prior pregnancy with no prior SAB 2= ⩾1 prior pregnancy with ⩾1 prior SAB
Plurality	Plurality for index pregnancy	0= Singleton 1= Multiple

Table 6. Variables included in the analysis for Manuscript 3: Exposure of caffeine intake and spontaneous abortion: An application of propensity score methods

Variable Name	Description	Variable Coding
Study Outcome		
Spontaneous abortion	Pregnancy loss at <20 completed weeks gestation	0= No SAB 1= SAB
Main Exposure		
Caffeine exposure from coffee and total caffeine	Calculated by self-reported daily consumption applied to Bunker's parameters for caffeine contents for caffeinated beverages.	
	Pre-pregnancy	Calculated pre-pregnancy consumption (mg/day)
	Early pregnancy	Calculated consumption (mg/day) at four weeks after LMP
	Current	Calculated consumption (mg/day) at the time of telephone interview
Covariates		
Maternal age	Maternal age at study enrollment	1= 18-24 years old 2= 25-29 years old 3= 30-34 years old 4= ⩾ 35 years old
Race/ethnicity	Mother's race/ethnicity	0= Non-Hispanic White 1= Non-Hispanic Black 2= Hispanic 3= Asian/Other
Marital	Marital status	0= Unmarried 1= Married
Education	Maternal education	0= ≤12 years 1= >12 - <16 years 2= 16+ years

Table 6 (cont). Variables included in the analysis for Manuscript 3: Exposure of caffeine intake and spontaneous abortion: An application of propensity score methods

Variable Name	Description	Variable Coding
Covariates		
Smoking	Cigarette smoking	0= Non-smoker 1= Smoked any
Alcohol	Alcohol use during pregnancy	0= No alcohol use 1= Any alcohol use
Vitamin use	Vitamin use	0= No 1= Yes
NVP symptoms	Symptoms of nausea and vomiting during pregnancy	0= both symptoms 1= nausea only 2= no symptoms
Parity	Parity	0= 0 1= 1 2= 2+
SAB history	Pregnancy loss history	0= No prior pregnancy 1= Prior pregnancy with no prior pregnancy loss 2= Prior pregnancy with prior pregnancy loss

3.9 Missing values

To increase the precision of our estimates, multiple imputation was considered as a tool to systematically impute values for covariates with missing values in some observations. Multiple imputation, in SAS PROC MI and PROC MIANALYZE,[85-87] was sought if covariate patterns of missing data were found to be "missing at random" (MAR) and where covariates of interest were missing greater than five percent of in the observation. If the data

were concluded to be "missing completely at random" (MCAR), complete case analysis was employed where observations with missing values were dropped from the final model.

3.10 Data analysis

3.10.1 Data management

RFTS employed a rigorous protocol for data collection and cleaning by the telephone interviewing staff and computer programmer. The screening interview was conducted in-house at UNC where each potential participants was assigned a unique study identification number; CATI (computer-assisted telephone interviewing) system was used for both baseline and follow-up interviews where this method of collecting telephone interview data minimized "false skip-patterns" and potential human errors associated with questionnaire skip patterns. Data collected from the completed baseline and follow-up telephone interviews were concatenated from ASCII files to SAS data sets in batches weekly from Battelle Centers for Public Health Research and Evaluation in Durham, NC. Descriptive statistics (frequencies and cross-tabulations) were performed to further detect any error in data coding and data entry.

3.10.2 Statistical methods

3.10.2.1 Analytic plan for study aim 1

Various maternal characteristics, health behaviors before and during pregnancy, and reproductive and medical histories have potential influences on the presence or absence of NVP symptoms. Logistic regression is a commonly used analytic technique in epidemiology to estimate the odds of an exposure of interest to an outcome of interest,

and results from logistic regression are easily interpretable for both case-control and cohort studies. The more prevalent the outcome, the adjusted odds ratio (OR) overestimates the risk ratio (RR) when it is greater than one or underestimates the RR when it is less than one.[88]

Aim 1 applied modified Poisson regression with robust error variance[89] models to model the log-risk between independent variables or possible risk factors for NVP. Although binomial regression, or log-binomial model, is generally preferred and more widely used; nevertheless, it has the tendency to be less stable than logistic regression[90] and has the potential problem with model non-convergence.[91] The use of Modified Poisson regression with robust error variance as an analytic technique reduces the problems often seen in logistic regression and binomial regression, and this allows for a direct an estimate of an association that best represented the true relative risk. We examined maternal characteristics including maternal age, race/ethnicity, marital status, and education. Maternal health behaviors included cigarette smoking and alcohol use; reproductive and medical histories consisted of age at menarche, parity, and previous pregnancy loss. As the main outcome variable of interest, NVP was categorized into subgroups to assess the association of the independent variables on symptom-specific NVP. A separate analysis was done to assess the effects of individual predictors on the time of symptom onset (typical or delayed) and duration (symptoms restricted within the first trimester or symptoms lasting beyond the first trimester) of NVP. Bivariate analyses were first performed to examine each covariate and outcomes in order to reduce the list of potential risk factors for NVP. Covariates were entered into the final models if they were found to be statistically significant ($\alpha \leqslant 0.20$) or have risk ratios (RR) greater than

two. Because we examined a highly prevalent outcome from a prospective cohort study, the β coefficient from the modified Poisson regression was a direct estimate of the log risk and the exponentiation of the β coefficients can directly produce a risk ratio, allowing the risk ratio to be the primary effect measure for the predictors and NVP.

3.10.2.2 Analytic plan for study aim 2

Aim #2 examined the association between nausea and vomiting in early pregnancy and spontaneous abortion. Because of the methodological sensitivity in studying spontaneous abortion, survival analysis was a well-suited analytic tool to handle the concerns with left truncation and to account for delay entry.

We used a discrete-time survival model with a logit link (a.k.a. discrete-time continuation ratio logistic survival model) to estimate week-specific odds ratios for the conditional probability of having SAB in a given week,[92-94] given that a woman was still pregnant at the beginning of that time. The main exposure of interest, NVP, was coded as a three-level ordinal categorical variable (with the understanding that linearity assumption is met) where the presence of both nausea symptoms and vomiting episodes was considered the referent category, and "no NVP symptoms" and "nausea only" are the other two index categories. It was also categorized into "short", "moderate", and "long" to evaluate the effects of NVP symptom duration and early pregnancy loss. Similar to study aim #1, maternal characteristics included maternal age, race/ethnicity, marital status, and education. Maternal health behaviors included cigarette smoking and alcohol use; reproductive and medical histories consisted of age at menarche, parity, plurality, and previous pregnancy loss. Effect measure modification was assessed by comparing stratum-specific estimates of each covariate of interest and with the likelihood ratio test

($p<0.10$). Covariates that were deemed to be non-modifiers were assessed to determine whether they were confounders.

We modeled the probability of having a pregnancy loss at a given week *j*, conditional on a subject not experiencing a loss prior to week *j*. Equation 3.1, defined as the hazard at time t_j, is the conditional probability of pregnancy loss for subject *i* at week *j*, where T_i is the gestational week of loss for subject *i*, *j* is the gestational week of pregnancy, and x_{ij} is the set of covariates of interest (the presence or absence of NVP) for subject *i* at gestational week *j*.

$$h(t_{ij}) = Pr(T_i = j \mid T_i \geq j, x_{ij}). \quad (3.1)$$

In turn, we expressed the odds of NVP and SAB as (Equation 3.2)

$$\text{logit}(h(t_{ij}) = \alpha_j + \beta_1 x_{1ij}' + \beta_2 x_{2ij}' + \beta_3 x_{3ij}' + \ldots.. + \beta_n x_{nij}', \quad (3.2)$$

where β describes the log odds between NVP and SAB, x_{ij}' is the effects of each covariate for subject *i* at week *j*, and α_j is the week-specific intercepts to account for the varying probabilities of SAB as pregnancy progresses. This model essentially treats time as a discrete factor by introducing one parameter α_j for each possible time of having a pregnancy loss at t_j. Exp(β) represents the odds ratio for having the outcome of interest given a one-unit increase in the predictor.

3.10.2.3 Analytic plan for study aim 3

Estimating propensity scores

Using logistic regression models, we first estimated the probability of coffee consumption (yes/no) and consumption of any caffeinated beverages (yes/no) for pre-pregnancy, four weeks post-LMP, and around the time of interview. Covariates that were selected to enter the logistic regression model included maternal age, maternal race/ethnicity, education, marital status, smoking, alcohol use, parity, symptoms of NVP, and pregnancy loss history. We then subclassified all propensity scores into propensity score quintiles, as proposed by D'Agostino,[17] and propensity score quintiles were used in discrete-time survival model to estimate the effects of caffeine intake at three time points and SAB.

Discrete-time survival model

For both traditional covariate adjustment and propensity score analyses, discrete-time survival models [93, 94] with a logit link (Equations 3.1 and 3.2) were also used to estimate week-specific odds ratios for the conditional probability of having SAB in a given week, given that a woman was still pregnant at the beginning of that time, with caffeine exposure from coffee only and from total caffeine at three time points as the main exposures of interest.

Caffeine exposure (mg/day) was examined as a categorical variable, with cutpoints in caffeine use set at "none", ">0 to <median", "⩾median", and ">75th percentile". Potential confounders that were used in the traditional models included maternal age, maternal race/ethnicity, education, marital status, alcohol use, and vitamin use.

In the propensity score models, we first examined caffeine effects within each propensity score quintile and used an interaction test to determine whether stratification was needed, based on p=0.05 as the significance level of interest. Propensity scores were kept as the main effects in the models if the test of interaction for each time window did not meet our criteria of <0.05; otherwise, stratification was necessary to examine quintile-specific estimates. Because of the properties of propensity score models, we did not include additional covariates in the propensity score models.

3.11 Software use

Data cleaning, data manipulation (e.g. variable recoding), and data analysis were conducted using SAS 9.1 (SAS Institute, Inc., Cary, NC)

3.12 IRB approval

As determined by The Public Health Institutional Review Board (IRB), this research does not constitute as human subjects research as defined under federal regulations, and therefore, this research does not required by IRB approval (Study # 06-0285).

CHAPTER 4: MANUSCRIPTS

4.1 Manuscript 1: Factors influencing the characteristics of nausea and vomiting in early pregnancy

Abstract

Background: The risk factors or markers for symptoms of nausea and vomiting in early pregnancy (NVP) have not been well established and no study focused on early pregnancy has examined potential risk factors for the subclasses of NVP. We examined potential risk factors (maternal demographic, behavioral, reproductive and medical histories) for the characteristics of NVP symptoms involving symptom severity, time of onset, and duration. **Methods:** Our study population consisted of 2,430 newly pregnant women (≤ 12 weeks' gestation) or women planning a pregnancy, who participated in a prospective cohort study on early pregnancy health, from 2000 to 2004 in three U.S. cities. Detailed data on NVP and other participant health information were collected through telephone interviews, ultrasound, and medical records. Modified Poisson regression with robust error variance models were used to generate the risk ratios for potential risk factors to NVP characteristics. **Results:** Most characteristics, with the exception of plurality, were not associated with having NVP. Dose-response gradient for increasing risk for delayed symptom onset with maternal age; increased risks were also found among Non-Hispanic Blacks [Risk ratio (RR): 4.3, 95% confidence interval (CI):

1.6, 11.6)], Hispanic women (RR=2.3, 95%CI: 0.4, 11.5), and women who began menstruating between 12 and 13 years old (RR=2.6, 95CI: 0.8, 8.7). Our findings also indicated that older maternal age and Non-Hispanic Black women were less likely to experience NVP symptoms lasting beyond the first trimester. **Conclusions:** Maternal age and maternal race/ethnicity were found to be associated with most aspects of NVP, while other factors were associated with selected characteristics of NVP. Further investigation is warranted in order to provide better understanding underlying biological and endocrinological mechanisms that may promote the onset of NVP symptoms.

4.1.1 Introduction

The onset of symptoms of nausea and vomiting during early pregnancy (NVP) often is the first indicator to a woman that she is pregnant. NVP is common, affecting 50% to 90% of pregnant women some time prior to 20 completed weeks' gestation.[1, 2] Timing of NVP onset, duration and severity of symptoms differ among women and among individual pregnancies of the same women. Hyperemesis gravidarum, the most severe form of NVP, affects 0.5% to 2% of pregnancies and is one of the most common reasons for hospitalization among pregnant women.[35] While there is no uniform definition for hyperemesis gravidarum, the usual diagnostic criteria are persistent vomiting, dehydration, ketonuria, electrolyte disturbance, and weight loss greater than five percent, with the absence of other disorders.[36]

Although having NVP is unpleasant and can be debilitating to pregnant women's daily lives, these symptoms have been hypothesized to act as a deterrent for maternal exposure from harmful substances that could potentially damage the embryos during the

critical stages of embryogenesis.[27] The etiology of NVP remains unclear. The secretion of human chorionic gonadotropin (hCG) and its rapid increase during the first trimester are considered by many to be the most likely contributor to NVP symptoms.[2, 28] Increased secretion of thyroxine, prostaglandin $E_{2,}$[30] estrogens,[1] prolactin and estradiol[41] have been proposed as contributing factors. Furthermore, two studies have found evidence that pregnancies of multiple gestation tend to exhibit higher levels of hCG, thus causing these women to experience more severe NVP.[95, 96]

NVP symptom characteristics can vary. Time of symptom onset typically begins at four to eight weeks' gestation, generally peaks in intensity between eight and twelve weeks, and then subsides by twenty weeks' gestation.[30, 37-39] Although NVP symptoms for most women are restricted to early pregnancy, some studies found 20% to 30% of pregnant women continued to be affected by these symptoms beyond 20 weeks' gestation.[1, 34]

No risk factors or markers for NVP symptoms have been established. Some previous studies have found that the presence of NVP is associated with older maternal age, occupation such as manual or service workers, parity, cigarette smoking, and infant gender,[32, 40, 43] although other studies have not replicated these associations.[31, 39, 40, 97] Chronic illnesses such as hypertension, liver and renal diseases,[40] maternal hormones,[41] vitamin use,[32] and stress[43] have been shown to be related to the risk of NVP. Finally, reproductive history including increasing gravidity, plurality (singleton or multiple births), and increasing number of prior miscarriages have also been found to increase the risk for NVP.[97]

One study reported that cigarette smoking or marijuana use was associated with delay the onset of NVP symptoms.[25] Another study found that compared to women with early onset, women who reported late onset of NVP symptoms that began after the first trimester were more likely to be less educated and to be African-Americans.[97] Another study[34] found late onset to be more common among women who are older, had higher parity, and gained less weight during pregnancy. Few have examined risk factors and markers for long symptom duration, but one study found that long symptom duration, defined as symptoms that lasted more than four months (independent of timing of onset), was more common among younger women, with multiple gestation, and among multigravidae.[97]

As part of a prospective pregnancy cohort study, we examined potential risk factors for the characteristics of NVP symptoms involving severity, time of onset and duration.

4.1.2 Methods

Data for this study were collected from 2000 to 2004 as part of a prospective cohort epidemiologic study (Right from the Start or RFTS) on drinking water disinfection by-products and spontaneous abortion (SAB). Details of this study are described elsewhere.[82, 83, 98] Briefly, we recruited pregnant women, from the communities and prenatal care clinics in three cities (Raleigh, North Carolina; Memphis, Tennessee; and Galveston, Texas), who 1) were ≥18 years old and were ≤12 weeks' gestation or 2) were between 18 and 45 and trying to conceive a pregnancy for no more than six months.

Participants were asked to take part in an endovaginal ultrasound assessment to confirm the gestational age and the viability of the fetus around eight weeks' gestation. Two telephone interviews were conducted, no later than 16 weeks' and between 22 and 26 weeks' gestation to ascertain information on maternal health behavior and current and past medical and reproductive histories, including NVP symptoms.

A woman was considered to have experienced symptoms of nausea if she reported having had "nausea or feeling sick to her stomach at any time" during the index pregnancy and she experienced vomiting episodes if she "had nausea so bad that she vomited". Time of onset, that is, start of NVP symptoms, and stop dates were collected separately for symptoms of nausea and vomiting episodes. Onset and stop dates were collected, noting the month, day, and year. Using the midpoint of the week, we imputed "days" for women who were unable to recall the exact days.

For the purposes of this study, symptom severity was classified into "nausea symptoms only", "nausea with vomiting episodes", and "no symptoms". Time of onset was subdivided into "typical onset", defined in our study as symptom onset that took place prior to the start of the second trimester (≤ 13 weeks gestation), and "delayed onset", which was defined as symptoms that started after the first trimester (>13 weeks gestation). Symptom duration was defined as symptoms restricted to the first trimester only and symptoms lasting beyond one trimester. Data relating to the onset and duration of NVP symptoms are presented in gestational weeks, as calculated based on a woman's last menstrual period (LMP).

Information on a wide range of potential risk factors for NVP was collected during the telephone interviews. Maternal characteristics included maternal age (<25

years, 25 to 29 years old, 30 to 34 years, ≥35 years), maternal race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, and Asian/other), marital status (married, other), and education (<12 years, >12 to <16 years, ≥16 years). Maternal health behaviors were cigarette smoking during pregnancy (none, any) and alcohol use during pregnancy (none, any); reproductive and medical histories included age at menarche (≤11 years, 12 to 13 years, ≥14 years), plurality (singleton, multiple), gravidity (primigravida, 2, 3, ≥4) and pregnancy loss history (no prior SAB, ≥1 pregnancy with no SAB, ≥1 pregnancy with ≥1 SABs).

Statistical analysis

With NVP symptoms being a prevalent outcome in our study population, we estimated risk ratios (RRs) instead of odds ratios (ORs), because the ORs would overestimate the RRs when greater than one or underestimate the RRs when it is less than one.[88] Log-binomial[91] and Poisson regression[99] are the recommended procedures estimating risk ratios, though log-binomial models have the tendency to be less stable than logistic regression.[90] To eliminate problems with model non-convergence with log-binomial regression,[91] we used the modified Poisson regression with robust error variance[89] models to generate the RRs for risk factors in relation to the timing of onset and two NVP symptoms outcomes.

Bivariate analyses were first conducted to examine each covariate and outcomes in order to reduce a list of potential risk factors for NVP; covariates were entered into the final models if they were found to be statistically significant (alpha ≤ 0.2) or have risk ratios (RR) greater than two. Based on our strategies, risk factors retained in the various

final analysis models included maternal age, maternal race and ethnicity, education, marital status, smoking, alcohol use, age at menarche, pregnancy loss history, plurality, and gravidity.

4.1.3 Results

We included a total of 2,430 women in the final analysis, with the majority of the study participants recruited from Raleigh (45.2%) or Memphis (37.3%). The mean age at enrollment was 27.8 years and the mean gestational age at enrollment was 54.8 days. Non-Hispanic White women made up over 50% of the study population and 32% were Non-Hispanic Black; nine percent were of Hispanic ethnicity with the majority of those (75.6%) from the Galveston area. Our study population consisted mostly of women who were college graduate (48.8%), married (66.0%), non-smokers (94.5%) or non-drinkers (97.6%), and first time mothers (49.5%). Only 23 women (0.9%) carried multiple gestation. Because of the similarities in characteristics between the two groups and the small number of women carrying multiples, we included these women with women carrying singletons.

The majority of the study participants (88.5%) reported experiencing some form of NVP symptoms with 35.3% having nausea symptoms only and 53.2% having nausea symptoms with vomiting episodes. Characteristics were similar in many respects for women who experienced any NVP symptoms and those who experienced no symptoms. Most maternal sociodemographic and behavioral characteristics, reproductive and medical histories were not associated with having either form of NVP symptoms, with the exception of plurality. Symptoms were more likely to occur among women carrying

multiple gestations (RR = 1.3, 95%CI: 1.2, 1.5 for experiencing nausea only and RR=1.2, 95%CI: 1.2, 1.3 for experiencing nausea with vomiting episodes) (Table 7).

Among women with NVP symptoms, the median time of onset for nausea symptoms was five gestational weeks and the median time of onset for vomiting episodes was seven weeks. Among women with any NVP, 1.1% (n = 23) reported delayed onset of symptoms, defined as any NVP symptoms that began after the first trimester (>13 weeks gestation). We then examined the time of onset in gestational weeks with respect to plurality of the study's index pregnancy. Regardless of plurality, onset for nausea symptoms and for vomiting episodes took place for most women within the first trimester. Women carrying multiple fetuses had a slight delay (in gestational weeks) of nausea symptom onset compared to women carrying singletons; nevertheless, the start of symptoms mostly concentrated between second and ninth week gestation (Figures 7 and 8).

A dose-response gradient was detected for increasing maternal age in relation to an increased risk of delayed symptom onset (Table 8), compared to women with typical symptom onset. The risk for delayed symptom onset was elevated, but imprecise, for Non-Hispanic Black women (RR = 4.3, 95% CI: 1.6, 11.6) and among Hispanic women (RR = 2.3, 95% CI: 0.4, 11.5). High school graduates were less likely than women with at least college degree to experience delayed symptom onset; however, there was a suggestion that women with some years of college were at an increased risk (RR =1.8, 95% CI: 0.6, 5.2) for delayed onset. Women who began menstruating at 12 and 13 years old were almost three-times (RR = 2.6, 95% CI: 0.8, 8.7) as likely as women who began menstruating at a later age (≥14 years old) to have their symptom onset delayed.

Compared to primigravidae, each additional pregnancy was shown to have an inverse relationship with delayed symptom onset.

The median duration for NVP symptoms was eight weeks. Few characteristics were associated with symptoms lasting beyond the first trimester (Table 9). The regression analysis showed that older (≥35 years old) (RR=0.8, 95% CI: 0.6, 1.0), Non-Hispanic Black women (RR=0.8, 95%CI: 0.7, 1.0), and women who consumed alcohol during their pregnancy (RR=0.6, 95%CI: 0.3, 1.1) were less likely to experience symptoms that lasted beyond the first trimester. Additionally, we saw suggestive evidence for increase in risk for longer duration with each successive pregnancy among multigravidae when compared to primigravidae.

4.1.4 Discussion

Eighty-nine percent of our study cohort reported having experienced symptoms of nausea with or without vomiting, which is higher than previously reported in prospective or retrospective cohort[6, 7, 41] and case-control studies.[97] This may be because the study recruited women trying to conceive and those very early in pregnancy, which better captured early pregnancy conditions.

We found little to no association between most characteristics we examined and the risk for having symptoms of NVP, inconsistent with some earlier studies[3, 5, 39, 97] but in agreement with others.[31, 39, 40, 97]

For majority of RFTS women, symptom onset for either nausea symptoms or vomiting episodes began at some point within the first trimester, primarily between the second and tenth week gestation, regardless of plurality. Among women with NVP

symptoms, we found increased, but imprecise, risk for delayed symptom onset among Non-Hispanic Blacks and among women of Hispanic ethnicity, which agree with findings from an earlier study;[97] however, our findings on decreasing risk with each additional previous pregnancy did not confirm findings from the same study. Our findings also suggested that women with some college education and among women who had at least one previous pregnancy with no prior spontaneous abortion to have a higher risk for delayed onset. While no earlier studies examined a woman's age at menarche and time of symptom onset, our findings suggested a modest increase in risk of delayed symptom onset for women who began their menarche at 12 and 13 years old. This evidence merits further investigations to uncover possible determinants for menarche within this age range and how these roles ultimately influence late symptom onset during pregnancy.

We divided NVP symptom duration into "symptoms restricted to the first trimester" and "symptoms lasting beyond the first trimester". Notably, older (≥35 years old) and Non-Hispanic Black women were found to have decreased risks for experiencing symptoms that lasted beyond the first trimester. Our findings of elevated risk among multigravidae (≥2) and among younger women (<25 years old), though marginal, supported the findings of one,[97] but not all studies.[34]

One strength of this study is the use of data from a prospective cohort study of reproductive age women who were identified early in their pregnancies or who were trying to conceive. We were able to collect information on potential risk factors and markers such as maternal behaviors, reproductive and medical histories, and information on symptoms of nausea and vomiting early in pregnancy. Unlike most previous studies, we were able to more precisely define the date the start and end for NVP symptoms

because we collected onset and end dates independently for symptoms of nausea and vomiting episodes, by the month, day, and year.

The information we collected allowed us to examine risk factors and markers for specific NVP symptom characteristics including severity, timing, and duration. We were able to examine effects of potential risk factors within the NVP subgroups, potentially providing clues to biologic and physiologic etiologic pathways for NVP symptoms. Furthermore, we found a tendency for young maternal age to increase the risk and a tendency for Non-Hispanic Black and older maternal age to decrease the risk for long symptom duration, but the opposite was indicated on these maternal characteristics for delayed symptom onset; in addition to the other characteristics and behaviors that were found to be more strongly associated with long symptom duration and delayed symptom onset.

Despite our enhanced protocol and intense effort to recruit women early into our study from communities and prenatal care clinics, limitations still exist with regard to data collection. Although we interviewed 97% of all study cohort to assess NVP information prospectively and 3% retrospectively due to pregnancy loss, we still had to impute the "day" for approximately one-fourth of women with NVP symptoms for either their onset and/or stop dates based on the responses provided for the "week in the month". With the inclusion of the imputed dates, the effect estimates for time of onset and symptom duration were unlikely to have differed greatly than if we were to generate these estimates using actual dates provide by the women. This is because we had asked them to recall these dates around events in their lives (i.e.: birthdays, holidays, etc.); therefore, information given in the "week of the month" ideally would approximate the

recall date(s). NVP information was collected from telephone interviews based on women's recall, rather than from daily diaries. Therefore, inaccurate participant self-report on timing of NVP and its characteristics cannot be ignored. A woman's experience with nausea symptoms is generally more subjective while her experience with vomiting episodes is more objective. Therefore, some women with nausea symptoms may have been misclassified as having no symptoms and vice versa, though these classification errors for experiencing any vomiting seem unlikely.

In RFTS, no biological specimens were collected so we were unable to measure the concentration of hCG and other hormone levels in maternal blood at the point of conception or from the point of a clinically recognized pregnancy, or to monitor its pattern throughout the entire gestation period. Consequently, we were unable to compare the overall hCG profile between singleton and multiple pregnancies. Furthermore, our study did not measure the changes in NVP symptoms; once a woman reported having either nausea only or nausea symptoms with accompanying vomiting episodes we made the assumption that her symptom severity to be constant (i.e.: always nausea only or always both symptoms). This assumption would have prohibited us from finding out other potential risk factors or markers or generating more precise estimates for the risk factors already measured because of its variability in its natural cycle and characteristics.

In this study, we collected a limited amount of information concerning medical treatment and weight loss and weight gain as a result of NVP; thereby we were unable to evaluate hyperemesis gravidarum. From our data, only 20% of 2,149 women with NVP symptoms self-reported having contacted their clinicians and 30 reported being hospitalized for NVP symptoms, thus considered severe enough symptoms to warrant

medical attention. Some of these cases may have been classified as hyperemesis gravidarum.

Our study found that maternal age and maternal race/ethnicity were associated with most aspects of NVP examined in this study, while other characteristics were found to be associated with only specific characteristics of NVP. While many hypotheses have been proposed, the etiology of this common pregnancy symptom remains unknown. Future studies should focus on further refinements in measuring NVP in order to provide a better understanding of the biological and endocrinological mechanisms between the characteristics evaluated in this study and other potential risk factors and markers to NVP symptoms.

4.1.5 References

Table 7. Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom severity: Right from the Start (2000-2004), n=2,430

	Symptom severity				
	No symptoms	Having nausea only		Having both symptoms of nausea and vomiting	
	n (%)	n (%)	Adj. RR†‡ (95% CI)	n (%)	Adj. RR†‡ (95% CI)
Maternal age					
<25 years old	76 (27.2)	160 (18.7)	1.0 (0.8, 1.3)	497 (38.5)	1.1 (1.0, 1.2)
25 to 29 years old	84 (30.1)	284 (33.1)	1.0	383 (29.7)	1.0
30 to 34 years old	70 (25.1)	279 (32.5)	1.0 (0.8, 1.2)	308 (23.9)	1.0 (0.9, 1.1)
≥35 years old	49 (17.6)	135 (15.7)	0.9 (0.8, 1.2)	103 (8.0)	0.8 (0.7, 0.9)
Race/ethnicity					
Non-Hispanic White	137 (49.1)	558 (65.1)	1.0	662 (51.3)	1.0
Non-Hispanic Black	112 (40.1)	215 (25.1)	0.9 (0.8, 1.0)	444 (34.4)	0.9 (0.9, 1.0)
Hispanic	22 (7.9)	57 (6.7)	1.0 (0.8, 1.1)	130 (10.1)	1.0 (0.9, 1.1)
Asian/Other	8 (2.9)	27 (3.2)	1.0 (0.8, 1.2)	30 (2.3)	1.1 (1.0, 1.2)
Education					
<12 years	94 (33.7)	174 (20.3)	0.9 (0.8, 1.0)	455 (35.3)	1.0 (1.0, 1.1)
>12 to <16 years	47 (16.9)	161 (18.8)	1.0 (0.9, 1.1)	313 (24.3)	1.1 (1.0, 1.2)
≥16 years	138 (49.5)	523 (61.0)	1.0	522 (40.5)	1.0
Marital status					
Married	169 (60.6)	663 (77.4)	1.0	770 (59.6)	1.0
Other	110 (39.4)	194 (22.6)	0.9 (0.8, 1.0)	321 (40.36)	1.0 (0.9, 1.0)

*NVP: Nausea and vomiting during early pregnancy

†Model adjusted for maternal age, race/ethnicity, education, marital status, alcohol use, age at menarche, pregnancy loss history, and plurality

‡Compared to women with no nausea and vomiting symptoms

Table 7 (cont). Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom severity: Right from the Start (2000-2004), n=2,430

	Symptom severity				
	No symptoms	Having nausea only		Having both symptoms of nausea and vomiting	
	n (%)	n (%)	Adj. RR†‡ (95% CI)	n (%)	Adj. RR†‡ (95% CI)
Smoking					
Any	16 (5.8)	31 (3.6)	-	87 (6.7)	-
None	263 (94.3)	827 (96.4)	-	1,204 (93.3)	-
Alcohol use					
Any	11 (4.0)	22 (2.6)	0.8 (0.7, 1.1)	25 (1.9)	0.9 (0.7, 1.1)
None	266 (96.0)	836 (97.4)	1.0	1,266 (98.1)	1.0
Age at menarche					
≤11 years old	59 (21.5)	177 (20.8)	1.1 (1.0, 1.2)	298 (23.2)	1.1 (1.0, 1.1)
12 to 13 years old	132 (48.0)	464 (54.6)	1.1 (1.0, 1.2)	674 (52.5)	1.1 (1.0, 1.1)
≥14 years old	84 (30.6)	209 (24.6)	1.0	312 (24.3)	1.0
Pregnancy loss (SAB) history					
No prior SAB	88 (31.5)	246 (28.7)	1.0	391 (30.3)	1.0
≥1 pregnancy with no SAB	150 (53.8)	429 (50.0)	1.0 (1.0, 1.1)	609 (47.2)	1.0 (1.0, 1.1)
≥1 pregnancy with ≥1 SAB	41 (14.7)	183 (21.33)	1.1 (1.0, 1.2)	291 (22.5)	1.1 (1.1, 1.2)

*NVP: Nausea and vomiting during early pregnancy

† Model adjusted for maternal age, race/ethnicity, education, marital status, alcohol use, age at menarche, pregnancy loss history, and plurality

‡Compared to women with no nausea and vomiting symptoms

Table 7 (cont). Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom severity: Right from the Start (2000-2004), n=2,430

	Symptom severity				
	No symptoms	Having nausea only		Having both symptoms of nausea and vomiting	
	n (%)	n (%)	Adj. RR†‡ (95% CI)	n (%)	Adj. RR†‡ (95% CI)
Plurality					
Singleton	279 (100.0)	852 (99.3)	1.0	1,274 (98.7)	1.0
Multiple	0 (0.0)	6 (0.70)	1.3 (1.2, 1.5)	17 (1.3)	1.2 (1.2, 1.3)
Gravidity					
Primigravid	93 (33.3)	253 (29.5)	-	408 (31.6)	-
2	89 (31.9)	283 (33.0)		377 (29.2)	-
3	45 (16.1)	171 (19.9)	-	256 (19.8)	-
≥4	52 (18.6)	151 (17.6)	-	250 (19.4)	-

*NVP: Nausea and vomiting during early pregnancy

† Model adjusted for maternal age, race/ethnicity, education, marital status, alcohol use, age at menarche, pregnancy loss history, and plurality

‡Compared to women with no nausea and vomiting symptoms

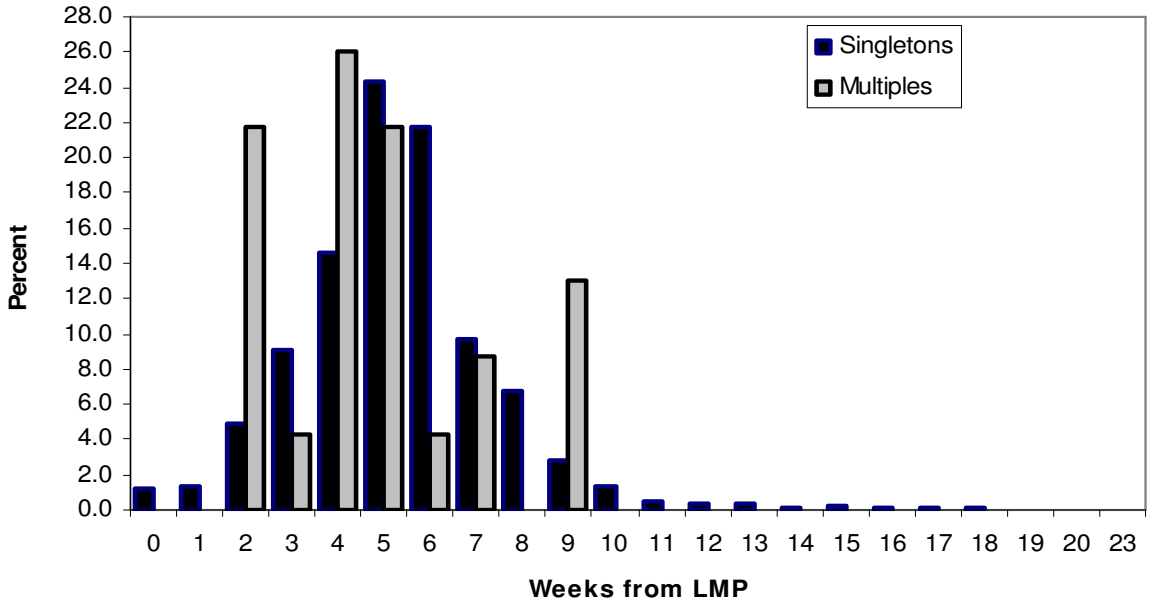


Figure 7. Onset of nausea symptoms by plurality

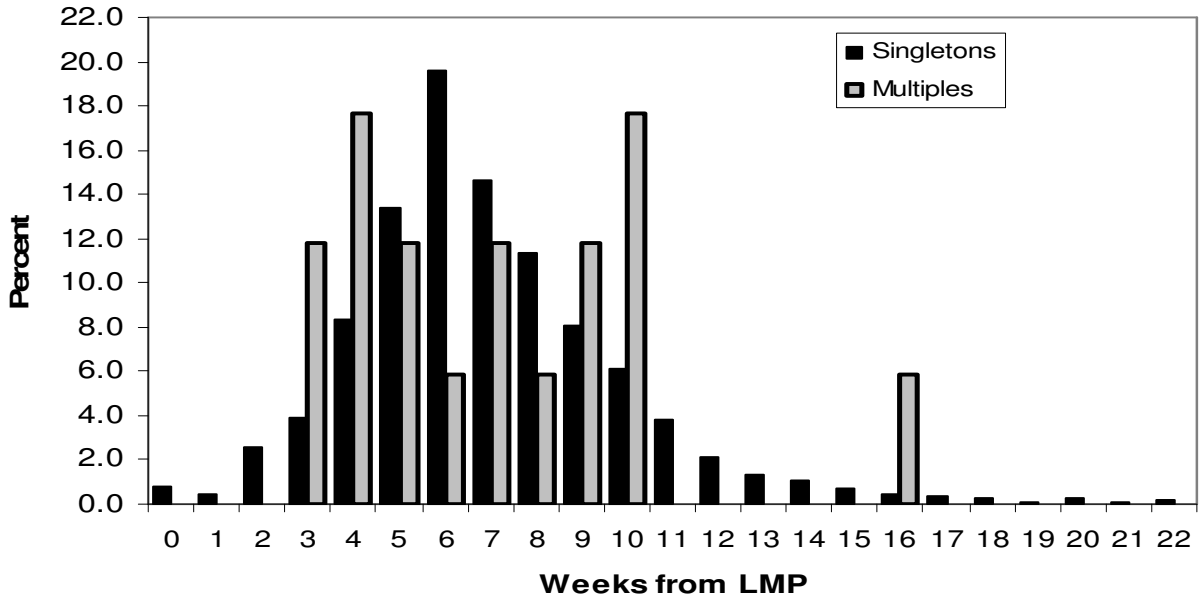


Figure 8. Onset of vomiting episodes by plurality

Table 8. Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom time of onset: Right from the Start (2000-2004), n=2,430

	Time of onset		
	Typical onset	Delayed onset	
	n (%)	n (%)	Adj. RR†‡ (95% CI)
Maternal age			
<25 years old	647 (30.5)	7 (30.4)	0.9 (0.3, 3.0)
25 to 29 years old	658 (31.0)	7 (30.4)	1.0
30 to 34 years old	581 (27.4)	6 (26.1)	1.3 (0.4, 4.2)
≥35 years old	235 (11.1)	3 (13.0)	1.7 (0.4, 6.9)
Race/ethnicity			
Non-Hispanic White	1,211 (57.2)	8 (34.8)	1.0
Non-Hispanic Black	643 (30.3)	13 (56.5)	4.3 (1.6, 11.6)
Hispanic	185 (8.7)	2 (8.7)	2.3 (0.4, 11.5)
Asian/Other	80 (3.8)	0 (0.0)	- §
Education			
<12 years	621 (29.3)	5 (21.7)	0.7 (0.2, 2.7)
>12 to <16 years	464 (21.9)	9 (39.1)	1.8 (0.6, 5.2)
≥16 years	1,035 (48.8)	9 (39.1)	1.0
Marital status			
Married	1,418 (66.9)	14 (63.6)	-
Other	703 (33.1)	8 (36.4)	-

*NVP: Nausea and vomiting during early pregnancy
†Model adjusted for maternal age, race/ethnicity, education, smoking, age at menarche, pregnancy loss history, plurality, and gravidity
‡Compared to women with typical symptom onset
§Cannot estimate due to small sample size

Table 8 (cont). Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom time of onset: Right from the Start (2000-2004), n=2,430

	Time of onset		
	Typical onset	Delayed onset	
	n (%)	n (%)	Adj. RR†‡ (95% CI)
Smoking			
Any	115 (5.4)	1 (4.4)	1.1 (0.1, 8.2)
None	2,006 (94.6)	22 (95.6)	1.0
Alcohol use			
Any	46 (2.2)	1 (4.4)	-
None	2,075 (97.8)	22 (95.7)	-
Age at menarche			
≤11 years old	471 (22.4)	3 (13.6)	1.0 (0.2, 5.1)
12 to 13 years old	1,118 (53.1)	16 (72.7)	2.6 (0.8, 8.7)
≥14 years old	518 (24.6)	3 (13.6)	1.0
Pregnancy loss (SAB) history			
No prior SAB	629 (29.7)	6 (26.1)	1.0
≥1 pregnancy with no SAB	1,024 (48.3)	13 (56.5)	1.6 (0.4, 6.5)
≥1 pregnancy with ≥1 SAB	468 (22.1)	4 (17.4)	1.2 (0.2, 7.5)

*NVP: Nausea and vomiting during early pregnancy
†Model adjusted for maternal age, race/ethnicity, education, smoking, age at menarche, pregnancy loss history, plurality, and gravidity
‡Compared to women with typical symptom onset
§Cannot estimate due to small sample size

Table 8 (cont). Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom time of onset: Right from the Start (2000-2004), n=2,430

	Time of onset		
	Typical onset	Delayed onset	
	n (%)	n (%)	Adj. RR†‡ (95% CI)
Plurality			
Singleton	2,098 (98.9)	23 (100.0)	1.0
Multiple	23 (1.1)	0 (0.0)	- §
Gravidity			
Primigravid	653 (30.8)	6 (26.1)	1.0
2	652 (30.7)	8 (34.8)	0.9 (0.5, 2.6)
3	420 (19.8)	5 (21.7)	0.8 (0.2, 2.6)
≥4	396 (18.7)	4 (17.4)	0.7 (0.1, 4.2)

*NVP: Nausea and vomiting during early pregnancy

†Model adjusted for maternal age, race/ethnicity, education, smoking, age at menarche, pregnancy loss history, plurality, and gravidity

‡Compared to women with typical symptom onset

§Cannot estimate due to small sample size

Table 9. Selected maternal characteristics by NVP[*] symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom duration: Right from the Start (2000-2004), n=2,430

	Symptom duration			
	Symptoms restricted to 1st trimester		Symptoms lasting beyond 1st trimester	
	n (%)	Adj. RR[*†] (95% CI)	n (%)	Adj. RR[†‡] (95% CI)
Maternal age				
<25 years old	486 (28.8)	1.1 (1.0 1.1)	167 (36.9)	1.2 (1.0, 1.3)
25 to 29 years old	530 (31.4)	1.0	133 (29.4)	1.0
30 to 34 years old	474 (28.1)	1.0 (1.0, 1.1)	113 (24.9)	1.0 (0.9, 1.2)
≥35 years old	198 (11.7)	0.9 (0.9, 1.0)	40 (8.8)	0.8 (0.6, 1.0)
Race/ethnicity				
Non-Hispanic White	986 (58.5)	1.0	233 (51.6)	1.0
Non-Hispanic Black	491 (29.1)	0.9 (0.9, 1.0)	163 (36.1)	0.8 (0.7, 1.0)
Hispanic	141 (8.4)	1.0 (0.9, 1.1)	45 (10.0)	0.9 (0.8, 1.1)
Asian/Other	69 (4.0)	1.0 (1.0, 1.1)	11 (2.4)	0.9 (0.6, 1.4)
Education				
<12 years	442 (26.2)	1.0 (0.9, 1.0)	182 (40.2)	1.1 (0.9, 1.3)
>12 to <16 years	368 (21.8)	1.0 (1.0, 1.1)	104 (23.0)	1.2 (1.0, 1.4)
≥16 years	877 (52.0)	1.0	167 (36.9)	1.0
Marital status				
Married	1,171 (69.4)	1.0	259 (57.2)	1.0
Other	516 (30.6)	1.0 (0.9, 1.0)	194 (42.8)	1.0 (0.9, 1.1)

[*]NVP: Nausea and vomiting during early pregnancy
[†]Model adjusted for maternal age, race/ethnicity, education, marital status, alcohol use, age at menarche, and gravidity
[‡]Compared to women with no nausea and vomiting symptoms

Table 9 (cont). Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom duration: Right from the Start (2000-2004), n=2,430

	Symptom duration			
	Symptoms restricted to 1st trimester		Symptoms lasting beyond 1st trimester	
	n (%)	Adj. RR*† (95% CI)	n (%)	Adj. RR†‡ (95% CI)
Smoking				
Any	86 (5.1)	-	30 (6.6)	-
None	1,602 (94.9)	-	423 (93.4)	-
Alcohol use				
Any	41 (2.4)	0.9 (0.8, 1.0)	6 (1.3)	0.6 (0.3, 1.1)
None	1,647 (97.6)	1.0	447 (98.7)	1.0
Age at menarche				
≤11 years old	355 (21.2)	1.0 (1.0, 1.1)	118 (26.3)	1.2 (1.0, 1.4)
12 to 13 years old	906 (54.0)	1.0	228 (50.9)	1.1 (1.0, 1.3)
≥14 years old	417 (24.9)	1.0 (1.0, 1.1)	102 (22.8)	1.0
Pregnancy loss (SAB) history				
No prior SAB	514 (30.5)	-	118 (26.1)	-
≥1 pregnancy with no SAB	810 (48.0)	-	227 (50.1)	-
≥1 pregnancy with ≥1 SAB	364 (21.6)	-	108 (23.8)	-

*NVP: Nausea and vomiting during early pregnancy

†Model adjusted for maternal age, race/ethnicity, education, marital status, alcohol use, age at menarche, and gravidity

‡Compared to women with no nausea and vomiting symptoms

Table 9 (cont). Selected maternal characteristics by NVP* symptoms, adjusted risk ratios (RR) and 95% confidence interval (CI) on maternal characteristics in relation with symptom duration: Right from the Start (2000-2004), n=2,430

	Symptom duration			
	Symptoms restricted to 1st trimester		Symptoms lasting beyond 1st trimester	
	n (%)	Adj. RR*† (95% CI)	n (%)	Adj. RR†‡ (95% CI)
Plurality				
Singleton	1,670 (98.9)	-	448 (98.9)	-
Multiple	18 (1.1)	-	5 (1.1)	-
Gravidity				
Primigravid	537 (31.8)	1.0	119 (26.3)	1.0
2	521 (30.9)	1.0 (1.0, 1.0)	139 (30.7)	1.1 (1.0, 1.2)
3	334 (19.8)	1.0 (1.0, 1.1)	91 (20.1)	1.2 (1.1, 1.3)
≥4	296 (17.5)	1.1 (1.0, 1.1)	104 (23.0)	1.3 (1.1, 1.5)

*NVP: Nausea and vomiting during early pregnancy

†Model adjusted for maternal age, race/ethnicity, education, marital status, alcohol use, age at menarche, and gravidity

‡Compared to women with no nausea and vomiting symptoms

4.2 Manuscript 2: Severity and duration of nausea and vomiting symptoms in early pregnancy and spontaneous abortion

Abstract

Background: Previous studies have shown an inverse association between the presence of nausea and vomiting in pregnancy (NVP) and spontaneous abortion (SAB). We evaluated this phenomenon more closely by examining the effects of symptom severity and duration on SAB. **Methods:** Our analysis included 2,430 women who participated in a prospective cohort study on pregnancy health between 2000 and 2004 in three U.S. cities. Detailed data on NVP and other participant health information were collected through telephone interviews, ultrasound assessments in early gestation, and medical records abstractions. Discrete-time continuation ratio logistic survival models were used to estimate week-specific pregnancy loss associated with the subclasses of NVP. **Results:** Having no NVP symptoms was found to increase risk for SAB [Adjusted odds ratio (OR) =3.2, 95% confidence interval (CI): 2.4, 4.3], compared to having any symptoms. Increasing maternal age was found to strengthen the risk for SAB among women without any symptoms. Reduced risks for SAB was found across all maternal age groups for those with longer symptom duration, but the effects were much stronger in the oldest maternal age group (OR=0.38, 95%CI: 0.24, 0.61 for moderate duration and OR=0.15, 95%CI: 0.06, 0.37 for long duration). **Conclusions:** Our results confirm that the absence of NVP symptoms is associated with early pregnancy loss. They also clearly indicate that longer symptom duration is protective against early loss, especially among women in the oldest maternal age group.

4.2.1 Introduction

Pregnancy is a time of many changes; these changes are hormonal and physiological, as a woman's body prepares for the growth of a fetus. Nausea and vomiting in early pregnancy (NVP) affects 50 to 90 percent of women,[4, 5, 26] and usually presents as one of the first signs of pregnancy.

The severity of symptoms experienced by women differs; some women may only experience symptoms of nausea, while other women experience both nausea and vomiting. Timing of symptom onset and duration vary both among women and across multiple pregnancies of a woman. Symptoms can begin as early as two to four weeks' gestation[30, 37] and end for most women by 20 weeks' gestation.[30, 100] However, 20% of women continue to experience symptoms beyond 20 weeks' gestation.[100] Hyperemesis gravidarium is the most severe form of NVP, in which women have prolonged NVP symptoms that cause maternal weight loss, electrolyte imbalance, and dehydration.[29]

The pathogenesis of NVP is poorly understood but has been attributed to the rise in the hormone hCG, human chorionic gonadotropin, and the trophoblastic activity and gonadotropin production in the early phases of pregnancy.[3] A number of maternal characteristics, including primiparity, younger maternal age, and lower education, have been associated with NVP, and higher maternal BMI has been identified as a risk factor for vomiting.[5]

Spontaneous abortion (SAB) is defined as a pregnancy loss before 20 completed weeks' gestation. Though it affects approximately 10% of all clinically-recognized pregnancies,[22] the actual rate of pregnancy loss is about two to five times higher, most

not clinically recognized.[101] SAB typically affects women of all ages and across all sociodemographic characteristics. Current knowledge on the causes of SAB is limited, but two consistent risk factors are older maternal age and history of previous SABs. Other potential risk factors found in epidemiologic studies have included occupational chemical exposures,[15] maternal stress, caffeine consumption,[12-14] uterine fibroids, [47, 48, 102, 103] incompetent cervix,[49] and drinking water disinfection by-products.[16]

Earlier epidemiologic studies have shown an inverse association between NVP and SAB.[3-8] The absence of NVP may be indicative of an impending pregnancy loss, where the presence of symptoms serves as a predictor of healthy pregnancy outcome. Additionally, an earlier analysis[104] found a dose-response pattern between the presence of nausea alone and the presence of nausea symptoms with vomiting episodes and decreased risk of SAB. However, many of these studies reveal methodologic limitations including differences in study design, selection criteria, sample size, ascertainment of NVP information and classification of NVP symptoms, and analytical approach.

As part of a prospective cohort study on drinking water disinfection by-products and early pregnancy loss, we obtained detailed information on symptoms of nausea and vomiting. Because we recruited women quite early in gestation or before they conceived, we were able to collect detailed prospective data on the pregnancy, onset of nausea symptoms and vomiting episodes, behavioral factors and other maternal characteristics. In this study, we focused on symptom severity and duration of NVP to evaluate whether different sub-classifications of NVP are associated with an increased risk of early pregnancy loss.

4.2.2 Methods

Participant recruitment

We identified and recruited women who were newly pregnant and women who were trying to conceive from Raleigh, North Carolina; Memphis, Tennessee; and Galveston, Texas. Women were recruited from public and private prenatal care providers and the communities between 2000 and 2004 for an epidemiologic study that examined the relationship between drinking water disinfection by-products and early pregnancy loss. Women were eligible to be in the study if they were at least 18 years old and pregnant at less than 12 weeks gestation or if they were between 18-45 years old and had been trying to conceive for no longer than six months. A screening interview was used to determine whether a pregnant woman was eligible to enroll (or "pre-enroll" for a woman who was trying to become pregnant). The recruitment process has been described in detail elsewhere.[82, 83, 98] An endovaginal ultrasound was performed around eight weeks' gestation (median gestational age at ultrasound = 9.3 weeks) to confirm pregnancy viability and to ascertain the gestational age of the fetus. Telephone interviews took place in two time intervals, at no later than 16 weeks' gestation and then between 20 and 25 weeks' gestation, to collect information on health behaviors, medical and reproductive history, current pregnancy history and symptoms, and water exposure characteristics.

Exposure to symptoms of nausea and vomiting during early pregnancy

Information on NVP was collected during two telephone interviews. Symptoms of nausea was defined as "nausea or feeling sick to your stomach at any time" during the index pregnancy, and a vomiting episode was defined as "had nausea so bad that you

vomited". We collected information on the timing of onset and ending dates separately for symptoms of nausea and for vomiting episodes. Questions that pertained to vomiting episodes were asked exclusively of women who experienced symptoms of nausea. Onset and end dates were collected as month, day, and year. For those who were unable to recall the exact days, we collected timing information with respect to "week in the month", and we imputed the day as the midpoint of the week. Symptoms had not subsided for 13% (164 out of 1,291) of women who had both symptoms of nausea and vomiting and 26% (222 out of 858) of women with nausea symptoms only at the time of the telephone interviews; therefore, we created a symptom ending date for these women at 23 weeks gestation, which was shown in the literature as a time during pregnancy when most NVP symptoms commonly end.[37]

We described symptoms of NVP using two categories: symptom severity and symptom duration. Absence or presence of NVP symptoms, whether nausea alone or with the addition of vomiting episodes, was used to classify symptom severity into 1) no symptoms, 2) nausea only, and 3) nausea symptoms with vomiting episodes. The length of symptoms in weeks was calculated separately for nausea symptoms and vomiting episodes. The length of these intervals were used to denote short ($<50^{th}$ percentile), moderate ($>50^{th}$ to $<75^{th}$ percentile), and long symptom duration ($\geq 75^{th}$ percentile), in weeks for nausea symptoms and vomiting episodes.

Health characteristics, health behavior, and pregnancy outcome assessment

Using the telephone interviews, a wide range of risk factors for early pregnancy loss were collected and considered as potential effect measure modifiers and

confounders. These covariates included maternal age, race and ethnicity, maternal education, marital status, smoking, alcohol use, age at menarche, parity, SAB history, and plurality. We found, in the main study,[82, 83] that the self-reported last menstrual period (LMP) in this cohort to be highly reliable; therefore, self-reported LMP was used to date the onset of pregnancy. Pregnancy outcome data were ascertained via participant self-report, with confirmation via medical records abstraction and/or with the presence or absence of corresponding vital records for the pregnancy. Early pregnancy loss or spontaneous abortion is defined in this study as a loss of a pregnancy before 20 completed weeks', or 140 days, gestation calculated from the self-reported LMP.

Multiple imputation

We determined the level of missing data by evaluating the data three ways: 1) the number of missing observations observed for each covariate, 2) the amount of "missingness" for each covariate cross-classified by the main exposure, and 3) the amount of "missingness" for each covariate cross-classified by the outcome. If missingness for any covariates of interest occurred in greater than five percent of the total observations and the covariate patterns of the missing data were found to be "missing at random" (MAR), then the missing values were imputed as a function of the existing data using multiple imputation (PROC MI and PROC MIANALYZE).[85-87] It was determined that the level of data "missingness" did not exceed our pre-determined criteria. As a result, complete case analysis was conducted for all analyses with less than two percent of women dropped due to missingness.

Statistical analysis

We used a discrete-time continuation ratio logistic survival model to estimate week-specific pregnancy loss in relation to symptoms of nausea and vomiting in early pregnancy. Adjusted odds ratios and the 95% confidence intervals were generated for the conditional probability of having a pregnancy loss in a given gestational week, conditional on a woman's pregnancy having survived to the beginning of that gestational week. Our cohort was followed from the time of enrollment into the study until 20 weeks' gestation, with possible pregnancy outcomes being a viable pregnancy that went on past 20 weeks, pregnancy loss, or participant dropout. Pregnancies that survived passed 20 completed gestational weeks were included for the entire period, and women whose pregnancies ended for other reasons (i.e.: ectopic pregnancies, molar pregnancies) or those who were lost-to-follow-up were censored at the end of the risk period.

Two distinct discrete-time survival models were used to estimate the sub-classes (severity and duration) of symptoms of NVP in relation to early pregnancy loss. The covariates that we considered were known to be predictors of pregnancy loss, including maternal age, race and ethnicity, maternal education, marital status, smoking, alcohol use, age at menarche, parity, pregnancy loss history, and plurality. Effect measure modifiers were evaluated by comparing stratum-specific estimates of each covariate of interest along with examination of the likelihood ratio test ($p<0.15$). Covariates that were deemed to be non-modifiers were then tested for confounding. Evaluation of potential confounders was conducted first by examining the association of each covariate of interest by the main exposure among all subjects and also by examining each covariate with the outcome among the unexposed group, along with using the change-in-estimate

approach, where covariates were retained in the final models if they changed the effect estimates for the exposure of interest by greater than 10% when removed from the model. In each model, maternal race and ethnicity, maternal education, marital status, alcohol use, age at menarche, parity, and pregnancy loss history were evaluated as potential confounders; maternal age was included as an effect measure modifier.

4.2.3 Results

A total of 2,430 women were included in the final analysis (78% of all eligible and enrolled) in the study, with majority of the study participants recruited from the Raleigh or Memphis study areas (82.4%). The mean age at enrollment for this study population was 27.8 years (SD=5.5) and the mean gestational age at enrollment was 54.8 days (SD=14.0). Although non-Hispanic White women made up over 50% of the study population, substantial proportions (31.8%) of study participants were non-Hispanic Black. Very few women in our study population self-identified as smokers or consumers of alcohol during pregnancy. Approximately 37% were considered as overweight or obese by the Institute of Medicine classification.[105] Almost half of the study participants were primigravida and 21% reported having a prior early pregnancy loss. Only 23 women had multiple gestations in this index pregnancy (Table 10). Because of the small number of women carrying multiples and the similar characteristics between the two groups, the analyses presented here combined women with singleton and multiple pregnancies. There were 260 (10.7%) SABs in this study population.

Eighty-eight percent of the women in this cohort experienced some form of NVP, with more than half (53.2%) reported experiencing symptoms of nausea accompanied by

vomiting episodes (Table 10). Onset for NVP symptoms was more likely to occur in early pregnancy, with 99.0% of initial nausea symptoms and 95.3% of initial vomiting episodes taking place in the first trimester. The median onset time for nausea symptoms was 5.7 weeks and that for vomiting episodes was 7.0 weeks from LMP.

Table 11 highlights the maternal sociodemographic characteristics, behavioral, and reproductive characteristics of the study population, stratified by sub-classes of NVP. Women who reported experiencing both symptoms of nausea and vomiting episodes tended to be younger (<30 years old), of Hispanic origin, and had completed fewer years of education. Reproductive characteristics were similar for women with any form of NVP, but women who reported experiencing no nausea symptoms at this index pregnancy tended to begin menstruating at a later age (≥14 years old), were nulliparous, or had no history of prior SAB if they had ≥1 prior pregnancy.

In our study cohort, most women with any form of NVP symptoms reported having either short (32.7%) or moderate (28.4%) length of symptom duration, where these intervals are equivalent to six weeks or less and between seven to ten gestational weeks, respectively, with an overall mean symptom duration of eight gestational weeks. Those who reported having long symptom duration tended to be younger (<30 years old), Black or of Hispanic ethnicity, had fewer years of education, and began menstruating at an early age (≤11 years old). Women with either short or moderate length of symptom duration were likely to be older (≥30 years old), White, have completed at least college, began menstruating between 12-13 years old, and had no prior SAB.

Overall, women with no symptoms of NVP had three times the odds for having a SAB (OR=3.2, 95% CI: 2.4, 4.3) compared to women with any form of NVP symptoms.

The remaining results of modeling the week-specific odds of a woman having an early pregnancy loss for the different forms of NVP symptoms are shown in Table 12. The association between NVP and pregnancy loss was strengthened with increasing maternal age, from OR=4.04 (95%CI: 2.14, 7.61) for women <25 years old to OR=10.92 (95%CI: 4.47, 26.67) for women 35 years and older. A more moderate pattern of association was found across maternal age groups for having nausea symptoms only. Maternal age also modified the association between duration and pregnancy loss. Longer symptom duration was found to have a much reduced odds for SAB for all maternal age groups, but the association was much stronger among the oldest maternal age group (OR=0.38, 95%CI: 0.24, 0.61 for moderate duration and OR=0.15, 95%CI: 0.06, 0.37 for long duration).

4.2.4 Discussion

We examined the associations between the subgroups of symptoms of nausea and vomiting and spontaneous abortion in a population of women recruited early in their pregnancy or women who were planning a pregnancy. The incidence of NVP (35.3% for nausea symptoms only, 53.2% for both symptoms) was within the range for nausea symptoms only (20% to 50%) and for both symptoms (48% to 80%) reported by previous cohort studies.[6, 32, 41] However, the incidence for experiencing nausea only and for experiencing both nausea symptoms and vomiting episodes are higher than those reported in other cohort studies in which recruitment took place in prenatal clinics.[6, 30] This may be the result of our study design, which collected symptom data early in pregnancy, increasing the completeness of reporting.

The effects of symptom severity and duration were found to be different across maternal age groups. It is known that the rate of spontaneous abortion begins to increase for women between ages 30 to 35, resulting from chromosomally aberrant and/or chromosomally normal losses,[106] and the absence of NVP symptoms is a marker for an increased risk in early pregnancy. Thus the effect of NVP appears to be magnified with advancing maternal age. Some of the odds ratios were imprecise within maternal age categories, and caution in interpretation is warranted.

To our knowledge, no previous study has evaluated the effects of NVP symptom duration on early pregnancy loss. Studies that evaluated symptom severity often classified symptoms of nausea and vomiting together as a single dichotomous factor. Additionally, no other studies have acquired such detailed information on symptom duration distinctively for nausea symptoms and for vomiting episodes. The few studies that collected such information have done so by either gathering descriptive data on duration exclusively for nausea with or without the presence of vomiting episodes or for the presence of both symptoms. Because we collected detailed information, we have been able to more accurately determine symptom duration. Furthermore, our assignment of a symptom ending date at 23 weeks gestation (for women whose symptoms had not subsided by the time of the telephone interview) should have minimal impact on the reported results because we later categorized symptom duration as short, moderate, and long. The duration of symptoms for our study cohort was found to be within the range for mean symptom duration (eight to twelve gestational weeks) found in the literature;[31, 32] nevertheless, a sizable proportion of our cohort reported their symptoms lasting for six

weeks or less, which may be attributed to our enhanced collection methods where misclassification of exposure was minimized.

Studies of SABs are subject to left truncation because women tend to enter the study at different gestational ages and women without a viable pregnancy prior to study enrollment, who would have otherwise been eligible to participate, are missing from the study population.[70, 71] Prenatal clinic recruitment, as conducted in many previous studies, allowed for rapid identification of pregnant women; however, a proportion of the source population would have been excluded due to pregnancy losses prior to recruitment. Effect estimates based on such data may introduce bias because study participation is differential by outcome. Data for this study came from a study population that was recruited early in pregnancy or as they were planning pregnancies, instead of from solely a population of women seeking prenatal care. Because we were able to identify these women early and follow them through their pregnancies, we prospectively collected detailed information on maternal behaviors, reproductive and medical histories, and other maternal characteristics at an earlier gestational age than previous studies; consequently, differential misclassification due to recall bias was minimized. Despite these improvements, limitations still exist in our study regarding the information collected on NVP symptoms. Our study did not attempt to collect NVP information daily, and errors on the timing of NVP and its characteristics could have occurred. Misclassification introduced by recall bias, could bias the association with early pregnancy loss; however, because NVP assessment took place prior to 71% of all pregnancy losses (the remaining 29% interviewed after loss), the potential for recall error would be limited. Furthermore, we imputed symptom stop dates for 18% of women with any NVP symptoms who

reported that their symptoms had not subsided at the time of the interviews; even if symptom duration for those women were calculated from an actual symptom stop date, the effect estimate for symptom duration would not have changed substantially, because we chose an end point in the gestational period that reflected a time when most symptoms would subside on average. Even though our study recruited women who were early in their gestation period and women who were not yet pregnant, we still could have missed capturing a small proportion of women with early pregnancy loss; thereby potentially producing more imprecise estimates (with wider confidence intervals) and underestimating the magnitude of effect on the exposures with respect to the outcome.

The mechanisms by which NVP predicts favorable pregnancy outcome are not known; nevertheless, several different mechanisms have been postulated to explain the reported feto-protective effect of NVP. These include reducing fetal exposure to potential teratogens present in the maternal diet,[7] improving the quality of maternal diets to favor the consumption of certain nutrients, increasing energy expenditure that alter hormonal balance in favor of maternal and fetal tissue growth,[7, 29] along with other nutritional and non-nutritional pathways. In our study, we were able to confirm previous findings that the absence of NVP symptoms is associated with pregnancy loss. We also were able, for the first time, to provide evidence that longer symptom duration is protective against early pregnancy loss. In addition to a more refined protocol and a large study population, our statistical methods accounted for left censoring by time of study entry, allowing us to generate more valid estimates of the associations between the exposures and early pregnancy loss. Along with complete assessment into the timing and occurrence of NVP,

a better understanding into the biological and physiological mechanisms of this common pregnancy phenomenon may help further explain this present association.

4.2.5 References

Table 10. Sociodemographic characteristics, selected maternal behavior, and reproductive histories of women in analysis: Right from the Start (2000-2004), n=2,430

	N	%
Study site		
Raleigh	1,097	45.1
Memphis	907	37.3
Galveston	426	17.5
Maternal age		
<25 years old	733	30.2
25 to 29 years old	752	31.0
30 to 34 years old	658	27.1
≥35 years old	287	11.8
Race/ethnicity		
Non-Hispanic White	1,359	56.0
Non-Hispanic Black	771	31.8
Hispanic	209	8.6
Asian/Other	89	3.7
Education		
<12 years	723	29.8
>12 to <16 years	521	21.5
≥16 years	1,185	48.8
Marital status		
Married	1,604	66.0
Other	825	34.0
Income		
≤$40,000/year	1,017	43.6
$40,001 to $80,000/year	773	33.2
>$80,000/year	541	23.2
Smoking		
None	2,296	94.5
<10 cigarettes/day	87	3.6
≥10 cigarettes/day	47	1.9
Alcohol use		
Any	58	2.4
None	2,370	97.6

Table 10 (cont). Sociodemographic characteristics, selected maternal behavior, and reproductive histories of women in analysis: Right from the Start (2000-2004), n=2,430

	N	%
BMI*		
Underweight (<19.8)	250	10.6
Normal weight (19.8-26.0)	1,235	52.2
Overweight (<26.0-29.0)	310	13.1
Obese (>29.0)	572	24.2
Age at menarche		
≤11 years old	534	22.2
12 to 13 years old	1,272	52.8
≥14 years old	605	25.1
Parity		
Nulliparous/none	1,202	49.5
1	766	31.5
2+	462	19.01
Pregnancy loss history		
No prior SAB	726	29.9
≥1 pregnancy with no SAB	1,189	48.9
≥1 pregnancy with ≥1 SAB	515	21.2
NVP† symptom severity		
No symptom	279	11.5
Nausea symptoms only	858	35.3
Nausea and vomiting	1,291	53.2
NVP symptom duration		
Short	701	32.7
Moderate	607	28.4
Long	833	38.9
Plurality		
Singleton	2,407	99.1
Multiple	23	0.9

*BMI: Body Mass Index, based on Institute of Medicine classifications
†NVP: Nausea and vomiting during pregnancy

Table 11. Selected maternal characteristics by NVP* symptoms: Right from the Start (2000-2004), n=2,430

	Symptom severity			Symptom duration		
	No symptom	Nausea only	Nausea and vomiting	Short	Moderate	Long
	n (%)	n (%)	N (%)	n (%)	n (%)	n (%)
Maternal age‡§						
<25 years old	76 (27.2)	160 (18.7)	497 (38.5)	179 (26.1)	131 (27.2)	182 (36.8)
25 to 29 years old	84 (30.1)	284 (33.1)	383 (29.7)	226 (32.9)	140 (29.1)	155 (31.4)
30 to 34 years old	70 (25.1)	279 (32.5)	308 (23.9)	206 (30.0)	151 (31.3)	105 (21.3)
≥35 years old	49 (17.6)	135 (15.7)	103 (7.9)	76 (11.1)	60 (12.5)	52 (10.5)
Race/ethnicity‡						
Non-Hispanic White	137 (49.1)	558 (65.1)	662 (51.3)	410 (59.8)	308 (63.9)	258 (52.2)
Non-Hispanic Black	112 (40.1)	215 (25.1)	444 (34.4)	205 (29.9)	113 (23.4)	156 (31.6)
Hispanic	22 (7.9)	57 (6.7)	130 (10.1)	53 (7.7)	36 (7.5)	58 (11.7)
Asian/Other	8 (2.9)	27 (3.2)	54 (4.2)	18 (2.7)	25 (5.2)	22 (4.4)
Education‡§						
<12 years	94 (33.7)	174 (20.3)	455 (35.3)	167 (24.3)	107 (22.2)	158 (32.0)
>12 to <16 years	47 (16.9)	161 (18.8)	313 (24.3)	150 (21.9)	105 (21.8)	117 (23.7)
≥16 years	138 (49.5)	523 (61.0)	522 (40.5)	369 (53.8)	270 (56.0)	219 (44.3)

*NVP: Nausea and vomiting during pregnancy
†SAB: Spontaneous abortion
‡p<0.05 for severity
§p<0.05 for duration

Table 11 (cont). Selected maternal characteristics by NVP* symptoms: Right from the Start (2000-2004), n=2,430

	Symptom severity			Symptom duration		
	No symptom	Nausea only	Nausea and vomiting	Short	Moderate	Long
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Marital status‡§						
Married	169 (60.6)	663 (77.4)	770 (59.6)	481 (70.1)	355 (73.7)	329 (66.6)
Other	110 (39.4)	194 (22.6)	521 (40.4)	205 (29.9)	127 (26.4)	165 (33.4)
Smoking‡						
None	263 (94.3)	827 (96.4)	1,204 (93.3)	652 (94.9)	460 (95.4)	466 (94.3)
<10 cigarettes/day	8 (2.9)	21 (2.5)	58 (4.5)	26 (3.8)	15 (3.1)	18 (3.6)
≥10 cigarettes/day	8 (2.9)	10 (1.2)	29 (2.3)	9 (1.3)	7 (1.5)	10 (2.0)
Alcohol use§						
Any	11 (4.0)	22 (2.6)	25 (1.9)	13 (1.9)	13 (2.7)	5 (1.0)
None	266 (96.0)	836 (97.4)	1,266 (98.1)	674 (98.1)	469 (97.3)	489 (99.0)
Age at menarche§						
≤11 years old	59 (21.5)	177 (20.8)	298 (23.2)	128 (18.8)	105 (21.9)	129 (26.4)
12 to 13 years old	132 (48.0)	464 (54.6)	674 (52.5)	380 (55.7)	264 (55.1)	243 (49.7)
≥14 years old	84 (30.6)	209 (24.6)	312 (24.3)	174 (25.5)	110 (23.0)	117 (23.9)

*NVP: Nausea and vomiting during pregnancy
†SAB: Spontaneous abortion
‡p<0.05 for severity
§p<0.05 for duration

Table 11 (cont). Selected maternal characteristics by NVP* symptoms: Right from the Start (2000-2004), n=2,430

	Symptom severity			Symptom duration		
	No symptom	Nausea only	Nausea and vomiting	Short	Moderate	Long
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Parity§						
Nulliparous/none	147 (52.7)	409 (47.7)	645 (50.0)	358 (52.1)	219 (45.4)	228 (46.2)
1	83 (29.8)	300 (35.0)	382 (29.6)	206 (30.0)	174 (36.1)	168 (34.0)
2+	49 (17.6)	149 (17.4)	264 (20.5)	123 (17.9)	89 (18.5)	98 (19.8)
Pregnancy loss history§						
No prior SAB	88 (31.5)	246 (28.7)	391 (30.3)	227 (33.0)	151 (31.3)	137 (27.7)
≥1 pregnancy with no SAB	150 (53.8)	429 (50.0)	609 (47.2)	308 (44.8)	236 (49.0)	235 (47.6)
≥1 pregnancy with ≥1 SAB	41 (14.7)	183 (21.3)	291 (22.5)	152 (22.1)	95 (19.7)	122 (24.7)
Plurality						
Singleton	279 (100.0)	852 (99.3)	1,274 (98.7)	677 (98.5)	479 (99.4)	487 (98.6)
Multiple	0 (0.0)	6 (0.7)	17 (1.3)	10 (1.5)	3 (0.6)	7 (1.4)
SAB†‡§						
Yes	76 (27.7)	114 (13.7)	70 (5.6)	87 (12.9)	60 (10.3)	33 (4.0)
No	198 (72.3)	720 (86.3)	1,185 (94.4)	589 (87.1)	522 (89.7)	790 (96.0)

*NVP: Nausea and vomiting during pregnancy
†SAB: Spontaneous abortion
‡p<0.05 for severity
§p<0.05 for duration

Table 12. Unadjusted and adjusted odds ratio on the association between NVP* symptom severity and duration and risk of pregnancy loss and adjusted odds ratios stratified by maternal age: Right from the Start (2000-2004)

	Main effects model		Models with an interaction term (maternal age)			
			<25 years old	25 to 29 years old	30 to 34 years old	≥ 35 years old
NVP characteristics	Unadj. OR† (95% CI)‡	Adj. OR (95% CI)	Adj. OR (95% CI)	Adj. OR (95% CI)	Adj. OR (95% CI)	Adj. OR (95% CI)
Symptom severity§						
No symptoms	5.58 (3.99, 7.81)	5.05 (3.56, 7.18)	4.04 (2.14, 7.61)	3.89 (1.89, 7.99)	6.47 (3.16, 13.24)	10.92 (4.47, 26.67)
Nausea only	2.34 (1.72, 3.17)	2.46 (1.79, 3.37)	2.01 (1.46, 2.76)	1.97 (1.37, 2.83)	2.54 (1.78, 3.64)	3.30 (2.11, 5.16)
Symptom duration‖						
Moderate	0.74 (0.52, 1.04)	0.72 (0.51, 1.02)	0.68 (0.48, 0.97)	0.70 (0.48, 1.02)	0.48 (0.31, 0.75)	0.38 (0.24, 0.61)
Long	0.31 (0.20, 0.46)	0.30 (0.20, 0.45)	0.46 (0.23, 0.94)	0.50 (0.23, 1.05)	0.38 (0.24, 0.61)	0.15 (0.06, 0.37)

*NVP: Nausea and vomiting during pregnancy

†OR: Odds ratio

‡CI: Confidence interval

§Compared to symptoms of both nausea and vomiting; models stratified by maternal age and adjusted for maternal race and ethnicity, education, marital status, alcohol use, age at menarche, parity, and pregnancy loss history

‖Compared to short symptom duration; models stratified by maternal age and adjusted for maternal race and ethnicity, education, marital status, alcohol use, age at menarche, parity, and pregnancy loss history

4.3 Manuscript 3: Studying the effects of maternal health behaviors in pregnancy: An application of propensity score methods

Abstract

Background: Caffeine is readily available and widely consumed in most countries. Issues surrounding caffeine consumption for pregnant women have generated much attention and controversy since the 1980's because of its potential negative impact during pregnancy. Most epidemiologic studies have relied on the use of observational data to estimate causal effects; however, this poses a challenge when inherent differences may exist across the exposure groups. Propensity score methods have been proposed as a well-suited alternative approach in reducing the limitations that often plague observational study data. We illustrated the use of propensity score methods in studying pregnancy health using a study that evaluated the association between caffeine exposure at three time points and spontaneous abortion (SAB). Results from propensity score models were also compared to estimates from traditional covariate adjustment models. **Methods:** We included 2,407 women with singleton pregnancies, who participated in a prospective cohort study on early pregnancy health between 2000 and 2004 in three U.S. cities. Information on caffeine consumption and other health data were collected extensively through participant telephone interviews, ultrasound assessments in early gestation, and medial records abstraction. Logistic regression model was used to estimate individual propensity scores. Week-specific pregnancy loss associated with caffeine exposure was estimated with discrete-time hazard models. **Results:** There were no overall meaningful differences in estimates from the traditional covariate adjustment models and those from the propensity score models. In the analysis of full study

population, there was little evidence for an increase in risk for SAB in coffee and caffeine exposure around the time of the telephone interview. When we restricted analysis to women who reported intake prospectively, pre-pregnancy and around time of interview consumptions were found to increase the risk for SAB. Among women who reported intake prospectively, no evidence for increase risk of SAB was found across three time points. **Conclusions:** This study was the first, to our knowledge, to apply propensity score methods in studying health behaviors in pregnancy. Given our results using both traditional and propensity score models, this indicates that our original models were robust enough to detect the differences on the risk of SAB across exposure groups. Moreover, this provides additional confirmation that our results from traditional covariate adjustment analysis were not as a result of biased estimation often found from using observational study data.

4.3.1 Introduction

Use of observational studies of maternal health exposures in pregnancy

The goal of many epidemiologic studies is to estimate the causal effect of broadly defined exposures on a health outcome,[53] and reproductive and perinatal health research is no exception. The use of randomized studies is not always justified or ethical in studying pregnancy health (e.g. randomizing pregnant women to receive potentially harmful exposures). Consequently, investigators often rely on observational research as an alternative for estimating causal effects. The causal interpretation from observational data is complicated because exposure groups may differ systematically with respect to

relevant observed covariates and, therefore, may not be directly comparable and such differences may lead to biased effect estimates.[17-20] Traditional covariate adjustment may not be sufficient to correct this bias in estimation due to a limited number of covariates for adjustment that can be used. Recently, attention has been paid to another approach to the control of confounding through the use of propensity scores.[17, 23]

The utility of propensity score methods in studying pregnancy health

Propensity score methods have been used in pharmacoepidemiology and other disciplines, but reproductive and perinatal epidemiologic studies have been slow to adopt this method. The concept of propensity scores estimated at baseline was developed by Rosenbaum and Rubin[23] to reduce bias in observational studies. In principle, the effect of an exposure can be measured among individuals with the same probability of being exposed, thus allowing control for confounding. Under the assumption that all relevant predictors of exposure have been adequately accounted for, the use of propensity scores to reduce bias is attractive, because individuals with the same propensity score have the same chance of being exposed even though in truth some were and some were not. Therefore, propensity score methods are often conceptualized as mimicking randomized trials (e.g. randomizing pregnant women to consume or not consume caffeine).[20]

In this paper we illustrate the use of propensity scores for studying the effects of health behaviors in pregnancy, a period of time in which many women may more carefully modulate their behaviors and exposures. We illustrate its application as an extension of an earlier study that evaluated the association between caffeine exposure at three time points and SAB.[24] Caffeine provides a useful example as it's consumption

around the time of pregnancy is influenced by a number of factors. Epidemiologic studies on caffeine consumption during pregnancy and risk of spontaneous abortion (SAB) have generated substantial public health interest since the 1980s; however conclusive evidence for or against an association has not been established.[21] SAB, which affects approximately 10% of all clinically-recognized pregnancies,[22] is defined as a pregnancy loss before 20 completed weeks' gestation. The concern about caffeine consumption during pregnancy is based on caffeine's ability to cross the placental membranes freely,[54] its decreased clearance rate in second and third trimesters,[55, 56] potential to influence cellular development by increasing cAMP,[107] and the interference with utero-placental circulation with increased catecholamines.[58] Methodological limitations in earlier epidemiologic studies,[21] including differences in measuring and quantifying exposure and control of confounding, pose a challenge to the interpretation of results from previous studies. To document the approach in estimating propensity scores and its use in regression modeling, a series of SAS codes (SAS Institute, Cary, NC) is provided at the end (Appendix 4).

Propensity Score

Under the assumption of strongly ignorable treatment or exposure assignment, a propensity score is defined as the conditional probability of exposure to a potential risk factor given an individual's set of observed covariates.[23] Each individual has a vector of observed covariates measured at baseline, X_i, and an indicator of exposure (or treatment), Z_i, where Z_i=1 if exposed and Z_i=0 if unexposed. The propensity score, $e(X_i)$ is the probability of exposure for a person with covariates X_i, that is,

$$e(X_i) = P(Z_i=1|X_i = x_i),$$

where we assume that the $Z_{i,}$ i =1, and n (number of study subjects) are independent conditional on the set of covariates X. The propensity score is a function of the observed covariates X such that the conditional distribution of X given the propensity score e(X) is the same for subjects with Z_i = 1 and with Z_i =0. Individuals with the same propensity score should have the same probability of being exposed, although they may have different sets of X's. Under the additional conditions of strongly ignorable treatment assignment X, (i.e., no systematic unobserved differences between subjects with Z=1 and Z=0), the estimated treatment effect at a fixed value of the propensity score is unbiased.

The propensity score is generally estimated from the observed data using multivariable logistic regression.[17] After they are estimated, propensity scores are used to control for selection bias or confounding. The three most common techniques that use propensity score include matching on propensity score, stratified analysis with propensity score as the stratification variable, and combining the use of conventional multivariable modeling with propensity scores.[17, 20] Matching is an approach that matches each exposed individual to at least one unexposed individual with similar propensity scores. Stratification is a modeling approach in which the investigator controls for propensity scores in propensity score-disease relationship, and the third approach allows the investigators to include the propensity score along with additional predictors for the outcome in the model.[17, 20] Because the use of propensity score methods allow

simultaneous control for confounding by several covariates, it is unnecessary to adjust for the same set of covariates that was used in propensity score estimation.

4.3.2 Methods

Study population

Data for this study came from a prospective cohort study on drinking water disinfection by-products and SAB (Right from the Start, RFTS). RFTS recruited and enrolled pregnant women at least 18 years old who were less than 12 weeks' gestation, and women between 18 to 45 years old who were trying to conceive for six months or less, from three U.S. cities (Raleigh, North Carolina; Memphis, Tennessee; and Galveston, Texas) from 2000 to 2004. Details of the study protocol can be found elsewhere.[83, 98] We collected detailed information on maternal characteristics, health behaviors, medical and reproductive histories, current pregnancy history, and pregnancy symptoms during two telephone interviews that took place at no later than 16 weeks' gestation and then between 20 and 25 weeks' gestation.

Assessment of caffeine consumption

Information on daily consumption of caffeinated beverages (coffee, iced tea, hot tea, and sodas) for a typical week was assessed during a telephone interview. Detailed questions were asked about all sources, including the number of cups consumed daily in a typical week, beverage sizes, and brewing method for coffee (brewed versus instant). We designated cups of coffee and hot tea as small (4 to 10 oz), medium (12 to 14 oz), large (16 to 24 oz); sizes for iced tea were designated as small (4 to 10 oz), medium (12 to 20

oz), and large (22 to 34 oz), and for sodas as small (8 to 12 oz), medium (14 to 22 oz), and large (24 to 34 oz). Reported consumption of small amounts less than one cup or glass was set to a small half cup or glass.

Information regarding changes in consumption for any caffeinated beverages, if any, was also collected as was the timing of changes down to the 'month', 'day', and 'year' for each beverage. When the 'day' was unknown but the "week in month" was reported, we imputed the day using the midpoint of the week; mid-point of the month was used to impute the day of event when only the month of the event was known.

Determination of caffeine exposure

With information collected at the time of the telephone interview, caffeine exposures from coffee alone and from all sources (mg/day) were calculated based on daily amount consumed (number of cups consumed and cup sizes), coffee brewing methods, and with the assignment of caffeine content[84] based on the midpoints of the cup sizes.

We then generated caffeine exposure (mg/day) at three time points [pre-pregnancy, four weeks after last menstrual period (LMP), and consumption at the time of the interview] for coffee only and total caffeine for each woman based on consumption assessed at the time of the interview and if applicable the self-reported consumption, prior to any change in consumption and the timing of change in intake. Total caffeine was the cumulative intake (mg/day) from caffeinated coffee, caffeinated iced and hot tea, and caffeinated sodas.

Probability of caffeine consumption during three time windows

We first estimated the probability of coffee consumption (yes/no) and consumption of any caffeinated beverages (yes/no) for three time points using logistic regression models. Covariates were selected to enter the logistic regression model included known and strong predictors of pregnancy loss, as well as other predictors of caffeine use, including maternal age, race and ethnicity, maternal education, marital status, smoking, alcohol use, parity, symptoms of nausea and vomiting (NVP), and pregnancy loss history. Estimated propensity scores for consumption for each subject were predicted probabilities from the logistic regression models.

After constructing the propensity scores, all individuals were subclassified into propensity score quintiles, as suggested by D'Agostino.[17] Because the propensity score is a scalar summary of the observed background covariates, women in the quintile can have different values for specific observed covariates, but, within the same stratum, women who consumed coffee or any caffeinated beverages and non-consumers would have similar distributions of the covariates. This subclassification was used in subsequent analysis models, with time to SAB as the outcome variable and caffeine exposure from coffee alone and from all sources as the main exposures.

Gestational age and outcome classifications

An endovaginal ultrasound was performed around eight weeks' gestation (median gestational age at ultrasound = 9.3 weeks) to confirm pregnancy viability and to ascertain the gestational age of the fetus.

We found the self-reported LMP in this cohort to be highly reliable and consistent with LMP dates assessed by ultrasound;[83, 108] therefore, self-reported LMP was used to date the onset of pregnancy. Pregnancy outcome data were ascertained via participant self-report and confirmed via medical records and/or corresponding vital or fetal death records. Early pregnancy loss or SAB is defined in this study as a loss of a pregnancy before 20 completed weeks' gestation as calculated from self-reported LMP.

Statistical analysis

We constructed discrete hazard models as unconstrained continuation odds ratio models to estimate the effects of exposure from coffee (or from all caffeine sources) on the odds of having a SAB at each gestational week, conditional on a woman's pregnancy having survived to the beginning of that gestational week.[92] We then compared results using the traditional approach of adjustment for confounders as covariates to propensity score models.

Caffeine exposure (mg/day) was examined as a categorical variable, with cutpoints in caffeine use (none, >0 to median, ≥median, and then separately for >75th percentile) generated in order to produce more easily interpretable estimates which allowed for comparison of results with earlier studies.

Potential confounders that were used in covariate adjustments for the traditional models included maternal age, race and ethnicity, maternal education, marital status, smoking, alcohol use, body mass index, vitamin use, NVP, pregnancy loss history, and induced abortion history, where covariates were retained in the final model if they

changed the effect estimates for the exposure by 10% or greater or if they were predictive of the outcome based on a p-value of <0.20.

In the propensity score models, we first examined caffeine effects within each propensity score quintile and used an interaction test to determine whether stratification was needed, based on p=0.05 as the significance level of interest. We found that the test of interaction for each time window did not meet our criteria of <0.05 (range: 0.10 to 0.98); therefore, propensity scores were adjusted as the main effects. Additionally, we did not adjust for any covariates in the propensity score models.

4.3.3 Results

We included a total of 2,407 women in the final analysis, of whom 258 (10.7%) had a SAB. Among those with pregnancy loss, 74 (29%) of loss occurred prior to the telephone interview and 184 (71%) took place after the interview. The majority of the study participants recruited were from Raleigh (45.2%) or Memphis (37.3%). The mean age at enrollment was 27.8 years and the mean gestational age at enrollment was 54.8 days. Non-Hispanic White women made up approximately 56% of the study population and 32% were Non-Hispanic Black; nine percent were of Hispanic ethnicity with the majority of these (76.0%) from the Galveston area. Our study population included many women who had completed at least college (49.0%), who were married (66.0%), non-smokers (94.6%), non-drinkers of alcoholic beverages during pregnancy (97.6%), or first time mothers (49.6%). Approximately 45% were considered as overweight or obese by the Institute of Medicine classification.[105] The majority of the study participants (88.4%) reported experiencing some form of NVP symptoms, with 35.4% having nausea

symptoms only and 53.0% having nausea symptoms with vomiting episodes. Approximately 79% of study population consumed caffeinated beverages prior to becoming pregnant, with 41% (n=777) reported consuming coffee. Coffee consumption decreased by more than 50% from pre-pregnancy to the time of the interview, whereas consumption of non-coffee caffeinated beverages dropped approximately 23% in the same time period. Among consumers, the median intake for coffee prior to pregnancy was 348.0 mg/day and the median intake declined to 200.8 mg/day by the time of the telephone interview. Median caffeine consumption among consumers decreased from 234.9 mg/day prior to pregnancy to 165.3 mg/day around the time of the telephone interview (data not shown).

In the analysis using the full study population, findings from propensity score models showed no substantial differences in the point estimates compared to results from the traditional multivariable outcome models, in which we saw that exposures from coffee alone and caffeine from all sources mostly showed no increase in risk for SAB. Most odds ratios (ORs) from the propensity score models moved closer to the null or remained the same as those obtained by the traditional models. Under the traditional covariate adjustments, there was little evidence for increase in risk for SAB for low (<200.8mg/day; OR=1.2, 95%CI: 0.7, 1.9) coffee exposure and some evidence for higher total caffeine exposure (OR=1.2, 95%CI: 0.9, 1.7 for above median and OR=1.3, 95%CI: 0.9, 1.9 for greater than 75th percentile) at the time of the interview. The ORs elevated slightly for total caffeine (1.3 and 1.4 respectively) under propensity score estimation, but such shift in the magnitude was minimal (Table 13). When the analyses were restricted to women who reported coffee and caffeine intake retrospectively (pregnancy loss

occurred before interview), we saw evidence of increase in risk for SAB in pre-pregnancy caffeine intake and for both coffee and total caffeine exposure around the time of the interview, though the results were imprecise in the traditional models. Propensity score model results also supported findings of a positive association in these time points, but the magnitude of effects were similar (Table 14). Among women who reported coffee and caffeine consumption prospectively (pregnancy loss occurred after interview), propensity score models yielded more precise estimates compared to the ORs from the traditional multivariable outcome models. We saw no evidence of an increase in risk for SAB with results from both modeling methods, with the exception of an imprecise OR (OR=3.0, 95%CI: 0.7, 13.2) for coffee consumption around the time of the interview. However, this OR dropped to 0.6 (95%CI: 0.2, 1.8) and precision of point estimates increased by 50% when the analysis was conducted using propensity score methods (Table 15).

4.3.4 Discussions

To our knowledge, this study is the first to apply propensity score methods in studying health behaviors in pregnancy. Pregnancy is a critical time in a woman's life; often a time when modification of health behaviors takes place for the health of the fetus or due to the advice of clinicians. Investigators are often interested in evaluating the effects of certain behaviors on pregnancy outcomes. Because studies of pregnancy health rely heavily on observational research study designs, women with exposure of interest may be systematically different than those who are unexposed, leading to potentially biased estimates of the maternal behaviors and outcome relationships.

We chose to re-visit the example of caffeine exposure and SAB due to its potential public health importance and conflicting results found in recent literature.[21] Data on caffeine exposure from observational studies may not be adequate in estimating caffeine effects on pregnancy outcome because some pregnant women may drastically modify their level of consumption during pregnancy. Because randomizing pregnant women to different levels of caffeine exposures cannot be done, the balancing property of propensity score method is an especially useful application over traditional multivariable analysis approach because this method mimics randomization. We applied the use of propensity score methods to our earlier study[24] which, like many others, relied on the use of traditional covariate adjustment methods to evaluate the caffeine-pregnancy loss relationship. With this new application we were able to compare the potentially biased estimates obtained from traditional models to the results from propensity score models; thereby providing us a glimpse on the effects of coffee and total caffeine on the risk of SAB if women were randomized to their exposure status, and that the observed covariates in consumers and non-consumers were evenly distributed.

Overall, we observed no meaningful differences between the estimates from the traditional models and those from the propensity score models. Results from the propensity score models supported our earlier findings for an increased risk in SAB for consumption of coffee alone and for total caffeine around the time of the interview, in the analyses among all women and of losses that took place prior to the interview (Tables 13 and 14). Nevertheless, in the analysis that was restricted to losses after the interview (Table 15), the single but highly imprecise estimate that showed a moderate increase in risk for SAB for high coffee consumption around the time of the interview from the

traditional model disappeared using the propensity score model that produced a more precise estimate not showing a positive association.

Propensity scores are estimated probabilities, based on each woman's measured characteristics, that she would consume or not consume caffeine; this implies that women with the same estimated propensity score have equal probability of 'becoming exposed'. The robust properties of the propensity scores allowed us to examine the association between caffeine exposure and SAB under the context of "virtual randomization", in which individuals in the consumer and non-consumer groups had overall similar risk factors that predict SAB. This form of grouping eliminated the potential biases as a result of the systematic differences between the exposure groups with respect to the observed covariates. We first examined the distribution of propensity scores by caffeine variables (using proc univariate); we then performed additional checks by examining the distribution of the covariates by each exposure group, stratified by propensity score quintiles.[20] The covariates we used did an adequate job in estimating the propensity of coffee and total caffeine consumption at three time points and we were able to confirm that balanced covariates were achieved across both groups (consumers and non-consumers) from the stratified analysis. Given we saw no substantial changes from the original results observed using traditional models compared to the estimates from the propensity score models, this indicated that our original models were robust enough to detect the differences on the risk of SAB among women who consumed caffeine and women who did not consume. This makes it less likely that our findings on the lack of positive association between caffeine intake and SAB in our earlier study are a result of biased estimation that often plagues observational studies.

Unlike other studies that recruited their study populations from prenatal care clinics, we identified and recruited a population of pregnant women early in gestation and women who were trying to conceive. With the use of a prospective cohort study design, we collected detailed information as pregnancy progresses to discern patterns of caffeinated coffee use and generated exposure for coffee alone and caffeine from all sources. Because of our timing in recruitment and collection of maternal health information, with caffeine assessment taking place prior to loss among 67% with loss, differential misclassification of their exposures due to potential recall bias was minimized. In the analysis of loss before interview, the increase in risk for SAB we observed for coffee and caffeine consumption around the time of interview could have come from reporting bias from early pregnancy loss or a return of pre-pregnancy consumption level due to the lack of influence from NVP symptoms.

Despite improvements in our study design, limitations still exist. Our study did not collect coffee, iced tea, hot tea, and soda consumption information weekly; rather, we placed assumptions that the reported consumption levels before and after change were maintained in the gestational weeks before and after the change. Consequently, we could precisely capture multiple changes in consumption that may have occurred throughout the pregnancy. A few practical limitations persist, despite the many advantages for applying propensity score methods on observational study data. Unlike true randomization, where the study design removes bias from both measured and unmeasured factors, residual bias may still persist because propensity scores can only be estimated from measured data and bias cannot be controlled for unmeasured covariates or for variables that were measured poorly.[20] Although many current literature addresses the

theoretical framework and the advantages of propensity score methods, very few have addressed other limitations such as issues with handling missing data and systematic variable selection.[17, 109]

The use of propensity score methods is a novel approach in studying health exposures in pregnancy. While it should not be regard as the sole method in the analyses of observational study data,[17, 20] its application is an important addition to traditional modeling methods for investigators to evaluate exposure and outcome relationships by reducing bias in estimates and for improving inferences on causal effects.

4.3.5 References

Table 13. Results using traditional and propensity score models. Coffee and caffeine consumption and the risk of pregnancy loss, contrasting none, below or equal to the median, above the median, and above the 75th percentile: all pregnancy loss (n=258)

	Traditional models*				Propensity score models	
Coffee Exposure (mg/day)	# Losses	Unadj. OR‡	Adj.† OR	95% CI§	Adj. OR	95%CI
Pre-pregnancy						
None	176	1.0	1.0	-	1.0	-
>0 - <348.0	32	0.8	0.9	(0.6, 1.3)	0.8	(0.5, 1.2)
≥348.0	49	0.9	0.8	(0.5, 1.1)	0.9	(0.6, 1.2)
>696.0	18	1.0	0.9	(0.5, 1.5)	1.0	(0.6, 1.7)
Four weeks post LMP#						
None	189	1.0	1.0	-	1.0	-
>0 - <348.0	27	0.8	0.8	(0.5, 1.2)	0.8	(0.5, 1.2)
≥348.0	42	1.0	0.9	(0.6, 1.3)	1.0	(0.7, 1.5)
>602.3	15	0.8	0.7	(0.4, 1.2)	0.8	(0.5, 1.4)
Time of telephone interview						
None	211	1.0	1.0	-	1.0	-
>0 - <200.8	19	1.1	1.2	(0.7, 1.9)	1.2	(0.7, 1.9)
≥200.8	27	1.2	1.0	(0.6, 1.6)	1.1	(0.7, 1.8)
>372.9	5	0.8	0.7	(0.3, 1.8)	0.7	(0.3, 1.8)

*Savitz et al. Caffeine and miscarriage risk. *Epidemiology* 2008;19:55-62

†Model adjusted for maternal age, race/ethnicity, maternal education, marital status, alcohol use, vitamin use, and symptoms of nausea and vomiting during early pregnancy

‡OR: Odds ratio

§CI: Confidence interval

#LMP: Last menstrual period

Table 13 (cont). Results using traditional and propensity score models. Coffee and caffeine consumption and the risk of pregnancy loss, contrasting none, below or equal to the median, above the median, and above the 75th percentile: all pregnancy loss (n=258)

	Traditional models*				Propensity score models	
Total Caffeine Exposure (mg/day)	# Losses	Unadj. OR‡	Adj.† OR	95% CI§	Adj. OR	95%CI
Pre-pregnancy						
None	52	1.0	1.0	-	1.0	-
>0 - <243.7	104	1.0	1.0	(0.7, 1.4)	1.0	(0.7, 1.4)
≥243.7	101	0.9	0.9	(0.6, 1.2)	1.0	(0.7, 1.4)
>513.2	48	0.9	0.8	(0.5, 1.2)	0.9	(0.6, 1.4)
Four weeks post LMP§						
None	70	1.0	1.0	-	1.0	-
>0 - <210.3	93	1.1	1.0	(0.7, 1.4)	1.0	(0.8, 1.4)
≥210.3	95	1.1	1.0	(0.7, 1.3)	1.0	(0.8, 1.4)
>463.1	47	1.0	0.9	(0.6, 1.3)	1.0	(0.7, 1.4)
Time of telephone interview						
None	102	1.0	1.0	-	1.0	-
>0 - <144.3	71	1.1	1.0	(0.7, 1.4)	1.1	(0.8, 1.5)
≥144.3	84	1.4	1.2	(0.9, 1.7)	1.3	(1.0, 1.8)
>273.2	46	1.5	1.3	(0.9, 1.9)	1.4	(1.0, 2.1)

*Savitz et al. Caffeine and miscarriage risk. *Epidemiology* 2008;19:55-62

†Model adjusted for maternal age, race/ethnicity, maternal education, marital status, alcohol use, vitamin use, and symptoms of nausea and vomiting during early pregnancy

‡OR: Odds ratio

§CI: Confidence interval

#LMP: Last menstrual period

Table 14. Results using traditional and propensity score models. Coffee and caffeine consumption and the risk of pregnancy loss, contrasting none, below or equal to the median, above the median, and above the 75th percentile: pregnancy loss before interview (n=74)

	Traditional models*				Propensity score models	
Coffee Exposure (mg/day)	# Losses	Unadj. OR‡	Adj.† OR	95% CI§	Adj. OR	95%CI
Pre-pregnancy						
None	45	1.0	1.0	-	1.0	-
>0 - <348.0	11	0.8	0.8	(0.3, 2.0)	0.7	(0.3, 1.7)
≥348.0	17	1.2	1.0	(0.6, 1.9)	1.1	(0.6, 2.1)
>696.0	7	1.8	1.5	(0.7, 3.5)	1.5	(0.7, 3.6)
Four weeks post LMP#						
None	50	1.0	1.0	-	1.0	-
>0 - <348.0	11	0.9	1.0	(0.4, 2.2)	0.9	(0.4, 2.1)
≥348.0	13	1.2	1.0	(0.5, 1.9)	1.1	(0.6, 2.2)
>602.3	4	0.9	0.8	(0.3, 2.3)	0.9	(0.3, 2.5)
Time of telephone interview						
None	53	1.0	1.0	-	1.0	-
>0 - <200.8	7	1.7	1.9	(0.8, 4.4)	1.6	(0.7, 3.9)
≥200.8	13	2.3	1.8	(0.9, 3.7)	2.1	(1.0, 4.2)
>372.9	2	1.4	1.1	(0.3, 4.5)	1.3	(0.2, 5.3)

*Savitz et al. Caffeine and miscarriage risk. *Epidemiology* 2008;19:55-62

†Model adjusted for maternal age, race/ethnicity, maternal education, marital status, alcohol use, vitamin use, and symptoms of nausea and vomiting during early pregnancy

‡OR: Odds ratio

§CI: Confidence interval

#LMP: Last menstrual period

Table 14 (cont). Results using traditional and propensity score models. Coffee and caffeine consumption and the risk of pregnancy loss, contrasting none, below or equal to the median, above the median, and above the 75th percentile: pregnancy loss before interview (n=74)

	Traditional models*				Propensity score models	
Total Caffeine Exposure (mg/day)	# Losses	Unadj. OR‡	Adj.† OR	95% CI§	Adj. OR	95%CI
Pre-pregnancy						
None	9	1.0	1.0	-	1.0	-
>0 - <243.7	33	1.6	1.6	(0.7, 3.4)	1.6	(0.8, 3.4)
≥243.7	31	1.3	1.2	(0.6, 2.6)	1.3	(0.6, 2.9)
>513.2	14	1.2	1.1	(0.4, 2.6)	1.2	(0.5, 2.9)
Four weeks post LMP§						
None	18	1.0	1.0	-	1.0	-
>0 - <210.3	28	1.2	1.2	(0.6, 2.2)	1.3	(0.7, 2.5)
≥210.3	28	1.1	1.0	(0.5, 2.0)	1.1	(0.6, 2.2)
>463.1	11	0.6	0.5	(0.2, 1.4)	0.6	(0.2, 1.5)
Time of telephone interview						
None	24	1.0	1.0	-	1.0	-
>0 - <144.3	18	1.3	1.1	(0.6, 2.2)	1.3	(0.7, 2.5)
≥144.3	31	2.1	1.9	(1.1, 3.5)	1.9	(1.0, 3.5)
>273.2	18	2.6	2.3	(1.2, 4.5)	2.2	(1.1, 4.4)

*Savitz et al. Caffeine and miscarriage risk. *Epidemiology* 2008;19:55-62

†Model adjusted for maternal age, race/ethnicity, maternal education, marital status, alcohol use, vitamin use, and symptoms of nausea and vomiting during early pregnancy

‡OR: Odds ratio

§CI: Confidence interval

#LMP: Last menstrual period

Table 15. Results using traditional and propensity score models. Coffee and caffeine consumption and the risk of pregnancy loss, contrasting none, below or equal to the median, above the median, and above the 75th percentile: pregnancy loss after interview (n=184)

	Traditional models*				Propensity score models	
Coffee Exposure (mg/day)	# Losses	Unadj. OR‡	Adj.† OR	95% CI§	Adj. OR	95%CI
Pre-pregnancy						
None	131	1.0	1.0	-	1.0	-
>0 - <348.0	21	1.0	1.2	(0.7, 2.1)	0.8	(0.5, 1.3)
≥348.0	32	1.0	0.8	(0.5, 1.3)	0.8	(0.5, 1.2)
>696.0	11	1.0	1.1	(0.5, 2.3)	0.8	(0.5, 1.6)
Four weeks post LMP#						
None	139	1.0	1.0	-	1.0	-
>0 - <348.0	16	0.9	1.1	(0.6, 2.1)	0.7	(0.4, 1.2)
≥348.0	29	1.0	0.8	(0.5, 1.2)	1.0	(0.6, 1.5)
>602.3	11	1.0	0.6	(0.3, 1.4)	0.8	(0.4, 1.5)
Time of telephone interview						
None	158	1.0	1.0	-	1.0	-
>0 - <200.8	17	1.0	0.7	(0.3, 1.6)	1.0	(0.6, 1.8)
≥200.8	9	1.0	1.0	(0.5, 1.8)	0.9	(0.5, 1.5)
>372.9	3	2.9	3.0	(0.7, 13.2)	0.6	(0.2, 1.8)

*Savitz et al. Caffeine and miscarriage risk. *Epidemiology* 2008;19:55-62

†Model adjusted for maternal age, race/ethnicity, maternal education, marital status, alcohol use, vitamin use, and symptoms of nausea and vomiting during early pregnancy

‡OR: Odds ratio

§CI: Confidence interval

#LMP: Last menstrual period

Table 15 (cont). Results using traditional and propensity score models. Coffee and caffeine consumption and the risk of pregnancy loss, contrasting none, below or equal to the median, above the median, and above the 75th percentile: pregnancy loss after interview (n=184)

	Traditional models*				Propensity score models	
Total Caffeine Exposure (mg/day)	# Losses	Unadj. OR‡	Adj.† OR	95% CI§	Adj. OR	95%CI
Pre-pregnancy						
None	43	1.0	1.0	-	1.0	-
>0 - <243.7	71	1.0	1.1	(0.7, 1.7)	0.8	(0.6, 1.2)
≥243.7	70	1.0	1.0	(0.6, 1.6)	0.9	(0.6, 1.3)
>513.2	34	1.1	0.8	(0.5, 1.5)	0.8	(0.5, 1.3)
Four weeks post LMP§						
None	52	1.0	1.0	-	1.0	-
>0 - <210.3	65	0.9	1.1	(0.7, 1.7)	1.0	(0.7, 1.4)
≥210.3	67	0.9	0.8	(0.5, 1.3)	1.0	(0.7, 1.5)
>463.1	36	0.9	0.8	(0.5, 1.3)	1.1	(0.7, 1.7)
Time of telephone interview						
None	78	1.0	1.0	-	1.0	-
>0 - <144.3	53	0.8	0.9	(0.6, 1.4)	1.0	(0.7, 1.4)
≥144.3	53	1.1	1.1	(0.7, 1.7)	1.2	(0.8, 1.7)
>273.2	28	1.1	1.1	(0.6, 1.8)	1.2	(0.8, 1.9)

*Savitz et al. Caffeine and miscarriage risk. *Epidemiology* 2008;19:55-62

†Model adjusted for maternal age, race/ethnicity, maternal education, marital status, alcohol use, vitamin use, and symptoms of nausea and vomiting during early pregnancy

‡OR: Odds ratio

§CI: Confidence interval

#LMP: Last menstrual period

CHAPTER 5: DISCUSSION

5.1 Summary of findings

In this study, we found that most maternal characteristics, with the exception of plurality, were not associated with having NVP. There was a dose-response gradient for increasing risk for delayed symptom onset with increasing maternal age; increased risk for delayed onset was also found among non-Hispanic Blacks, Hispanic women, or women who began menstruating between 12 and 13 years old. Additionally, our findings indicated that older maternal age and non-Hispanic Blacks were less likely to experience NVP symptoms lasting beyond the first trimester.

Findings of this study provided evidence that the lack of NVP symptoms is associated with an increased risk for SAB, compared to having any symptoms. Maternal age was found to modify the association between symptom severity and duration and SAB, with increasing maternal age found to strengthen the risk for SAB among women without any symptoms. Reduced risks for SAB were found across all maternal age groups for those with longer symptoms duration, but the effects were much stronger in the oldest maternal age group.

Propensity score methodology was applied in the third manuscript to study the effects of caffeine exposure (coffee only and total caffeine from all sources) and SAB. We found no overall meaningful differences between estimates from the traditional covariate adjustment models and the propensity score models. In the analysis of full study

population, there was a suggestion of increased risk for SAB for caffeine exposure from coffee alone and from all sources around the time of interview. In the analysis of women who reported intake retrospectively, pre-pregnancy and around time of interview consumptions were found to increase the risk for SAB. When we restricted analysis to women who reported intake prospectively, we saw no evidence for an increase in risk of SAB.

5.2 Strengths and limitations

5.2.1 Prospective data collection on maternal health exposures and behaviors

One strength of this study is the use of maternal exposure and behavioral data that were collected prospectively from women who were early in their pregnancy. Extensive information was collected via two telephone interviews and included maternal health behaviors (e.g. caffeine consumption), reproductive and medical histories, current pregnancy history and symptoms (e.g. NVP symptoms), and water exposure characteristics. RFTS collected information on the timing of onset and ending dates separately for symptoms of nausea and for vomiting episodes. NVP onset and ending dates were collected as month, day, and year, as well as timing information in "week in the month". Detailed information was also collected on consumption of caffeinated beverages (coffee, iced tea, hot tea, and sodas). Questions were asked about all four sources, including the number of cups consumed daily, beverage sizes, and brewing method for coffee. Furthermore, information regarding changes in consumption (timing and amount) was also collected.

In this study, we collected information directly from study participants early in pregnancy (i.e. when recall is more likely accurate); however, because variation in symptoms of NVP is known to exist daily or weekly, one main draw back in this study is that NVP data were not collected daily. A more extensive mode of data collection (e.g. daily diaries) would have captured the severity, timing of onset, and symptom duration of NVP symptoms more accurately. This form of data collection should also be applied to caffeine exposure assessments. Given we did not collect coffee, iced tea, hot tea, and soda consumption information weekly; rather, we placed assumptions that the reported consumption levels before and after change were maintained throughout the pregnancy. Consequently, we were unable to precisely capture the variations in consumption that may have taken place throughout the pregnancy.

5.2.2 Subclassification of nausea and vomiting symptoms

With our extensive collection of information on NVP symptoms, another strength of this study is our ability to examine NVP with respect to symptom severity, timing of symptom onset, and symptom duration. No other study of NVP and SAB to date has examined the effects of NVP subgroups, and only one other recent study[42] explored potential maternal health factors that may potentially influence the NVP subgroups.

5.2.3 Prospective follow-up on study subjects and outcome assessment

Because of the study design, another strength of this study is the prospective

follow-up of each woman to determine pregnancy outcome. Many previous studies have solely relied on hospital records or vital records to estimate gestational age and to ascertain pregnancy outcomes. Unlike previous studies, we confirmed current pregnancy status (e.g. still pregnant, had SAB, etc.) with each woman at the start of each telephone interview, and we utilized three different sources (participant self-report, medical chart abstraction, and vital and fetal death record linkages) for the confirmation and verification of pregnancy outcome. This approach is especially important for the identification of SABs in a timely manner. Moreover, first trimester endovaginal ultrasounds were used to confirm the viability of pregnancy and to ascertain gestational age of the fetus. We found in this study that self-reported LMP to be highly reliable in dating pregnancy onset; therefore estimated gestational age in this study population was primarily determined based on women's self-reported LMP and secondarily with ultrasound dating. If both self-reported LMP and LMP from ultrasound data were available, ultrasound data were available to adjust LMP estimates of gestational age if they differed for more than seven days or when LMP dates were unavailable or deemed invalid.

Although this study followed women prospectively through their pregnancies, limitations still exist. Our study population, unlike those used in previous studies, represents a selective group of women who were very early in their pregnancy or were actively trying to conceive a pregnancy. Consequently, they may be more apt to modify their behavior and adopt a healthier lifestyle prior to or around the time of study enrollment. Although we ascertained information on behavior prior to change (along with the date of change), it is possible that this still leads to an underestimation of the true

exposure and a bias in estimates. Maternal age of RFTS study participants were found to be comparable to the general population in the same geographic location but different in other maternal characteristics, where RFTS women were more likely to be highly educated (≥ 16 years), non-Hispanic White (Raleigh and Memphis) and Hispanic (Galveston), and nulliparous.[83]

5.2.4 Use of discrete-time continuation ratio logistic survival model

Studies of SABs are subject to left truncation because women tend to enroll into a study at various gestational ages and women without a viable pregnancy prior to study enrollment, who would have otherwise been eligible to participate, are missing from the study population.[70, 71] Analytic techniques used previously in many studies on maternal exposures and SAB posed limitations, especially with the use of logistic regression in prospective cohort studies. Because logistic regression does not account for left truncation and assumes equal follow-up time for all subjects, it is especially problematic when the left truncation is differential with respect to exposure status; thus fitting logistic regression models would create a biased estimate of effect from SAB data. In this study, we used discrete-time continuation ratio logistic regression model (a.k.a. discrete hazard model), with is a marked improvement in statistical methods over previous studies. With this technique, we were able to handle the concerns with left truncation and to account for delay entry by modeling the probability of a woman having a pregnancy loss at a given week, conditional on her not having experienced a loss in the previous week.

5.2.5 Application of propensity score methods in studying pregnancy health

The application of propensity score methods in studying pregnancy health is the final strength of this study. Although the propensity scores have been widely used in pharmacoepidemiology, reproductive and perinatal epidemiologic studies have been slow to adopt this method. Because causal interpretation from observational data has limitations and the effect estimates using traditional covariate adjustments may be biased, propensity score methods allowed for the distribution of the covariates to be balanced between exposure groups, mimicking "virtual randomization" to an extent. We illustrated the use of propensity scores in examining caffeine exposure and SAB and we were able to compare results using propensity scores and findings from traditional multivariable regression analysis. Although this method reduces bias known to exist in observational data, it is not without limitations. Unlike randomized studies where measured and unmeasured covariates are balanced across exposure groups, residual bias may linger because propensity scores can only be estimated from measured data; therefore, bias from unmeasured covariates cannot be accounted for. Furthermore, other issues concerning systematic variable selection and handling of missing data[110-112] must be addressed. Despite its current limitations, propensity score methods provide a useful addition to examine factors in pregnancy health that otherwise would not be appropriate for randomization.

5.3 Public health implications

5.3.1 Symptoms of NVP

Little is known about the natural history of early pregnancy, and pregnant women may not know what to expect with regard to symptoms. Symptoms of NVP is a common phenomenon that affects majority of pregnant women; symptoms can differ in severity, timing of onset, and duration, both among individual women and among each pregnancy of the same woman. Secretion of hCG and other hormones, as well as the rapid rise in their levels have been proposed as the most likely contributors. Few published studies suggested several maternal characteristics as potential risk factors for having NVP, but other studies yield conflicting results on these factors.

The presence of NVP, though it can be unpleasant and potentially be disruptive in a pregnant women's daily life, is a marker for favorable pregnancy outcomes. However, NVP may lead to serious consequence for a small percent of pregnant women, as evidenced by hyperemesis gravidarum. Our study did not evaluate hyperemesis gravidarum or medical treatment for NVP, but our data indicated that 20% of women in our study population contacted their clinicians and a handful was hospitalized due to NVP. Findings from this study suggest that maternal age and maternal race/ethnicity are associated with most aspects of NVP, while other maternal characteristics are associated with selected NVP subtypes. Also, our findings confirmed that the risk for SAB increases with nausea alone and even more so for women no NVP symptoms; there was additional evidence that suggested that longer symptom duration reduced the risk for SAB and maternal age may play a vital role in these effects. There are no formal treatments for NVP symptoms and clinical management varies depending on the severity of the

symptoms; nevertheless, careful management of these symptoms is critical for pregnant women and their pregnancies.

5.3.2 Caffeine consumption

Caffeine is readily available and widely consumed in most countries. Since the 1980s, issues surrounding caffeine consumption for pregnant women have generated much public health attention and controversy. Caffeine is a central nervous stimulant and impacts a pregnant woman differently; as evidence by its ability to cross placental membranes and decrease in clearance in second and third trimesters. Epidemiological data have suggested caffeine consumption may be associated with adverse pregnancy outcomes, notably early pregnancy loss. However, due to differences in study designs and other methodological limitations in previous epidemiologic studies, the association between caffeine exposure and SAB has not been firmly established.

Most studies to date have not been able to identify whether the effects of caffeine, if any, are restricted to certain peri-conceptional or pregnancy time window, or have been able to identify a threshold or a "maximum exposure limit" so that caffeine consumption does not affect the risk of SAB. We were able to examine a wide range of caffeine exposure (coffee alone and total caffeine from coffee, iced and hot tea, and sodas) at three time windows using two forms of statistical methods. While an increased risk of SAB was suggested for consumption around the time of interview, findings from this study showed very little evidence that overall caffeine consumption increases the risk for SAB.

5.4 Conclusions

This study addressed two important issues concerning pregnant women. Using a prospective cohort study design, exposure and outcome assessment methods were greatly improved in many ways over those used in previous studies. The prospective follow-up of study participants allowed us to ascertain exposure data early in pregnancy and capture information from the study questionnaire in a very detailed manner.

While many hypotheses have been proposed, the etiology of NVP symptoms remains largely unknown. Along with maternal characteristics, behaviors, and daily exposures, various intricate biological and physiological processes may be triggered, which in turn, could dictate whether a pregnant woman is more prone to NVP symptoms. Given its complexity, narrowing down specific causes for the onset of NVP symptoms is more challenging. RFTS did not collect biological specimens; nevertheless, our findings were able to provide additional insight that women with certain maternal traits were more prone to exhibit specific NVP characteristics.

With our findings, along with results found in previous studies, the most definitive conclusion we can make is that the lack of NVP symptoms is associated with an increased risk for SAB. Compared to having both symptoms of nausea and vomiting, our data also indicated an increased risk for SAB with nausea symptoms alone, but the magnitude of effect was not as strong as pregnancies without any symptoms. To our knowledge, we were the first study that examined the association between symptom duration (length in weeks of gestation) with SAB, and the results provided evidence that longer symptom duration decreased the risk for SAB. With the exception of a few notable

theories, the mechanisms by which NVP predicts favorable pregnancy outcome are still not known. Future studies should focus on further refinement in measuring the timing and occurrence of NVP in order to provide a better understanding of the biological and endocrinological mechanisms between the characteristics evaluated in this study and other potential risk factors to NVP symptoms, and its relationship with pregnancy outcomes.

Finally, our findings in this study of largely null results were based on enhanced exposure assessments compared to previous studies where positive associations for caffeine exposure and SAB were detected. We then applied propensity score methods to further confirm our findings, which suggested that our null findings were not as a result of biased estimation from using observational data. Because methodological limitations from past and recent studies provide challenges to draw definitive conclusions about the effects of caffeine exposure on the risk of SAB, future studies must strive to overcome current methodological challenges by integrating detailed exposure assessments (e.g. daily diaries) and close follow-up of pregnancy outcomes. Until investigators can find ways to overcome these limitations, conclusions about the influence of caffeine exposure and SAB would likely remain inconclusive.

APPENDIX 1. QUESTIONS ON NAUSEA AND VOMITING DURING EARLY PREGNANCY IN THE BASELINE QUESTIONNAIRE: RIGHT FROM THE START

Questionnaire shown in Appendix 1 was abstracted from the Right from the Start Baseline Questionnaire (Section H, Current Pregnancy History, pp. 43-46), as prepared by Savitz (co-Principal Investigator) and colleagues at the Department of Epidemiology, The University of North Carolina at Chapel Hill (CATI Version 2, October 25, 2001. Text updated to match CATI on December 16, 2002).

Section H Current pregnancy history
This next set of questions is about your current pregnancy

Nausea

H3a. Some women experience days or weeks of nausea or feeling sick to their stomach when they are pregnant while other women don't. At any time during <u>this</u> pregnancy, have you had a feeling of nausea?
no → skip to H7a. yes don't know/refused

H3b. When did you first notice the nausea?
month: ____ day: _____ *[if doesn't remember day ask H3c.]* year: _____
❑ doesn't remember/refused

H3c. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?
___1st ____2nd ___ 3rd ____ 4th ___ 5th
❑ doesn't remember/refused

H3d. Has the nausea gotten worse?
yes →skip to H4a. no don't know/refused

H3e. Has this feeling of nausea stopped completely?
yes no → Skip to H4a. don't know/refused

H3f. When did the nausea stop completely?
month: ____ day: _____ *[if doesn't remember day ask H3c.]* year: _____
❑ doesn't remember/refused

H3g. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?
___1st ____2nd ___ 3rd ____ 4th ___ 5th
❑ doesn't remember/refused

Vomiting

H4a. During this pregnancy, have you had nausea so bad that you vomited?
Yes
No → skip to H5.
don't know/refused

H4b. When did you start vomiting?
month: ____ day: _____ *[if doesn't remember day ask H4c.]* year: _____
❑ doesn't remember/refused

H4c. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?
___1st ____2nd ___ 3rd ____ 4th ___ 5th
❑ doesn't remember/refused

H4d. When your vomiting was the worst, how many days a week did you vomit?
_____ # day(s)

H4e. When your vomiting was the worst, how many times each day did you vomit? Did you vomit once a day, 2 times a day, 3 times a day or more than 3 times a day?
❑ once a day
❑ 2 times a day
❑ 3 times a day
❑ more than 3 times a day
❑ don't know

H4f. Has the vomiting stopped completely?
no → skip to H5. yes don't know/refused

H4g. When did you stop vomiting?
month: ____ day: _____ *[if doesn't remember day ask H4h.]* year: _____
❑ doesn't remember/refused

H4h. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?
___1st ____2nd ___ 3rd ____ 4th ___ 5th

- ❑ doesn't remember/refused

H5. Has/did the nausea and/or vomiting caused/cause you to _____?
[Program out questions that don't apply, ie. if she doesn't drink coffee (C1a. and C1e. equal 0) then don't ask the question about coffee]

a. change the amount of food you eat	Yes → PROBE	No	Don't know
	Did it cause you to eat more or less?		
b. change the amount of coffee you drink	Yes → More or less?	No	Don't know
c. change the amount of alcohol you drink	Yes → More or less?	No	Don't know
d. change the amount of soda you drink	Yes → More or less?	No	Don't know
e. change the amount of water you drink	Yes → More or less?	No	Don't know
f. change the amount of exercise you do	Yes → More or less?	No	Don't know
g. change the number of cigarettes you smoke	Yes → More or less?	No	Don't know
h. change the number of hours you work	Yes → More or less?	No	Don't know
i. change your normal daily activities	Yes	No	Don't know
j. not be able to take your prenatal vitamin or supplements	Yes	No	Don't know
k. to change anything else in your life?	Yes → PROBE	No	Don't know
	Specify:		

H6a. Have/Did you contact(ed) a doctor or nurse specifically because of the nausea and/or vomiting? *[do not count regular prenatal visit, only count if she specifically called about the nausea or vomiting]*
no yes don't know/refused

H6b. Did you take/Are you taking any medicine, vitamins, or supplements to help stop the nausea and/or vomiting?
no → skip to H6d. yes don't know/refused → skip to H6d.

H6c. What was/is the name of the medicine you took/are taking?
_____

H6d. Were you put in the hospital for the nausea and/or vomiting?
no yes don't know/refused

H6e. During the time [you were experiencing/you've been experiencing] the nausea and/or vomiting,
[did you lose weight / have you lost weight],
[did you fail to gain weight/have you failed to gain weight], or
[did you gain weight/have you been gaining weight] as desired?
lost weight
didn't gain weight → skip to H7a
gained → skip to H7a
don't know/refused → skip to H7a.
other: (specify) _____

H6f. How many pounds have/did you lost/lose?
_____ #lbs

APPENDIX 2. QUESTIONS ON NAUSEA AND VOMITING DURING EARLY PREGNANCY IN THE FOLLOW-UP QUESTIONNAIRE: RIGHT FROM THE START

Questionnaire shown in Appendix 2 was abstracted from the Right from the Start Still-Pregnant Follow-Up (Section H, Current Pregnancy History, pp. 23-24), as prepared by Savitz (co-Principal Investigator) and colleagues at the Department of Epidemiology, The University of North Carolina at Chapel Hill (CATI Version 2, October 15, 2001. Text updated to match CATI on December 16, 2002).

Section H Current pregnancy history
This next set of questions is about your current pregnancy

Nausea

H3a. Some women experience days or weeks of nausea or feeling sick to their stomach when they are pregnant while other women don't. At any time during <u>this</u> pregnancy, have you had a feeling of nausea?
no → skip to H7a. yes don't know/refused

H3b. As best as you can remember, when did you first notice the nausea?
month: ____ day: _____ *[if doesn't remember day ask H3c.]* year: _____
❑ doesn't remember/refused

H3c. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?
___1st ____2nd ___ 3rd____ 4th ___ 5th
❑ doesn't remember/refused

H3e. Has this feeling of nausea stopped completely?
yes no → Skip to H4a. don't know/refused

H3f. When did the nausea stop completely?
month: ____ day: _____ *[if doesn't remember day ask H3c.]* year: _____
❑ doesn't remember/refused

H3g. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?

___1st ____2nd ___ 3rd ____ 4th ___ 5th

- ❑ doesn't remember/refused

Vomiting

H4a. During this pregnancy, have you had nausea so bad that you vomited?

Yes No → skip to H5. don't know/refused

H4b. As best as you can remember, when was the first time that you vomited?

month: ____ day: _____ *[if doesn't remember day ask H4c.]* year: _____

- ❑ doesn't remember/refused

H4c. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?

___1st ____2nd ___ 3rd ____ 4th ___ 5th

- ❑ doesn't remember/refused

H4f. Has the vomiting stopped completely?

no → skip to H5. yes don't know/refused

H4g. When was the last time you vomited?

month: ____ day: _____ *[if doesn't remember day ask H4h.]* year: _____

- ❑ doesn't remember/refused

H4h. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?

___1st ____2nd ___ 3rd ____ 4th ___ 5th

- ❑ doesn't remember/refused

H6a. Have/Did you contact(ed) a doctor or nurse specifically because of the nausea and/or vomiting? *[do not count regular prenatal visit, only count if she specifically called about the nausea or vomiting]*

no yes don't know/refused

H6b. Did you take/Are you taking any medicine, vitamins, or supplements to help stop the nausea and/or vomiting?

no → skip to H6d. yes don't know/refused → skip to H6d.

H6c. What was/is the name of the medicine you took/are taking?

_____

H6d. Were you put in the hospital for the nausea and/or vomiting?

no yes don't know/refused

H6e. During the time [you were experiencing/you've been experiencing] the nausea and/or vomiting,
[did you lose weight / have you lost weight],
[did you fail to gain weight/have you failed to gain weight], or
[did you gain weight/have you been gaining weight] as desired?
lost weight
didn't gain weight → skip to Pain
gained → skip to Pain
don't know/refused → skip to Pain
other: (specify) _____

H6f. How many pounds have/did you lost/lose?
_____ #lbs

APPENDIX 3. QUESTIONS ON COFFEE CONSUMPTION IN THE BASELINE QUESTIONNAIRE: RIGHT FROM THE START

Questionnaire shown in Appendix 3 was abstracted from the Right from the Start Baseline Questionnaire (Section C, Health Behaviors, pp. 9-13), as prepared by Savitz (co-Principal Investigator) and colleagues at the Department of Epidemiology, The University of North Carolina at Chapel Hill (CATI Version 2, October 25, 2001. Text updated to match CATI on December 16, 2002).

The next questions are about caffeine intake. I will ask you about how much coffee, tea, and soda you usually drink per day.

Ca. For the questions about caffeine intake, please think about what you drank over the past week when answering what you currently drink in a typical day. Try to answer these questions about your usual habits even if some days you drank a different amount than on other days.

Was this past week a typical week for you, meaning that you weren't on vacation or there wasn't anything unusual that would affect your coffee, tea or soda use?
Yes
No → then, for all the following questions please think about a different week that you would consider a more typical week for you.

Caffeinated coffee
[interviewer notes: include hot and/or cold coffee]

C1a. Currently, how many cups of caffeinated coffee do you usually drink per day?
____ # per day ____<1 per day → *skip to C1c.* none→ *skip to C1c.*
____ don't know/refused → *skip to C1c.*

C1asize. Are those cups usually small like a tea cup, about 4-10 oz; medium like a coffee mug, about 12-14oz; or large like an travel mug or oversized coffee mug, about 16-24oz?
Sm Med Lg Other: specify _____ don't know/refused

C1b. Do you usually drink instant or brewed coffee? *[brewed includes espresso]*
brewed instant both *[only if 50/50]* don't know/refused

C1c. In the past 4 months, since [date], have you changed the amount of caffeinated coffee you drink?
yes no → *skip to C2a.* don't know/refused → *skip to C2a.*

C1d. When did you change the amount of caffeinated coffee you drink?
month: ____ day: _____ *[if doesn't remember day ask C1e.]* year: ___
- ❑ doesn't remember/refused

C1e. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?
___1st ____2nd ___ 3rd ____ 4th ___ 5th
- ❑ doesn't remember/refused

C1f. Before you changed, how many cups of caffeinated coffee did you usually drink per day?
_____ # per day ____<1 per day none

C1g. Were those cups usually small like a tea cup, about 4-10 oz; medium like a coffee mug, about 12-14oz; or large like an travel mug or oversized coffee mug, about 16-24oz?
Sm Med Lg Other: specify _____ don't know/refused

Hot caffeinated tea
Now I'm going to ask you about caffeinated tea.

C2a. Currently, how many cups of hot caffeinated tea do you usually drink per day?
____ # per day ____<1 per day → *skip to C2b.* none → *skip to C2b.*
Don't know/refused → *skip to C2b.*

C2asize. Are those cups usually small like a tea cup, about 4-10 oz; medium like a coffee mug, about 12-14oz; or large like an travel mug or oversized coffee mug, about 16-24oz?
Sm Med Lg Other: specify _____ don't know/refused

C2b. In the past four months, have you changed the amount of hot caffeinated tea you drink?
yes no → *skip to C3a.* don't know/refused → *skip to C3a.*

C2c. When did you change the amount of hot caffeinated tea you drink?
month: ____ day: _____ *[if doesn't remember day ask C2d.]* year: ___
- ❑ doesn't remember/refused

C2d. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?

___1st ____2nd___ 3rd ____ 4th___ 5th

❑ doesn't remember/refused

C2e. Before you changed, how many cups of hot caffeinated tea did you usually drink per day?

_____ # per day ____<1 per day none

C2f. Were those cups usually small like a tea cup, about 4-10 oz; medium like a coffee mug, about 12-14oz; or large like an travel mug or oversized coffee mug, about 16-24oz?

Sm Med Lg Other: specify _____ don't know/refused

Iced caffeinated tea

[interviewer notes: both brewed and instant]

C3a. Currently, how many glasses of iced caffeinated tea do you usually drink per day?

____ # per day ____<1 per day → *skip to C3b*. none → *skip to C3b.*

Don't know/refused → *skip to C3b*.

C3asize. Are those glasses usually small like a juice glass, about 4-10oz; medium like a water glass, about 12-20; or large like a giant size drink at the movies/Fast food, about 22-34oz?

Sm Med Lg Other: specify _____ don't know/refused

C3b. In the past four months, have you changed the amount of iced caffeinated tea you drink?

yes no → *skip to C4a*. don't know/refused → *skip to C4a.*

C3c. When did you change the amount of iced caffeinated tea you drink?

month: ____ day: _____ *[if doesn't remember day ask C3d.]* year: ___

❑ doesn't remember/refused

C3d. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?

___1st ____2nd___ 3rd ____ 4th___ 5th

❑ doesn't remember/refused

C3e. Before you changed, how many glasses of caffeinated tea did you usually drink per day?

_____ # per day ____<1 per day none

C3f. Were those glasses usually small like a juice glass, about 4-10oz; medium like a water glass, about 12-20; or large like a giant size drink at the movies/Fast food, about 22-34oz?

Sm Med Lg Other: specify _____ don't know/refused

Decaffeinated drinks
[interviewer notes: for both C4a and C4b. include hot or cold drinks]

C4a. Currently, how many cups of <u>decaffeinated</u> coffee do you usually drink per day?
____ # per day ____<1 per day→ *skip to C4b.* none→ *skip to C4b.*
Don't know/refused→ *skip to C4b.*

C4asize. Are those cups usually small like a tea cup, about 4-10 oz; medium like a coffee mug, about 12-14oz; or large like an travel mug or oversized coffee mug, about 16-24oz?
Sm Med Lg Other: specify _____ don't know/refused

C4b. Currently, how many cups of <u>decaffeinated</u> tea do you usually drink per day?
____ # per day ____<1 per day → *skip to C5a.* none→ *skip to C5a.*
don't know/refused→ *skip to C5a.*

C4bsize. Are those cups usually small like a tea cup, about 4-10 oz; medium like a coffee mug, about 12-14oz; or large like an travel mug or oversized coffee mug, about 16-24oz?
Sm Med Lg Other: specify _____ don't know/refused

Sodas/soft drinks
The next questions are about sodas and soft drinks.

C5a. Currently, do you drink one or more cans, bottles or glasses of soda or soft drinks per day, including tonic water, club soda, soda water, seltzer and caffeinated water?
yes no → *skip to C5g.* don't know/refused → *skip to C5g.*

C5b. What brand or type of soda do you drink most often? Anything else *[Up to 4 brands, those she usually/regularly drinks. If store brand, include name of store and type of soda, ie. cola.]*
_____ _____

C5c. Is (brand) diet? *[for each of up to 4 types/brands she drinks ask C5c.]*
yes no don't know/refused

C5d. Is (brand) caffeine free? *[for each of up to 4 types/brands she drinks ask C5d.]*
Yes→ *if yes for all sodas skip to C5g* No Don't know/refused

[for each type/brand she drinks that <u>is</u> caffeinated or if she doesn't know if it's caffeinated, ask C5e-C5f.]

C5e. How many cans, bottles or glasses of _____ do you usually drink per day?
____ # per day ____<1 per day → *skip to C5g.*

C5f. Are those usually small like a can of soda, about 8-12oz, medium like a small soda bottle about 14-22oz or large like a giant drink at the movies or fast food, about 24-34oz?
sm med lg Other _____ don't know/refused

[Ask C5g-C5j about caffeinated drinks as a whole, not for each individual brand.]

C5g. The next questions are <u>only</u> about <u>caffeinated</u> sodas. In the past 4 months, have you changed the amount of caffeinated soft drinks/sodas you usually drink?
yes no → skip to C6a. don't know/refused → skip to C6a.

C5h. When did you change the amount of caffeinated soft drinks/sodas you drink?
month: ____ day: _____ *[if doesn't remember day ask C5i.]* year: ___
❑ doesn't remember/refused

C5i. Do you remember what week in [month] that was, the first, second, third, fourth, or fifth?
___1st ____2nd___ 3rd ____ 4th___ 5th
❑ doesn't remember/refused

C5j. Before you changed, how many cans, bottles or glasses of caffeinated soda or soft drinks did you usually drink per day?
_____ # per day ____<1 per day → *skip to C6a.* none → *skip to C6a.*

C5k. Were those usually small like a can of soda, about 8-12oz, medium like a small soda bottle about 14-22oz or large like a giant drink at the movies or fast food, about 24-34oz?
sm med lg Other _____ don't know/refused

APPENDIX 4. SAS CODES FOR MANUSCRIPT 3

The following SAS codes were used by the authors to estimate the propensity scores using logistic regression and to model caffeine exposure on SAB with propensity scores using unconstrained continuation odds ratio model. Steps that were used to estimate propensity scores and quintile assignments were adapted from paper by D'Agostino.[17]

Sample data

ID	x	z1	z2	z3	y
303	214.9	1	0	1	0
1122	98.2	1	1	0	0
1210	1540.3	0	1	1	0
3486	785.4	0	1	1	1
7832	0	1	1	0	0
10268	468.9	1	0	0	1

Step 1: Estimating propensity scores using a saturated logistic regression model

```
proc logistic data=final descending;
        model y= x z1 z2 z3….zn;
output out=pscoffpreg p=predict;
run;
```

x= main exposure
y= outcome
z1 …z_n = confounders

Step 2: Creating quintiles based on the estimated propensity scores

```
proc rank groups=5 data=pscoffpreg out=quintile;
        ranks rnks;
var predict;
data qtile; set quintile; quintile=rnks + 1;
run;
proc freq data=quintile; tables rnks; run;
proc freq data=qtile; tables quintile; run;
```

Step 3: Fitting the SAB model

```
proc logistic data=final descending;
class quintile/param=ref;
model abort = ind5 ind6 ind7 ind8 ind9 ind10 ind11 ind12 ind13 ind14
        ind15 ind16 ind17 ind18 ind19 ind20 quintile x z1 z2 z3;
contrast 'OR for x in any given quintile' x 1/estimate=exp;
title 'propensity scores for exposure on SAB'; run;
```

REFERENCES

1. Jarnfelt-Samsioe A. Nausea and vomiting in pregnancy: a review. Obstet Gynecol Surv 1987;42:422-7.

2. Huxley RR. Nausea and vomiting in early pregnancy: its role in placental development. Obstet Gynecol 2000;95:779-82.

3. Jarnfelt-Samsioe A, Samsioe G, Velinder GM. Nausea and vomiting in pregnancy--a contribution to its epidemiology. Gynecol Obstet Invest 1983;16:221-9.

4. Tierson FD, Olsen CL, Hook EB. Nausea and vomiting of pregnancy and association with pregnancy outcome. Am J Obstet Gynecol 1986;155:1017-22.

5. Klebanoff MA, Koslowe PA, Kaslow R, Rhoads GG. Epidemiology of vomiting in early pregnancy. Obstet Gynecol 1985;66:612-6.

6. Weigel MM, Reyes M, Caiza ME, et al. Is the nausea and vomiting of early pregnancy really feto-protective? J Perinat Med 2006;34:115-22.

7. Weigel MM, Weigel RM. Nausea and vomiting of early pregnancy and pregnancy outcome. An epidemiological study. Br J Obstet Gynaecol 1989;96:1304-11.

8. Weigel RM, Weigel MM. Nausea and vomiting of early pregnancy and pregnancy outcome. A meta-analytical review. Br J Obstet Gynaecol 1989;96:1312-8.

9. Kline J, Stein Z, Susser M. Conception to Birth: Epidemiology of Prenatal Development. New York: Oxford University Press, 1989.

10. Kallen B. Epidemiology of Human Reproduction. Boca Raton, FL: CRC Press, 1988.

11. Ness RB, Grisso JA, Hirschinger N, et al. Cocaine and tobacco sse and the risk of spontaneous abortion. N Engl J Med 1999;340:333-339.

12. Kline J, Levin B, Silverman J, et al. Caffeine and spontaneous abortion of known karyotype. Epidemiology 1991;2:409-17.

13. Fenster L, Eskenazi B, Windham GC, Swan SH. Caffeine consumption during pregnancy and spontaneous abortion. Epidemiology 1991;2:168-74.

14. Wen W, Shu XO, Jacobs DR, Jr., Brown JE. The associations of maternal caffeine consumption and nausea with spontaneous abortion. Epidemiology 2001;12:38-42.

15. Mendola P, Buck GM, Vena JE, Zielezny M, Sever LE. Consumption of PCB-contaminated sport fish and risk of spontaneous fetal death. Environ Health Perspect 1995;103:498-502.

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

17. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in Medicine 1998;17:2265-2281.

18. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health 2006;60:578-586.

19. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Statistics in Medicine 2004;23:2937-2960.

20. Sturmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. Journal of Clinical Epidemiology 2006;59:437.e1-437.e24.

21. Signorello LB, McLaughlin JK. Maternal caffeine consumption and spontaneous abortion: a review of the epidemiologic evidence. Epidemiology 2004;15:229-39.

22. Wilcox AJ, Weinberg CR, Wehmann RE, Armstrong EG, Canfield RE, Nisula BC. Measuring early pregnancy loss: laboratory and field methods. Fertil Steril 1985;44:366-74.

23. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41-55.

24. Savitz DA, Chan RL, Herring AH, Howards PP, Hartmann KE. Caffeine and miscarriage risk. Epidemiology 2008;19:55-62.

25. Sayle AE, Wilcox AJ, Weinberg CR, Baird DD. A prospective study of the onset of symptoms of pregnancy. J Clin Epidemiol 2002;55:676-80.

26. Klebanoff MA, Mills JL. Is vomiting during pregnancy teratogenic? Br Med J (Clin Res Ed) 1986;292:724-6.

27. Sherman PW, Flaxman SM. Nausea and vomiting of pregnancy in an evolutionary perspective. Am J Obstet Gynecol 2002;186:S190-7.

28. Mori M, Amino N, Tamaki H, Miyai K, Tanizawa O. Morning sickness and thyroid function in normal pregnancy. Obstet Gynecol 1988;72:355-9.

29. Coad J, Al-Rasasi B, Morgan J. Nutrient insult in early pregnancy. Proc Nutr Soc 2002;61:51-9.

30. Gadsby R, Barnie-Adshead AM, Jagger C. A prospective study of nausea and vomiting during pregnancy. Br J Gen Pract 1993;43:245-8.

31. Zhou Q, O'Brien B, Relyea J. Severity of nausea and vomiting during pregnancy: what does it predict? Birth 1999;26:108-14.

32. Kallen B, Lundberg G, Aberg A. Relationship between vitamin use, smoking, and nausea and vomiting of pregnancy. Acta Obstet Gynecol Scand 2003;82:916-20.

33. Behrman CA, Hediger ML, Scholl TO, Arkangel CM. Nausea and vomiting during teenage pregnancy: effects on birth weight. J Adolesc Health Care 1990;11:418-22.

34. Linseth G, Vari P. Nausea and vomiting in late pregnancy. Health Care Women Int 2005;26:372-86.

35. Gazmararian JA, Petersen R, Jamieson DJ, et al. Hospitalizations during pregnancy among managed care enrollees. Obstet Gynecol 2002;100:94-100.

36. Koren G, Boskovic R, Hard M, Maltepe C, Navioz Y, Einarson A. Motherisk-PUQE (pregnancy-unique quantification of emesis and nausea) scoring system for nausea and vomiting of pregnancy. Am J Obstet Gynecol 2002;186:S228-31.

37. Lacroix R, Eason E, Melzack R. Nausea and vomiting during pregnancy: A prospective study of its frequency, intensity, and patterns of change. Am J Obstet Gynecol 2000;182:931-7.

38. Wilcox AJ, Weinberg CR, O'Connor JF, et al. Incidence of early loss of pregnancy. N Engl J Med 1988;319:189-94.

39. O'Brien B, Zhou Q. Variables related to nausea and vomiting during pregnancy. Birth 1995;22:93-100.

40. Zhang J, Cai WW. Severe vomiting during pregnancy: antenatal correlates and fetal outcomes. Epidemiology 1991;2:454-7.

41. Lagiou P, Tamimi R, Mucci LA, Trichopoulos D, Adami HO, Hsieh CC. Nausea and vomiting in pregnancy in relation to prolactin, estrogens, and progesterone: a prospective study. Obstet Gynecol 2003;101:639-44.

42. Louik C, Hernandez-Diaz S, Werler MM, Mitchell AA. Nausea and vomiting in pregnancy: maternal characteristics and risk factors. Paediatric and Perinatal Epidemiology 2006;20:270-278.

43. Davis M. Nausea and vomiting of pregnancy: an evidence-based review. J Perinat Neonatal Nurs 2004;18:312-28.

44. Neugebauer R, Kline J, O'Connor P, et al. Determinants of depressive symptoms in the early weeks after miscarriage. Am J Public Health 1992;82:1332-9.

45. Neugebauer R, Kline J, Shrout P, et al. Major depressive disorder in the 6 months after miscarriage. JAMA 1997;277:383-8.

46. Frost M, Condon JT. The psychological sequelae of miscarriage: a critical review of the literature. Aust N Z J Psychiatry 1996;30:54-62.

47. Benson CB, Chow JS, Chang-Lee W, Hill JA, 3rd, Doubilet PM. Outcome of pregnancies in women with uterine leiomyomas identified by sonography in the first trimester. J Clin Ultrasound 2001;29:261-4.

48. Eldar-Geva T, Meagher S, Healy DL, MacLachlan V, Breheny S, Wood C. Effect of intramural, subserosal, and submucosal uterine fibroids on the outcome of assisted reproductive technology treatment. Fertil Steril 1998;70:687-91.

49. Lidegaard O. Cervical incompetence and cerclage in Denmark 1980-1990. A register based epidemiological survey. Acta Obstet Gynecol Scand 1994;73:35-8.

50. Margaret Weigel M, Reyes M, Caiza ME, et al. Is the nausea and vomiting of early pregnancy really feto-protective? J Perinat Med 2006;34:115-22.

51. Greenland S, Brumback B. An overview of relations among causal modelling methods. Int. J. Epidemiol. 2002;31:1030-1037.

52. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol 2002;155:176-84.

53. Hernan MA. A definition of causal effect for epidemiological research. J Epidemiol Community Health 2004;58:265-271.

54. Goldstein A, Warren R, Aldridge A, et al. Passage of caffeine into human gonadal and fetal tissue. Biochem Pharmacol 1962;11:166-8.

55. Aldridge A, Bailey J, Neims AH. The disposition of caffeine during and after pregnancy. Semin Perinatol 1981;5:310-4.

56. Brazier JL, Ritter J, Berland M, Khenfer D, Faucon G. Pharmacokinetics of caffeine during and after pregnancy. Dev Pharmacol Ther 1983;6:315-22.

57. Weathersbee PS, Lodge JR. Caffeine: its direct and indirect influence on reproduction. J Reprod Med 1977;19:55-63.

58. Kirkinen P, Jouppila P, Koivula A, Vuori J, Puukka M. The effect of caffeine on placental and fetal blood flow in human pregnancy. Am J Obstet Gynecol 1983;147:939-42.

59. Czeizel AE, Puho E. Association between severe nausea and vomiting in pregnancy and lower rate of preterm births. Paediatr Perinat Epidemiol 2004;18:253-9.

60. Bech BH, Nohr EA, Vaeth M, Henriksen TB, Olsen J. Coffee and fetal death: a cohort study with prospective data. Am J Epidemiol 2005;162:983-90.

61. Matijasevich A, Barros FC, Santos IS, Yemini A. Maternal caffeine consumption and fetal death: a case-control study in Uruguay. Paediatric and Perinatal Epidemiology 2006;20:100-109.

62. Infante-Rivard C. Caffeine intake and small-for-gestational-age birth: modifying effects of xenobiotic-metabolising genes and smoking. Paediatric and Perinatal Epidemiology 2007;21:300-309.

63. Dlugosz L, Belanger K, Hellenbrand K, Holford TR, Leaderer B, Bracken MB. Maternal caffeine consumption and spontaneous abortion: a prospective cohort study. Epidemiology 1996;7:250-5.

64. Fenster L, Hubbard AE, Swan SH, et al. Caffeinated beverages, decaffeinated coffee, and spontaneous abortion. Epidemiology 1997;8:515-23.

65. Cnattingius S, Signorello LB, Anneren G, et al. Caffeine intake and the risk of first-trimester spontaneous abortion. N Engl J Med 2000;343:1839-45.

66. Tolstrup JS, Kjaer SK, Munk C, et al. Does caffeine and alcohol intake before pregnancy predict the occurrence of spontaneous abortion? Hum Reprod 2003;18:2704-10.

67. Infante-Rivard C, Fernandez A, Gauthier R, David M, Rivard GE. Fetal loss associated with caffeine intake before and during pregnancy. JAMA 1993;270:2940-2943.

68. Parazzini F, Chatenoud L, Di Cintio E, et al. Coffee consumption and risk of hospitalized miscarriage before 12 weeks of gestation. Hum. Reprod. 1998;13:2286-2291.

69. Weng X, Odouli R, Li D-K. Maternal caffeine consumption during pregnancy and the risk of miscarriage: a prospective cohort study. American Journal of Obstetrics and Gynecology 2008;198:279.e1-279.e8.

70. Howards PP, Hertz-Picciotto I, Poole C. Conditions for bias from differential left truncation. .Am. J. Epidemiol. 2007;165:444-452.

71. Hertz-Picciotto I, Swan SH, Neutra RR, Samuels SJ. Spontaneous abortions in relation to consumption of tap water: an application of methods from survival analysis to a pregnancy follow-up study. Am J Epidemiol 1989;130:79-93.

72. Waller DK, Spears WD, Gu Y, Cunningham GC. Assessing number-specific error in the recall of onset of last menstrual period. Paediatric and Perinatal Epidemiology 2000;14:263-267.

73. Savitz DA, Terry JW, Jr., Dole N, Thorp JM, Jr., Siega-Riz AM, Herring AH. Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. Am J Obstet Gynecol 2002;187:1660-6.

74. Brown JE, Jacobs DR, Jr., Barosso GM, et al. Recruitment, retention and characteristics of women in a prospective study of preconceptional risks to reproductive outcomes: experience of the Diana Project. Paediatr Perinat Epidemiol 1997;11:345-58.

75. Axelsson G, Rylander R, Molin I. Outcome of pregnancy in relation to irregular and inconvenient work schedules. Br J Ind Med 1989;46:393-8.

76. Kline J, Levin B, Silverman J, et al. Caffeine and spontaneous abortion of known karyotype. Epidemiology 1991;2:409-17.

77. Infante-Rivard C, Fernandez A, Gauthier R, David M, Rivard GE. Fetal loss associated with caffeine intake before and during pregnancy. JAMA 1993;270:2940-3.

78. Fenster L, Windham GC, Swan SH, Eskenazi B. Miscarriage, caffeine, and the epiphenomena of pregnancy: the causal model. Epidemiology 1991;2:313.

79. Mills JL, Holmes LB, Aarons JH, et al. Moderate caffeine use and the risk of spontaneous abortion and intrauterine growth retardation. JAMA 1993;269:593-7.

80. Parazzini F, Bocciolone L, Fedele L, Negri E, La Vecchai C, Acaia B. Risk Factors for Spontaneous Abortion. Int. J. Epidemiol. 1991;20:157-161.

81. George L, Granath F, Johansson AL, Olander B, Cnattingius S. Risks of repeated miscarriage. Paediatr Perinat Epidemiol 2006;20:119-26.

82. Savitz DA, Singer PC, Herring AH, Hartmann KE, Weinberg HS, Makarushka C. Exposure to drinking water disinfection by-products and pregnancy loss. Am J Epidemiol 2006;164:1043-1051.

83. Savitz DA, Singer PC, Hartmann KE, et al. Drinking Water Disinfection By-Products and Pregnancy Outcome.: Awwa Research Foundation, 2005.

84. Bunker ML, McWilliams M. Caffeine content of common beverages. J Am Diet Assoc 1979;74:28-32.

85. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons, Inc., 1987.

86. Schafer JL. Analysis of Incomplete Multivariate Data. New York: Chapman and Hall, 1997.

87. Yuan YC. Multiple Imputation for Missing Data: Concepts and New Development. Rockville: SAS Institute, Inc.

88. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690-1.

89. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159:702-6.

90. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. Am J Epidemiol 2005;162:199-200.

91. Wacholder S. Binomial regression in GLIM: estimating risk ratios and risk differences. Am J Epidemiol 1986;123:174-84.

92. Cole SR, Ananth CV. Regression models for unconstrained, partially or fully constrained continuation odds ratios. Int. J. Epidemiol. 2001;30:1379-1382.

93. Rodriguez G. Multilevel Modeling. Generalized Linear Model: WWS509: Princeton University

94. Singer J, Willett J. Applied Longitudinal Data Analysis: Modeling Change and Event Occurence. New York: Oxford University Press, 2003.

95. Grun JP, Meuris S, De Nayer P, et al. The thyrotrophic role of human chorionic gonadotrophin (hCG) in the early stages of twin (versus single) pregnancies. Clin Endocrinol (Oxf) 1997;46:719-25.

96. Basso O, Olsen J. Sex ratio and twinning in women with hyperemesis or pre-eclampsia. Epidemiology 2001;12:747-9.

97. Louik C, Hernandez-Diaz S, Werler MM, Mitchell AA. Nausea and vomiting in pregnancy: maternal characteristics and risk factors. Paediatr Perinat Epidemiol 2006;20:270-8.

98. Promislow JH, Makarushka CM, Gorman JR, Howards PP, Savitz DA, Hartmann KE. Recruitment for a community-based study of early pregnancy: the Right From The Start study. Paediatr Perinat Epidemiol 2004;18:143-52.

99. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003;157:940-3.

100. Broussard CN, Richter JE. Nausea and vomiting of pregnancy. Gastroenterol Clin North Am. 1998;27:123-151.

101. Forbes LS. The evolutionary biology of spontaneous abortion in humans. Trends in Ecology and Evolution 1997;12:446-450.

102. Ramzy AM, Sattar M, Amin Y, Mansour RT, Serour GI, Aboulghar MA. Uterine myomata and outcome of assisted reproduction. Hum Reprod 1998;13:198-202.

103. Stovall DW, Parrish SB, Van Voorhis BJ, Hahn SJ, Sparks AE, Syrop CH. Uterine leiomyomas reduce the efficacy of assisted reproduction cycles: results of a matched follow-up study. Hum Reprod 1998;13:192-7.

104. Chan R, Savitz D, Herring A. Nausea and vomiting in early pregnancy on caffeine consumption. Society for Pediatric and Perinatal Epidemiologic Conference. Seattle, WA, 2006.

105. Medicine Io. Nutrition during pregnancy. Part I. Weight gain; Part II. Nutrient supplements. Washington, DC: National Academy Press, 1990.

106. Kline J, Stein Z, Susser M. Conception to Birth: Epidemiology of Prenatal Development. New York: Oxford University Press, 1989.

107. Weathersbee PS, Lodge JR, Kirkinen P, et al. Caffeine: its direct and indirect influence on reproduction. J Reprod Med 1977;19:55-63.

108. Savitz DA, Singer PC, Herring AH, Hartmann KE, Weinberg HS, Makarushka C. Exposure to drinking water disinfection by-products and pregnancy loss. Am. J. Epidemiol. 2006;164:1043-1051.

109. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association 1984;79:516-524.

110. Gustafson P. Measurement error and misclassification in statistics and epidemiology: impacts and Bayesian adjustments. Boca Raton: Chapman and Hall/CRC Press, 2004.

111. Rubin DB, Schenker N. Multiple imputation in health-are databases: An overview and some applications. Statistics in Medicine 1991;10:585-598.

112. Schmid CH, Rosner B. A Bayesian approach to logistic regression models having measurement error following a mixture distribution. Stat Med 1993;12:1141-53.