2018

How Fine Particulate Matter Modifies Preterm Birth Risks in Korea

Hyun Jin Choi
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral dissertation by

Hyun Jin Choi

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Cheryl Cullen, Committee Chairperson, Public Health Faculty
Dr. Leslie Elliott, Committee Member, Public Health Faculty
Dr. Mehdi Agha, University Reviewer, Public Health Faculty

Chief Academic Officer
Eric Riedel, Ph.D.

Walden University
2018

Abstract

How Fine Particulate Matter Modifies Preterm Birth Risks in Korea

by

Hyun Jin Choi

BS, Utah State University, 2002

BS, Yonsei University, 2000

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

August 2018

Abstract

Despite the increasing interest in preterm birth risk associated with maternal exposure to the current level of fine particulate matter ($PM_{2.5}$) in Korea, there is little information on differences in $PM_{2.5}$ exposure and its impact on preterm birth. This study was designed to examine the effects of Korea's air quality on preterm birth, including the possibility of moderation and mediation. This work was also designed to investigate manipulable factors for $PM_{2.5}$ exposure. The theoretical framework of this quantitative and observational study included the social ecological theory and systems theory. The conceptual framework of this partially ecologic and retrospective cohort study included the social ecological model and Rothman's sufficient component cause model. Data of 19,371 Korean women who gave birth in 2015 were analyzed by logistic regression and multiple regression, including testing for moderation and mediation. An increase in $PM_{2.5}$ exposure by 10 $\mu g/m^3$ in the 3rd week before childbirth increased the likelihood of preterm birth by 6.52 times. Moderation and mediation by $PM_{2.5}$ did not exist between sociodemographic factors and gestational age but existed between socioeconomic and energy policy factors and gestational age. The most influential factor for $PM_{2.5}$ exposure was unemployment rate at the organizational level. These results show the need for socioeconomic interventions to reduce $PM_{2.5}$ exposure more effectively. These findings indicate that prenatal care should be addressed with a socioeconomic- and energy-policy-sensitive approach to lower preterm birth due to severe air pollution in Korea. This study has the potential to change people's perceptions of threats from $PM_{2.5}$ exposure, which could lead to behavior changes.

How Fine Particulate Matter Modifies Preterm Birth Risks in Korea

by

Hyun Jin Choi

BS, Utah State University, 2002

BS, Yonsei University, 2000

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

August 2018

Acknowledgments

I would like to express my deepest gratitude to Dr. Cheryl Cullen, who led me with her strong leadership and extensive knowledge, motivated me to make progress in my doctoral dissertation and provided me with constructive feedback on how to improve my work, particularly with regard to the relevancy of research concepts to the profession and social change. I would like to thank Dr. Leslie Elliott, who shared her knowledge of the concepts of research and encouraged me to look at things from different perspectives. My special thanks to Dr. Mehdi Agha, who taught me Advanced Biostatistics at Walden University. Thanks to all professors at Walden University.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Introduction

Preterm birth is defined as a birth having a gestational age of 36 weeks plus six days or less at delivery of the infant, missing the final weeks of pregnancy (Symanski et al., 2014). Preterm birth is a significant risk factor for infant morbidity and mortality, as well as morbidity and mortality in later life since the final weeks of pregnancy are necessary for the complete development of the vital organs (Leal et al., 2016; Wallace et al., 2016). The economic burden associated with preterm birth includes the increased maternal delivery costs in the short-term because of the likelihood of a surgical cesarean section birth instead of vaginal birth, prolonged hospital length of stay, and slower recovery time after discharge to home (Trasande et al., 2016; Yoon, 2009). Care of the infant may include both short-term and long-term costs. Higher morbidity and mortality, based on gestational age, are associated with intensive care, medications, the potential for physical or learning disabilities, and chronic health concerns increasing social and financial burden (Ha et al., 2014; Leal et al., 2016; Trasande et al., 2016).

The incidence of preterm birth in Korea rose 44% from 4.8% in 2005 to 6.9% in 2015 (Korean Statistical Information Service [KOSIS], 2016). This indicates that the given efforts to reduce the prevalence rate of preterm birth have not been effective at this time and marks the need for a new approach that addresses emerging health threats so that Korean leaders can establish efficacious preventive measures. One of the current health threats in Korea is exposure to fine particulate matter ($PM_{2.5}$). $PM_{2.5}$ refers to particulate matter less than 2.5 μm in aerodynamic diameter. Maternal $PM_{2.5}$ exposure is associated with the increase in oxidative stress and intrauterine inflammation, which may trigger pathways leading to preterm birth (Ha et al., 2014; Lee at al., 2013; Menon, 2014). The 2015 annual mean concentration of $PM_{2.5}$ in Korea was 26 µg/m^3, which is 2.6 times the World Health Organization (WHO) safe level of 10 µg/m^3 (Korea Ministry of Environment [KME], 2016; WHO, 2016). The 2015 annual exposure to $PM_{2.5}$ exceeded the WHO limit in many areas of Korea as shown in Figure 1.

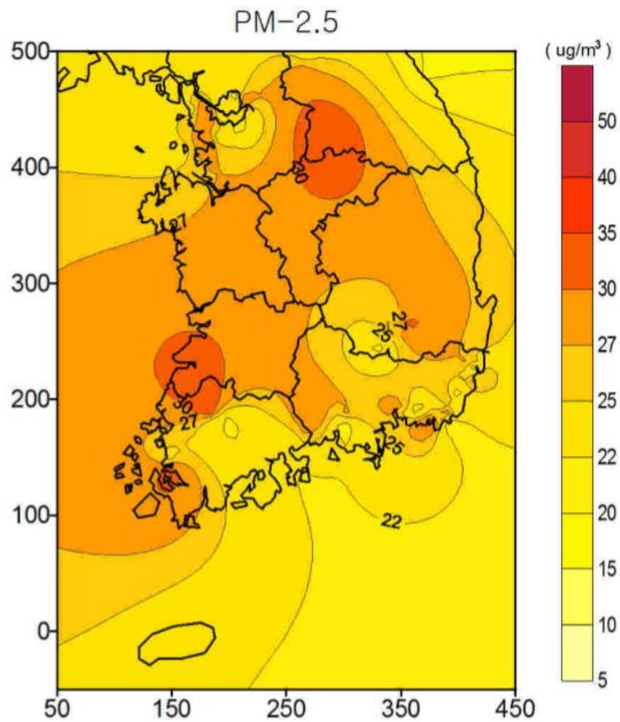

Figure 1. Spatial distribution of 2015 annual mean concentrations of $PM_{2.5}$ in Korea. Adapted from "Annual Report of Air Quality in Korea, 2015" by KME, 2016, p. 19 Retrieved from http://library.me.go.kr/search/DetailView.ax?sid=1&cid=5618423.

The standard for $PM_{2.5}$ in Korea is 25 µg/m^3, which is 2.5 times the WHO guideline of 10 µg/m^3. This 2.5-fold $PM_{2.5}$ regulation may give rise to severe air pollution in Korea and may not protect vulnerable populations, especially pregnant women (KME, 2016; Lombardo & Buckeridge, 2007). The incidence of preterm birth associated with maternal exposure to $PM_{2.5}$ has been increasing (Laurent et al., 2016; Stieb et al., 2016; Trasande et al., 2013). There is growing interest in examining the effects the current level of $PM_{2.5}$ has on the occurrence of preterm birth in Korea so that leaders can make informed decisions on setting air quality standards to protect public health.

In Korea, sociodemographic characteristics that may be correlated with preterm birth and maternal $PM_{2.5}$ exposure have changed (Erickson, Ostry, Chan, & Arbour, 2016; Kim et al., 2005; Leal et al., 2016; Messer et al., 2008; OECD, 2015; Yi, Kim, & Ha, 2010). Evidence regarding the sociodemographic status of mothers associated with preterm birth and $PM_{2.5}$ exposure is needed to demonstrate the need for financing initiatives targeting vulnerable populations, which may maximize favorable health outcomes at the overall level.

The social change impact of the findings from this study may include influencing people's perceptions of the threat from $PM_{2.5}$ exposure, leading to behavior change and identification of where to intervene to control $PM_{2.5}$ exposure and associated preterm birth risks. Findings from this study may also contribute to reducing inequities in $PM_{2.5}$ exposure and its associated health disparities in Korea. With no scholarly evidence on the effects of current levels of $PM_{2.5}$ on the occurrence of preterm delivery, no effort will be taken to curb the situations of severe air pollution, which may result in the continued rise in the rate of preterm births in Korea.

This chapter introduces identified gaps in the literature and my chosen research problems. I discuss the alignment among research components including the purpose of the study, research questions, hypotheses, theory, concept, and design. Further, I describe the nature of this study, the operationalization of variables, threats to validity, and strategies to mitigate them. The subsequent sections outline my assumptions and the scope and delimitations of this study. The relationship between this research and social change is then presented, and the chapter concludes with a summary and transition to the next chapter.

Background

Recent Research on Preterm Birth

Researchers have investigated biological and physiological etiologies of preterm birth based on medical factors associated with excessive medicalization—an increase in health care utilization, including preconception care (Goossens et al., 2016; Leal et al., 2016; Tellapragada et al., 2016). Their efforts have helped improve the clinical understanding of individual differences in susceptibility to preterm birth according to medical variables. However, few researchers have examined risk factors beyond individual medical factors, including social, physical environmental, and policy factors. In studying the association between medical factors and preterm birth, researchers have regarded the covariates associated with both independent and dependent variables only as their potential confounders. Few researchers have examined the potential moderation and mediation by the covariates to account for the preterm birth risk. Research efforts are needed to understand factors other than medical ones associated with preterm birth, including the possibility of moderation and mediation.

Consequences of Inadequate Thresholds for $PM_{2.5}$

Setting a high threshold for $PM_{2.5}$ may decrease the perceived threat of $PM_{2.5}$ (Parker et al., 2004). When the perceived threat is lower than the actual risk, the benefit of interventions is likely to be underestimated, which may not trigger timely action. Thresholds play a critical role in making decisions and implementing measures to mitigate damage to public health (Lombardo & Buckeridge, 2007). Thresholds provide public reassurance that no health threat exists in cases below the threshold value (Lombardo & Buckeridge, 2007). Thresholds should be determined based on empirical or theoretical evidence (Lombardo & Buckeridge, 2007). The thresholds for $PM_{2.5}$ in Korea is 25 $\mu g/m^3$, which is 2.5 times the WHO guideline. A gap in the literature exists regarding the risk of preterm birth when maternal exposure to $PM_{2.5}$ is between 10 and 25 $\mu g/m^3$ compared to exposure of less than 10 $\mu g/m^3$.

Assessment of $PM_{2.5}$ Exposure

Researchers have used databases of various air pollution monitoring systems to assess exposure to $PM_{2.5}$. The $PM_{2.5}$ data from air quality databases are not measured at the individual exposure domains of the study participants. Researchers have used surrogate $PM_{2.5}$ concentrations measured at $PM_{2.5}$ monitoring stations closest to the study participants' residences to calculate individual $PM_{2.5}$ exposure (DeFranco et al., 2016; Ha et al., 2014; Laurent et al., 2016). Many researchers have attempted to fill a gap in this field of environmental research by improving the accuracy of individual $PM_{2.5}$ exposure estimates. They have proposed various methods to infer better individual exposures levels including modified kriging, Bayesian, geographic information system (GIS), modeled estimates based on satellite data, and land-use regression methods (DeFranco et al., 2016; Ha et al., 2014; Hyder et al., 2014; Laurent et al., 2016; Leem et al., 2006; Suh et al., 2009). Researchers have aimed at yielding approximations of $PM_{2.5}$ concentration close to the true value of $PM_{2.5}$ measured at individual exposure domains (Hao et al., 2016; Hyder et al., 2014; Laurent et al., 2016). This focus on accuracy in individual exposure assessment emphasized individual differences in response to $PM_{2.5}$ exposure and individual-focused behavior changes promoting avoidance of $PM_{2.5}$ exposure while overlooking strategies leading to macro-level changes that contribute to reducing social and political disparities in $PM_{2.5}$ exposure. Considering that it is often not possible to decrease exposure to polluted air through individual behavior change, I instead focused on social and political determinants generating group differences in $PM_{2.5}$ exposure including proximity to and the number of coal power plants.

This study could provide evidence regarding the effects of Korea's air quality on preterm birth, including potential moderation and mediation by the air quality in an independent sociodemographic effect on gestational age. The study could also provide evidence regarding variables that explain differences in maternal exposure to $PM_{2.5}$. The study is needed to identify where to intervene to reduce social and political disparities in $PM_{2.5}$ exposure and to enhance monitoring of health disparities associated with $PM_{2.5}$. This study is also needed to develop strategies to curb the situations of severe air pollution and reduce the incidence of preterm birth in Korea.

Problem Statement

A total of 30,408 Korean infants were born preterm in 2015, and the medical costs required for them is estimated at ₩1 trillion ($900 million) from birth to the first 3 years (KOSIS, 2016; Kim, Lee, Kim, & Park, 2016; Yoon, 2009). The problem is that the incidence of preterm birth in Korea rose 44%, from 4.8% in 2005 to 6.9% in 2015 (KOSIS, 2016). Another problem is that Korea is one of the world's most air-polluted countries (OECD Better Life Index.org, 2016). Annual exposure to $PM_{2.5}$ in 2015 exceeded WHO safe levels in many areas of Korea (KME, 2016). The evidence of an association between maternal exposure to $PM_{2.5}$ and preterm birth has been increasing (Hao et al., 2016 & Laurent et al., 2016). Despite increasing concerns about the preterm birth risk associated with maternal exposure to the current level of $PM_{2.5}$, there is a dearth of information regarding what variables can explain the differences in $PM_{2.5}$ exposure and how the air quality in Korea modifies preterm birth risk.

Purpose of the Study

This was a quantitative study to determine the current best predictors of preterm birth and $PM_{2.5}$ exposure. I quantified moderation and mediation by $PM_{2.5}$ in the relationship of sociodemographic factors of maternal age, maternal education level, and maternal occupation status, to gestational age. The study was designed to examine the preterm birth risk associated with maternal $PM_{2.5}$ exposure, including the possibility of moderation and mediation. Further, I sought to investigate social ecological factors modeled at the individual, community, organizational, and policy levels that account for differences in maternal exposure to $PM_{2.5}$.

Research Questions and Hypotheses

I developed the following research questions and corresponding hypotheses:

RQ1: Do individual, organizational, and policy-level factors explain community-level $PM_{2.5}$ exposure?

$H_0$1: There are no statistically significant effects of individual, organizational, and policy-level factors on community-level $PM_{2.5}$ exposure.

H_a1: There are statistically significant effects of individual, organizational, and policy-level factors on community-level $PM_{2.5}$ exposure.

RQ2: What is the relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors?

$H_0$2: There is no statistically significant relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors.

H_a2: There is a statistically significant relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors.

RQ3: What effect, if any, does the interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, and occupation have on gestational age?

$H_0$3: The interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, occupation has no statistically significant effect on gestational age.

H_a3: The interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, occupation has a statistically significant effect on gestational age.

RQ4: What effect, if any, do sociodemographic factors, such as maternal age, education level, and occupation have on gestational age through $PM_{2.5}$ exposure?

$H_0$4: Sociodemographic factors, such as maternal age, education level, occupation have no statistically significant effects on gestational age through $PM_{2.5}$ exposure.

H_a4: Sociodemographic factors, such as maternal age, education level, occupation have statistically significant effects on gestational age through $PM_{2.5}$ exposure.

Theoretical Framework

The social ecological theory (SET) and systems theory provided a theoretical framework for this study. The SET offers ways of identifying contributing factors to individual or group differences in preterm birth from a multi-level perspective consisting of individual, community, organizational, and policy levels (Watt et al., 2016 & Erickson et al., 2016). This theory posits that determinants of preterm birth operate at different social-ecological levels, and that individual or group differences in preterm birth are results of complex multi-level interactions (Erickson et al., 2016). The SET is a useful approach to understanding differences in exposure to $PM_{2.5}$ at the community level and what leads to those differences between communities (Glouberman, & Millar, 2003 & Vandenbroucke, Broadbent, & Pearce, 2016).

The systems theory explains how the interrelated constructs function (Trochim, Cabrera, Milstein, Gallagher, & Leischow, 2006). This theory emphasizes the links between the individual constructs to predict a more reliable outcome of the complex system and is useful when investigating complex health issues (Peters, 2014 & Trochim et al., 2006). The systems thinking approach would hold that the relationship between the sociodemographic factors and preterm birth may change based on the levels of $PM_{2.5}$ exposure.

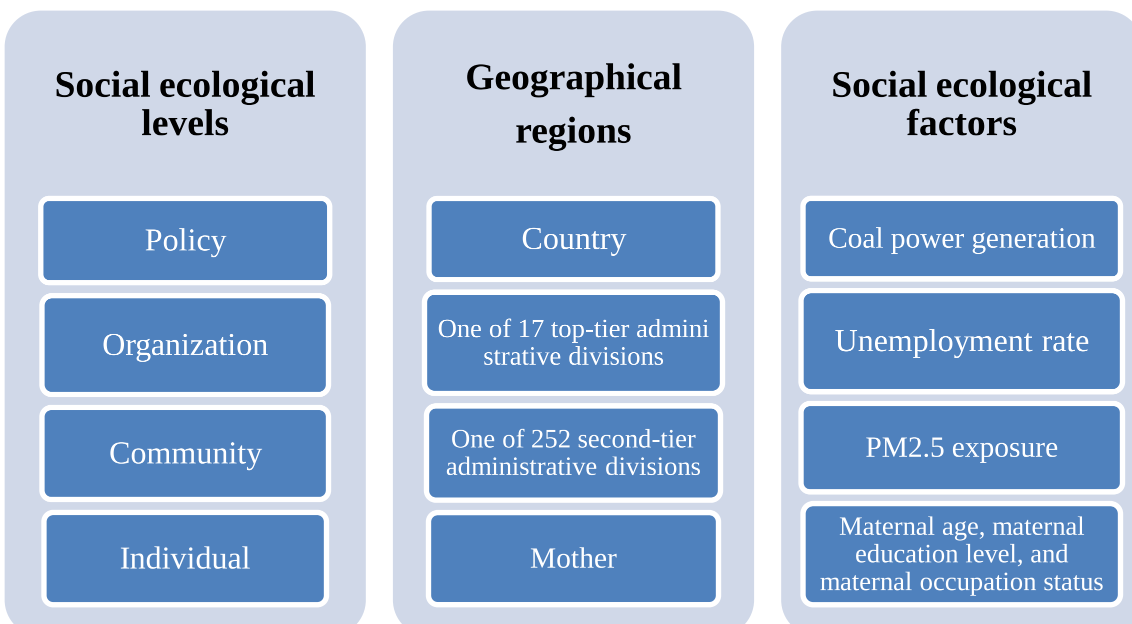

Figure 2. Social ecological factors.

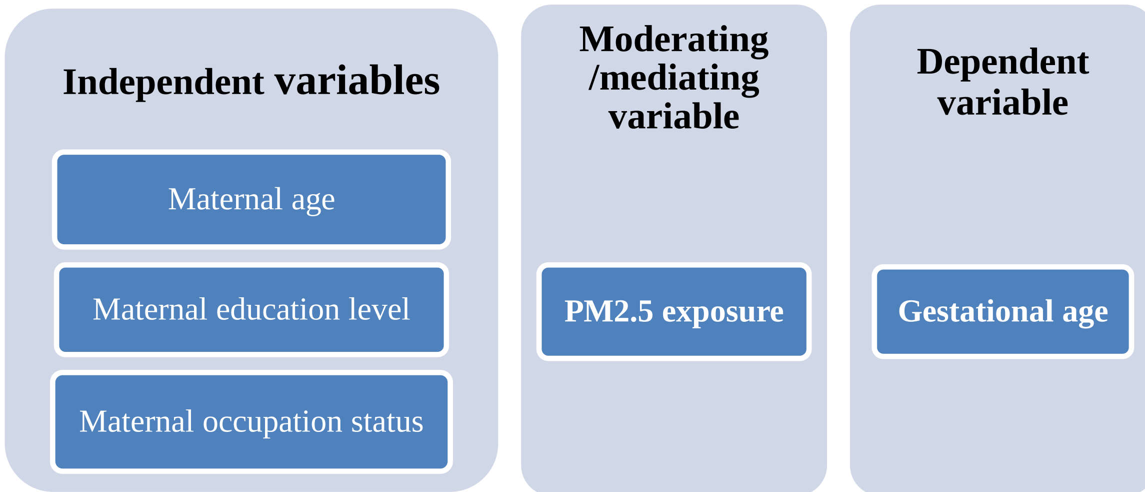

Figure 3. Systems theory perspective approach theoretical framework.

Conceptual Framework

The social ecological model (SEM) provided a conceptual framework for designing and evaluating the meteorological variable in this research. Specifically, SEM offers a framework for measuring meteorological variables while reducing the possibilities of measurement errors. In this investigation, I classified maternal exposure to $PM_{2.5}$ as a population-level factor based on the SEM because it cannot be accounted for by measures of individual attribute.

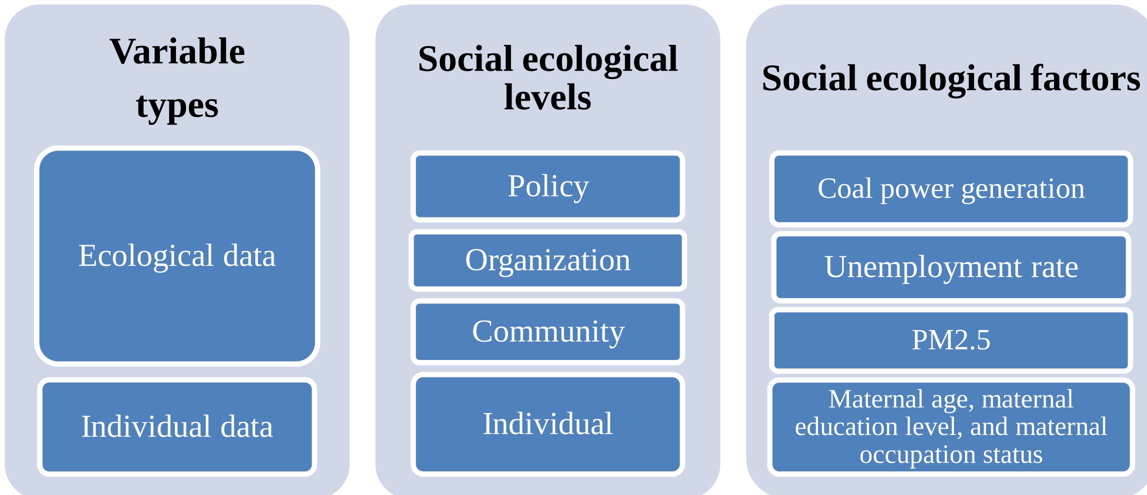

Figure 4. Types of the variables based on the social ecological model.

I also used Rothman's sufficient component cause model (RSCCM) as a conceptual framework for this research (Szklo, & Nieto, 2014). This model allows researchers to test hypotheses that can better correspond to current tendencies of significant health risks, including emerging health threats and new knowledge (Szklo, & Nieto, 2014). Researchers can construct a causal pie either by including emerging health threats or by excluding obsolete risk factors (Vandenbroucke, Broadbent, & Pearce, 2016). The causal pie thus formed is transformed into a deductive test of the causal hypothesis (Rothman & Greenland, 2005).

In this investigation, an emerging health concern, namely $PM_{2.5}$ exposure was an unknown component cause in the causal pie. Known risk factors formed part of the causal pie along with $PM_{2.5}$ exposure. This model provided the framework for testing the significance of overall joint effects of the components on the occurrence of preterm birth (see Rothman & Greenland, 2005). The outcome of preterm birth in the model depends on what component causes are in the model and whether their level is enough to affect the occurrence of preterm birth. I formed and tested the causal pie to determine whether $PM_{2.5}$ is a complementary component cause of maternal sociodemographic factors for preterm birth in Korea. The results of this study could show sufficient cause models that fit in the Korean population and indicate the most optimized preventive strategies for preterm birth in Korea.

Nature of the Study

This was a quantitative, correlational observation, partially ecologic, retrospective cohort study. The quantitative and correlational observation design aligns with the goal of answering research questions involving prevention or direct and indirect causal paths among the variables (Szklo, & Nieto, 2014). The partially ecologic study design is consistent with constructing individual level and ecological level variables based on the SET in this study. This study had a retrospective cohort design to assess historical exposures using existing data, including Vital Statistics of Korea for Live Birth (VSKLB), Korea Ministry of Environment Database (KMED), Economically Active Population Survey Statistics (EAPSS), and Statistics of Electric Power in Korea (SEPK). The factors that account for differences in $PM_{2.5}$ exposure were assessed using multiple linear regression. The association between maternal $PM_{2.5}$ exposure and preterm birth was evaluated using multiple logistic regression. The modified gestational age by $PM_{2.5}$ exposure was examined by moderation and mediation analysis using SPSS.

Definitions

Birth weight: The infant's weight measured at birth (KOSIS, 2016).

Duration of cohabitation: The period of time from the day the parents begin to live together to the birth of the infant, regardless of marriage status (KOSIS, 2016).

Education: The highest education institution completed based on the accredited education institutions by the minister of education (KOSIS, 2016.).

Coal power generation: A quantified capacity of a form of energy generated from coal. The amount of power generated from coal is measured as the unit of megawatts (KEPCO, 2016).

Full term: The gestational age of infants born between 37 and 40 weeks (Park, Kim, Lee, & Kim, 2008).

Gestational age: The time interval passed between the conception and birth of the newborn (Leem et al., 2006).

Marital status: The status of whether a person is legally married or unmarried (KOSIS, 2016).

Maternal age: The age of the mother at delivery. It is represented as the first six digits of the resident registration number generated based on the individual's date of birth (KOSIS, n.d.). For example, the first six digits of the person's resident registration number for a person born on July 9, 1995, would be 950709.

Preterm: The gestational age of infants born between 23 and 36 weeks (Cloherty, 2012).

Social ecological characteristics: Individual, community, organizational, and policy level characteristics of mothers. Social ecological variables in this study include maternal age, occupation, education, unemployment rate, and coal power generation

(Erickson at al., 2016; Glouberman, & Millar, 2003; Ji & Kang, 2016; Messer et al., 2008).

Sociodemographic characteristics: The social and demographic characteristics of the mother. Standard sociodemographic variables include age, education level, and occupation (Patra, Greene, Patel, & Meier, 2016).

Unemployed: People who have no work despite their efforts to find a job (KOSIS, 2016).

Unemployment rate: Percentage of unemployed persons to total number of unemployed, paid or self-employed (KOSIS, 2016).

Assumptions

This study required the following assumptions. First, I assumed that the maternal residence was the same as that reported at birth during the last 4 weeks of pregnancy. This assumption was necessary because I used data from a cross-sectional survey. Second, because it was not possible to cross-check with relevant data from other available sources, I assumed that the self-reported information on maternal education and occupation was offered in good faith and honestly to provide accurate data. Third, I assumed that the unemployment rate and coal powered generation remained the same as the reported monthly data over the same month. The daily value was equal to the monthly data for the month to which it belongs. This assumption was required because the data was collected at the monthly aggregation level.

Scope and Delimitations

This investigation was limited to Korean women for whom $PM_{2.5}$ data for the last 4 weeks of pregnancy were available. This study targeted women of Korean ethnicity who gave birth in 2015. Additional parameters were nulliparous, singleton births of women aged 20 to 44. Multiple gestations, preterm birth history, abortion, and advanced maternal age are known risk factors for preterm delivery (Leal et al., 2016), and women with one or more of those factors were excluded from the study. The chosen variables in this study included preterm birth, gestational age, $PM_{2.5}$ exposure, maternal age, maternal education level, maternal occupation status, unemployment rate, and coal power generation. The other variables for descriptive statistics were duration of cohabitation, infant birth weight, infant gender, and maternal marital status.

The study population was constrained to the residents of the sub-tier administrative areas in which air quality monitors for $PM_{2.5}$ were placed. $PM_{2.5}$ monitors were selectively located in high-density population areas such as Seoul, Busan, and Incheon (KME, 2016). The study sample may not adequately represent the entire population of pregnant woman in Korea, which limits the generalizability of study results, which should be applied only to women residing in the city.

Limitations

This study was limited by the secondary data available, which dictated the framing of the research questions, study sample, measures, variables, statistical analyses, and validity (see Smith et al., 2011). For example, secondary data from the birth record could not provide some of the variables known to be bi-variately correlated with $PM_{2.5}$

exposure and preterm birth, such as maternal smoking habits, drug use, alcohol consumption, and prenatal care utilization (Erickson et al., 2016; Goossens et al., 2016; Leal et al., 2016). All known confounders could not be controlled in this study, which may have misled some of my conclusions. To address the limitation of residual confounding, I set specific criteria for inclusion, which included an age range of 20-44 years, nulliparous, and singleton gestation. The other strategy was to use stratification from the statistical analyses to perform the analysis within a homogeneous subgroup of ages: 20 to 29, 30 to 39, and 40 to 44 (see Osborne, 2015).

Birth records based on self-reporting could not foreclose the risk of response bias (Hao et al., 2016 & Laurent et al., 2016). A significant amount of missing and erroneous data may produce biases, including additive effects (Langkamp, Lehman, & Lemeshow, 2010). To address the limitation of missing and incomplete data, I excluded birth records containing missing, unknown, or incomplete values from the analysis.

In compliance with the Personal Information Protection Act, geocoded home address was not available. The variables to be measured based on the place of residence such as maternal exposure to $PM_{2.5}$, unemployment rate, and coal power generation were assessed as area attributes in which the participant resides, at an aggregate level. Aggregated data was assigned to individual participants of this study, which may have led to information bias (see Szklo, & Nieto, 2014). I considered ecologic fallacy when interpreting the results of this study because aggregated data at an ecologic level may not accurately quantify or represent the individual level data (Szklo, & Nieto, 2014). The results of this study may not be replicable in cases where a different level of aggregation, such as different administrative units is used.

Significance

This research fills gaps in the literature by providing evidence regarding the possibilities of adequate management of the modified preterm birth risk by $PM_{2.5}$ exposure, which may contribute to a decreased preterm birth risk at an overall level. The results of this study may provide a more reliable prediction on the likelihood of preterm birth by examining whether indirect effects exist (see Ferguson et al., 2017; Peters, 2014). This research was unique because I applied SET and systems theory to the configuration and evaluation of variables and focused on the interactions between the variables (Buzi et al., 2015; Peters, 2014). The findings of this study may allow for assessment of the influence of each social-ecological level and identification of promising social-ecological level for public health initiatives.

This study was original because it was the first effort to examine impacts of $PM_{2.5}$ exposure on preterm birth in Korea since the air quality standard for $PM_{2.5}$ took effect in 2015 (KME, 2016). The results of this study may affect the evaluation process of the air quality standard for $PM_{2.5}$ in Korea. Insights from this study may be useful for public health professionals, healthcare providers, and policymakers when developing solutions to reduce preterm birth and drive the public health spending for those vulnerable to preterm birth and $PM_{2.5}$ exposure. As a result, efficiency in public health spending may be improved and inequities in $PM_{2.5}$ exposure and its attributable preterm birth may be decreased.

Summary

This study could provide evidence regarding the effect of air quality in Korea on the occurrence of preterm birth, including the moderation and mediation by the air quality. SET provided the theoretical framework I used to construct the research questions with the goal of understanding the variables that account for the differences in exposure to $PM_{2.5}$ and preterm birth. The SEM offered a conceptual framework for forming and evaluating the variables of this study. Rothman's sufficient component cause model provided a conceptual framework for determining whether $PM_{2.5}$ is a complementary component causes for preterm birth of maternal sociodemographic factors.

This study was a quantitative, observational, partially ecologic, retrospective cohort study using secondary data including those from VSKLB, KMED, EAPSS, and SEPK. The study variables included preterm birth, gestational age, maternal $PM_{2.5}$ exposure, maternal age, maternal education level, maternal occupation status, unemployment rate, and coal power generation. The other variables for descriptive statistics were duration of cohabitation, infant birth weight, infant gender, and maternal marital status. The variables including coal power generation, $PM_{2.5}$ exposure, and unemployment rate were measured at the ecological level.

In Chapter 2, I detail the theoretical and conceptual framework of this study, as well as links between the key variables, based on the current literature. Specifically, I describe the interactions of this study and the literature to ensure new insights that may lead to social change.

Chapter 2: Literature Review

Introduction

The incidence of preterm birth in Korea has increased 44% in less than a decade (KOSIS, 2016). As of 2015, annual exposure to $PM_{2.5}$ exceeded the WHO safe level in many parts of Korea (KME, 2016). Despite increasing evidence of an association between maternal exposure to $PM_{2.5}$ and preterm birth, there exists a dearth of information on the preterm birth risk associated with maternal exposure to the current level of $PM_{2.5}$ in Korea (Hao et al., 2016 & Laurent et al., 2016). I conducted this research to examine the preterm birth risk associated with maternal exposure to $PM_{2.5}$ in Korea's air, including the modified preterm birth risk due to the potential moderation, and mediation by Korea's air quality. I also investigated the variables influencing differences in maternal $PM_{2.5}$ exposure for the purpose of reducing the disparities in $PM_{2.5}$ exposure.

In my review of the current literature, I found original studies designed to test methods to improve the accuracy of individual exposure estimation to $PM_{2.5}$ in order to predict preterm birth risk (Ha et al., 2014; Hao et al., 2016; Hyder et al., 2014; Laurent et al., 2016). I also found that, in keeping with SET, some researchers criticized the traditional approach of investigating fragmented risk factors by focusing on medical heterogeneity in preterm birth because traditional studies have not explicitly addressed potential interactions between multi-level factors (Glouberman, & Millar, 2003; Vandenbroucke, Broadbent, & Pearce, 2016). The potential modification of preterm birth risk by $PM_{2.5}$ exposure remains to be studied. Research regarding what sociodemographic determinants are complementary component causes of maternal $PM_{2.5}$ exposure to generate preterm birth is needed in order to implement practical preventive measures.

This chapter includes descriptions of my literature search strategy, selected theoretical and conceptual frameworks, and the linkage between the study variables from the existing relevant studies.

Literature Search Strategy

The literature review consisted of three elements: research on preterm birth, research on the link between $PM_{2.5}$ and preterm birth, and SET. My aim was to identify the current tendencies of research on preterm birth. The results of the literature review showed what factors link to preterm birth and what remains to be studied to enrich understanding of causes of preterm birth. I also undertook this literature review to synthesize existing evidence of the links between (a) risk factors for preterm birth and the SET, (b) sociodemographic status and preterm birth, (c) $PM_{2.5}$ and preterm birth, and (d) sociodemographic status and $PM_{2.5}$ exposure. The results of the literature review helped to develop hypotheses and recommendations for this research.

I performed the literature search using the MEDLINE and Google Scholar databases as summarized in Table 1. Key search terms included: *fine particulate matter, gestational age, Korean, preterm birth, risk factor, and social ecological theory*. I searched them separately and in combination using Boolean operators as shown in Table 1. I also manually searched the reference lists of selected studies.

Table 1

Search Strategy

Databases	Search modes	Publication type	Publication year	Peer reviewed selected	Search terms
Medline	Booleans	Journal article	2016-2017	Yes	(preterm birth OR gestational age) AND risk factor
Google Scholar	Booleans	All	Any time	No	(preterm birth OR gestational age) AND social ecological theory
Medline	Booleans	Journal article	2012-2017	Yes	(preterm birth OR gestational age) AND fine particulate matter
Google Scholar	Booleans	All	Any time	No	fine particulate matter AND social ecological theory
Medline	Booleans	Journal article	2012-2017	Yes	((preterm birth OR gestational age) AND fine particulate matter AND Korean
Google Scholar	Booleans	All	Any time	No	((preterm birth OR gestational age) AND fine particulate matter AND Korean

Theoretical Foundation

Social Ecological Theory (SET)

SET provided the theoretical framework I used to construct research questions and study variables. SET was developed in the 1980s and has an origin in social psychology (McLeroy, Bibeau, Steckler, & Glanz, 1988). SET covers multi-level measures of data from individual to group/area level data (McLeroy et al., 1988). It posits that variables involving disease and health tend to group corresponding to one of the multi-levels, which consist of individual, community, organizational, and policy levels (McLeroy et al., 1988). This theory also posits that complex multi-level interactions within and across layers exist and generate differences in disease and health (McLeroy et al., 1998). SET emphasizes the need for evaluation of multiple levels of influence and facilitates analysis of the impact of multilevel factors simultaneously (Sallis et al., 2006).

SET has been widely used to determine the contributing factors and their interactions associated with the occurrence of a disease (Mendez et al., 216; Sallis et al., 2006; Vos et al., 2014). Researchers have previously used it to determine group/area level risk factors associated with adverse birth outcomes, including income inequality, unemployment rate, and crime rates (Brumberg, & Shah, 2015; Mendez et al., 216; Vos et al., 2014). Previous researchers have also examined the influence of built environments in residences on the development of a disease or maintenance of health. For example, preventive strategies targeting obesity-related preterm birth that promoted physical activity were successful when addressing the multi-levels of influence based on SET (VCAA, 2010). SET explains how people's behaviors could be changed to have good health, most commonly by things like making physical activity more a normal and safe part of individuals lives, or by removing nuisances like pollution, trash, and noise (VCAA, 2010). Studies based on SET could cover individual and ecological level factors simultaneously and include proposals for macro-level actions to remove environmental hazards beyond individual control (Brumberg, & Shah, 2015; Mendez et al., 216; Vos et al., 2014).

I selected SET for this study because it is a useful approach to understand disparities in disease and health, including the difference in exposure. This theory describes how the difference in exposure results from complex multi-level interactions at the individual, community, organizational, and policy levels (Erickson at al., 2016; Parker, Baldwin, Israel, & Salinas, 2004). SET provides a framework for (a) identifying the variables that relate to the difference in exposure, and (b) understanding how their interactions influence disease and health eventually. I selected SET for this study to help identify promising levels of intervention that can increase the efficiency of public health spending and contribute to reducing gaps in health disparity. Specifically, I used it to identify the targets of change at each of the ecological levels for the purpose of reducing group differences in $PM_{2.5}$ exposure followed by disparities in health status.

Links between SET and preterm birth. Variables investigated in recent studies of preterm birth have included: maternal age, height, vitamin D concentration, serum zinc concentrations, Fontan operation, shigellosis, autoimmune hepatitis, sideropenic anemia, pre-pregnancy obesity, gestational weight gain, preterm birth history, periodontitis, oligohydramnios, presence of Nugent's intermediate vaginal flora, gestational diabetes mellitus, maternal education, marital status, employment status, occupation, ethnicity, income inequality, domestic violence, immigrant/refugee status, teen pregnancy, inadequate prenatal care, airborne particulate matter, disaster, and organochlorine pesticides (Averett, & Fletcher, 2016; El Rafei et al., 2016; DeFranco et al., 2016; Derraik et al., 2016; Huang et al., 2016; Leal et al., 2016; Maslow et al., 2016; Miliku et al., 2016; Miller et al., 2016; Oftedal et al., 2016; Perveen & Soomro, 2016; Prunet et al., 2017; Pundi et al., 2016; Parisot et al., 2016; Stokkeland et al., 2016; Tyagi et al., 2016; Tellapragada et al., 2016; Wallace et al., 2016; Wang et al., 2016). The variables can be stratified based on SET as Table 2 shows.

Table 2

Classification of Risk Factors for Preterm Birth Based on SET

Social ecological level	*N*	%	Risk factors
Intrapersonal	19	68%	Age
			Height
			Vitamin D concentration
			Serum zinc concentrations
			Fontan operation
			Shigellosis
			Autoimmune hepatitis
			Sideropenic anemia
			Prepregnancy obesity
			Gestational weight gain
			Preterm birth history
			Periodontitis
			Oligohydramnios
			Nugent's intermediate vaginal flora
			Gestational diabetes mellitus
			Maternal education
			Maternal employment status
			Maternal occupation
			Ethnicity
Community	3	11%	Particulate matter
			Disaster
			Organochlorine pesticides
Organizational	4	14%	Domestic violence
			Income inequality
			Inadequate prenatal care
			Immigrants or refugees
Policy	2	7%	Air quality standards
			Immigration policies and systems

Table 2 shows that the variables associated with preterm birth can be located on one of the four social ecological levels. This table also indicates that there is little research on the association of risk factors outside the individual level with preterm birth. The other issue that has not been adequately addressed in studies of preterm birth is that many did not examine the moderating and mediating effects that may exist. Based on the SET, McLeroy et al. (1988) criticized the approach that focused on biological or psychological etiologies while neglecting the role of social and physical environmental factors on disease and health. They noted that this individual focus often aligns with a victim-blaming ideology.

This criticism regarding neglecting the role of social and physical environmental factors in disease and health is still valid to studies of preterm birth risks. Given such, further research is warranted to address the role of social and physical environmental factors on disease and health. SET relates to the present study as it pinpoints where to study when examining the influence of physical environment and $PM_{2.5}$ exposure on preterm birth. Researchers guided by SET have suggested a shift from an emphasis on fragmented risk factors for a disease to an emphasis on cross-level interactions with a broader multilevel view (Glouberman, & Millar, 2003).

Use of the SET in this research. I used SET to construct research questions and the variables for this study. The study variables were located variously on the individual, community, organization, and policy layers. SET highlights interactions that impact exposures and outcomes between individual characteristics and the factors located on upper layers such as economic, social, environmental, and political resources. SET emphasizes the significance of social and policy actions leading to changes in community, organization, and policy. I designed the research question guided by SET to explore whether the cross-level interactions of $PM_{2.5}$ exposure influence gestational age.

Systems Theory

Systems theory was developed by Ludwig von Bertalanffy and has been adopted in various fields including biology, physics, psychology, anthropology, and management (Peter, 2014). Systems theory provides a framework that describes the components of systems and the interactions between these components (Trochim et al., 2006). Systems theory posits that the outcome of the system is the result of aggregated effects of the components of the system (Peter, 2014). Even systems with exactly the same outcome cannot be expected to consist of the same components. This theory holds that a system has an ability to respond to internal and external change (Peter, 2014). Even the same system can be expected to yield different results according to its response to changes.

I used systems theory to explain that sociodemographic effects on preterm birth would be changed by the $PM_{2.5}$ exposure levels (Trochim et al., 2006). The association between sociodemographic factors and gestational age may differ between $PM_{2.5}$ exposure levels. This theory warranted my examination of the modification of preterm birth risk by $PM_{2.5}$ exposure. I designed the research question associated with this theory to explore whether $PM_{2.5}$ exposure mediates the independent sociodemographic effects on preterm birth.

Conceptual Framework

Social Ecological Model (SEM)

I used SEM to derive the practical concept of designing and evaluating meteorological variables in this research. This concept has helped to improve the validity and reliability of estimating variables by suggesting how to measure the meteorological variables. SEM was useful in designing and evaluating maternal exposure to $PM_{2.5}$ in this research on the association between ecologic level $PM_{2.5}$ exposure and preterm birth.

Measurement error of maternal $PM_{2.5}$ exposure. Researchers have used various spatial modeling to estimate individual exposure to $PM_{2.5}$ based on databases of various air pollution monitoring systems due to the spatial distance between the air quality monitoring sites and the individual exposure domain of the study participants (Ha et al., 2014; Hyder et al., 2014; Laurent et al., 2016). In addition to the spatial adjustment, accurate measurement of $PM_{2.5}$ exposure at individual level requires further information including physiological characteristics lung capacity and a profile of behavior like indoor and outdoor activities (Ha et al., 2014; Hyder et al., 2014; Laurent et al., 2016). The estimates of individual-level $PM_{2.5}$ exposure based on the closest monitor data without details on a profile of individual behaviors and physiological characteristics could not guarantee a high level of reliability for exposure measurement. Measurement error and misclassification could not be excluded from the estimates of $PM_{2.5}$ exposure at individual level since no alternative could substitute directly measured individual $PM_{2.5}$ exposure with no measurement error.

Links between SEM and $PM_{2.5}$ exposure. Often, exposure to polluted air cannot be avoided through changes in individual behavior (Trasande et al., 2016). The difference in exposure to ambient pollutants and meteorological conditions can be better explained by organizational, community, and policy determinants, including air quality standards and regulations on emissions of atmospheric hazards like motor vehicle and coal-fired power plants rather than individual characteristics (Trasande et al., 2016).

In this environmental research field, meteorological data, including $PM_{2.5}$, atmospheric pressure, temperature, season and humidity tend to be grouped into community layers based on SEM (Erickson et al., 2016; Giorgis-Allemand et al., 2017; Hao, Strosnider, Balluz, & Qualters, 2016; He et al., 2016; Lavigne et al., 2016; McLeroy et al., 1988; Strand, Barnett, & Tong, 2011). A number of researchers measured meteorological variables at area-level to examine its possible associations with preterm birth (Erickson et al., 2016; Giorgis-Allemand et al., 2017; Hao et al., 2016; He et al., 2016; Lavigne et al., 2016; Strand, Barnett, & Tong, 2011).

One of the benefits of using SEM in this investigation was to reduce the probability of measurement errors in $PM_{2.5}$ exposure compared to individual-level $PM_{2.5}$ exposure estimates by measuring $PM_{2.5}$ exposure at the community level. The $PM_{2.5}$ exposure estimates based on the closest monitor data were more reliable as community-level measures than individual-level measures (Erickson et al., 2016; Giorgis-Allemand et al., 2017; Hao, Strosnider, Balluz, & Qualters, 2016; He et al., 2016). Another benefit of SEM in this work was to assess group differences in exposure to atmospheric pollutants and to drive macro-level initiatives to reduce social and political disparities in atmospheric pollutants.

Rothman's Sufficient Component Cause Model (RSCCM)

I used RSCCM to examine moderation and mediation by $PM_{2.5}$ in the relationship between maternal sociodemographic factors and gestational age. Component causes refer to a cluster of risk factors having the joint effect for a sufficient cause of a disease (Szklo, & Nieto, 2014). Sufficient cause means a set of minimum component causes that can produce disease (Szklo, & Nieto, 2014). Rothman and Greenland (2005) described that more than one causal mechanisms could exist for a disease and complete understanding of causal mechanisms in their entirety is impossible. The advantage of applying RSCCM in developing a preventive strategy is that disease prevention is possible by eliminating one of the single component causes from the established sufficient causal pie (Szklo, & Nieto, 2014). Without a complete understanding of possible causal mechanisms though, practical preventive strategies can be developed using the Rothman's causality model. This model was useful in this research, given that few causal mechanisms for preterm birth were discovered and the understanding of the mechanisms underlying preterm birth was still incomplete (El Rafei et al., 2016; Oftedal et al., 2016; Perveen & Soomro, 2016; Tellapragada et al., 2016; Wallace et al., 2016).

This model allows researcher to compose a causal pie and transform the causal pie into a deductive test of hypothesis (Rothman, & Greenland, 2005). This approach is more pragmatic and flexible because it allows for the formation of causal pies that better align with current tendencies of significant health risks, including emerging health threats and newly identified risk factors (Rothman, & Greenland, 2005). The outcome of interest in a causal pie depends on the composition of component causes and their sufficient levels (Rothman, & Greenland, 2005). Another benefit of RSCCM was to warrant my examination of the unknown component cause for preterm birth, including maternal $PM_{2.5}$ exposure (see Rothman, & Greenland, 2005). RSCCM explains that the influence of $PM_{2.5}$ exposure on preterm birth depends on the presence or absence of its complementary component causes (Szklo, & Nieto, 2014).

Use of RSCCM for this study. I selected RSCCM for this study to examine the unknown component cause for preterm birth, like maternal $PM_{2.5}$ exposure levels, and to identify whether $PM_{2.5}$ would be a complementary component causes for preterm birth of maternal sociodemographic factors. I used the RSCCM to create the causal pie in Figure 5 to explore whether $PM_{2.5}$ exposure moderates and mediates the sociodemographic effects on preterm birth. I then transformed the causal pie into hypotheses: The interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, maternal education level, maternal occupation status, would have no statistically significant effect on gestational age; the interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, maternal education level, maternal occupation status, would have a statistically significant effect on gestational age; sociodemographic factors, such as maternal age, maternal education level, maternal occupation status, would have no statistically significant effects on gestational age through $PM_{2.5}$ exposure; and sociodemographic factors, such as maternal age, maternal education level, maternal occupation status, would have statistically significant effects on gestational age through $PM_{2.5}$ exposure.

In this study's causal pie, Figure 5, I included the sociodemographic variables as the component causes, and regarded the variable having the strongest correlation with gestational age as a necessary component caus. I also treated $PM_{2.5}$ exposure as an unknown component cause in the model. As guided by the sufficient cause model, I statistically tested the impact of $PM_{2.5}$ exposure on gestational age. The results of the statistical analysis may help to find optimal prevention strategies for women vulnerable to preterm birth by showing the minimum set of conditions that can lead to preterm births in Korea.

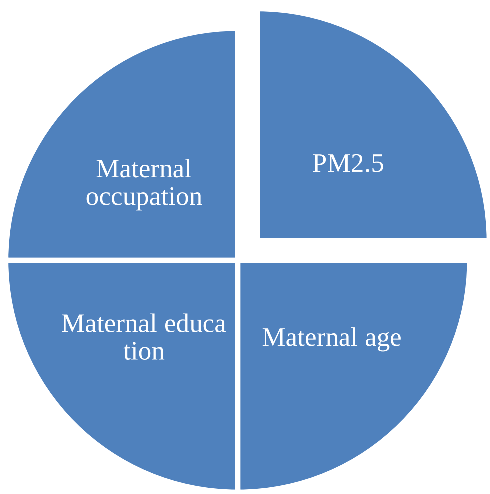

Figure 5. Causal pie for preterm birth to test unknown cause $PM_{2.5}$ exposure.

Key Variables and Concepts

Sociodemographic Status and Preterm Birth

Previous researchers have consistently concluded that sociodemographic status is a significant predictor of preterm birth (Oftedal et al., 2016 & Savitri et al., 2016). In recent studies of risk estimate of preterm birth, sociodemographic factors, associated with preterm birth have included: maternal age, marital status, ethnicity, educational attainment, parity, race, and household income (Laurent et al., 2016; Lee et al., 2013; Ha et al., 2014; Hao et al., 2016; Stieb et al., 2016; Symanski et al., 2014). In my literature review, I confirmed statistically significant correlations between maternal age, education, occupation, or marital status and preterm birth in the Korean population (Ha et al., 2004; Kim et al., 2007; Leem et al., 2006; Suh et al., 2009).

In recent studies of the relationship between maternal age and gestational age, the researchers found a U-shaped association and reported that the risk for women aged ≥ 40 years increased by 24% over the age group below 40 years (Kenny et al., 2013 & Restrepo-Méndez et al., 2015). Low maternal education was associated with increased risk of preterm birth in Norway (Oftedal et al., 2016). In several studies, the researchers found that maternal occupation was a predictor of preterm birth (Casas et al., 2015; Prunet et al., 2017; von Ehrenstein et al., 2014). Maternal employment status during pregnancy influenced preterm birth in these studies, although the results regarding the employment status associated with increased risk of preterm birth were inconsistent across studies (Casas et al., 2015; von Ehrenstein et al., 2014). Food industry workers showed an increased risk of preterm birth in a study, while professionals, managers, and engineers had lowest rates of preterm birth from the other study (Casas et al., 2015; Prunet et al., 2017). The inconsistency in prediction of preterm birth based on sociodemographic characteristics across the recent relevant studies may result from the diverse methods of measuring sociodemographic variables, constructing potential confounders, and unmeasured moderation and mediation effects (Casas et al., 2015; Prunet et al., 2017; von Ehrenstein et al., 2014).

Through the literature review, I confirmed the link between sociodemographic status and preterm birth. Given such, this research was warranted to examine the potential moderation or mediation by a covariate in the link between each sociodemographic variable and preterm birth to understand the possible paths among the variables and develop prevention strategies addressing these pathways.

$PM_{2.5}$ and Preterm Birth

Evidence of the link between $PM_{2.5}$ exposure and preterm birth has been increasing (Hao et al., 2016; Laurent et al., 2016; Trasande et al., 2013). In the existing relevant studies, most researchers found a statistically significant association between maternal exposure to $PM_{2.5}$ and preterm birth (Hao et al., 2016; Laurent et al., 2016; Stieb et al., 2016; Trasande et al., 2013). The estimates of preterm birth risk based on maternal $PM_{2.5}$ exposure varied with the risk or odds ratios ranging from 1.011-1.133 or 0.96-0.97 (Hao et al., 2016 & Trasande et al., 2013). The deleterious effects of $PM_{2.5}$ on preterm birth prevailed at high levels of $PM_{2.5}$ exposure (Hao et al., 2016; Laurent et al., 2016; Ha et al., 2014; Stieb et al., 2016; Trasande et al., 2013). The available evidence in the literature was sufficient to justify further research on preterm birth that examines the moderation and mediation by the Korea's air quality.

$PM_{2.5}$ Exposure and Sociodemographic Status

The association between sociodemographic status and $PM_{2.5}$ exposure levels has not been explicitly evidenced because there are only a few studies regarding the link between them in the literature. From the existing relevant studies, Stieb et al. (2016) found statistically significant differences in $PM_{2.5}$ exposure according to education and income. Higher education or lower income were associated with increased exposure to $PM_{2.5}$ (Stieb et al., 2016). Rotko et al. (2000) observed inverse proportional associations between education, occupational status, and age and the likelihood of exposure to $PM_{2.5}$. In their study, individuals with a lower occupational status, education, and younger age were more likely to have $PM_{2.5}$ exposure than those with upper occupational status, higher education, and older age (Rotko et al., 2000). Because of little research regarding the relationship between sociodemographic variables and $PM_{2.5}$ exposure, in this study I handled it as having no evidence of a relationship between $PM_{2.5}$ exposure and sociodemographic status. In this study, I decided whether to perform a mediation analysis on these variables based on the analysis results of this study population.

Research in Korea

$PM_{2.5}$ and preterm birth. There were no studies of preterm birth risk associated with maternal exposure to $PM_{2.5}$ in Korea. In all studies of Korean women, the researchers evaluated the aggregated level data at PM_{10} with no separation between PM_{10} and $PM_{2.5}$ (Ha et al., 2004; Kim et al., 2007; Leem et al., 2006; Suh et al., 2009). In all relevant studies in Korea, the researchers used data from 1998 to 2004 (Ha et al., 2004; Kim et al., 2007; Leem et al., 2006; Suh et al., 2009). The results found from the decade ago data could not provide valid evidence to develop current strategies of preterm birth prevention because of changes in the variables related to preterm birth including the air pollution, population demographics, social and economic factors (see Leal et al., 2016).

The variables relate to both preterm birth and $PM_{2.5}$ exposure differed from those of a decade ago according to the rapid social changes and economic growth in Korea (OECD Better Life Index, 2015).

Sociodemographic status, $PM_{2.5}$ exposure, and preterm birth. Researchers found factors that were bi-variately associated with particulate matter and preterm birth as follows: maternal age, education level, occupation, marital status, infant gender, birth order, parity, birth season, birth month, and maternal history of stillbirth (Ha et al., 2004; Kim et al., 2007; Leem et al., 2006; Suh et al., 2009). In the studies included in my literature review, the researchers consistently controlled for maternal age, infant gender, and maternal education (Ha et al., 2004; Kim et al., 2007; Leem et al., 2006; Suh et al., 2009). These factors of maternal age, infant gender, and maternal education were correlated with preterm birth and PM_{10} exposure in Korea and should be considered as covariates in the future study of air pollution's effect on preterm birth in Korea. The association between maternal age, education, and occupation and preterm birth was fairly consistent across studies. Although previous researchers have found links between several sociodemographic determinants and preterm birth, changes in sociodemographic characteristics over the past decade warranted my research based on data from 2015 (Ha et al., 2004; Kim et al., 2007; Leem et al., 2006; OECD Better Life Index, 2015; Suh et al., 2009).

Methodologic Considerations

Methodologies in the literature. In my literature review, I found that prevailed methodologies were the quantitative deductive approach, positivism philosophical foundations, and correlational or quasi-experimental design. These findings have helped to determine the methodology that can increase the comparability of existing and future research on preterm birth. Quantitative, positivism and observational methods were reliable and valid when studying the factors associated with the likelihood of preterm birth (Burkholder, Cox, & Crawford, 2016). In this study, I used these methods to obtain meaningful and comparable new insights in alignment with the current literature.

Acute effects of $PM_{2.5}$ on preterm birth. Symanski et al. (2014) found a significant association between $PM_{2.5}$ exposure and mild, moderate, and severe preterm birth during the last relevant 4 weeks of pregnancy, but not during the rest of the pregnancy. These findings could provide evidence of short lag times of $PM_{2.5}$ exposure to its effect on preterm birth which would be less than 4 weeks. In another study, the researchers observed lag times of 0 to 3 days between maternal $PM_{2.5}$ exposure and the effect on preterm birth, which provided evidence of the acute or short-term effect of $PM_{2.5}$ on preterm birth (Lima, Nascimento, Medeiros, & Santos, 2014). In the study, the exposure period was the last 4 weeks of pregnancy and was not affected by differences in the length of pregnancy between individuals.

Summary and Conclusions

Preterm birth risk factors can be located on one of the four social ecological levels. SET describes how cross-level interactions of $PM_{2.5}$ exposure affect gestational age. SET also indicates that there is little research on the risk factors outside the individual level for preterm birth, including social, environmental, and policy factors. In my literature review, I confirmed the link between sociodemographic status and preterm birth. When studying the association between sociodemographic factors and preterm birth, researchers considered the covariates related to both independent and dependent variables only as potential confounders. In the recent relevant studies, the variation in the estimates of preterm birth based on sociodemographic characteristics might result from the unmeasured moderation and mediation effects.

SEM presented a practical concept in this research to design and evaluate maternal exposure to $PM_{2.5}$. RSCCM has helped examine the unknown component cause for preterm birth, including maternal $PM_{2.5}$ exposure, and determine whether $PM_{2.5}$ is a complementary component cause of maternal sociodemographic factors for preterm birth.

This study was a quantitative, observational, partially ecologic, retrospective cohort study. In this research, I measured maternal $PM_{2.5}$ exposure as a population-level attribute at the community level. My aim was to explore whether the cross-level interactions of PM2.5 exposure influence gestational age. I also explored the possibility of moderation and mediation in the link between each sociodemographic variable and gestational age to understand the possible pathways between variables. The findings of this work could be useful in developing prevention strategies for preterm birth.

In Chapter 3, I detail the research design, study population, sampling, and methodology of this study, as well as data analysis plan. I also describe ethical concerns related to this study.

Chapter 3: Research Method

Introduction

The primary purpose of this work was to investigate the effect of maternal $PM_{2.5}$ exposure on preterm birth, including the potential moderation and mediation by $PM_{2.5}$ exposure in the relationships between sociodemographic factors and gestational age. I designed this study to determine whether $PM_{2.5}$ is a complementary component cause of maternal sociodemographic factors for preterm birth in Korea. In this investigation, I also examined which variables could explain differences in $PM_{2.5}$ exposure.

This chapter consists of four sections. In the first part of this chapter, I describe the research design and rationale. In the second section, I give details on the target population, sample size, instrumentation, operationalization of study variables, and data analysis plan. A discussion of the threats to validity is followed by the methodology section. Lastly, I address ethical concern related to this study and close with a summary.

Research Design and Rationale

The dependent variables of interest in this study were preterm birth and gestational age. The independent variables were based on SET, including (a) individual factors such as maternal age, maternal education level, and maternal occupation status; (b) community factor such as maternal $PM_{2.5}$ exposure; (c) organizational factor such as unemployment rate; and (d) policy factor such as coal power generation. In investigations for moderation and mediation, I regarded $PM_{2.5}$ exposure as a potential moderator and mediator in the relationship between maternal sociodemographic factors of the independent variables and gestational age of the dependent variable.

This was a quantitative, observational, partially ecologic, retrospective cohort study. The quantitative approach is useful for identifying what variables could predict the occurrence of preterm birth (Burkholder, Cox, & Crawford, 2016). Both the increase in the incidence of preterm birth and the high exposure levels to $PM_{2.5}$ in many areas of Korea are the research problem I addressed in this study. This problem aligns with a call to action for eliminating preterm birth risks. A quantitative approach was best to determine the current best predictors of preterm birth. It was the appropriate approach to test the theory of the cross-level interactions to change the occurrence of preterm birth. This approach helped me identify pathways that influence preterm birth and propose preventive strategies on preterm birth.

I used correlational observation, which is suitable to answer research questions involving prevention or causal paths (Aschengrau, & Seage, 2014). The theories of the potential direct and indirect effects of maternal $PM_{2.5}$ exposure on preterm birth that I tested in this study were grounded in the literature as discussed in the previous chapter.

This was a partially ecologic study in which both individual level and ecological level variables were involved. I retrospectively applied the ecological level assessment of exposure in this study. The variables measured at an ecological level were as follows: Maternal $PM_{2.5}$ exposure, unemployment rate, and coal power generation. The variables on upper levels of the SEM are population-level factors such as individual population members' attributes and environmental measures (Aschengrau, & Seage, 2014). These factors could not be accounted for by measures of individual attributes. They could be measured as a population-level attribute (Erickson et al., 2016). The variables on upper layers of the SEM represent characteristics of the area to which each belongs and are measured as collective attributes (Aschengrau, & Seage, 2014). The variables I measured at an individual-level in this study were socio-demographic factors composed of maternal age, maternal education level, and maternal occupation status. This design could explicitly address potential cross-level interactions influencing preterm birth. It could also allow for assessment of social and political determinant of $PM_{2.5}$ exposure, leading to social and political actions to reduce disparities both in $PM_{2.5}$ exposure and preterm birth.

This study had a retrospective cohort design. The retrospective cohort approach is useful to examine the association between the incidence of the outcome of interest and a suspected exposure (Szklo, & Nieto, 2014). This approach allows researchers to assess historical exposures from existing data. Retrospective cohort studies can be completed expeditiously by saving time for follow-up (Aschengrau, & Seage, 2014). This design is beneficial for student researchers because it requires fewer resources and less cost to study than prospective design.

Methodology

Population

The target population for this study was defined as Korean pregnant women who resided in Korea and gave singleton live birth in 2015. A live birth refers to a product of conception which is thoroughly extracted from the mother, regardless of the completed weeks of gestation, and breathes, or shows other vital signs such as pulse or heartbeat (Statistics Korea [KOSTAT], 2016). Such a product is considered live born and is registered and counted as a live birth in VSKLB (KOSTAT, 2016). According to the VSKLB for the year of 2015, a total of 421,988 singleton live births were filed, which is equal to the target population size of this study.

Sampling and Sampling Procedures

The study population was selected from the VSKLB data file by a two-step approach consisting of stratified single-stage cluster sampling, followed by screening using inclusion criteria. At the first step, I selected stratified and cluster sampling of two basic probability sampling techniques in conjunction with one another. This stratified single-stage cluster sampling was conducted to form a pre-sample population to proceed with screening. For the stratified sampling, each stratum was defined as one of the mutually exclusive 17 top-tier administrative districts. This method enabled the sample to embed the various subgroups in the sample (see Teddlie, & Yu, 2007). One or more valid $PM_{2.5}$ monitoring stations were installed in only the 12 top-tier administrative districts (KME, 2016). Not all top-tier administrative districts proceeded with additional sampling, of which 12 were selected for cluster sampling, mitigating the possibilities of misclassification by measurement error in $PM_{2.5}$ exposure in this study. For the cluster sampling, I defined each cluster as one of the mutually exclusive 252 sub-tier administrative areas, called SI in the major cities, GU in urban areas, or GUN in rural areas. One cluster with the largest number of births was selected for each of the 12 strata using a probability proportional to size sampling method, and I included all births in the chosen 12 clusters in this analysis. This method allows for a greater chance of sampling large clusters with high population density (Teddlie, & Yu, 2007). There was no overlap of subpopulations between clusters because each population was divided into one of the clusters according to their registered residential address.

The final study sample was drawn through screening the pre-sampled population with eligibility criteria specified as follow.

Inclusion criteria. Pregnant women who meet all of the following criteria were selected in this study.

1. The women were aged 20 to 44 years.
2. The women provided information on the sub-tier administrative district where they resided.
3. Delivery occurred at a gestational age of 23 to 40 weeks.
4. The women had nulliparous and singleton pregnancies.
5. The women gave birth to infants weighing between 2,500 and 5,000g.
6. The women were of Korean ethnicity.
7. The women gave birth between February 1, 2015 and December 31, 2015.

I set Criteria 1,3 through 6 to restrict and mitigate potential confounding. Criteria 2 was set to comply with regulations prescribed by the Personal Information Protection Act and the guideline of Health Insurance Portability and Accountability Act (HIPAA). I set Criteria 6 because Koreans are ethnically homogeneous. Approximately 96% of the people in Korea was composed of Korean ethnic group (KOSIS, 2015). This criterion hardly limited the representativeness of the study population. I also set Criteria 7 to ensure a high level of the content validity due to the amendment of birth report form as of February 1, 2015.

The minimum number of subjects for this analysis was 1,084. The sample size was calculated by a priori power analysis using G*Power 3.1 with the alpha value of 0.05, power of 0.95, and an assumed effect size of 0.20. The analysis was expected to have a 95% power to detect a significant difference between groups as the effect size was a small value of 0.20 and the alpha value was 0.05.

Archived Data

I used a secondary dataset after linking files of the VSKLB, KMED, EAPSS, and SEPK.

VSKLB. The VSKLB included the variables of interest for this study as follow: date of birth, completed weeks of gestation at delivery, gender and birth weight of the infant, area level maternal residential address, maternal age, maternal occupation, maternal education level, maternal marital status, ethnicity, parity, and duration of cohabitation.

Every birth in Korea must be reported to the local family registration office at the domicile by the Family Relationships Act Art.45 (R.O. Korea) (National Law Information Center [NLIC], 2016). The father or mother of each newborn is duty bound to report the birth, if legally married (NLIC, 2016). If a child is born to a couple with no legal marriage, the mother is obligated to report the birth (NLIC, 2016). The birth registration must be completed within 30 days of birth (NLIC, 2016). The registration of birth could be processed through the online reporting system or a visit to the local family registration office at the domicile (NLIC, 2016). The reporter is asked to fill out the birth report form and attach a birth certificate (NLIC, 2016).

Upon receipt of the report of birth, the registrar in the local family registration office verifies the information with the birth certificate or other available sources (NLIC, 2016). The registrar raises queries or requests additional evidence as needed to verify the reported information, and after confirmation registers the birth into the local family registration system (KOSTAT, 2016). The vital events registered in each local family registration system are then incorporated into the national real-time web-based vital statistics system (KOSTAT, 2016).

Access to the VSKLB required permission from the KOSTAT. The KOSTAT is the national statistical agency and acts as the national data hub (KOSTAT, 2016). The KOSTAT operates the real-time web-based vital statistics system and produces the national-level vital statistics (KOSTAT, 2016). The KOSTAT runs a website dedicated to the national data repository, MicroData Integrated Service (MDIS; KOSTAT, 2016). The process of obtaining access to the data set involved submitting a remote access service application, including a research synopsis, and confidentiality agreement (KOSTAT, 2016). Upon getting approval, I could only access the dataset remotely and could not download the data set to my personal computer (KOSTAT, 2016). I could analyze the dataset on the virtual server while remotely accessing the dataset (KOSTAT, 2016). I could download only the outputs of statistical analyses approved by the KOSTAT (KOSTAT, 2016).

KMED. The KMED contained the factors necessary for the assessment of maternal exposure to $PM_{2.5}$ in this study like daily mean concentrations of $PM_{2.5}$ in 2015. In this study, I calculated maternal exposure to $PM_{2.5}$ based on the averaged $PM_{2.5}$

concentrations at the maternal residential area for the last 4 weeks of pregnancy. I linked the KMED to the VSKLB. The linkage of the data files provided the factors for the assessment of maternal $PM_{2.5}$ exposure during the last 4 weeks of pregnancy as follows: date of birth, area level maternal residential address, and daily mean concentrations of $PM_{2.5}$ measured at 100 valid monitoring stations in Korea.

The Korea Ministry of Environment (KME) launched $PM_{2.5}$ monitoring and alert system in 2015 (KME, 2015). The $PM_{2.5}$ concentrations in Korea are available only for the period from January 1, 2015, to present. The KME operates a website dedicated to air quality in Korea in order to respond to the increasing public concern about air pollution and interest in the good quality of the environment (Air Korea.or.kr, 2013). The website offers the nationwide real-time air quality information also disseminates warning and forecasting of $PM_{2.5}$ levels (Air Korea.or.kr, 2013). The database of $PM_{2.5}$ concentration exists on the website (Air Korea.or.kr, 2013). The website is the public domain, and any users could download the desired dataset directly from the website to the personal computers without any permissions to gain access to the data.

EAPSS. The EAPSS contained the variable of interest for this study such as unemployment rates of each month in the 17 top-tier administrative divisions of Korea. In this study, I calculated the averaged unemployment rate of the maternal residential area across the last 4 weeks of pregnancy by the linkage of the data files between the EPASS and VSKLB. The linkage of the data files provided parameters needed to compute an averaged unemployment rate of the maternal residential area across the last 4 weeks of pregnancy as follows: date of birth, area level maternal residential address, and the unemployment rates in the months to which the last 4 weeks of gestation belong.

The EAPSS provided essential data for analyzing and establishing employment policies in Korea (KOSIS, 2016). The sample frame of the EAPSS employed the same as the population census and updated annually through replacing the households from old to new (KOSIS, 2016). A stratified two-stage self-weighting design by geographic region was used (Hwang, Kim, & Kim, 2017). The applied sample fraction was 0.2% of the total population in each of the city/province (KOSIS, 2016). The first sampling unit was the wide-area enumeration district, and the second stage sampling unit was the household/member of the selected enumeration district (KOSIS, 2016). In principle, 20 households/members were sampled in each sampled enumeration district and surveyed (Hwang et al., 2017). A total of 33,000 households/members from 1,647 sample enumeration districts were recruited per month (KOSIS, 2016). This survey included respondents who belonged to the sampled household and resided in Korea and were 15 years or older on the 15th day of the month at the response (KOSIS, 2016). Soldiers, social service personnel, duty police officers, and prison inmates were excluded from this survey (KOSIS, 2016).

The data collection tool employed for the EAPSS is the EAP survey (KOSIS, 2016). The EAPSS database was located in the MDIS domain. No permission was needed to gain access to the EAPSS dataset (KOSIS, 2016). For this investigation, I downloaded the 2015 EAPSS data set.

SEPK. The SEPK contained the variable of interest for this study such as the amount of electricity generated from coal power plants of each month and locations of the coal plants. In this study, I calculated the coal power generation for the last 4 weeks of pregnancy based on the date of birth and monthly amount of electricity generated from coal power plants. The linkage between the SEPK and VSKLB data files provided the factors for the measurement of coal power generation during the last 4 weeks of pregnancy as follows: date of birth and monthly amount of electricity generation from coal power plants.

In this study, I calculated the proximity to coal power plants as the weighted average distance according to the amount of coal-powered electricity generation of each plant from the monitoring station in maternity residential area to all coal-fired power plants. The linkage among the SEPK, KMED, and VSKLB data files provided the factors for the measurement of the proximity to coal power plants as follows: area level maternal residential address, locations of $PM_{2.5}$ monitoring stations, monthly amount of electricity generation from coal power plants, and places of coal power plants.

The SEPK provides 30 items related to the generation, facility, purchase, and sales of electric power (KEPCO, 2016). The SEPK helps to have an overall understanding of the current status of Korean energy industry including the generation and development of electric power resources (KEPCO, 2016). The SEPK provides evidence of decisions on energy policy in Korea and data sources for research on developing the efficient power technologies (Jamasb, Nepal, & Timilsina, 2017).

Korea Electric Company, Ltd. (KEPCO) manages the generation, transmission, and supply of electric power in accordance with the Korea Electric Power Corporation Act (Ji & Kang, 2016). The Korean electric industry has been composed of the KEPCO and five subsidiaries of KEPCO since the Korean electric industry reform in 2001(Jamasb et al., 2017). The KEPCO and its five subsidiaries had an aggregate installed generating capacity of 207,334 GWh as on December 31, 2015, which is 100% of the gross generation of electric power from coal-fired power plants in Korea in 2015 (KEPCO, 2016).

The KEPCO is responsible for the management of SEPK from data collection and analysis through release (KEPCO, 2016). The data management tool used in the KEPCO is the SEPK management system. The KEPCO releases electric power statistics every quarter and annually (KEPCO, 2016). For this investigation, I downloaded the 2015 SEPK publication from the KEPCO website.

Instrumentation and Operationalization of Constructs

VSKLB. The instruments of data collection for the VSKLB as follow:

Birth report form. It is Korea standard format for self-report of birth (NLIC, 2016). The current version had been effective as of February 1, 2015 (NLIC, 2016). This form is only available in Korean and the birth notification can be done in Korean (NLIC, 2016). The birth notification and family registration procedures were integrated in 1970, and after that, the birth report form contains all items related to family relationships and birth events (KOSTAT, 2016). The items related to family registration include the characteristics of the newborn and parent as follows: name, gender of the infant, birth date and time, place of birth, type of birth (hospital, home, etc.), parent's marital status, resident registration number of the parent, and parental domicile or current address (Cho et al., 2013). The items related to birth event are as follows: gestational age at delivery, parity, birth order, and the infant's birth weight, the nationality of the parents, parental education, occupation, and duration of cohabitation regardless of legal marriage (Suh et al., 2009). All the items of the birth report form have measured by either scale or categorical levels.

Birth certificate. A valid certificate proving the fact of the live birth by a hospital, doctor or midwife who attended at the birth or that issued by a doctor who examined the infant (NLIC, 2016). Variables available for this research from the birth certificates include infant gender, birth weight, birth order, gestational age at birth, parental age, parental educational level and occupation (Kim et al., 2007). All the items of the birth certificate are measured at the scale or categorical level.

Reliability and validity of the instruments for the VSKLB. The level of reliability was high based on a low rate of 0.9% for the late reporting in 2005 (Cho et al., 2013 & KOSIS, 2016). The level of reliability has increased with the uniformity of medical practice in the assessment of anthropometric, clinical, and medical factors including gestational age, birth weight, parity, and birth order (Bodnar et al., 2014; Martin, Osterman, Kirmeyer, & Gregory, 2015; Suh et al., 2009; Vogel et al., 2014). Some researchers demonstrated that the extensive use of ultrasound measurements in gestational age estimation allowed for more precise and reliable information on gestational age (Martin et al., 2015 & Suh et al., 2009). The level of reliability was high due to the cross-check with relevant data obtained from other available systems or sources. The local family registrar verified the reported information such as age and residential address across the resident registration card, driver's license, passport, certificate of seal impression, or other identification prescribed by the Supreme Court Regulations (Cho et al., 2013; Hwang, Kim, & Kim, 2017; KOSIS, 2016). The registrar confirmed the consistency of the reported maternal marital status with the record from the family registration system (Cho et al., 2013; Hwang, Kim, & Kim, 2017; KOSIS, 2016; Mikkelsen et al., 2015).

The level of external validity increased with the enactment of the Family Relationships Act, which contributed to reducing missing reports (Cho et al., 2013; KOSIS, 2016; Mikkelsen et al., 2015). The level of external validity was high with the high level of reporting completeness (Cho et al., 2013; KOSIS, 2016; Mikkelsen et al., 2015). Some researchers found that the external validity increased as the online family registration system became more accessible (Cho et al., 2013; Mikkelsen et al., 2015; Pu, Gao, Fan, & Wang, 2016).

The level of content validity was high with the adoption of the United Nations (UN) principles and recommendations in developing the questionnaire (KOSIS, 2016 & Vogel et al., 2016). Researchers concluded that the standardized birth report form over the Korea contributes to the high level of the content validity (Kim et al., 2007; Leem et al., 2006; Suh et al., 2009). On the other hand, some researchers expressed concerns about the external validity because births born to foreign parent were excluded from the family registration system (Cho et al., 2013; Hwang, Kim, & Kim, 2017; Vogel et al., 2016).

KMED for $PM_{2.5}$. The instruments of measuring $PM_{2.5}$ concentrations from air quality database as follow:

Weight concentration or equivalent automatic measurement methods. These methods are in use to measure $PM_{2.5}$ mass concentrations from monitoring stations (KME, 2016). Beta attenuation monitor was chosen in measuring and monitoring $PM_{2.5}$ mass level in Korea (KME, 2016). Using the Beta attenuation monitor method, $PM_{2.5}$ sample is a collection of particulate matter less than 2.5μm floating in the air for a certain period of time (Liu et al., 2013; Loh & Choi, 2017). Beta-rays penetrate the $PM_{2.5}$ sample and measure the mass concentration of $PM_{2.5}$ (Liu et al., 2013; Loh & Choi, 2017). The daily mean concentration of $PM_{2.5}$ is the average of the mass concentrations of $PM_{2.5}$ continuously measured throughout the day (KME, 2016).

Measurement sites. To date, a total of 100 valid monitoring stations are located nationwide (KMED, 2016). The location of the monitors allows for maximizing spatial representativeness in proportion to population density and measuring the overall level by averting the effects of extraneous local factors (KME, 2016 & Loh & Choi, 2017). The spatial measurements of KMED for $PM_{2.5}$ is limited to the 100 regions in which $PM_{2.5}$ monitors are placed.

Reliability and validity of the instruments for the KMED for $PM_{2.5}$. The level of internal validity was high with the automation of the Beta attenuation monitor method (Liu et al., 2013; Schweizer, Cisneros, & Shaw, 2016; Toro et al., 2015). The level of reliability was high with the high time resolution and data accuracy of the Beta attenuation monitor method in comparisons to the gravimetric mass measurement method and light-scattering method (Liu et al., 2013; Schweizer, Cisneros, & Shaw, 2016; Triantafyllou et al., 2016). The level of internal validity was high with the uniformity of operational procedures in the calibration, measurement, and validation of the data across all the monitoring stations (Toro et al., 2015). On the other hand, 55% of the monitors concentrated on the city, such as 23 in Seoul, 17 in Busan, 8 in Daegu, and 7 in Incheon (KMED, 2016 & Loh & Choi, 2017). The level of external validity needs to be improved by expanding the air quality monitoring networks to areas not installed yet and reducing the inequity of air quality monitoring service (Loh & Choi, 2017). Several researchers have found that there was little difference in adverse health effects according to the chemical profile of $PM_{2.5}$ per se (Tran et al. 2000, & Maynard & Maynard, 2002). They also supported that mass concentration of $PM_{2.5}$ is a more influential determinant accounting for adverse health effects (Tran et al. 2000, & Maynard & Maynard, 2002).

EAPSS. The instrument of data collection from EAPSS as follow:

EAPS. This survey is designated for the national statistics by the Statistics Act (KOSTAT, 2016). Any persons who are requested to answer have an obligation on response with good faith in accordance with the Statistics Act Art. 32 and 34 (R.O. Korea) (KOSTAT, 2016). The current version was effective as of January 01, 2015 (KOSIS, 2016). The primary purpose of this survey is to collect necessary data for analyzing the macroeconomic status and developing policies of human resources (KOSIS, 2016). The Statistics Korea, Bureau of Social Statistics, Employment Statistics

Division is responsible for the operation of this survey (Hwang et al., 2017). The data is collected via 15 minutes long personal interviews (Hwang et al., 2017). It measures labor supply, employment structure, available labor hours and degree of human resource utilization (Hwang et al., 2017). Once the interview survey is completed, the information entered into the PDA is transmitted to the regional Statistics Office and then passed to the National Statistics Office (KOSIS, 2016).

Reliability and validity of the EAPSS. The level of external validity was high with the rigid two-stage stratified cluster random sampling procedures (KOSIS, 2016 & Pu, Gao, Fan, & Wang, 2016). The representativeness of the sample increased with the replacement of 900 households every month (Hwang et al., 2017 & KOSIS, 2016). The level of external validity was high with the high response rate, which was 95% in 2015 (KOSIS, 2016 & Szklo, & Nieto, 2014). The recruitment tactics including promotional materials and return presents helped to increase the response rate (Hwang, Kim, & Kim, 2017 & KOSIS, 2016). The degree of content validity was high with the reasonable procedure like the questionnaire development using pilot surveys (KOSIS, 2016). The level of reliability was high with the in-person interview procedures and the continued training of the technical manual for interviewers (Cerin et al., 2016 & KOSIS, 2016). The level of reliability was high with the lowest level of geographic disaggregation and frequency to optimize the estimate from 16 cities/provinces every month (Hwang et al., 2017; KOSIS, 2016; Rodriguez-Rodriguez et al., 2017).

SEPK. The instrument of data collection from SEPK as follows:

SEPK management system. KEPCO has acted as the provider of national statistics relating to the Korean electricity industry activity from 1961 (KEPCO, 2016). KEPCO operates SEPK management system, to analyze, communicate, and disseminate the electric power industry activities (KEPCO, 2016). The SEPK management system helps maintain statistical consistency and efficiently operate electric power statistics (KEPCO, 2016). KEPCO collaborates with its five subsidiaries to manage and collect data of facilities, resources, performance, purchase, and sales status (KEPCO, 2016). They utilize SEPK management system to process and manipulate data and aggregate and validate data to produce national electric power statistics.

Reliability and validity of the SEPK. The level of reliability was high as the basis of Korea's energy policy formulation process (Han, Yoo, & Kwak, 2004; Ji, & Kang, 2016; KEPCO, 2016). The internal validity increased with the development of the technical management system and operational manual (Ji, & Kang, 2016; Oh, 2015; Streimikiene, & Siksnelyte, 2016). The SEPK management system has long been regarded as a sustainable, powerful, and practical statistics management system (Erdogdu, 2014; Ji, & Kang, 2016; Oh, 2015; Streimikiene, & Siksnelyte, 2016). The level of content validity was high with the reasonable process of data item development addressing issues for each task according to the degree of urgency, importance, and influence for each task (Ji, & Kang, 2016 & KEPCO, 2016). The level of external validity was high with the monopolized role of KEPCO and its five subsidiaries as they are responsible for the overall electricity generation of Korea (Han, Yoo, & Kwak, 2004; Jamasb, Nepal, & Timilsina, 2017; Ji, & Kang, 2016; KEPCO, 2016).

Operationalization

Dependent variables. The dependent variables are as follows:

Gestational age. It was defined as the number of weeks completed from the conception to birth of the newborn (Leem et al., 2006). I used the recorded gestational age in the VSKLB for the gestational age variable in this study. I made a subset with gestational age between 23 and 40 weeks in the VSKLB data set. The gestational age variable was interval scale, and its range was between 23 and 40 years.

Preterm birth. It was defined as a birth with a gestational age of 23 to 36 weeks (Cloherty, 2012). Based on the recorded gestational age in the VSKLB, I dummy-coded the preterm birth variable as 0 if the gestational age was between 37 weeks to 40 weeks and 1 if the gestational age was between 23 weeks to 36 weeks. The variable preterm birth was dichotomous nominal.

Independent variables. The independent variables are as follows, in a hierarchical order based on the SEM:

Individual factors. They were defined as variables located on the innermost layer based on the SEM. I re-coded the individual level factors in the VSKLB data set as follows: maternal age into three categories as 0 if the ages were in 20 to 29, 1 if the ages were in 30 to 39, 2 if the ages were in 40 to 44; maternal education level into two categories as 0 if mothers had ≤ 12 years of education, 1 if mothers had >12 years of education; maternal occupation status into two categories as 0 for unemployed and 1 for employed.

Community factor. It was defined as a variable located on the second layer based on the SEM. This included maternal $PM_{2.5}$ exposure. Maternal $PM_{2.5}$ exposure was measured as continuous data and categorical data. I calculated the continuous data based on the average and variance of $PM_{2.5}$ exposures for each exposure window. The analysis included five exposure windows, including the 1st, 2nd, 3rd, and 4th weeks before giving birth, and the last 4 weeks of pregnancy. I created the categorical data with dummy coding. I classified the averaged $PM_{2.5}$ exposure as 0 for < 10 $\mu g/m^3$, 1 for 10-24.99 $\mu g/m^3$, and 2 for ≥ 25 $\mu g/m^3$ for each exposure window. Cutoffs of 10 and 25 $\mu g/m^3$ used to categorize average exposures were the 2015 WHO and KME guidelines of annual average exposure to $PM_{2.5}$, respectively. The categorical data also included dichotomous $PM_{2.5}$ exposures, either exposed or unexposed, which was defined as above or below the current Korea Air Quality Standard of 25 $\mu g/m^3$ based on the average $PM_{2.5}$ exposure during the last 4 weeks of pregnancy.

The exposure measurement period was the same for all subjects as 1 week and 4 weeks and computed based on the date of birth of the infant. I dealt with the missing $PM_{2.5}$ measurements by replacing them with an average concentration of the weeks before and after. I used area level address of maternal residence recorded in the VSKLB database to assign one of the 252 sub-tie administrative districts. The sub-tier administrative district was the lowest geographical unit for which community-level data were available (KOSIS, 2016). It represents ecologic groups of 10,001 to 616,982 people and the area of GU, SI, and GUN was respectively 3 to 223 km^2, 73 to 1202 km^2, and 33 to 1522 km^2 (KOSIS, 2016). Linked data files for KMED and VSKLB provided data such as the daily average concentration of $PM_{2.5}$ measured at 100 monitoring sites throughout Korea, the date of birth, and the sub-tier administrative district of maternal residence. I assessed the maternal $PM_{2.5}$ exposure variable as either interval scale or ordinal categorical, depending on the research question.

Organizational factor. It was defined as variable located on the third layer based on the SEM. This included the unemployment rate measured by the average unemployment rate in the maternal residential area during the last 4 weeks of pregnancy. The maternal residential area was defined as one of the 17 top-tier administrative divisions where the woman lived. I evaluated the unemployment rate variable as a weighted average of the monthly data of unemployment rate taken at the months of the last 4 weeks of pregnancy. Linked data files of the EAPSS and VSKLB provided monthly averages of the unemployment rate and date of birth. The variable was measured as the ratio scale.

Policy factor. It was defined as variable located on the outermost layer based on the SEM. This referred to coal power generation measured by the gross generation of electricity from coal source in Korea for the last 4 weeks of pregnancy. I calculated the coal power generation variable as a weighted average of monthly coal-powered electricity generation in the months of the last 4 weeks of pregnancy. The linked data files of the SEPK and VSKLB provided monthly averages of gross generation of electricity from coal source in Korea and date of birth. I have expressed the variable as the interval scale level.

Covariates. Covariates included interpersonal and environmental factors and infant characteristics.

I re-coded interpersonal factors in the VSKLB into the dichotomous categorical level as follows: duration of cohabitation (< one year, ≥ one year); maternal marital status (married or unmarried).

I evaluated the environmental factor as follows: proximity to coal power plants (the weighted average distance according to the amount of coal-powered electricity generation of each plant from the monitoring station in maternity residential area to all coal-fired power plants).

I assessed the newborn characteristics as follows: birth weight (2,500 g to less than 3,000 g, 3,000 g to less than 4,000 g, 4,000 g to 5,000 g); infant gender (Male or Female).

Table 3

Operationalization of Study Variables

Variable name	Variable type	Measurement level	Operationalization
		Dependent variables	
Gestational age	Individual data	Scale-interval	Range: 23-40 weeks
Preterm birth	Individual data	Nominal-dichotomous	0 = Term birth 1 = Preterm birth
		Independent variables	
Individual factors			
Maternal age	Individual data	Ordinal-categorical	0 = 20-29 years 1 = 30-39 years 2 = 40-44 years
Maternal education level	Individual data	Nominal-dichotomous	0 = ≤ 12 years 1 = > 12 years
Maternal occupation status	Individual data	Nominal-dichotomous	0 = unemployed 1 = employed
Community factor			
$PM_{2.5}$ exposure	Ecological data		
Average $PM_{2.5}$		Scale-interval	Range: Min/max of the mean concentration of $PM_{2.5}$ for each exposure window
Variance $PM_{2.5}$		Scale-interval	Range: Min/max of the variance of $PM_{2.5}$ for each exposure window

Categorical $PM_{2.5}$		Ordinal-categorical	0 = < 10 µg/m^3 1 = 10-24.99 µg/m^3 2 = ≥ 25 µg/m^3
Dichotomous $PM_{2.5}$		Nominal-dichotomous	0 = < 25 µg/m^3 1 = ≥ 25 µg/m^3
Organizational factor			
Unemployment rate	Ecological data	Scale-ratio	Range: Min/max of the averaged unemployment rate for the last 4 weeks of pregnancy
Policy factor			
Coal power generation	Ecological data	Scale-interval	Range: Min/max of the averaged gross generation of electricity for the last 4 weeks of pregnancy
	Covariates		
Infant birth weight	Individual data	Ordinal-categorical	0 = 2.50-2.99 kg 1 = 3.00-3.99 kg 2 = 4.00-5.00 kg
Duration of cohabitation	Individual data	Nominal-dichotomous	0 = < 1year 1 = ≥ 1year
Infant gender	Individual data	Nominal-dichotomous	0 = Male 1 = Female
Maternal marital status	Individual data	Nominal-dichotomous	0 = Unmarried 1 = Married
Proximity to coal fired power plants	Ecological data	Scale-interval	Range: Min/max of the averaged distance

Data Analysis Plan

For this study, software used for analyses was the statistical program SPSS, version 24. In the data cleaning procedures, birth records containing missing, unknown, or incomplete values were excluded from the study data set. In the data screening procedures, using the eligibility criteria specified in the population section, I created a subset made up of subjects eligible for this study from the VSKLB data file, and then linked the subset with other data files such as KMED, EAPSS, and SEPK data files. Records with all of the maternal age of 20-44 years, the gestational age of 23-40 weeks, nulliparous pregnancies, singleton births, birthweight of 2,500g-5,000g, and Korean ethnicity were included in the subset.

In the descriptive statistics, the continuous variables, which was symmetric distribution, was summarized as mean, standard deviation (SD), variance, and the asymmetric continuous variable were expressed as median, range, and interquartile range (IQR). Categorical or dichotomous variables were described by the number of cases, percentage, mode, and Index of qualitative variation (IQV) in a table.

In preliminary statistics, I used chi-square and Wilcoxon-Mann Whitney tests to determine whether there was a difference between the unexposed and exposed groups according to study variables. I also utilized one-way ANOVA, Kruskal Wallis, and Pearson correlation tests to examine the association between variables and identify potential confounders. I used bivariate regression analyses between sociodemographic variables, $PM_{2.5}$ exposure, and gestational age to test significant direct effects between the variables and to assess whether prerequisites of moderation and mediation analysis.

The research questions and hypotheses are restated as follows:

RQ1: Do individual, organizational, and policy-level factors explain community-level $PM_{2.5}$ exposure?

$H_0$1: There are no statistically significant effects of individual, organizational, and policy-level factors on community-level $PM_{2.5}$ exposure.

H_a1: There are statistically significant effects of individual, organizational, and policy-level factors on community-level $PM_{2.5}$ exposure.

RQ2: What is the relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors?

$H_0$2: There is no statistically significant relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors.

H_a2: There is a statistically significant relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors.

RQ3: What effect, if any, does the interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, and occupation have on gestational age?

$H_0$3: The interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, occupation has no statistically significant effect on gestational age.

*H*a3: The interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, occupation has a statistically significant effect on gestational age.

RQ4: What effect, if any, do sociodemographic factors, such as maternal age, education level, and occupation have on gestational age through $PM_{2.5}$ exposure?

$H_0$4: Sociodemographic factors, such as maternal age, education level, occupation have no statistically significant effects on gestational age through $PM_{2.5}$ exposure.

H_a4: Sociodemographic factors, such as maternal age, education level, occupation have statistically significant effects on gestational age through $PM_{2.5}$ exposure.

In inferential statistics, I used multiple regression analysis to determine the effects of individual, organizational, and policy level factors on community-level $PM_{2.5}$ exposure. Multiple linear regression analysis was selected for two reasons: (a) the dependent maternal $PM_{2.5}$ exposure variable was continuous scale, and (b) the linear model could serve explicit interpretability with unstandardized coefficients, unlike odds ratios. I set the statistical significance at the $p < .05$. I performed a multiple logistic regression analysis to examine the association between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors. I selected the multiple logistic regression analysis for two reasons: (a) the dependent preterm birth variable was dichotomous, and (b) the logistic model could provide odds ratios from the combination of categorical and continuous variables (Warner, 2013). I performed the statistical analysis after controlling for covariates that were associated with maternal $PM_{2.5}$ exposure and preterm birth and expected to be potential confounders. I assessed the statistical significance of the odds ratio by 95% CI. Inference was based on statistical significance at the 5% level. I performed moderation analysis to determine whether the interaction between community-level $PM_{2.5}$ exposure and individual level sociodemographic factors significantly predict gestational age. I conducted mediation analyses on the effects of sociodemographic variables on gestational age through $PM_{2.5}$ exposure. I used the mediation analysis guided by Preacher and Hayes's approach, and tested statistical significance of the indirect effects of mediation via bootstrapping using SPSS. I selected the Preacher and Hayes's approach because it could deal with the non-normal distribution of the indirect effect through bootstrapping (Preacher, & Hayes, 2008). In this analysis, I set the bootstrap command at 5,000 sampling replications and 95% CI (Hayes, 2012).

Threats to Validity

Threats to internal validity. Threats to internal validity are as follows:

History. During the entire pregnancy, the mother might be exposed to stressors other than the study interest, which may interfere with evaluation on the impact of exposure to study variables, including $PM_{2.5}$. The influence of other stressors on the development of preterm birth could not be ruled out. It may not be possible to conclude that the preterm birth observed in this study was indeed due to differences in the exposure measured in this study. To avoid these potential threats, I equally applied a short averaging exposure period of 1 week and 4 weeks to all subjects, reducing the likelihood of exposure to stressors other than this study interest (Cutts, Izurieta, & Rhoda, 2013).

Threats to external validity. The study sample was composed of people residing in the area where the $PM_{2.5}$ monitor was installed. Over 50% of monitors were installed in the city (Loh & Choi, 2017). The urban population differs from the rural population in sociodemographic characteristics (Laurent et al., 2016). The findings from the study of urban population could not be generalizable to the rural population. $PM_{2.5}$ monitors are in the process of expanding their installations to rural areas, so the findings from this study may not be replicable. The results of this study may differ from those obtained after $PM_{2.5}$ monitor construction is completed. To mitigate these threats, I used the stratified single-stage cluster sampling method to ensure that (a) the probability of each subgroup being included in the sample was the same, and (b) various subgroups could be included in this study sample (Teddlie, & Yu, 2007).

Ethical Procedures

Secondary data analysis using existing de-identified data, such as licensed microdata, did not require regulatory authority (e.g., Ministry of Drug and Food Safety (MFDS)) approval. I sought the Walden University Institutional Review Board (IRB) approval as mandated for all research at Walden University to comply with the University's ethical standards and U.S. federal regulations (WU, n.d.). I have received a certificate from National Institute of Health (NIH) *Protecting Human Research Participants* course. I have signed a data use agreement regarding the use of licensed microdata to comply with the laws and regulations of the governing bodies associated with the data recipient's educational program (Appendix A). In accordance with the laws and regulations of the governing bodies associated with the KOSTAT and the data recipient, I completed a confidentiality agreement. I, as a data user, adhered to use the licensed microdata for the approved purposes only and obtain KOSTAT's review and approval to download SPSS outputs (KOSIS, 2015).

This study required information on the date of birth and residential address, which is considered as a personal identifier by the Personal Information Protection Act and the guideline of Health Insurance Portability and Accountability Act (HIPAA) (KOSIS, 2015 & Salazar et al., 2015). Employment of these data in this study poses ethical concerns about potential intrusion of privacy and breach of confidentiality. To address ethical concerns related to the date of birth, I used the date of birth only to identify the end date of pregnancy to assess maternal exposure to $PM_{2.5}$ during the last 4 weeks of pregnancy. Certain individuals can be identified only if their date of birth is not in and of itself, but in combination with other information. Licensed microdata did not contain any information that could make it possible to identify a specific individual by combining it with the date of birth. I employed area-level addresses in this study. The area level addresses could provide only information on a sub-tier administrative division where the mother lived among the 252 sub-tier administrative districts called SI, GU or GUN. The full address, including the address following the SI, GU or GUN address was not obtained.

The licensed microdata cannot be downloaded from the server. Data downloaded from the public domain was stored on a password-protected computer, and the data and documents are destroyed confidentially after five years (WU, n.d.).

Summary

This study was a quantitative, observational, partially ecologic, retrospective cohort study using pre-existing data including those from VSKLB, KMED, EAPSS, and SEPK. In this study, I examined the effects of maternal $PM_{2.5}$ exposure on preterm birth, including the possibility of moderation and mediation. I also worked to investigate what variables to have the effects on maternal exposure to $PM_{2.5}$. Specifically, I sought to determine whether maternal $PM_{2.5}$ exposure at the community level modifies gestational age by its interacting with the individual level sociodemographic factors such as maternal age, maternal education level, and maternal occupation status.

This investigation was limited to Korean women for whom $PM_{2.5}$ data of the last 4 weeks of pregnancy was available. This study targeted women of Korean ethnicity who gave birth in 2015. Additional parameters were nulliparous, singleton births of women aged 20 to 44. I selected the study population according to stratified single-stage cluster sampling and then screened by inclusion criteria. I obtained the dependent variables, gestational age, and preterm birth from the VSKLB. I obtained the independent variables as follows: (a) $PM_{2.5}$ exposure from the linkage of the KMED and VSKLB data files, (b) maternal sociodemographic variables from the VSKLB data set, (c) unemployment rate from the linked data files of the EAPSS and VSKLB, and (d) coal power generation from the linkage of the SEPK and VSKLB data files. In the statistical analysis, I used the SPSS, version 24. I displayed the descriptive statistics in tables. I used One-way ANOVA, Kruskal Wallis, and Pearson correlation for preliminary statistics, and multiple logistic regression, multiple linear regressions including testing for moderation and mediation for inferential statistics.

In Chapter 4, I report the results of the statistical analysis and findings of this study.

Chapter 4: Results

Introduction

In this study, I investigated the effects of $PM_{2.5}$ exposure on preterm birth. Specifically, I worked to determine whether maternal $PM_{2.5}$ exposure at the community level moderates and mediates the association between individual sociodemographic factors such as maternal age, maternal education level, maternal occupational status, and gestational age. I also sought to determine the individual, organizational, and policy level factors that influence the difference in $PM_{2.5}$ exposure at the community level.

The research questions and hypotheses were as follows:

RQ1: Do individual, organizational, and policy-level factors explain community-level $PM_{2.5}$ exposure?

$H_0$1: There are no statistically significant effects of individual, organizational, and policy-level factors on community-level $PM_{2.5}$ exposure.

H_a1: There are statistically significant effects of individual, organizational, and policy-level factors on community-level $PM_{2.5}$ exposure.

RQ2: What is the relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors?

$H_0$2: There is no statistically significant relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors.

H_a2: There is a statistically significant relationship between community-level maternal $PM_{2.5}$ exposure and preterm birth after controlling for individual, organizational, and policy-level factors.

RQ3: What effect, if any, does the interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, and occupation have on gestational age?

H_03: The interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, occupation has no statistically significant effect on gestational age.

H_a3: The interaction between community-level $PM_{2.5}$ exposure and individual-level sociodemographic factors, such as maternal age, education level, occupation has a statistically significant effect on gestational age.

RQ4: What effect, if any, do sociodemographic factors, such as maternal age, education level, and occupation have on gestational age through $PM_{2.5}$ exposure?

H_04: Sociodemographic factors, such as maternal age, education level, occupation have no statistically significant effects on gestational age through $PM_{2.5}$ exposure.

H_a4: Sociodemographic factors, such as maternal age, education level, occupation have statistically significant effects on gestational age through $PM_{2.5}$ exposure.

This chapter is organized into three parts. In each section, I describe the data collection process and the results of the data analysis. I conclude the chapter with a summary and a transition to the next chapter.

Data Collection

Data collection began with Walden's IRB approval (# 03-06-18-0601853), and the data use agreement after KOSTAT's authorization for remote data access (Appendix A). The KOSTAT provided a virtual server containing data for 2015 birth records, $PM_{2.5}$ measurements, unemployment rates, and power generation statistics. There were no differences in data collection compared to the data collection plan for this study approved by Walden's IRB.

In the VSKLB data file, I selected study sample members using a two-step approach consisting of stratified single-stage cluster sampling and screening by inclusion criteria. A total of 12 strata were selected in the stratified sampling procedure because one or more valid $PM_{2.5}$ monitoring stations were installed in only the 12 top-tier administrative districts (KME, 2016): Seoul, Busan, Daejeon, Daegu, Gwangju, Gangwon Province, Gyeonggi Province, Gyeongsangnam Province, Incheon, Jeju, Jeollanam Province, and Ulsan (see Table 4). In the cluster sampling procedure, one cluster with the largest number of births was chosen for each of the 12 strata using a probability proportional to size sampling method. Twelve clusters including Gangseo-gu, Jin-gu, Dalseo-gu, Seo-gu, Buk-gu, Yuseong-gu, Nam-gu, Bundang-gu, Wonju-si, Yeosu-si, Gimhae-si, and Jeju-si, were selected for screening in this study, as shown in Table 4.

In the screening procedure, pregnant women who met all inclusion criteria were selected for this study. Records with all of the maternal age of 20-44 years, gestational age of 23-40 weeks, nulliparous pregnancies, singleton births, birth weight of 2,500g - 5,000g, and Korean ethnicity were included in the subset. As a result of all the sampling procedures, a total of 19,371 pregnant women were recruited from the VSKLB data file as shown in Table 4.

Table 4

Sampling Scheme

Strata	Clusters	Number of subjects
Seoul	Gangseo-gu	2,644
Busan	Jin-gu	1,329
Daegu	Dalseo-gu	1,921
Incheon	Seo-gu	1,907
Gwangju	Buk-gu	1,387
Daejeon	Yuseong-gu	1,492
Ulsan	Nam-gu	1,320
Gyeonggi-do	Bundang-gu	2,138
Gangwon-do	Wonju-si	989
Jeollanam-do	Yeosu-si	884
Gyeongsangnam-do	Gimhae-si	1,901
Jeju	Jeju-si	1,459
	Total	19,371

Note. Each stratum is defined as one of the 17 top-tier administrative districts; Each cluster is defined as one of the 252 sub-tier administrative areas which are called either SI for the major cities, GU in urban areas or GUN in rural areas.

I linked the subset made from the VSKLB data file with other data files including those from KMED, EAPSS, and SEPK. In the data cleaning procedure, no birth records were excluded from the study data set because there was no birth record that contained missing, unknown, or incomplete values. There were no missing data in VSKLB, EAPSS, and SEPK data files. In the KMED file, missing $PM_{2.5}$ measurements of 0.8%

were handled by replacement with the mean concentration of the weeks before and after. The final data set of this study was constructed with 19,371 subjects. The sample size was increased from 1,084 to 19,371 due to sampling from existing data that already has a fixed sample size. I adopted the increased sample size, taking into account the low frequencies of the exposure and outcome, specifically, the small size of the exposed group in this cohort design (see Aschengrau, & Seage, 2014). The sample size of 19,371 subjects far exceeded the minimum sample size of 1,084 required for the analysis calculated in the previous chapter.

The study sample was representative of the target population of this investigation. The number of pregnant women who gave singleton live births in 2015 was 421,988. Of these, 169,841 pregnant women met all inclusion criteria used to minimize confounding in this study. I assessed the representativeness of this study sample for the population of these 169,841 pregnant women. The preterm birth rate was the same at 2.6% in both this study sample (*n* = 496) and the population group of 169,841 pregnant women (*n* = 4,488), and the study sample can be said to be representative of the population of 169,841 pregnant women. I calculated the post-hoc power for the increased final study sample size with the same input parameters using G*Power, version 3.1. The size of this sample was large enough not to underestimate the effect of maternal exposure to $PM_{2.5}$, with 100% power in this cohort study. The study sample showed an adequate level of representation allowing for generalization of the study results. This sample provided an appropriate size for all statistical analyses planned in this study, even with the low frequencies of the exposure and outcome.

Results

This section includes the results of the descriptive, preliminary, and inferential statistics for this study sample.

Descriptive Analyses

After sampling, I included a total of 19,371 eligible pregnant women in this study analysis. Table 5 summarizes the frequency distribution of categorical characteristics of 19,371 mothers in this study. In this population, the distribution of categorical characteristics, including infant's gender and maternal occupation status, was almost even across the categories based on a high variation of 99.8%. The duration of cohabitation showed a fairly heterogeneous distribution in the sample, with a relatively high variation of 94.1%, and the most typical duration of cohabitation was more than one year (n = 12,035, 62.1%). The most homogeneous distribution was observed in maternal marital status, with a variation of 4.7%, and nearly every mother of this study sample was married (n = 19,138, 98.8%). Most of the 19,371 mothers had full-term births (n = 18,875, 97.4%), and the distribution of preterm birth was relatively homogeneous in this study sample based on the variation of 10.1%. Of the 19,371 mothers included in this study, 34.4% (n = 6,671) of the mothers were between the ages of 20 to 29 years old; 63.8% (n = 12,360) were age 30 to 39 years; and 1.8% (n = 340) were 40 through 44 years old. The majority of mothers had more than 12 years of education (84.4%; n = 16,353), and 15.6% (n = 3,018) had less than 12 years of education. For the birth weight of the newborn, the birth weight of 3.00-3.99 kg was the largest subgroup (72.4%, n =

14,017); followed by the birth weight of 2.50-2.99 kg (24.4%, *n* = 4,728); and the birth weight of 4.00 to 5.00 kg (3.2%, *n* = 626).

Table 5

Frequency Distribution of Categorical Characteristics of the Study Population

Variables	*N*	%	Mode	IQV
Duration of cohabitation	19,371	100	≥ 1 year	.941
< 1year	7,336	37.9		
≥ 1	12,035	62.1		
Infant birth weight	19,371	100	3.00-3.99 kg	.623
2.50-2.99 kg	4,728	24.4		
3.00-3.99	14,017	72.4		
4.00-5.00	626	3.20		
Infant gender	19,371	100	Male	.998
Male	10,149	52.4		
Female	9,222	47.6		
Maternal age	19,371	100	30-39 years	.711
20-29 years	6,671	34.4		
30-39	12,360	63.8		
40-44	340	1.8		
Maternal education level	19,371	100	> 12 years	.527
≤ 12 years	3,018	15.6		
> 12	16,353	84.4		
Maternal marital status	19,371	100	Married	.047
Unmarried	233	1.2		
Married	19,138	98.8		
Maternal occupation status	19,371	100	Unemployed	.998
Unemployed	10,106	52.2		
Employed	9,265	47.8		
Preterm birth	19,371	100	Full term	.101
No	18,875	97.4		
Yes	496	2.6		

Note. Acronyms: IQV: Index of qualitative variation; *N*: Number of cases.

Table 6 includes descriptive statistics on continuous variables of this study sample. Variables, including gestational age and proximity to coal plants, were reported as median, interquartile range (IQR), and range because the distributions were skewed. The remaining data showed a symmetrical distribution and are expressed as mean ($\bar{x}$), variance (s^2), and standard deviation (s). The gestational age was between 28 and 40 weeks, with a median of 39 weeks. The mothers' proximity to coal plants ranged from 1.29 to 3.29 $\times 10^2$ km, and median distance was 1.84 $\times 10^2$ km. The mean concentration of $PM_{2.5}$ for the last 4 weeks of pregnancy was 23.01 μg/m^3 in this study population. The most heterogeneous average $PM_{2.5}$ exposure in the entire cohort appeared at the 4th week before childbirth ($\bar{x}$ = 22.568 μg/m^3, s^2 = .023, s = .153), with the most homogeneous average $PM_{2.5}$ occurring during the last 4 weeks of pregnancy. ($\bar{x}$ = 23.014 μg/m^3, s^2 = .012, s = .111). The mean of the $PM_{2.5}$ variance over the last 4 weeks of pregnancy was 121.20 in this population. The most heterogeneous variance of $PM_{2.5}$ exposure in the sample appeared at the 3rd week before childbirth ($\bar{x}$ = 82.243, s^2 = .188, s =.434), with the most homogeneous variance of $PM_{2.5}$ exposure occurring in the last 4 weeks of pregnancy ($\bar{x}$ = 121.199, s^2 = .086, s = .293). The mean value of the average $PM_{2.5}$ exposure of the whole sample per exposure window was: the last 4weeks of pregnancy, 23.01 μg/m^3; the 1st week before childbirth, 22.23 μg/m^3; the 2nd week before childbirth, 22.29 μg/m3; the 3rd week before childbirth, 22.51 μg/m^3; the 4th week before childbirth, 22.57 μg/m^3. The mean value of the variance $PM_{2.5}$ exposure of the entire sample over each exposure window was: the last 4 weeks of pregnancy, 121.20; the 1st week before childbirth, 82.95; the 2nd week before childbirth, 79.62; the 3rd week before childbirth, 82.24; the 4th week before childbirth, 82.83. The highest average exposure and variance exposure were observed during the last 4 weeks of pregnancy, reflected by 23.01 μg/m^3 and 121.20, respectively. The mean values of coal power generation and unemployment rate were 17.096×10^3 GWh and 3.505, respectively, with the variability of .001 and .995, respectively.

Table 6

Descriptive Statistics on Continuous Characteristics of the Study Population

	M	Variation	Skewness	Kurtosis
Coal power generation [a]	1.2329	.001 (.02690)	.156	-.926
Gestational age	39.00	2 (28-40)	-1.135	3.021
Proximity to coal plants	1.8400	.80 (1.29-3.29)	1.063	.353
Unemployment rate	3.5053	.995 (.99774)	.258	-.410
Average $PM_{2.5}$				
Last 4 weeks of pregnancy [a]	1.3620	.012 (.11086)	-.025	.122
The 1st week before childbirth [a]	1.3470	.023 (.15008)	-.119	-.001
The 2nd week before childbirth [a]	1.3481	.022 (.14930)	-.077	-.025
The 3rd week before childbirth [a]	1.3524	.023 (.15195)	-.106	-.064
The 4th week before childbirth [a]	1.3535	.023 (.15276)	-.101	-.035
Variance $PM_{2.5}$				
Last 4 weeks of pregnancy [a]	2.0835	.086 (.29287)	-.053	-.171
The 1st week before childbirth [a]	1.9188	.184 (.42862)	-.089	-.277
The 2nd week before childbirth [a]	1.9010	.185 (.43041)	-.090	-.242
The 3rd week before childbirth [a]	1.9151	.188 (.43372)	-.105	-.261
The 4th week before childbirth [a]	1.9182	.188 (.43314)	-.095	-.280

Note. Acronyms: IQR: Interquartile range; M: Mean or Median; Data are displayed as median, IQR (range) or mean, variance (standard deviation) depending on the distribution; [a]: The data was logarithmically transformed due to the skewed distribution.

I examined the differences in the characteristics between the unexposed and exposed groups in this cohort study. Women in this study were divided into two groups, either exposed or unexposed, which was defined as above or below the current Korea Air Quality Standard of 25 $\mu g/m^3$ based on the average $PM_{2.5}$ exposure during the last 4-week period of pregnancy. I used chi-square tests and Wilcoxon-Mann Whitney tests to compare differences in categorical and continuous characteristics between the two groups at a p value of 0.05. The Wilcoxon-Mann Whitney test was chosen for non-normally distributed continuous variables, including coal power generation, proximity to coal plants, and unemployment rate. Testing for normality on each data was performed using the Kolmogorov-Smirnov (K-S) test at a p value of 0.05. The p value of all data was 0.000, and the distribution of all continuous variables deviated significantly from normal (see Frankfort-Nachmias, & Leon-Guerrero, 2015).

Table 7 shows chi-square and Wilcoxon-Mann Whitney test results. Of the overall sample of 19,371 mothers, 12,071 (86.4%) mothers were grouped as unexposed while exposed group consisted of 7,300 (2.6%) mothers. No significant differences between the two groups were found in terms of duration of cohabitation, χ^2 [1] = .711; p =.399, infant birth weight, χ^2 [2] = 2.844; p =.241, infant gender, χ^2 [1] = .120; p =.729, maternal age, χ^2 [2] = .608; p =.738, maternal education level, χ^2 [1] = .156; p =.693, maternal marital status, χ^2 [1] = .012; p =.913, and maternal occupation status, χ^2 [1] = .213; p =.645. Significant differences between the two groups were detected in coal power generation, proximity to coal plants, and unemployment rate, U = 42,772,932, p = .001, U = 37,848,605, p = .000, and U = 37,848,605, p = .000, respectively. In this population, the mean ranks coal power generation and proximity to coal plants were higher in the unexposed group, and the mean rank unemployment rate was higher in the exposed group. This meant that exposed mothers had a higher unemployment rate than unexposed mothers, while their proximity to coal plants and coal power generation were relatively low. The differences in baseline characteristics between the exposed and unexposed groups identified in Table 7 were appropriately adjusted in inferential statistics.

Table 7

Differences in Characteristics Between Unexposed and Exposed Groups

	Unexposed (n = 12,071)	Exposed (n = 7,300)	p
Duration of cohabitation			.399
< 1year	4,599 (38.1%)	2,737 (37.5%)	
≥ 1	7,472 (61.9%)	4,563 (62.5%)	
Infant birth weight			.241
2.50-2.99 kg	2,949 (24.4%)	1,779 (24.4%)	
3.00-3.99	8,752 (72.5%)	5,265 (72.1%)	
4.00-5.00	370 (3.1%)	256 (3.5%)	
Infant gender			.729
Male	6,336 (52.5%)	3,813 (52.2%)	
Female	5,735 (47.5%)	3,487 (47.8%)	
Maternal age			.738
20-29 years	4162 (34.5%)	2,509 (34.4%)	
30-39	7,704 (63.8%)	4,656 (63.8%)	
40-44	205 (1.7%)	135 (1.8%)	
Maternal education level			.693
≤ 12 years	1,871 (15.5%)	1,147 (15.7%)	
> 12	10,200 (84.5%)	6,153 (84.3%)	
Maternal marital status			.913
Unmarried	146 (1.2%)	87 (1.2%)	
Married	11,925 (98.8%)	7,213 (98.8%)	
Maternal occupation status			.645
Unemployed	6,282 (52.0%)	3,824 (52.4%)	
Employed	5,789 (48.0%)	3,476 (47.6%)	
Coal power generation [a]	9792.55	9509.81	.001
Proximity to coal plants	10200.50	8835.24	.000
Unemployment rate	8639.98	11415.67	.000

Note. Chi-square tests were used for categorical variables; Wilcoxon-Mann Whitney tests were used for continuous variables; Categorical variables were displayed in frequency and percent; Continuous variables were expressed in mean rank; $p < 0.05$ indicates significant association; [a]: The data was logarithmically transformed.

Research Question 1

I ran a multiple linear regression analysis to identify factors associated with $PM_{2.5}$ exposure. Covariates were included in the regression model if they accounted for significant differences in $PM_{2.5}$ exposure or contributed to changes in the variance of the model, or either of the slopes in the prediction of $PM_{2.5}$ exposure. Table 8 displays the associations between covariates and $PM_{2.5}$ exposure. I used one-way ANOVA tests to examine the association of $PM_{2.5}$ exposure with infant birth weight and maternal age because the data met the assumption of normality (Frankfort-Nachmias, & Leon-Guerrero, 2015). I utilized Kruskal Wallis tests to assess the association of $PM_{2.5}$ exposure with duration of cohabitation, infant gender, maternal education level, occupation status, and maternal marital status because the data did not meet the assumption of normality (Frankfort-Nachmias, & Leon-Guerrero, 2015). I evaluated the assumption of normality for the $PM_{2.5}$ exposure distribution for each categorical covariate at a *p* value of 0.05 using the Kolmogorov-Smirnov (K-S) test. I used Pearson correlation tests to examine the association between $PM_{2.5}$ exposure and continuous covariates because all data met the assumption of linearity (Frankfort-Nachmias, & Leon-Guerrero, 2015).

Kruskal Wallis tests showed statistically significant differences in $PM_{2.5}$ exposure according to maternal education level, $\chi 2[1] = 5.403$; $p = 0.020$. The mean rank $PM_{2.5}$ exposure was higher among mothers with less than 12 years of education. The Pearson correlation tests indicated a significant weak correlation between $PM_{2.5}$ exposure and all continuous covariates, including coal power generation, proximity to coal plants, and unemployment rate, *r* = .029, -.112, and .287, respectively, and all *p* values of .000. The covariates, such as maternal education level, coal power generation, proximity to coal plants, and unemployment rate, were included in the linear regression model since they accounted for significant differences in $PM_{2.5}$ exposure. I performed a series of multiple linear regression analyses to evaluate the change in the adjusted R square of the model for each covariate added. The covariates, including duration of cohabitation, maternal age, and maternal occupation status, were selected for the linear regression model because they contributed to the change in the variance of the model or the change in either of the coefficients of predictors. I have categorized the unemployment rate as a dichotomous variable, above or below the mean, to mitigate the serious positive autocorrelation found using the Durbin-Watson test, *d* = 0.005 (Wagner, 2016). The dichotomous unemployment rate was included in the regression model because it was significantly associated with $PM_{2.5}$ exposure, such as the Kruskal Wallis test results, $\chi2[1] = 1037$; $p < 0.00001$.

Table 8

Association Between $PM_{2.5}$ Exposure [a] *and Covariates*

Covariates added	*p*	Adj. R^2	Slope changed
Maternal education level	.020		
Coal power generation [a]	.000		
Proximity to coal plants	.000		
Unemployment rate	.000		
Unemployment rate [b]	.000		
Duration of cohabitation	.105	.087	Yes
Infant birth weight	.093	.087	No
Infant gender	.908	.087	No
Maternal age	.494	.087	Yes
Maternal marital status	.820	.087	No
Maternal occupation status	.160	.088	Yes

Note. $p < 0.05$ indicates significant association; One-way ANOVA, Kruskal Wallis and Pearson correlation tests were used as appropriate; [a]: The data was logarithmically transformed; [b]: The data was categorized as a dichotomous variable.

The final multiple linear regression analysis model consisted of $PM_{2.5}$ exposure as a dependent variable, and duration of cohabitation, maternal age, maternal education level, maternal occupation status, coal power generation, proximity to coal plants, and unemployment rate, as the independent variables as shown in Table 9. In this multiple linear regression analysis model, the assumption of multicollinearity was satisfied with the values of the variance inflation factor (VIF) of 1.007 to 1.317 (Frankfort-Nachmias, & Leon-Guerrero, 2015). The values of Cook's distance ranged from .000 to .001, and no undue influence was found in this model (see Frankfort-Nachmias, & Leon-Guerrero, 2015). This model could not rule out the possibilities of many ties in the data, as confirmed in the Durbin-Watson test, $d = 0.007$ (see Wagner, 2016). The fitness of the overall regression model to the data was good with a p value of < 0.00001, and the predictors were, as a combination, significant within this model with an F-test result of F (7, 19,363) = 154.839. The level of $PM_{2.5}$ exposure estimates by predictors such as duration of cohabitation, maternal age, maternal education level, maternal occupation status, coal power generation, proximity to coal plants, and unemployment rate was low, $R = .230$. The combination of the predictors included in this model explained 5.3% of the variability in $PM_{2.5}$ exposure.

Table 9 shows the results of multiple linear regression analysis. The variables, such as maternal education level, maternal occupation status, and unemployment rate, significantly predicted $PM_{2.5}$ exposure. An equation for the multiple regression analysis was as follows: predicted $PM_{2.5}$ exposure = 1.285 - .007 × Maternal education level - .006 × Maternal occupation status + .050 × Unemployment rate. Multiple linear regression analysis showed a reduction of 1.016 $\mu g/m^3$ in $PM_{2.5}$ exposure to mothers who had received more than 12 years of education compare to mothers with less than 12 years of education. A 1.122 $\mu g/m^3$ increase in $PM_{2.5}$ exposure was associated with a higher unemployment rate than the sample mean, and employed mothers showed reduced $PM_{2.5}$ exposure by 1.014 $\mu g/m^3$. When comparing the strength of effect between significant predictors of $PM_{2.5}$ exposure regardless of the direction of effect, I found that unemployment rate at the organizational level was 7 and 8 times stronger than individual-level factors, maternal age, and maternal education level, respectively (B = .050, -.007, and -.006). The unemployment rate at the organizational level can be said to be the most influential factor of $PM_{2.5}$ exposure in this population.

Table 9

Multiple Linear Regression Analysis Assessing the Factors Influencing $PM_{2.5}$ Exposure [a]

	Coefficient	SE	*p*
Individual level			
Maternal age	-.002	.002	.241
Maternal education level	-.007	.002	.003
Maternal occupation status	-.006	.002	.001
Organizational level			
Unemployment rate [b]	.050	.002	.000
Policy level			
Coal power generation [a]	.054	.029	.063
Covariates			
Duration of cohabitation	-.001	.002	.521
Proximity to coal plants	-.001	.002	.363

Note. $p < 0.05$ indicates significant association; Acronyms: SE: Standard error; [a]: The data was logarithmically transformed; [b]: The data was categorized as a dichotomous variable.

Research Question 2

I performed a multiple logistic regression analysis to assess whether maternal $PM_{2.5}$ exposure was associated with preterm birth. I selected covariates for multivariate modeling if they met one or more of the following criteria: (a) They contributed to changing the variance in preterm birth up to more than 2-fold, or (b) they changed either of the odds ratio estimates for $PM_{2.5}$ exposure by more than 2-fold. I utilized the Cox & Snell R^2 method to evaluate the variation in preterm birth, also used a series of multiple logistic regression analyses for each covariate to assess the extent to which the odds ratios of $PM_{2.5}$ exposures and preterm birth changed. Table 10 displays the covariates selected for multivariate modeling. Infant birth weight was adjusted for because the variance in preterm birth changed by more than 2-fold by adding the covariate, infant birth weight. I also added the covariate, coal power generation in the model because it changed either of the odds ratio estimates for $PM_{2.5}$ exposures by more than 2-fold.

Table 10

Covariates to be Included in the Multiple Logistic Regression Model

Covariates	Cox & Snell R^2	Min.- Max. ratios of change in the ORs for $PM_{2.5}$ [a]
None (Referent)	.001	Referent
Coal power generation	.001	.79-3.00
Proximity to coal plants	.001	.98-1.04
Unemployment rate	.001	.98-1.00
Duration of cohabitation	.001	1.00-1.01
Infant birth weight	.023	.93-1.06
Infant gender	.002	.98-1.08
Maternal age	.001	1.00-1.03
Maternal education level	.001	.99-1.01
Maternal marital status	.001	.98-1.02
Maternal occupation status	.001	1.00 (No change)

Note. $p < 0.05$ indicates significant association; $PM_{2.5}$ exposure was measured as continuous data by averaging exposure over each exposure window, and categorical data by categorizing the average exposure to more or less than 10 and 25 $\mu g/m^3$ for each exposure window; Acronyms: OR: Odds ratio; [a]: Continuous data were logarithmically transformed.

The multiple logistic regression analysis model consisted of dichotomous preterm birth as a dependent variable and $PM_{2.5}$ exposures as independent variables while controlling for covariates, including coal power generation and infant birth weight. Maternal $PM_{2.5}$ exposure was averaged for the 1st, 2nd, 3rd, and 4th weeks before childbirth and during the last 4 weeks of pregnancy. The averaged $PM_{2.5}$ levels were categorized as < 10, 10-24.99, and ≥ 25 μg/m^3. In this analysis, I modeled maternal $PM_{2.5}$ exposure as a continuous variable and a categorical variable. I examined the risk estimates for a 10 μg/m^3 increase in $PM_{2.5}$ and three $PM_{2.5}$ categories of < 10, 10-24.99, and ≥ 25 μg/m^3. I assessed the association between $PM_{2.5}$ and preterm birth for five exposure periods, including the 1st, 2nd, 3rd, and 4th weeks before childbirth and the last 4 weeks of pregnancy. I used cutoffs of 10 and 25 μg/m^3 to categorize average exposures since they were the 2015 WHO and KME guidelines for an annual average exposure to $PM_{2.5}$, respectively. I set the statistical significance of association at the $p = 0.05$ level. I performed testing for the assumption of the goodness of fit using the Hosmer-Lemeshow test. I found a good fit of the model, $p = .558$ and this model indicated an acceptable level of prediction (see Frankfort-Nachmias, & Leon-Guerrero, 2015). This logistic regression model was statistically significant, with $\chi 2$ [17] = 456.483; $p < 0.00001$. In this model, the variance in preterm birth was 2.3 %, and 97.4% of cases were correctly classified.

Table 11 shows the adjusted ORs and 95% CIs for the associations between maternal exposure to $PM_{2.5}$ and preterm birth. A significant association was found in $PM_{2.5}$ exposure during the 3rd week before childbirth (adj*OR* = 6.520, 95% CI: 1.061; 40.074 for a 10 μg/m^3 increase in $PM_{2.5}$ exposure). $PM_{2.5}$ exposure during the 3rd week before childbirth was positively associated with the likelihood of preterm birth. The highest risk estimate for preterm birth was found in the exposure during the 3rd week before childbirth. For a 10 μg/m^3 increase in maternal $PM_{2.5}$ exposure during the 3rd week before childbirth, the risk of preterm birth increased by 6.52 times. No significant association between average exposure and preterm birth was found in the other exposure windows: adj*OR* = 5.337, 95% CI: .738; 38.577 for the 1st week before childbirth; adj*OR* = 4.657, 95% CI:.723; 29.979 for the 2nd week before childbirth; adj*OR* = 3.335, 95% CI:.495; 22.479 for the 4th week before childbirth, and; adj*OR* = .002, 95% CI:.000; 1.257 for the last 4 weeks of pregnancy, for each μg/m^3 increase in $PM_{2.5}$ mean concentration.

All women in this study reported an average $PM_{2.5}$ exceeding the WHO safe level of 10 μg/m^3 during the last 4 weeks of pregnancy. Of these, 37.7% had an average $PM_{2.5}$ higher than the KME standard of 25 μg/m^3. During the 4th week before childbirth, mothers with a $PM_{2.5}$ exposure of 10-24.99 and ≥ 25 μg/m3 had a 55% and 10.9% lower risk of preterm birth than those with a $PM_{2.5}$ exposure of < 10 μg/m^3, respectively (adj*OR* = .450, 95% CI: .098; 2.053 and adj*OR* = .891, 95% CI: .657; 1.210). The odds ratios were not statistically significant as all 95% CIs contained the null value of 1.0. During the 3rd week before childbirth, mothers with a $PM_{2.5}$ exposure of 10-24.99 μg/m^3 had a 68.9% lower risk of preterm birth than those with a $PM_{2.5}$ exposure of < 10 μg/m^3, whereas mothers with a $PM_{2.5}$ exposure of ≥ 25 μg/m^3 had 1.222 times higher risk of preterm birth than those with a $PM_{2.5}$ exposure of < 10 μg/m^3 (adj*OR* = .311, 95% CI: .040; 2.432 and adj*OR* = 1.222, 95% CI: .897; 1.663). The odds ratios were not statistically significant as all 95% CIs contained the null value of 1.0. During the 2nd week before childbirth, mothers with a $PM_{2.5}$ exposure of 10-24.99 and ≥ 25 μg/m^3 had 1.02 and 1.276 times higher risk of preterm birth than those with a $PM_{2.5}$ exposure of < 10 μg/m^3, respectively (adj*OR* = 1.020, 95% CI: .280; 3.716 and adj*OR* = 1.276, 95% CI: .935; 1.741). The risk estimates were not statistically significant as all 95% CIs contained the null value of 1.0. There was no significant association between categorized exposure and preterm birth for all exposure windows.

Table 11

Adjusted ORs and 95% CIs for the Effect of Maternal $PM_{2.5}$ Exposure on Preterm Birth

Exposure periods	*n*	Preterm births (%)	Adj. OR	95% CI
The 4th week before childbirth				
6.00-9.99	188	2 (1.1%)	Referent	
10.00-24.99	11,447	289 (2.5%)	.450	.098, 2.053
25.00-66.38	7,736	205 (2.6%)	.891	.657, 1.210
Per 10 μg/m^3 [a]	19,371	496 (2.6%)	3.335	.495, 22.479
The 3rd week before childbirth				
6.00-9.99	191	1 (0.5%)	Referent	
10.00-24.99	11,468	306 (2.7%)	.311	.040, 2.432
25.00-66.38	7,712	189 (2.5%)	1.222	.897, 1.663
Per 10 μg/m^3 [a]	19,371	496 (2.6%)	6.520	1.061, 40.074
The 2nd week before childbirth				
6.00-9.99	181	3 (1.7%)	Referent	
10.00-24.99	11,767	324 (2.8%)	1.020	.280, 3.716
25.00-66.38	7,423	169 (2.3%)	1.276	.935, 1.741
Per 10 μg/m^3 [a]	19,371	496 (2.6%)	4.657	.723, 29.979
The 1st week before childbirth				
6.00-9.99	208	4 (1.9%)	Referent	
10.00-24.99	11,728	307 (2.6%)	.981	.305, 3.156
25.00-66.38	7,435	185 (2.5%)	1.112	.819, 1.511
Per 10 μg/m^3 [a]	19,371	496 (2.6%)	5.337	.738, 38.577
Last 4 weeks of pregnancy				
10.76-24.99	12,071	313 (2.6%)	Referent	
25.00-56.10	7,300	183 (2.5%)	.855	.619, 1.181
Per 10 μg/m^3 [a]	19,371	496 (2.6%)	.002	.000, 1.257

Note. Adjusted for coal power generation and infant birth weight; OR: Odds ratio; $PM_{2.5}$ exposure was measured as continuous data by averaging exposure over each exposure window, and categorical data by categorizing the average exposure to more or less than 10 and 25 μg/m^3 for each exposure window; [a]: The data was logarithmically transformed.

I evaluated the association between variance exposures and preterm birth using multiple logistic regression analysis according to the same principles as assessing the association between average exposure and preterm birth. Table 12 shows covariates selected for the regression model. Infant birth weight and infant gender were adjusted for in the model because it contributed to a significant change of the variance of the model, and coal power generation was controlled for in the model because it changed the odds of preterm birth for $PM_{2.5}$ by more than 2-fold.

Table 12

Covariates to be Included in the Multiple Logistic Regression Model

Covariates	Cox & Snell R^2	Min.- Max. ratios of change in the ORs for $PM_{2.5}$ [a]
None (Referent)	.000	Referent
Coal power generation	.000	.79-.89
Proximity to coal plants	.000	.92-.98
Unemployment rate	.000	.99-1.00
Duration of cohabitation	.000	.99-1.00
Infant birth weight	.022	.84-1.16
Infant gender	.002	.96-1.05
Maternal age	.001	1.00-1.03
Maternal education level	.000	.99-1.00
Maternal marital status	.001	.98-1.00
Maternal occupation status	.000	.99-1.01

Note. $p < 0.05$ indicates significant association; Acronyms: OR: Odds ratio; [a]: The data were logarithmically transformed.

The multiple logistic regression analysis model consisted of dichotomous preterm birth as a dependent variable and variance in $PM_{2.5}$ exposures as independent variables, and covariates such as coal power generation, infant birth weight, and infant gender. Maternal $PM_{2.5}$ exposure was measured as the variance of exposure over the 1st, 2nd, 3rd, 4th weeks before childbirth and the last 4 weeks of pregnancy. Maternal $PM_{2.5}$ exposure was modeled as a continuous variable. Testing for the assumption of the goodness of fit was performed using the Hosmer-Lemeshow test. I found a good fit of the model, p = .054 and this model provided an acceptable level of prediction (see Frankfort-Nachmias, & Leon-Guerrero, 2015). This logistic regression model was statistically significant, with $\chi 2$ [9] = 495.045; $p < 0.00001$. In this model, the variance in preterm birth was 2.5 %, and 97.4% of cases were correctly classified.

Table 13 shows the adjusted ORs and 95% CIs for the associations between variance exposures and preterm birth. There was a significant protective association between the variance exposure and preterm birth during the last 4 weeks period of pregnancy (adj*OR* = .442, 95% CI: .212; .920). For a 10 $\mu g/m^3$ increase in the variance of $PM_{2.5}$ exposure over the last 4-weeks period of pregnancy, the risk of preterm birth decreased by 55.8%. The association between variance exposures and preterm birth was null in all of the one-week periods (adj*OR* = 1.138, 95% CI: .853; 1.518 for the 1st week before childbirth; adj*OR* = 1.302, 95% CI: .971; 1.745 for the 2nd week before childbirth; adj*OR* = 1.175, 95% CI: .884; 1.561 for the 3rd week before childbirth, and; adj*OR* = 1.194, 95% CI: .893; 1.598 for the 4th week before childbirth, per the variance increase by 10).

Table 13

Adjusted ORs and 95% CIs for the Effect of Maternal $PM_{2.5}$ Exposure Variance on Preterm Birth

Exposure windows	OR	95% CI
The 1st week before childbirth [a]	1.138	.853, 1.518
The 2nd week before childbirth [a]	1.302	.971, 1.745
The 3rd week before childbirth [a]	1.175	.884, 1.561
The 4th week before childbirth [a]	1.194	.893, 1.598
Last 4 weeks of pregnancy [a]	.442	.212, .920

Note. Adjusted for coal power generation, infant birth weight, and infant gender; OR: Odds ratio; [a]: The data was logarithmically transformed.

Research Question 3

I utilized a moderation analysis to determine the effects of individual-level characteristics, such as maternal age, maternal education level, and maternal occupation status, by $PM_{2.5}$ exposure on gestational age. Covariates were included in the multiple linear regression analysis when they were associated with gestational age or contributed to improving the quality of prediction of the model. I evaluated the associations between variables using Kruskal Wallis, Pearson, and Spearman correlation tests as appropriate. Specifically, I performed (a) Kruskal Wallis tests if the data did not meet the assumption of normality, (b) Pearson correlation tests when the data met the assumption of linearity, and (c) Spearman correlation tests if the data did not meet the assumption of linearity (see Frankfort-Nachmias, & Leon-Guerrero, 2015). I carried out Kruskal Wallis tests to examine the associations between gestational age and categorical variables, including duration of cohabitation, infant birth weight, infant gender, maternal age, maternal education level, maternal occupation status, and maternal marital status. I conducted Pearson correlation tests to assess the correlation between gestational age, average $PM_{2.5}$ exposure during the 4th week before childbirth, and unemployment rate. I used Spearman correlation tests to evaluate the correlation between gestational age and the remaining continuous variables.

Table 14 displays the results of these tests. Significant associations of gestational age were found in terms of infant birth weight, $\chi^2[2] = 1168.308$; $p < 0.00001$, infant gender, $\chi^2[1] = 63.301$; $p < 0.00001$, maternal age, $\chi^2[2] = 51.953$; $p < 0.00001$, maternal education level, $\chi^2[1] = 27.591$; $p < 0.00001$, and maternal occupation status, $\chi^2[1] = 11.875$; $p = .001$. The mean rank gestational age was higher as the infant birth weight increased and maternal age decreased. The mean rank gestational age was higher in female infants, in mothers with more than 12 years of education, and in employed mothers. Significant weak correlations ranging from -.015 to .018 were found between gestational age and the 4th week $PM_{2.5}$ level, proximity to coal plants, and unemployment rate. I included covariates, such as infant birth weight, infant gender, proximity to coal plants, and unemployment rate, in the multiple linear regression analysis model because they were significantly associated with gestational age. I also added covariates, including coal power generation, maternal marital status, average exposure to $PM_{2.5}$ during the 3rd week before childbirth, and variance in $PM_{2.5}$ exposure during the 4th week before childbirth, in the regression model because they contributed to improving the level of the estimate of gestational age.

Table 14

Covariates to be Adjusted in the Analysis

	p	R
Infant birth weight	.000	
Infant gender	.000	
Maternal age	.000	
Maternal education level	.000	
Maternal occupation status	.001	
Proximity to coal plants	.043	
Unemployment rate	.015	
Duration of cohabitation	.545	.277
Maternal marital status	.090	.278
Coal power generation [a]	.067	.278
Exposure periods	*$PM_{2.5}$ variance*	
Last 4 weeks of pregnancy [a]	.427	.277
The 1st week before childbirth [a]	.748	.277
The 2nd week before childbirth [a]	.756	.277
The 3rd week before childbirth [a]	.318	.277
The 4th week before childbirth [a]	.991	.278
Exposure periods	*$PM_{2.5}$ average*	
Last 4 weeks of pregnancy [a]	.433	.277
The 1st week before childbirth [a]	.947	.277
The 2nd week before childbirth [a]	.575	.277
The 3rd week before childbirth [a]	.306	.278
The 4th week before childbirth [a]	.035	

Note. Kruskal Wallis, Pearson, and Spearman correlation tests were used as appropriate; $p < 0.05$ indicates significant association; [a]: The data were logarithmically transformed.

I considered variables including coal power generation, infant birth weight, infant gender, maternal marital status, proximity to coal plants, and unemployment rate, the 4th week variance $PM_{2.5}$ exposure, and 3rd week average $PM_{2.5}$ exposure as covariates in the multiple linear regression analysis, Model 1, Table 15. Model 1 consisted of gestational age as a dependent variable and maternal age, maternal education level, and maternal occupation status, and 4th week average $PM_{2.5}$ exposure as independent variables.

In Table 15, I performed a multiple linear regression analysis with Model 1 to predict gestational age from individual and community level factors, such as maternal age, education level, occupation status, and the 4th week average $PM_{2.5}$ exposure. I adjusted for coal power generation, infant birth weight, infant gender, maternal marital status, proximity to coal plants, the 4th week variance $PM_{2.5}$ exposure, the 3rd week average $PM_{2.5}$ exposure, and unemployment rate in the model. The data of this model met the assumption of multicollinearity, VIF of 1.005 to 2.055. No evidence of undue influence correlation between the residuals was in this model, and Cook’s distance ranged from .000 to .035, and the Durbin-Watson statistic was 2.000. The level of estimates of gestational age was weak by the variables included in this model, $R = .278$ and the coefficient of determination was low at 7.7%. The model fit was good, F (12, 19,358) = 135.414, $p = .000$.

I ran another multiple linear regression using Model 2 in Table 15 to evaluate the effects of interactions of individual level factors with $PM_{2.5}$ exposure on gestational age. I included the interaction variables between individual and community level factors in this regression model. This model did not meet the assumption of multicollinearity with VIF values between 1.005 and 89.273. The VIF value of 89.273 indicates that the independent variables are highly correlated with each other (see Warner, 2013). Serious multicollinearity was not a problem in this analysis because this analysis was to predict gestational age as a combination of variables (see Warner, 2013). Cook's distance ranged from .000 to .030, and the Durbin-Watson statistic was 2.000, so there was no evidence of undue influence in this model. There was no change in the level of estimates of gestational age and the explained variability in gestational age, compared with Model 1, based on R = .278 and adjusted R^2 = .077. The overall model fit was good, F (14, 19,356) = 116.101, p = .000.

Model 1, Table 15 shows the factors that affected gestational age. A change to the older age group was associated with a decrease in gestational age of .056 weeks (B = -.056, p = .000). A .108 week increase in gestational age was associated with a change to a more educated group (B = .108, p = .000). A 10 μg/m^3 increase in $PM_{2.5}$ exposure during the 4th week before childbirth was associated with a reduction of .157 weeks in gestational age (B = -.157, p = .023). In Model 1, the variables including maternal age, maternal education level, and $PM_{2.5}$ exposure, significantly predicted gestational age. I used significant predictors of gestational age, including maternal age, maternal education level, and $PM_{2.5}$ exposure to generate interaction variables. I conducted moderation analyses with $PM_{2.5}$ exposure as the potential moderator in the relationship of gestational age with maternal age and maternal education level. Model 2, Table 15 shows the results of the moderation analyses. There were no interactions found to exist. None of the interactions were statistically significant, including those of maternal age and $PM_{2.5}$

exposure, and maternal education level and $PM_{2.5}$ exposure, $p = .702$ and $p = .535$, respectively.

Table 15

Moderation Analysis Explaining Gestational age

	Model 1		Model 2	
	B	*p*	B	*p*
Individual level				
Maternal age	-.056	.000	-.006	962
Maternal education	.108	.000	.223	.233
Maternal occupation status	.030	.050	.030	.050
Community level				
$PM_{2.5}$	-.157	.023	-.060	.685
Interaction				
Maternal age × $PM_{2.5}$ [a]			-.037	.702
Maternal education × $PM_{2.5}$ [a]			-.085	.535

Note. Adjusted for coal power generation, the 4th week variance $PM_{2.5}$ exposure, infant birth weight, infant gender, maternal marital status, proximity to coal plants, the 3rd week average $PM_{2.5}$ exposure, and unemployment rate; $p < 0.05$ indicates significant association; [a]: The data were logarithmically transformed.

Research Question 4

I ran mediation analyses to evaluate what indirect effect if any, sociodemographic factors, such as maternal age, education level, and occupation status would have on gestational age via $PM_{2.5}$ exposure using the PROCESS tool developed by Andrew F. Hayes (Hayes, 2012). In Figure 6, statistical significance for both “a” and “b” is a prerequisite for using the PROCESS tool to calculate the indirect effect of the predictor on the outcome via a mediator (Hayes, 2012). According to the Preacher and Hayes approach, I measured the indirect effect of the predictor on the outcome via the mediator by multiplying a and b, "*ab*" (Hayes, 2012). In Figure 6, "a" represents the bi-variate effect of the predictor on the mediator, and "b" means the effect of the mediator on the outcome after controlling for the predictor (Hayes, 2012).

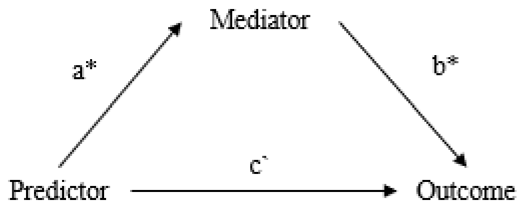

Figure 6. DAG analysis of testing for mediation. *: Association is significant at the 0.05 level; Acronym: DAG: Directed acyclic graph.

I used a series of bivariate regression and Kruskal Wallis tests to examine the strength and direction of the association between potential predictors, mediators, and outcomes. No sociodemographic factors, such as maternal age, maternal education level, and maternal occupation status, predicted $PM_{2.5}$ exposure significantly. This implies that the sociodemographic effects on gestational age could not be mediated by $PM_{2.5}$ exposure. I carried out mediation analyses with $PM_{2.5}$ exposure as the potential mediator in the relationship between organizational and policy level factors and gestational age, as shown in Figure 7 and 8. I used Pearson correlation tests to identify the covariates to be adjusted in each mediation analysis.

Figure 7 shows the indirect effect of the organizational level factor on gestational age through $PM_{2.5}$ exposure while adjusting for proximity coal plants. There was a significant indirect effect of unemployment rate mediated by $PM_{2.5}$ exposure in gestational age (ab = -.0068), with the 95 % confidence interval ranging from -.0114 to -.0021, statistically different from zero. Mothers living in areas with higher unemployment rates reported more $PM_{2.5}$ exposure (a = .0458, p = .0000), which led to a decrease in gestational age (b = -.1485, p = .0043). I also confirmed this significant mediation in the Sobel Test, Z = -2.84643, p = 0.00442.

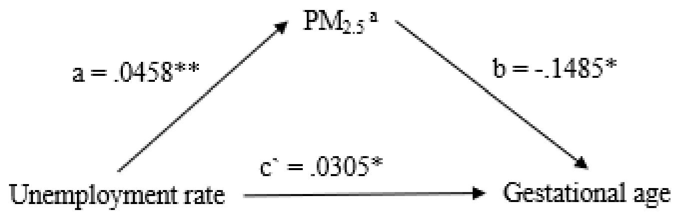

Figure 7. DAG analysis of mediation by $PM_{2.5}$ in the relationship between unemployment rate and gestational age after controlling for proximity to coal plants; *: Association is significant at the 0.05 level; **: Association is significant at the 0.01 level; [a]: The data were logarithmically transformed; Acronym: DAG: Directed acyclic graph.

Figure 8 displays the results of the mediation analysis in the PROCESS tool with $PM_{2.5}$ exposure as the potential mediator in the relationship between coal power generation and gestational age while controlling for the unemployment rate. There was a significant indirect effect of coal power generation mediated by $PM_{2.5}$ exposure in gestational age (ab = -.0470), with the 95 % confidence interval ranging from -.0855 to -.0110, statistically different from zero. The more coal power generation, the more $PM_{2.5}$ exposure increased (a = .3428, p = .0000), which translated to a reduction in gestational age (b = -.1372, p = .0082). I also confirmed this significant mediation in the Sobel Test, Z = -2.52774, p = 0.01148.

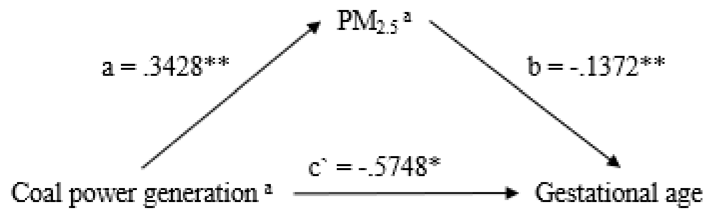

Figure 8. DAG analysis for mediation by $PM_{2.5}$ in the relationship between coal power generation and gestational age while controlling for unemployment rate; *: Association is significant at the 0.05 level; **: Association is significant at the 0.01 level; [a]: The data were logarithmically transformed; Acronym: DAG: Directed acyclic graph.

Summary

In the study population, maternal education level, maternal occupation status, and unemployment rate predicted maternal $PM_{2.5}$ exposure. The greatest influencing factor was unemployment rate at the organizational level. The change to the higher unemployment rate than the mean was associated with a 1.122 $\mu g/m^3$ increase in $PM_{2.5}$ exposure. All pregnant women in this study were exposed to $PM_{2.5}$ above the WHO safe level of 10 $\mu g/m^3$ during the last 4 weeks of pregnancy. A 10 $\mu g/m^3$ increase in maternal $PM_{2.5}$ exposure during the 3rd week before childbirth was associated with 6.52 times increase in the risk of preterm birth. When I modeled ORs with three categories of $PM_{2.5}$ exposure of < 10, 10-24.99, and ≥ 25 $\mu g/m^3$, the odds of preterm birth were not significant for all exposure windows. There was a significant protective association between $PM_{2.5}$ exposure and preterm birth when assessing exposure with the variance. For a 10 increase in the variance of $PM_{2.5}$ exposure over the last 4 weeks period of pregnancy, there was a 55.8% reduction in the risk of preterm birth. In this population, individual-level sociodemographic factors, such as maternal age, maternal education level, and maternal occupation status, did not interact with and were not mediated by

$PM_{2.5}$, and there were no indirect effects on gestational age of individual-level sociodemographic factors by $PM_{2.5}$. $PM_{2.5}$ mediated the association of gestational age with organizational and policy level factors, such as unemployment rate and coal power generation. In this population, there were harmful indirect effects on gestational age of the unemployment rate and coal power generation by $PM_{2.5}$.

This chapter described the statistical test results of the study sample. In Chapter 5, I discuss the findings of this study in terms of accordance and contrast with the literature, limitations of the study, implications of social change, and recommendations for future research. I close with a conclusion.

Chapter 5: Discussion, Conclusions, and Recommendations

Introduction

I conducted this study to determine the current best predictors of preterm birth and $PM_{2.5}$ exposure. Specifically, I examined the preterm birth risk associated with maternal $PM_{2.5}$ exposure, including the possibility of moderation and mediation. In this quantitative, correlational observation and partially ecologic study, I examined possible direct and indirect paths among the individual and ecological level variables. The retrospective cohort design of this study allowed for assessing historical exposures using secondary data. In the study, I worked to identify where to intervene to reduce differences in $PM_{2.5}$ exposure and to enhance monitoring of health disparities associated with $PM_{2.5}$. This study was performed to develop strategies to curb situations of severe air pollution and reduce the incidence of preterm birth in Korea. I sought to provide empirical evidence that would be useful in evaluating and improving the air quality standard for $PM_{2.5}$ in Korea.

In this study population, factors affecting maternal $PM_{2.5}$ exposure included maternal education level, maternal occupation status, and unemployment rate. The most influential factor was unemployment rate at the organizational level, which was eight to nine times stronger than individual-level factors, including maternal education level and maternal occupation status. A significant association between $PM_{2.5}$ exposure and preterm birth was observed in this population, specifically in the 3rd week before childbirth. A 10 $\mu g/m^3$ increase in maternal $PM_{2.5}$ exposure during the 3rd week before childbirth was associated with 6.52 times increase in the likelihood of preterm birth.

When assessing $PM_{2.5}$ exposure based on variance, I found a significant protective association between $PM_{2.5}$ exposure and preterm birth for the last 4 weeks of pregnancy. Women with a 10 $\mu g/m^3$ increase in the variance of $PM_{2.5}$ exposure had a 55.8% lower risk of preterm birth.

All pregnant women in this study were exposed to $PM_{2.5}$ above the WHO safe level of 10 $\mu g/m^3$ during the last 4 weeks of pregnancy. In risk estimates modeled with three categories of $PM_{2.5}$ exposure of < 10, 10-24.99, and ≥ 25 $\mu g/m^3$, the odds of preterm birth were not significant for all exposure windows. The odds of preterm birth were not significantly lower or higher in women with a $PM_{2.5}$ exposure of 10-24.99, and ≥ 25 $\mu g/m^3$ than women with a PM exposure of < 10 $\mu g/m^3$. No evidence of moderation and mediation by $PM_{2.5}$ exposure in the relationship between maternal age, maternal educational level, and maternal occupation status and gestational age was observed in this study population. Indirect adverse effects of the unemployment rate and coal power generation due to the mediation by $PM_{2.5}$ were detected as ab = -.0068 and ab = -.0470, respectively.

Interpretation of the Findings

Research Question 1

In this population, maternal education level, maternal occupation status, and unemployment rate significantly predicted maternal $PM_{2.5}$ exposure. Mothers who were less educated, unemployed, and of a lower socioeconomic status than the mean showed vulnerability to $PM_{2.5}$ exposure during the last 4 weeks of pregnancy. This evidence focuses on a variable that accounts for differences in maternal $PM_{2.5}$ exposure. When comparing the present study with previous studies, there was an overlap in the factors associated with disparities in maternal $PM_{2.5}$ exposure during pregnancy, including maternal education level and unemployment rate, but no researchers had examined a possible vulnerability due to maternal occupation status (Hannam et al., 2014, Lavigne et al., 2016, Salihu et al. 2012, Stieb et al. 2016).

Salihu et al. (2012) reported an opposite trend between $PM_{2.5}$ exposure and maternal education level in a study of 103,961 women in Florida. In this study, I found similar trends in that $PM_{2.5}$ exposure was higher in less educated women in Korea. This finding is different from those of Stieb et al. (2016) and Lavigne et al. (2016) who observed that $PM_{2.5}$ exposure increased as the proportion of women with secondary education increased in Canada. Unlike my observation in this study that an increase in $PM_{2.5}$ exposure is associated with an increasing unemployment rate, Salihu et al. (2012) reported no significant association between unemployment rate and maternal $PM_{2.5}$ exposure. Salihu et al. (2012) reported a significant difference in $PM_{2.5}$ exposure according to the marital status of the mother in the United States, and Hannam et al. (2014) found that $PM_{2.5}$ exposure was higher in young British mothers. In this study, there was no significant difference in $PM_{2.5}$ exposure among Korean women according to maternal age and maternal marital status.

The inconsistency between previous studies and this study might be due to differences in co-existing confounders and effect modifiers across studies (Hannam et al., 2014, Lavigne et al., 2016, Salihu et al. 2012, Stieb et al. 2016). Coexisting confounders and effect modifiers varied from population to population based on race, economic development, energy policy, and air pollution (Braveman et al., 2015, Erickson et al., 2016, Yorifuji et al., 2013). Further, these factors were highly correlated each other (Braveman et al., 2015, Erickson et al., 2016, Yorifuji et al., 2013). Many studies demonstrated that the population of low-income countries was associated with low socioeconomic status, a specific race, more existing diseases, and co-exposure to multiple air pollutants (Braveman et al., 2015, Erickson et al., 2016, Hao et al., 2016, Lavigne et al. 2016, Yorifuji et al., 2013). Residual confounding and unmeasured effect modification might lead to misinterpretation of the effect of multiple predictor variables on $PM_{2.5}$ exposure.

Given that I modeled variables according to SEM, the results of this study should be interpreted in the context of SEM. This study indicated that the most influential variable for $PM_{2.5}$ exposure existed at the organizational level. The unemployment rate at the organizational level (B = .050) indicated eight to nine times stronger effects on $PM_{2.5}$ exposure than individual level factors such as maternal education level and maternal occupation status (B = -.007 and -.006). The results of this study are difficult to compare with other published results due to differences in the statistical tests used. In other studies, chi-square, Mann–Whitney U, and two independent sample *t* tests were used to examine whether there were differences between exposed and unexposed groups according to variables (Hannam et al., 2014, Lavigne et al., 2016, Salihu et al. 2012, Stieb et al. 2016). These statistical tests were unable to analyze the strength and direction of the effects of the variables on $PM_{2.5}$ exposure (Frankfort-Nachmias, & Leon-Guerrero, 2015). The criteria for dividing exposed and unexposed groups varied from study to study, contributing to discrepancies in results (Frankfort-Nachmias, & Leon-Guerrero, 2015).

My interpretation of this study's results in the context of SEM showed that intervention strategies at the organizational level might be more effective at reducing maternal $PM_{2.5}$ exposure than changes in sociodemographic factors at the individual level. Given the results of this study, it could be said that socioeconomic disparities in $PM_{2.5}$ exposure would exist in Korea as the difference in $PM_{2.5}$ exposure in Korea was largely attributable to socioeconomic factors, including the unemployment rate at the organizational level. These results could provide pertinent information for intervention strategy to protect groups vulnerable to $PM_{2.5}$ exposure including where to intervene effectively.

Research Question 2

This study's findings indicated that maternal $PM_{2.5}$ exposure significantly affected the risk of preterm birth. In the United States and Canada, there was a wealth of evidence showing a significant association between $PM_{2.5}$ and preterm birth (DeFranco et al., 2016, Ha et al., 2014, Hao et al., 2016, Kloog et al., 2012, Laurent et al., 2016, Lee et al., 2013, Stieb et al., 2016, Symanski et al, 2014, Trasande et al., 2013). In studies from the United States, researchers observed positive associations between $PM_{2.5}$ and preterm birth. The adjusted odds ratio for preterm birth was 1.026 in a study from Florida (95% CI: 1.012; 1.039 per a 2.8 $\mu g/m^3$ increase in $PM_{2.5}$ for the 3rd trimester); 1.011 in a study from Georgia (95% CI: 1.006; 1.017 per a 2.01 $\mu g/m^3$ increase in $PM_{2.5}$ for the total pregnancy); 1.06 in a study from Massachusetts (95% CI: 1.01; 1.13 per a 10 $\mu g/m^3$

increase in $PM_{2.5}$ for the total pregnancy); 1.133 in a study from California (95% CI: 1.118; 1.148 per an IQR of 6.45 μg/m^3 increase in $PM_{2.5}$); 1.10 in a study from Pittsburgh (95% CI: 1.01; 1.20 per a 4.0 μg/m^3 increase in $PM_{2.5}$ for the 1st trimester); and 1.30 in a study from Texas (95% CI: 1.17; 1.45 per a 10 μg/m^3 increase in $PM_{2.5}$ for the 3rd trimester; Ha et al., 2014, Hao et al., 2016, Kloog et al., 2012, Laurent et al., 2016, Lee et al., 2013, Symanski et al, 2014). DeFranco et al. (2016) reported that pregnant women with high $PM_{2.5}$ exposures > 15 μg/m^3 had a 1.19-fold greater risk of preterm birth (adj*OR* = 1.19; 95 % CI: 1.09; 1.30).

Similar to the studies mentioned above, this study also confirmed a significant odds ratio of preterm birth, which was greater than 1.0 (adj*OR* = 6.520, 95 % CI: 1.061; 40.074 per a 10 μg/m^3 increase in $PM_{2.5}$ for the 3rd week before childbirth). I further observed a positive association between $PM_{2.5}$ exposure and preterm birth by modeling $PM_{2.5}$ exposure with three categories of < 10, 10-24.99, and ≥ 25 μg/m^3. Although the odds of preterm birth were not statistically significant in all exposure categories, women in the higher exposure category tended to have a higher *OR* of preterm birth than women in the lower exposure category. The higher the exposure category, the higher the risk of preterm birth for Korean women.

Although all studies included in the literature review hypothesized the deleterious effects of $PM_{2.5}$ on preterm birth, some researchers observed that maternal $PM_{2.5}$ exposure had a protective effect on preterm birth (Stieb et al., 2016 & Trasande et al., 2013). Stieb et al. (2016), in a Canadian study, reported an adjusted odds ratio of preterm birth of .96 (95% CI: 0.93; 0.99 per a 10 μg/m^3 increase in $PM_{2.5}$). Risk estimates for preterm birth in Canadian women were reduced by 4% for every 10 μg/m^3 increase in $PM_{2.5}$ exposure. Trasande et al. (2013), in a study in the United States with a birth rate of 26.9%, found that the risk of preterm birth decreased 3% per each one μg/m^3 increase in averaged $PM_{2.5}$ over the month of birth (adj*OR* = 0.97, 95% CI: 0.96; 0.98).

Some researchers reported no significant link between $PM_{2.5}$ and preterm birth, although many researchers, including myself, have reported that an odds ratio of preterm birth greater than 1.0 (Fleischer et al., 2014, Hannam et al., 2014, Hyder et al., 2014, Johnson et al., 2016, Salihu et al., 2012). In a study using World Health Organization Global Survey database from various regions of the world including Africa, Asia, and Latin America, Fleischer et al. (2014) found null results in all countries except China because the 95% CI for the adjusted odds ratio of preterm birth included a null value of 1.0. In studies from the United Kingdom and the United States, the adjusted odds ratios for preterm birth were not significant based on the 95% CIs of 0.85-1.12 for the 4th quartile in $PM_{2.5}$ estimates using a spatio-temporal (S-T) air pollution model by quartiles, .98-1.01 with 2.41 μg/m^3 increase in $PM_{2.5}$ based on satellite-based exposure estimates, .90-1.05 with $PM_{2.5}$ exposure for the 1st trimester, .98–1.07 in a single pollutant model of $PM_{2.5}$ (Hannam et al., 2014, Hyder et al., 2014, Johnson et al. 2016; Salihu et al. 2012).

The main reason for the inconsistency mentioned above might be the difference in $PM_{2.5}$ exposure estimation modeling for each study. Specifically, ground-based measurements had a greater potential for misclassification leading to the null association compared to satellite-based measurements (Hyder 2014). The range of $PM_{2.5}$ levels varied across studies. Analysis of low $PM_{2.5}$ levels was more likely to indicate null-biased associations than high levels (Fleischer et al., 2014, Hao et al., 2016; Laurent et al., 2016; Ha et al., 2014; Stieb et al., 2016; Trasande et al., 2013). Researchers modeled $PM_{2.5}$ exposure estimates as exposure assessments at different time points during pregnancy, including pre-, early-, late-, and whole pregnancy. The fetal development period relevant to $PM_{2.5}$ was early in pregnancy and before birth (Nachman et al., 2016, Symanski et al., 2014, Trasande et al., 2013). Exposure estimation modeling with less relevant fetal development periods might mislead the results toward the null in the analyses. Another possibility of exposure estimation modeling differences for each study was that the length of the exposure assessment period varied from study to study. The shorter the measurement period, the greater the chance of measurement errors in exposure estimates as the variation in exposure increased (Rappazzo et al., 2015). Analysis of narrow exposure windows was more likely to mislead results. Another reason for the inconsistency mentioned above could be that the average concentration of $PM_{2.5}$ might be neither a critical factor leading to adverse outcomes or the best measure assessing the effect of $PM_{2.5}$. Rather, other aspects of $PM_{2.5}$ exposure, including the composition of $PM_{2.5}$ toxic chemicals, the number of $PM_{2.5}$ species, and the variance of $PM_{2.5}$, might be involved in pathways leading to deleterious consequences (DeFranco et al. 2016, Hao et al., 2016, Lee et al., 2013, Miliku et al., 2016, Rappazzo et al., 2015, Stokkeland et al., 2016, Wang et al., 2016).

By examining more sophisticated exposure windows, including the one-week period of pregnancy, in this study, I found various effects at each refined exposure period. In the analyses of different lengths of exposure, regardless of statistical significance, the increase in 1-week exposure tended to be proportional to the risk of preterm birth, whereas the increase in 4-week exposure showed an opposing trend to preterm birth risk. In the analyses, I found statistical significance between preterm birth and $PM_{2.5}$ exposure for the 3rd week before childbirth, but not during the other exposure periods. These findings are consistent with the results of Hannam et al. (2014) and Kloog et al. (2012) who observed a null association between $PM_{2.5}$ and preterm birth for exposure during the last month and 30 days prior to childbirth, respectively. In a study of pregnant women residing in Harris County, Texas, Symanski et al. (2014) presented results for every 4-week period of pregnancy. $PM_{2.5}$ exposure during the last 4-week period of pregnancy showed significant odds ratios greater than 1.0 for preterm birth (Symanski et al., 2014). Trasande et al. (2013) assessed $PM_{2.5}$ exposure by averaging $PM_{2.5}$ concentrations obtained from the closest monitor at the birth hospital over the month of birth. $PM_{2.5}$ exposure at the month of birth showed significant odds ratios less than 1.0 for preterm birth (Trasande et al., 2013).

Few researchers investigated the association between $PM_{2.5}$ and preterm birth by evaluating weekly exposure or the variance in exposure. In this study, I found a significant odds ratio of 6.520 (95% CI = 1.061; 40.074) per 10 $\mu g/m^3$ increase in $PM_{2.5}$ over the 3rd week before childbirth. These findings indicated that the most vulnerable period for preterm birth was the 3rd week before childbirth. Increased $PM_{2.5}$ exposure by 10 $\mu g/m^3$ during that exposure window of vulnerability increased the risk of preterm birth by 6.52 times. By examining the variance in exposure, in this study, I found the effect of the variance in exposure on preterm birth. The result for the last 4 weeks of pregnancy showed a significant protective association between the variance in exposure and preterm birth (adj*OR* = .442, 95% CI: .212; .920). I interpreted this finding as suggesting that subsequent $PM_{2.5}$ exposure should be minimized when exposed to high levels of $PM_{2.5}$ to reduce the risk of preterm birth.

Research Question 3 and 4

In this study, I examined maternal $PM_{2.5}$ exposure as a moderator to the association between sociodemographic factors and gestational age in Korean women. In the analysis of this population, I found that gestational age was independently associated with maternal sociodemographic characteristics and $PM_{2.5}$ exposure but not with interactions. Previous researchers have examined sociodemographic factors as effect modifiers to the association between $PM_{2.5}$ and preterm birth. In a Canadian study, Lavigne et al. (2016) observed that the magnitude of the association between $PM_{2.5}$ and preterm birth varied according to maternal diabetes status. The results of a Japanese study indicated that there was a difference in the relationship between air pollution exposure and preterm birth when stratified by maternal diabetes and hypertension status (Yorifuji et al., 2013). Hao et al. (2016), in their sample of 511,658 births from Georgia, reported that the effect modification of $PM_{2.5}$ and preterm birth by maternal education level was evident in that the association was strongest in mothers who were less educated.

In this study, I examined whether maternal $PM_{2.5}$ exposure mediated the sociodemographic effects on gestational age. In the analysis of this population, the sociodemographic effects on gestational age were unable to be mediated by $PM_{2.5}$ exposure as maternal sociodemographic factors were not directly associated with maternal $PM_{2.5}$ exposure. In the mediation analysis of this population, I found that coal power generation and unemployment rate had an indirect adverse effect on gestational age by the presence of maternal $PM_{2.5}$ exposure as a mediator. Based on my recent literature review, I found no study of possible mediation by $PM_{2.5}$ in the relationship between sociodemographic factors and gestational age.

I generated Research Questions 3 and 4 of this study based on Rothman's sufficient component cause model. Given such, the analysis results of Research Questions 3 and 4 should be interpreted in the context of Rothman's sufficient component causal model. In this population, maternal $PM_{2.5}$ exposure at the community level was not a significant complementary component cause for preterm birth of individual-level sociodemographic factors, such as maternal age, maternal education level, and maternal occupation status. The mediation analysis indicated that maternal $PM_{2.5}$ exposure at the community level was a significant complementary component cause for preterm birth of organizational and policy level factors, such as unemployment rate and coal power generation. In this data, Korean women were more vulnerable to the effects of unemployment and coal development on gestational age in the presence of maternal $PM_{2.5}$ exposure.

The results of the moderation and mediation analysis indicated that maternal $PM_{2.5}$ exposure did not modify the effects of sociodemographic factors, such as maternal age, maternal education level, and maternal occupation status, on gestational age in Korean women. This study could provide evidence that maternal $PM_{2.5}$ exposure was not part of the mechanism by which sociodemographic factors adversely affected gestational age.

Limitations of the Study

There are some limitations in this study. In this work, I used secondary data from the National Vital Statistics database that provided limited data on individual characteristics. I could not control for certain confounders, such as maternal smoking status, alcohol use, and gestational complications in the analysis (Erickson et al., 2016; Goossens et al., 2016; Leal et al., 2016). The residual confounding might be plausible in this study, but it could be minimal in Korea because of the low rate of smoking, drinking, and complications during pregnancy. The rate of smoking, drinking, and complications during pregnancy were 0.4%, 1.6%, and 2.5%, respectively, in households with children from 0 to 2 years of age, based on the 2013 Korean Survey of Child conducted every five years by the Ministry of Health, Welfare, and Family Affairs (KOSIS, 2015).

The availability of a valid $PM_{2.5}$ monitoring station was limited to 12 of the 17 top-tier administrative districts in Korea (KME, 2016). Not all top-tier administrative districts proceeded with additional sampling, of which 12 were selected for cluster sampling. The selected final cluster consisted of 8 GUs of urban areas and 4 SIs of major cities. Rural areas were not included in this study, which could lead to selection bias. The findings of this study should be applied only to pregnant women residing in urban areas and major cities.

This study has a possibility of the Ecologic fallacy, given that I assessed some variables, such as maternal $PM_{2.5}$ exposure, unemployment rate, and coal power generation, based on area-level estimations. I calculated these variables based on the area level address, including the sub-tier administrative district and top-tier administrative district where the pregnant woman resided, not the home address. These variables, measured at the area level, might not precisely quantify or represent individual-level measures based on home addresses (Szklo, & Nieto, 2014). To address the Ecologic fallacy, I modeled the variables of this study using SEM as a conceptual framework. I classified each variable as one of the area levels, including community, organizational, and policy levels, based on the SEM. This conceptual framework was useful in reducing the chances of improperly assigning aggregate data to individual participants (Greenland, 2001). In this study, I was able to investigate the area level impact of $PM_{2.5}$, unemployment rate, and coal-power on preterm birth.

In this study, I assumed that unemployment rate and coal power generation remained constant over the same month. I also assumed that there was no change in the place of residence and occupation of the pregnant woman for the last 4 weeks period of pregnancy before delivery. These assumptions could lead to non-differential misclassification and nullify the results if the unemployment rate and coal power generation fluctuated in the same month, or the place of residence and occupation of the pregnant woman changed for the last 4 weeks period of pregnancy. Data on the unemployment rate and coal power generation had aggregate attributes (KOSIS, 2016 & KEPCO, 2016). For data on the unemployment rate and coal power generation, monthly data was the most common data aggregation period and had the shortest measurement period among the published data (KOSIS, 2016 & KEPCO, 2016). Some researchers who assessed the likelihood of misclassification due to residential mobility during pregnancy reported that the lowest mobility was seen at the 3rd trimester or late pregnancy and that most residential movements occurred in the same exposure area over a short distance (Bell, & Belanger, 2012, & Chen, Bell, Caton, Druschel, & Lin, 2010).

Recommendations

In this work, I assessed maternal $PM_{2.5}$ exposure at the community level based on SEM. A difference exists inherently between community-level exposure and individual-level exposure (Chang et al., 2015, Hyder et al., 2014, Szklo, & Nieto, 2014). In further work, researchers should measure maternal $PM_{2.5}$ exposure based on the $PM_{2.5}$ concentration of an individual's blood sample instead of using secondary data, including ground-based or satellite-based measurements, to investigate the association between individual-level exposure and preterm birth. The analysis of individual-level exposure based on blood $PM_{2.5}$ concentration may be beneficial in mitigating possible measurement errors imposed on any of the exposure estimation models.

In this investigation, I examined the association between $PM_{2.5}$ and preterm birth based on assessing the average and variance of $PM_{2.5}$ concentrations. In most studies, the researchers also evaluated exposure by averaging $PM_{2.5}$ measurements. In the published studies, the association between $PM_{2.5}$ and preterm birth was inconsistent, which contributed to underestimating the harmful effects of $PM_{2.5}$ on preterm birth (Fleischer et al., 2014, Hao et al., 2016, Hyder et al., 2014, Johnson et al., 2016, Laurent et al., 2016). A possible reason for the inconsistent results may be that the average concentration of $PM_{2.5}$ might be neither a critical factor leading to adverse outcomes or the best measure assessing the effect of $PM_{2.5}$. This reason marks the need for further analysis to evaluate maternal $PM_{2.5}$ exposure based on the new factor or new measurements, taking into account other aspects of $PM_{2.5}$ exposure that may be involved in pathways leading to deleterious consequences including the composition of $PM_{2.5}$ toxic chemicals, the number of $PM_{2.5}$ species and the variance of $PM_{2.5}$ (DeFranco et al. 2016, Hao et al., 2016, Lee et al., 2013, Miliku et al., 2016, Rappazzo et al., 2015, Stokkeland et al., 2016, Wang et al., 2016). The analysis of new measurements may be helpful in identifying certain aspects of $PM_{2.5}$ that are detrimental to preterm birth.

Researchers suggested that maternal $PM_{2.5}$ exposure during pregnancy was associated with the increase in oxidative stress and intrauterine inflammation, which triggered pathways leading to preterm birth (Ha et al., 2014; Lee at al., 2013; Menon, 2014). It is informative to identify potential biomarkers associated with preterm birth, including maternal medical conditions and inflammatory markers (DeFranco et al. 2016, Ha et al., 2016, Lee et al., 2013, Miliku et al., 2016, Stokkeland et al., 2016, Wang et al., 2016). Inflammatory markers of preterm birth could offer hypotheses for further analysis to investigate the association between $PM_{2.5}$ and specific inflammatory markers. The analysis of biomarkers would allow researchers to reveal the mechanism by which inflammation induced by $PM_{2.5}$ increased the risk of preterm birth.

In this study, I assessed maternal $PM_{2.5}$ exposure only for the last 4 weeks of pregnancy to examine the association of preterm birth with $PM_{2.5}$. In further investigation of Korean women, the researchers should measure $PM_{2.5}$ exposure for each trimester of pregnancy and entire pregnancy, to provide more comparable results to other published studies. In this investigation, I evaluated maternal $PM_{2.5}$ exposure as a refined exposure window only for the last 4 weeks of pregnancy to examine the association of preterm birth with $PM_{2.5}$. Additional analysis of refined period exposures in early pregnancy may be informative, to identify the critical timing of fetal development associated with $PM_{2.5}$ in early pregnancy.

In this analysis, the study sample members were selected from residents of 8 GUs in urban areas and residents of 4 SIs in major cities with high population density. Pregnant women living in rural areas with low population density were not selected in this study. The population density was correlated with area-level socioeconomic characteristics including the unemployment rate and air pollution (Grantz et al., 2016 & Pinault, Crouse, Jerrett, Brauer, & Tjepkema, 2016). Analysis of high population density areas may differ from that of low population density areas. This study of residents in highly populated areas was unable to offer adequate information on intervention strategies to protect population groups vulnerable to the risk induced by $PM_{2.5}$ living in less populated areas. One way to address selection bias would be to examine pregnant women after completion of $PM_{2.5}$ monitor installation throughout the country, including in rural areas. Additional analysis after installing $PM_{2.5}$ monitors nationwide may be beneficial to provide generalizable results for Korean pregnant women.

Implications

Positive social change potentials included in this study are at both the individual and social levels. In this study, maternal $PM_{2.5}$ exposure increased the risk of preterm birth. The finding of this study could lead to an increase in people's perceptions of threats from $PM_{2.5}$ exposure, followed by changes in their health-related behavior. People, especially pregnant women, could be motivated to avoid $PM_{2.5}$ exposure by wearing masks, using air purifiers, and limiting outdoor activities when there is an air pollution warning (Air Korea.or.kr, 2016, & Malmqvist et al., 2013). Evidence from this study showed that subsequent $PM_{2.5}$ exposure should be minimized when exposed to high levels of $PM_{2.5,}$ based on that the lower the risk of preterm birth as the variance in exposure increased. This information could be beneficial to people by providing specific action guidelines to mitigate harmful health impact of $PM_{2.5}$.

The SEM adopted as a conceptual framework of this study helped to examine the factors at various levels, such as individual and ecological levels simultaneously. Unlike most other studies focused on whether there was a difference between exposed and unexposed groups according to the variables, this study focused on significant predictors of $PM_{2.5}$ exposure, including direction and strength of the effect, using multiple linear regression analysis instead of the chi-square and Mann-U tests. The results of this study indicated that the difference in $PM_{2.5}$ exposure at the organizational level by socioeconomic status, including unemployment rate, was significantly greater than that by individual-level factors. The study's findings mark the need for the need for socioeconomic interventions to reduce exposure to environmental hazards beyond individual control, minimize group differences in exposure, and improve monitoring of health imbalances associated with environmental pollutants.

The observation that all pregnant women in this study were exposed to $PM_{2.5}$ above the WHO safe level of 10 $\mu g/m^3$ during the last 4 weeks of pregnancy demonstrated severe air pollution in Korea. The analysis of this study indicated that the most influential factor for $PM_{2.5}$ exposure was unemployment rate at the organizational level and there were the socioeconomic disparities in $PM_{2.5}$ exposure in Korea. This study provided evidence that intervention strategies at the organizational level could be more effective at reducing maternal $PM_{2.5}$ exposure than changes in sociodemographic factors at the individual level. The insights gained from this investigation could be useful for environmental professionals and policymakers to establish a more effective strategy to improve the $PM_{2.5}$ monitoring system in Korea and minimize differences in $PM_{2.5}$ exposure. This information could help public health professionals enhance monitoring of health disparities associated with $PM_{2.5}$.

This work was the first effort to investigate the effect of $PM_{2.5}$ on preterm birth since the beginning of air quality monitoring of $PM_{2.5}$ in Korea. This could provide first scholarly evidence that would be useful to evaluate and improve the current air quality standard for $PM_{2.5}$ in Korea. By modeling $PM_{2.5}$ exposure with three categories of < 10, 10-24.99, and ≥ 25 $\mu g/m^3$ based on WHO and KME guidelines, I determined whether maternal $PM_{2.5}$ exposure of 10-24.99 or ≥ 25 $\mu g/m^3$ increased the risk of preterm birth compared to the exposure of < 10 $\mu g/m^3$. Although the odds of preterm birth were not statistically significant in all exposure categories, women in the higher exposure category tended to have a higher OR of preterm birth than women in the lower exposure category. The higher the exposure category, the higher the risk of preterm birth for Korean women. The observation of this study could support evidence-based decisions to lower the $PM_{2.5}$ exposure limit and guide a more stringent safety level for $PM_{2.5}$ exposure in Korea.

This investigation is important because maternal $PM_{2.5}$ exposure had adverse effects on fetal development in Korea. Attempts to undermine the rise in preterm birth rates have not yet been successful, given that the preterm birth rate in 2015 increased by 44 % compared to a decade ago (KOISIS, 2016). A possible reason for the little progress in reducing preterm birth rates in Korea could be the lack of understanding of the preterm birth risk due to the emerging health threat, $PM_{2.5}$ based on the results of this study. The results of this study are informative for public health professionals to develop more effective strategies to lower preterm birth rates in Korea. This study provided evidence of the association between $PM_{2.5}$ and preterm birth in a middle-income country. This study could be useful for further meta-analysis to mitigate selection bias by including countries at all stages of development, from low-income countries to high-income countries given that there was little research on middle-income countries including Korea.

This investigation was the first attempt to determine the timing of fetal development associated with $PM_{2.5}$ by examining short-term exposures of 1 week and 4 weeks. This study's results showed a specific effect on the preterm birth of each one-week exposure for the last 4 weeks period of pregnancy. This information is actionable as it specifically pointed out the 3rd week before childbirth as a critical fetal development period relevant to $PM_{2.5}$. This investigation would be beneficial for researchers looking for background information on associations of adverse birth outcomes by assessing exposure to air pollution based on sophisticated exposure windows.

The analysis of this study was one of the few trials to assess the indirect impact of sociodemographic factors on gestational age by considering $PM_{2.5}$ as a moderator and mediator. In this study, there was no significant indirect sociodemographic impact on gestational age by $PM_{2.5}$, suggesting that $PM_{2.5}$ was not part of the mechanism by which sociodemographic factors adversely affected gestational age. This evidence would be useful to researchers investigating the possible pathways by which $PM_{2.5}$ increases the risk of preterm birth.

This study's findings showed harmful indirect effects on gestational age of the organizational and policy level factors by $PM_{2.5}$. $PM_{2.5}$ in which socioeconomic and energy policy factors such as unemployment rate and coal power generation passed through to affect adverse birth outcomes might lead to greater vulnerability to socioeconomic and energy policy effects on preterm birth risks. Pregnant women in Korea were more vulnerable to the effects of unemployment and coal development on gestational age in the presence of $PM_{2.5}$. Mothers with a high unemployment rate and high coal output were more likely to get adverse consequences for fetal development including preterm birth. Intervention strategies at the organizational and policy level are warranted to mitigate the adverse consequences of air pollution on pregnancy outcomes based on the finding of this study. Insights from this investigation suggested that prenatal care should be addressed in socioeconomic and energy policy sensitive approaches (Braveman et al., 2015, & Lorch, & Enlow, 2015). These findings could be useful information for health care providers, health insurance organizations, and policymakers to educate pregnant women and their caregivers, modify health care and insurance systems according to air pollution levels, socioeconomic status, and corresponding health care needs, and improve emergency preparedness and response to air pollution.

Conclusion

In this study population, the most influential factor for PM2.5 exposure was unemployment rate at the organizational level. These findings show the need for socioeconomic interventions to reduce $PM_{2.5}$ exposure more effectively. This study's results indicated the lower the risk of preterm birth as the variance in exposure increased, or as the average in exposure decreased. According to this study's findings, the more critical fetal development period relevant to $PM_{2.5}$ was the 3rd week before childbirth. In this study population, $PM_{2.5}$ was not part of the mechanism by which sociodemographic factors adversely affected gestational age, but it was part of the path passed through by socioeconomic and energy policy factors such as unemployment rate and coal power generation to affect adverse birth outcomes. These findings indicate that prenatal care should be addressed with a socioeconomic- and energy-policy-sensitive approach to lower preterm birth due to severe air pollution in Korea.

References

Air Korea. (2013.). *Air Korea information*. Retrieved from http://www.airkorea.or.kr/airinfo

Aschengrau, A., & Seage, G. R., III. (2014). *Essentials of epidemiology in public health*. Burlington, MA: Jones & Bartlett.

Averett, S. L., & Fletcher, E. K. (2016). Prepregnancy obesity and birth outcomes. *Maternal and Child Health Journal, 20*(3), 655-664. doi:10.1007/s10995-015-1865-0

Battelle Memorial Institute, and Center for International Earth Science Information Network. (2013). *Global annual average $PM_{2.5}$ grids from MODIS and MISR aerosol optical depth (AOD), 2001-2010*. Retrieved from http://sedac.ciesin.columbia.edu/data/set/sdei-global-annual-avg-pm2-5-2001-2010/maps.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182. Retrieved from http://psycnet.apa.org/fulltext/1987-13085-001.html

Bell, M. L., & Belanger, K. (2012). Review of research on residential mobility during pregnancy: Consequences for assessment of prenatal environmental exposures. *Journal of Exposure Science and Environmental Epidemiology, 22*(5), 429-438. doi:10.1038/jes.2012.42

Bodnar, L. M., Abrams, B., Bertolet, M., Gernand, A. D., Parisi, S. M., Himes, K. P., & Lash, T. L. (2014). Validity of birth certificate-derived maternal weight data. *Paediatric and Perinatal Epidemiology, 28*(3), 203-212. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/ppe.12120

Brumberg, H. L., & Shah, S. I. (2015). Born early and born poor: An eco-bio-developmental model for poverty and preterm birth. *Journal of Neonatal-Perinatal Medicine, 8*(3), 179-187. doi:10.3233/NPM-15814098

Braveman, P. A., Heck, K., Egerter, S., Marchi, K. S., Dominguez, T. P., Cubbin, C. . . . Curtis, M. (2015). The role of socioeconomic factors in black–white disparities in preterm birth. *American Journal of Public Health*, *105*(4), 694-702. doi:10.2105/AJPH.2014.302008

Burkholder, G. J., Cox, K. A., & Crawford, L. M. (2016). *The scholar-practitioner's guide to research design*. Baltimore, MD: Laureate Publishing.

Buzi, R., Smith, P., Kozinetz, C., Peskin, M., & Wiemann, C. (2015). A Socioecological framework to assessing depression among pregnant teens. *Maternal & Child Health Journal, 19*(10), 2187-2194. doi:10.1007/s10995-015-1733-y

Casas, M., Cordier, S., Martínez, D., Barros, H., Bonde, J. P., Burdorf, A. . . .Fernandez, M. (2015). Maternal occupation during pregnancy, birth weight, and length of gestation: Combined analysis of 13 European birth cohorts. *Scandinavian Journal of Work, Environment & Health, 41*(4), 384-96.

Centers for Disease Control and Prevention. (CDC). (2016). *Preterm birth*. Retrieved from https://www.cdc.gov/reproductivehealth/maternalinfanthealth

/pretermbirth.htm

Cerin, E., Zhang, C. J., Barnett, A., Sit, C. H., Cheung, M. M., Johnston, J. M. . . .Lee, R. S. (2016). Associations of objectively-assessed neighborhood characteristics with older adults' total physical activity and sedentary time in an ultra-dense urban environment: findings from the ALECS study. *Health & Place, 42*, 1-10. doi:10.1016/j.healthplace.2016.08.009

Chang, H. H., Warren, J. L., Darrow, L. A., Reich, B. J., & Waller, L. A. (2015). Assessment of critical exposure and outcome windows in time-to-event analysis with application to air pollution and preterm birth study. *Biostatistics, 16*(3), 509-521. doi:10.1093/biostatistics/kxu060

Chen, Bell, Caton, Druschel, & Lin, 2010). Residential mobility during pregnancy and the potential for ambient air pollution exposure misclassification. *Environmental Research, 110*(2), 162-168. doi:10.1016/j.envres.2009.11.001

Cho, J. H., Choi, S. K., Chung, S. H., Choi, Y. S., & Bae, C. W. (2013). Changes in neonatal and perinatal vital statistics during last 5 decades in Republic of Korea: compared with OECD nations. *Neonatal Medicine, 20*(4), 402-412. doi:10.5385/nm.2013.20.4.402

Cloherty, John P. (2012). *"Care of the extremely low birth weight infant". Manual of Neonatal care* (7th ed.). Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Cutts, F. T., Izurieta, H. S., & Rhoda, D. A. (2013). Measuring coverage in MNCH: design, implementation, and interpretation challenges associated with tracking vaccination coverage using household surveys. *PLoS Medicine*, *10*(5), e1001404. doi:10.1371/journal.pmed.1001404

DeFranco, E., Moravec, W., Xu, F., Hall, E., Hossain, M., Haynes, E. N. . . . Chen, A. (2016). Exposure to airborne particulate matter during pregnancy is associated with preterm birth: A population-based cohort study. *Environmental Health: A Global Access Science Source*, 156. doi:10.1186/s12940-016-0094-3

Derraik, J. B., Lundgren, M., Cutfield, W. S., & Ahlsson, F. (2016). Maternal height and preterm birth: a Study on 192,432 Swedish women. *PLoS O, 11*(4), e0154304. doi:10.1371/journal.pone.0154304

El Rafei, R., Abbas, H. A., Charafeddine, L., Nakad, P., Al Bizri, A., Hamod, D., & Yunis, K. A. (2016). Association of pre-pregnancy body mass index and gestational weight gain with preterm births and fetal size: An Observational Study from Lebanon. *Paediatric and Perinatal Epidemiology, 30*(1), 38-45. doi:10.1111/ppe.12249

Erickson, A. C., Ostry, A., Chan, L. H., & Arbour, L. (2016). The reduction of birth weight by fine particulate matter and its modification by maternal and neighbourhood-level factors: A multilevel analysis in British Columbia, Canada. *Environmental Health, 15*(1), 51-63. doi:10.1186/s12940-016-0133-0

Erdogdu, E. (2014). Investment, security of supply and sustainability in the aftermath of three decades of power sector reform. *Renewable and Sustainable Energy Reviews, 31*, 1-8. doi:10.1016/j.rser.2013.11.014.

Ferguson, K. K., Chen, Y. H., VanderWeele, T. J., McElrath, T. F., Meeker, J. D., & Mukherjee, B. (2017). Mediation of the relationship between maternal phthalate exposure and preterm birth by oxidative stress with repeated measurements across pregnancy. *Environmental Health Perspectives, 125*(3), 488-494. doi:10.1289/EHP282

Fleischer, N. L., Merialdi, M., van Donkelaar, A., Vadillo-Ortega, F., Martin, R. V., Betran, A. P., & Souza, J. P. (2014). Outdoor air pollution, preterm birth, and low birth weight: Analysis of the world health organization global survey on maternal and perinatal health. *Environmental Health Perspectives, 122*(4), 425-430. doi:10.1289/ehp.1306837

Frankfort-Nachmias, C., & Leon-Guerrero, A. (2015). *Social statistics for a diverse society (7th ed.).* Thousand Oaks, CA: Sage Publications.

Gee, G. C., & Payne-Sturges, D. C. (2004). Environmental health disparities: a framework integrating psychosocial and environmental concepts. *Environmental Health Perspectives*, *112*(17), 1645-1653. doi:10.1289/ehp.7074

Getzen, T. E. (2013). *Health economics and financing* (5th ed.). Hoboken, NJ: John Wiley and Sons.

Giorgis-Allemand, L., Pedersen, M., Bernard, C., Aguilera, I., Beelen, R. M., Chatzi, L. . . . Estarlich, M. (2017). The influence of meteorological factors and atmospheric pollutants on the risk of preterm birth. *American Journal of Epidemiology, 185*(4), 247-258. doi:10.1093/aje/kww141

Glouberman, S., & Millar, J. (2003). Evolution of the determinants of health, health policy, and health information systems in Canada. *American Journal of Public Health, 93*(3), 388-392. doi:10.2105/AJPH.93.3.388

Goossens, J., Delbaere, I., Dhaenens, C., Willems, L., Van Hecke, A. . . . Verhaeghe, S. (2016). Preconception-related needs of reproductive-aged women. *Midwifery, 33*, 64-72. doi:10.1016/j.midw.2015.10.012

Grantz, K. H., Rane, M. S., Salje, H., Glass, G. E., Schachterle, S. E., & Cummings, D. A. (2016). Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. *Proceedings of the National Academy of Sciences, 113*(48), 13839-13844. doi:10.1073/pnas.1612838113

Ha, E. H., Lee, B. E., Park, H. S., Kim, Y. S., Kim, H., Kim, Y. J. . . . Park, E. A. (2004). Prenatal exposure to PM10 and preterm birth between 1998 and 2000 in Seoul, Korea. *Journal of Preventive Medicine and Public Health, 37*(4), 300-305. Retrieved from https://www.jpmph.org/journal/view.php?year=2004&vol=37&page=300

Ha, S., Hu, H., Roussos-Ross, D., Haidong, K., Roth, J., & Xu, X. (2014). The effects of air pollution on adverse birth outcomes. *Environmental Research, 134*, 198-204. doi:10.1016/j.envres.2014.08.002

Han, S. Y., Yoo, S. H., & Kwak, S. J. (2004). The role of the four electric power sectors in the Korean national economy: An input–output analysis. *Energy Policy, 32*(13), 1531-1543. doi:10.1016/S0301-4215(03)00125-3

Hannam, K., McNamee, R., Baker, P., Sibley, C., & Agius, R. (2014). Air pollution exposure and adverse pregnancy outcomes in a large UK birth cohort: Use of a novel spatio-temporal modelling technique. *Scandinavian Journal of Work, Environment & Health, 40*(5), 518-530. doi:10.5271/sjweh.3423

Hao, H., Chang, H. H., Holmes, H. A., Mulholland, J. A., Klein, M., Darrow, L. A., & Strickland, M. J. (2016). Air pollution and preterm birth in the U.S. state of Georgia (2002-2006): Associations with concentrations of 11 ambient air pollutants estimated by combining community multiscale air quality model (CMAQ) simulations with stationary monitor measurements. *Environmental Health Perspectives*, *124*(6), 875-880. doi:10.1289/ehp.1409651

Hao, Y., Strosnider, H., Balluz, L., & Qualters, J. R. (2016). Geographic variation in the association between ambient fine particulate matter ($PM_{2.5}$) and term low birth weight in the United States. *Environmental Health Perspectives*, *124*(2), 250-255. doi:10.1289/ehp.1408798

Hayes, A. F., & Preacher, K. J. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879-891. doi:10.3758/BRM.40.3.879

Hayes, A. F. (2012). *PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling*. Retrieved from http://www.afhayes.com/ public/process2012.pdf

He, J. R., Liu, Y., Xia, X. Y., Ma, W. J., Lin, H. L., Kan, H. D. . . . Xia, H. M. (2016). Ambient temperature and the risk of preterm birth in Guangzhou, China (2001–

2011). *Environmental Health Perspectives, 124*(7), 1100-1106. doi:10.1289/ehp.1509778

Huang, A., Ji, Z., Zhao, W., Hu, H., Yang, Q., & Chen, D. (2016). Rate of gestational weight gain and preterm birth in relation to prepregnancy body mass indices and trimester: A follow-up study in China. *Reproductive Health, 13*(1), 93-100. doi:10.1186/s12978-016-0204-2

Hwang, W. J., Kim, J. A., & Kim, H. G. (2017). Women's health status working at traditional marketplaces and their needs for public health care services. *Journal of Korean Academy of Community Health Nursing, 28*(1), 44-54. doi:10.12799/jkachn.2017.28.1.44

Hyder, A., Lee, H. J., Ebisu, K., Koutrakis, P., Belanger, K., & Bell, M. L. (2014). $PM_{2.5}$ exposure and birth outcomes: use of satellite-and monitor-based data. *Epidemiology, 25*(1), 58-67. doi:10.1097/EDE.0000000000000027

Jamasb, T., Nepal, R., & Timilsina, G. R. (2017). A quarter century effort yet to come of age: A survey of electricity sector reform in developing countries. *Energy Journal, 38*(3), 195-234. doi:10.5547/01956574.38.3.tjam

Johnson, S., Bobb, J. F., Ito, K., Savitz, D. A., Elston, B., Shmool, J. C. . . . Matte, T. (2016). Ambient fine particulate matter, nitrogen dioxide, and preterm birth in New York City. *Environmental Health Perspectives, 124*(8), 1283-1290. doi:10.1289/ehp.1510266

Ji, B. S., & Kang, S. J. (2016). An analysis on electric power policy through ACMS model: focused on the power generation subsidiary company of the Korea Electric

Power Corporation. *The Transactions of the Korean Institute of Electrical Engineers P, 65*(1), 31-40. doi:10.5370/KIEEP.2016.65.1.031

Kenny, L. C., Lavender, T., McNamee, R., O'Neill, S. M., Mills, T., & Khashan, A. S. (2013). Advanced maternal age and adverse pregnancy outcome: Evidence from a large contemporary cohort. *PloS One, 8*(2), e56583. doi:10.1371/journal.pone.0056583

Kim, B.Y., Lee, Y. R., Kim, Y. K., & Park, S.Y. (2016). *The analysis of medical support system and effectiveness for preterm infants*. Retrieved from http://www.prism.go.kr/homepage/entire/retrieveEntireDetail.do;jsessionid=2541FCAE1D7DAE973F49EADEF669F88A.node02?cond_research_name=&cond_research_start_date=&cond_research_end_date=&research_id=1351000-201600305&pageIndex=31&leftMenuLevel=160

Kim, O. J., Ha, E. H., Kim, B. M., Seo, J. H., Park, H. S., Jung, W. J. . . . Kim, H. (2007). PM_{10} and pregnancy outcomes: A hospital-based cohort study of pregnant women in Seoul. *Journal of Occupational and Environmental Medicine, 49*(12), 1394-1402. doi:10.1097/JOM.0b013e3181594859

Kloog, I., Melly, S. J., Ridgway, W. L., Coull, B. A., & Schwartz, J. (2012). Using new satellite based exposure methods to study the association between pregnancy $PM_{2.5}$ exposure, premature birth and birth weight in Massachusetts. *Environmental Health, 11*(1), 1140-1148. doi:10.1186/1476-069X-11-40

Korea Electric Power Corporation. (KEPCO). (2016). *Annual report* . Retrieved from https://home.kepco.co.kr/kepco/EN/C/htmlView/ENCBHP003.do?menuCd=EN030204

Korea Ministry of Environment. (n.d.). *Air quality standards and air pollution level.* Retrieved from http://eng.me.go.kr/eng/web/index.do?menuId=252

KOrean Statistical Information Service. (2015). *Rate of smoking, drinking, complications during pregnancy*. Retrieved from http://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=TX_117_2013_H173&vw_cd=MT_ZTITLE&list_id=117_11774_2013_05&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE#

KOrean Statistical Information Service. (2015). *Statistical database*. Retrieved from http://kosis.kr/eng/statisticsList/statisticsList_01List.jsp

Langkamp, D. L., Lehman, A., & Lemeshow, S. (2010). Techniques for handling missing data in secondary analyses of large surveys. *Academic Pediatrics, 10*(3), 205–210. doi:10.1016/j.acap.2010.01.005

Lavigne, E., Yasseen, A. S., Stieb, D. M., Hystad, P., Van Donkelaar, A., Martin, R. V. . . . Weichenthal, S. (2016). Ambient air pollution and adverse birth outcomes: Differences by maternal comorbidities. *Environmental Research, 148,* 457-466. doi:10.1016/j.envres.2016.04.026

Laurent, O., Hu, J., Li, L., Kleeman, M. J., Bartell, S. M., Cockburn, M. . . . Wu, J. (2016). A statewide nested case-control study of preterm birth and air pollution by source and composition: California, 2001-2008. *Environmental Health Perspectives, 124*(9), 1479-1486. doi:10.1289/ehp.1510133

Lee, B. E., Park, H. S., Kim, Y. J., Park, E. A., Hong, Y. C., & Ha, E. H. (2004). Air pollution exposure and health effects in fetus. *Journal of Preventive Medicine and Public Health, 37*(4), 291-299. doi:10.1016/j.mrgentox.2008.07.006

Lee, P., Roberts, J. M., Catov, J. M., Talbott, E. O., & Ritz, B. (2013). First trimester exposure to ambient air pollution, pregnancy complications and adverse birth outcomes in Allegheny County, PA. *Maternal and Child Health Journal, 17*(3), 545-555. doi:10.1007/s10995-012-1028-5

Leal, M. C., Esteves-Pereira, A. P., Nakamura-Pereira, M., Torres, J. A., Theme-Filha, M., Domingues, R. M. . . . Gama, S. G. (2016). Prevalence and risk factors related to preterm birth in Brazil. *Reproductive Health, 13*(Suppl 3), 127-138. doi:10.1186/s12978-016-0230-0

Leem, J. H., Kaplan, B. M., Shim, Y. K., Pohl, H. R., Gotway, C. A., Bullard, S. M. . . . Tylenda, C. A. (2006). Exposures to air pollutants during pregnancy and preterm delivery. *Environmental Health Perspectives, 114*(6), 905-910. doi:10.1289/ehp.8733

Lima, T. A. C. D., Nascimento, L. F. C., Medeiros, A. P. P. D., & Santos, V. D. P. (2014). Association between maternal exposure to particulate matter and premature birth. *Revista Ambiente & Água, 9*(1), 27-36. doi:10.4136/ambi-agua.1262

Liu, C. N., Awasthi, A., Hung, Y. H., Gugamsetty, B., Tsai, C. J., Wu, Y. C., & Chen, C. F. (2013). Differences in 24-h average $PM_{2.5}$ concentrations between the beta attenuation monitor(BAM) and the dichotomous sampler(Dichot). *Atmospheric environment, 75*, 341-347. doi:10.1016/j.atmosenv.2013.04.062

Lombardo, J. S., & Buckeridge, D. L. (Eds.). (2007). *Disease surveillance: A public health informatics approach. Hoboken,* NJ: Wiley-Interscience.

Lucas, R. (2012). *Frameworks for causal inference in epidemiology*. Retrieved from https://www.intechopen.com/books/epidemiology-current-perspectives-on-research-and-practice/frameworks-for-causal-inference-in-epidemiology/

Malmqvist, E., Jakobsson, K., Tinnerberg, H., Rignell-Hydbom, A., & Rylander, L. (2013). Gestational diabetes and preeclampsia in association with air pollution at levels below current air quality guidelines. *Environmental Health Perspectives, 121*(4), 488–493. doi:10.1289/ehp.1205736

Martin, J. A., Osterman, M. J., Kirmeyer, S. E., & Gregory, E. C. (2015). Measuring gestational age in vital statistics data: transitioning to the obstetric estimate. *National Vital Statistics Reports, 64*(5), 1-20. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/?term=26047089

Maslow, C. B., Caramanica, K., Li, J., Stellman, S. D., & Brackbill, R. M. (2016). Reproductive outcomes following maternal exposure to the events of September 11, 2001, at the World Trade Center, in New York City. *American Journal of Public Health, 106*(10), 1796-1803. doi:10.2105/AJPH.2016.303303

Maxwell, N. I. (2014). *Understanding environmental health: How we live in the world* (2nd ed.). Burlington, MA: Jones & Bartlett Learning, L.L.C.

Mendez, D. D., Thorpe, R. J., Amutah, N., Davis, E. M., Walker, R. E., Chapple-McGruder, T., & Bodnar, L. (2016). Neighborhood racial composition and poverty in association with pre-pregnancy weight and gestational weight gain. *SSM-Population Health, 2*, 692-699. doi:10.1016/j.ssmph.2016.09.008

Menon, R. (2014). Oxidative stress damage as a detrimental factor in preterm birth pathology. *Frontiers in Immunology,* 5, 567-581 doi:10.3389/fimmu.2014.00567

Messer, L. C., Vinikoor, L. C., Laraia, B. A., Kaufman, J. S., Eyster, J., Holzman, C. . . . O'Campo, P. (2008). Socioeconomic domains and associations with preterm birth. *Social Science & Medicine, 67*(8), 1247-1257. doi:10.1016/j.socscimed.2008.06.009

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., De Savigny, D., Lozano, R., & Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *The Lancet*, *386*(10001), 1395-1406. doi:10.1016/S0140-6736(15)60171-4

Miller, L. S., Robinson, J. A., & Cibula, D. A. (2016). Healthy immigrant effect: Preterm births among immigrants and refugees in Syracuse, NY. *Maternal and Child Health Journal, 20*(2), 484-493. doi:10.1007/s10995-015-1846-3

Miliku, K., Vinkhuyzen, A., Blanken, L. M., McGrath, J. J., Eyles, D. W., Burne, T. H. . . . Jaddoe, V. W. (2016). Maternal vitamin D concentrations during pregnancy, fetal growth patterns, and risks of adverse birth outcomes. *American Journal of Clinical Nutrition, 103*(6), 1514-1522. doi:10.3945/ajcn.115

McLeroy, K. R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly, 15*(4), 351-377. doi:10.1177/109019818801500401

Nachman, R. M., Mao, G., Zhang, X., Hong, X., Chen, Z., Soria, C. S. . . . Biswal, S. (2016). Intrauterine inflammation and maternal exposure to ambient $PM_{2.5}$ during preconception and specific periods of pregnancy: the Boston birth cohort. *Environmental Health Perspectives (Online), 124*(10), 1608-1615. doi:10.1289/EHP243

National Law Information Center. (2016). *Law search*. Retrieved from http://www.law.go.kr/eng/engLsSc.do

Noh, B. K., & Choi, K. H. (2017). Development of IoT-based $PM_{2.5}$ measuring device. *Journal of the Korean Society of Safety, 32*(1), 21-26. doi:10.14346/JKOSOS.2017.32.1.21

OECD Better Life Index. (2015). *Environment*. Retrieved from http://www.oecdbetterlifeindex.org/topics/environment/

Oftedal, A., Busterud, K., Irgens, L. M., Haug, K., & Rasmussen, S. (2016). Socio-economic risk factors for preterm birth in Norway 1999-2009. *Scandinavian Journal of Public Health, 44*(6), 587-592. doi:10.1177/1403494816653288

Osborne, J. W. (2015). *Best practices in logistic regression*. Thousand Oaks, CA: SAGE Publications.

Park, S. J., Kim, B. W., Lee, J. J., & Kim, M. H. (2008). Changes in the birth weight of full term newborns in Korean population over the past 5 years. *Korean Journal of Perinatology, 19*(2), 174-180. Retrieved from https://komci.org/GSResult.php?RID=0044KJP%2F2008.19.2.174&DT=1

Parisot, M., Jolivet, A., Boukhari, R., & Carles, G. (2016). Shigellosis and pregnancy in French Guiana: Obstetric and neonatal complications. *American Journal of Tropical Medicine and Hygiene, 95*(1), 26-30. doi:10.4269/ajtmh.15-0669

Patra, K., Greene, M. M., Patel, A. L., & Meier, P. (2016). Maternal education level predicts cognitive, language, and motor outcome in preterm infants in the second year of life. *American Journal of Perinatology, 33*(08), 738-744. doi:10.1055/s-0036-1572532

Parker, E. A., Baldwin, G. T., Israel, B., & Salinas, M. A. (2004). Application of health promotion theories and models for environmental health. *Health Education & Behavior, 31*(4), 491-509. doi:10.1177/1090198104265601

Perveen, S., & Soomro, T. K. (2016). Sideropaenic anaemia: Impact on perinatal outcome at tertiary care hospital. *Journal of The Pakistan Medical Association, 66*(8), 952-956. Retrieved from http://jpma.org.pk/full_article_text.php?article_id=7856

Pinault, L., Crouse, D., Jerrett, M., Brauer, M., & Tjepkema, M. (2016). Spatial associations between socioeconomic groups and NO_2 air pollution exposure within three large Canadian cities. *Environmental Research, 147*, 373-382. doi:10.1016/j.envres.2016.02.033

Peters, D. H. (2014). The application of systems thinking in health: Why use systems thinking?. *Health Research Policy and Systems, 12*(1), 51-57. doi:10.1186/1478-4505-12-51

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879-891. doi:10.3758/BRM.40.3.879

Prunet, C., Delnord, M., Saurel-Cubizolles, M. J., Goffinet, F., & Blondel, B. (2017). Risk factors of preterm birth in France in 2010 and changes since 1995: Results from the French National Perinatal surveys. *Journal of Gynecology Obstetrics and Human Reproduction, 46*(1), 19-28. doi:10.1016/j.jgyn.2016.02.010

Pu, X., Gao, G., Fan, Y., & Wang, M. (2016). Parameter estimation in stratified cluster sampling under randomized response models for sensitive question survey. *PLoS One, 11*(2), e0148267. doi:10.1371/journal.pone.0148267

Pundi, K. N., Pundi, K., Johnson, J. N., Dearani, J. A., Bonnichsen, C. R., Phillips, S. D. . . . Cetta, F. (2016). Contraception practices and pregnancy outcome in patients after Fontan operation. *Congenital Heart Disease, 11*(1), 63-70. doi:10.1111/chd.12291

Rao, D., Kumar, S., Mohanraj, R., Frey, S., Manhart, L. E., & L Kaysen, D. (2016). The impact of domestic violence and depressive symptoms on preterm birth in south India. *Social Psychiatry and Psychiatric Epidemiology, 51*(2), 225-232. doi:10.1007/s00127-015-1167-2

Restrepo-Méndez, M. C., Lawlor, D. A., Horta, B. L., Matijasevich, A., Santos, I. S., Menezes, A. B. . . . Victora, C. G. (2015). The association of maternal age with birthweight and gestational age: A cross-cohort comparison. *Paediatric and Perinatal Epidemiology, 29*(1), 31-40. doi:10.1111/ppe.12162

Rodriguez-Rodriguez, V., Rojo-Perez, F., Fernandez-Mayoralas, G., Morillo-Tomas, R., Forjaz, J., & Prieto-Flores, M. E. (2017). Active ageing index: application to Spanish regions. *Journal of Population Ageing, 10*(1), 25-40. doi:10.1007/s12062-016-9171-1

Rothman, K. J., & Greenland, S. (2005). Causation and causal inference in epidemiology. *American Journal of Public Health, 95*(S1), S144–S150. doi:10.2105/AJPH.2004.059204.

Rotko, T., Koistinen, K., HAeNNINEN, O. T. T. O., & Jantunen, M. (2000). Sociodemographic descriptors of personal exposure to fine particles ($PM_{2.5}$) in EXPOLIS Helsinki. *Journal of Exposure Science and Environmental Epidemiology, 10*(4), 385-393. doi:10.1038/sj.jea.7500104

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: current practices and new recommendations. *Social & Personality Psychology Compass, 5*(6), 359-371. doi:10.1111/j.1751-9004.2011.00355.x

Salazar, L , Crosby, R. A., & DiClemente, R. J. (2015). *Research methods in health promotion (2nd ed.*). San Francisco: Jossey-Bass.

Salihu, H. M., Ghaji, N., Mbah, A. K., Alio, A. P., August, E. M., & Boubakari, I. (2012). Particulate pollutants and racial/ethnic disparity in feto-infant morbidity outcomes. *Maternal and Child Health Journal, 16*(8), 1679-1687. doi:10.1007/s10995-011-0868-8

Sallis, J. F., Cervero, R. B., Ascher, W., Henderson, K. A., Kraft, M. K., & Kerr, J. (2006). An ecological approach to creating active living communities. *Annual Review of Public Health, 27*, 297-322. doi:10.1146/annurev.publhealth.27.021405.102100

Savitri, A. I., Zuithoff, P., Browne, J. L., Amelia, D., Baharuddin, M., Grobbee, D. E., & Uiterwaal, C. S. (2016).. Does pre-pregnancy BMI determine blood pressure during pregnancy? A prospective cohort study. *BMJ Open, 6*(8), e011626. doi:10.1136/bmjopen-2016-011626

Schweizer, D., Cisneros, R., & Shaw, G. (2016). A comparative analysis of temporary and permanent beta attenuation monitors: The importance of understanding data and equipment limitations when creating $PM_{2.5}$ air quality health advisories. *Atmospheric Pollution Research, 7*(5), 865-875. doi:10.1016/j.apr.2016.02.003

Smith, A. K., Ayanian, J. Z., Covinsky, K. E., Landon, B. E., McCarthy, E. P., Wee C. C., & Steinman, M. A. (2011). Conducting high-value secondary dataset analysis: An introductory guide and resources. *Journal of General Internal Medicine, 26*(8), 920–929. doi:10.1007/s11606-010-1621-5

Stacey, T., Prady, S., Haith-Cooper, M., Downe, S., Simpson, N., & Pickett, K. (2016). Ethno-specific risk factors for adverse pregnancy outcomes: Findings from the born in Bradford cohort study. *Maternal and Child Health Journal, 20*(7), 1394-1404. doi:10.1007/s10995-016-1936-x

Statistics Korea. (KOSTAT). (2016). *Survey outline*. Retrieved from http://kostat.go.kr/portal/eng/surveyOutline/8/1/index.static

Stieb, D. M., Chen, L., Beckerman, B. S., Jerrett, M., Crouse, D. L., Omariba, D. R. . . . Dugandzic, R. M. (2016). Associations of pregnancy outcomes and $PM_{2.5}$ in a national Canadian study. *Environmental Health Perspectives, 124*(2), 243-249. doi:10.1289/ehp.1408995

Stokkeland, K., Ludvigsson, J. F., Hultcrantz, R., Ekbom, A., Höijer, J., Bottai, M., & Stephansson, O. (2016). Increased risk of preterm birth in women with autoimmune hepatitis–a nationwide cohort study. *Liver International*, *36*(1), 76-83. doi:10.1111/liv.12901

Strand, L. B., Barnett, A. G., & Tong, S. (2011). The influence of season and ambient temperature on birth outcomes: A review of the epidemiological literature. *Environmental Research, 111*(3), 451-462. doi:10.1016/j.envres.2011.01.023

Streimikiene, D., & Siksnelyte, I. (2016). Sustainability assessment of electricity market models in selected developed world countries. *Renewable and Sustainable Energy Reviews, 57*, 72-82. doi:10.1016/j.rser.2015.12.113

Suh, Y. J., Kim, H., Seo, J. H., Park, H., Kim, Y. J., Hong, Y. C., & Ha, E. H. (2009). Different effects of PM_{10} exposure on preterm birth by gestational period estimated from time-dependent survival analyses. *International Archives of Occupational and Environmental Health, 8*2(5), 613-621. doi:10.1007/s00420-008-0380-7

Symanski, E., Davila, M., McHugh, M. K., Waller, D. K., Zhang, X., & Lai, D. (2014). Maternal exposure to fine particulate pollution during narrow gestational periods and newborn health in Harris County, Texas. *Maternal and Child Health Journal*, *18*(8), 2003-2012. doi:10.1007/s10995-014-1446-7

Szklo, M., & Nieto, F. J. (2014). *Epidemiology: Beyond the basics* (3rd ed.). Sudbury, MA: Jones and Bartlett.

Tellapragada, C., Eshwara, V. K., Bhat, P., Acharya, S., Kamath, A., Bhat, S. . . . Mukhopadhyay, C. (2016). Risk factors for preterm birth and low birth weight among pregnant Indian women: A hospital-based prospective study. *Journal of Preventive Medicine and Public Health*, *49*(3), 165-175. doi:10.3961/jpmph.16.022

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, *1*(1), 77–100. doi:10.1177/2345678906292430

Toro, R., Campos, C., Molina, C., & Leiva-Guzmán, M. A. (2015). Accuracy and reliability of Chile's national air quality information system for measuring particulate matter: Beta attenuation monitoring issue. *Environment International*, *82*, 101-109. doi:10.1016/j.envint.2015.02.009

Trasande, L., Malecha, P., & Attina, T. M. (2016). Particulate matter exposure and preterm birth: Estimates of US attributable burden and economic costs. *Environmental Health Perspectives, 124*(12), 1913-1918. doi:10.1289/ehp.1510810

Trasande, L., Wong, K., Roy, A., Savitz, D. A., & Thurston, G. (2013). Exploring prenatal outdoor air pollution, birth outcomes and neonatal health care utilization in a nationally representative sample. *Journal of Exposure Science & Environmental Epidemiology*, *23*(3), 315-321. doi:10.1038/jes.2012.124

Triantafyllou, E., Diapouli, E., Tsilibari, E. M., Adamopoulos, A. D., Biskos, G., & Eleftheriadis, K. (2016). Assessment of factors influencing PM mass concentration measured by gravimetric & beta attenuation techniques at a suburban site. *Atmospheric Environment*, *131*, 409-417. doi:10.1016/j.atmosenv.2016.02.010

Trochim, W. M., Cabrera, D.A., Milstein, B., Gallagher, R.S., & Leischow, S.J. (2006). Practical challenges of systems thinking and modeling in Public Health. *American Journal of Public Health*, *96*(3), 538–546. doi:10,2105/AJPH,2005,066001

Tyagi, V., Mustafa, M. D., Sharma, T., Banerjee, B. D., Ahmed, R. S., Tripathi, A. K., & Guleria, K. (2016). Association of organochlorine pesticides with the mRNA expression of tumour necrosis factor-alpha (TNF-α) & cyclooxygenase-2 (COX-2) genes in idiopathic preterm birth. *Indian Journal of Medical Research, 143*(6), 731-738. doi:10.4103/0971-5916.191986

Vandenbroucke, J. P., Broadbent, A., & Pearce, N. (2016). Causality and causal inference in epidemiology: The need for a pluralistic approach. *International Journal of Epidemiology*, *45*(6), 1776-1786. doi:10.1093/ije/dyv341

United States Environmental Protection Agency. (US EPA). (2017). *Particulate matter (PM) pollution*. Retrieved from https://www.epa.gov/pm-pollution.

Victorian Curriculum and Assessment Authority. (2010). *Social ecological model and physical activity*. Retrieved from http://www.vcaa.vic.edu.au/documents/vce/physicaledu/socialecologicalmodelandphysicalactivity.doc

Villeneuve, P. J., Goldberg, M. S., Burnett, R. T., van Donkelaar, A., Chen, H., & Martin, R. V. (2011). Associations between cigarette smoking, obesity, sociodemographic characteristics and remote-sensing-derived estimates of ambient $PM_{2.5}$: Results from a Canadian population-based survey. *Occupational and Environmental Medicine, 68*(12), 920-927. doi:10.1136/oem.2010.062521

Vogel, J. P., Chawanpaiboon, S., Watananirun, K., Lumbiganon, P., Petzold, M., Moller, A. B., ... & Tunçalp, O. (2016). Global, regional and national levels and trends of preterm birth rates for 1990 to 2014: Protocol for development of World Health Organization estimates. *Reproductive Health, 13*(1), 76-85. doi:10.1186/s12978-016-0193-1

von Ehrenstein, O. S., Wilhelm, M., Wang, A., & Ritz, B. (2014). Preterm birth and prenatal maternal occupation: The role of Hispanic ethnicity and nativity in a population-based sample in Los Angeles, California. *American Journal of Public Health, 104*(S1), S65-S72. doi:10.2105/AJPH.2013.301457

Vos, A. A., Posthumus, A. G., Bonsel, G. J., Steegers, E. A., & Denktaş, S. (2014). Deprived neighborhoods and adverse perinatal outcome: A systematic review and meta-analysis. *Acta Obstetricia et Gynecologica Scandinavica, 93*(8), 727-740. doi:10.1111/aogs.12430

Walden University. (n.d.). *Research ethics & compliance: Guides and FAQs*. Retrieved from http://academicguides.waldenu.edu/researchcenter/orec/guides.

Wallace, M. E., Mendola, P., Chen, Z., Hwang, B. S., & Grantz, K. L. (2016). Preterm Birth in the context of increasing income inequality. *Maternal and Child Health Journal, 20*(1), 164-171. doi:10.1007/s10995-015-1816-9

Wang, H., Hu, Y., Hao, J., Chen, Y., Wang, Y., Zhu, P. . . . Xu, D. (2016). Maternal serum zinc concentration during pregnancy is inversely associated with risk of preterm birth in a Chinese population. *Journal of Nutrition, 146*(3), 509-515. doi:10.3945/jn.115.220632

Warner, R. M. (2013*). Applied statistics: From bivariate through multivariate techniques* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Watt, M., Eaton, L., Dennis, A., Choi, K., Kalichman, S., Skinner, D., & Sikkema, K. (2016). Alcohol use during pregnancy in a South African community: Reconciling knowledge, norms, and personal experience. *Maternal and Child Health Journal, 20*(1), 48-55. doi:10.1007/s10995-015-1800-4

World Health Organization. (2014). *Ambient (outdoor) air pollution in cities database 2014*. Retrieved from http://www.who.int/phe/health_topics/outdoorair/databases/cities-2014/en/

Yi, O., Kim, H., & Ha, E. (2010). Does area level socioeconomic status modify the effects of PM_{10} on preterm delivery?. *Environmental Research, 110*(1), 55-61. doi:10.1016/j.envres.2009.10.004

Yoon, H. S. (2009). Statistics and medical cost of preterm in Korea. *Hanyang Medical Reviews, 29*(4), 386-390. doi: 10.7599/hmr.2009.29.4.386

Yorifuji, T., Naruse, H., Kashima, S., Takao, S., Murakoshi, T., Doi, H., & Kawachi, I. (2013). Residential proximity to major roads and adverse birth outcomes: A hospital-based study. *Environmental Health, 12*(1), 34-45. doi:10.1186/1476-069X-12-34

Appendix A: Data Use Agreement

DATA USE AGREEMENT

This Data Use Agreement ("Agreement"), effective as of 8 December 2017 ("Effective Date"), is entered into by and between Hyun Jin Choi("Data Recipient") and Statistics Korea ("Data Provider"). The purpose of this Agreement is to provide Data Recipient with access to a Limited Data Set ("LDS") for use in research in accord with laws and regulations of the governing bodies associated with the Data Provider, Data Recipient, and Data Recipient's educational program. If there is a discrepancy between the laws, the agreement is subject to stricter laws.

1. Definitions. Unless otherwise specified in this Agreement, all capitalized terms used in this Agreement not otherwise defined have the meaning established for purposes of the "HIPAA Regulations" codified at Title 45 parts 160 through 164 of the United States Code of Federal Regulations, as amended from time to time.

2. Preparation of the LDS. Data Provider shall prepare and furnish to Data Recipient a LDS in accord with laws and regulations of the governing bodies associated with the Data Provider, Data Recipient, and Data Recipient's educational program.

Data Fields in the LDS. **No direct identifiers such as names may be included in the Limited Data Set (LDS).** The researcher will also not name the organization in the doctoral project report that is published in Proquest. In preparing the LDS, Data Provider or designee shall include the **data fields specified as follows**, which are the minimum necessary to accomplish the research: Vital Statistics for Live Birth Data Files for period 1 February 2015 to 31 December 2015..

3. Responsibilities of Data Recipient. Data Recipient agrees to:

   a. Use or disclose the LDS only as permitted by this Agreement or as required by law;

   b. Use appropriate safeguards to prevent use or disclosure of the LDS other than as permitted by this Agreement or required by law;

   c. Report to Data Provider any use or disclosure of the LDS of which it becomes aware that is not permitted by this Agreement or required by law;

   d. Require any of its subcontractors or agents that receive or have access to the LDS to agree to the same restrictions and conditions on the use and/or disclosure of the LDS that apply to Data Recipient under this Agreement; and

   e. Not use the information in the LDS to identify or contact the individuals who are data subjects.

4. Permitted Uses and Disclosures of the LDS. Data Recipient may use and/or disclose the LDS for its research activities only.

5. Term and Termination.

   a. Term. The term of this Agreement shall commence as of the Effective Date and shall continue for so long as Data Recipient retains the LDS, unless sooner terminated as set forth in this Agreement.

   b. Termination by Data Recipient. Data Recipient may terminate this agreement at any time by notifying the Data Provider and returning or destroying the LDS.

   c. Termination by Data Provider. Data Provider may terminate this agreement at any time by providing thirty (30) days prior written notice to Data Recipient.

   d. For Breach. Data Provider shall provide written notice to Data Recipient within ten (10) days of any determination that Data Recipient has breached a material term of this Agreement. Data Provider shall afford Data Recipient an opportunity to cure said alleged material breach upon mutually agreeable terms. Failure to agree on mutually agreeable terms for cure within thirty (30) days shall be grounds for the immediate termination of this Agreement by Data Provider.

   e. Effect of Termination. Sections 1, 4, 5, 6(e) and 7 of this Agreement shall survive any termination of this Agreement under subsections c or d.

6. Miscellaneous.

   a. Change in Law. The parties agree to negotiate in good faith to amend this Agreement to comport with changes in federal law that materially alter either or both parties' obligations under this Agreement. Provided however, that if the parties are unable to agree to mutually acceptable amendment(s) by the compliance date of the change in applicable law or regulations, either Party may terminate this Agreement as provided in section 6.

   b. Construction of Terms. The terms of this Agreement shall be construed to give effect to applicable federal interpretative guidance regarding the HIPAA Regulations.

   c. No Third Party Beneficiaries. Nothing in this Agreement shall confer upon any person other than the parties and their respective successors or assigns, any rights, remedies, obligations, or liabilities whatsoever.

d. Counterparts. This Agreement may be executed in one or more counterparts, each of which shall be deemed an original, but all of which together shall constitute one and the same instrument.

e. Headings. The headings and other captions in this Agreement are for convenience and reference only and shall not be used in interpreting, construing or enforcing any of the provisions of this Agreement.

IN WITNESS WHEREOF, each of the undersigned has caused this Agreement to be duly executed in its name and on its behalf.

DATA PROVIDER	DATA RECIPIENT
Signed: 민지현	Signed: [signature]
Print Name: Ji Hyun Min	Print Name: Hyun Jin Choi
Print Title: Microdata provision official of Statistics Korea	Print Title: Ph.D. Student