Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Angela Lynn Largen Date

Spatial Variation in Preterm Birth and Risk Factors in Five Counties Around Atlanta, GA

By

Angela Lynn Largen

Master of Public Health

Epidemiology

_____

Michael Kramer, Ph.D

Faculty Thesis Advisor

Spatial Variation in Preterm Birth and Risk Factors in Five Counties Around Atlanta, GA

By

Angela Lynn Largen

Bachelor of Science

University of Alaska, Fairbanks, 2006

Faculty Thesis Advisor: Michael Kramer, Ph.D

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2013

Abstract

Spatial Variation in Preterm Birth and Risk Factors in Five Counties Around Atlanta, GA

By Angela Largen

Background: Preterm birth is a major contributor to infant morbidity and mortality and is a continuing burden in the US despite modern medical care. Many risk factors are implicated in the progression of preterm birth. Health interventions exist for some of these factors, and identifying spatial regions containing high proportions of women with these modifiable risk factors can inform targeted interventions to those regions.

Objective: The aim of this study was to identify areas with distinct spatial variation in preterm birth within the five core counties around Atlanta, and determine the relative contribution of significant modifiable risk factors in these areas. Additionally, this study will ascertain if differences exist between African-American women and white women in spatial patterns of preterm birth and risk factors.

Methods: Birth data from the Georgia Birth File was obtained for 2005-2007. Preterm birth prevalence was smoothed using spatial Empirical Bayes by census tract, and areas with distinct spatial variation of preterm birth were selected. Regions were identified for distinctly low and high preterm birth prevalence among all women, black women only, and white women only. Attributable risks for modifiable factors were calculated to determine their relative contribution to preterm birth by area.

Results: Differences were found in locations of high and low prevalence areas for white and black women. For the area of high preterm birth in all women, attributable risk values were greatest for short interpregnancy interval (4.6%), tobacco use (3.3%), and pregnancy associated hypertension (1.3%). In the high prevalence area for black women, the greatest attributable risk was due to short interpregnancy interval (5.1%), tobacco use (3.2%), and advanced maternal age (2.6%). In the area of high prevalence for white women only, short interpregnancy interval (9.5%) and eclampsia (1.4%) had the greatest attributable risk.

Conclusion: This study highlights the existence of spatial variation of preterm birth within the five counties around Atlanta, and that contributions of modifiable risk factors to preterm birth vary by location and by race. These findings suggest health interventions should be targeted at specific risk factors in defined locations to have the greatest impact on preterm birth prevention.

Spatial Variation in Preterm Birth and Risk Factors in Five Counties Around Atlanta, GA

By

Angela Lynn Largen

Bachelor of Science

University of Alaska, Fairbanks, 2006

Faculty Thesis Advisor: Michael Kramer, Ph.D

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology

2013

Acknowledgements

I would like to thank my thesis advisor, Dr. Michael Kramer, for his advice and timely feedback on this project, and for his suggestions on project design.

Many thanks to my family for their encouragement, and a special thanks to Jay for editing and statistical suggestions.

Table of Contents

1. Introduction

Preterm birth is defined as birth at less than 37 weeks of gestation (1, 2). According to the CDC, overall prevalence in the United States increased from 11.6% in 2000 to 12.7% in 2005, with the greatest increase seen in late preterm birth infants (34 to 36 weeks of gestation) (3). In the state of Georgia, preterm births accounted for 13.4% of all births in 2008, the 12th highest rate in the nation that year (4). The US government continues to recognize preterm birth as a significant national problem, and for the past 30 years has included it in the Healthy People objectives. Healthy People 2020 aims to reduce preterm birth from the 2007 rate of 12.7% to 11.4%, a modest 10% rate reduction (5).

Despite medical advances, prematurity accounts for a significant portion of infant morbidity and mortality, and contributes to health care costs in excess of $26 billion dollars annually in the US (1, 6). Though a leader in healthcare spending, US country indicators of health often lag behind those of other developed countries, and even some developing countries. Infant mortality is one such indicator, and in 2006 the United States was ranked 39th in infant mortality by the World Health Organization, tied with Lithuania, Serbia, Slovakia, and Thailand (7). From 2000 to 2005 the infant death rate remained constant at around 6.8 deaths per 1,000 live births. This lack of decline in infant mortality is thought to be attributed to the increase in preterm births during this period (3). While the preterm birth rate declined slightly in 2006 (2), it remains a significant burden in the United States and is the leading cause of death among infants (8). Currently, 65-68.6% of infants who die are preterm (8, 9), with about 36% of total infant deaths attributed to causes related to preterm birth (9, 10). Preterm infants that survive face a greater risk of developmental disabilities and illness than infants carried to term. These include language and learning disabilities, ADHD, behavioral and social problems, cerebral palsy, mental retardation, vision and hearing problems, asthma, reoccurring infections, and poor growth (1).

Numerous risk factors contributing to preterm birth have been identified, spanning genetics, biology, social interactions, and economic circumstance. However, it is difficult to establish appropriate preventative measures as multiple risks may act in concert and can vary between women and groups of women. Evaluation of large, urban areas that group women together may overlook sub-areas with specific risk profiles that differ from the larger, aggregate area (11). Using geographic information systems to identify patterns in health and location of risk factors has become increasingly prevalent in the literature. The importance of the location of maternal residence on pregnancy outcomes has been implicated in many studies. One group of researchers looked at spatio-temporal patterns in birth weight across term and preterm infants to identify areas with increasing low birth weight and possible causes in those areas (12). Other studies have looked spatial distribution of specific factors thought to influence birth outcomes such as prenatal care (13), and residential racial isolation (14). Through all of these studies, the connection between location and birth outcomes is emphasized.

This project will address the issue of preterm birth in the five core counties around Atlanta, Georgia (Fulton, DeKalb, Gwinnett, Cobb, and Clayton county) by examining the contribution of individual and neighborhood level risk factors to preterm birth in this area. Risk factors previously identified in the literature to be significantly associated with preterm birth will be included. Specifically, population attributable risk percentages (PAR%) of modifiable factors will be calculated to further estimate the magnitude of preterm birth cases associated with each risk factor. Spatial variation in preterm birth across the study area will be assessed to identify areas of high and low preterm birth and to compare differences in attributable risks of significant factors between locations. As suggested by South et al., identification of specific risk factors in defined geographic areas can inform more targeted interventions to reduce preterm birth rates (11).

Research questions

- In the 5 core counties around Atlanta, GA, where are distinct areas of high and low prevalence of preterm birth, and what are the population attributable risk percentages for significant modifiable risk factors?
- Are there differences in preterm prevalence rates and attributable risk profiles when the geographic scale changes from counties to census tracts?
- Do locations of high preterm birth prevalence areas differ among African-American and white women within the 5 core counties in Georgia and do they have different significant risk factors in these high prevalence areas?

2. Literature Review

2.1 Preterm Birth Risk Factor Overview

The obstetric mechanisms that precede early parturition include preterm premature rupture of the membranes (PPROM) surrounding the fetus, spontaneous rupture of intact membranes, or indicated preterm birth (9), however the factors contributing to these causes are numerous and often poorly understood. Several biological and environmental factors have been indicated in this progression, and the literature describing risks to early parturition suggest an interaction of multiple variable types. Determining risk factors leading to prematurity is essential for identifying women who are at an increased risk for early birth. Through early identification of high risk women, preventative measures can be implemented.

2.2 Risk Factors During Pregnancy

Biological changes can occur during pregnancy that can contribute to preterm birth. Maternal weight change has been implicated as a major independent risk factor (15). Insufficient weight gain during pregnancy has been suggested to increase the risk of preterm birth. A systematic review performed by the Research Triangle Institute found a fourfold increase in preterm birth among women of normal pre-pregnancy weight with the lowest pregnancy weight gain (16). Excessive weight gain during pregnancy has not been implicated as a risk factor for early birth, and in some cases a protective effect has been suggested (OR 0.54, CI: 0.52-0.57, among over 300,000 women) (17). However, other maternal and fetal complications can arise from too much weight gained during gestation (18).

Additional biological changes such as gestational diabetes and pregnancy associated hypertension are linked to preterm birth (1). Hypertension during pregnancy is more commonly associated with pre-eclampsia, in which both high blood pressure and protein in the urine occur after 20 weeks of gestation. It is one of the most common complications of pregnancy and is present in up to 26%-29% of pregnancies among women with no prior history of child birth (19).

Pre-eclampsia is often a basis for indicated preterm birth (9), and in one large study, posed the greatest risk for preterm birth yielding an adjusted odds ratio of 5.07 (CI: 4.70-5.36), with many of these women undergoing cesarean section (17).

The presence of infection is widely recognized as a risk for early birth, occurring in around 40% of preterm cases (20). Infections leading to preterm birth can include intrauterine, urinary tract, and systemic infections, as well as bacterial vaginosis (21). Bacterial vaginosis, a shift in the normal flora which leads to a greater presence of anaerobic bacteria, has been associated with a 2-fold increase in preterm birth (22).

Inter-pregnancy interval has also been documented as a factor for preterm birth (9, 23, 24). According to a meta-analysis by Conde-Agudelo et al., short inter-pregnancy intervals (< 18 months between pregnancies) and long intervals (> 59 months between pregnancies) carry an increased risk of preterm birth compared to intervals between 18 and 59 months. The greatest risk was seen in smaller intervals (<18 months, adjusted OR 1.40 (CI: 1.24-1.58)) as compared to longer intervals (>59 months, adjusted OR 1.20 (CI: 1.17-1.24) (24). Multiple studies suggest the greatest risk of preterm birth occurs with intervals of less than 6 months (23, 25). Additionally, in a review paper by Zhu of three studies, the optimal inter-pregnancy interval to reduce the risk of preterm birth, along with low birth weight, and small for gestational age births, was 18-23 months between births (23). Multiple gestation pregnancies are also at high risk for preterm birth, but the biological pathways leading to preterm birth in this group may be different than those of singleton births (26).

Lifestyle choices during gestation can impact the risk of delivering early. Tobacco use during pregnancy can cause a small increase in the likelihood of early parturition due to spontaneous preterm birth, with a relative risk of 1.2 to 1.6 in women who smoke as compared to women who do not smoke during pregnancy (27). However, it is suggested that there is greater risk for indicated rather than spontaneous preterm birth among women who smoke due to potential deleterious health effects to the fetus (28). The amount of alcohol consumed during

pregnancy can increase risk as well. In a recent study from the Danish National Birth Cohort, women who drank at least 4 drinks per week were at a greater risk for delivering early, with the highest risk seen in women who had 7 or more drinks per week. However, no risk was demonstrated in women who consumed less than 4 drinks per week (29). Conversely, engaging in physical activity during pregnancy can decrease the risk of preterm birth (29, 30).

2.3 Maternal Risk Factors Prior to Pregnancy

Many maternal risk factors for early delivery are related to poor overall health status prior to conception. Both low- and high- prepregnancy weight can influence time of delivery. In a recent meta-analysis of 78 studies, an increased likelihood of preterm birth among underweight women was found across all three types of preterm birth (PPROM, spontaneous, and induced), with a relative risk of 1.21 (CI: 1.10-1.28) in 32 studies and 1.29 (CI: 1.10-1.46) in 14 studies (31). The impact of high prepregnancy weight varies across the three types of preterm birth. Hendler et al. reported that obese women had lower rates of spontaneous preterm births, while indicated preterm birth increased with increasing BMI (32). In a review of maternal obesity and pregnancy, Ramachenderan et al. confirms these results across numerous studies, reporting no significant increase in spontaneous preterm birth among obese women despite experiencing an increased rate of induced preterm birth due to medical concerns (18). Regardless of the correlation of increased obesity with decreased spontaneous prematurity, obesity may be a precursor to other risk factors such as hypertension and diabetes. A history of chronic hypertension prior to pregnancy has been shown to increase the risk for preterm birth in general (OR 1.64; CI: 1.15-2.36) (33) and indicated preterm birth specifically (OR 8.1; CI: 6.2-10.6) (34) when compared to healthy women.

Pregestational diabetes has also been implicated in early parturition. Sibai et al. found an increased risk for both indicated (OR 4.8; CI: 3.0-7.5) and spontaneous preterm birth (OR 2.1; CI: 1.4-2.0) among a group of women with diabetes before pregnancy as compared to a group of

healthy controls (34). A later study confirmed this risk with an adjusted odds ratio of 2.54 for overall preterm birth in women with chronic diabetes as compared to healthy women (17). Additionally, among a group of women with pregestational diabetes, frequency of preterm birth increased with increasing severity of diabetes (35).

2.4 Demographic and Socioeconomic Factors

Maternal demographic risk factors include age, education, and race. Both extremes of the maternal age spectrum are well established risks for preterm birth (1), and, in the last 20 years incidence of total births among older women (≥35 years) have been steadily increasing (36). This age cohort likely accounts for a significant portion of indicated preterm births, as advancing age often necessitates early removal of the fetus due to medical complications (36). In addition, increasing age carries a greater likelihood of other risks such as diabetes and hypertension. Maternal education is suggested to be associated with preterm birth; women who do not complete high school carry a greater risk (37).

African-American women in the United States are twice as likely to deliver preterm than the general population, and are three to four times as likely to deliver very preterm as compared to white women (9, 33). There may be a biological component to the disparity between racial groups. For example, a higher rate of bacterial vaginosis, a known predictor of preterm birth, was significantly associated with income levels in black women yet not for white women (38), which may explain some of the disparity between African-American and white preterm birth rates. However, more complex socioeconomic factors are likely involved in producing this difference. The influence of race on birth outcomes is thought to be mediated by an interaction of a number of individual and neighborhood level risk factors (39).

While demographic differences have been shown to contribute to early birth, these and other factors may be influenced by a woman's socio-economic status and other environmental attributes, thus obscuring exact causal pathways. Women from disadvantaged backgrounds are

more likely to deliver preterm (40). Early birth is often more prevalent in disadvantaged groups irrespective of race, levels of socio-economic status show gradients of increased preterm birth with decreasing socio-economic affluence as measured by percent poverty and maternal education (40). Confirming this, in a study examining preterm birth incidence in the UK, infants whose mothers resided in the most deprived areas (as defined by an index of deprivation) were twice as likely to be born very preterm than those in the least deprived areas (41). Despite overall increases in prematurity with poverty, rates of preterm birth among black women in the United States are still at least twice as high across all levels of poverty and education as white women (40). Additionally, as reported by Messer et al., among a Raleigh, NC birth cohort, residing in areas with high rates of violent crime was associated with a moderate increase in preterm birth in non-Hispanic white and black women (42).

Since it has been suggested that prematurity is impacted by both individual and socio-contextual factors, studies modeling only one of these groups may lead to erroneous conclusions about the impact of both of these effects on early parturition (1). Additionally, the racial disparities in preterm birth in the US are not explained by individual level factors alone (43). A study examining the effects of neighborhood level risks and maternal smoking on birth outcomes found that while smoking caused an increase in preterm births across varying levels of neighborhood poverty, the only areas with statistically significant greater risk were neighborhoods that were predominantly African-American, suggesting that more complex social interactions are involved and certain neighborhoods are more greatly affected (44). Living in areas with high rates of crime has been associated with preterm birth outcomes, and African-American and white women tend to live in areas with different in crime rates, though the increase in preterm birth in these areas may be due more to the increase in neighborhood deprivation rather than crime rates and increased stress due to crime (43).

2.5 Spatial Analysis and Preterm Birth

Mapping health data can be a useful tool to convey spatial information. Maps can support the allocation of public health services by providing a persuasive visual representation of areas most at risk for a condition or in need of health services to key stakeholders (45). Further, it is suggested that risk factors that vary with location could contribute to the overall preterm birth patterns observed across spatial areas (46). Spatial analyses have been used among researchers to investigate preterm birth and its etiologies. Examples of applications of GIS towards health questions surrounding preterm birth include maternal distance to specialized pre- and postnatal care (13, 47), neighborhood level racial isolation (14), and maternal exposure to environmental pollutants (48).

Dividing large areas into smaller regions can aid in identifying patterns in health data that are not obvious at the macro level. The choice of spatial scale, however, can have a significant impact on the results, as utilizing different boundaries and areal sizes can impact the statistical inferences drawn from the data, potentially distorting the real magnitude of the effect and making comparisons to other areas difficult (49, 50). Some studies use predefined boundaries such as census tract and blocks, which are useful for discussing health information with stakeholders since they are commonly used and are familiar to many people. However, since these divisions are based on city and town administrative areas and the relative sizes and population densities vary, measurement bias can be introduced when using these boundaries in health studies (51). To get around this some researchers use novel methodologies to define spatial boundaries. For example, Cutchin et al. accounted for the environment of the area, and composition and perceptions of the individuals who reside in it to form neighborhood boundaries as opposed to using politically defined boundaries such as zip codes or census tracts that do not necessarily reflect the health and behavior patterns of the area (52).

In another study, South et al. explored preterm birth risk factors in one county in Ohio using spatial scanning of point data to define high and low areas of preterm birth instead of using

aggregate data in a tract or census group. It was found that some spatially distinct areas with high proportions of preterm birth had different risk factor profiles: one area of high preterm birth prevalence displayed a greater risk of early parturition due to smoking than other areas with similar rates of preterm birth (11). This research group later extended their study to include multiple counties in Ohio. Using spatial analysis they identified areas that would benefit most from specific interventions for preterm birth in urban areas (53).

Interventions exist to prevent preterm births for certain known etiologies. Identifying women that can benefit from these may be difficult since many studies focus on risk factors across large groups of women, potentially ignoring subpopulations with greater exposure to specific risks. Breaking down large regions and determining factors specific to those areas has great utility, as it can allow for targeted interventions in geographically specific areas (11), thus creating a greater public health impact through targeted preventative measures.

3. Methods

3.1 Study Population

The population for this retrospective cohort study was drawn from the Georgia Birth File from the Office of Health Indicators for Planning, Department of Community Health, including birth and maternal individual level risk factor information for Clayton, Cobb, DeKalb, Fulton, and Gwinnett counties observed over the 3 year period of 2005-2007. A total of 165,812 live birth events were reported during this time. Preterm birth in multi-fetus pregnancies is often caused by different mechanisms than in singleton pregnancies (26); therefore 5,808 birth events were excluded due to multiple gestations, resulting in 160,004 singleton births for the analysis. Census tract level risk factor estimates were taken from the American Community Survey provided by the US Census Bureau for the 5 year period of 2005-2009. This study was approved by the Emory Institutional Review Board.

3.2 Variable Selection

Selection and categorization of significant covariates implicated in the progression of preterm birth was informed by the literature. Individual level variables were limited to those collected from the birth certificate. Univariate distributions were examined for all variables. The outcome variable, preterm birth, was dichotomized using week of gestation: < 37 weeks for preterm birth, and ≥ 37 weeks for term birth using birth certificate reported last menstrual period. Prevalence proportions were calculated for the overall, black only, and white only populations, as well as for each census tract, using the number of preterm births in the numerator and total births in that group in the denominator.

Maternal race and ethnicity were combined into a single variable with 4 levels including non-Hispanic white (reference category), non-Hispanic black, Hispanic, and other (Asian, American Indian/ Alaska Native, Native Hawaiian/ Pacific Islander, or multiracial) (11). Maternal age was separated into three groups including <18 years, 18-34 years, and > 34 years (1, 54).

Marital status was categorized as either single or married. The mother's educational attainment was categorized as less than high school, completed high school or GED, and at least some college (37). Pre-pregnancy risk factors, including diabetes, chronic hypertension, and previous preterm or small for gestational age birth, were dichotomized as either present or absent (1, 33, 34). Prior fetal death and prior induced termination were each dichotomized as either at least 1 or none (55, 56). Smoking was categorized as either any amount of smoking or no smoking during pregnancy (27). Alcohol use was divided into two levels, including maternal consumption of less than 4 drinks per week and consumption of 4 or more drinks per week (29). Maternal disease states occurring during pregnancy, including pregnancy-associated hypertension and eclampsia, were dichotomized as either present or absent. Interpregnancy interval was categorized as <18 months, 18-59 months, and >59 months for women with a prior pregnancy, or categorized as primipara if the current pregnancy was their first (24). A deprivation index (DI) was used as the neighborhood level risk factor and was separated into quartiles of women (11, 42, 43). The referent category for the DI was the lowest quartile (least deprived). The deprivation index was previously calculated at the census tract level and is a combined measure of neighborhood poverty, education, employment, and housing (57).

3.3 Selection of Areas for Analysis

Maternal residence was previously matched to census tracts in the 5 counties of interest. Residence census tract and birth information (term or preterm) were imported into OpenGeoDa 1.0.1 software (58). Census tract preterm birth prevalence proportions were smoothed using the spatial Empirical Bayes method with queen contiguity as described by Anselin (59). Smoothing preterm prevalence can decrease the amount of instability associated with rates in a particular area (59) and allows for better determination of contiguous groups of census tracts with distinctly different preterm birth prevalence rates. These smoothed rates were then transferred to ArcGIS 9.3 software (60) where they were mapped and separated into quintiles. Contiguous regions of

census tracts in the highest, lowest, and middle quintiles were selected for analysis. While all census tracts were smoothed over preterm birth prevalence and grouped into one of the five quintiles, only contiguous regions of high, low, and middle prevalence were selected for further analysis; contiguous regions that had the greatest number of overall birth events were chosen for further study to maximize the number of individuals for the analysis. Hotspots of preterm birth were those areas in the highest preterm birth prevalence quintile, while cold spots were areas in the lowest preterm prevalence quintile.

The preterm birth prevalence quintiles were calculated to identify areas of spatial variation of preterm birth in the five core counties of Atlanta, as well as to identify local preterm prevalence areas that were distinctly different from the overall prevalence of the entire area. The method of using the highest and lowest quintiles to identify these areas of hot and cold preterm birth prevalence was adopted from a study by South et al. (11). Other methods of determining hot and cold spots of preterm birth prevalence exist. One such example is the LISA statistic using GeoDa software, which measures spatial autocorrelation through calculation of local Moran's I for spatial units and their significance, yielding an output that identifies the location of clusters (61). While this method would produce hot and cold clusters of preterm birth prevalence through a more rigorous method identifying statistically significant spatial variation, it would be more difficult to identify middle prevalence areas using this statistic.

3.4 Statistical Analysis

All statistical analyses were carried out using SAS version 9.2. Complete case analysis was performed and observations with missing covariates were deleted from the eligible data set. A multivariable logistic regression model was fit using the risk factors mentioned above regressed on the binary birth outcome (preterm or term). An overall model fit, including all races and all areas, was performed first with all available risk factors identified from the literature to verify their significance in this population. Only significant covariates were kept in subsequent

models of sub-areas of high and low prevalence. Factors were considered significant with a p-value of <0.05. After the initial model was fit over the entire study area population, separate models were fit to each of the identified high, middle, and low prevalence sub-regions described above. This allowed for comparison of significant risk factor profiles between the total aggregate area and hot and cold spots of preterm birth prevalence.

To further assess the impact of these risk factors on preterm birth, adjusted attributable risk fractions were calculated using equation 1, and the number of preventable cases were then obtained using equation 2, taken from South et al. (11). By determining the fraction of births that are due to a specific risk while controlling for other potentially confounding risk factors, the magnitude of a particular covariate in a region can be quantified to help inform targeted interventions in that region (11). The number of preventable cases indicates the number of preterm births that could have been prevented by early detection and subsequent intervention for a given risk factor. Risk factors were designated as modifiable if they are able to be attenuated by either a specific medical intervention or through education.

(1) $AR = pp(\frac{RR-1}{RR})$, pp = proportion of cases exposed to risk factor in region, RR = adjusted relative risk

(2) $Preventable\ cases = AR \times (preterm\ births\ in\ region)$

To compare risk factor profiles among black and white women to that of the total population, the analyses outlined above were repeated with non-Hispanic black mothers alone, and non-Hispanic white mothers alone. Deprivation index quartiles were recalculated to reflect each of these populations alone. Prevalence was also recalculated to reflect only these racial populations and to find high and low areas specific to these groups.

Due to the high number of missing covariate values and subsequent deletion of observations in the complete case analysis, a separate analysis using multiple imputation was performed across the entire population (all races), and results compare with the complete case analysis. Data for the population were first imputed using PROC MI. The Markov Chain Monte Carlo (MCMC) method was used to determine imputed values and multiple chains were used (as opposed to a single chain) to reduce correlation among the imputed values (62). Five imputations were performed. It is suggested that few imputations are sufficient in most cases: the relative efficiency of 5 imputations for data with 30% to 50% missingness is approximately 91% to 94% of the theoretical maximum (62). Adjusted odds ratios were calculated by fitting logistic regression models with the imputed datasets. PROC MIANALYZE was then used to combine these five regression results. To determine the proportion of women with each risk factor for use in calculating attributable risk, means were computed from the imputation results for each categorical variable. When estimating proportions for categorical variables from imputed data, means are suggested to be superior to rounding and then computing proportions (63).

4. Results

4.1 Descriptive statistics for the total study population (all race/ethnicities, African-American only, and white only)

From the 160,004 eligible singleton births, 96,543 (60%) had complete information for all covariates; regression models and subsequent analyses were performed using only these complete cases. Prevalence of each missing covariate is shown in Table 1. Distributions of smoothed preterm birth prevalence rates using the spatial Empirical Bayes method are mapped in Figures 1-3 for the entire study area, including smoothed preterm birth rates among all race and ethnic groups in the study region (Figure 1), rates among non-Hispanic black mothers only (Figure 2), and rates among non-Hispanic white mothers only (Figure 3). Overall prevalence percentages of preterm birth for the entire population, black women only, and white women only were 11.8%, 15.8%, and 8.8%, respectively. Smoothed rates by census tract ranged from 5.9% to 25.5% over the entire study population, 4.2% to 26.2% among non-Hispanic black women, and 0% to 41.7% among non-Hispanic white women. Low, middle, and high prevalence quintiles were <9.51%, 10.64% to 12.96%, and >15.87% for the entire study population, <13.17%, 14.57% to 15.76%, and >18.04% for the black population, and <7.32%, 8.43% to 9.45%, and >11.66% for the white population.

Prevalence of risk factors for the three study populations are shown in Table 2. Among all races and ethnicities, maternal age ranged from 10 to 53 years old, with the majority between 18 and 34 (80%). Almost half of the study population had attended at least some college, increasing to about 75% of the population among non-Hispanic white women. Several indicated risk factors were higher among the African American mothers than among the overall population and white women only. Chronic hypertension was about twice as high in black women compared to white women; however, this risk factor was less than 2% of the study population. Diabetes prevalence was about the same in each group (over 2%). Short interpregnancy interval (less than 18 months between pregnancies) occurred in 6% of the total population. Over 50% of women

gave birth for the first time across all three groups. Behavioral risk factors such as smoking and drinking at least 4 alcoholic beverages per day were low among all populations, < 5% and < 0.1%, respectively. Tobacco use, however, was about twice as high among white women as black women. The average neighborhood deprivation index was higher for black women, about twice that of the entire population and about three times that of white mothers only.

4.2 Logistic regression models for overall populations

Results of the overall logistic regression models for each population are listed in Table 3, this includes adjusted odds ratios and confidence intervals of the risk factors in each of the three overall study populations. For all races and ethnicities combined, all factors except alcohol use were found to be significantly associated with preterm birth; therefore alcohol use was not included in subsequent analyses.

Significant risk factors were similar across the total, black, and white study groups, however there was some variation. For both the total and black only populations, maternal age of 35 or older was significant (OR 1.17; CI: 1.10-1.24, and OR 1.23; CI: 1.12-1.34), but was not significant in the overall white population ($p = 0.11$). Education level was significant across all three groups; some college or higher had a protective effect when compared to a high school education alone, while not completing high school was a risk factor for the total and African-American population, but not for white women. Maternal race was not included in the regression models for black and white women separately, but among all women in the study area, non-Hispanic black race was associated with greater odds of preterm birth when compared to white women (OR 1.59; CI: 1.50-1.68).

Pre-pregnancy factors such as prior fetal death and induced termination were not risk factors for the white population, but both were significant risk factors for the total and black populations. An interpregnancy interval of less than 18 months was a risk factor for all three populations. A high interpregnancy interval of greater than 59 months was weakly associated

with preterm birth for the overall population (OR 1.08; CI: 1.01-1.16) and for the white population (OR 1.28; CI: 1.10-1.50), but not for black women. Prior preterm or small for gestational age birth and chronic hypertension were significant across all populations with relatively large effect sizes, ranging from an odds ratio of 2.15 to 2.80 for hypertension and from 3.35 to 3.53 for previous preterm or small birth.

Risk factors during pregnancy indicated some differences between black and white mothers. Tobacco use was not significant for white women, but was a significant risk for African-American women, with an adjusted odds ratio of 1.52 (CI: 1.30-1.77). In addition to chronic hypertension, pregnancy-associated hypertension was a significant risk for all populations, with similar effect sizes ranging from an odds ratio of 2.38 to 2.50. The largest effect size in all three sub-analyses was for eclampsia with an odds ratio of 4.80 (CI: 3.90-5.90) for the entire population, 5.49 (CI: 4.14-7.27) for black women, and 5.00 (CI: 3.36-7.43) for white women. While an interpregnancy interval of <18 months was significant across all populations, a long interval (>59 months) was not significant for African-American women.

The neighborhood level risk factor (deprivation index) was significant for the total population; the highest quartile (most deprived) was significantly associated with preterm birth with an adjusted OR of 1.13 (CI: 1.06-1.21). Among the black population, both the third and fourth highest quartiles were significant with adjusted odds ratios of 1.10 (CI: 1.02-1.20) and 1.23 (CI: 1.30-1.34), respectively. Level of deprivation was not significant for the white population.

4.3 Areas of spatially distinct prevalence rates: Combined race and ethnicity groups

From the smoothed preterm birth rates among all races shown in Figure 1, the three areas highlighted in Figure 4 were chosen to represent areas of high, low, and mid-level prevalence rates. Table 4 displays the adjusted odds ratios of all risk factors for each prevalence sub-area, with significant risk factors outlined in red. All factors except alcohol use were significant for the overall study area. Several of the risk factors significant in the overall region were not significant

in one or more of the high, middle, or low prevalence regions. There was some heterogeneity between significant characteristics ($p < 0.05$) for the spatially distinct sub-areas of low, middle, and high prevalence rates (see Table 4 for all comparisons). In the regression model for the low prevalence area, the addition of all four deprivation index levels caused non-convergence of the model due to an insufficient number of mothers in some of the categories, therefore the third and fourth highest deprivation quartiles were removed to allow a stable model to be generated. In both the middle and high prevalence regions deprivation index was not a statistically significant factor.

Modifiable risk factors that were significant for all three sub-regions included prior preterm or small birth, eclampsia, and a short interpregnancy interval. The effect size for eclampsia was one of the largest for all three groups, ranging from an odds ratio of 4.81 to 6.31, however the confidence intervals for this variable were relatively large indicating a considerable amount of uncertainty in the estimates. Chronic hypertension was not significant for the middle prevalence region, but was significant for both the low and high prevalence regions, with respective odds ratios of 3.24 (CI: 1.74-6.04) and 2.84 (CI: 1.72-4.67). Any tobacco use during pregnancy was significant in both the middle and high regions, but not the low prevalence area. Pregnancy associated hypertension was significant in the low and high regions with relatively large effect sizes (OR 2.73, CI: 1.85-4.03; OR 2.16, CI: 1.51-3.08), but not the middle prevalence region.

The number of preventable preterm births for each potentially modifiable risk factor is shown in Table 5; variables that were significant in the regression models are highlighted in red. While eclampsia had one of the greatest odds ratio effect sizes for all three sub-regions, other variables contributed greater numbers of preventable preterm births. In the high prevalence area, the greatest number of preventable preterm births were seen with short interpregnancy interval, any tobacco use, and pregnancy associated hypertension, with 70 (4.6%), 51 (3.3%), and 28 (1.3%), respectively. The top three contributing factors for the middle prevalence region were short interpregnancy interval (n=13, 3.0%), young maternal age (n=12, 2.7%), and any tobacco

use (n=9, 2.1%). For the low prevalence area, the three main variables were advanced maternal age (n=44, 9.4%), pregnancy associated hypertension (n=27, 5.7%), and previous preterm or small birth (n=21, 4.4%). The top three for the entire five county population were short interpregnancy interval (n=389, 3.4%), pregnancy associated hypertension (n=360, 3.2%), and advanced maternal age (n=261, 2.3%).

4.4 Areas of spatially distinct prevalence rates: Non-Hispanic black mothers only

From the smoothed preterm birth rates for black mothers only in Figure 2, spatially distinct areas are highlighted in Figure 5 (high, low, and middle prevalence rates). The corresponding adjusted odds ratios for each area are listed in Table 6, and the number of preventable preterm births due to modifiable risk factors are listed Table 7. The high prevalence region (> 18.04% preterm births) shared most significant modifiable risk factors with the entire African-American population, though chronic diabetes was significant in the entire black population only and not among the sub-populations. Both low and middle prevalence areas had fewer significant modifiable risk factors. In both the middle and high prevalence regions, deprivation index was not a statistically significant factor, however both prior preterm birth and eclampsia were significant in these areas. Deprivation index, prior preterm or small birth, and eclampsia were not comparable in the low rate area as they were not included in the model due to convergence issues.

Among the three sub-regions, the modifiable risk factors chronic hypertension and pregnancy associated hypertension were significant in all three areas. These two variables had the greatest effect on preterm birth in the low prevalence area with odds ratios of 6.83 (CI: 2.79-16.99) for chronic hypertension and 4.08 (CI: 1.97-8.49) for pregnancy associated hypertension. Eclampsia and previous preterm or small for gestational age birth had the greatest effect sizes in the middle prevalence region (OR 11.86, CI: 3.90-36.04, and OR 6.35, CI: 2.80-14.40). These two variables were also strongly associated with preterm birth in the high prevalence area, with an odds ratio of 6.74 (CI: 3.25-13.97) for eclampsia and 2.26 (CI: 1.29-3.96) for previous preterm

or small birth. Other variables that were significant for the middle and high but not for the low prevalence region include increased maternal age and short interpregnancy interval. Tobacco use was only significant in the high prevalence region.

The low prevalence region had the fewest number of preventable preterm births; for two the significant risk factors in this regions, chronic hypertension, and pregnancy associated hypertension, the number of preventable cases included 9 (6.9%), and 9 (6.5%) preterm births, respectively. Advanced maternal age contributed a relatively large number of preventable preterm births (n=14, 10.8%), but was not a significant variable in the regression model. In contrast to the low prevalence region, the high prevalence region had a large number of preventable preterm births, a total of 225 among significant risk factors. The greatest contributors for the high prevalence region included short interpregnancy interval (n=73, 5.1%), any tobacco use (n=45, 3.2%), and advanced maternal age (n=37, 2.6%).The middle prevalence region had fewer preventable preterm births (n=99), the major contributors included advanced maternal age (n=34, 9.1%), short interpregnancy interval (n=19, 5.0%), and pregnancy associated hypertension (n=15, 4.1%).

While there were far fewer numbers of preventable preterm births in the low and middle prevalence regions than in the area with high prevalence, this is to be expected since the total numbers of preterm births in these groups were much lower in these areas. In the low and middle areas there were 132 and 375 total preterm births, compared to 1,424 in the high prevalence area.

4.5 Areas of spatially distinct prevalence rates: Non-Hispanic white mothers only

From the smoothed preterm birth rates for white mothers only shown in Figure 3, spatially distinct areas are highlighted in Figure 6 (high, low and middle prevalence rates). Since this group had multiple distinct areas of high preterm birth, two high prevalence regions were chosen in order to compare significant risk factors in different locations with the same smoothed

preterm birth prevalence. Adjusted odds ratios for each prevalence area are listed in Table 8 and corresponding population attributable risks are located in Table 9. Due to homogeneity of certain risk factors in the subgroups of white women, some risk factors caused non-convergence of the logistic regression model and were taken out of the final model for that specific area. These are notated in Table 8.

The logistic regression model for the entire study area with non-Hispanic white mothers yielded far fewer significant risk factors compared to all races combined. Potentially modifiable risk factors in this group include diabetes, chronic and pregnancy associated hypertension, a previous preterm or small for gestational age birth, eclampsia, and both short and long interpregnancy interval. The low prevalence region selected had only two significant factors, including eclampsia with an odds ratio of 6.24 (CI: 1.51-25.79), and a long interpregnancy interval with an odds ratio of 2.38 (CI: 1.17-4.83). The middle prevalence region also had only two significant risk factors; these were different from the low region and included previous preterm or small birth (OR 4.47, CI: 2.36-8.46) and pregnancy associated hypertension (OR 2.79, CI: 1.93-4.03). High prevalence areas 1 and 2 each had different significant factors. For area 1, some college or higher had a slight protective effect, while prior induced termination was the only significant risk contributing to preterm birth (OR 4.25, CI: 1.39-12.97). High prevalence area 2 had two significant modifiable risk factors, eclampsia and short interpregnancy interval, however there was a large amount of variance associated with eclampsia (OR 21.42, CI: 1.18-388.67).

The number of potentially preventable preterm births attributable to specific modifiable risk factors ranged totaled 354 preterm births for the entire white population, 11, 38, 0 and 7 for the low, middle, and two high prevalence areas, respectively. For all white women, pregnancy associated hypertension contributed the greatest number, with 107 (4.1%) preventable preterm births. This was followed by prior preterm or small birth with 59 births (2.3%), and a long interpregnancy interval with 57 births (2.2%). The significant factors, long interpregnancy interval and pregnancy associated hypertension, contributed 8 (5.5%) and 3 (1.9%) preventable

births to the low prevalence region. Among all sub-regions, the middle prevalence area resulted in highest number of preventable preterm births, with 26 (6.4%) for pregnancy associated hypertension and 12 (2.8%) for prior preterm or small birth. For high prevalence area 1, there were no significant modifiable risk factors; for area 2, short interpregnancy interval contributed 6 (9.5%) and eclampsia contributed 1 (1.4%) potentially preventable preterm birth. Tobacco use also contributed 6 preventable preterm births in this area, but was not a significant variable in the regression model.

4.6 Multiple imputation results: Combined race and ethnicity groups

Adjusted odds ratios from the multiple imputation analysis, for all races by prevalence area, are shown in Table 10, and attributable risks are shown in Table 11. Odds ratios from the imputation analysis were similar to those in the complete case analysis, and a majority of the confidence intervals had no or only a slight reduction in width (Table 12). There were a few differences in significant variables. In the entire study area, all of the same variables were significant between the two methods, with the exception of the race/ethnicity category 'Other,' which was significant for the imputation analysis but not for the complete case analysis. For the low prevalence area, all of the same variables were significant except for the race/ethnicity African-American category, which was significant for the imputation (OR 1.35, CI: 1.08-1.68) but not for the complete case (OR 1.26, CI: 0.93-1.69). Two variables had different significance in the middle prevalence region: African-American race was significant in the imputed data (OR 1.41, CI: 1.12-1.79) but not in the complete case data (OR 1.31, CI: 0.99-1.75), while prior induced termination was significant for the complete case analysis (OR 1.57, CI: 1.09-2.26) it was not for the imputation data (OR 1.28, CI: 0.92-1.78). In the high prevalence area, advanced maternal age and marital status of single were both significant for the imputed data (OR 1.33, CI: 1.12-1.57 and OR 1.23, CI: 1.08-1.40) yet not for the complete case data (OR 1.22, CI: 0.98-1.52 and OR 1.13, CI: 0.96-1.32).

The two analysis methods produced similar adjusted attributable risk percents, however since a greater number of births were included in the imputation analysis, the number of preventable preterm births was higher for almost all modifiable risk factors. In the entire study area, the potentially modifiable variables that contributed the most to preterm birth included short interpregnancy interval with 634 (3.4%) preventable preterm births, pregnancy associated hypertension with 581 (3.1%) preventable preterm births, and advanced maternal age with 509 (2.7%) preterm births. For the low prevalence area, advanced maternal age (n=65, 7.9%), pregnancy associated hypertension (n=37, 4.5%), and previous preterm or small for gestational age birth (n=30, 3.6%) contributed the most to preventable preterm births. For the high prevalence area, short interpregnancy interval (n=612, 5.0%), tobacco use (n=341, 2.8%), and advanced maternal age (n=276, 2.2%) contributed the most preventable preterm births.

5. Discussion

5.1 Comparing the entire study area to hot spots

This study found areas with significant spatial variation in preterm birth prevalence within five major counties in Georgia, and determined the population attributable risks for potentially modifiable risk factors in these areas. Contiguous groups of census tracts of high preterm birth prevalence, such as those identified by this study, may benefit from increased public health campaigns targeted at specific risk factors in those defined areas.

Among all ethnicities and races, there were almost three times as many preterm births in the high prevalence region than in either the middle and low prevalence areas, contributing about 13% of all preterm births in the entire study area. There was some variation in significant risk factors between the sub-areas, however the greatest difference between the high, middle, and low prevalence regions was the total number of preventable preterm births. The greatest contributor to preventable preterm births was short interpregnancy interval in the high prevalence region with 70 births, short interpregnancy interval in the middle prevalence region with 13 births, and advanced maternal age in the low prevalence region with 44 births. The greater relative numbers of preventable preterm births in the high prevalence region suggests that public health interventions be focused in this area with specific attention to short interpregnancy interval, tobacco use, and pregnancy associated hypertension. While there are high numbers of preventable preterm births among these variables in the entire population, using the high prevalence region to focus on these risk factors could have a large impact on reducing preterm births due to these risks at potentially lower cost.

5.2 Prevalence and risk factors in black women compared to white women

The overall prevalence of preterm birth among all ethnicities in this study was 11.8%; preterm birth was higher for black mothers individually (15.7%) and lower for white mothers individually (8.9%). This is consistent with national preterm birth estimates which indicate higher

rates for black women as compared to white women, with national US prevalence rates of 16.8% and 10.5%, respectively (as of 2011) (64). The physical locations of hotspots and cold spots of preterm birth varied between black and white women. This could be due to variation in relative numbers of black and white women residing in these areas and the life and social circumstances that resulted in their residence there; it could also be due to social and economic factors in these areas that preferentially impact one racial group more than another.

In this study, the high prevalence regions for African-American women yielded significantly more preventable preterm births among modifiable risk factors, suggesting these areas would benefit more from public health interventions targeted at African-American women. Among the highest quintile regions, black women also experienced a greater prevalence of preterm birth (>18.04%) than white women (>11.66%). The total numbers of preterm births for white women in the high prevalence regions was also relatively low (38 and 68 preterm births) compared to the high prevalence region for black women (1,424 preterm births).

Among the significant modifiable risk factors for all races and ethnicities, short interpregnancy interval (SII) (<18 months), pregnancy associated hypertension (PAH), and advanced maternal age contributed the greatest number of potentially preventable preterm births, contributing 389 (3.4%), 360 (3.2%), and 261 (2.3%) preterm births, respectively. When African-American and white mothers were investigated separately, the top three attributable risks were the same for black mothers as the entire population, with short interpregnancy interval, pregnancy associated hypertension, and advanced maternal age, contributing 231 (3.9%), 193 (3.2%), and 156 (2.6%) preterm births, respectively. The similarity in risk factor profiles for the black population alone and the overall population is likely due to the relatively high proportion of African-American women living in the Atlanta area. The three risk factors that contributed the most to preterm births for white mothers also included pregnancy associated hypertension with 107 (4.1%) risk attributable preterm births, however, previous preterm/ small birth and long

interpregnancy interval (>59 months) were two times higher than short interpregnancy interval in this group, with 59 (2.3%) risk attributable births.

For some areas, such as both high prevalence regions for white mothers only, low numbers of preventable preterm births may suggest the cost to benefit ratio of public health campaigns targeted at these specific groups of women may be too high. In general, public health efforts may be better focused on areas that include high risk women, such as those identified by the overall study population, or the high prevalence region for African-American women.

5.3 Multiple imputation vs. complete case analysis

Complete case analysis is often used for data with missing information but may cause bias and decreased precision in the results (65). Multiple imputation provides an alternate method for analyzing data with missing information. In this study, imputation of the birth data among all races led to a slight increase in precision for most covariates (ascertained through decreased width of confidence intervals) likely due to the increased sample size with the imputed data set. It is possible that there may be bias introduced through violations of the assumptions of multiple imputation, however, this cannot be tested without external data.

Part of the goal of this study was to determine the contribution of modifiable risk factors to preterm birth through calculation of population attributable risk. The analysis methods yielded similar attributable risk percentages for all areas, with the most variation between analyses seen in the high prevalence region. However, since the number of birth events eligible for analysis increased at least two fold in every region, the number of preventable preterm births also increased. For example, in the overall study area for all races combined, the percent of preventable preterm births due to pregnancy associated hypertension was similar for the complete case and imputation analyses (3.2% and 3.1%, respectively), however the number of preventable preterm births was 360 and 581. This suggests that while imputation may give more precise odds ratio estimates in this study, when comparing the relative contribution of risk factors to preterm

birth, a complete case analysis may be sufficient to identify important factors to assist in targeting interventions in this area.

5.4 Strengths

One strength of this study was that the sample size for the total population of women was large enough to provide good estimates for all risk factors in the overall model. Additionally, while issues of spatial scale may exist, the use of census tracts as opposed to point data allows for easier communication about areas of concern to interested parties. In addition, by employing a spatial smoothing technique, instabilities in prevalence rates between census tracts were likely reduced.

Finally, doing separate analyses on black and white women allowed for comparisons of the location hotspots of preterm birth and relative burden in these two groups, as well as a comparison of major contributing risk factors in these women.

5.5 Limitations

There are several limitations to this project. First, low numbers of preterm births in some risk factor categories for white women led to their elimination in the regression model. Both high prevalence regions included in this group had risk factors that caused non-convergence of the regression model and could not be included due to low numbers of women in those categories.

Because this project involved comparing the relative contribution of each risk factor to the preterm birth total in a specific area and population, traditional best fit model selection was not performed. This could potentially diminish the magnitude of effect of certain risk factors important in the causal pathway of preterm birth and their overall contribution to the number of preventable preterm births.

Further, the individual level data for the mothers was obtained from birth records, so potential maternal risk factors such as weight, presence of certain infections, and other health

conditions that may impact the time of birth were not included in this study. Additionally, some studies have suggested that there is substantial underreporting of some maternal illnesses and characteristics on birth certificates, potentially reducing their apparent contribution to preterm birth. Included in these are diabetes and hypertension (both pre-pregnancy and gestational) (66), as well as smoking (67), which could explain the lack of significance of smoking among white women in this study.

Lastly, the modifiable areal unit problem (MAUP) is a common dilemma in studies using some aspect of spatial analysis. Since the choice of spatial scale can be varied with the same data, the decision of where boundaries are drawn can impact the statistical results (49). In this investigation, the five counties were divided into census tracts to determine distinct areas of preterm birth prevalence; dividing the area into larger or smaller areas may have impacted the boundaries of these areas of distinct preterm birth prevalence, the prevalence rates themselves, or the magnitude of effect of each risk factor.

References

1. Behrman RE, Butler AS, Institute of Medicine (U.S.). Committee on Understanding Premature Birth and Assuring Healthy Outcomes. *Preterm birth : causes, consequences, and prevention*. Washington, D.C.: National Academies Press; 2007.
2. CDC. Reproductive Health, Preterm Birth. Atlanta: Centers for Disease Control and Prevention; 2011. (http://www.cdc.gov/reproductivehealth/maternalinfanthealth/PretermBirth.htm). (Accessed June 12, 2011 2011).
3. Macdorman MF, Mathews TJ. Recent trends in infant mortality in the United States. *NCHS Data Brief* 2008(9):1-8.
4. Martin JA, Hamilton BE, Sutton PD, et al. Births: final data for 2008. *Natl Vital Stat Rep* 2010;59(1):1, 3-71.
5. USDHHS. Maternal, Infant, and Child Health Objectives. Washington DC; 2011. (http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=26). (Accessed June 29, 2011).
6. Centers for Disease C. Preterm Birth. (http://www.cdc.gov/reproductivehealth/maternalinfanthealth/PretermBirth.htm). (Accessed October 5, 2013).
7. WHO. Core Health Indicators. (http://apps.who.int/whosis/database/core/core_select.cfm). (Accessed June 1, 2011).
8. Callaghan WM, MacDorman MF, Rasmussen SA, et al. The contribution of preterm birth to infant mortality rates in the United States. *Pediatrics* 2006;118(4):1566-73.
9. Goldenberg RL, Culhane JF, Iams JD, et al. Epidemiology and causes of preterm birth. *Lancet* 2008;371(9606):75-84.

10. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2007 period linked birth/infant death data set. *Natl Vital Stat Rep* 2011;59(6):1-30.

11. South AP, Jones DE, Hall ES, et al. Spatial Analysis of Preterm Birth Demonstrates Opportunities for Targeted Intervention. *Matern Child Health J* 2011.

12. English PB, Kharrazi M, Davies S, et al. Changes in the spatial pattern of low birth weight in a southern California county: the role of individual and neighborhood level factors. *Soc Sci Med* 2003;56(10):2073-88.

13. Charreire H, Combier E. Poor prenatal care in an urban area: a geographic analysis. *Health & place* 2009;15(2):412-9.

14. Anthopolos R, James SA, Gelfand AE, et al. A spatial measure of neighborhood level racial isolation applied to low birthweight, preterm birth, and birthweight in North Carolina. *Spatial and spatio-temporal epidemiology* 2011;2(4):235-46.

15. Ray JG, Vermeulen MJ, Shapiro JL, et al. Maternal and neonatal outcomes in pregestational and gestational diabetes mellitus, and the influence of maternal obesity and weight gain: the DEPOSIT study. Diabetes Endocrine Pregnancy Outcome Study in Toronto. *QJM* 2001;94(7):347-56.

16. Viswanathan M, Siega-Riz AM, Moos MK, et al. Outcomes of maternal weight gain. *Evid Rep Technol Assess (Full Rep)* 2008(168):1-223.

17. Rosenberg TJ, Garbers S, Lipkind H, et al. Maternal obesity and diabetes as risk factors for adverse pregnancy outcomes: differences among 4 racial/ethnic groups. *Am J Public Health* 2005;95(9):1545-51.

18. Ramachenderan J, Bradford J, McLean M. Maternal obesity and pregnancy complications: a review. *Aust N Z J Obstet Gynaecol* 2008;48(3):228-35.

19. Sibai BM. Management of late preterm and early-term pregnancies complicated by mild gestational hypertension/pre-eclampsia. *Semin Perinatol* 2011;35(5):292-6.

20. Averbuch B, Mazor M, Shoham-Vardi I, et al. Intra-uterine infection in women with preterm premature rupture of membranes: maternal and neonatal characteristics. *Eur J Obstet Gynecol Reprod Biol* 1995;62(1):25-9.

21. Pararas MV, Skevaki CL, Kafetzis DA. Preterm birth due to maternal infection: Causative pathogens and modes of prevention. *Eur J Clin Microbiol Infect Dis* 2006;25(9):562-9.

22. Guaschino S, De Seta F, Piccoli M, et al. Aetiology of preterm labour: bacterial vaginosis. *BJOG* 2006;113 Suppl 3:46-51.

23. Zhu BP. Effect of interpregnancy interval on birth outcomes: findings from three recent US studies. *Int J Gynaecol Obstet* 2005;89 Suppl 1:S25-33.

24. Conde-Agudelo A, Rosas-Bermudez A, Kafury-Goeta AC. Birth spacing and risk of adverse perinatal outcomes: a meta-analysis. *JAMA* 2006;295(15):1809-23.

25. de Weger FJ, Hukkelhoven CW, Serroyen J, et al. Advanced maternal age, short interpregnancy interval, and perinatal outcome. *Am J Obstet Gynecol* 2011;204(5):421 e1-9.

26. Martin JA, Osterman MJ, Sutton PD. Are preterm births on the decline in the United States? Recent data from the National Vital Statistics System. *NCHS Data Brief* 2010(39):1-8.

27. Cnattingius S. The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. *Nicotine Tob Res* 2004;6 Suppl 2:S125-40.

28. Goldenberg RL, Culhane JF. Prepregnancy health status and the risk of preterm delivery. *Arch Pediatr Adolesc Med* 2005;159(1):89-90.

29. Andersen AM, Olsen J. The Danish National Birth Cohort: selected scientific contributions within perinatal epidemiology and future perspectives. *Scand J Public Health* 2011;39(7 Suppl):115-20.

30. Jukic AM, Evenson KR, Daniels JL, et al. A Prospective Study of the Association Between Vigorous Physical Activity During Pregnancy and Length of Gestation and Birthweight. *Matern Child Health J* 2011.

31. Han Z, Mulla S, Beyene J, et al. Maternal underweight and the risk of preterm birth and low birth weight: a systematic review and meta-analyses. *Int J Epidemiol* 2011;40(1):65-101.

32. Hendler I, Goldenberg RL, Mercer BM, et al. The Preterm Prediction Study: association between maternal body mass index and spontaneous and indicated preterm birth. *Am J Obstet Gynecol* 2005;192(3):882-6.

33. Haas JS, Fuentes-Afflick E, Stewart AL, et al. Prepregnancy health status and the risk of preterm delivery. *Arch Pediatr Adolesc Med* 2005;159(1):58-63.

34. Sibai BM, Caritis SN, Hauth JC, et al. Preterm delivery in women with pregestational diabetes mellitus or chronic hypertension relative to women with uncomplicated pregnancies. The National institute of Child health and Human Development Maternal-Fetal Medicine Units Network. *Am J Obstet Gynecol* 2000;183(6):1520-4.

35. Sibai BM, Caritis S, Hauth J, et al. Risks of preeclampsia and adverse neonatal outcomes among women with pregestational diabetes mellitus. National Institute of Child Health and Human Development Network of Maternal-Fetal Medicine Units. *Am J Obstet Gynecol* 2000;182(2):364-9.

36. Martin JA, Hamilton BE, Sutton PD, et al. Births: final data for 2007. *Natl Vital Stat Rep* 2010;58(24):1-85.

37. Luo ZC, Wilkins R, Kramer MS. Effect of neighbourhood income and maternal education on birth outcomes: a population-based study. *CMAJ* 2006;174(10):1415-20.

38. Paul K, Boutain D, Manhart L, et al. Racial disparity in bacterial vaginosis: the role of socioeconomic status, psychosocial stress, and neighborhood characteristics, and possible implications for preterm birth. *Soc Sci Med* 2008;67(5):824-33.

39. O'Campo P, Burke JG, Culhane J, et al. Neighborhood deprivation and preterm birth among non-Hispanic Black and White women in eight geographic areas in the United States. *Am J Epidemiol* 2008;167(2):155-63.

40. Kramer MS, Seguin L, Lydon J, et al. Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly? *Paediatr Perinat Epidemiol* 2000;14(3):194-210.

41. Smith LK, Draper ES, Manktelow BN, et al. Socioeconomic inequalities in very preterm birth rates. *Arch Dis Child Fetal Neonatal Ed* 2007;92(1):F11-4.

42. Messer LC, Kaufman JS, Dole N, et al. Violent crime exposure classification and adverse birth outcomes: a geographically-defined cohort study. *Int J Health Geogr* 2006;5:22.

43. Messer LC, Kaufman JS, Dole N, et al. Neighborhood crime, deprivation, and preterm birth. *Ann Epidemiol* 2006;16(6):455-62.

44. Nkansah-Amankra S. Neighborhood contextual factors, maternal smoking, and birth outcomes: multilevel analysis of the South Carolina PRAMS survey, 2000-2003. *J Womens Health (Larchmt)* 2010;19(8):1543-52.

45. Roberts EM, English PB, Wong M, et al. Spatially continuous local rate modeling for communication in public health: a practical approach. *Journal of public health management and practice : JPHMP* 2008;14(6):562-8.

46. Kramer MR, Williamson R. Multivariate Bayesian spatial model of preterm birth and cardiovascular disease among Georgia women: Evidence for life course social determinants of health. *Spatial and spatio-temporal epidemiology* 2013;6:25-35.

47. Pilkington H, Blondel B, Papiernik E, et al. Distribution of maternity units and spatial access to specialised care for women delivering before 32 weeks of gestation in Europe. *Health & place* 2010;16(3):531-8.

48. Leem JH, Kaplan BM, Shim YK, et al. Exposures to air pollutants during pregnancy and preterm delivery. *Environmental health perspectives* 2006;114(6):905-10.

49. Flowerdew R, Manley DJ, Sabel CE. Neighbourhood effects on health: does it matter where you draw the boundaries? *Soc Sci Med* 2008;66(6):1241-55.

50. Stafford M, Duke-Williams O, Shelton N. Small area inequalities in health: are we underestimating them? *Soc Sci Med* 2008;67(6):891-9.

51. Kramer MR, Cooper HL, Drews-Botsch CD, et al. Metropolitan isolation segregation and Black-White disparities in very preterm birth: a test of mediating pathways and variance explained. *Soc Sci Med* 2010;71(12):2108-16.

52. Cutchin MP, Eschbach K, Mair CA, et al. The socio-spatial neighborhood estimation method: an approach to operationalizing the neighborhood concept. *Health & place* 2011;17(5):1113-21.

53. Hall ES, South AP, Jones DE, et al. Spatial analysis in support of community health intervention strategies. *AMIA Annual Symposium proceedings / AMIA Symposium AMIA Symposium* 2012;2012:311-20.

54. Mason SM, Kaufman JS, Daniels JL, et al. Black preterm birth risk in nonblack neighborhoods: effects of Hispanic, Asian, and non-Hispanic white ethnic densities. *Ann Epidemiol* 2011;21(8):631-8.

55. Shah PS, Zao J. Induced termination of pregnancy and low birthweight and preterm birth: a systematic review and meta-analyses. *BJOG* 2009;116(11):1425-42.

56. Edlow AG, Srinivas SK, Elovitz MA. Second-trimester loss and subsequent pregnancy outcomes: What is the real risk? *Am J Obstet Gynecol* 2007;197(6):581 e1-6.

57. Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *Journal of urban health : bulletin of the New York Academy of Medicine* 2006;83(6):1041-62.

58. Anselin L. OpenGeoDa. Spatial Analysis Laboratory (SAL), Department of Geography, University of Illinois, Urbana-Champaign.

59. Anselin L. GeoDa 0.9 User's Guide. Urbana-Champaign, IL: Spatial Analysis Laboratory (SAL), Department of Agricultural and Consumer Economics, University of Illinois, 2003.

60. ESRI. ArcMAP 9.3. Redlands, CA, 1999-2008.

61. Waller LA, Gotway CA. *Applied Spatial Statistics for Public Health Data*. New Jersey: John Wiley & Sons, Inc; 2004.

62. Inc. SI. SAS/STAT® 9.3 User's Guide: The MI Procedure. Cary, NC: SAS Institute Inc., 2011.

63. Allison PD. Imputation of Categorical Variables with PROC MI, Paper 113-30. Cary, NC: SAS Institute Inc: SAS Institute Inc.

64. Martin JA, Hamilton BE, Ventura SJ, et al. Births: Final Data for 2011. *Natl Vital Stat Rep* 2013;62(1).

65. Klebanoff MA, Cole SR. Use of multiple imputation in the epidemiologic literature. *Am J Epidemiol* 2008;168(4):355-7.

66. Lydon-Rochelle MT, Holt VL, Cardenas V, et al. The reporting of pre-existing maternal medical conditions and complications of pregnancy on birth certificates and in hospital discharge data. *Am J Obstet Gynecol* 2005;193(1):125-34.

67. Land TG, Landau AS, Manning SE, et al. Who underreports smoking on birth records: a Monte Carlo predictive model with validation. *PLoS One* 2012;7(4):e34853.

Tables

Table 1. Prevalence of missing covariates

	All races N = 160,004		Non-Hispanic black N = 60,906		Non-Hispanic white N = 47,345	
	N	%	N	%	N	%
Race/ ethnicity	2,968	1.9	-	-	-	-
Education	9,428	5.9	2,393	3.9	2,898	6.1
Marital status	268	0.2	103	0.2	88	0.2
Diabetes	54,427	34.0	20,725	34.0	15,715	33.2
Hypertension (chronic)	54,427	34.0	20,725	34.0	15,715	33.2
Previous preterm/ small birth	54,427	34.0	20,725	34.0	15,715	33.2
Prior fetal death	22,453	14.0	8,792	14.4	5,604	11.8
Prior induced termination	22,529	14.1	8,816	14.5	5,625	11.9
Alcohol use	21,121	13.2	8,195	13.5	5,156	10.9
Tobacco use	20,495	12.8	7,919	13.0	4,874	10.3
Hypertension (pregnancy)	54,427	34.0	20,725	34.0	15,715	33.2
Interpregnancy interval	66	0.0	25	0.0	13	0.0
Eclampsia	54,427	34.0	20,725	34.0	15,715	33.2

Table 2. Prevalence of risk factors for preterm birth among study populations

	All races N = 96,543		Non-Hispanic black N = 38,049		Non-Hispanic white N = 29,232	
	N	%	N	%	N	%
Preterm birth	11,373	11.8	5,971	15.7	2,596	8.9
Maternal demographic factors						
Age						
< 18	3,243	3.4	1,829	4.8	366	1.3
18-34	78,104	80.9	31,372	82.5	21,747	74.4
≥ 35	15,196	15.7	4,848	12.7	7,119	24.4
Education						
< High school	22,081	22.9	6,184	16.3	2,219	7.6
High school/ GED	27,443	28.4	14,497	38.1	4,820	16.5
Some college or higher	47,019	48.7	17,368	45.7	22,193	75.9
Race/ ethnicity						
Non-Hispanic black	38,049	39.4	-	-	-	-
Non-Hispanic white	29,232	30.3	-	-	-	-
Hispanic	22,709	23.5	-	-	-	-
Other	6,553	6.8	-	-	-	-
Single	40,892	42.4	23,969	63.0	4,985	17.1
Pre-pregnancy risk factors						
Diabetes	2,286	2.4	882	2.3	642	2.2
Hypertension (chronic)	837	0.9	543	1.4	210	0.7
Previous preterm/small birth	833	0.9	359	0.9	327	1.1
Prior fetal death	15,314	15.9	6,823	17.9	5,210	17.8
Prior induced termination	9,856	10.2	6,344	16.7	2,168	7.4
During pregnancy risk factors						
Alcohol use						
≥ 4 drinks per week	36	0.0	14	0.0	16	0.1
Any tobacco use	2,514	2.6	992	2.6	1,403	4.8
Hypertension (pregnancy)	2,312	2.4	1,056	2.8	933	3.2
Eclampsia	401	0.4	213	0.6	113	0.4
Interpregnancy interval						
< 18 months	5,843	6.1	2,784	7.3	1,306	4.5
18-59 months	28,803	29.8	9,942	26.1	10,014	34.3
> 59 months	11,789	12.2	5,184	13.6	2,446	8.4
Primipara	50,108	51.9	20,139	52.9	15,466	52.9
Neighborhood level risk factors[a]						
Deprivation Index quartile						
<25% (least deprived)	23,432	24.3	9,337	24.5	7,116	24.3
25-49.9%	23,780	24.6	9,448	24.8	7,782	26.6
50-74.9%	24,307	25.2	9,341	24.6	7,118	24.4
≥ 75%	25,024	25.9	9,923	26.1	7,216	24.7

a – Quartiles of individuals relative to all 5 counties, recalculated for blacks only and whites only

Table 3. Adjusted odds ratios for preterm birth in the overall study populations

	All races N=96,543		Black N=38,049		White N=29,232	
	aOR	95% CI	aOR	95% CI	aOR	95% CI
Maternal demographic factors						
Age						
< 18	1.18	(1.06, 1.31)	1.01	(0.87, 1.17)	0.97	(0.69, 1.37)
18-34	Ref.		Ref.		Ref.	
≥ 35	1.17	(1.10, 1.24)	1.23	(1.12, 1.34)	1.09	(0.98, 1.20)
Education						
< High school	1.11	(1.04, 1.17)	1.19	(1.09, 1.30)	1.13	(0.95, 1.34)
Completed high school/ GED	Ref.		Ref.		Ref.	
Some college or higher	0.84	(0.80, 0.89)	0.88	(0.82, 0.94)	0.84	(0.75, 0.94)
Race/ ethnicity						
Non-Hispanic black	1.59	(1.50, 1.68)	-	-	-	-
Non-Hispanic white	Ref.		-	-	-	-
Hispanic	0.97	(0.90, 1.04)	-	-	-	-
Other	1.01	(0.91, 1.11)	-	-	-	-
Single	1.20	(1.14, 1.26)	1.20	(1.12, 1.28)	1.33	(1.18, 1.50)
Pre-pregnancy risk factors						
Diabetes	1.35	(1.20, 1.51)	1.20	(1.01, 1.43)	1.37	(1.08, 1.74)
Hypertension (chronic)	2.71	(2.31, 3.16)	2.80	(2.32, 3.38)	2.15	(1.50, 3.08)
Previous preterm/small birth	3.53	(3.03, 4.11)	3.35	(2.69, 4.17)	3.50	(2.69, 4.55)
Prior fetal death	1.24	(1.18, 1.31)	1.38	(1.29, 1.48)	1.07	(0.96, 1.19)
Prior induced termination	1.13	(1.06, 1.20)	1.09	(1.01, 1.17)	1.11	(0.95, 1.28)
During pregnancy risk factors						
Alcohol use						
≥ 4 drinks per week	1.47	(0.62, 3.46)	-	-	-	-
Any tobacco use	1.39	(1.25, 1.55)	1.52	(1.30, 1.77)	1.10	(0.92, 1.31)
Hypertension (pregnancy)	2.50	(2.26, 2.76)	2.39	(2.08, 2.75)	2.38	(2.00, 2.84)
Eclampsia	4.80	(3.90, 5.90)	5.49	(4.14, 7.27)	5.00	(3.36, 7.43)
Interpregnancy interval						
< 18 months	1.62	(1.49, 1.75)	1.61	(1.44, 1.79)	1.50	(1.25, 1.81)
18-59 months	Ref.		Ref.		Ref.	
> 59 months	1.08	(1.01, 1.16)	1.04	(0.95, 1.15)	1.28	(1.10, 1.50)
Primipara	1.09	(1.04, 1.15)	1.09	(1.01, 1.17)	1.20	(1.09, 1.32)
Neighborhood level risk factor[a]						
Deprivation Index quartile						
<25% (least deprived)	Ref.		Ref.		Ref.	
25-49.9%	1.04	(0.97, 1.10)	1.07	(0.98, 1.16)	1.04	(0.93, 1.17)
50-74.9%	1.03	(0.96, 1.10)	1.10	(1.02, 1.20)	1.08	(0.96, 1.22)
≥ 75%	1.13	(1.06, 1.21)	1.23	(1.13, 1.34)	0.97	(0.86, 1.10)

Significant factors ($p < 0.05$) highlighted in red

a – Quartiles of individuals relative to all 5 counties, recalculated for blacks only and whites only

Table 4. Adjusted odds ratios for preterm birth for all races by prevalence area

	Entire region N=96,543		Low prevalence area N=5,871		Middle prevalence area N=3,497		High prevalence area N=8,075	
	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI
Maternal demographic factors								
Age								
<18	1.18	(1.06, 1.31)	0.61	(0.18, 2.10)	2.53	(1.41, 4.53)	1.07	(0.85, 1.36)
18-34	Ref.		Ref.		Ref.		Ref.	
≥35	1.17	(1.10, 1.24)	1.38	(1.11, 1.72)	1.09	(0.82, 1.44)	1.22	(0.98, 1.52)
Education								
<High school	1.11	(1.04, 1.17)	0.88	(0.55, 1.42)	0.75	(0.53, 1.07)	1.16	(1.01, 1.34)
Completed high school/ GED	Ref.		Ref.		Ref.		Ref.	
Some college or higher	0.84	(0.80, 0.89)	0.82	(0.61, 1.11)	1.06	(0.81, 1.39)	0.90	(0.77, 1.05)
Race/ ethnicity								
Non-Hispanic black	1.59	(1.50, 1.68)	1.26	(0.93, 1.69)	1.31	(0.99, 1.75)	1.95	(1.31, 2.92)
Non-Hispanic white	Ref.		Ref.		Ref.		Ref.	
Hispanic	0.97	(0.90, 1.04)	0.97	(0.69, 1.38)	1.23	(0.89, 1.70)	1.29	(0.82, 2.03)
Other	1.01	(0.91, 1.11)	0.76	(0.57, 1.03)	1.19	(0.77, 1.82)	1.01	(0.44, 2.34)
Single	1.20	(1.14, 1.26)	1.63	(1.22, 2.18)	1.28	(1.00, 1.64)	1.13	(0.96, 1.32)
Pre-pregnancy risk factors								
Diabetes	1.35	(1.20, 1.51)	2.00	(1.33, 3.00)	1.05	(0.49, 2.25)	1.40	(0.91, 2.16)
Hypertension (chronic)	2.71	(2.31, 3.16)	3.24	(1.74, 6.04)	1.42	(0.45, 4.42)	2.84	(1.72, 4.67)
Previous preterm/small birt	3.53	(3.03, 4.11)	4.97	(3.04, 8.10)	2.89	(1.08, 7.74)	2.22	(1.41, 3.49)
Prior fetal death	1.24	(1.18, 1.31)	1.09	(0.85, 1.39)	1.12	(0.85, 1.48)	1.22	(1.06, 1.41)
Prior induced termination	1.13	(1.06, 1.20)	0.92	(0.63, 1.35)	1.57	(1.09, 2.26)	1.08	(0.92, 1.26)
During pregnancy risk factors								
Any tobacco use	1.39	(1.25, 1.55)	0.56	(0.22, 1.42)	1.69	(1.03, 2.77)	1.73	(1.37, 2.18)
Hypertension (pregnancy)	2.50	(2.26, 2.76)	2.73	(1.85, 4.03)	1.46	(0.74, 2.86)	2.16	(1.51, 3.08)
Eclampsia	4.80	(3.90, 5.90)	4.81	(2.26, 10.24)	5.52	(1.66, 18.37)	6.31	(3.23, 12.35)
Interpregnancy interval								
<18 months	1.62	(1.49, 1.75)	1.71	(1.09, 2.71)	1.57	(1.04, 2.35)	1.55	(1.27, 1.89)
18-59 months	Ref.		Ref.		Ref.		Ref.	
>59 months	1.08	(1.01, 1.16)	1.05	(0.73, 1.49)	0.92	(0.65, 1.29)	1.16	(0.95, 1.40)
Primigravida	1.09	(1.04, 1.15)	1.64	(1.30, 2.06)	1.05	(0.83, 1.33)	0.95	(0.82, 1.09)
Neighborhood level risk factors [a]								
Deprivation Index quartile								
<25% (least deprived)	Ref.		Ref.		Ref.		Ref.	
25-49.9%	1.04	(0.97, 1.10)	1.29	(0.78, 2.12)	0.92	(0.72, 1.18)	0.99	(0.33, 2.98)
50-74.9%	1.03	(0.96, 1.10)	-	-	0.92	(0.59, 1.43)	0.95	(0.32, 2.86)
≥75%	1.13	(1.06, 1.21)	-	-	0.79	(0.56, 1.11)	1.12	(0.38, 3.31)

Significant factors ($p < 0.05$) highlighted in red

a – Quartiles of individuals relative to all 5 counties

Table 5. Risk attributable preterm births for potentially modifiable factors among all races by prevalence area

	Entire population N = 11,373		Low prevalence area N=470		Middle prevalence area N=450		High prevalence area N=1,512	
	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)
Maternal age								
< 18	80	0.7	2	0.4	12	2.7	9	0.6
≥ 35	261	2.3	44	9.4	6	1.4	23	1.5
Diabetes	99	0.9	16	3.5	0	0.1	9	0.6
Hypertension (chronic)	158	1.4	11	2.4	1	0.3	18	1.2
Previous preterm/ small birth	200	1.8	21	4.4	4	0.9	17	1.1
Any tobacco use	123	1.1	4	0.8	9	2.1	51	3.3
Hypertension (pregnancy associated)	360	3.2	27	5.7	3	0.8	28	1.8
Eclampsia	149	1.3	11	2.4	4	0.9	19	1.3
Interpregnancy interval								
< 18 months	389	3.4	11	2.3	13	3.0	70	4.6
> 59 months	110	1.0	2	0.4	5	1.2	28	1.8

Significant risk factors ($p < 0.05$) highlighted in red

Table 6. Adjusted odds ratios for preterm birth for non-Hispanic black women by prevalence area

	Entire region N=38,049		Low prevalence area N=1,153		Middle prevalence area N=2,717		High prevalence area N=6,867	
	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI
Maternal demographic factors								
Age								
<18	1.01	(0.87, 1.17)	1.61	(0.53, 4.88)	0.69	(0.35, 1.36)	1.12	(0.88, 1.43)
18-34	Ref.		Ref.		Ref.		Ref.	
≥35	1.23	(1.12, 1.34)	1.55	(0.99, 2.42)	1.71	(1.26, 2.31)	1.41	(1.12, 1.77)
Education								
<High school	1.19	(1.09, 1.30)	1.34	(0.58, 3.11)	1.17	(0.74, 1.85)	1.16	(1.00, 1.34)
Completed high school/ GED	Ref.		Ref.		Ref.		Ref.	
Some college or higher	0.88	(0.82, 0.94)	1.01	(0.62, 1.64)	0.87	(0.67, 1.13)	0.86	(0.74, 1.02)
Single	1.20	(1.12, 1.28)	1.26	(0.81, 1.96)	1.11	(0.86, 1.43)	1.09	(0.91, 1.31)
Pre-pregnancy risk factors								
Diabetes	1.20	(1.01, 1.43)	1.10	(0.49, 2.51)	0.98	(0.53, 1.80)	1.33	(0.85, 2.09)
Hypertension (chronic)	2.80	(2.32, 3.38)	6.83	(2.75, 16.99)	3.19	(1.65, 6.15)	2.25	(1.35, 3.76)
Previous preterm/small birth	3.35	(2.69, 4.17)	-		6.35	(2.80, 14.40)	2.26	(1.29, 3.96)
Prior fetal death	1.38	(1.29, 1.48)	1.60	(1.04, 2.48)	1.22	(0.93, 1.60)	1.29	(1.11, 1.49)
Prior induced termination	1.09	(1.01, 1.17)	0.89	(0.44, 1.83)	1.24	(0.92, 1.68)	1.19	(1.01, 1.39)
During pregnancy risk factors								
Any tobacco use	1.52	(1.30, 1.77)	0.77	(0.17, 3.42)	1.05	(0.30, 3.65)	1.58	(1.25, 1.99)
Hypertension (pregnancy)	2.39	(2.08, 2.75)	4.08	(1.97, 8.49)	2.42	(1.47, 3.98)	2.04	(1.42, 2.94)
Eclampsia	5.49	(4.14, 7.27)			11.86	(3.90, 36.04)	6.74	(3.25, 13.97)
Interpregnancy interval								
<18 months	1.61	(1.44, 1.79)	1.47	(0.60, 3.57)	1.98	(1.28, 3.07)	1.58	(1.29, 1.94)
18-59 months	Ref.		Ref.		Ref.		Ref.	
>59 months	1.04	(0.95, 1.15)	1.04	(0.54, 1.99)	0.96	(0.66, 1.39)	1.18	(0.96, 1.44)
Primigravida	1.1	(1.01, 1.17)	1.22	(0.76, 1.95)	1.48	(1.11, 1.98)	0.93	(0.80, 1.08)
Neighborhood level risk factors [a]								
Deprivation Index quartile								
<25% (least deprived)	Ref.		Ref.		Ref.		Ref.	
25-49.9%	1.07	(0.98, 1.16)	0.90	(0.46, 1.75)	0.88	(0.65, 1.19)	0.86	(0.59, 1.27)
50-74.9%	1.10	(1.02, 1.20)	-		1.03	(0.76, 1.39)	0.89	(0.63, 1.25)
≥75%	1.2	(1.13, 1.34)	1.00	(0.48, 2.11)	0.89	(0.46, 1.74)	0.93	(0.68, 1.27)

Significant factors ($p < 0.05$) highlighted in red

a – Quartiles of individuals relative to all 5 counties among black population only

Table 7. Risk attributable preterm births for potentially modifiable factors among African-Americans only by prevalence area

	Entire population N=5,971		Low prevalence area N=132		Middle prevalence area N=375		High prevalence area N=1,424	
	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)
Maternal age								
< 18	3	0.0	3	2.0	-7	-1.8	14	1.0
≥ 35	156	2.6	14	10.8	34	9.1	37	2.6
Diabetes	31	0.5	1	0.6	0	-0.1	8	0.5
Hypertension (chronic)	122	2.0	9	6.5	10	2.7	14	1.0
Previous preterm/ small birth	102	1.7	0	0.0	11	2.9	12	0.9
Any tobacco use	84	1.4	-1	-0.5	0	0.0	45	3.2
Hypertension (pregnancy associated)	193	3.2	9	6.9	15	4.1	26	1.8
Eclampsia	97	1.6	0	0.0	10	2.7	18	1.3
Interpregnancy interval								
< 18 months	231	3.9	3	1.9	19	5.0	73	5.1
> 59 months	32	0.5	1	0.5	-3	-0.7	30	2.1

Significant risk factors ($p < 0.05$) are highlighted in red

Table 8. Adjusted odds ratios for preterm birth for non-Hispanic white women by prevalence area

	Entire region N=29,232		Low prevalence area N=2,304		Middle prevalence area N=4,674		High prevalence area #1 N=271		High prevalence area #2 N=574	
	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI
Maternal demographic factors										
Age										
<18	0.97	(0.69, 1.37)	ł		0.59	(0.22, 1.57)	ł		0.64	(0.20, 2.01)
18-34	Ref.		Ref.		Ref.		Ref.		Ref.	
≥35	1.09	(0.98, 1.20)	1.01	(0.68, 1.50)	1.21	(0.93, 1.56)	0.85	(0.27, 2.70)	1.35	(0.47, 3.87)
Education										
<High school	1.13	(0.95, 1.34)	0.85	(0.25, 2.88)	1.30	(0.83, 2.04)	1.97	(0.70, 5.50)	1.78	(0.98, 3.24)
Completed high school/ GED	Ref.		Ref.		Ref.		Ref.		Ref.	
Some college or higher	0.84	(0.75, 0.94)	0.64	(0.34, 1.21)	0.87	(0.66, 1.14)	0.36	(0.14, 0.93)	1.29	(0.51, 3.23)
Single	1.33	(1.18, 1.50)	1.73	(1.00, 2.99)	1.02	(0.74, 1.42)	1.16	(0.49, 2.75)	2.28	(1.22, 4.25)
Pre-pregnancy risk factors										
Diabetes	1.37	(1.08, 1.74)	1.57	(0.59, 4.18)	1.03	(0.55, 1.92)	1.89	(0.24, 14.69)		
Hypertension (chronic)	2.15	(1.50, 3.08)	1.41	(0.28, 7.23)	2.11	(0.92, 4.87)	3.62	(0.12, 111.05)	3.17	(0.25, 39.96)
Previous preterm or small birth	3.50	(2.69, 4.55)	3.16	(0.82, 12.14)	4.47	(2.36, 8.46)	15.36	(0.65, 362.14)	1.19	(0.09, 15.43)
Prior fetal death	1.07	(0.96, 1.19)	1.01	(0.63, 1.60)	1.03	(0.79, 1.33)	0.61	(0.20, 1.86)	1.14	(0.44, 2.98)
Prior induced termination	1.11	(0.95, 1.28)	1.11	(0.62, 1.99)	0.99	(0.61, 1.61)	4.25	(1.39, 12.97)	0.99	(0.51, 1.90)

Significant factors ($p < 0.05$) highlighted in red

a – Quartiles of individuals relative to all 5 counties among white population only

ł – Addition of variable caused non-convergence of regression model, therefore it was not included in the final model.

Table 8 (cont.). Adjusted odds ratios for preterm birth for non-Hispanic white women by prevalence area

	Entire region N=29,232		Low prevalence area N=2,304		Middle prevalence area N=4,674		High prevalence area #1 N=271		High prevalence area #2 N=574	
	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI
During pregnancy risk factors										
Any tobacco use	1.10	(0.92, 1.31)	1.82	(0.56, 5.87)	0.59	(0.31, 1.12)	0.52	(0.14, 1.98)	1.42	(0.76, 2.62)
Hypertension (pregnancy)	2.38	(2.00, 2.84)	1.94	(0.72, 5.22)	2.79	(1.93, 4.03)	2.08	(0.28, 15.36)	ǂ	
Eclampsia	5.00	(3.36, 7.43)	6.24	(1.51, 25.79)	3.46	(0.96, 12.51)	ǂ		21.42	(1.18, 388.67)
Interpregnancy interval										
<18 months	1.50	(1.25, 1.81)	1.89	(0.80, 4.46)	1.40	(0.84, 2.32)	2.06	(0.55, 7.75)	2.83	(1.09, 7.34)
18-59 months	Ref.		Ref.		Ref.		Ref.		Ref.	
>59 months	1.28	(1.10, 1.50)	2.38	(1.17, 4.83)	0.92	(0.61, 1.40)	0.52	(0.11, 2.37)	0.52	(0.14, 2.01)
Primigravida	1.20	(1.09, 1.32)	1.15	(0.75, 1.75)	1.14	(0.90, 1.44)	1.53	(0.60, 3.90)	1.15	(0.60, 2.19)
Neighborhood level risk factors [a]										
Deprivation Index quartile										
<25% (least deprived)	Ref.		Ref.		Ref.		Ref.		ǂ	
25-49.9%	1.04	(0.93, 1.17)	0.85	(0.53, 1.36)	0.99	(0.75, 1.32)	0.41	(0.05, 3.46)	ǂ	
50-74.9%	1.08	(0.96, 1.22)	0.91	(0.50, 1.67)	0.98	(0.74, 1.31)	1.31	(0.53, 3.20)	ǂ	
≥75%	0.97	(0.86, 1.10)	0.80	(0.51, 1.28)	ǂ		ǂ		ǂ	

Significant factors ($p < 0.05$) highlighted in red

a – Quartiles of individuals relative to all 5 counties among white population only

ǂ – Addition of variable caused non-convergence of regression model, therefore it was not included in the final model.

Table 9. Risk attributable preterm births for potentially modifiable factors among white mothers only by prevalence area

	Entire population N= 2,596		Low prevalence area N=136		Middle prevalence area N=409		High prevalence area #1 N=38		High prevalence area # N=68	
	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)
Maternal age										
< 18	-1	-0.1	0	0.0	-4	-0.9	ł		-2	-3.3
≥ 35	50	1.9	1	0.4	16	3.9	-1	-2.3	1	1.9
Diabetes	22	0.9	2	1.3	0	0.1	1	2.5	ł	
Hypertension (chronic)	21	0.8	1	0.4	4	0.9	1	1.9	1	1.0
Previous preterm/ small birth	59	2.3	2	1.5	12	2.8	1	2.5	0	0.2
Any tobacco use	14	0.6	2	1.3	-8	-1.9	-5	-12.2	6	8.2
Hypertension (pregnancy associated)	107	4.1	2	1.8	26	6.4	1	2.7	ł	
Eclampsia	37	1.4	3	1.9	3	0.7	ł		1	1.4
Interpregnancy interval										
< 18 months	51	2.0	3	2.4	6	1.4	3	6.8	6	9.5
> 59 months	57	2.2	8	5.5	-3	-0.7	-3	-7.4	-3	-4.0

Significant factors ($p < 0.05$) highlighted in red

ł – Addition of variable caused non-convergence of regression model, therefore it was not included in the final model.

Table 10. Imputation results: Adjusted odds ratios for preterm birth for all races by prevalence area

	Entire region N=160,004		Low prevalence area N=9,798		Middle prevalence area N=27,580		High prevalence area N=65,785	
	aOR	95% CI	aOR	95% CI	aOR	95% CI	aOR	95% CI
Maternal demographic factors								
Age								
<18	1.12	(1.03, 1.22)	1.00	(0.45, 2.23)	2.03	(1.24, 3.34)	1.07	(0.88, 1.29)
18-34	Ref.		Ref.		Ref.		Ref.	
≥35	1.19	(1.14, 1.24)	1.30	(1.10, 1.53)	1.07	(0.85, 1.34)	1.33	(1.12, 1.57)
Education								
<High school	1.10	(1.05, 1.15)	0.85	(0.56, 1.30)	0.83	(0.61, 1.12)	1.16	(1.03, 1.30)
Completed high school/ GED	Ref.		Ref.		Ref.		Ref.	
Some college or higher	0.83	(0.79, 0.87)	0.83	(0.66, 1.05)	1.00	(0.79, 1.26)	0.81	(0.72, 0.92)
Race/ ethnicity								
Non-Hispanic black	1.58	(1.51, 1.65)	1.35	(1.08, 1.68)	1.41	(1.12, 1.79)	1.37	(1.05, 1.79)
Non-Hispanic white	Ref.		Ref.		Ref.		Ref.	
Hispanic	0.97	(0.91, 1.02)	0.95	(0.72, 1.25)	1.16	(0.89, 1.53)	0.92	(0.67, 1.26)
Other	1.10	(1.02, 1.18)	0.87	(0.71, 1.08)	0.97	(0.68, 1.39)	0.91	(0.55, 1.52)
Single	1.21	(1.16, 1.25)	1.66	(1.34, 2.06)	1.11	(0.91, 1.37)	1.23	(1.08, 1.40)
Pre-pregnancy risk factors								
Diabetes	1.40	(1.25, 1.57)	1.85	(1.23, 2.77)	1.13	(0.54, 2.36)	1.51	(0.96, 2.36)
Hypertension (chronic)	3.18	(2.74, 3.70)	3.69	(2.07, 6.58)	2.00	(0.80, 5.03)	3.12	(1.98, 4.91)
Previous preterm or small	4.24	(3.71, 4.84)	4.76	(2.98, 7.63)	2.95	(1.25, 6.96)	3.06	(2.08, 4.50)
Prior fetal death	1.22	(1.16, 1.28)	1.22	(1.00, 1.48)	1.16	(0.92, 1.46)	1.19	(1.06, 1.34)
Prior induced termination	1.18	(1.12, 1.25)	0.96	(0.70, 1.30)	1.28	(0.92, 1.78)	1.13	(0.99, 1.28)
During pregnancy risk factors								
Any tobacco use	1.40	(1.27, 1.55)	0.99	(0.52, 1.86)	1.72	(1.12, 2.63)	1.63	(1.34, 1.98)
Hypertension (pregnancy)	2.61	(2.34, 2.90)	2.68	(1.81, 3.96)	1.83	(0.97, 3.47)	2.32	(1.73, 3.12)
Eclampsia	6.37	(5.23, 7.76)	4.73	(2.32, 9.67)	6.93	(2.25, 21.34)	8.03	(4.16, 15.53)
Interpregnancy interval								
<18 months	1.63	(1.53, 1.73)	1.53	(1.07, 2.17)	1.48	(1.04, 2.10)	1.62	(1.39, 1.90)
18-59 months	Ref.		Ref.		Ref.		Ref.	
>59 months	1.08	(1.02, 1.14)	1.25	(0.97, 1.60)	0.96	(0.72, 1.28)	1.09	(0.93, 1.28)
Primipara	1.12	(1.08, 1.17)	1.36	(1.14, 1.61)	1.20	(0.99, 1.47)	0.96	(0.86, 1.08)
Neighborhood level risk factors [a]								
Deprivation Index quartile								
<25% (least deprived)	Ref.		Ref.		Ref.		Ref.	
25-49.9%	1.05	(1.00, 1.10)	1.14	(0.79, 1.63)	0.97	(0.79, 1.19)	0.68	(0.37, 1.22)
50-74.9%	1.06	(1.00, 1.11)	-	-	0.93	(0.64, 1.34)	0.74	(0.41, 1.33)
≥75%	1.15	(1.09, 1.21)	-	-	0.83	(0.63, 1.09)	0.78	(0.44, 1.39)

Significant factors ($p < 0.05$) highlighted in red

a – Quartiles of individuals relative to all 5 counties

Table 11. Imputation results: Risk attributable preterm births for potentially modifiable factors among all races by prevalence area

	Entire population N= 18,856		Low prevalence area N= 826		Middle prevalence area N=3,215		High prevalence area N=12,250	
	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)	N	Adjusted AR (%)
Maternal age								
< 18	92	0.5	0	0.0	66	2.1	62	0.5
≥ 35	509	2.7	65	7.9	35	1.1	276	2.2
Diabetes	180	1.0	22	2.6	8	0.2	103	0.8
Hypertension (chronic)	293	1.6	18	2.1	18	0.6	166	1.4
Previous preterm/ small birth	346	1.8	30	3.6	36	1.1	176	1.4
Any tobacco use	200	1.1	0	0.0	65	2.0	341	2.8
Hypertension (pregnancy associated)	581	3.1	37	4.5	45	1.4	259	2.1
Eclampsia	260	1.4	15	1.8	35	1.1	164	1.3
Interpregnancy interval								
< 18 months	634	3.4	15	1.8	76	2.4	612	5.0
> 59 months	177	0.9	21	2.5	-15	-0.5	130	1.1

Significant factors ($p < 0.05$) highlighted in red

Table 12. Percent change in precision between complete case and imputation analysis

	Entire study area			Low prevalence area			Middle prevalence area			High prevalence area		
	CI Width		Percent	CI Width		Percent	CI Width		Percent	CI Width		Percent
	Complete case method	Imputation method	gain in precision	Complete case method	Imputation method	gain in precision	Complete case method	Imputation method	gain in precision	Complete case method	Imputation method	gain in precision
Maternal demographic factors												
Age												
< 18	0.25	0.19	23.88	1.92	1.79	7.01	3.12	2.10	32.77	0.52	0.40	21.52
18-34												
≥ 35	0.14	0.11	21.93	0.61	0.43	29.11	0.62	0.50	19.40	0.54	0.45	17.91
Education												
< High school	0.13	0.11	20.78	0.87	0.74	15.36	0.54	0.52	5.13	0.33	0.27	17.67
Completed high school/ GED												
Some college or higher	0.09	0.07	14.59	0.50	0.39	21.29	0.58	0.48	17.84	0.28	0.20	28.75
Race/ ethnicity												
Non-Hispanic black	0.19	0.15	21.10	0.75	0.59	21.12	0.76	0.67	11.22	1.61	0.74	54.07
Non-Hispanic white												
Hispanic	0.14	0.11	18.71	0.69	0.53	23.55	0.81	0.64	21.41	1.21	0.59	51.13
Other	0.20	0.16	20.19	0.46	0.37	18.96	1.05	0.72	31.80	1.91	0.98	48.81
Single	0.11	0.09	20.74	0.95	0.72	23.99	0.64	0.46	28.17	0.36	0.32	11.27
Pre-pregnancy risk factors												
Diabetes	0.32	0.32	-1.10	1.66	1.54	7.63	1.77	1.82	-3.18	1.25	1.40	-12.17
Hypertension (chronic)	0.85	0.96	-13.32	4.30	4.51	-4.92	3.96	4.24	-6.87	2.95	2.93	0.65
Previous preterm/small birth	1.08	1.13	-4.76	5.06	4.65	8.11	6.66	5.71	14.32	2.07	2.42	-16.85
Prior fetal death	0.13	0.12	10.86	0.54	0.48	11.02	0.63	0.55	13.00	0.35	0.28	20.11
Prior induced termination	0.14	0.14	2.36	0.73	0.60	17.63	1.17	0.85	26.75	0.34	0.29	12.85
During pregnancy risk factors												
Any tobacco use	0.31	0.27	11.20	1.20	1.34	-11.59	1.74	1.51	13.36	0.81	0.64	20.96
Hypertension (pregnancy associated)	0.50	0.55	-10.60	2.19	2.15	1.73	2.12	2.50	-18.01	1.57	1.39	11.15
Eclampsia	2.00	2.53	-26.50	7.98	7.36	7.84	16.72	19.09	-14.21	9.12	11.37	-24.65
Interpregnancy interval												
< 18 months	0.26	0.20	20.11	1.62	1.10	31.88	1.31	1.06	18.98	0.61	0.51	16.32
18-59 months												
> 59 months	0.15	0.12	20.72	0.76	0.63	17.27	0.63	0.56	11.84	0.45	0.35	22.62
Primigravida	0.11	0.09	19.50	0.76	0.47	37.97	0.51	0.48	5.19	0.27	0.22	20.04
Neighborhood level risk factors [a]												
Deprivation Index quartile												
<25% (least deprived - referent)												
25-49.9%	0.13	0.10	22.07	1.34	0.83	37.83	0.46	0.40	13.10	2.65	0.85	67.90
50-74.9%	0.14	0.11	20.33	-	-	-	0.85	0.70	17.80	2.55	0.92	63.90
≥ 75%	0.16	0.12	21.50	-	-	-	0.55	0.46	15.63	2.93	0.94	67.84

Figures

Figure 1. Distribution of smoothed preterm birth prevalence rates among all races/ethnicities

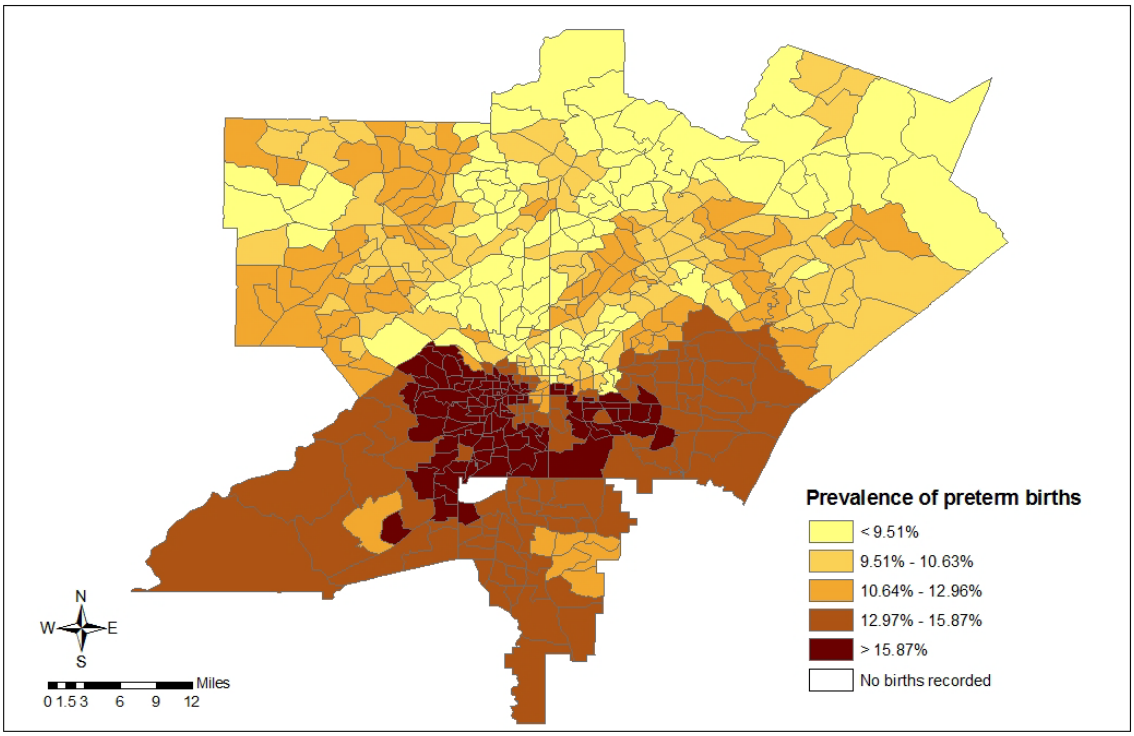

Figure 2. Distribution of smoothed preterm birth prevalence among non-Hispanic black mothers

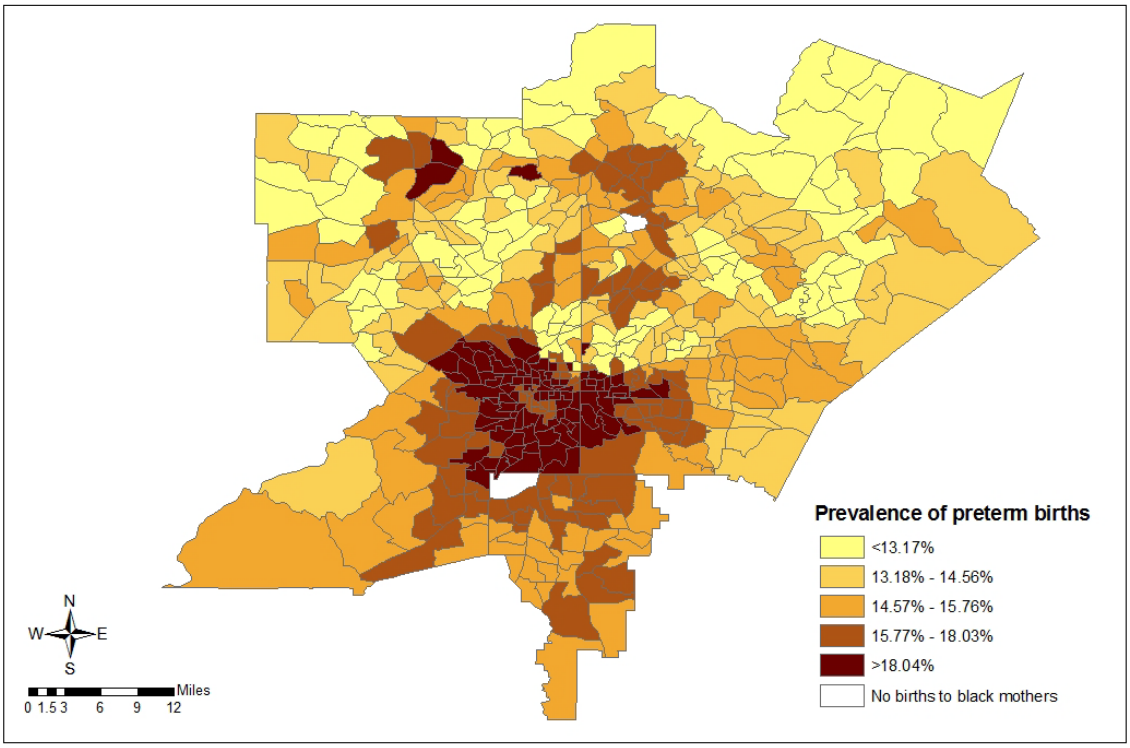

Figure 3. Distribution of smoothed preterm birth prevalence among non-Hispanic white mothers

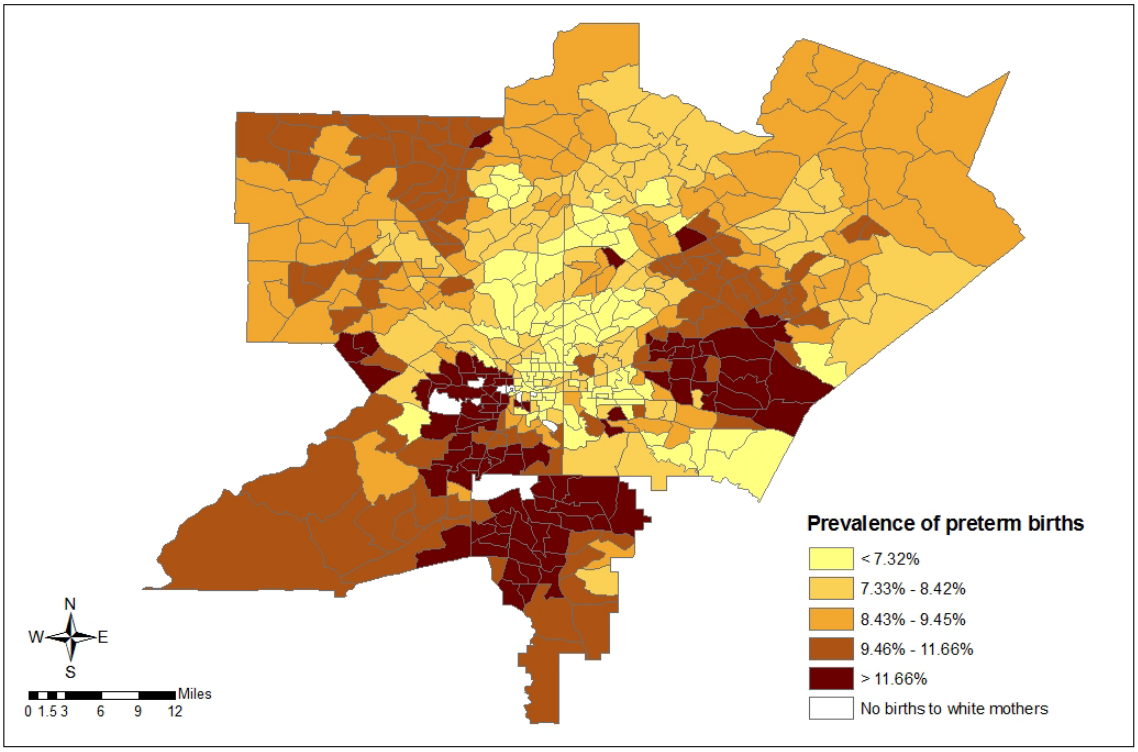

Figure 4. High, low and middle prevalence regions among all races/ethnicities

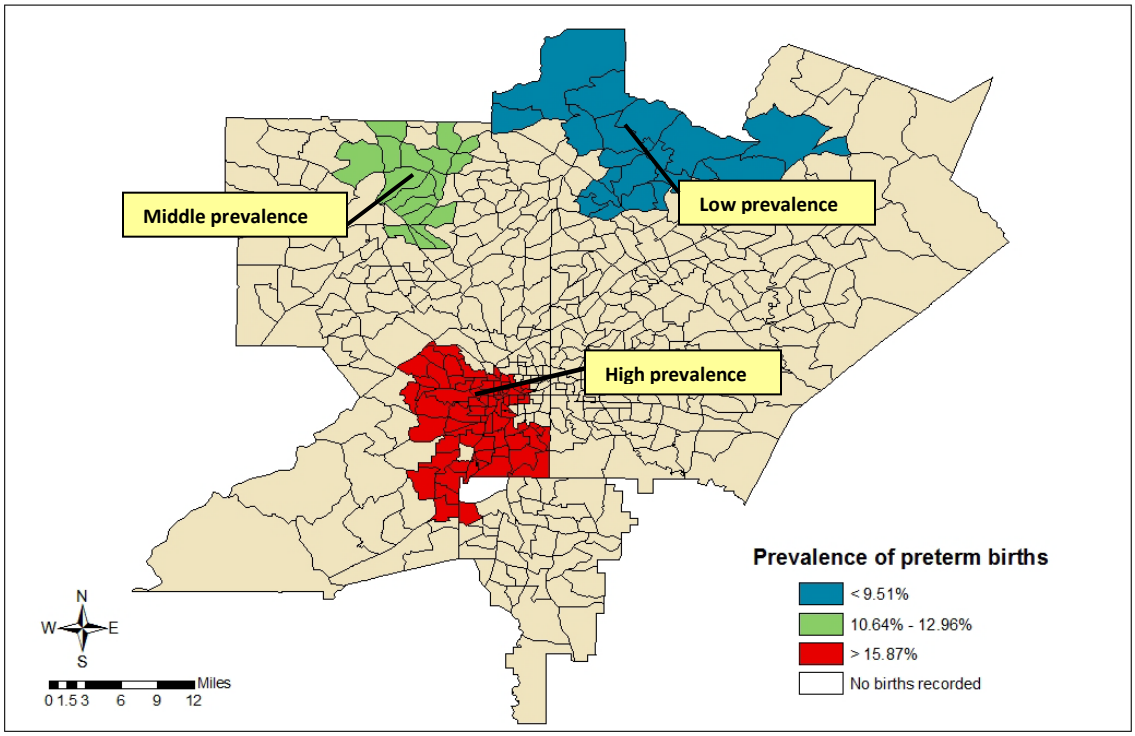

Figure 5. High, low and middle prevalence regions among all non-Hispanic black mothers

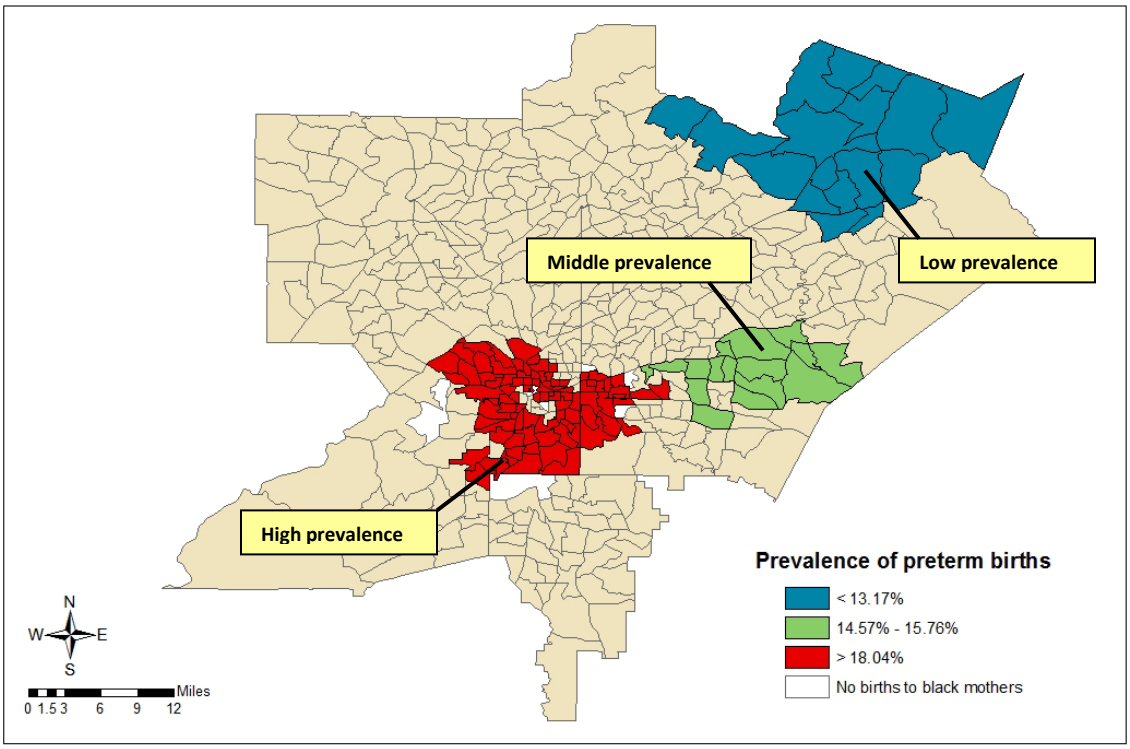

Figure 6. High, low and middle prevalence regions among all non-Hispanic white mothers