Distribution Agreement

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

Signature:

_____ _____

Audrey L. Flak Date

$PM_{2.5}$ exposure in early life and childhood asthma incidence in a retrospective birth cohort

By

Audrey Lauren Flak, M.P.H.

Doctor of Philosophy

Epidemiology

_____

Lyndsey Darrow, Ph.D.
Advisor

_____

Matthew Strickland, Ph.D., M.P.H.
Committee Member

_____

Mitchel Klein, Ph.D.
Committee Member

_____

Craig Hansen, Ph.D.
Committee Member

_____

Carolyn Drews-Botsch, Ph.D., M.P.H.
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

$PM_{2.5}$ exposure in early life and childhood asthma incidence in a retrospective birth cohort

By

Audrey Lauren Flak

M.P.H., Emory University, 2010

B.S., Vanderbilt University, 2008

Advisor: Lyndsey Darrow, Ph.D.

An abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Epidemiology

2016

Abstract

$PM_{2.5}$ exposure in early life and childhood asthma incidence in a retrospective birth cohort

By Audrey Flak

It is well established that urban air pollution exacerbates pre-existing asthma, and the literature suggests that it may also play a role in asthma development. This dissertation investigated the association between exposure to $PM_{2.5}$ (particulate matter ≤2.5μm in diameter) in early life and childhood asthma incidence in the Kaiser Air Pollution and Pediatric Asthma Study (KAPPA). The KAPPA Study is a birth cohort of 24,608 children born between 2000 and 2010 enrolled in Kaiser Permanente Georgia.

Asthma case definitions vary widely among studies using medical records to define disease. In Aim 1, we examined 15 case definitions of incident asthma in early life. Choice of case definition had a large impact on the estimate of asthma incidence by age 3 and the ability to predict asthma at school age. These results informed our decision to designate one asthma diagnosis plus one asthma-related medication dispensing as the primary outcome definition in subsequent aims.

In Aims 2 and 3 we assessed the association between prenatal and first year of life exposure to $PM_{2.5}$ and asthma incidence by ages 2 through 6. The impact of exposure to both primary $PM_{2.5}$ from traffic emissions and total $PM_{2.5}$ was explored. In adjusted models, an increase of 1 $\mu g/m^3$ of traffic $PM_{2.5}$ during the first year of life was associated with a 2.7% to a 5.8% absolute increase in risk of asthma, depending on the follow-up age (Risk Difference(95%CI) age 2=0.027(0.003,0.050); age 3=0.037(0.004,0.070); age 4=0.037(-0.007,0.082); age 5=0.058(0.004,0.112); age 6=0.036(-0.029,0.101)). An increase of 1 $\mu g/m^3$ of total $PM_{2.5}$ was associated with a 0.4% to 1.8% increase in risk of asthma (RD(95%CI) age 2=0.008(-0.002,0.017); age 3=0.007(-0.006,0.020); age 4=0.004(-0.014,0.022); age 5=0.018(-0.005,0.041); age 6=0.018(-0.011,0.046)). Risk differences were smaller for the association of $PM_{2.5}$ exposure during pregnancy. Across aims, we observed little evidence of additive interaction between $PM_{2.5}$ and child race, child sex, maternal asthma, or city region of residence.

This dissertation provides some evidence for an association between $PM_{2.5}$ exposure in early life and childhood asthma incidence. Our results highlight the impact of case definition on estimates of asthma incidence in early childhood in a medical record setting.

$PM_{2.5}$ exposure in early life and childhood asthma incidence in a retrospective birth cohort

By

Audrey Lauren Flak

M.P.H., Emory University, 2010

B.S., Vanderbilt University, 2008

Advisor: Lyndsey Darrow, Ph.D.

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Epidemiology

2016

Acknowledgements

I am indebted to the members of my dissertation committee, for their dedication to this work and for re-framing the way I think about epidemiology. Specifically, I thank: Lyndsey Darrow for her mentorship and time, for helping me to grow as an epidemiologist, for always pushing me to take analyses one step further and teaching me to be cautious in my interpretations; Matt Strickland for the opportunity to work on this study and his unwavering encouragement and feedback throughout the program; Mitch Klein for being a wonderful and patient teacher, for answering my hundreds of questions about epidemiology, and for always coming up with new ways to try to understand our data; Carey Drews-Botsch for her thoughtful feedback on all of my drafts and for allowing me to part of the T32 community at Emory; Craig Hansen for providing us access to these data, and remaining involved in this project even after moving to Australia.

I thank the Department of Epidemiology, the Department of Environmental Health, Jena Black, Julie Gazmararian, and the Emory SOPHIA group for supporting me throughout the PhD program. I thank the Department of Epidemiology faculty for their dedication to teaching inside and outside of the classroom.

I thank my family and friends for their love and support. In particular, I thank my parents for being my number one champions and for always reminding me to keep things in perspective.

I acknowledge the financial support for this dissertation: NIOSH Environmental Epidemiology Training Grant 5T01OH008609-06, NIH Reproductive, Perinatal, and Pediatric Epidemiology Training Grant 5T32HD052460 and U.S. EPA grant R834799 *[This publication's contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US EPA. Further, US EPA does not endorse the purchase of any commercial product or services mentioned in the publication.]*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Asthma is a chronic disease characterized by a narrowing of the airways that causes shortness of breath, trouble breathing, coughing, and chest tightness. It greatly affects quality of life and can result in hospitalizations and, in severe cases, death. Asthma is the most common chronic disease among children and has had a consistently high prevalence in the United States and worldwide over the past several decades. Since the lungs develop over an extended period, beginning at six weeks after conception and continuing through childhood, both pre- and postnatal exposure may be of importance in asthma incidence. While it is well established that ambient air pollution exacerbates pre-existing asthma, a growing body of literature suggests it may also play a role in asthma development (Bråbäck et al. 2009, Bowatte et al. 2015).

This dissertation will explore the impact of exposure to particulate matter less than or equal to 2.5 micrometers ($PM_{2.5}$) during pregnancy and the first year of life on childhood asthma incidence in the Kaiser Air Pollution and Pediatric Asthma (KAPPA) Study. The KAPPA Study is a retrospective birth cohort of children born between 2000 and 2010 enrolled in Kaiser Permanente Georgia. The air pollution data for this work will come from the EPA-funded Southeastern Center for Air Pollution and Epidemiology which is one of four national Clean Air Research Centers in the United States (Southeastern Center for Air Pollution and Epidemiology (SCAPE) 2010). This dissertation is divided into 3 main aims which address the following objectives:

1. Assess the impact of different approaches to using medical records to estimate the cumulative incidence of asthma by age 3. Determine the validity of these early-life

asthma case definitions, which exclusively use information available in medical records, in predicting asthma at school age.

2. Examine the association between $PM_{2.5}$ from traffic at the residential location during pregnancy and the first year of life and cumulative asthma incidence by ages 2 through 6
3. Examine the association between total $PM_{2.5}$ at the residential location during pregnancy and the first year of life and cumulative asthma incidence by ages 2 through 6

Together, the findings of this dissertation will contribute to our knowledge of the use of medical records to define asthma in early life, and the associations between $PM_{2.5}$ exposure, overall and specifically from traffic sources, in key developmental windows and childhood asthma incidence.

CHAPTER 2

Literature Review

CHILDHOOD ASTHMA

Asthma is a chronic inflammatory disease characterized by a reversible narrowing of the airways caused by airway edema, hyperresponsiveness, and bronchoconstriction (National Asthma Education and Prevention Program. Expert Panel Report 3 2007). This tightening causes recurrent shortness of breath, trouble breathing, coughing, and chest tightness. Asthma is a remarkably heterogeneous disease with phenotypes and response to therapies differing considerably between patients. Childhood asthma specifically not only effects current health, but can also impact lung development through mechanisms such as airway remodeling (Grol et al. 1999, Strunk et al. 2006, Durrani et al. 2011). While frequency of respiratory symptoms often change as a child gets older, the lungs of an individual who had childhood asthma may never recover from such remodeling and will always be susceptible to respiratory hyperresponsiveness (Yoshikawa et al. 2011, Gershwin et al. 2012). Treatment of asthma involves use of medications to prevent asthma symptoms, medications to target current symptoms, and avoidance and management of situations known to exacerbate the disease in the individual. Asthma greatly affects quality of life, and when uncontrolled can result in hospitalizations and death.

Asthma is the most common chronic disease among children and has had a consistently high prevalence in the United States and worldwide over the past several decades (Anandan et al. 2010, Akinbami et al. 2012). Asthma rates are highest in the United States and other “Westernized” countries (Gold et al. 2005). In the United States in 2010, 9.5% of children ages 0 – 17 years, approximately 7 million children, had asthma with this prevalence increasing at a rate of 1.4% per year (Moorman JE et al. 2012). 58.3% of these

children reported at least one asthma attack in the past year. Asthma is responsible for substantial stress on the healthcare system between physician visits, emergency department visits, asthma management programs, and hospitalizations. Combining direct and indirect costs of asthma, it is estimated that it costs $3.2 billion per year to treat children with asthma. This estimate includes both healthcare costs and indirect costs such as school and work absences (Selgrade et al. 2006).

Asthma Diagnosis

Asthma is typically diagnosed based on recurrence of symptoms, exclusion of alternate diagnoses, response of symptoms to specific medications, physical examination, and spirometry testing if the patient is over the age of five (National Asthma Education and Prevention Program. Expert Panel Report 3 2007). Spirometry testing is crucial for diagnosis because it allows the physician to determine the reversibility of the airway obstruction. In children, conditions that must be ruled out before diagnosing the condition as asthma include upper respiratory disease, bronchitis, endobronchial disease, bronchial/tracheal compression, tracheomalacia, and cystic fibrosis (Bush 2007, National Asthma Education and Prevention Program. Expert Panel Report 3 2007, Hedlin et al. 2012). Medication-wise, if a patient's symptoms do not respond to a high dose of inhaled steroids it is likely that their condition is not asthma. However, syndromes besides asthma will also respond to this treatment so a response does not confirm that the condition is asthma (Bush 2007). The use of biomarkers to diagnose pediatric asthma is an area of active research and includes investigation of the use of exhaled NO, exhaled breath condensate, and biomarkers found in urine (Taylor 2011, Hedlin et al. 2012).

Diagnosis of asthma in children under the age of five is particularly difficult. Fifty to eighty percent of children who have asthma experience symptoms before the age of five

(National Asthma Education and Prevention Program. Expert Panel Report 3 2007). Asthma symptoms in children of these ages are variable and not specific (Global Initiative for Asthma (GINA) 2009). Cough and wheeze, trademark symptoms of asthma in this age group, are common in young children with and without asthma. To make asthma diagnosis more complicated, objective testing, such as spirometry, that is crucial for diagnosis in older individuals is not available for use in children under the age of five. In children of these ages it is important for the provider to determine whether the child's symptoms include actual "wheeze," symptom triggers and history, family history, and if there are any physical deformities that may be causing the symptoms (Bush 2007). The difficulty of diagnosing asthma in this age group often leads to underdiagnosis of the disease which can prevent a child with asthma from receiving proper treatment (National Asthma Education and Prevention Program. Expert Panel Report 3 2007, Eigen 2008).

Despite the difficulties in diagnosing asthma at young ages, it is not uncommon for children to be diagnosed before reaching the age of five. National Health Interview Survey data show that between 2004 and 2005 the prevalence of asthma in children ages 0 to 4 years was 6.2% so diagnoses do occur in non-negligible amounts in this age group (Akinbami et al. 2009). In certain populations and subgroups the prevalence is higher, even when only considering children up until age 2 or 3 (Young et al. 1994, Reichman et al. 2008). There is not complete consensus on the validity of making asthma diagnoses at such young ages. Most researchers and clinicians argue that these diagnoses are necessary in order for a child to have their condition managed and treated appropriately (Hovland et al. 2012). Others disagree arguing that an official asthma diagnosis should be deferred until a child is old enough for objective testing (Brand et al. 2008). There is currently no widely accepted youngest age at which an asthma diagnosis is considered valid. Given respiratory

symptoms and patient history consistent with asthma some physicians will diagnose asthma in infants as young as a couple of months of age.

Wheeze

Wheeze is the most important indicator that a young child either currently has asthma or may go on to develop asthma. However, not all children with asthma experience this symptom and not all children with wheeze have or will ever have asthma. The European Respiratory Society (ERS) Task Force reviewed studies on early life symptoms in children with asthma and concluded that among children with persistent asthma, about 25% wheeze by the age of 6 months and 75% wheeze by the age of 3 years. Among children with recurrent wheeze, it is thought that about half have asthma (Stewart 2008). Wheeze is defined as a high pitched whistling sound that occurs during exhalation that often has an almost musical quality (National Asthma Education and Prevention Program. Expert Panel Report 3 2007, Brand et al. 2008). It primarily occurs among young children and is not common after the age of 6. It is only one of several types of noisy breathing that can occur in infants and young children. A child wheezes when their airways narrow and as a result they have limited expiratory flow (Brand et al. 2008). There is currently no way to differentiate a child whose wheeze is an early sign of asthma from those whose wheeze will be short-lived and is not indicative of future respiratory problems (Hovland et al. 2012).

The narrowing of the airways that can result in symptoms of wheeze can be caused by many different things. Wheeze frequently occurs among children with allergies and those with upper or lower respiratory tract illnesses such as bronchiolitis. In some children wheeze is caused by abnormalities in the structure of a child's airways such as bronchomalacia or cystic fibrosis (Brand et al. 2008). The etiology of wheeze, particularly with respect to allergic pathways, may differ greatly between persistent and transient

wheezers and among children of different ages (Wright 2002). Wheeze is treated by many of the same medications used to treat asthma such as inhaled short-acting beta-agonists and oral corticosteroids.

There are several classifications of wheeze. A report by the ERS Task Force divides wheeze into five categories based on occurrence (whether the wheeze is transient, persistent, or late-onset) and cause (whether it is episodic and associated with viral cold, or whether it is multiple-trigger) (Brand et al. 2008). Recurrent wheeze is a frequently used classification whose definition is variable. It usually describes at least two or three wheeze episodes during a specific time period such as within the first couple of years of life.

Early Life Respiratory Symptoms and Future Disease

A clinically important area of research is trying to determine which young children with respiratory symptoms such as wheeze will go on to develop asthma. The ability to make this prediction has the potential to improve quality of life among the individual and target them for interventions such as removal of exacerbating exposures (Fouzas et al. 2013). Currently, four asthma prediction models exist that each use early life symptoms and child and family characteristics to predict asthma development: The Asthma Predictive Index (API) (Castro-Rodriguez et al. 2000), the Isle of Wright score (Kurukulaaratchy et al. 2003), the ECA severity score (Devulapalli et al. 2008), and the PIAMA risk score (Caudri et al. 2009). As detailed in a 2013 critical review of these models by Fouzas and Brand, (Fouzas et al. 2013) these models all use symptoms before the age of five to predict asthma development by around age ten (exact ages of symptoms and prediction differ between models). While all of the models use some of the same underlying characteristics for their prediction, the specific inputs vary greatly between the models without even one characteristic that is shared by all four models. The most commonly included inputs

between models are frequency of wheezing episodes and parental history of asthma (Fouzas et al. 2013).

Among these risk prediction models, the API is the oldest and also the most widely recognized and used. Developed using the Tucson Children's Respiratory Study and published for the first time in 2000 (Castro-Rodriguez et al. 2000), this model includes both loose and stringent indices based on wheezing, physician diagnosis of eczema and allergic rhinitis, parental asthma, and blood test results. The blood test results are used to determine percent of white blood cells that are eosinophils. Eosinophils are white blood cells that become active during infections and allergic reactions and are common among individuals with asthma (Williams 2004, National Asthma Education and Prevention Program. Expert Panel Report 3 2007). The clinical value of the API has recently been questioned by both an editorial in January 2011 and a review article from March 2013. A first criticism brought up by Brand in 2011 is the lack of independent validation studies of the scale (Brand 2011). In a response to this editorial, the authors of the scale acknowledged the lack of such studies and concluded that the scale should only be used in populations similar to that in which it was developed (Castro-Rodriguez et al. 2011). Beyond validation, the diagnostic capabilities of the scale have also been questioned. The loose index has been ruled out as a useful test based on its inability to predict future asthma or lack thereof as evidenced by its positive and negative likelihood ratios (Fouzas et al. 2013). The stringent index fares better, but its utility is also questioned based on its high false negative rate (calculated by Fouzas and Brand as 73-85%) and high negative likelihood ratio (Fouzas et al. 2013). Additionally, the requirement of a blood test makes it difficult for this index to be used routinely (Fouzas et al. 2013).

The criticisms of the API represent the criticisms Fouzas and Brand have brought up for all four existing models – mainly the lack of validation of these scales and insufficient

negative predictive power (proportion of individuals predicted not to develop asthma who actually do not develop asthma) (Fouzas et al. 2013). While these tools continue to be validated and used – most recently plans for an external validation of the PIAMA scale were published (Hafkamp-de Groen et al. 2012) – some in the field question the utility of current predictive models and even the ability to develop a better one in the future. In their 2013 article Fouzas and Brand conclude that, "widespread use of the currently available asthma predictive models in clinical practice is not justified" and go on to question whether it is possible for a clinically useful prediction model to be developed (Fouzas et al. 2013). Previously, Roberts also questioned the potential for development of a model that is more accurate than current models while also being clinically useful (Roberts 2009). Roberts brings up the multifactorial nature of asthma in his argument and the number of variables that may potentially be needed in order to develop a more accurate predictive scale.(Roberts 2009) These sentiments have been echoed by the developers of the API (Castro-Rodriguez 2010).

Asthma Phenotypes and Allergic Sensitization

Asthma is a heterogeneous disease with several subphenotypes. Two ways to classify asthma cases are 1) into atopic asthma (also known as allergic or extrinsic asthma) and nonatopic asthma (aka instrinsic asthma) and 2) into eosinophilic asthma and noneosinophilic (or neutrophil) asthma. While these two categorizations are related, they use different characteristics for classification.

Atopic and nonatopic asthma are the most commonly used asthma classifications. Atopic asthma is commonly defined as asthma occurring in an individual with signs of atopy, for example an individual who is sensitized to common inhaled allergens (e.g. birch, pollen, and mold). Allergic sensitization is one of the strongest predictors that an individual

will develop asthma. Atopic asthma is thought to be the result of gene-environment interactions which impact innate immune system development particularly during early life (Holt et al. 2005). Indoor and outdoor allergen exposure and subsequent atopic sensitization is strongly associated with asthma development in what is thought by some to be a causal way (Gaffin et al. 2009, Baxi et al. 2010). These allergens include animal dander, cockroach, rodent, and dust mite. Among young children who wheeze, presence of atopic sensitization can help predict who will go on to develop asthma (Sly et al. 2008). Gaffin and Phipatanakul posit that the allergen exposure leads to allergen sensitization which in turn leads to atopic asthma (Gaffin et al. 2009). It is possible that early life sensitization is simply a marker of an asthmatic phenotype and does not play a causal role in asthma development. Alternatively, the relationship may indeed be causal. If this is the case, one possible mechanistic explanation is that inflammation (caused by virus or allergens) during infancy, a period of lung growth and airway remodeling, can disrupt tissue differentiation programs which in turn impacts respiratory function and can result in asthma (Kusel et al. 2007).

Much research has focused on distinguishing atopic and nonatopic asthma phenotypes and the prevalence of each. Allergic sensitization can be determined using the biomarker immunoglobulin E (IgE). Allergen-specific IgE antibodies are created upon contact to an allergen and play a crucial role in the allergic response. They are found in higher quantities in atopic individuals than in nonatopic individuals. Atopy status can be determined by examining blood IgE levels (using a multiallergen or specific IgE antibody screen) or by using an allergen-specific IgE skin prick test. Use of a multiallergen IgE test is currently the mostly widely accepted method to make this determination (Szefler et al. 2012). While the proportion of asthma attributable to atopy differs depending on study population and methods, it is likely that it is somewhere close to half. A review by Pearce and colleagues in 1999 estimated that less than half of all asthma cases are attributable to

atopy (Pearce et al. 1999). In the ten studies on children included in the review prevalences ranged from 30 to 57 percent. From NHANES III skin prick data (from 1988 to 1994) on asthmatic children ages 6 to 19, it was calculated that 55.2% of cases are attributable to atopy (Arbes Jr et al. 2007). Two reports using more recent NHANES data (2005-2006) have found that 62.1% of all asthma cases are atopic (Gergen et al. 2009), and that 66% of children ages 2-12 with asthma also have atopy (Wells et al. 2013) where atopy was defined as at least one positive specific IgE result.

A second way to categorize asthma is via immune pathway into eosinophilic asthma and noneosinophilic asthma (also known as neutrophil asthma). Asthma resulting from these different inflammation pathways may have different responses to medication and may also be caused by different risk factors (Drews et al. 2009). This categorization classifies cases by the proportion of cells that are eosinophils in induced sputum (Pizzichini et al. 1996, Gibson et al. 2000). In 2002, a review of the literature by Douwes and others concluded that 46% of asthmatic children have an eosinophilic phenotype which is comparable, but slightly lower than the percentage in adult asthmatics (Douwes et al. 2002). The authors also point out that it is possible that some asthmatic children have a phenotype resulting from a blend of the two inflammatory mechanisms (Douwes et al. 2002). Eosinophilic asthma (in adults and children) is associated with more severe asthma phenotypes than noneosinophilic asthma. In children, eosinophilic asthma is associated with lower pulmonary function (Lee et al. 2013).

Studies have come to different conclusions about the relation between eosinophilic asthma and atopic asthma. Some research has found that inflammation in nonatopic asthmatics is primarily neutrophilic (Drews et al. 2009). A study by Lee in 2012 found that the majority of individuals with both eosinophilic and noneosinophilic asthma were atopic –

reporting 87% atopy in the eosinophilic group and 77% atopy in the noneosinophilic group (Lee et al. 2013).

Hygiene Hypothesis and Asthma

While early life allergen exposure that results in allergic sensitization is associated with increased risk of asthma, exposure to microbial agents and sources of allergens such as pets early in life may be protective. The hygiene hypothesis states that lack of exposure to infection and microbial agents in early childhood is responsible for the increases in allergic disease, such as atopic asthma, in industrialized countries (Strachan 1989). This hypothesis speculates that avoidance of these exposures impacts the development of an individual's immune system making them more susceptible to allergic diseases later in life. The specific immunologic mechanism through which this occurs has not fully been elucidated (Romagnani 2004).

The consensus of recent research on these associations is that decreased exposures in childhood may explain some of the increases in asthma prevalence in industrialized countries, but that they are not solely responsible for the trend (Ramsey et al. 2005, Brooks et al. 2013). While the hygiene hypothesis specifically relates to allergic (atopic) asthma, some exposures of the hygiene hypothesis such as microbial exposure are also seen to have a protective effect against nonatopic asthma (Brooks et al. 2013). The findings on associations between hygiene hypothesis exposures as a whole and asthma development do not all show the same protective effect. In a 2004 article, Ramsey and Celedón reviewed the literature on asthma and several exposures relevant to the hygiene hypothesis (i.e. household size, crowding, mycobacterial infections, vaccinations, gastrointestinal and parasitic infections, farming and endotoxin exposure, antibiotics, probiotics, and intestinal flora) (Ramsey et al. 2005) and saw inconsistent findings between each exposure and

asthma. There are several potential explanations for these varying results. These associations may differ for atopic and nonatopic asthma, for individuals with high and low risk (specifically those with and without family history of asthma), and may be dependent on the specific timing of exposure (Ramsey et al. 2005, Brooks et al. 2013).

Asthma Disparities

There are striking disparities in childhood asthma prevalence with the highest rates occurring among African Americans and children living in poverty (McDaniel et al. 2006, Akinbami et al. 2012). Between 2008 and 2010 prevalence of current asthma among children ages 0 to 17 was almost twice as high among black children as among white children (16.0% vs. 8.2% respectively). In general, asthma prevalence is lower in Hispanic and Latino children than in white children with the exception of Puerto Rican children who between 2008 and 2010 had a prevalence of 16.9% (Moorman JE et al. 2012). Data from the same source show a higher asthma prevalence among children living below the federal poverty threshold than among children living at 450% or above the poverty level (12.4% and 8.1% respectively) (Moorman JE et al. 2012).

Racial disparities appear to persist when accounting for socioeconomic status and socioeconomic disparities appear to persist when accounting for race. A study by McDaniel and others in 2006 analyzing National Health Interview Survey (NHIS) data from 1997 through 2003 found that racial asthma disparities cannot be explained by child or family characteristics. In fact, at every income asthma prevalence was higher among non-Hispanic black children than among non-Hispanic white children (McDaniel et al. 2006). This relationship was also seen when examining NHIS data from 1997 alone (Smith et al. 2005), NHANES data (Roberts 2002), and in smaller studies in Rhode Island (Pearlman et al. 2006), Michigan (Nelson et al. 1997), and Los Angeles County (Simon et al. 2003), but not in NHIS

data from 1993 to 1996 (Akinbami et al. 2002). Data from the McDaniels study also demonstrate the strong inverse relationship between income and asthma prevalence.

Among children with asthma, African American children and those from the lowest socioeconomic groups suffer the worst outcomes. Children from both of these groups are more likely to have experienced an asthma attack in the past 12 months (Moorman JE et al. 2012), have activity limitations (Simon et al. 2003), have asthma-related emergency department visits and are less likely to have their disease managed through regular check-ups and medication (Crocker et al. 2009, Kim et al. 2009). Emergency department visits among children with current asthma were three times as common among black children than among white children (Moorman JE et al. 2012). While health insurance status undoubtedly can explain some of the differences in morbidity, it does not account for all of the disparity. Differences in access to healthcare and use of care may also be causing some of this disparity (Akinbami et al. 2002, Pearlman 2012). Additionally, among individuals with equal access to care the quality of care received, which includes important components such as access to management medications, may be vastly different (Simon et al. 2003, Gold et al. 2005, Cabana et al. 2007). Beyond the healthcare system, factors such as housing quality, exposure to environmental toxins and allergens, stress, and segregation may all contribute to these differences (Williams et al. 2009, Lamb et al. 2011, Pearlman 2012). One of many pathways through which lower quality housing may be associated with asthma morbidity is exposure to cockroach allergen (Gold et al. 2005). This allergen is found at higher levels in housing that is in disrepair (Rauh et al. 2002) and a large body of research shows that exposure to it is strongly associated with asthma morbidity (Rosenstreich et al. 1997, Gruchalla et al. 2005, Wang et al. 2009).

Asthma Risk Factors

Beyond the previously discussed racial, socioeconomic, and allergic risk factors, there are many other known risk factors for childhood asthma. Childhood asthma occurs more frequently among males than females, with this sex ratio inverting by adolescence and adulthood (King et al. 2004, Almqvist et al. 2008). This phenomena is supported by the most recent U.S. surveillance data (2008-2010) where among individuals ages 0-17 years prevalence was higher in males than females (11.1% vs. 7.8% respectively) and among individuals 18 years and older the prevalence was higher in females than males (9.7% vs. 5.7% respectively) (Moorman JE et al. 2012). Family history of asthma, and more specifically parental history, is a strong risk factor for childhood asthma (King et al. 2004) which explains its inclusion in guidelines on how to diagnose childhood asthma. Family history and racial risk factors for asthma have spurred research that has found several genetic risk factors for asthma development (Bracken et al. 2002). This field will continue to grow as genetic methods expand with the most recent advances being made by genome-wide association studies (GWAS) (Tamari et al. 2013). As described in a review by Duijts in 2012 there are thought to be several fetal origins of childhood asthma such as being born at low birth weight (Duijts 2012). Most research indicates that exclusive breast feeding and birth by vaginal delivery (in comparison to cesarean section) are both protective against asthma development potentially through influence on immune system development (Gdalevich et al. 2001, Simon et al. 2003, Cho et al. 2013). The role that diet during pregnancy and fetal nutrition play in asthma development remain to be elucidated (Global Initiative for Asthma (GINA) 2009, Duijts 2012). Childhood obesity/high body mass index is an additional risk factor for asthma development (Noal et al. 2011, Papoutsakis et al. 2013).

The lungs develop over an extended period starting during gestation and continuing through childhood. This extended development means both pre- and postnatal exposures may be of importance in asthma incidence. Developing lungs are susceptible to harm by

many inhaled toxicants, the most well-known of which is environmental tobacco smoke (ETS) (Yost et al. 1989, Pinkerton et al. 2000, Wang et al. 2008). ETS exposure can be harmful both pre- and postnally, but research suggests that its greatest impact may be *in utero* (Wang et al. 2008). *In utero* smoke exposure is associated with decreased lung function and an increased risk of developing asthma. The effects of pre- and postnatal ETS highlight the importance of studying the impact of other toxicants in both of these developmental periods. Biological susceptibility and one such toxicant, air pollution, will be discussed in depth in a subsequent section of this proposal.

In conclusion, here is a list of risk factors of childhood asthma: African American race, low socioeconomic status, allergic sensitization, male sex, family history, genetic factors, low birth weight, maternal smoking during pregnancy, delivery by cesarean section, absence of breast feeding, high body mass index, and environmental tobacco smoke.

Asthma Diagnosis for Research

Given the diversity of disorders which can fall under the umbrella term of "asthma," asthma ascertainment for research is challenging. The most frequently used methods of asthma diagnosis for research are: self-report, use of ICD diagnosis codes from patients' medical records, and classification based on medical record information such as ICD diagnoses, symptoms, and asthma-related medications.

Self-report

Self-report is the most common method of determining whether an individual has (or has ever had) asthma and is used in diverse areas of the asthma research field (Litonjua et al. 1999, Castro-Rodriguez et al. 2000, Akinbami et al. 2002, Smith et al. 2005, Gordian et al. 2006, Pearlman et al. 2006, Jerrett et al. 2008). This information is ascertained via

interview or questionnaire either by a standardized asthma questionnaire which includes questions about symptoms and doctor diagnoses or by one question such as “Do you have asthma?” or “Has a doctor ever diagnosed you with asthma?” The logistical advantages of self-report are outweighed by the lack of reliability of the resulting disease classification. Low sensitivity has been found between self-report of asthma via questionnaire and doctor diagnosed asthma (Toren et al. 1993). For childhood asthma identification, there is poor agreement between maternal report of a child having asthma and doctor diagnosed asthma (Miller et al. 2001). This agreement varies based on race/ethnicity, family income, and educational attainment. In one population it was found that parents tended to under-report asthma in their children when comparing diagnosis to a diagnosis from the medical record using a combination of physician diagnosis and symptoms (Yoo et al. 2007).

ICD Codes

Use of ICD codes is common. Asthma covers ICD-9 codes 493.XX, with 493.0X used specifically for extrinsic (atopic) asthma. Given differences between what different doctors will diagnose as “asthma” there is concern about the diversity of disorders which may be included under this code. A study by Juhn 2011 examined which children in a cohort would be classified as having asthma comparing classification by ICD-9 codes alone to classification using a method developed by Yunginger and colleagues in 1992 (Yunginger et al. 1992). The classification scheme by Yunginger uses information available in the medical record to categorize a child as either having no asthma, probable asthma, or definite asthma. It uses ICD-9 diagnosis codes in addition to information on type and history of symptoms such as cough and wheeze. While the classification has been used by other studies, it is not apparent whether it has been validated as a reliable method of identifying children with asthma. The study found that while ICD diagnosis alone identifies a child that

is likely to have asthma, lack of such a diagnosis does not preclude asthma indicated by symptoms in the medical record (Juhn et al. 2011). The authors accordingly concluded that use of ICD codes alone underestimates asthma prevalence.

Medical Record Diagnosis Algorithm

Researchers using medical record data (typically from Health Maintenance Organizations (HMOs)) normally use an algorithm to classify patients as asthmatic or asthma-free based on inpatient asthma diagnoses (i.e. hospital admissions), outpatient asthma diagnoses (i.e. clinic, emergency department), and asthma management medication prescriptions. These algorithms vary considerably from study to study in the types of diagnoses (i.e. primary or secondary) that are considered for classification and the number and types of medications. Table 2.1 details algorithms used in a sample of literature on this topic. The criteria range from easy to difficult to satisfy. For example, in Lieu 1998, in the absence of an ICD diagnosis, a child must have one asthma–related medication (other than a beta-agonist) in the past 12 months to be classified as asthmatic (Lieu et al. 1998). In Verstraeten et al. 2003 and Maher et al. 2004, in the absence of an ICD diagnosis, a child must have 5 asthma-related medications to be classified as asthmatic (with the exception of specific medication combinations which necessitate fewer dispensings) (Verstraeten et al. 2003, Maher et al. 2004).

Studies have been completed examining: 1) classification differences between prescription algorithms (Osborne et al. 1995) and 2) the percent of children diagnosed by prescription information who also report physician-diagnosed asthma (Peled et al. 2006). When researchers base their diagnoses on prescriptions alone they must take precautions not to classify children with well-controlled asthma as not having asthma (Cockcroft et al. 1996). When asthma is well-controlled medications such as β-agonists, generally used to

treat asthma symptoms, may not have been prescribed recently. There should be other medications in the medical record to make it apparent a child has asthma, but only if researchers are looking for a wide enough range of medications when making their determinations.

The origins of classification schemes used by researchers are somewhat mysterious. It appears to be common for researchers to base their algorithms on the algorithms of one or two previous studies in their specific field or to use algorithms developed by HMOs for insurance purposes. Another common practice is to use the Healthcare Effectiveness Data and Information Set (HEDIS) guidelines for "persistent asthma" as an inspiration for algorithm development. The HEDIS guidelines are performance measures used in the managed care industry. According to the guidelines an individual needs one of the following to qualify for persistent asthma: 1) four asthma medications dispensings 2) one emergency department visit or hospitalization with principal diagnosis of asthma or 3) four asthma outpatient visits with two or more asthma medication dispensings (Schatz et al. 2006). Given these guidelines are for "persistent asthma" researchers relax them for a diagnosis of asthma that is not necessarily persistent (for example by requiring two medication dispensings instead of four). This conversion of the HEDIS guidelines to asthma algorithms for research seems arbitrary and not based on use of the final algorithm in any prior research.

Table 2.1. Medical record asthma classification schemes

Article	Population	Criteria Needed	Criteria				
			Inpatient Diagnosis	Outpatient Diagnosis	Any ICD-9 Diagnosis	Asthma Medications	Other
Caudri 2009 (Caudri et al. 2009)	Birth cohort *Children*	1 at age 7 and 1 at age 8			Report of doctor diagnosed asthma	Prescription for inhaled steroids	≥1 wheezing episode
Celedón 2004 (Celedón et al. 2004)	HMO in Boston *Children 2-5*	1	1 hospitalization	2 ambulatory or 1 emergency department		2 dispensings in past 12 months	
Clark 2010 (Clark et al. 2010)	British Columbia *Children*	1	1 hospitalization	2 primary care diagnoses in 12 months			
Finkelstein 2000 (Finkelstein et al. 2000)	HMO *Children 2-18*	1	1 hospitalization	1 ambulatory visit or 1 emergency department encounter		1 asthma dispensing	
Firoozi 2010 (Firoozi et al. 2010)	Quebec *Adults*	Both			493 (except 493.2)	1 in past 2 years	
Grana 1997 (Grana et al. 1997)	U.S. Healthcare HMO *All ages*	1			493.00-493.99	2	Procedure (i.e. asthma care/control program, or theophylline)
Lafata 2002 (Lafata et al. 2002)	HMO *Children 5-14*	1	1 hospitalization	2 outpatient (ambulatory or emergency department)			
Lieu 1998 (Lieu et al. 1998)	Kaiser Northern California *Children*	1	1 hospitalization	1 emergency department visit		1 (if β-agonist also needed diagnosis or other medication)	

Table 2.1. Medical record asthma classification schemes *(Continued)*

Maher 2004 (Maher et al. 2004)	HMO *Children* *(0-18 months)*	1				Any of the following: *(a)* 1 diagnosis and 1 medication dispensing *(b)* 1 dispensing inhaled β-agonist and 1 dispensing cromolyn/ nedrocromil sodium or inhaled corticosteroids *(c)* 5 medication dispensings	
Milton 1998 (Milton et al. 1998)	HMO *Adults (15-55)*	1	1 hospitalization with primary diagnosis	1 emergency department visit for asthma	1 occupational asthma diagnosis (specific to research question)	1 outpatient diagnosis and 1 of the following dispensings: 2 β-agonist inhalers, 1 β-agonist inhaler with theophylline, 1 steroid or cromolyn inhaler, 1 steroid taper, outpatient IV theophylline or nebulized β-agonists treatment	
Peled 2006 (Peled et al. 2006)	HMO *Children*	1				2 dispensings in 1 year	
Schatz 2004 (Schatz et al. 2004)	Kaiser Southern California *Children & Adults*	1	1 hospitalization	1 emergency department or clinic		2 dispensings in 1 year (some medications excluded e.g. oral corticosteroids)	
Schatz 2006 (Schatz et al. 2006)	Kaiser Southern California Persistent Asthma *Adults (18-56)*	1	1 hospitalization with primary diagnosis	1 emergency department primary diagnosis or 4 outpatient and 2 dispensings		4 dispensings	

Table 2.1. Medical record asthma classification schemes *(Continued)*

Shields 2002 (Shields et al. 2002)	5 providers including 1 HMO *Children 2-18*	1	1 hospitalization	2 ambulatory primary diagnoses or 1 emergency department		2 asthma medications with ≥1 asthma visit and 1 β-agonist	
Shields 2003 (Shields et al. 2003)	HMO and primary care clinician plan *Children >2 years*	2 total: 1 diagnosis and 1 additional criterion	Any 2 visits with primary diagnosis in 1 year	1 hospitalization or emergency department visit with primary diagnosis	ICD-9-CM 493.00-493.99	2 β-agonist or 1 β-agonist and 1 asthma visit	
Sinclair 2006 (Sinclair et al. 2006)	Kaiser Atlanta *Adults*	1	2 primary diagnosis	2 primary or secondary diagnoses		3	
Siwik 2005 (Siwik et al. 2005)	Health Alliance Plan HMO *Children 6-8*	1	1 hospitalization	2 outpatient or 1 emergency department		2 bronchodilator or 1 anti-inflammatory (prescription claim)	
Verstraeten 2003 (Verstraeten et al. 2003)	HMO *Children*	1			1 ICD-9 code + 1 medication prescription	5 total or 1 β-agonist and 1 cromolyn prescription	
Vollmer 2005 (Vollmer et al. 2005)	Kaiser Northwest *Adults*	2 total: 1 utilization and 1 medication	1 inpatient with primary diagnosis	1 emergency department or 1 outpatient or urgent care visit		2 β-agonist inhaler dispensings or 1 other medication dispensing	
Zeiger 2008 (Zeiger et al. 2008) Zeiger 2010 (Zeiger et al. 2010)	*Adults*	1	1 hospitalization	1 emergency department or 1 outpatient		2 dispensings (excluding oral corticosteroids)	

AIR POLLUTION

Air pollution is a mixture of gases and particles in the air originating from natural and anthropogenic sources. Natural pollutant sources include biogenic processes (e.g. plant decomposition), volcanoes, lightning, and forest fires. Anthropogenic pollutant sources include coal combustion (i.e. coal burning power plants, motor vehicles), pesticide applications, and industrial leaks. Pollutants can be either primary (emitted by a source), secondary (resulting from chemical reactions in the atmosphere), or both. Nitrogen dioxide (NO_2) is an example of a pollutant that is both primary and secondary. It is emitted from combustion processes, but is also formed from the reaction of nitric oxide (NO) and ozone (O_3). Ozone, on the other hand, is strictly a secondary pollutant.

The two main categories of air pollution are gas-phase and particulate. There are hundreds of different gas-phase substances in the atmosphere, very few of which threaten human health. Health effects are determined by compound toxicity, atmospheric residence time and concentration. Two important gas-phase pollutants are sulfur dioxide (SO_2) and ozone (O_3). SO_2 is a gaseous pollutant produced from combustion of coal and other sulfur-containing fuels. It also has natural sources such as volcanic eruptions and can be created by atmospheric reactions. SO_2 is associated with a wide range of respiratory symptoms and outcomes in children and other populations (U.S. Environmental Protection Agency 2013). O_3 is a secondary pollutant and is one of the most toxic components of air pollution. It is the main constituent of smog and reaches its highest concentrations on sunny days. Additional gas-phase pollutants (nitric oxide (NO), nitrogen dioxide (NO_2), and carbon monoxide (CO)) will be discussed in subsequent sections.

Particulate matter (PM) is matter suspended in air such as smoke, soot, sea salt, and pollen. The specific constituents of PM vary by region. For instance, in areas near the ocean sea salt leads to higher amounts of chloride in PM than in PM from other regions. Source

apportionment can be completed to determine source contributions to particulate matter samples. Particulate matter can be classified based on composition, for example $PM_{2.5}$ nitrate and $PM_{2.5}$ elemental carbon. Classifications based on composition are important since the specific components can dictate what effect PM has on human health outcomes. Particulate matter can also be classified based on aerodynamic diameter. Such classifications are: ultrafine <100 nm, fine 0.1-2.5 μm, coarse >2.5 - 10 μm, $PM_{2.5}$ ≤2.5 μm, and PM_{10} ≤10 μm. These classifications are also important when considering health outcomes. Particles of different sizes deposit in different parts of the respiratory tract which can dictate potential biological effect. Coarse particles cannot move past the upper airways, while smaller particles can end up in lung alveoli. Mouth breathing, in children and others, increases particulate matter exposure because air avoids the nasal passage whose mucus and hair act as barriers to prevent matter from ending up further into the respiratory system.

Air pollution varies spatially and temporally. Spatial variation is determined by emission sources, pollutant chemistry, and transport of pollutants and precursors. Pollutants with considerable spatial heterogeneity include carbon monoxide (CO), nitrogen oxides (NO_X) and PM elemental carbon (EC). More spatially homogenous pollutants which vary on a larger regional scale include $PM_{2.5}$, sulfate (SO_4), and ozone (O_3). Temporal variation is driven by human activities and meteorology. For example, in most regions, ozone is highest during the summer months because it is a secondary pollutant and sunlight is necessary for its formation. Pollutant variations are not independent and many co-vary based on common sources and participation in the same chemical reactions.

Air pollution has serious effects on human health and the environment. It has been linked to a broad range of negative health outcomes and impacts several biological systems including respiratory, cardiovascular, reproductive, and central nervous systems.

Anthropogenic pollutant sources are of the most concern for human health outcomes. In the United States, the National Air Quality Standards (NAAQS) set standards for six criteria pollutants which are known to be harmful to the environment and human health: CO, lead, NO_2, O_3, sulfur dioxide (SO_2), and particulate matter 2.5 and 10 ($PM_{2.5}$ and PM_{10}). The Harvard Six Cities Study helped identify the health effects of particulate matter specifically. It was a landmark longitudinal study started in 1974 which found a strong linear association between particulate matter and all-cause mortality when adjusting for other risk factors (Dockery et al. 1993). The results were particularly surprising because an increase in mortality was seen with particulate matter concentrations that were not what we would consider to be high.

Traffic-Related Pollutants

Traffic-related air pollution refers to pollutants produced by vehicular traffic and includes both particulate and gaseous air pollution. Important traffic-related pollutants are carbon monoxide (CO), particulate matter (PM), nitric oxide (NO), nitrogen dioxide (NO_2), and $PM_{2.5}$ elemental carbon (EC). Primary traffic pollutants react and form secondary pollutants (e.g. ozone and nitrates). Vehicle emissions depend on many factors such as vehicle type and age, fuel type (i.e. diesel, gasoline, alternative fuel), maintenance, and driving conditions. While we normally focus on combustion emissions, vehicles also produce non-combustion emissions (e.g. tire wear and resuspended road dust) (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010). Pollutant concentrations near roadways depend on amount and type of traffic, wind patterns, and meteorology. Exhaust concentrations can decline substantially even within a couple of hundred meters of the roadway. Traffic-related air pollution can decrease concentrations of other pollutants near

roadways. For example, ozone is lowest near roadways because it is scavenged by NO to create NO_2.

Strides have been made to reduce vehicle emissions and miles traveled, both which directly impact vehicle-related pollution. Such strides include regulation of certain emission components (e.g. CO, hydrocarbons (HC), NO_X, and PM), vehicle inspection legislation, and production and use of cleaner fuels. Despite these advances, traffic continues to be a significant contributor to global air pollution and the size of the global motor vehicle fleet continues to increase with gross domestic product (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010).

Nitrogen Oxides (NO_X)

Transportation is the major producer of NO_X now that there are emissions requirements on stationary coal combustion. NO_X is the sum of NO and NO_2 and is commonly used for reporting since NO and NO_2 can rapidly interconvert due to changes in ozone and sunlight (Harrison 1999). NO_2 is more toxic than NO and is one of the most important gas-phase emissions from a public health perspective. It has harmful health effects and is also a precursor to ozone. Research shows NO_2 to be associated with respiratory infections, lung function, and potentially even increased death rates (Godish 2004).

Traffic-Related Particulate Matter

Traffic is a considerable contributor to particulate air pollution. This matter results from both vehicular combustion and non-combustion emissions. $PM_{2.5}$ elemental carbon (EC) and $PM_{2.5}$ organic carbon (OC) are carbonaceous aerosols and classes of traffic emitted particulate matter. EC is light absorbing while OC is light scattering (Szidat S et al. 2009).

Particles containing elemental carbon are on average fine and ultrafine mode particles enabling lung penetration. Consequently, EC is associated with negative respiratory outcomes such as decreased lung function in children and lung cancer (Barraza-Villarreal et al. 2011, Garshick et al. 2012). EC is primarily created from fuel combustion and makes up a small portion of $PM_{2.5}$ mass in the United States, ranging from 5% to 14% (Godish 2004). In the United States, up to a third of $PM_{2.5}$ mass consists of OC which is produced from traffic emissions as well as biogenic sources (Godish 2004).

Carbon Monoxide (CO)

Carbon monoxide is a gas-phase pollutant that is deadly at high concentrations well beyond what is observed in ambient air. It is a by-product of incomplete fossil fuel combustion. Due to transportation emissions, concentrations are highest in metropolitan areas and have a strong diurnal pattern peaking at rush hour. Concentrations are highest near major roadways and diminish rapidly with increased distance from roadways. CO also has yearly variation with the highest concentrations in the winter due to less efficient sink processes which function to remove CO from the atmosphere (Godish 2004).

Pollution Data & Health Research

For research and regulation purposes, air quality data can either be measured or modeled. One source of measured air quality data in the United States is the network of local and state air pollution monitors. The U.S. Environmental Protection Agency Air Quality System aggregates and makes accessible data from pollution monitors across the country. Measured data can also be collected using personal exposure monitors to capture individual exposure. While this is a cumbersome process, its advantage is that it can fully capture exposure from the many environments with distinct pollution profiles that an individual

may spend time in throughout the day. An advantage of modeled data is that it can estimate pollution concentrations in locations where there are not existing pollution monitors while taking into account meteorology and other factors. One Eulerian air pollution model, Community Multiscale Air Quality (CMAQ), will be discussed in depth later on in this proposal (Chapter 3. Data Sources).

Historically, air pollution health research has focused on the effects of individual pollutants. This approach is still the predominant one, for example, with studies looking at the effects of 12 different pollutants and pollutant categories: CO, NO_2, SO_2, O_3, PM_{10}, $PM_{2.5}$, $PM_{2.5-10}$ (PM with aerodynamic diameter from 2.5 to 10 μm), $PM_{2.5}$ sulfate, $PM_{2.5}$ nitrate, $PM_{2.5}$ elemental carbon, $PM_{2.5}$ organic carbon, $PM_{2.5}$ water-soluble metals (chromium, copper, iron, manganese, nickel, and vanadium) (Darrow et al. 2011). The fact that we monitor individual pollutants rather than pollutant mixtures lends itself to this approach. Additionally, results from these studies are useful for the creation of regulations and single pollutant standards. More recently there has been an effort to characterize the effects of multi-pollutant mixtures rather than individual pollutants. A 2010 article by Dominici and colleagues describes this approach, its strengths and challenges (Dominici et al. 2010). The objective of this research is to identify multi-pollutant combinations which pose a threat to human health. The rationale behind this approach is that we are exposed to pollutant mixtures rather than individual pollutants in isolation. The health effects of such mixtures may be different than the sum of effects of individual pollutants (Dominici et al. 2010). There are many ways to implement this approach, for example by focusing on pollutants from specific sources or by grouping pollutants by hypothesized mechanism of effect.

PRENATAL AND EARLY-LIFE AIR POLLUTION EXPOSURE AND ASTHMA INCIDENCE

Biological Susceptibility and Potential Mechanisms

Environmental exposures during gestation and early life have the potential to impact the maturation of the respiratory and immune systems, both of which are important in asthma development. The respiratory system develops over an extended period, beginning six weeks after conception and continuing through adolescence. During this development, repair mechanisms are not as adept at responding to environmental insults as those in mature adult lungs (Kajekar 2007). The significance of the prenatal window in particular is supported by research which shows that children who develop asthma by the age of seven have 40% of their associated lung deficit at birth (Bisgaard et al. 2012). Lung development after birth is primarily composed of growth of additional bronchioles and alveoli and is critical for the lungs to meet the increasing metabolic demands of a growing child (Pinkerton et al. 2000, Moore et al. 2003, Wang et al. 2008). Environmental exposures during this time, particularly during gestation and the first 2 years of life, can also have an impact on the development of a child's immune programming and response (Peden 2000). The development that occurs during these periods is paired with children's greater exposure to air pollution relative to adults due to increased ventilation rates and more time spent outdoors (American Academy of Pediatrics 1999, Pinkerton et al. 2000).

There are several mechanisms by which air pollution may lead to asthma development which make the potential of a causal relationship biologically possible. The following two mechanisms are particularly compelling (Gowers et al. 2012): 1) *Oxidative stress*: Some components of air pollution are free radicals while others can participate in reactions that result in free radicals. Oxidative stress is the result of an imbalance of free radical production in the lungs and antioxidant defenses available to deal with the free radicals (Kelly 2003). This can result in inflammatory reactions and airway injury. The Health Effects Institute (HEI) Report concluded that oxidative stress is one mechanism through which air pollution impacts human health (HEI Panel on the Health Effects of

Traffic-Related Air Pollution 2010). 2) *Airway remodeling*: Airway remodeling is when structural changes in the airways occur caused by prolonged inflammation. Such changes in the lungs may lead to asthma development. Additional potential mechanisms of effect are through changing inflammatory mediators and immunological responses, and increasing allergen sensitization (Gowers et al. 2012). One hypothesis is that air pollution exposures trigger irritative inflammatory changes in the airways instead of allergic changes (Gruzieva et al. 2013). If this was the case we would expect to see a stronger association between air pollution and nonatopic asthma than among air pollution and atopic asthma.

A growing area of research is the exploration of the role of gene-environment interactions in asthma development. It is well established that some individuals are genetically pre-disposed to asthma. Recent research suggests that certain genetic factors may make an individual more susceptible to air pollution's effects (McLeish et al. 2007, Melen et al. 2008). In particular, genes related to antioxidative systems and inflammatory responses are of interest. Some research has indicated that polymorphisms in genes of both of these types modify the asthma risk resulting from ozone exposure (Yang et al. 2009).

Conclusions of Key Review Articles

While it is well established that air pollution exposures exacerbate pre-existing asthma, a growing body of literature suggests that this exposure may also play an important role in asthma development (Tzivian 2011, Gowers et al. 2012). A number of recent review articles and meta-analyses have examined the association between air pollution exposure and asthma incidence, the majority of which focus on the effects of traffic-related air pollution.

Two formal reports that review published studies have been produced on the topic. The first report, published in 2010 by HEI, concluded from their review of the literature that

residential proximity to busy roads is a risk factor for childhood asthma development (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010). The authors concluded that the evidence for this relationship to be causal was borderline between "sufficient" and "suggestive but not sufficient." A more recent report was published in 2012 by the UK Department of Public Health's Advisory Committee on the Medical Effects of Air Pollutants (Gowers et al. 2012). It concluded that asthma induction is not associated with community-level air pollution. The report did find evidence that residential proximity to traffic-related air pollution may play a role in asthma induction particularly for individuals living near busy roads with heavy goods vehicle traffic. The authors concluded that air pollution likely plays a small role in asthma induction when comparing it to other causal factors. Among other sources, this report relied on the results of a systematic review and meta-analysis by Anderson and colleagues to come to these conclusions (Anderson et al. 2013). The main meta-analysis in the Anderson article included assessed the outcome of incidence and prevalence of asthma and wheeze in children and adults. This analysis found a significant association between NO_2 exposure and this heterogeneous outcome (OR (95% CI): 1.07 (1.02, 1.13)). Meta-analyses looking at NO_2 and more restricted outcomes (i.e. only children, only asthma, and only incidence) were also significant, but no one analysis included all three of these outcome restrictions. The article concluded that air pollution exposure impacts asthma incidence.

To date, four review articles have examined the association between traffic-related air pollution and asthma occurrence, two of which included meta-analyses on the association. In 2008, a review by Salam and others examined publications between 2006 and 2007 and concluded that residential proximity to busy roads is associated with asthma development and exacerbation (Salam et al. 2008). In 2009, a review by Bråbäck and Forsberg examined results from 15 papers on 10 cohort studies with the outcomes of

asthma and allergic sensitization (Bråbäck et al. 2009). They concluded that in children traffic exhaust is associated with the development of respiratory symptoms. While they made no formal conclusions regarding asthma incidence, the majority of included studies that examined this outcome saw increased incidence associated with pollution exposure. In 2012 Gasana and colleagues conducted a review and meta-analysis which included 3 articles examining traffic-related air pollution asthma incidence (Gasana et al. 2012). Separate meta-analyses found positive associations between NO_2, PM_{10}, and $PM_{2.5}$ exposure and asthma incidence, but only the association with NO_2 was statistically significant (OR (95% CI): 1.14 (1.06, 1.26)). Most recently, in 2015 Bowatte and colleagues conducted a systematic review and meta-analysis on the association between traffic-related air pollution and childhood asthma, using only results from birth cohorts (Bowatte et al. 2015). They identified five birth cohorts with results on the association between longitudinal childhood exposure to air pollution and asthma development. Their meta-analyses found some evidence of an association between nitrogen oxides and asthma, particulate matter and asthma, and black carbon and asthma. However, many of their meta-analyses showed evidence of heterogeneity. All five birth cohorts included in their meta-analyses are discussed in the next section of this chapter and included in Table 2.2.

While these reviews and reports are key to our understanding of air pollution and asthma, none of them focus specifically on air pollution exposure during prenatal and early life periods.

Results of Key Studies

Table 2.2 summarizes nine key studies on prenatal and early life traffic-related air pollution exposure and childhood asthma incidence. The majority of these studies are either birth cohorts or studies nested within birth cohorts, such as nested case control studies.

Exposure estimation ranged from fairly basic to sophisticated. On the basic side, some studies defined exposure using distance of residence to roadway or area truck route density. This information was either calculated using residential address or obtained via self-report in the case of Zhou et al. 2013 where participants were asked questions such as "[Do] cars often or continuously pass by your house?" (Zhou et al. 2013). On the sophisticated side of exposure assessment, many studies used land-use regression models. When possible, studies broke down exposure into pollution components: NO_2, NO, particulate matter, and black (elemental) carbon or soot. As with many asthma studies, the majority of these studies defined asthma incidence using parental report of asthma obtained via questionnaire. Two studies that used alternate asthma diagnosis methods were completed by Clark and colleagues in 2010 and Carlsten and colleauges in 2011. Clark et al. used physician billing records and hospital discharge records to determine asthma diagnosis (Clark et al. 2010). Carlsten et al. had a pediatric allergist blinded to child exposure examine each child at age seven and use a symptom-based algorithm to determine asthma status (Carlsten et al. 2011).

Out of the nine articles in Table 2.2, seven of them came to the conclusion that prenatal or early life traffic-related air pollution was associated with asthma development. Significant associations were found by at least one study for every pollutant examined ($PM_{2.5}$, PM_{10}, NO, NO_2, NO_x, and black carbon) as well as for exposure metrics such as distance from roadway. While some of the strongest associations were seen for particulate matter exposure (in particular by Carlsten et al. 2011 and Gruzieva et al. 2013), this was not consistent across all studies. The strongest effects from these nine studies were seen in Carlsten et al. 2011 which was completed in a population of children at high risk for asthma based on family history of asthma or allergic diseases such as atopic dermatitis, seasonal or food allergies. These results contribute to our knowledge that air pollution can be

particularly harmful to susceptible or vulnerable populations. Krämer et al. and Oftedal et al. were the two studies which found no association between traffic-related air pollution and asthma development (Krämer et al. 2009, Oftedal et al. 2009). Krämer et al. did not see an overall effect of pollution on asthma, but did see an effect when restricting analyses to residents of Munich, the area with the highest pollution in their study. The results in Oftedal et al. indicated a preventive effect of air pollution, but were not statistically significant. The authors speculated that their results could be explained by selection issues and diagnostic misclassification of asthma.

In addition to the nine studies in Table 2.2, there have been many additional studies on related associations. Several articles have examined the association between prenatal and early life traffic-related air pollution exposure and allergic and respiratory outcomes such as wheeze, respiratory symptoms, allergic sensitization, eczema and lung function. Others have looked at traffic-related air pollution exposure and asthma development, but focused on exposure outside of the prenatal and early life periods.

This literature review has focused on the long-term respiratory effects of chronic exposure to air pollution from traffic. Fewer studies have assessed whether chronic exposure to air pollution from all sources, not just traffic, is associated with asthma incidence. However, there is some evidence of an association between total air pollution exposure and asthma incidence. A 2002 study in Japan saw evidence of an association between annual-average NO_2 concentrations and asthma development, but no association between PM_{10} concentrations and asthma development (Shima et al. 2002). A more recent study found that exposure to $PM_{2.5}$ during pregnancy was associated with asthma incidence by age 6, but only among boys (Leon Hsu et al. 2015).

Effect Modification by Atopy and Socioeconomic Status

There is some indication in the literature that air pollution exposure has a stronger effect on asthma development among children without atopy than among children with atopy. These results align with results from research which has shown stronger associations of air pollution (indoor and outdoor) and secondhand smoke with respiratory symptoms, asthma prevalence and asthma symptom exacerbation in nonatopic children (Strachan et al. 1998, Hirsch et al. 1999, Kattan et al. 2007). The results suggesting the same trend with air pollution and asthma incidence are very preliminary and not entirely consistent. Gehring and colleagues in 2010 saw that the effect of air pollution on asthma incidence may be limited to nonatopic asthma where this phenotype was defined as asthma without evidence of sensitization (Gehring et al. 2010). However, this study had small numbers of children with each phenotype. The results of Gruzieva and colleagues in 2013 also suggest that the effect may be strongest for nonallergic asthma particularly at older ages (where nonallergic asthma was defined the same way as in Gehring et al. 2010). At age eight, the association with both first year of life NO_x and PM_{10} was much stronger for nonallergic asthma than for allergic asthma. They also saw a difference at four years of age, but the effect estimates were more similar (Gruzieva et al. 2013). An additional study by McConnell in 2006 saw larger effects of distance from current residence to major road on asthma prevalence and wheeze among children without a history of allergic symptoms (McConnell et al. 2006). Nevertheless, results on the topic as a whole are inconsistent. Other studies have found either no difference in effect between sensitized and non-sensitized children, or even a stronger association between pollution and atopic asthma than pollution and nonatopic asthma (Annesi-Maesano et al. 2007, Brauer et al. 2007).

The small body of research that has examined whether the impact of air pollution exposure on asthma differs by socioeconomic status (SES) has come to disparate conclusions. SES is an important confounder of the relation between air pollution and

asthma. Air pollution is spatially heterogeneous with the highest pollution often occurring in low SES areas. As discussed previously, asthma occurs at higher rates among individuals from low SES backgrounds. A study by Shankardass and colleagues in 2009 looked at whether children from families with low SES and with high parental stress were more susceptible to the impact of traffic-related air pollution on asthma development (Shankardass et al. 2009). They found that children from families with these characteristics were more susceptible, as evidenced by larger hazard ratios of effect. The authors hypothesized that individuals with higher levels of stress may respond differently to oxidative burden and that this may explain both results as low SES environments typically have higher levels of stress. Two additional studies have found that the effect of air pollution on childhood asthma hospitalizations and ambulatory visits is greater for children of low SES (Neidell 2004, Burra et al. 2009). The authors of one study hypothesized that pollution could be a potential mechanism of the effect of SES on asthma (Neidell 2004). Two recent studies did not find effects to differ by SES, one looking at the impact of air pollution on asthma incidence, and the other looking at the impact of air pollution on asthma medication sales (sales of short-acting β-agonists specifically) (Zmirou et al. 2004, Laurent et al. 2009). While this is in no way an exhaustive look at all studies of effect modification by SES, these results indicate that this is a relationship worth studying.

Limitations of Previous Research

A major limitation of much of the previous research in this area is the absence of high-quality data on both air pollution and also asthma. There is a limited amount of research that focuses specifically on air pollution exposure during pregnancy and early life despite the importance of these windows developmentally. For exposure assessment, the reliance on residence at birth is problematic. Exposure estimation based solely on residence

(at any time point) ignores the fact that our daily movements mean our location is not always in close proximity to residence. This can lead to misclassification of exposure particularly if the places one spends time have substantially different pollution profiles than that at one's residence. The use of residential address may be less of an issue when studying exposure during infancy, but more of an issue once a child is attending school and spending less time at home. The reliance solely on address at birth is also problematic since it is common for families to move during pregnancy and the first few years of a child's life. Reasons for moving during this time include the need to accommodate a growing family, changing financial resources, and the desire to relocate to neighborhoods better suited for children (Saadeh et al. 2013).

Asthma ascertainment in many studies suffer from the use of self-reported asthma and from the lack of information about timing of onset. Asthma is a heterogeneous outcome. Different studies classify asthma in different ways making it possible they are studying disparate biological outcomes. Many studies also lack adequate control for important potential confounders such as SES and family history of asthma. Varying asthma classification schemes and control for different potential confounders limit the ability to meaningfully compare results between studies. Future studies should examine whether some individuals are more susceptible to air pollution's effects, for example based on gender, SES, or genetic background (Salam et al. 2008).

Table 2.2. Summary of key articles on prenatal and early life traffic-related air pollution and childhood asthma incidence

Carlsten 2011 (Carlsten et al. 2011) – Vancouver, Canada		
Study Details • Birth cohort of children born in 1995 at high risk for asthma development • Children randomized to study arms: control arm with usual care, intervention arm with care aimed at asthma prevention • 100 intervention children, 84 control Exposure • Estimated exposure during birth year to traffic-related NO, NO_2, black carbon, and $PM_{2.5}$ • Land use regression model. Estimated annual pollutant concentrations using samples from multiple sites in the study area in 2003, residential address at birth, and 55 descriptive variables	Outcome • Asthma diagnosed at age 7 by a pediatric allergist blinded to child information • Asthma definition: ≥2 episodes of cough (each lasting ≥2 weeks), ≥2 distinct episodes of wheeze (each lasting ≥1 week), plus ≥1 of the following: nocturnal cough in the absence of a cold (at least once a week), hyperpnoea-induced cough or wheeze, or response to drug treatment Analysis • Logistic regression for IQR increase of exposure • Covariates: maternal education, maternal, paternal or sibling asthma, atopy at 1 year, intervention status, gender, ethnicity	Main Findings for Asthma **OR (95% CI) control arm** $PM_{2.5}$: 4.1 (1.2, 13.8) NO: 1.7 (1.0, 2.8) NO_2: 1.9 (0.9, 4.0) Black carbon: 1.2 (0.6, 2.3) **OR (95% CI) intervention arm** $PM_{2.5}$: 2.1 (0.6, 7.1) NO: 0.8 (0.4, 1.6) NO_2: 1.3 (0.6, 2.9) Black carbon: 1.1 (0.5, 2.5)
Clark 2010 (Clark et al. 2010) – British Columbia		
Study Details • Population-based nested case control study. 5 controls matched to each incident asthma case on sex, and month and year of birth • Cohort: births in 1999 and 2000. Inclusion: on provincial medical plan, live in area during pregnancy and first year of life. 3,482 cases; 17,410 controls Exposure • Average prenatal and first year of life exposure • Exposures to NO, NO_2, $PM_{2.5}$, black carbon assigned from residential zip code using land use regression (LUR) models, monitoring data, distance to stationary sources *[Separate analyses on point source exposure instead of LUR not in table]*	Outcome • Records-based asthma diagnosis using physician billing records from primary care visits and hospital discharge records • Asthma definition: ≥2 primary care diagnoses in 12 month period, or 1 hospital admission for asthma Analysis • Conditional logistic regression. OR for IQR increase. Prenatal and first year of life exposure included in same model when not highly correlated. • Covariates: native status, breast feeding, maternal smoking and age, income, education, birth weight, gestational length	Main Findings for Asthma **OR (95% CI) in utero exposure** NO: 1.05 (1.02, 1.09) NO_2: 1.02 (0.97, 1.07) $PM_{2.5}$: 1.02 (1.00, 1.03) Black carbon: 1.08 (1.02, 1.15) Road proximity: 0.97 (0.82, 1.15) **OR (95% CI) first year of life exposure** NO: 1.03 (1.00, 1.07) NO_2: 1.13 (1.04, 1.23) $PM_{2.5}$: 1.01 (0.99, 1.03) Black carbon: 1.14 (1.01, 1.29) Road proximity: 1.01 (0.84, 1.22)

Table 2.2. Summary of key articles on prenatal and early life traffic-related air pollution and childhood asthma incidence *(Continued)*

Gehring 2010 (Gehring et al. 2010) – The Netherlands		
Study Details • PIAMA Study Prospective Birth Cohort • Lifestyle intervention assessed for allergy and asthma prevention. Children followed from years 1 to 8 (n=3,863). • Natural history arm included allergic and nonallergic mothers. Intervention arm included only allergic mothers. Exposure • Estimated long term average ambient pollution concentration at birth address using land-use regression models (NO_2, $PM_{2.5}$, soot) • Used monitoring data from 40 study sites, and data on traffic, road, and population density near monitoring location for estimation	Outcome • Parental report of asthma diagnosis at ages 1 to 8 collected yearly via questionnaire • Atopic asthma = prevalent asthma in past 12 months and being sensitized. Nonatopic asthma = prevalent asthma in the past 12 months without sensitization. Analysis • Longitudinal logistic regression using GEE (correlation between yearly outcomes). OR per IQR increase. • Sensitivity analyses: potential confounding by region, impact of study design, interaction between air pollution and whether a child moved • Covariates: gender, study arm, parental allergies and education, maternal smoking during pregnancy, breastfeeding, siblings, home environment, day care, region, Dutch nationality, health characteristics	Main Findings for Asthma **OR (95% CI) for asthma in first 8 years** $PM_{2.5}$: 1.26 (0.97, 1.63) NO_2: 1.17 (0.96, 1.42) Soot: 1.17 (0.95, 1.42) **Crude OR (95% CI) nonatopic asthma** $PM_{2.5}$: 1.85 (0.92, 3.73) NO_2: 2.98 (1.21, 7.37) Soot: 2.06 (0.99, 4.30) **Crude OR (95% CI) atopic asthma** $PM_{2.5}$: 0.95 (0.64, 1.40) NO_2: 1.00 (0.63, 1.58) Soot: 0.97 (0.64, 1.46)
Gruzieva 2013 (Gruzieva et al. 2013) – Sweden		
Study Details • Swedish birth cohort BAMSE. Children born between 1994 and 1996 (n=3,633) Exposure • First year of life exposure, current exposure, average exposure since last follow-up. • Roadway PM_{10} and NO_X during first year of life • Calculated time weighted outdoor air pollution exposure using residence, daycare, and school addresses, and emission data. Gaussian dispersion model used to calculate traffic-PM_{10} and traffic-NO_X.	Outcome • Asthma ages 1, 2: 3+ episodes of wheeze and inhaled corticosteroids treatment or bronchial hyper-reactivity in absence of respiratory infection • Asthma ages 4, 8, 12: 4+ episodes of wheeze or 1+ episode and prescribed inhaled corticosteroids • Incident asthma = first questionnaire when fulfill asthma requirements • Allergic asthma = fulfill asthma requirements and evidence of allergen sensitization in blood IgE levels Analysis • Multinomial logistic regression, GEE. Modeled exposure as 5^{th} to 95^{th} percentile difference. • Results for asthma at ages 1, 2, 4, 8, and 12 • Assessed time-exposure interaction and effect modification by sex and allergic heredity • Covariates: municipality, SES, heredity (parental asthma, asthma medication, hay fever, allergies), year house built	Main Findings for Asthma **OR (95% CI) first year of life NO_X** Overall effect: 1.21 (0.79, 1.84) 1 year: 0.85 (0.44, 1.62) 2 years: 0.96 (0.51, 1.80) 4 years allergic: 1.5 (0.4, 5.1) 4 years nonallergic: 2.4 (1.0, 5.6) 8 years allergic: 0.8 (0.2, 2.4) 8 years nonallergic: 2.6 (0.9, 8.1) 12 years: 1.87 (1.01, 3.44) **OR (95% CI) first year of life PM_{10}** Overall effect: 1.34 (0.80, 2.23) 1 year: 0.79 (0.39, 1.62) 2 years: 1.14 (0.57, 2.25) 4 years allergic: 1.4 (0.3, 6.8) 4 years nonallergic: 1.6 (0.5, 5.3) 8 years allergic: 1.1 (0.3, 3.8) 8 years nonallergic: 3.8 (0.9, 16.2) 12 years: 2.39 (1.18, 4.86)

Table 2.2. Summary of key articles on prenatal and early life traffic-related air pollution and childhood asthma incidence *(Continued)*

Krämer 2009 (Krämer et al. 2009) – Germany		
Study Details • German birth cohorts: GINIplus, LISAplus. Children followed-up at ages 1, 2, 3, 4 and 6 (n=2059) Exposure • Estimated traffic-related air pollution at birth address: soot ($PM_{2.5}$ absorbance), NO_2 • Whether distance to major road <50 meters • Land-use regression models used residence at birth to model air pollution using 2002 measurements	Outcome • Parental report of physician-diagnosed asthma Analysis • Cox regression - exposure from birth address • Covariates: parental allergy and education, gender, maternal smoking in pregnancy, smoking in home, furry animal contact in first year of life, elder siblings, gas cooking, home dampness, indoor molds, living on a farm	Main Findings for Asthma **RR (95% CI) from Cox regression asthma by age 6** Soot: 1.16 (0.87, 1.54) NO_2: 1.17 (0.86, 1.58) <50m: 0.86 (0.66, 1.14)
Oftedal 2009 (Oftedal et al. 2009) – Oslo, Norway		
Study Details • Cross-sectional study of children from the Oslo Birth Cohort who resided in Oslo in the first year of life and during the year prior to the questionnaire • Cohort: children born 1992-1993. Followed up at ages 9 to 10 (n=2,871) Exposure • NO_2 exposure during first year of life and before asthma onset calculated based on residence • EPISODE dispersion model (3D Eulerian/Lagrangian dispersion model) used to estimate NO_2 exposure • Distance from birth residence to major roadway	Outcome • Parental report of physician-diagnosed asthma via questionnaire • Onset categorized into: before age 4, before age 9-10, between age 4 and 9-10 Analysis • Modeled IQR increase of exposure with Cox proportional hazard model and logistic regression. Completed some models smoothing exposure with cubic splines. • Assessed interaction with: sex, parental atopy, skin prick test results, cohort indicator • Covariates: sex, parental education and atopy (maternal or paternal history of asthma, hay fever, or eczema), maternal smoking during pregnancy and marital status, cohort population	Main Findings for Asthma **RR (95% CI) from Cox PH regression** NO_2 1st year life: 0.82 (0.67, 1.02) NO_2 before onset: 0.82 (0.67, 1.02) Distance to major road: 0.99 (0.90, 1.08) **OR (95% CI) from logistic regression** NO_2 1st year life: 0.81 (0.65, 1.02) Distance to major road: 0.98 (0.89, 1.08) *[Also presented results for onset <4 years of age, ≥4 years of age]*

Table 2.2. Summary of key articles on prenatal and early life traffic-related air pollution and childhood asthma incidence *(Continued)*

Patel 2011 (Patel et al. 2011) – New York City, NY, USA		
Study Details • Columbia University Birth Cohort. Enrollment 1998-2006 (n=727) • Maternal inclusion: nonsmoking, Dominican or African American ethnicity, residence in Northern Manhattan or South Bronx Exposure • Exposure during prenatal period, ages 1, 2, 3, and 5 collected for 250 m around home at each age • Modeling variables: proximity to highway, roadway density, truck route density, four-way intersection density, number of city bus stops, stationary source proximity, percentage of commercial building area	Outcome • Maternal report via questionnaire of physician-diagnosed asthma in the past 12 months • Questionnaires completed every 3 months between birth and age 2 and every 6 months at ages 2-3, 4-5. Analysis • Generalized estimating equations • Covariates: sex, ethnicity, smoker in home, income, residential cockroach/mouse allergen, age, age by GIS variable interaction, wheeze in past 12 months, positive indoor allergen-specific IgE, total IgE from birth to age 5	Main Findings for Asthma **OR (95% CI) prenatal exposure, outcome by age 5** Highway proximity: 2.07 (1.28, 3.34) Roadway density: 1.43 (1.01, 2.02) Truck route density: 1.11 (0.74, 1.65) **OR (95% CI) exposure year 1, outcome at age 5** Highway proximity: 1.14 (0.76, 1.70) Roadway density: 1.04 (0.73, 1.48) Truck route density: 0.92 (0.63, 1.35) *[Also presented results for age 1, 2, and 3, and other exposure metrics]*
Zhou 2013 (Zhou et al. 2013) – France		
Study Details • EDEN mother-child cohort study (n=1,765 mother-child pairs) Exposure • In utero and first year of life traffic-related air pollution exposure assessed via questionnaire administered at time of outcome assessment. • Questions: "[Is] your house located near a bus stop or a passageway of trucks?" "Do…cars often or continuously pass by your house?" "Do you live less than 200 meters away from a road with heavy traffic?"	Outcome • Maternal report via questionnaire at ages 4, 8, and 12 months. Analysis for asthma in the first year of life. • Doctor diagnosed asthma; Doctor diagnosed asthma with wheezing and/or history of bronchiolitis Analysis • Logistic regression, GEE, adjusted population attributable risk (aPAR) • Covariates: study centre, maternal factors (occupation, age, pre-pregnancy BMI), birth weight, cesarean delivery, preterm birth, breast feeding, siblings, gender, family history of asthma, eczema, allergic rhinitis, food allergy	Main Findings for Asthma **In utero exposure** Trend (OR>1) between exposure and asthma at age 1. *[Data not reported]* **OR (95% CI) first year of life exposure** 1.71 (1.08, 2.72) **aPAR(%) (95% CI)** 13.52 (2.38, 20.53)

Table 2.2. Summary of key articles on prenatal and early life traffic-related air pollution and childhood asthma incidence *(Continued)*

Zmirou 2004 (Zmirou et al. 2004) – France		
Study Details • Vesta multi-center pair-matched case control study (n=195 matched pairs). Conducted 1998-2000 • Incident asthma cases (ages 4-14) matched at time of diagnosis with control without asthma or chronic respiratory symptoms on. Matched on: city, age, gender Exposure • Cumulative index of lifelong traffic exposure. • Traffic Density = (I/D). Where I= traffic density on index road (road within 300 m of location resulting in largest I/D ratio), D=distance to road. • Weighted value calculated using all residence, day care, and school locations for each child.	Exposure *(continued)* • Cumulative index calculated using I/D values weighted according to age, gender, and city specific averages of time spent at home and school Outcome • Physician-diagnosed incident asthma Analysis • Conditional logistic regression. Lifelong exposures, exposures from ages 0-3 years • Covariates: social class, environmental tobacco smoke during pregnancy, exposure to maternal smoking at home, day care, gas cooking, months with pets and humidity at home, personal allergy	Main Findings for Asthma **OR (95% CI) lifelong exposure** Tertile 1: reference Tertile 2: 0.95 (0.50, 1.82) Tertile 3: 0.82 (0.43, 1.59) **OR (95% CI) exposure before age 3** Tertile 1: reference Tertile 2: 1.48 (0.73, 3.02) Tertile 3: 2.28 (1.14, 4.56) **Exposure before age of 3: modeling log transformed exposure as a quantitative predictor** 1.30 (1.04, 1.62)

CHAPTER 3

Methods

DATA SOURCES

Air Pollution Data

The air pollution data for this work will come from the EPA-funded Southeastern Center for Air Pollution and Epidemiology (Southeastern Center for Air Pollution and Epidemiology (SCAPE) 2010) which is one of four national Clean Air Research Centers in the United States. Housed by Emory University and Georgia Institute of Technology, SCAPE is composed of renowned air pollution and environmental health experts. A major goal of SCAPE is to assess impacts of air pollution mixtures using a variety of approaches to mixture characterization, including approaches that group pollutants by source (e.g. source apportionment), and hypothesized biological mechanism of effect (i.e. oxidative potential). Two separate pollutant datasets created by researchers in SCAPE's Air Quality Core and Biostatistics Core will be used in this dissertation, one estimates $PM_{2.5}$ from traffic (RLINE), and one estimates total $PM_{2.5}$ (CMAQ-RLINE Fusion). The creation of these datasets are described in the methods sections of Chapter 5 (RLINE) and Chapter 6 (CMAQ-RLINE Fusion) of this dissertation. The CMAQ-RLINE Fusion dataset uses inputs from the Community Multiscale Air Quality Model (CMAQ). A description of CMAQ is included here.

Community Multiscale Air Quality Model (CMAQ)

CMAQ is a multiple pollutant model developed by the U.S. Environmental Protection Agency that can be used for air quality research, regulation, and forecasting purposes. A very brief overview of the model is described here. For a more in depth explanation please refer to the EPA CMAQ documentation and a 2006 paper reviewing CMAQ updates (the

primary references for this summary) (Byun DW et al. 1999, Byun et al. 2006). CMAQ is an Eulerian emissions-based chemical transport model which uses a "one atmosphere" approach which takes into account multiple pollutants at multiple spatial scales. Eulerian models (like CMAQ) estimate air pollution using a grid system supplying each grid with pollution estimates for time points of interest. CMAQ contains separate modeling systems for meteorology, emissions, and chemical transport. It is in a continuous state of development with system changes occurring regularly. A concise explanation of CMAQ comes from Chapter 1 of the original EPA documentation (Byun DW et al. 1999):

> *CMAQ is a multi-pollutant, multiscale air quality model that contains state-of-science techniques for simulating all atmospheric and land processes that affect the transport, transformation, and deposition of atmospheric pollutants and/or their precursors on both regional and urban scales. It is designed as a science-based modeling tool for handling all the major pollutant issues (including photochemical oxidants, particulate matter, acidic, and nutrient deposition) holistically.*

Several inputs are used by the CMAQ pollution simulation, the two main classes of which are meteorology and emissions. CMAQ is designed to use information from multiple meteorological models and can currently use information from the Fifth Generation Penn State/National Center for Atmospheric Research Mesoscale Model (MM5), the Regional Atmospheric Modeling System (RAMS) and the Weather Research and Forecast Model (WRF). MM5 has been used since the first release of CMAQ. Before it is run, initial and boundary conditions are specified. This process includes the determination of terrain height, land use specifications, meteorology background fields and lateral boundary conditions. The meteorology model itself uses observations on four state variables (wind,

pressure, humidity and temperature) along with equations on thermodynamics, moisture, and momentum. The model is flexible with different options available to researchers depending on what assumptions they want to make, simulation scales, and CPU and memory available for processing. Among other options, there are multiple schemes for radiation cooling, convective parameterization, and planetary boundary layers. After the meteorology data are simulated from MM5 or the other models, it enters the Meteorology-Chemistry Interface Processor (MCIP) where it is converted into a form that can be used by the other CMAQ modules.

In the initial release of CMAQ, the emissions inputs were supplied by the Models-3 Emission Processing and Projection System based on the Emission Modeling System-95 (MEPPS EMS-95). This system is highly user-driven with several opportunities for researchers to determine process specifics. MEPPS uses emission inventory data from point, area, biogenic, and mobile sources and meteorological information. The emissions data come from pollution control agencies, and the meteorological data come from MCIP. These data are imported into MEPPS and first allocated spatially and temporally into pollution grids and then speciated. The data are processed by the Emissions-Chemistry Interface Processor (ECIP) to create hourly emissions data that can be used by the chemical transport model. CMAQ can now also use information from the SMOKE (Sparse Matrix Operator Kernel Emission) modeling system. When run, SMOKE replaces both MEPPS and ECIP.

In addition to simulated meteorology and emissions information, CMAQ uses input data on land use, concentration fields (used as a base for the chemical transport model), temperature, aerosol number density, and vertical ozone profiles. These inputs involve several computations to be useful to the system such as calculation of plume rise, photolysis rates, and cloud parameters.

All data inputs are used by the CMAQ Chemistry Transport Model (CTM) to simulate pollutants at multiple scales. The modeling system is designed to provide good pollution estimates under a wide range of situations and to be adaptable to different user demands and experimentation. The "one atmosphere" approach simultaneously accounts for chemical interactions of multiple pollutants at dynamic temporal and spatial scales. The majority of CTM components can be classified into eight modules: advection, photolysis, cloud aqueous process, diffusion, process analysis, gas-phase chemistry, plume in grid, and aerosol. The gas phase chemistry and aerosol chemistry and dynamics modules are both particularly fundamental to the modeling of pollutant interactions. The aerosol chemistry module models both species from primary emissions as well as secondary species and considers $PM_{2.5}$ and PM_{10}. The impact on aerosol chemistry by clouds is taken into account by the cloud module. This component of CTM models cloud physics and chemistry and their impact on the reactions modeled in CTM and ultimately on pollutant concentrations. The plume-in-grid technique is used for the simulation of plume growth (horizontal and vertical), rise, and transport of plumes resulting from major point source pollutants. The governing equations used by CTM use a generalized coordinate system which allows specific processes to use the coordinate systems best fitted to their calculations while also allowing for adaption to coordinates used by meteorological models. This generalized coordinate system is also used by the transport algorithms that model advection and vertical and horizontal diffusion. The many components of CTM produce the final speciated spatial and temporally allocated air pollution estimates.

Kaiser Permanente Georgia

Health outcomes and covariate data for children and mothers will come from Kaiser Permanente Georgia (KPGA) Health Maintenance Organization (HMO) electronic medical

records. KPGA offers comprehensive medical services to 240,000 members in the metropolitan Atlanta area and has a database of electronically linked administrative and clinical records. Detailed member-level data include information on HMO enrollment history, residence, primary care visits, drug prescriptions, emergency department visits, and hospital admissions. A key advantage of this data source is that it includes information on all encounters with the medical system, not just emergency department visits and hospitalizations. For my analyses, data are available for 24,608 children born between 2000 and 2010 who have been insured by Kaiser since birth through at least the first year of life.

Information on asthma diagnosis and asthma-related medication dispensings will be used for asthma classification. The list of medications we will consider as "asthma-related" is provided in Table A2 in Appendix A. This list was created by a team of physicians and researchers for the Medication Exposure in Pregnancy Risk Evaluation Program (MEPREP) on which researchers at KPGA collaborated. It should be comprehensive of all medications used to treat childhood asthma. All medications on this list have also been confirmed as used for asthma by a pediatric allergist at Emory University (Dr. Karen DeMuth).

Georgia Birth Certificates

Supplementary covariate data will come from Georgia birth certificates. Variables requested include infant birth characteristics (e.g., preterm delivery, reduced birth weight), maternal characteristics (e.g., age, smoking status), paternal characteristics, and sociodemographic variables (e.g., race, maternal education). Birth certificate data have already been linked to Kaiser Permanente data for previous projects.

Data Preparation

A number of data cleaning steps will be necessary in order to transform datasets received from KPGA and the state of Georgia into an analytic dataset that we can use for analyses. This section describes how we will complete two of the more complicated data cleaning steps, determining child race and ethnicity, and first year of life residence.

Child race and ethnicity

Not all children enrolled in KPGA have race information available in their medical record. We will use information from the following sources to determine child race: child race from KPGA, maternal race from KPGA, maternal race from child birth certificates, and paternal race from child birth certificates. All children will have their race classified as black, white, other, or unknown. The other category will include Asian, American Indian, Alaska Native Hawaiian or other Pacific Islander, and children identifying with more than one racial group. We will use the following algorithm to assign child race.

1. For children matched to birth certificates: if mother and father birth certificate race are the same it will be assigned as child race. If father race missing on birth certificate, use mother's race on birth certificate for child's race
2. If no birth certificate is available, or birth certificate race is missing for mother, or birth certificate mother and father race is not concordant: use child race information from KPGA
3. If child has no race information in KPGA, and mother and father birth certificate race is not concordant: classify child's race as 'Other'
4. If child has no race information in KGPA, and mother birth certificate race missing, but have father birth certificate race: use father's race from birth certificate for child's race

5. If no birth certificate is available, or birth certificate race is missing for both mother and father, and no child KPGA race information is available: use mother KPGA race information
6. For all remaining children: classify race as unknown

Child ethnicity will be determined using the same data sources as used to determine child race. The KPGA ethnicity variable includes the categories 'yes' and 'no or unknown.' Subsequently, the child ethnicity variable used in our analyses will have the same categories. A child will be classified as Hispanic if they have any evidence of Hispanic ethnicity, i.e. Hispanic is listed in any of the following places: birth certificate mother ethnicity, birth certificate father ethnicity, child KPGA ethnicity, or mother KPGA ethnicity.

First year of life residence

Residential information for the first year of life will be used to estimate first year of life $PM_{2.5}$ exposure. These residences will first be geocoded to a 250 meter by 250 meter grid system of Atlanta for which air pollution data are available. The child residence dataset from KPGA has multiple observations per child, with each observation listing the 250 meter grid of the residence, and start and stop dates for that residence. For many children, residential history information from KPGA is incomplete; some children have no residential information available, while others are missing residential data for certain periods of time. We will assign $PM_{2.5}$ exposure for all children with at least one residence with a start date between their date of birth (DOB) and their first birthday. In order to complete assignment, we need to determine child residence for every day in the first year of life period (DOB to the day before the first birthday). We will use the following algorithm to complete this assignment:

1. Combine any consecutive residences with the same 250 meter grid assignment. For example, if there are 2 consecutive residences with the same grid assignment, the clean dataset will have one residence listed, with the earliest start date, and latest stop date out of these two observations.
2. Remove gaps between residences starting in the first year of life: calculate the median between the end date of the residence before the gap and the start date of the residence after the gap. The end date of the residence before the gap will be moved to 1 day before the median. The start date of the residence after the gap will be moved to the median.
3. For children whose first residence in the first year of life starts after their DOB, the start date of their first residence will be moved up to their DOB.
4. For children whose last residence starting in the first year of life ends before the day before their first birthday:
    a. If the child has no residences between their first and second birthdays, the end date of their last residence before the first birthday will be extended to the day before their first birthday
    b. For children with a residence between their first and second birthdays, the first residence after their first birthday will be considered in the assignment algorithm. The median between the end date of their last residence before their first birthday and the start date of the first residence after their first birthday will be calculated.
        i. If the median falls after the child's first birthday, the end date of their last residence before their first birthday will be moved to the day before their first birthday.

ii. If the median falls before their first birthday, the end date of their last residence before the first birthday will be moved to the day before the median. The first residence after their first birthday will be set to the endpoints: start date = median between actual start date and the end date of the last residence before the first birthday; end date = day before their first birthday.

Prenatal residence will be determined using the same algorithm, instead using maternal residential history during pregnancy (between the estimated day of conception and the day before the date of birth).

Socioeconomic Data

Information on neighborhood-level socioeconomic status will come from Georgia Department of Public Health (DPH). Georgia DPH has created innovative SES characterizations at block group spatial resolution using data on 25 variables from the 2010 Census. These characterizations classify block groups into four major categories and eighteen demographic clusters based on age, income, family structure, education, housing, and employment (Demographic Clusters of Georgia 2012). The characterizations were created using TwoStep Clustering and discriminant analysis in SPSS statistical software (Zhou 2012). SPSS TwoStep Clustering begins by creating pre-clusters using a sequential clustering approach and then groups the pre-clusters to form the final clusters (SPSS 2001). Discriminant analysis is a multivariate procedure which tests differences between groups and then determines which variables are necessary for classifying inter-group differences (UCLA Institute for Digital Research and Education 2013). First, the major clusters were created by using the TwoStep Clustering approach followed by discriminant analysis. The

demographic clusters were then created using the same approach, but within the already determined major clusters. Descriptions of all eighteen clusters are provided by DPH Office of Health Indicators for Planning (OHIP) and are included in Table A1 in Appendix A.

While *a priori* we plan to use Georgia Department of Public Health's demographic clusters to characterize neighborhood socioeconomic status, we will also explore individual variables from the American Community Survey. American Community Survey 2010 data (5 year estimates, averaging data from 2006-2010) on the following variables will be pulled via Social Explorer at census tract level (U.S. Census Bureau 2010): median household income, median year house built, median house value, percent less than high school education, percent unemployment, and percent of families in poverty.

ANALYTIC METHODS

The majority of methods for this dissertation are not included in this section and instead are described in each of the main dissertation chapters (Chapter 4, Chapter 5, and Chapter 6). All dissertation components will be completed using a retrospective birth cohort. This cohort includes all Kaiser Permanente Georgia Health Plan members with a date of birth between January 1, 2000 and December 31, 2010 (inclusive of endpoint dates) who were born in Kaiser and enrolled the following 12 months after their date of birth. Gaps of enrollment of up to 45 days within this 12 month period are permitted. This cohort will be "closed" in that it will only contain members born into the cohort. Members will be lost to follow-up due to death (very rare) and the ending of Kaiser health insurance coverage (very common). The cohort study design will allow for the calculation of risks. All analyses will be performed using SAS 9.3 (SAS Institute, Cary, NC).

This dissertation section describes a couple of analytic methods used in the preparation of data for this dissertation, but that are not described elsewhere.

Spatial Data Linkage

This project has a substantial spatial component. All spatial data will be linked using a 250 meter by 250 meter grid network of Atlanta created by SCAPE researchers. Data from Kaiser Permanente Georgia, Georgia DPH, and SCAPE will be linked in order to complete spatial assignment of air pollution exposure and SES characterization retrospectively using residential history. This linkage will result in a grid assignment for each residence, a SES characterization for each grid, and pollution estimates for each grid. All mapping will be completed using an ellipsoid geodatum (base model of the earth) and the WGS84 coordinate system (World Geodetic System 1984 revision). The linkage will be completed in the following ways:

1) Mother and child residential addresses have already been geocoded to an ellipsoid geodatum by Kaiser Permanente using Yaddress (Yurisoft, Salano Beach, CA). For this project, researchers at Kaiser will use a point-in-polygon approach in ArcGIS (ESRI, Redlands, CA) to map each mother and child residence to a 250 meter grid.
2) Georgia DPH SES characterizations are mapped at U.S. Census block group spatial resolution also using an ellipsoid geodatum. The 250 meter grids were created to nestle within Census tracts (multiple grids per tract) which will facilitate the assignment of a SES characterization for each grid using ArcGIS.
3) CMAQ pollution estimates are created using a spherical geodatum. For the purposes of exposure assignment, CMAQ data coordinates were translated to an ellipsoid model and the WGS84 coordinate system using Python software. This computationally intensive conversion was completed by SCAPE researchers and led by Dr. Heather Holmes at Georgia Tech and Dr. John Pearce at Emory University. After the conversion of CMAQ to an ellipsoid model, CMAQ grids can be linked to our

250 meter grids which will allow us to complete exposure assignment using the CMAQ-RLINE fusion dataset.

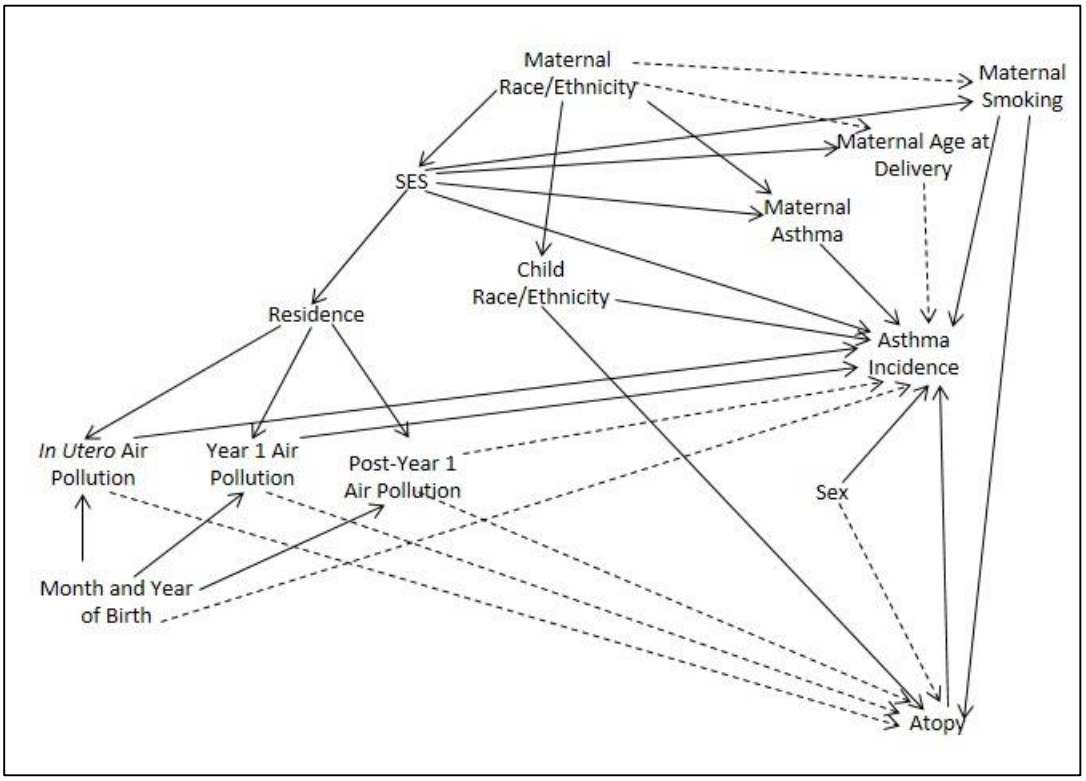
Power Calculations

Power calculations were calculated for Cox proportional hazard models. Based on our initial calculations, we expected to have excellent power to detect modest hazard ratios. In this calculation, we anticipated that 17,000 children would remain enrolled in Kaiser Permanente Georgia through the age of three and 14% of these would be diagnosed with asthma. Based on that scenario, we would have had 80% power to detect a hazard ratio of 1.08 per standard deviation increase in pollutant concentration. (Estimates were prepared using PASS statistical software(PASS Statistical Software. Version 8.0.7.) with α=0.05 and R^2=0.5 for the other model covariates, and assuming a normal distribution of the annually averaged pollutant concentrations in the 250 meter grids).

Directed Acyclic Graph

The directed acyclic graph (DAG) in Figure 3.1 was created when developing the study questions and analytic plan for this dissertation. It summarizes relationships that are relevant to our study questions and helped inform modeling decisions. Dotted lines represent relationships that are not as well established.

Figure 3.1. Directed acyclic graph of prenatal and first year of life air pollution exposure and childhood asthma incidence



CHAPTER 4

Please note, this chapter was submitted for publication to the journal Pediatric Allergy and Immunology *on December 5, 2015. At the time of submission of this dissertation to Laney Graduate School the chapter was still under review by the journal.*

Evaluating early-life asthma definitions as a marker for subsequent asthma in an electronic medical record setting

Audrey L. Flak [1], Matthew J. Strickland [2], Karen A. Freedle [3], Mitchel Klein [2], Carolyn Drews-Botsch [4], Craig Hansen [5, 6], Lyndsey A. Darrow [4]

[1] Department of Epidemiology, Rollins School of Public Health and Laney Graduate School, Emory University, Atlanta, Georgia

[2] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia

[3] Department of Pediatrics, Emory University School of Medicine, Atlanta, Georgia, USA

[4] Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia

[5] Kaiser Permanente Georgia Center for Clinical and Outcomes Research, Atlanta, Georgia

[6] South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia

ABSTRACT

Background: Case definitions for asthma incidence in early life vary between studies using medical records to define disease. This study assessed the impact of different approaches to using medical records on estimates of asthma incidence by age 3 and determined the validity of early-life asthma case definitions in predicting school-age asthma. *Methods:* Asthma diagnoses and medications by age 3 were used to classify 7,103 children enrolled in Kaiser Permanente Georgia according to 15 different definitions of asthma. School-age asthma was defined as an asthma diagnosis between ages 5 and 8. Sensitivity (probability of asthma by 3 given school-age asthma), specificity (probability of no asthma by 3 given no school-age asthma), positive and negative predictive value (probability of (no) school-age asthma given (no) asthma by 3), and likelihood ratios (combining sensitivity and specificity) were used to determine predictive ability. *Results:* 9.0% to 35.2% of children were classified as asthmatic by age 3 depending on asthma case definition. Concordance of asthma classification by age 3 and school-age asthma status ranged from 71.4% to 79.9%. Early-life asthma classifications were more specific than sensitive and were better at identifying children who would not have school-age asthma (negative predictive values: 79.1% to 86.6%) than at predicting children who would have school-age asthma (positive predictive values: 43.5% to 71.5%). *Conclusions:* Choice of case definition had a large impact on the estimate of asthma incidence. While ability to predict school-age asthma was limited, several of the case definitions performed similarly to clinical asthma prediction tools.

INTRODUCTION

Asthma development often begins early in life with an estimated 50-80 percent of children who have asthma experiencing symptoms before age five (National Asthma Education and Prevention Program. Expert Panel Report 3 2007). It is difficult to diagnose asthma in young children due to variable and non-specific symptoms and lack of reliable objective testing. It is also challenging to distinguish children who will experience persistent asthma throughout childhood from those with transient wheeze. Despite these complications, extensive research focuses on asthma in early childhood, requiring investigators to develop case definitions for incident asthma in early life.

There is both clinical and research interest in using early life respiratory symptoms to identify children who will experience persistent asthma in later childhood. Previous studies have created and evaluated the performance of clinical asthma prediction tools to identify children in early life who are at high risk of having persistent asthma or wheeze at school age. The Asthma Prediction Index (API), which consists of both loose and stringent indices, was developed in 2000 using the Tucson Children's Respiratory Cohort and is a popular clinical prediction tool (Castro-Rodriguez et al. 2000). Additional prediction indices developed using birth cohorts include the Isle of Wright score and the PIAMA (Prevention and Incidence of Asthma and Mite Allergy) risk score (Kurukulaaratchy et al. 2003, Caudri et al. 2009). These predictive indices require information that can be prospectively collected by clinicians in the interest of patient care such as results of blood work and skin prick tests. The inclusion of detailed clinical parameters make these indices not well-suited for use in large retrospective studies that rarely have access to such information on all individuals. It is unknown whether retrospective studies that lack detailed clinical information can also predict who will experience persistent asthma at school age. This

ability would be valuable since large studies have the potential to shed light on causes of asthma that may be missed by smaller clinical studies with less statistical power.

There is tremendous variability between case definitions for early-life incident asthma among studies using medical records or administrative claims data to define disease. Case definitions differ in the quantity and types of diagnoses and medications required to classify a child as asthmatic. In 2005, Dombkowski and colleagues used Medicaid data to assess differences between six prevalent asthma case definitions for use in surveillance among children ages 18 years and younger (Dombkowski et al. 2005). Similar to earlier research in a different Medicaid population (Buescher et al. 1999), they found that childhood asthma prevalence was highly dependent on the definition used. It is uncertain whether the same variability in estimated disease proportion would be observed for case definitions intended to classify incident asthma in non-Medicaid pediatric populations, such as among children enrolled in health maintenance organizations.

The goals of the present study are to fill some of this knowledge gap by comparing different cumulative incident asthma case definitions in the first three years of life and assessing their ability to predict asthma at school-age. Specifically, this study addresses the following objectives in a birth cohort of children enrolled in Kaiser Permanente Georgia: 1) Assess the impact of different approaches to using medical records to estimate the cumulative incidence of asthma by age three. 2) Determine the validity of these early-life asthma case definitions, which exclusively use information available in medical records, in predicting school-age asthma. This analysis seeks to identify a case definition for asthma in early life that minimizes disease misclassification when used as a proxy for asthma at school age.

METHODS

The Kaiser Air Pollution and Pediatric Asthma (KAPPA) Study is a birth cohort of children born between 2000 and 2010 enrolled in Kaiser Permanente Georgia (KPGA) Health Maintenance Organization (HMO) for at least the first year of life. KPGA is an integrated health care system that provides medical care services to approximately 240,000 members in the metropolitan Atlanta area. KAPPA was developed to assess the effects of air pollution exposure in infancy on childhood asthma incidence. Among the 18,488 children who were born between 2000 and 2007 and were age six or older at the time of the KAPPA study, this analysis was completed using the subgroup of 7,103 children enrolled in KPGA continuously from birth until at least age six (allowing up to 90 day enrollment gaps). We used information from KPGA electronic medical records and administrative databases to examine 15 different case definitions for early-life incident asthma. Table 4.2 includes the case definitions assessed (definitions of terms are provided in Table 4.1). Several of these case definitions are used either exactly or with slight variations (e.g. modified medication list, specific timing of events, different child ages) in previous studies (Getahun et al. 2010, Dawood et al. 2011, Goyal et al. 2011, Li et al. 2011, Quinto et al. 2011, Black et al. 2013, Gold et al. 2014).

Incident asthma in early life was classified for each child using events from the medical record between birth and age three. We then individually assessed the ability of each of the 15 definitions of early-life incident asthma to predict school-age asthma status, defined as at least 1 asthma diagnosis (ICD-9 code 493.XX) between ages five and eight. Although asthma diagnoses at school age are subject to measurement error, they are more reliable than earlier diagnoses and indicate evidence of continued asthma morbidity (National Asthma Education and Prevention Program. Expert Panel Report 3 2007). Predictive ability was measured using sensitivity (probability of incident asthma by age 3 given school-age asthma), specificity (probability of no incident asthma by age 3 given no

school-age asthma), positive predictive value (probability of school-age asthma given incident asthma by age 3), and negative predictive value (probability of no school-age asthma given no incident asthma by age 3). Likelihood ratio tests, which combine sensitivity and specificity to assess overall prediction accuracy, were also calculated: positive likelihood ratio (sensitivity divided by one minus specificity) and negative likelihood ratio (one minus sensitivity divided by specificity) (Gallagher 1998). All analyses were completed in SAS 9.3 (Cary, NC).

RESULTS

In this cohort of 7,103 children (Table 4.3), 1,705 children (24.0%) had an asthma diagnosis recorded between ages five and eight ("school-age"). Using diagnoses and medication dispensings in the first three years of life, 2,719 children (38.3%) were classified as asthmatic by at least one case definition. Cumulative asthma incidence by age three ranged from 9.0% (definition 4) to 35.2% (definition 7) depending on the case definition used (Table 4.4). The definitions which produced the lowest asthma incidence, which had the lowest sensitivity and highest specificity, were best able to predict who had and did not have school-age asthma. For example, definition four (≥3 asthma diagnoses), resulted in the lowest estimate of early-life asthma incidence (9.0%) and had the highest overall concordance between asthma incidence by age three and school-age asthma (79.9%). This predictive success of the most stringent definitions is attributable to the fact that the majority of children were not classified as asthmatic at school age.

Concordance of asthma classification by age three and school-age asthma status was similar between different case definitions, ranging from 71.4% to 79.9%, but the extent to which the asthma cases or non-cases at school age were misclassified varied. Overall, the tests were more specific than sensitive. In this population, with a prevalence of school-age

asthma of 24%, the early-life asthma classifications were far superior at ruling out school-age asthma (negative predictive values ranged from 79.1% to 86.6%) than they were at predicting school-age asthma (positive predictive values ranged from 43.5% to 71.5%). Across definitions, the positive and negative likelihood ratios would generally be considered as having poor to moderate predictive ability for a clinical test (Gallagher 1998). We saw no evidence that prediction ability was dependent on child characteristics (Table 4.5).

The impact of adding additional information to the case definition was mixed. Consider for example definition 8, at least 1 asthma diagnosis and 1 medication dispensing. Making the definition more complex by additionally classifying a child as asthmatic if they had 1 asthma-related ED visit or hospitalization or 3 asthma diagnoses (definition 15) resulted in little predictive benefit by any examined metric. However, changing definition 8 by specifying that the medication had to be a controller (definition 13), sharply decreased the percent of children classified as asthmatic by age 3 and resulted in a marked increase in specificity and positive predictive value. Similar results were found specifying medication type in other definitions.

DISCUSSION

Using electronic medical records from a large HMO in the southeastern U.S., we systematically examined different ways to classify asthma in early life and evaluated which case definitions were best able to predict children who will have evidence of asthma at school-age. In this population, choice of case definition had a large impact on the estimate of asthma incidence in early life. Dombkowski and colleagues reached a similar conclusion when examining case definitions for prevalent asthma in a cohort of children enrolled in Medicaid (Dombkowski et al. 2005). For example, concordance estimated by kappa

statistics between asthma classifications using events before age five and an asthma diagnosis in the subsequent year ranged from 0.28 to 0.40. We examined different asthma classifications than the Dombkowski study and had more time between initial classification and later disease status, but observed similar kappa values ranging from 0.21 to 0.38 (calculated from values in Table 4.4) (McHugh 2012).

While none of the case definitions we examined consistently identified children who would be diagnosed with asthma at school age, their performance was comparable to that of clinical asthma prediction tools. When using events by age three to predict active asthma at age six, the loose Asthma Predictive Index (API) has a sensitivity of 56.6%, and a specificity of 80.8%, which are very similar to the sensitivity and specificity of our definitions 1, 10, and 12. When validated at the same age, the stringent API has an almost identical sensitivity and specificity as our case definition 4 (stringent API sensitivity 27.5%, specificity 96.3%) (Castro-Rodriguez et al. 2000). Similarities in performance also exist between our definitions and other clinical prediction tools. For example, when using a cut point of a severity score of 6, the Environmental and Childhood Asthma (ECA) severity index has almost identical prediction metrics to our case definition 14 (ECA sensitivity 51.5%, specificity 88.1%, positive predictive value 54.3%, negative predictive value 86.8%) (Devulapalli et al. 2008). The generally poor predictive ability of our case definitions and of clinical prediction tools reflects the complex and often transient nature of early life respiratory symptoms (Brand et al. 2008, Fouzas et al. 2013).

There is a high prevalence of school-age asthma in this cohort, with almost a quarter of children receiving at least one asthma diagnosis between ages five and eight. This prevalence is higher than Georgia state estimates; in 2010 it was estimated that among children ages five to nine years in Georgia 13.7% had current asthma and 20.4% had ever been diagnosed with asthma (U.S. Centers for Disease Control and Prevention et al. 2011).

The higher prevalence in our population can likely be explained partially by the use of medical records for classification which in comparison to parental report yields higher prevalence estimates for childhood asthma (Yoo et al. 2007). Additionally, prevalence of asthma diagnoses has been found to be higher among insured than uninsured children (Coker et al. 2012).

The 15 case definitions for early life asthma that we examined are a subset of the many potential definitions one could choose. We did not examine definitions that use only information on medications, and not diagnoses, to determine whether a child has asthma. These definitions were excluded because medications used to treat asthma are also used to treat other conditions. We also did not examine incident asthma case definitions that considered whether a diagnosis was classified as primary in the medical record, because in our dataset we were unable to determine primary status for 83.9% of asthma diagnoses given to children in our cohort. While we are referring to this outcome as early-life asthma given the use of asthma ICD-9 diagnoses, we are cognizant that respiratory conditions before age six are not typically called asthma and continued wheezing may be a more appropriate term for these outcomes.

This analysis has several strengths and limitations. The KAPPA study is uniquely positioned to examine early-life asthma case definitions due to access to medical records on over 7,000 children insured by KPGA from birth until at least age six. The record-based classification used in this study, instead of the commonly used parental report, prevents recall bias from impacting study results. The use of medication dispensings, rather than medication prescriptions, is a strength of this analysis since dispensings are expected to align more closely with actual medication intake. Limitations of using medical record data are the inability to account for variations in provider practices and lack of information on indication for medications. There is undoubtedly some misclassification of asthma status

among children between ages five and eight. Even though reliability of asthma diagnoses increases as children age, this outcome is not perfect in determining asthma status at school age, particularly since it was determined using ICD-9 codes. Our analyses were restricted to children in the KAPPA cohort who were followed until age six. Results were comparable if we restricted the cohort to children enrolled through age eight. While prevalence of an early life asthma diagnosis was similar between children in our analysis and children lost to follow-up (22.5% vs. 21.2%), it is possible that loss to follow-up impacted our findings. Positive and negative predictive values are directly dependent on asthma prevalence; one would expect these values to differ when examining the performance of these case definitions in a population with a different school-age asthma prevalence. Other prediction metrics may also vary in different populations, particularly outside of an HMO setting.

There is no perfect way to classify asthma status using medical records, particularly in early childhood. Given the challenges of asthma diagnosis in early life, misclassification in asthma research is unavoidable. Our analysis indicated that choice of case definition had a large impact on the estimate of asthma incidence in early life. This dependence has implications for the comparability of findings between studies that use different case definitions for childhood asthma. The results of this analysis emphasize the importance of completing sensitivity analyses to assess the impact of case definition choice on research results and to facilitate better comparisons across studies. Among the early-life asthma case definitions we examined, there was not an obvious choice as to which was best at predicting school-age asthma. Several of our case definitions performed similarly to clinical asthma prediction tools, showing that asthma diagnoses and medications in early life can be used to predict asthma at school age with as much accuracy as can be obtained with some detailed clinical tools. Despite their limitations, clinical asthma prediction tools have proved useful in asthma research, for example in studies to identify lung function biomarkers and develop

effective asthma therapies (Castro-Rodriguez 2011). The comparable predictive ability of our early-life asthma definitions combined with the unique advantages of large record based studies highlight the potential for record based studies to continue to advance our knowledge about asthma etiology.

Table 4.1. Diagnosis and medication definitions

Outcome	Definition
Asthma diagnosis	ICD-9 code 493.XX
Wheeze diagnosis	ICD-9 code 786.07
Acute asthma diagnosis	a) emergency department or inpatient asthma diagnosis *or* b) asthma diagnosis with status asthmaticus or acute exacerbation (ICD-9 codes 493.01, 493.02, 493.11, 493.12, 493.21, 493.22, 493.91, 493.92)
Atopic dermatitis	ICD-9 code 691.8
Allergic rhinitis	ICD-9 code 477.X
Asthma controller[a]	Aminophylline, beclomethasone diproprionate**,** budesonide, budesonide/formoterol fumarate, cromolyn sodium, fluticasone propionate, fluticasone/sameterol, mometasone furoate, montelukast sodium, salmeterol xinafoate, theophylline anhydrous, tiotropium bromide, triamcinolone acetonide
Asthma reliever	Albuterol, albuterol sulfate, ipratropium bromide, ipratropium/albuterol sulfate, levalbuterol, metaproterenol sulfate
Asthma-related medication	Dispensing of any asthma controller or reliever

[a] Underlined medications contain a steroid

Table 4.2. Early-life asthma case definitions

Case Definition	Criteria Needed	Asthma diagnosis (n)	Other diagnoses	Asthma-related medication dispensings required (n)
1	Any	1	1 wheeze diagnosis	
2	All	1		
3	All	2		
4	All	3		
5	All	1	1 atopic dermatitis or 1 allergic rhinitis diagnosis	
6	Any	2	1 acute asthma diagnosis	
7	Any	1		2
8	All	1		1
9	All	1		2
10	Any	1		2 (at least 1 steroid)
11	All	1		2 (at least 1 steroid)
12	Any	1		1 controller
13	All	1		1 controller
14	All	1		2 reliever or 1 controller
15	Any	3	1 asthma-related ED visit or hospitalization	1 if in same year as 1 asthma diagnosis

These are the minimum required events for each case definition using events by age 3. Only 1 diagnosis per day counted. ED=emergency department. Definitions of all terms are included in Table 4.1.

Table 4.3. Cohort characteristics (n=7,103)

Characteristic	N (%)
Sex	
Female	3,474 (48.9)
Male	3,629 (51.1)
Race/Ethnicity	
Black	3,004 (42.3)
White	2,847 (40.1)
Other Race[a]	691 (9.7)
Missing Race	561 (7.9)
Hispanic Ethnicity	359 (5.1)
Maternal Education	
<12th grade	91 (1.3)
High School/GED	737 (10.4)
Some College or more	4,330 (61.0)
Missing Education	1,945 (27.4)
Kaiser Permanente Enrollment Duration [b]	
Enrolled through age 6	7,103 (100.0)
Enrolled through age 8	4,075 (57.4)
Year of Birth	
2000 – 2001	2,273 (32.0)
2002 – 2003	2,130 (30.0)
2004 – 2005	1,482 (20.9)
2006 – 2007	1,218 (17.1)

[a] Includes the following racial groups: Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander, children identifying with more than one racial group
[b] Enrollment through age 6 part of inclusion criteria. Children enrolled through age 8 are a subset of children enrolled through age 6. Reduction in sample size across follow-up reflects shorter follow-up time available for children born in later years of the study (e.g., a child born in 2005 could be at most 8 years old at the time of medical record data abstraction) as well as HMO enrollment attrition over time.

Table 4.4. Early asthma classifications (using events by age 3) and prediction of school-age asthma (at least 1 asthma diagnosis between ages 5 and 8) among children enrolled in Kaiser Permanente Georgia (n=7,103)

Asthma definition applied to age 0-3 years	% Meeting definition by age 3	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Positive Likelihood Ratio	Negative Likelihood Ratio	% Correctly predicted asthma status at school age
1. 1 asthma or wheeze diagnosis	29.4%	57.6%	79.5%	47.0%	85.6%	2.8	0.5	74.2%
2. 1 asthma diagnosis	22.5%	49.9%	86.2%	53.2%	84.5%	3.6	0.6	77.5%
3. 2 asthma diagnoses	13.2%	35.4%	93.7%	64.1%	82.1%	5.6	0.7	79.7%
4. 3 asthma diagnoses	9.0%	26.8%	96.6%	71.5%	80.7%	7.9	0.8	79.9%
5. *Atopic Asthma:* 1 asthma diagnosis AND (1 atopic dermatitis OR allergic rhinitis diagnosis)	7.4%	19.5%	96.5%	63.5%	79.1%	5.5	0.8	78.0%
6. 2 asthma diagnoses OR 1 acute asthma diagnosis	14.2%	36.8%	93.0%	62.3%	82.3%	5.2	0.7	79.5%
7. 1 asthma diagnosis OR 2 medication dispensings	35.2%	63.8%	73.8%	43.5%	86.6%	2.4	0.5	71.4%
8. 1 asthma diagnosis AND 1 medication dispensing	21.7%	49.2%	87.0%	54.4%	84.4%	3.8	0.6	77.9%
9. 1 asthma diagnosis AND 2 medication dispensings	19.8%	46.7%	88.7%	56.7%	84.0%	4.1	0.6	78.6%
10. 1 asthma diagnosis OR 2 medication dispensings 1 of which must be a steroid	24.0%	52.2%	84.8%	52.1%	84.9%	3.4	0.6	77.0%
11. 1 asthma diagnosis AND 2 medication dispensings 1 of which must be a steroid	11.7%	31.8%	94.6%	65.1%	81.5%	5.9	0.7	79.5%
12. 1 asthma diagnosis OR 1 controller dispensing	24.4%	52.8%	84.6%	52.0%	85.0%	3.4	0.6	77.0%
13. 1 asthma diagnosis AND 1 controller dispensing	12.1%	32.8%	94.5%	65.2%	81.7%	5.9	0.7	79.7%
14. 1 asthma diagnosis AND (2 reliever dispensings OR 1 controller dispensing)	19.9%	47.0%	88.7%	56.8%	84.1%	4.2	0.6	78.7%
15. Any of the following: a) 1 asthma diagnosis AND 1 medication dispensing in the same year, b) 1 asthma-related ED visit or hospitalization, c) 3 asthma diagnoses	21.6%	49.3%	87.1%	54.7%	84.5%	3.8	0.6	78.0%

These are the minimum required events for each case definition. Only 1 diagnosis per day counted. ED=emergency department. 1,705 children in cohort (24%) have an asthma diagnosis between ages 5 and 8. *Sensitivity:* probability of incident asthma by age 3 given school-age asthma. *Specificity*: probability of no incident asthma by age 3 given no school-age asthma. *Positive predictive value*: probability of school-age asthma given incident asthma by age 3. *Negative predictive value*: probability of no school-age asthma given no incident asthma by age 3. *Positive likelihood ratio*: sensitivity divided by one minus specificity. *Negative likelihood ratio*: one minus sensitivity divided by specificity. Positive likelihood ratios >10 and negative likelihood ratios <0.1 are considered to be indicative of case definitions with high predictive value. Positive likelihood ratios between 2 and 10, and negative likelihood ratios between 0.5 and 0.1 indicate case definitions that may have some predictive value (Gallagher 1998).

Table 4.5. Positive and negative likelihood ratios for prediction of school-age asthma by individual characteristics

	Positive Likelihood Ratio					Negative Likelihood Ratio				
Case Definition	Full Cohort	White	Black	Male	Female	Full Cohort	White	Black	Male	Female
1	2.8	2.7	2.8	2.5	3.1	0.5	0.6	0.5	0.5	0.6
2	3.6	3.5	3.7	3.2	4.1	0.6	0.6	0.5	0.6	0.6
3	5.6	5.1	6.6	4.8	6.9	0.7	0.7	0.7	0.7	0.7
4	7.9	7.4	9.5	6.8	9.4	0.8	0.8	0.7	0.7	0.8
5	5.5	5.5	5.4	5.1	5.9	0.8	0.9	0.8	0.8	0.8
6	5.2	4.7	6.1	4.4	6.3	0.7	0.7	0.6	0.7	0.7
7	2.4	2.4	2.3	2.2	2.6	0.5	0.5	0.4	0.5	0.5
8	3.8	3.7	3.9	3.4	4.3	0.6	0.6	0.5	0.6	0.6
9	4.1	4.1	4.2	3.7	3.7	0.6	0.6	0.6	0.6	0.6
10	3.4	3.4	3.5	3.0	3.9	0.6	0.6	0.5	0.5	0.6
11	5.9	5.4	7.0	5.4	6.4	0.7	0.7	0.7	0.7	0.7
12	3.4	3.4	3.5	3.0	3.9	0.6	0.6	0.5	0.5	0.6
13	5.9	5.4	6.8	5.5	6.3	0.7	0.7	0.7	0.7	0.7
14	4.2	4.1	4.1	3.7	4.7	0.6	0.6	0.6	0.6	0.6
15	3.8	3.7	3.9	3.4	4.3	0.6	0.6	0.5	0.6	0.6

Case definitions included in Table 4.2. School-age asthma defined as at least 1 asthma diagnosis between ages 5 and 8. *Positive likelihood ratio*: sensitivity divided by one minus specificity. *Negative likelihood ratio*: one minus sensitivity divided by specificity. Positive likelihood ratios >10 and negative likelihood ratios <0.1 are considered to be indicative of case definitions with high predictive value. Positive likelihood ratios between 2 and 10, and negative likelihood ratios between 0.5 and 0.1 indicate case definitions that may have some predictive value (Gallagher 1998).

CHAPTER 5

Prenatal and first year of life exposure to primary $PM_{2.5}$ from traffic and childhood asthma incidence in a birth cohort

INTRODUCTION

Air pollution exposure during pregnancy and early life may play an important role in the development of childhood asthma. Exposures during both of these critical windows can impact immune programming and response, and respiratory system development which begins six weeks after conception and continues through adolescence (Peden 2000). During this period of growth, repair mechanisms are not as adept at responding to environmental insults as those in mature adult lungs (Kajekar 2007). Previous research showing that children who develop asthma by the age of seven have 40% of their associated lung deficit at birth highlights the importance of the prenatal window in particular (Bisgaard et al. 2012). After birth, important structural changes in the lungs occur, such as the growth of additional bronchioles and alveoli, which are critical for meeting the increasing metabolic demands of a growing child (Pinkerton et al. 2000, Moore et al. 2003, Wang et al. 2008). The development that occurs during this period is paired with children's greater exposure to ambient air pollution relative to adults due to increased ventilation rates and more time spent outdoors (American Academy of Pediatrics 1999, Pinkerton et al. 2000). Exposure to particulate matter equal to or less than 2.5 micrometers in diameter ($PM_{2.5}$) may be particularly detrimental to lung development because particles of this size are small enough to end up in lung alveoli. Among children, mouth breathing increases particulate matter exposure because air avoids the blocking mechanisms in the nasal passage and can subsequently end up further into the respiratory system.

It is well established that exposure to ambient air pollution exacerbates childhood asthma, and growing evidence suggests it may also play a role in asthma development (Tzivian 2011, Gowers et al. 2012). The impact of pollution from traffic is specifically of interest. Four review articles summarizing evidence from previous studies have suggested an association between traffic-related air pollution and residential proximity to busy roads with respiratory symptoms and asthma incidence (Salam et al. 2008, Bråbäck et al. 2009, Gasana et al. 2012, Bowatte et al. 2015). Results of individual studies on the association between prenatal and first year of life $PM_{2.5}$ exposure from traffic and asthma incidence have ranged from null to suggestive of a positive association (Clark et al. 2010, Gehring et al. 2010, Carlsten et al. 2011). A study completed by Clark and colleagues in 2010, attempted to determine the relative importance of $PM_{2.5}$ exposure during the prenatal and first year of life periods by including estimates of $PM_{2.5}$ exposure in each window in the same regression model (Clark et al. 2010). However, they were unsuccessful at teasing out these effects due to the high correlation between the exposures in their data.

In this study, we examine the association between prenatal and first year of life exposure to primary $PM_{2.5}$ from traffic emissions and childhood asthma incidence at ages 2 through 6 in a birth cohort of children enrolled in Kaiser Permanente Georgia. We aim to determine the impact of exposure during these two separate periods, and also whether one is of relatively more importance (i.e. if exposure in one of these periods has a stronger association with childhood asthma).

METHODS

Data Sources/Study Population

The Kaiser Air Pollution and Pediatric Asthma (KAPPA) Study is a birth cohort of children born between January 1, 2000 and December 31, 2010 enrolled in Kaiser Permanente Georgia (KPGA) Health Maintenance Organization (HMO) for at least the first year of life (allowing up to 90 day gaps in enrollment). KAPPA members were followed from birth until September 2013 or until their enrollment in Kaiser ended if that occurred first. KPGA is an integrated health care system that provides medical care services to approximately 240,000 members in the metropolitan Atlanta area. KAPPA was developed to assess the effects of air pollution exposure in infancy on childhood asthma incidence. Detailed individual information on the 24,608 children in the KAPPA cohort is available from KPGA on demographic characteristics, residential history, HMO enrollment, diagnoses, and medication dispensings. KAPPA includes 21,791 children (88.6%) who are linked to mothers in the KPGA system. For these children, maternal information is also available. Among children linked to mothers, 85.3% are also linked to Georgia birth certificates from which information is available on pregnancy and parental characteristics. In the KAPPA cohort, maternal linkage allows determination of maternal asthma status and estimation of prenatal air pollution exposure. It also allows us to establish which children have siblings in the KAPPA cohort and account for this lack of independence in analyses. For the 2,817 children not linked to mothers, we assumed they had no siblings in the cohort.

Prenatal and First Year of life $PM_{2.5}$ Exposure Estimates

$PM_{2.5}$ data were modeled by colleagues at Georgia Institute of Technology using a research line-source dispersion model for near-surface releases (RLINE) (Zhai et al. 2015). RLINE was developed in 2013 by the U.S. Environmental Protection Agency's Office of Research and Development specifically for health studies of traffic pollution. RLINE uses vehicle emissions data to model primary $PM_{2.5}$ from traffic emissions and implements a

steady-state Gaussian plume model to create a smooth modeled exposure surface (Community Modeling and Analysis System 2015). Traffic emissions were estimated by Atlanta Regional Commission for 2011 for each section of the roadway in the Atlanta metropolitan area using information about traffic patterns and vehicle emissions. In addition to emissions data, RLINE incorporates meteorological data to predict pollutant dispersion patterns. For example, wind patterns and weather conditions are important in determining whether pollutants are blown away from the roadway, and in what direction, or whether they stay near the roadway after being emitted. AERMET, the pre-processor for meteorological data before they are used in AERMOD (a widely used pollution dispersion model), was used to generate the meteorological data for RLINE (U.S. Environmental Protection Agency 2015). Wind speed, wind direction, and atmospheric stability were the main meteorological factors that impacted RLINE estimates.

Due to its sharp spatial gradient RLINE often over-estimates pollution values around roadways. In order to compensate this tendency, the modeled pollutant concentrations from RLINE were scaled using estimates of traffic $PM_{2.5}$ source impacts created by a chemical mass balance (CMB) approach and based on monitoring values. The estimated RLINE $PM_{2.5}$ values were calibrated using the following regression equation:

$$\Delta PM_{2.5} = 10^{[0.32 \log(RLINE_{PM_{2.5}}) - 0.05]}$$

where the regression coefficients were chosen based on the difference between the estimated $PM_{2.5}$ values from RLINE and the estimated $PM_{2.5}$ source impacts from CMB. The final calibrated RLINE estimates for 2011 are shown in Figure 5.1. Please note, the air quality data for this work are still being developed and may change before the publication

of results. One such potential change is the use of emissions data at different spatial and temporal resolution.

We used these 2011 $PM_{2.5}$ data at a 250 meter grid resolution to estimate pollution exposure in all years of the study (2000-2011 for first year of life exposure, 1999-2010 for prenatal exposure), based on the assumption that the spatial pattern of roadway impacts did not change substantially over our study period. The prenatal period was defined as the period between the estimated start of gestation and the day before a child's birth. For the 18,583 children linked to birth certificates the start of the gestation period was determined by first counting back the number of weeks gestation, using gestational age from the birth certificate, from the date of birth. Then, to account for the obstetric convention of starting the gestational week count on the day of the last menstrual period, the start date of the prenatal period was moved forward 2 weeks. For the remaining 6,025 children the start of the prenatal period was defined as 38 weeks before the date of birth (assuming a full term gestational age of 40 weeks, with conception occurring at day 14, per obstetric convention). The first year of life was defined as the period between the child's birth date and the day before their first birthday. Maternal residential location during the prenatal period and child residential location during the first year of life was used to estimate pollution exposure during each of these windows. If a mother moved during pregnancy or a child moved during the first year of life their pollution exposure was calculated as a weighted average based on the amount of time spent at each residence.

The spatial domain of available RLINE $PM_{2.5}$ data does not cover the entire region in which KAPPA mothers and children live (see Figure B1 in Appendix B). Children residing outside the RLINE pollution region at any time during the first year of life were excluded. For analyses of prenatal exposure, children whose mothers resided outside the RLINE pollution region during pregnancy were excluded.

Asthma Classification

In the KAPPA study, we define asthma as at least one asthma diagnosis (ICD-9 493.XX) and one asthma-related medication dispensing (including both steroid and non-steroid asthma controllers and relievers) after the first year of life. The following medications were considered asthma-related: *aminophylline, albuterol, albuterol sulfate, beclomethasone diproprionate, budesonide, budesonide/formoterol fumarate, cromolyn sodium, fluticasone propionate, fluticasone/sameterol, ipratropium bromide, ipratropium/albuterol sulfate, levalbuterol, metaproterenol sulfate, mometasone furoate, montelukast sodium, salmeterol xinafoate, theophylline anhydrous, tiotropium bromide,* and *triamcinolone acetonide*. Classifications of individual asthma medications are provided in Table A2 in Appendix A. We are assessing cumulative asthma incidence, so once a child is classified as having asthma, they are classified as asthmatic at every subsequent age. The time of asthma incidence is defined as the time at which a child has satisfied both criteria: received both an asthma diagnosis and asthma-related medication dispensing.

Diagnoses and medications during the first year of life were ignored due to the overlap between these events and the first year of life exposure window, and also due to concerns about the reliability of asthma diagnoses this early in life. Among the 1,453 children with an asthma diagnosis during the first year of life (5.9% of the cohort), 65.7% had at least one asthma diagnosis after the first year of life, and 61.0% had both an asthma diagnosis and an asthma-related medication dispensing after the first year of life. Among the 6,467 children with an asthma-related medication dispensing during the first year of life (26.3% of the cohort), 61.2% had at least one asthma-related medication dispensing after the first year of life, and 39.8% had both an asthma diagnosis and an asthma-related medication dispensing after the first year of life. Children with asthma diagnoses or asthma-

related medication dispensings in the first year of life without both an asthma diagnosis and asthma-related medication dispensing after the first year of life were considered non-asthmatic in all analyses. Among children included in our analyses of first year of life $PM_{2.5}$ exposure, follow-up was similar between children with diagnoses or medications in the first year of life and children without diagnoses or medications in the first year of life (32.1% vs. 27.4% followed until at least age 6).

Socioeconomic Status Characterization

Potential confounding by socioeconomic status (SES) is a major concern in studies of the respiratory effects of residential air pollution due to the spatial comparison in pollution values (as opposed to a temporal comparison). SES is an important factor in determining residence which often results in residential air pollution varying markedly by socioeconomic level. Additionally, asthma rates vary across socioeconomic groups with the highest rates occurring among children in poverty. Since SES frequently determines both residential pollution exposure and childhood asthma incidence, it can confound the association between pollution and asthma. We assessed potential confounding by SES by exploring the impact of four sets of relevant covariates: individual socioeconomic characteristics, novel demographic characterizations at block group spatial resolution, census tract characteristics from the American Community Survey, and distance from the Atlanta city center.

Among the variables available from Kaiser medical records and child birth certificates, the following individual characteristics may be markers of socioeconomic status and were assessed: child race and ethnicity, maternal education, paternal education, and maternal marital status.

The impact of neighborhood-level SES was assessed using novel demographic clusters created by Georgia Department of Public Health from data on 25 variables from the 2010 U.S. Census (Georgia Department of Public Health et al. 2013-2015). These innovative SES characterizations combine information on age, income, family structure, housing, education and employment to classify census block groups into eighteen minor demographic clusters (Demographic Clusters of Georgia 2012, Zhou 2012) (Figure 5.2; full descriptions of the clusters are provided in Table A1 in Appendix A). These minor demographic clusters can be grouped into four major categories ranging broadly from high to low SES: A – highest SES, B – second highest SES, suburban and urban, C – rural, average to lower than average SES, D – lowest SES. These clusters provide more nuanced descriptions of neighborhoods than can be achieved by any one variable from the U.S. Census. Demographic cluster assignments were completed using child residential location at birth.

While *a priori* we planned to use Georgia Department of Public Health's demographic clusters to control for neighborhood SES, we also explored using individual variables from the American Community Survey to assess the sensitivity of results to alternative characterizations of neighborhood SES. American Community Survey 2010 data (5-year estimates, averaging data from 2006-2010) on the following variables were pulled via Social Explorer at census tract level (U.S. Census Bureau 2010): median household income, median year house built, median house value, percent less than high school education, percent unemployment, and percent of families in poverty.

After exploring maps of the demographic clusters and conducting descriptive analyses of the cohort, we added distance from the Atlanta city center to the list of variables aimed at controlling for SES. While the novel demographic clusters characterize block group level SES, we noted that in the same demographic cluster could include inner city areas as

well as areas far outside the urban core (Figure 5.2), and that there may be important differences between individuals living in the same demographic cluster in different areas of Atlanta. Factors that may differ between individuals inside and outside the city core include frequency of healthcare use and environmental exposures other than air pollution (e.g., cockroach allergen, agricultural activity) and may be risk factors for asthma diagnoses not already controlled for. For that reason, we also explored controlling for area of the city by dividing the metropolitan Atlanta area into four regions: inside the city core (referred to as "metro Atlanta" and defined as inside the I-285 highway perimeter of the metropolitan area), within five miles of the city core, five to ten miles from the city core, and more than ten miles from the city core (see Figure B2 in Appendix B). City region assignment was completed using 250 meter grid of child residence at birth.

Analytic Approach

We assessed the impact of first year of life and prenatal air pollution exposure on cumulative incidence of asthma in subsequently longer time periods: 1-2 year risk of asthma, 1-3 year risk, 1-4 year risk, 1-5 year risk, and 1-6 year risk. Each analysis only included children enrolled in KPGA for the entire risk period allowing 90 day gaps in enrollment (e.g., a child whose KPGA enrollment ends after age 2 was only included in the 1-2 year analysis). The outcome in each analysis was defined as events by the age of interest among children enrolled until that age. For example, analyses in the age 5 cohort examine outcomes by the fifth birthday among children enrolled until at least their fifth birthday. This approach allowed us to calculate potentially differential effects by age of follow-up. The risk difference was used as the primary measure of association. This approach was completed using a binomial generalized linear regression model with an identity link function:

$$E\big(P(Y=1|X)\big) = \beta_0 + \beta_1(first\ year\ of\ life\ PM_{2.5}) + \Sigma_{i=1}^{i=p}\delta_i(covariate_i)$$

where the dichotomous outcome Y represents asthma incidence in the time period of interest. These models were implemented using generalized estimating equations with an exchangeable correlation structure in order to account for correlation between siblings in the cohort. Since variance is dependent on probability in the binomial distribution it is expected that models of this type will not meet the homoscedasticity assumption necessary for linearly modeling probabilities. Robust variance estimation was used to account for this potential heteroscedasticity and potential misspecification of the working correlation matrix. We assessed potential for confounding by maternal asthma (defined as at least one maternal asthma diagnosis (ICD-9 493.XX) during a mother's enrollment in KPGA), maternal age, birth year, and by the socioeconomic factors described previously. Potential effect measure modification was also of interest, particularly by child race, and was assessed on the additive scale by the addition of interaction terms into relevant models and also by running stratified models.

We first completed model building, assessing all potential confounding variables and interaction of interest, examining only $PM_{2.5}$ exposure during the first year of life. We then assessed the impact of prenatal exposure by using the final adjusted first year of life exposure model and substituting the prenatal exposure estimate for the first year of life exposure estimate. Lastly, we ran a single model containing both prenatal and first year of life exposure estimates in order to determine the relative importance of exposure during each of the developmental periods (contingent on enough variation between the two estimates).

Due to the known convergence issues of binomial generalized linear regression models, several modeling decisions were made with the goal of minimizing problems with convergence by maximizing sample size and limiting the number of model parameters. All minor demographic clusters with less than 150 individuals were combined with the closest demographic cluster within its major demographic category. In order to include individuals with missing covariate data we used an indicator variable for missing covariates (race, ethnicity, maternal asthma, and maternal education) in analyses rather than excluding children with missing data. Since this practice can result in bias, we completed a sensitivity analysis using only complete cases (Greenland et al. 1995, Glymour et al. 2008). We did not impute missing data because we had limited covariates that could be used to guide the imputation. Maternal age was included in analyses using a linear spline with cut-points at the tertiles of the age distribution, in order to maximize the sample size in each category and allow for potential non-linear effects. Additionally, all covariates that did not impact the association between air pollution and asthma were dropped from analyses.

Additional Analyses and Sensitivity Analyses

Since asthma in early life is often transient, we completed two analyses aimed at assessing whether $PM_{2.5}$ is associated with more lasting asthma phenotypes. In the first analysis, we classified children as having "persistent asthma," if they had incident asthma (at least 1 asthma diagnosis and 1 asthma-related medication dispensing) and also had evidence of asthma in the past year (at least 1 asthma diagnosis or 1 asthma-related medication dispensing) at each follow-up age. We took two different approaches to handling children with prior incident asthma, but no persistent asthma at a given follow-up age: (1) we included these children in the reference (i.e., non-diseased) group and (2) we excluded these children from the analysis. In the second analysis to assess the outcome of

continued asthma symptoms, we assessed cumulative incidence of asthma by age 2 restricted to the cohort of children enrolled until their fifth birthday. In this analysis, the outcome of interest was children who were classified as having incident asthma by their second birthday, and also had at least one asthma diagnosis or asthma-related medication dispensing between their second and fifth birthdays. The comparison non-disease group included children without incident asthma by age 5 and children with incident asthma at age 2 without any asthma medications or diagnoses between ages 2 and 5.

There is immense variability between case definitions for early-life incident asthma among studies using medical records or administrative claims data to define disease. In Chapter 4 of this dissertation we examined 15 ways to define asthma in early-life and concluded that choice of asthma case definition impacted estimates of asthma incidence and that different case definitions had varying success in predicting asthma at school age. We are interested in whether the asthma definition selected also impacts our estimate of the association between $PM_{2.5}$ and asthma. To assess this potential, we estimated the association between $PM_{2.5}$ and asthma incidence by age 5 using 14 of the case definitions from Chapter 4 to see if associations differ when changing case definitions. We did not assess the association with case definition 5, atopic asthma, since more detailed clinical data than we have access to are necessary to reliably separate atopic and non-atopic asthma phenotypes.

Sensitivity analyses were completed to determine whether excluding certain children from the cohort impacted model results. We completed analyses excluding children missing race information or missing at least 90 days of residence data during exposure windows. We also completed analyses excluding children who were not linked to mothers or for whom their maternal match was deemed less reliable. Linkage of children to mothers was completed by KPGA using birth certificates and Kaiser medical records. Maternal

matches were considered unreliable if they were completed using incomplete medical record information and not confirmed by a birth certificate, or if birth certificates and medical records included discrepant maternal information. In order to assess whether associations were different among children linked to birth certificates and not linked to birth certificates, we completed an analysis restricted to children for whom birth certificates were available. To assess the potential for selection bias, we completed analyses assessing the association between $PM_{2.5}$ and asthma at earlier follow-up ages restricted to children followed until at least age 6.

The risk difference is our primary measure of association of interest. Given data on a full cohort (i.e. denominator data), we wanted to take advantage of the opportunity to estimate risk. Some advocate additive models as more appropriate to assess biologic interaction, leading us to choose the risk difference over the risk ratio (Greenland S et al. 2008, Vanderweele et al. 2014). Additionally, no previous studies on air pollution and pediatric asthma incidence have assessed risk differences. However, for secondary analyses we were interested in multiplicative effects so we completed sensitivity analyses using hazard ratios, risk ratios, and odds ratios which are more commonly used measures of association. These analyses provided some results that were more comparable to those from previous studies. The risk ratios and odds ratios were computed using log binomial linear regression and logistic regression, respectively, both implemented using generalized estimating equations and an exchangeable correlation matrix to account for correlation between siblings in the cohort. The hazard ratios were calculated with Cox proportional hazards regression, after first checking the proportional hazards assumption, using the robust sandwich estimator to adjust standard errors for the lack of independence between siblings.

RESULTS

Descriptive Results

Among the 24,608 children in the KAPPA cohort, 23,865 (97.0%) have information on residential location during the first year of life. Among these children, 23,100 resided in the region for which we have RLINE $PM_{2.5}$ estimates for the entire first year of life. A subgroup of 19,951 of these children also have information on maternal residential location during the prenatal period with all prenatal residences located within the pollution data region. The number of children followed to each age decreased with each year of follow-up (Table 5.1). This was partially due to the fact that children born in later years of the study were not yet old enough for the older follow-up ages. For example, a child born in 2010 could be at most 3 years old at the time the KAPPA cohort was defined in 2013. Among children who could be potentially followed until each age, losses due to HMO enrollment attrition increased over time ranging from 22.3% of children lost to follow-up by age 2 to 61.6% of children lost to follow-up by age 6 (see Table B1 in Appendix B). There is no evidence that $PM_{2.5}$ exposure differed between children who remained insured by KPGA and children whose HMO enrollment ended, but there is some evidence that children with asthma were more likely to continue KPGA enrollment than children without asthma (see Table B2 in Appendix B).

This is a racially diverse cohort with 34.6% of children classified as of African American race, and 12.0% of children identifying with other, non-white racial groups (Table 5.1). As expected from an HMO cohort, this is a cohort with relatively high socioeconomic status. More than half of children are born to mothers who attended at least some college and 62.3% of children are born in census block groups of the highest (out of four levels) socioeconomic status (Georgia Department of Public Health major demographic cluster A) (Table 5.1, Table 5.2).

In the KAPPA cohort, proximity to the city center is the strongest determinant of ambient $PM_{2.5}$ levels at the residence. Since RLINE estimates primary $PM_{2.5}$ from traffic emissions, the highest pollution estimates in this region are along highways and inside the city core (Figure 5.1). Consequently, there is a strong pollution gradient among children in the KAPPA cohort with the highest estimates among children who live inside the city center ("metro Atlanta"), the second highest estimates among children with live within 10 miles of the city core, and the lowest estimates among children living more than 10 miles from the city core (Table 5.2). The spatial pattern of demographic clusters in metro Atlanta determine the pattern of $PM_{2.5}$ exposure among children in each cluster (Figure 5.2, Table 5.2). Because demographic clusters belonging to major category B (classified as average to high SES) are predominately located inside the city core and track along the major highways of Atlanta, there is the highest pollution exposure among children in this group. The lowest pollution estimates are seen among children in demographic clusters belonging to major category C which are rural clusters; 90.4% of children at residences in major category C live outside metropolitan Atlanta.

Relationships between $PM_{2.5}$ and individual and census tract characteristics are different when averaging over all city regions compared to stratifying by city region. In the KAPPA cohort, children of white race have the lowest $PM_{2.5}$ estimates because the majority of these children live more than 10 miles from the city center (Table 5.1, Figure B3 in Appendix B, Table 5.3). However, when stratifying by city region, white children have higher $PM_{2.5}$ estimates than black children in every city region. The correlation between census tract characteristics (i.e., median household income, and percent of families living in poverty) and first year of life $PM_{2.5}$ exposure also varies in different parts of the city (Table 5.3). Among all children, as census tract level income increases and poverty decreases, first year of life $PM_{2.5}$ exposure decreases. However, among children born in metro Atlanta and

children born more than 10 miles from the city center, the opposite is true with a positive correlation between income and $PM_{2.5}$ and a negative correlation between poverty and $PM_{2.5}$.

This cohort has a high burden of asthma diagnoses and asthma-related medication dispensings, with 32.4% of children receiving both a diagnosis and medication by the age of 6 (Table 5.4). Consistent with other U.S. populations, the highest asthma rates are among male children, children of black race, and children whose mothers have asthma (Table 5.5). By age 6, 45.4% of children whose mothers have asthma are classified as having asthma compared to 29.7% of children whose mothers do not have asthma. There were no consistent trends in asthma by birth year or maternal education. At every follow-up age, there is the least asthma observed in children whose residence at birth is classified as demographic cluster B (the cluster with the highest $PM_{2.5}$ estimates) and the most asthma among children whose residence at birth is classified as demographic cluster C (the cluster with the lowest $PM_{2.5}$ estimates). Geographically, asthma incidence at every follow-up age increases when moving away from the city center, with the lowest asthma incidence among children living inside the city center and the highest asthma incidence among children living more than 10 miles from the city center (Table 5.5).

Association between first year of life $PM_{2.5}$ and asthma incidence

Table 5.6 and Figure 5.4 contain results from statistical models assessing the association between first year of life traffic $PM_{2.5}$ and asthma incidence, each model adjusting for a different set of covariates. All risk differences were calculated for a change of 1 $\mu g/m^3$ of primary $PM_{2.5}$ from traffic which is equivalent to moving from the 3^{rd} to the 97^{th} percentile of the exposure distribution in the cohort (Figure 5.3). When unadjusted for covariates, an increase in $PM_{2.5}$ is associated with a decrease in childhood asthma incidence

at all follow-up ages (Model 0). For example, an increase of 1 μg/m^3 of primary $PM_{2.5}$ from traffic is associated with a 3.7% decrease in absolute risk of asthma by age 4 (RD (95% CI): -0.037 (-0.067, -0.007)). Results changed little when adjusting for individual-level covariates (Model 1). Maternal age, maternal education, and child ethnicity were not significant predictors of childhood asthma, and did not appear to be confounding the association between $PM_{2.5}$ and asthma so were dropped from the model (Model 2). The additional individual-level variables of paternal education and maternal marital status were also not found to be confounders of the association of interest (Model A1 includes these variables, shown in Table B3 in Appendix B).

Moving to models with both individual-level and demographic control (Table 5.6, Figure 5.4, Models 3-5), adding any form of demographic control to the model moved risk differences between $PM_{2.5}$ and asthma in the positive direction. Out of the demographic variables assessed (major demographic cluster, minor demographic cluster, and American Community Survey variables), minor demographic cluster had the largest impact on the association between traffic $PM_{2.5}$ and asthma incidence and was retained in final models. Adding city region to models already adjusted for individual characteristics and minor demographic cluster had a big impact on the estimated association between $PM_{2.5}$ and childhood asthma (Table 5.6, Figure 5.4, Models 6-8). Results were fairly similar when dividing the Atlanta region into 4 areas (Model 6) compared to diving the metropolitan Atlanta area into 3 regions (Model 7), so dividing the city into 3 regions was used in the final model in order to limit model parameters. In the final adjusted model (Model 7) an increase of 1 μg/m^3 of $PM_{2.5}$ is associated with a 2.7% to a 5.8% increase in risk of childhood asthma depending on follow-up age with confidence intervals at ages 2, 3, and 5 excluding the null. These final adjusted models control for child sex, child race, maternal asthma, birth year, minor demographic cluster, and city region.

In our final adjusted models, the strongest predictor of childhood asthma was maternal asthma. When controlling for all other covariates, at age 6, the absolute risk of asthma is 15.2% higher among children whose mothers have asthma compared to children whose mothers do not have asthma or for whom we were unable to determine maternal asthma status (Table 5.7, RD (95% CI): 0.152 (0.117, 0.187)). Other strong risk factors for asthma were male sex and black race. At age 6, the risk of asthma was 8.6% higher in male children than female children and 6.0% higher in black children than white children, when controlling for all other covariates (RD (95% CI): 0.086 (0.064, 0.108); 0.060 (0.030, 0.089)). Correlation between siblings increased with follow-up age from 0.16 in the age 2 analysis to 0.24 in the age 6 analysis.

A priori, we were interested in whether there was effect measure modification using the risk difference of the association between $PM_{2.5}$ and asthma incidence by race and maternal asthma. Stratifying crude model results showed different associations between $PM_{2.5}$ and asthma for white and black children, but adjusted associations were similar and there was no evidence of statistical interaction (see Tables B4 and B5 in Appendix B). For child sex, the associations between $PM_{2.5}$ and asthma were stronger among males than among females in adjusted models at ages 2 through 5. However, there was no evidence of statistical interaction by sex (see Tables B6 and B7 in Appendix B). For maternal asthma, both unadjusted and adjusted associations were stronger among children whose mothers do not have asthma than among children whose mothers have asthma, but there was no evidence of statistical interaction (see Tables B8 and B9 in Appendix B). Given the different descriptive patterns by city region, we also assessed whether the association between $PM_{2.5}$ and asthma differed by city region. Stratified results showed different associations between $PM_{2.5}$ and asthma in both unadjusted and adjusted models. However, there was no evidence of statistical interaction and the small sample sizes in the individual regions makes it

probable that differences may be due to random error (see Tables B10 and B11 in Appendix B).

Association between prenatal $PM_{2.5}$ and asthma incidence

The distributions of prenatal and first year of life $PM_{2.5}$ exposure in the KAPPA cohort are visually identical (Figure 5.3). There is a moderate amount of residential mobility in the KAPPA cohort, with 18.2% of children moving during the first year of life and 18.6% of mothers moving during the prenatal period. Among children with both first year of life residence data and prenatal residence data, 36.0% changed residence at least once during the period from conception to the first birthday. Despite this mobility, prenatal and first year of life pollution estimates are highly correlated (Spearman correlation coefficient = 0.93) and 64.0% of children have identical $PM_{2.5}$ estimates for the two periods. The high correlation between these estimates is partially due to the lack of temporal variability in the RLINE $PM_{2.5}$ data used to estimate exposure.

Both unadjusted and adjusted risk differences for the association between prenatal $PM_{2.5}$ and childhood asthma are very similar to the estimates of the association between first year of life $PM_{2.5}$ and childhood asthma (Table 5.8, Figure 5.5). In the adjusted models, with the exception of the age 2 analysis, effect estimates are slightly closer to the null value in the prenatal models than in the first year of life models. For example, in the final adjusted model an increase of 1 $\mu g/m^3$ of primary $PM_{2.5}$ from traffic emissions in the prenatal period is associated with a 2.5% increase in asthma risk by age 4, and the same change in $PM_{2.5}$ in the first year of life is associated with a 3.7% increase in asthma risk by age 4, with confidence intervals for both estimates including the null (RD (95% CI): prenatal 0.025 (-0.020, 0.070); first year of life 0.037 (-0.007, 0.082)).

Because of the high correlation between prenatal and first year of life pollution estimates in this cohort, these data are not well-suited to tease out the separate effects of $PM_{2.5}$ in each of these periods or to decide which exposure window is more important. Including both exposure estimates in the same statistical model results in instability which precludes meaningful interpretation of effects.

Additional analyses and sensitivity analyses

In general, analyses using more detailed case definitions that take into account continued asthma morbidity resulted in stronger associations between first year of life $PM_{2.5}$ and childhood asthma. When examining persistent asthma, defined as a child with incident asthma who also has evidence of asthma morbidity in the past year, results were fairly similar to the incident asthma results. Most results using the persistent asthma definition showed a stronger association between $PM_{2.5}$ and asthma particularly when excluding children with incident but without persistent asthma from analyses (Table 5.9). For example, at age 4 an increase of 1 $\mu g/m^3$ of $PM_{2.5}$ is associated with a 3.7% increase in absolute risk of incident asthma (RD (95% CI): 0.037 (-0.007, 0.082)). When using the definition of persistent asthma the increased risk is larger at 4.9%, and increases to 5.1% when excluding children with incident but not persistent asthma from the reference group (RD (95% CI): 0.049 (0.010, 0.088), 0.051 (0.010, 0.092)).

Stronger associations were also seen in analyses using events later in childhood to re-define asthma outcomes at age 2 completed among the cohort of 8,592 children enrolled until age 5. The risk difference between first year of life $PM_{2.5}$ and asthma was larger when defining an asthma case as a child with incident asthma at age 2 and at least one asthma diagnosis or medication between ages 2 and 5 than when defining the outcome as just incident asthma at age 2 (RD (95% CI): 0.029 (-0.011, 0.069) vs. 0.004 (-0.0319, 0.0404)).

These models were controlled for covariates in the final adjusted models, but controlled for neighborhood SES using major demographic cluster instead of minor demographic cluster due to model convergence difficulties. The minor demographic clusters break the region into finer categories subsequently providing tighter control for neighborhood SES than the major demographic clusters. Consequently, there may be more residual confounding by SES in these models than in other statistical models. In previous models, the use of major demographic cluster resulted in smaller, but fairly similar, risk differences to those from models with minor demographic cluster (Table 5.6 Model 3 vs. Model 4).

We completed analyses at age 5 using alternative case definitions of asthma (Table 5.10). Conclusions about the association between traffic $PM_{2.5}$ and childhood asthma at age 5 were consistent across case definitions with all risk differences above zero and all confidence intervals excluding the null. Exact estimates of the association varied from the lowest estimate using case definition 3 (2 asthma diagnoses RD(95% CI) 0.050 (0.002, 0.097)) to the highest estimate using case definition 10 (1 asthma diagnosis or 2 medication dispensings 1 of which must be a steroid RD(95% CI) 0.079 (0.023, 0.135). The estimate of the association using case definition 8 (the case definition used in all other analyses) fell in the middle of the estimates from other case definitions (RD (95% CI) 0.058 (0.004, 0.112)).

Results of sensitivity analyses excluding children with missing race information, missing 90 or more days of residence data during the first year of life, with no maternal match or an unreliable maternal match, or for whom birth certificates were unavailable were very similar to results from models including the entire cohort (see Table B12 in Appendix B). The frequency of each of these events was relatively rare and each analysis excluded less than 15% of the cohort (with the exception of the birth certificate analyses which excluded closer to 25% of the cohort). The analyses excluding children with missing race information and without maternal matches used a complete case approach to missing

data in these variables. The similarity of these results to the results of main analyses suggest that the use of an indicator for missing data did not bias results.

We completed some analyses examining asthma at earlier follow-up ages among children followed until at least age 6. Among the 6,628 children enrolled until age 6, the risk difference for the association between first year of life $PM_{2.5}$ and asthma by age 2 was 0.006 (95% CI -0.039, 0.051), asthma by age 3 was 0.017 (-0.035, 0.069), asthma by age 4 was 0.033 (-0.025, 0.091), and asthma by age 5 was 0.048 (-0.014, 0.110). These estimates show smaller increases in risk associated with $PM_{2.5}$ and have less precision due the smaller sample size than our main model results, but all estimates are in the same direction as those from the main models.

Alternative measures of association

Unadjusted models and adjusted models with all covariates from the final adjusted risk difference model (child sex, race, maternal asthma, birth year, minor demographic clusters and city region) were completed using log binomial linear regression, logistic regression, and Cox proportional hazards regression. The log binomial linear regression and logistic regression models were completed using the same cohorts of children as the risk difference models. The only difference in the modeling was the use of different link functions (i.e. log and logit instead of identity). Cox proportional hazards models were completed using the 22,987 children in the KAPPA cohort with information on first year of life $PM_{2.5}$ exposure data and who were enrolled in Kaiser until at least their first birthday. Cox models were completed for each follow-up age taking into account the timing of asthma incidence defining failure time by age in days. For example, at age 3 the outcome of interest was asthma incidence between the first and third birthdays. Children who had not been classified as asthmatic by their third birthday were censored on their third birthday (or at

the time their enrollment in Kaiser ended if that occurred before their third birthday). This same approach was used at each follow-up age. Figure B4 in Appendix B shows a survival curve for the age 6 analysis which examined outcomes between the first and sixth birthdays.

The proportional hazards assumption was assessed using the age 6 analysis for all variables in the final adjusted first year of life model: first year of life $PM_{2.5}$, sex, race, maternal asthma, birth year, minor demographic clusters, and city region. The assumption was first tested graphically by plotting survival curves (see Figure B5 in Appendix B). Plots for race, sex, and minor demographic cluster suggested that these variables may not meet the proportional hazards assumption, so time dependent variables were included for each of these variables in separate Cox proportional hazards models adjusting for other model covariates. Likelihood ratio tests of these extended Cox models indicated that it may be important to include interaction terms between sex and time, and race and time, in the model, but that the other time-dependent variables were unnecessary (see Table B13 in Appendix B). The final Cox proportional hazards model was run including the following sets of time dependent variables: (1) none, (2) time dependent variables with race, (3) time dependent variables with sex, and (4) time dependent variables with both sex and race (see Table B13 in Appendix B). The results for the association between first year of life $PM_{2.5}$ and childhood asthma incidence were almost identical from the four models, so the reduced model without any time dependent variables was chosen as the final model.

Table 5.11 includes the unadjusted and adjusted risk difference, risk ratio, odds ratio, and hazard ratio results for the association between first year of life $PM_{2.5}$ from traffic and childhood asthma incidence. At all follow-up ages, crude estimates using all measures of association indicate an inverse association between $PM_{2.5}$ and asthma and adjusted estimates using all measure of association indicate a positive association between $PM_{2.5}$ and

asthma. For example, in adjusted models, at age 4 for an increase of 1 μg/m^3 of $PM_{2.5}$, the risk difference model indicates a 3.7% increase in absolute asthma risk, the risk ratio model indicates a 1.18 times higher asthma risk, and the odds ratio model indicates a 1.25 times higher odds of asthma (RD(95% CI) 0.037 (-0.007, 0.082); RR(95% CI) 1.18 (0.98, 1.42); OR(95% CI) 1.25 (0.97, 1.62)) with p-values for the estimates almost identical between the different models. As anticipated given the lack of a rare outcome, at every follow-up age the risk ratios are closer to the null value of 1 than the odds ratios (Greenland S et al. 2008). At age 4 the Cox proportional hazards model results indicate higher asthma hazard associated with an increase of 1 μg/m^3 of $PM_{2.5}$ from traffic (HR (95% CI): 1.31 (1.10, 1.57)), but has a much smaller p-value than the other models. This is likely due to the increased power resulting from including all children enrolled until the first birthday in the model, rather than only including children enrolled until their fourth birthday as in the age 4 other analyses.

DISCUSSION

In this study we assessed the association between primary $PM_{2.5}$ from traffic emissions at the residential location in early life and childhood asthma incidence in a cohort of children enrolled in Kaiser Permanente Georgia. Results from this study provide some evidence of an association between $PM_{2.5}$ exposure from traffic and asthma incidence. However, results were dependent on which variables are included in our models with results of some models providing little evidence of an association between traffic $PM_{2.5}$ and asthma. Similar to the difficulties in previous studies, the high correlation between estimates of prenatal and first year of life exposure prevented us from determining the relative importance of exposure during each of these periods. Future work in this cohort

that incorporates temporal variation in pollution levels at each residence will be better suited to tease out the independent effects of these exposure windows.

Major strengths of the KAPPA study include the availability of comprehensive medical record data for outcome classification and the availability of estimates of pollution exposure for both prenatal and early life periods calculated using high quality residential history data and fine scale $PM_{2.5}$ estimates. Access to data on a full cohort (i.e. denominator data) allows for estimation of risk. The use of information on both individual and neighborhood characteristics allowed us to explore possible confounding by a variety of socioeconomic factors operationalized in different ways. We investigated study hypotheses in the KPGA population, a primarily urban population in the southeastern U.S. with high rates of asthma and access to healthcare. It is possible results may not generalize well to markedly different populations. Nonetheless, relationships observed in our population have public health importance given the large number of people in the U.S. represented by our study population.

A limitation of the KAPPA study is incomplete information on race and familial asthma history, both important predictors of childhood asthma. We conducted sensitivity analyses restricted to subjects with race information and information on maternal asthma status, which yielded results consistent with the main analyses (see Table B12 in Appendix B). The study could also have benefited from supplementary information not available in Kaiser medical records or from birth certificates. For example, information on individual socioeconomic attributes such as income would have allowed for more control of confounding by individual-level SES. However, the use of maternal education provided some control. Also, given that this is an insured population with high levels of education, we anticipate that individuals are more exchangeable on individual-level factors than in other populations, making us less concerned about confounding by individual-level SES.

Information on early life environment, such as exposure to secondhand smoke exposure, would have also been useful since these exposures may be related to both asthma and SES raising concerns about potential confounding. Additionally, maternal matches were determined using family units within the KPGA system, so they do not necessarily represent biologic relationships (i.e. they could represent adoptive parents or step-parents). Prenatal pollution exposure estimates and determination of maternal asthma status are only accurate for children whose maternal match is their biological mother. However, maternal asthma was the strongest risk factor for childhood asthma in our data and it seems unlikely we would see this association if a large proportion of mothers were not biologic links.

The KAPPA study has high loss to follow-up rates particularly at later ages which impacted study power and exacerbated model convergence issues. We completed a number of analyses to address whether the association between $PM_{2.5}$ and asthma was different among children lost to follow-up and retained in the cohort. Loss to follow up was not associated with $PM_{2.5}$ exposure and the crude relationship between $PM_{2.5}$ and asthma was similar between children lost to follow-up and children retained in the cohort (see Table B2 in Appendix B). In addition to this crude analysis, we completed an analysis examining the association between traffic $PM_{2.5}$ and childhood asthma at earlier follow-up ages among the children followed until at least age 6 (presented in results section). Effect estimates were smaller and less precise in this analysis, but showed positive associations between $PM_{2.5}$ and asthma as in our main models. These results lessened our concerns about the impact selection bias may have had on our study results.

The use of fine scale $PM_{2.5}$ data from RLINE for this project allowed us to create highly resolved estimates of pollution exposure from traffic at the residential location. Compared to other dispersion models, RLINE is relatively new, and still in development. Advantages of RLINE include the use of new formulations for plume spread of near surface

releases and the use of a wind meander algorithm (Snyder et al. 2013). One limitation of RLINE is that it was developed for flat roadways and does not adjust for hills or take into account barriers that can impact small scale meteorology (e.g., roadside vegetation, buildings). The biggest shortcoming of our exposure assignment is the application of 2011 RLINE data to all years of our study. These data fail to account for temporal trends in pollution and potential changes in the spatial distribution of the road network. While we would expect that a high pollution area in 2011 would also be a high pollution area in other years of our study, for example a residence near a major highway, this may not be true for all parts of the metropolitan Atlanta area due to possible changes in traffic patterns. Moreover, the use of one year of pollution data limited the variability between prenatal and first year of life pollution estimates because the exposure contrasts were strictly spatial, and many mothers lived at the same location during pregnancy that their children lived at during the first year of life. This lack of variability prevented us from determining the relative impact of exposure in each of these developmental periods. While the RLINE model used information about $PM_{2.5}$ dispersion and there are biological reasons why $PM_{2.5}$ specifically may cause asthma, we cannot rule out the possibility that $PM_{2.5}$ is acting as a surrogate of another unmeasured pollutant from traffic that impact asthma incidence. Since we are not controlling for the effect of other pollutants, and there is a high correlation between pollutants produced by vehicles, our effect estimates may be picking up the impact of these other pollutants on asthma incidence. Future work in the KAPPA cohort will examine effects of NO_X and CO and attempt to separate the impacts of individual pollutants.

For exposure estimation, our use of residential data with information on changes in residence is an improvement over previous studies that have used residence at birth as a proxy for residence during the entire prenatal and first year of life periods. This residential mobility can be substantial, as 36% of our cohort changed residence at least once during the

period from conception to the first birthday. Although our research question concerns ambient particulate matter exposure, if it is causally associated with our outcome it would be through the pathway of personal exposure. Personal exposure to pollution is determined by ambient pollution levels as well as factors such as housing air exchange rate and time-activity patterns (i.e. time spent outdoors and in different locations). For the prenatal period, there is some evidence from the literature indicating high correlations between estimates of pollutant exposures based on maternal residence alone and those incorporating information on maternal time-activity patterns (Iniguez et al. 2009). Two separate studies have found that, on average, pregnant women spend over 65% of their day at home helping to limit misclassification due to time-activity patterns in comparison to other populations (Iniguez et al. 2009, Nethery et al. 2009). We would also expect infants to spend more time at home during the first year of life than they would later in adolescence. Furthermore, the question, "Do traffic emissions near the maternal residence during pregnancy or near the infant residence during the first year of life cause asthma?" is a relevant causal question and has clear implications for behavioral modification; a small change of residential location has the potential to dramatically decrease exposure (e.g., moving away from a busy road).

The percent of children in this cohort with asthma is much higher than in other populations with 32.4% of children being classified as asthmatic by age 6. For example, in 2010 in the state of Georgia it was estimated that among children ages five to nine years 13.7% had current asthma and 20.4% had ever been diagnosed with asthma (U.S. Centers for Disease Control and Prevention et al. 2011). These high asthma rates are likely due to a number of factors. It has been found previously that, in comparison to use of parental report, use of medical records for classification results in a higher prevalence estimate for childhood asthma and that prevalence of an asthma diagnosis is higher among insured than

uninsured children (Yoo et al. 2007, Coker et al. 2012). A high proportion of our cohort is of African American race which is a known risk factor for asthma (Moorman JE et al. 2012). Additionally, we focused on cumulative incidence and our outcome group likely included children with transient wheeze diagnosed as asthma who did not go on to have asthma symptoms later in childhood.

Measurement error in the outcome of interest stemming from inherent difficulty in early-life asthma diagnosis is a limitation of our study. We conducted extensive analyses to explore the sensitivity of the results to decisions about outcome classification. The choice to define asthma incidence as at least one asthma diagnosis and one asthma-related medication dispensing was made after a thorough investigation of different ways to define asthma in our cohort with counsel from a pediatric allergist (see dissertation chapter 4). This definition insures that a child classified as asthmatic has had a doctor diagnose their respiratory condition as asthma, and also has some evidence of respiratory symptoms requiring treatment for which they filled a medication prescription. In supplementary analyses we saw that conclusions about the association between traffic $PM_{2.5}$ and asthma incidence by age 5 were consistent if we used any of the other case definitions of asthma from dissertation chapter 4 (Table 5.10, excluding the case definition for atopic asthma). The use of medical records for disease determination, instead of the frequently used method of maternal report, prevents some types of bias such as recall bias from impacting disease classification (Miller et al. 2001). However, asthma is particularly difficult to diagnose in early life due to transient and non-specific symptoms and the lack of reproducible objective testing. Asthma-related medications we used for disease classification are also used to treat other respiratory conditions and can be used for diagnostic rather than treatment purposes (e.g., a child's response to steroids or bronchodilators can help determine whether or not their condition is asthma). Our outcome

is cumulative incidence of asthma, so the outcome group at all ages includes children with asthma diagnoses made early in life which are the least accurate. However, at later follow-up ages our outcome group is expected to contain a greater proportion of definitive asthma cases and have less misclassification. In analyses using an outcome of persistent asthma we saw stronger associations at follow-up ages 4 and 5 consistent with better disease classification (Table 5.9).

The KAPPA study was not designed to differentiate between the effects of $PM_{2.5}$ on atopic and non-atopic asthma phenotypes and subsequently our data are not well-suited for this endeavor. In order to reliably separate these phenotypes, information on skin prick testing and blood IgE levels is necessary. In the absence of these data, we did not attempt to determine whether the association between $PM_{2.5}$ and childhood asthma is different for atopic and non-atopic phenotypes.

The risk difference, an absolute measure of effect, was chosen as the main measure of association for this study after careful consideration. Risk differences lend themselves well to the investigation of additive interaction which is thought to be more relevant to public health than multiplicative interaction (Greenland S et al. 2008, Vanderweele et al. 2014). The risk difference has grown in popularity recently, but is often not used due to the need for denominator data in order to estimate risk and the known model convergence issues of this measure. The best solution for model convergence in our analyses was to limit model parameters either by dropping unnecessary variables, or combining categories. Birth year was included in our final adjusted models in order to have some time control due to the temporal patterns of both asthma and pollution. However, given the lack of a trend in either $PM_{2.5}$ exposure or asthma incidence in our cohort, birth year did not appear to be a strong confounder (Table 5.1, Table 5.5). Since dropping birth year had little impact on model results, it was frequently dropped to aid in convergence for models that would not

converge otherwise (see Table B3 in Appendix B). The use of a robust Poisson model, which is thought to help risk difference models converge, was tried as an alternative, but did not aid in model convergence and in fact made convergence problems worse in some models (Spiegelman et al. 2005). Sensitivity analyses showed our results were not sensitive to modeling and data cleaning decisions, such as including children with gaps in residence data and using an indicator group for missing values in analyses (see Table B12 in Appendix B).

Confounding by SES-related factors acted in the opposite direction from what we expected *a priori* in this study population. We anticipated our unadjusted models would be biased in the positive direction since children from the lowest SES groups often have the most asthma and also live in the most polluted areas. Surprisingly, our unadjusted results showed a negative association between $PM_{2.5}$ and asthma. Descriptive analyses revealed a strong spatial pattern in our data; increasing distance from the city center was associated with increasing asthma rates and decreasing $PM_{2.5}$. Looking at demographic clusters, children living in the cluster with the highest $PM_{2.5}$ had the least asthma (cluster B), and children living in the cluster with the lowest $PM_{2.5}$ had the most asthma (cluster C). It seems likely to us that our crude results are confounded by SES and also by factors other than air pollution that change as distance from the city center increases (e.g., health care utilization). When we control for SES and city region our risk differences move in a positive direction. Out of all of the variables we assessed, city region had the largest impact on our results and may be an important factor to consider in other studies of traffic $PM_{2.5}$ and asthma in large metropolitan areas with a similar pattern of pollution and urban sprawl.

The results of this study suggest that in the KAPPA cohort, the risk of asthma increases an absolute 2.7% to 5.8% (depending on follow-up age) with every increase of 1 $\mu g/m^3$ of primary $PM_{2.5}$ from traffic emissions. However, we note that our effect estimates

are greatly impacted by which covariates are included in our models. For comparison, here are the results of three previous studies assessing the association between $PM_{2.5}$ from traffic and childhood asthma incidence, all using the odds ratio as a measure of association [OR (95% CI)]: Carlsten et al. 3.1 (1.3, 7.4) calculated for an IQR increase (IQR = 4.1 μg/m^3), Clark et al. 1.01 (0.99, 1.03) for 1 μg/m^3, Gehring et al. 1.28 (1.10, 1.49) for an IQR increase (IQR = 3.2 μg/m^3)(Clark et al. 2010, Gehring et al. 2010, Carlsten et al. 2011). These results range from null to suggestive of a positive association. For a more direct comparison, using the odds ratio as a measure of association to examine asthma incidence by age 6, in the KAPPA study we observed an odds ratio (95% CI) of 1.06 (0.94, 1.18) calculated for an IQR increase (IQR = 0.37 μg/m^3). While these odds ratios are calculated for different changes in $PM_{2.5}$ and are adjusted for different factors, many of them are suggestive of a positive association between $PM_{2.5}$ and childhood asthma incidence. Our study, one of the largest to date to examine these associations, and one of the few to use medical records for disease classification, adds to the growing body of results providing some evidence that exposure to traffic in early life may have lasting impacts on respiratory health. In future work with the KAPPA study using time resolved $PM_{2.5}$ data and examining the impact of other pollutants, we hope to better elucidate the association between residential air pollution and asthma in this cohort of children and to tease out the differential impacts of exposure during pregnancy and the first year of life.

Table 5.1. KAPPA cohort characteristics

	Children with first year of life $PM_{2.5}$ data		Children with prenatal $PM_{2.5}$ data [a]	
Characteristic	n (%)	Mean first year of life traffic $PM_{2.5}$	n (%)	Mean prenatal traffic $PM_{2.5}$
Cohort	23,100	1.17	19,951	1.18
Children with siblings in cohort	7,313 (31.7)	1.15	7,123 (35.7)	1.17
Sex				
Female	11,330 (49.1)	1.17	9,748 (48.9)	1.18
Male	11,770 (51.0)	1.17	10,203 (51.1)	1.18
Race/Ethnicity				
Black	7,995 (34.6)	1.17	7,220 (36.2)	1.19
White	9,034 (39.1)	1.12	8,467 (42.4)	1.13
Other [b]	2,771 (12.0)	1.27	2,643 (13.3)	1.29
Unknown Race	3,300 (14.3)	1.20	1,621 (8.1)	1.24
Hispanic Ethnicity	1,839 (8.0)	1.19	1,759 (8.8)	1.21
Maternal Education				
<12th grade	285 (1.2)	1.20	280 (1.4)	1.23
High School/GED	2,605 (11.3)	1.10	2,524 (12.7)	1.12
Some College or more	13,442 (58.2)	1.16	13,113 (65.7)	1.18
Missing	6,768 (29.3)	1.20	4,034 (20.2)	1.23
Maternal Asthma				
Yes	2,488 (10.8)	1.16	2,419 (12.1)	1.18
No	17,998 (77.9)	1.17	17,532 (87.9)	1.18
Missing	2,614 (11.3)	1.20	0	--
Kaiser Permanente Enrollment Duration [c]				
Enrolled until age 2	17,960 (77.8)	1.17	15,631 (78.4)	1.18
Enrolled until age 3	14,251 (61.7)	1.17	12,434 (62.3)	1.18
Enrolled until age 4	10,999 (47.6)	1.17	9,620 (48.2)	1.19
Enrolled until age 5	8,592 (37.2)	1.17	7,521 (37.7)	1.19
Enrolled until age 6	6,629 (28.7)	1.17	5,806 (29.1)	1.19
Birth Year				
2000	2,456 (10.6)	1.20	2,054 (10.3)	1.22
2001	2,369 (10.3)	1.18	1,977 (9.9)	1.20
2002	2,266 (9.8)	1.17	1,946 (9.8)	1.19
2003	2,185 (9.5)	1.16	1,929 (9.7)	1.18
2004	2,138 (9.3)	1.15	1,871 (9.4)	1.17
2005	2,023 (8.8)	1.15	1,741 (8.7)	1.17
2006	2,198 (9.5)	1.15	1,935 (9.7)	1.16
2007	2,216 (9.6)	1.16	1,919 (9.6)	1.17
2008	2,101 (9.1)	1.16	1,835 (9.2)	1.17
2009	1,585 (6.9)	1.18	1,403 (7.0)	1.19
2010	1,563 (6.8)	1.19	1,341 (6.7)	1.19

[a] A subset of children with first year of life $PM_{2.5}$ data [b] Includes Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander, and children identifying with more than one racial group [c] Reduction in sample size across follow-up reflects the shorter follow-up time available for children born in later years of the study (e.g., a child born in 2010 could be at most 3 years old at the time KAPPA follow-up ended in September 2013) as well as HMO enrollment attrition over time.

Table 5.2. Distribution of prenatal and first year of life traffic $PM_{2.5}$ by demographic clusters and city region

	Children with first year of life $PM_{2.5}$ data		Children with prenatal $PM_{2.5}$ data	
Characteristic	n (%)	Mean first year of life traffic $PM_{2.5}$	n (%)	Mean prenatal traffic $PM_{2.5}$
Major Demographic Cluster				
A	14,401 (62.3)	1.08	12,626 (63.3)	1.10
B	2,293 (9.9)	1.53	1,925 (9.7)	1.55
C	1,130 (4.9)	0.98	974 (4.9)	1.00
D	5,272 (22.8)	1.28	4,423 (22.2)	1.30
Minor Demographic Cluster				
A.1	3,065 (13.3)	1.22	2,747 (13.8)	1.23
A.2	2,243 (9.7)	1.19	1,976 (9.9)	1.20
A.3	9,093 (39.4)	1.01	7,903 (39.6)	1.03
B.1	1,080 (4.7)	1.58	896 (4.5)	1.59
B.3/B.4	1,213 (5.3)	1.49	1,029 (5.2)	1.51
C.1/C.2	856 (3.7)	0.89	735 (3.7)	0.92
C.3/C.4	274 (1.2)	1.26	239 (1.2)	1.27
D.1	2,453 (10.6)	1.16	2,083 (10.4)	1.18
D.3	631 (2.7)	1.37	520 (2.6)	1.38
D.4	1,436 (6.2)	1.37	1,204 (6.0)	1.39
D.5	450 (2.0)	1.41	378 (1.9)	1.43
D.6/D.7	302 (1.3)	1.40	238 (1.2)	1.40
City Region				
Metro Atlanta [a]	2,425 (10.5)	1.51	2,030 (10.2)	1.52
≤10 miles from metro Atlanta	9,894 (42.8)	1.27	8,449 (42.4)	1.29
>10 miles from metro Atlanta	10,781 (46.7)	1.00	9,472 (47.5)	1.02

[a] Metro Atlanta defined as inside the I-285 perimeter of Atlanta (Figure B2). Minor demographic clusters with less than 150 individuals were combined with the closest demographic cluster within its major category for analyses (reflected in table).

Table 5.3 Traffic $PM_{2.5}$ exposure and census tract household income and poverty by city region

	All Regions	Metro Atlanta	≤10 miles from metro Atlanta	>10 miles from metro Atlanta
n (row %)				
All Races	23,100	2,425 (10.5)	9,894 (42.8)	10,781 (46.7)
Black	7,995	884 (11.1)	4,901 (61.3)	2,210 (27.6)
White	9,034	851 (9.4)	2,204 (24.4)	5,979 (66.2)
Other Race [a]	2,771	275 (9.9)	1,221 (44.1)	1,275 (46.0)
Unknown Race	3,300	415 (12.6)	1,568 (47.5)	1,317 (39.9)
Mean first year of life traffic $PM_{2.5}$				
All Children	1.17	1.51	1.27	1.00
Black	1.17	1.46	1.21	0.96
White	1.12	1.55	1.33	0.98
Other Race [a]	1.27	1.57	1.37	1.11
Unknown	1.20	1.53	1.29	1.00
Spearman Correlation with first year of life $PM_{2.5}$				
Median household income [b]	ρ = -0.183	ρ = 0.259	ρ = -0.125	ρ = 0.268
Percent families poverty [b]	ρ = 0.175	ρ = -0.257	ρ = 0.106	ρ = -0.100

[a] Includes Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander, and children identifying with more than one racial group. [b] American Community Survey 2010 data (5 year estimates) at census tract level, prepared by Social Explorer

Table 5.4. Asthma incidence and prenatal and first year of life traffic $PM_{2.5}$ exposure by each follow-up age

	Children with first year of life $PM_{2.5}$ data; First year of life traffic $PM_{2.5}$			Children with prenatal $PM_{2.5}$ data; Prenatal traffic $PM_{2.5}$		
Cohort	Asthma n (%)	Mean $PM_{2.5}$ among cases	Mean $PM_{2.5}$ among non-cases	Asthma n (%)	Mean $PM_{2.5}$ among cases	Mean $PM_{2.5}$ among non-cases
Age 2	1,994 (11.1)	1.15	1.17	1,731 (11.1)	1.17	1.19
Age 3	2,627 (18.4)	1.15	1.17	2,316 (18.6)	1.17	1.19
Age 4	2,650 (24.1)	1.16	1.17	2,309 (24.0)	1.17	1.19
Age 5	2,465 (28.7)	1.16	1.17	2,132 (28.4)	1.18	1.19
Age 6	2,149 (32.4)	1.17	1.17	1,854 (31.9)	1.18	1.19

Asthma incidence (at least one asthma diagnosis (ICD-9 493.XX) and one asthma-related medication dispensing after the first year of life) calculated among children enrolled until each follow-up age (see Table 5.1 for number enrolled until each age). For example, the age 5 cohort examines asthma incidence by the 5th birthday among children enrolled until at least their 5th birthday.

Table 5.5. Percent of children with incident asthma at each follow-up age by covariates

	Incident asthma by follow-up age (%)				
Characteristic	Age 2	Age 3	Age 4	Age 5	Age 6
Cohort	11.1	18.4	24.1	28.7	32.4
Sex					
Female	8.8	14.9	20.0	24.2	27.9
Male	13.3	21.8	28.0	33.1	36.8
Race/Ethnicity					
Black	12.5	20.2	26.1	31.2	34.8
White	10.7	17.9	23.1	26.6	29.8
Other [a]	8.0	15.4	20.6	26.9	31.8
Unknown	11.4	17.3	23.9	28.1	33.4
Hispanic Ethnicity	12.1	18.3	21.9	28.1	31.7
Maternal Education					
<12th grade	11.9	18.1	27.1	31.0	31.3
High School/GED	10.4	17.9	22.2	27.1	31.1
Some College or more	11.2	18.8	24.0	28.4	31.8
Missing	11.2	17.9	25.0	29.9	34.2
Maternal Asthma					
Yes	16.8	26.1	33.1	40.4	45.4
No	10.3	17.4	22.5	26.4	29.7
Missing	10.8	17.0	24.7	31.1	36.3
Birth Year					
2000	12.7	19.7	25.8	32.3	36.7
2001	10.4	18.8	23.8	27.3	30.0
2002	11.7	19.6	26.2	29.0	33.0
2003	10.2	18.5	24.1	28.4	32.5
2004	11.6	17.8	22.0	25.6	28.9
2005	11.8	19.6	24.8	29.2	31.9
2006	11.1	18.0	25.5	29.1	32.3
2007	10.5	17.0	22.3	27.9	33.5
2008	10.7	17.3	23.0	29.9	—
2009	9.9	17.9	21.7	—	—
2010	11.0	17.0	—	—	—
Major Demographic Cluster					
A	11.4	18.9	24.2	29.0	32.6
B	8.6	13.9	19.5	23.8	27.6
C	12.0	19.8	28.0	29.6	35.0
D	11.1	18.7	24.7	29.6	33.2
City Region					
Metro Atlanta [b]	7.8	14.1	20.2	25.2	28.2
≤10 miles from metro Atlanta	11.1	18.5	24.1	28.7	32.8
>10 miles from metro Atlanta	11.9	19.3	25.0	29.5	33.1

[a] Includes Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander, and children identifying with more than one racial group. [b] Metro Atlanta defined as inside the I-285 perimeter of Atlanta (Figure B2). "—" = children born in this year not eligible for follow-up age. For example, a child born in 2010 could be at most 3 years old at the time KAPPA follow-up ended in September 2013.

Table 5.6. Risk differences for first year of life traffic $PM_{2.5}$ and childhood asthma incidence in models adjusting for different sets of covariates

	Cohort	MODEL 0		MODEL 1		MODEL 2	
Unadjusted models & models with individual-level covariates		RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
	Age 2	-0.0221 (-0.0390, -0.0053)	0.01	-0.0222 (-0.0399, -0.0044)	0.01	-0.0194 (-0.0359, -0.0029)	0.02
	Age 3	-0.0316 (-0.0552, -0.0080)	<0.01	-0.0290 (-0.0544, -0.0036)	0.03	-0.0277 (-0.0511, -0.0043)	0.02
	Age 4	-0.0370 (-0.0673, -0.0066)	0.02	-0.0326 (-0.0651, -0.0001)	0.05	-0.0326 (-0.0629, -0.0024)	0.03
	Age 5	-0.0184 (-0.0553, 0.0186)	0.33	-0.0195 (-0.0586, 0.0195)	0.33	-0.0175 (-0.0539, 0.0190)	0.35
	Age 6	-0.0287 (-0.0725, 0.0150)	0.20	-0.0297 (-0.0757, 0.0163)	0.21	-0.0305 (-0.0735, 0.0126)	0.17
	Cohort	MODEL 3		MODEL 4		MODEL 5	
Models with individual-level & demographic covariates		RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
	Age 2	-0.0074 (-0.0271, 0.0123)	0.46	0.0076 (-0.0136, 0.0287)	0.48	0.0026 (-0.0189, 0.0241)	0.81
	Age 3	-0.0020 (-0.0302, 0.0263)	0.89	0.0099 (-0.0201, 0.0400)	0.52	0.0064 (-0.0237, 0.0366)	0.68
	Age 4	-0.0137 (-0.0496, 0.0223)	0.46	0.0065 (-0.0332, 0.0462)	0.75	0.0058 (-0.0330, 0.0447)	0.77
	Age 5	0.0094 (-0.0338, 0.0525)	0.67	0.0148 (-0.0329, 0.0626)	0.54	0.0217 (-0.0246, 0.0680)	0.36
	Age 6	-0.0051 (-0.0558, 0.0456)	0.84	-0.0062 (-0.0625, 0.0502)	0.83	0.0134 (-0.0402, 0.0671)	0.62
	Cohort	MODEL 6		MODEL 7: **FINAL MODEL**		MODEL 8	
Models with individual-level, demographic & spatial covariates		RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
	Age 2	0.0278 (0.0034, 0.0522)	0.03	0.0265 (0.0027, 0.0503)	0.03	0.0208 (-0.0026, 0.0441)	0.08
	Age 3	0.0389 (0.0053, 0.0726)[a]	0.02	0.0369 (0.0040, 0.0698)[a]	0.03	0.0260 (-0.0077, 0.0597)	0.13
	Age 4	0.0411 (-0.0048, 0.0871)	0.08	0.0373 (-0.0073, 0.0819)	0.10	0.0226 (-0.0202, 0.0654)	0.30
	Age 5	0.0630 (0.0074, 0.1187)	0.03	0.0578 (0.0035, 0.1122)	0.04	0.0519 (0.0003, 0.1035)	0.05
	Age 6	0.0440 (-0.0224, 0.1105)	0.19	0.0359 (-0.0289, 0.1008)	0.28	0.0393 (-0.0225, 0.1011)	0.21

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] birth year dropped in order for model to converge

List of Model Covariates

Model 0: unadjusted model

Model 1: child sex, child race, child ethnicity, maternal asthma, birth year, maternal age, maternal education

Model 2: child sex, child race, maternal asthma, birth year

Model 3: child sex, child race, maternal asthma, birth year, major demographic cluster

Model 4: child sex, child race, maternal asthma, birth year, minor demographic cluster

Model 5: child sex, child race, maternal asthma, birth year, median household income, median year structure built, median house value, percent less than high school, percent families in poverty

Model 6: child sex, child race, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤5 miles from metro Atlanta, >5–10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 7 (**FINAL MODEL**): child sex, child race, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 8: child sex, child race, maternal asthma, birth year, median household income, median year structure built, median house value, percent less than high school, percent families in poverty, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 5.7. Full adjusted model results for first year of life traffic $PM_{2.5}$ and asthma incidence at ages 2, 4, and 6

	Age 2 n = 17,958 Correlation = 0.16		Age 4 n = 10,998 Correlation = 0.20		Age 6 n = 6,628 Correlation = 0.24	
Parameter	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Intercept	0.0538 (0.0215, 0.0862)	<0.01	0.1304 (0.0710, 0.1898)	<0.01	0.2488 (0.1639, 0.3336)	<0.01
Traffic $PM_{2.5}$	0.0265 (0.0027, 0.0503)	0.03	0.0373 (-0.0073, 0.0819)	0.10	0.0359 (-0.0289, 0.1008)	0.28
Male sex	0.0433 (0.0346, 0.0521)	<0.01	0.0782 (0.0626, 0.0938)	<0.01	0.0859 (0.0639, 0.1079)	<0.01
Black race	0.0236 (0.0115, 0.0357)	<0.01	0.0299 (0.0091, 0.0508)	<0.01	0.0595 (0.0300, 0.0890)	<0.01
Unknown/other race	0.0024 (-0.0091, 0.0140)	0.68	0.0011 (-0.0207, 0.0229)	0.92	0.0409 (0.0077, 0.0740)	0.02
Maternal asthma	0.0609 (0.0444, 0.0774)	<0.01	0.1041 (0.0780, 0.1302)	<0.01	0.1519 (0.1165, 0.1873)	<0.01
Birth year: 2001	-0.0164 (-0.0362, 0.0035)	0.11	-0.0145 (-0.0462, 0.0171)	0.37	-0.0636 (-0.1032, -0.0240)	<0.01
Birth year: 2002	-0.0087 (-0.0288, 0.0113)	0.39	0.0077 (-0.0242, 0.0396)	0.64	-0.0339 (-0.0737, 0.0059)	0.10
Birth year: 2003	-0.0208 (-0.0401, -0.0016)	0.03	-0.0149 (-0.0464, 0.0166)	0.35	-0.0430 (-0.0832, -0.0027)	0.04
Birth year: 2004	-0.0109 (-0.0308, 0.0090)	0.28	-0.0334 (-0.0645, -0.0022)	0.04	-0.0693 (-0.1125, -0.0262)	<0.01
Birth year: 2005	-0.0065 (-0.0272, 0.0141)	0.53	-0.0044 (-0.0382, 0.0294)	0.80	-0.0446 (-0.0902, 0.0011)	0.06
Birth year: 2006	-0.0165 (-0.0359, 0.0029)	0.10	-0.0015 (-0.0368, 0.0338)	0.93	-0.0519 (-0.0958, -0.0081)	0.02
Birth year: 2007	-0.0181 (-0.0380, 0.0019)	0.08	-0.0339 (-0.0675, -0.0002)	0.05	-0.0321 (-0.0829, 0.0186)	0.21
Birth year: 2008	-0.0197 (-0.0400, 0.0006)	0.06	-0.0246 (-0.0584, 0.0092)	0.15	-	-
Birth year: 2009	-0.0286 (-0.0486, -0.0087)	<0.01	-0.0435 (-0.0839, -0.0030)	0.04	-	-
Birth year: 2010	-0.0142 (-0.0362, 0.0079)	0.21	-	-	-	-
Metro Atlanta	-0.0379 (-0.0581, -0.0177)	<0.01	-0.0593 (-0.0961, -0.0225)	<0.01	-0.0808 (-0.1363, -0.0253)	<0.01
≤10 mi from metro Atlanta	-0.0173 (-0.0299, -0.0047)	<0.01	-0.0250 (-0.0468, -0.0032)	0.02	-0.0301 (-0.0619, 0.0017)	0.06
Demographic cluster A.2	0.0131 (-0.0055, 0.0316)	0.17	0.0323 (-0.0009, 0.0655)	0.06	-0.0073 (-0.0549, 0.0402)	0.76
Cluster A.3	0.0201 (0.0051, 0.0351)	<0.01	0.0430 (0.0160, 0.0699)	<0.01	0.0052 (-0.0342, 0.0445)	0.80
Cluster B.1	-0.0189 (-0.0406, 0.0028)	0.09	-0.0476 (-0.0886, -0.0067)	0.02	-0.0631 (-0.1282, 0.0020)	0.06
Cluster B.3/B.4	0.0199 (-0.0053, 0.0450)	0.12	0.0556 (0.0097, 0.1015)	0.02	0.0204 (-0.0445, 0.0853)	0.54
Cluster C.1/C.2	0.0214 (-0.0068, 0.0497)	0.14	0.0569 (0.0029, 0.1109)	0.04	-0.0217 (-0.0936, 0.0502)	0.55
Cluster C.3/C.4	0.0209 (-0.0211, 0.0628)	0.33	0.1240 (0.0447, 0.2034)	<0.01	0.1449 (0.0328, 0.2570)	0.01
Cluster D.1	0.0138 (-0.0057, 0.0333)	0.17	0.0430 (0.0086, 0.0774)	0.01	0.0002 (-0.0493, 0.0498)	0.99
Cluster D.3	0.0121 (-0.0178, 0.0420)	0.43	0.0493 (-0.0053, 0.1039)	0.08	0.0326 (-0.0465, 0.1116)	0.42
Cluster D.4	0.0047 (-0.0173, 0.0267)	0.68	0.0322 (-0.0085, 0.0729)	0.12	-0.0102 (-0.0688, 0.0484)	0.73
Cluster D.5	0.0333 (-0.0062, 0.0729)	0.10	0.0236 (-0.0415, 0.0886)	0.48	-0.0139 (-0.0997, 0.0718)	0.75
Cluster D.6/D.7	0.0051 (-0.0373, 0.0476)	0.81	0.0450 (-0.0257, 0.1156)	0.21	0.0814 (-0.0202, 0.1829)	0.12

RD = Risk Difference, CI = Confidence Interval, p = p-value, - = variable not included in model, mi = miles
Reference group: female sex, white race, no or unknown maternal asthma, birth year 2000, >10 miles from metro Atlanta, demographic cluster A.1

Table 5.8. Risk differences for prenatal and first year of life traffic $PM_{2.5}$ and asthma incidence, unadjusted and final adjusted models

Cohort (adjusted model n)	First Year of Life $PM_{2.5}$ Unadjusted Model		First Year of Life $PM_{2.5}$ Final Adjusted Model	
	RD (95% CI)	p	RD (95% CI)	p
Age 2 (n=17,958)	-0.0221 (-0.0390, -0.0053)	0.01	0.0265 (0.0027, 0.0503)	0.03
Age 3 (n=14,249)	-0.0316 (-0.0552, -0.0080)	<0.01	0.0369 (0.0040, 0.0698)[a]	0.03
Age 4 (n=10,998)	-0.0370 (-0.0673, -0.0066)	0.02	0.0373 (-0.0073, 0.0819)	0.10
Age 5 (n=8,591)	-0.0184 (-0.0553, 0.0186)	0.33	0.0578 (0.0035, 0.1122)	0.04
Age 6 (n=6,628)	-0.0287 (-0.0725, 0.0150)	0.20	0.0359 (-0.0289, 0.1008)	0.28
Cohort (adjusted model n)	Prenatal $PM_{2.5}$ Unadjusted Model		Prenatal $PM_{2.5}$ Final Adjusted Model	
	RD (95% CI)	p	RD (95% CI)	p
Age 2 (n=15,629)	-0.0239 (-0.0420, -0.0058)	<0.01	0.0298 (0.0051, 0.0544)	0.02
Age 3 (n=12,432)	-0.0365 (-0.0617, -0.0113)	<0.01	0.0250 (-0.0087, 0.0587)	0.15
Age 4 (n=9,619)	-0.0377 (-0.0700, -0.0054)	0.02	0.0252 (-0.0199, 0.0703)	0.27
Age 5 (n=7,520)	-0.0266 (-0.0656, 0.0124)	0.18	0.0307 (-0.0240, 0.0854)	0.27
Age 6 (n=5,805)	-0.0308 (-0.0773, 0.0157)	0.19	0.0216 (-0.0432, 0.0864)	0.51

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] birth year dropped in order for model to converge

Final adjusted models control for child sex, child race, maternal asthma, birth year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 5.9. Risk differences for first year of life traffic $PM_{2.5}$ and persistent asthma

Cohort	Children with persistent asthma [n (%)]	Including children with incident, but not persistent asthma as non-cases		Excluding children with incident, but not persistent asthma	
		RD (95% CI)	p	RD (95% CI)	p
Age 2	1,994 (11.1)	0.0265 (0.0027, 0.0503)[b]	0.03	0.0265 (0.0027, 0.0503)[b]	0.03
Age 3	2,196 (15.4)	—	—	—	—
Age 4	1,965 (17.9)	0.0490 (0.0098, 0.0882)	0.01	0.0507 (0.0095, 0.0920)	0.02
Age 5	1,629 (19.0)	0.0538 (0.0068, 0.1008)	0.02	0.0589 (0.0082, 0.1096)	0.02
Age 6	1,350 (20.4)	0.0284 (-0.0274, 0.0843)	0.32	0.0343 (-0.0275, 0.0962)	0.28

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge even if dropped birth year from the model, [b] Identical analysis to incident asthma at age 2

Persistent asthma defined as a child meeting the incident asthma classification (at least 1 asthma diagnosis (ICD-9 493.XX) and 1 asthma-related medication dispensing) with evidence of asthma in the past year (at least 1 asthma diagnosis or 1 asthma-related medication dispensing)

Models adjust for child sex, child race, maternal asthma, birth year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 5.10. First year of life traffic $PM_{2.5}$ and incident asthma by age 5 among children enrolled through age 5 (n=8,592), comparing different outcome definitions

Outcome Definition	n(%) meeting definition	UNADJUSTED		ADJUSTED	
		RD (95% CI)	p	RD (95% CI)	p
1. 1 asthma or wheeze diagnosis	3,082 (35.9)	-0.0121 (-0.0517, 0.0275)	0.55	0.0671 (0.0090, 0.1252)	0.02
2. 1 asthma diagnosis	2,570 (29.9)	-0.0124 (-0.0501, 0.0253)	0.52	0.0740 (0.0186, 0.1294)	<0.01
3. 2 asthma diagnoses	1,757 (20.5)	-0.0081 (-0.0409, 0.0247)	0.63	0.0495 (0.0019, 0.0971)	0.04
4. 3 asthma diagnoses	1,311 (15.3)	0.0117 (-0.0174, 0.0408)	0.43	0.0527 (0.0101, 0.0954)	0.02
6. 2 asthma diagnoses OR 1 acute asthma diagnosis	1,856 (21.6)	-0.0130 (-0.0465, 0.0206)	0.45	0.0539 (0.0050, 0.1028)	0.03
7. 1 asthma diagnosis OR 2 medication dispensings	3,322 (38.7)	-0.0336 (-0.0739, 0.0067)	0.10	0.0593 (0.0004, 0.1181)	0.05
8. 1 asthma diagnosis AND 1 medication dispensing *(KAPPA study definition)*	2,465 (28.7)	-0.0184 (-0.0553, 0.0186)	0.33	0.0578 (0.0035, 0.1122)	0.04
9. 1 asthma diagnosis AND 2 medication dispensings	2,168 (25.2)	-0.0167 (-0.0522, 0.0189)	0.36	0.0655 (0.0139, 0.1172)	0.01
10. 1 asthma diagnosis OR 2 medication dispensings 1 of which must be a steroid	2,685 (31.3)	-0.0212 (-0.0594, 0.0170)	0.28	0.0790 (0.0231, 0.1349)	<0.01
11. 1 asthma diagnosis AND 2 medication dispensings 1 of which must be a steroid	1,388 (16.2)	0.0064 (-0.0239, 0.0367)	0.68	0.0579 (0.0144, 0.1013)[a]	<0.01
12. 1 asthma diagnosis OR 1 controller dispensing	2,715 (31.6)	-0.0229 (-0.0612, 0.0155)	0.24	0.0742 (0.0180, 0.1303)	<0.01
13. 1 asthma diagnosis AND 1 controller dispensing	1,434 (16.7)	-0.0003 (-0.0311, 0.0305)	0.98	0.0526 (0.0085, 0.0967)[a]	0.02
14. 1 asthma diagnosis AND (2 reliever dispensings OR 1 controller dispensing)	2,181 (25.4)	-0.0196 (-0.0552, 0.0161)	0.28	0.0617 (0.0098, 0.1135)	0.02
15. Any of the following: a) 1 asthma diagnosis AND 1 medication dispensing in the same year, b) 1 asthma-related ED visit or hospitalization, c) 3 asthma diagnoses	2,450 (28.5)	-0.0124 (-0.0496, 0.0248)	0.51	0.0655 (0.0108, 0.1202)	0.02

[a]Birth year dropped in order for model to converge. These are the minimum required events for each case definition. Definition numbers align with numbers from Dissertation Chapter 4 (excluded definition 5 (atopic asthma)). Only 1 diagnosis per day counted. ED = emergency department; Asthma diagnosis = ICD-9 code 493.XX; Wheeze diagnosis = ICD-9 code 786.07; Acute asthma diagnosis = a) emergency department or inpatient asthma diagnosis *or* b) asthma diagnosis with status asthmaticus or acute exacerbation (ICD-9 codes 493.01, 493.02, 493.11, 493.12, 493.21, 493.22, 493.91, 493.92); Asthma controller (underlined medications contain a steroid) = Aminophylline, beclomethasone diproprionate**,** budesonide, budesonide/formoterol fumarate, cromolyn sodium, fluticasone propionate, fluticasone/sameterol, mometasone furoate, montelukast sodium, salmeterol xinafoate, theophylline anhydrous, tiotropium bromide, triamcinolone acetonide; Asthma reliever = Albuterol, albuterol sulfate, ipratropium bromide, ipratropium/albuterol sulfate, levalbuterol, metaproterenol sulfate; Medication dispensing = dispensing of any asthma controller or reliever. Adjusted models control for child sex, child race, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 5.11. First year of life traffic $PM_{2.5}$ and incident asthma, comparing different measures of association

	UNADJUSTED		**ADJUSTED**	
Cohort	RISK DIFFERENCES			
	RD (95% CI)	p	RD (95% CI)	p
Age 2	-0.0221 (-0.0390, -0.0053)	0.01	0.0265 (0.0027, 0.0503)	0.03
Age 3	-0.0316 (-0.0552, -0.0080)	<0.01	0.0369 (0.0040, 0.0698)[a]	0.03
Age 4	-0.0370 (-0.0673, -0.0066)	0.02	0.0373 (-0.0073, 0.0819)	0.10
Age 5	-0.0184 (-0.0553, 0.0186)	0.33	0.0578 (0.0035, 0.1122)	0.04
Age 6	-0.0287 (-0.0725, 0.0150)	0.20	0.0359 (-0.0289, 0.1008)	0.28
Cohort	RISK RATIOS			
	RR (95% CI)	p	RR (95% CI)	p
Age 2	0.83 (0.71, 0.97)	0.02	1.37 (1.09, 1.72)	<0.01
Age 3	0.86 (0.75, 0.97)	0.02	1.27 (1.04, 1.54)	0.02
Age 4	0.87 (0.76, 0.98)	0.03	1.18 (0.98, 1.42)	0.09
Age 5	0.94 (0.83, 1.07)	0.37	1.20 (0.99, 1.46)	0.06
Age 6	0.92 (0.81, 1.06)	0.24	1.09 (0.90, 1.33)	0.39
Cohort	ODDS RATIOS			
	OR (95% CI)	p	OR (95% CI)	p
Age 2	0.81 (0.68, 0.96)	0.02	1.42 (1.09, 1.85)	<0.01
Age 3	0.82 (0.70, 0.96)	0.02	1.35 (1.06, 1.73)	0.02
Age 4	0.82 (0.70, 0.97)	0.02	1.25 (0.97, 1.62)	0.08
Age 5	0.92 (0.77, 1.10)	0.36	1.33 (1.00, 1.75)	0.05
Age 6	0.88 (0.72, 1.08)	0.22	1.16 (0.85, 1.58)	0.35
Cohort	HAZARD RATIOS [b]			
	HR (95% CI)	p	HR (95% CI)	p
Age 2	0.79 (0.68, 0.92)	<0.01	1.33 (1.06, 1.67)	0.01
Age 3	0.81 (0.71, 0.91)	<0.01	1.29 (1.07, 1.56)	<0.01
Age 4	0.82 (0.73, 0.92)	<0.01	1.31 (1.10, 1.57)	<0.01
Age 5	0.83 (0.75, 0.93)	<0.01	1.30 (1.10, 1.54)	<0.01
Age 6	0.83 (0.75, 0.93)	<0.01	1.28 (1.08, 1.51)	<0.01

RD = Risk Difference, RR = Risk Ratio, OR = Odds Ratio, HR = Hazard Ratio, all calculated for 1 $\mu g/m^3$
CI = Confidence Interval, p = p-value, [a] birth year dropped in order for model to converge, [b] Hazard ratios calculated using the 22,987 children in the KAPPA cohort with first year of life $PM_{2.5}$ estimates enrolled in Kaiser Permanente Georgia until at least their first birthday. The outcome of interest was asthma incidence between the first birthday and the birthday of each age cohort. For example, in the age 4 analysis the outcome of interest is asthma between the first and fourth birthdays.

Adjusted models control for child sex, child race, maternal asthma, birth year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Figure 5.1. 2011 RLINE-modeled primary $PM_{2.5}$ from traffic emissions ($\mu g/m^3$)

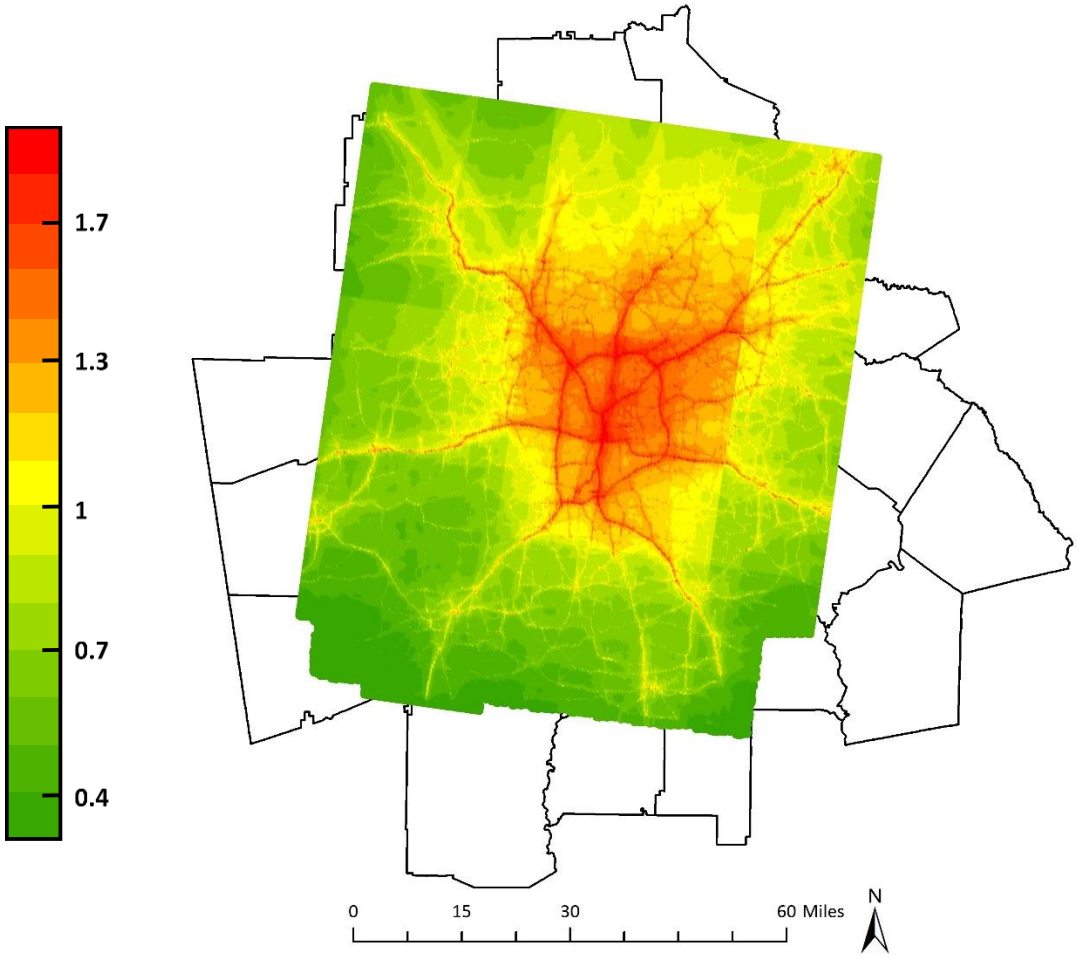
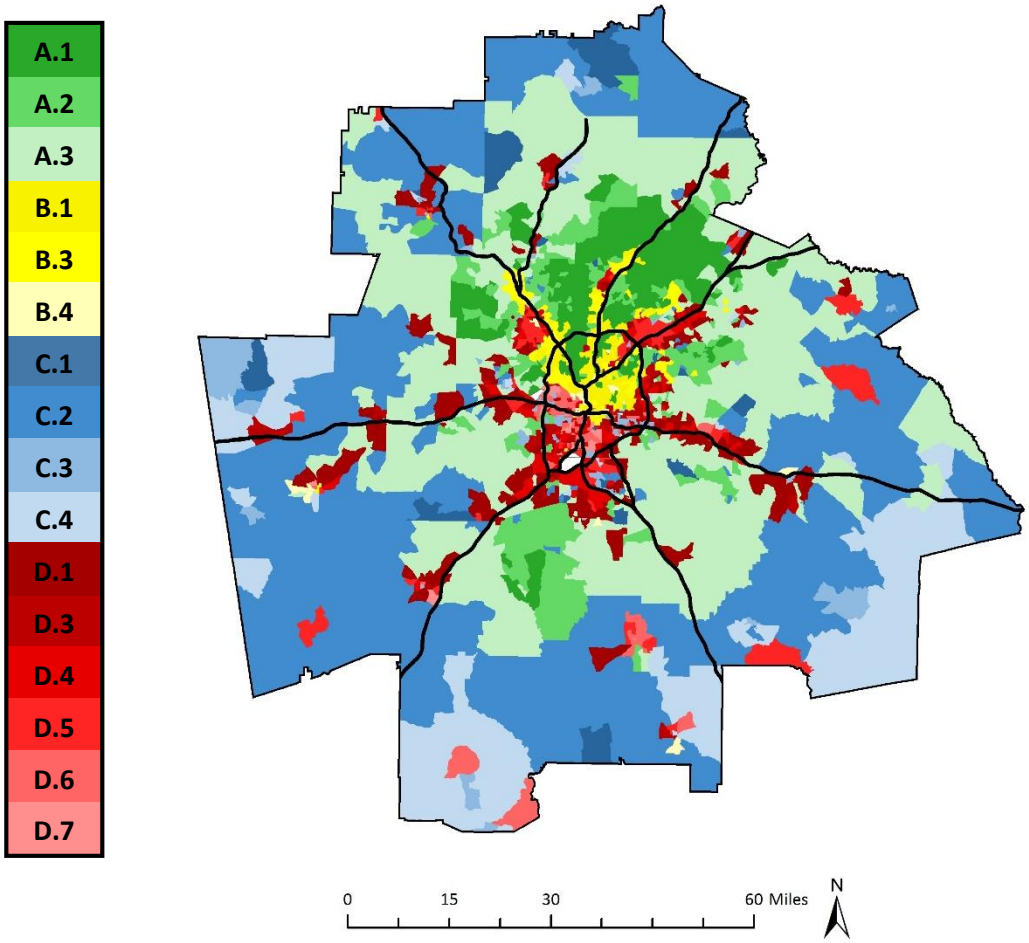


Figure 5.2. Minor demographic clusters in the 29 county metropolitan Atlanta area



Minor demographic cluster descriptions are included in Table A1 located in Appendix A.

Reference: Demographic Clusters of Georgia. Georgia Department of Public Health: Office of Health Indicators for Planning: September 2012.

Figure 5.3. Distribution of prenatal (n=19,951) and first year of life (n=23,100) traffic $PM_{2.5}$ assigned exposure in the KAPPA cohort

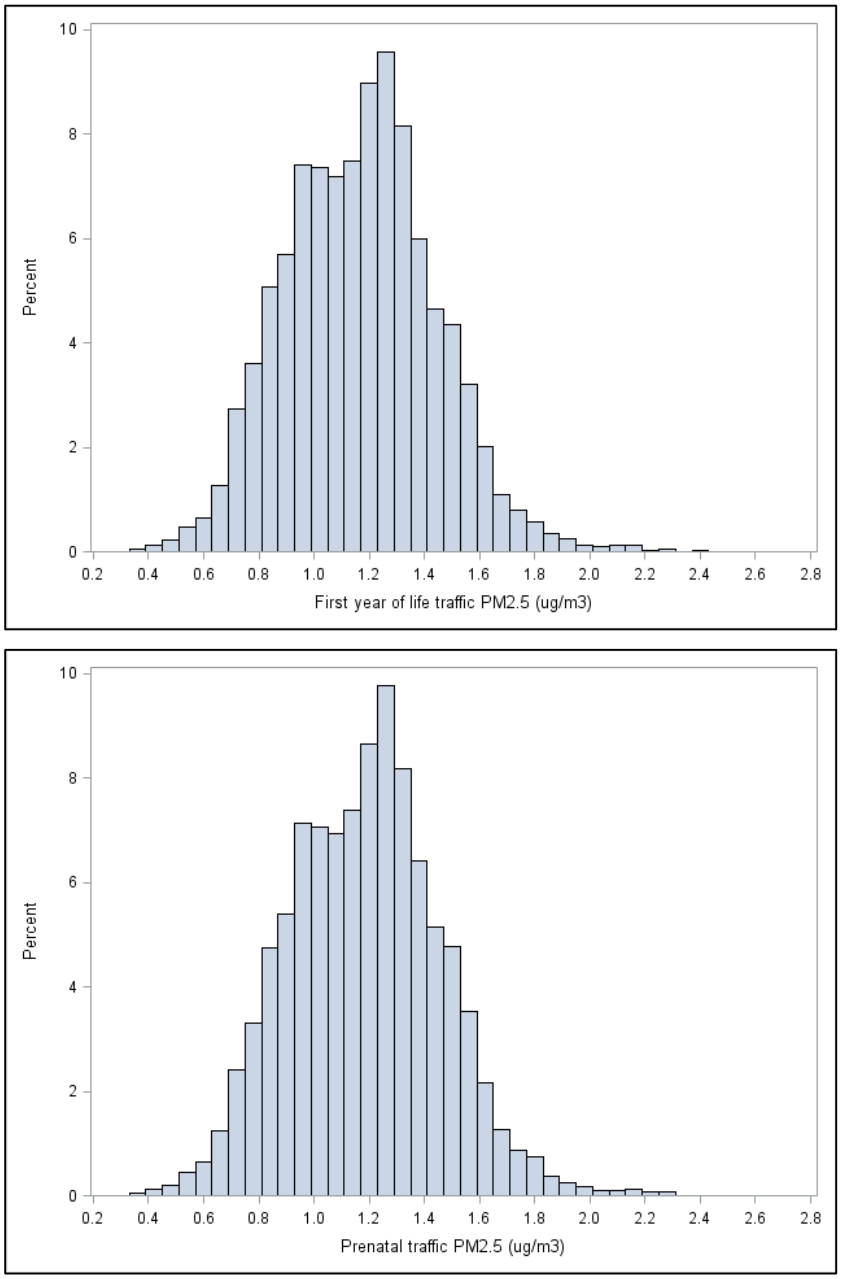
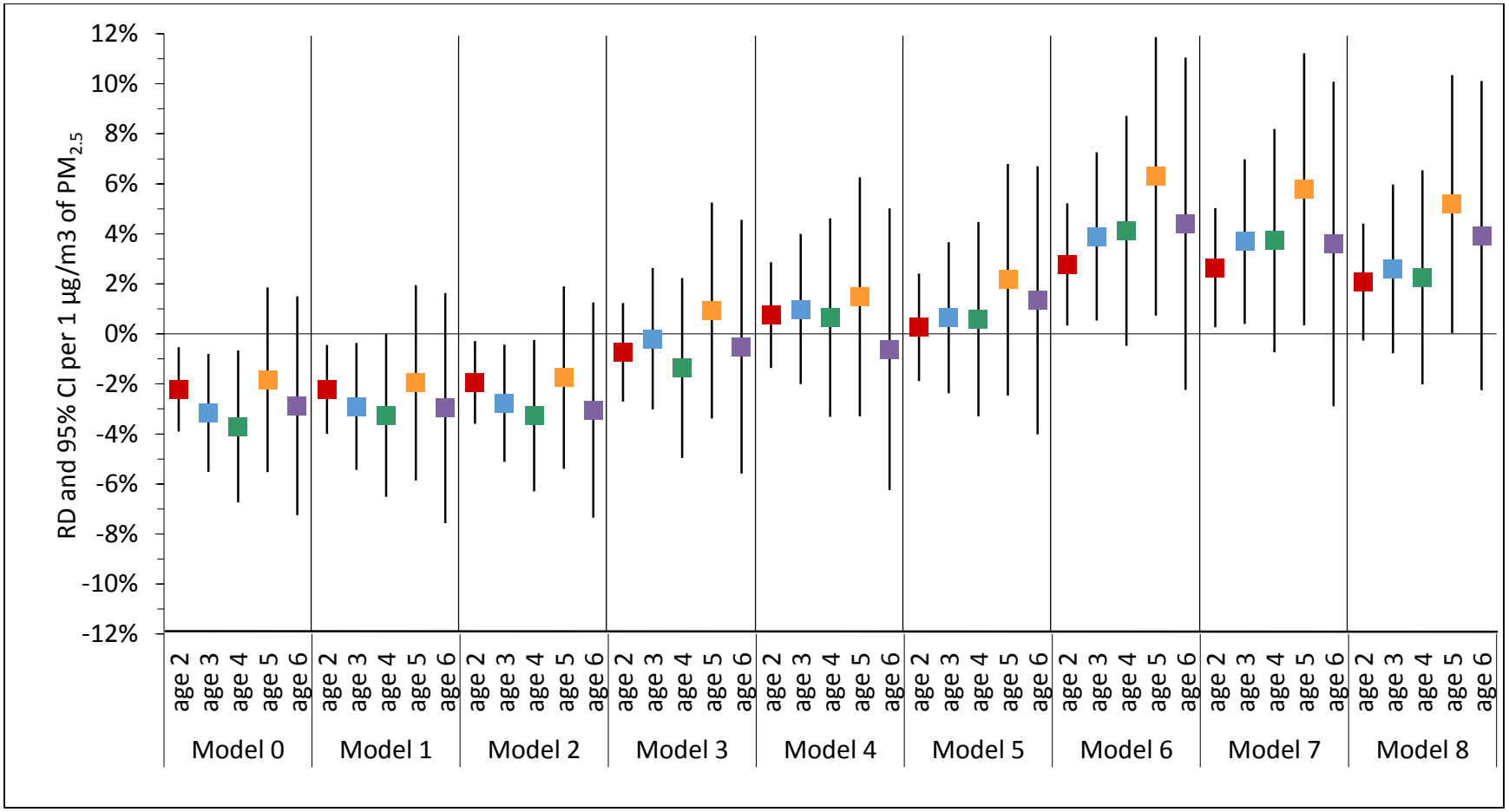


Figure 5.4. Risk differences for first year of life traffic $PM_{2.5}$ and asthma incidence in models adjusting for different sets of covariates



RD = Risk Difference, CI = Confidence Interval

List of Model Covariates
Model 0: unadjusted model

Model 1: child sex, child race, child ethnicity, maternal asthma, birth year, maternal age, maternal education

Model 2: child sex, child race, maternal asthma, birth year

Model 3: child sex, child race, maternal asthma, birth year, major demographic cluster

Model 4: child sex, child race, maternal asthma, birth year, minor demographic cluster

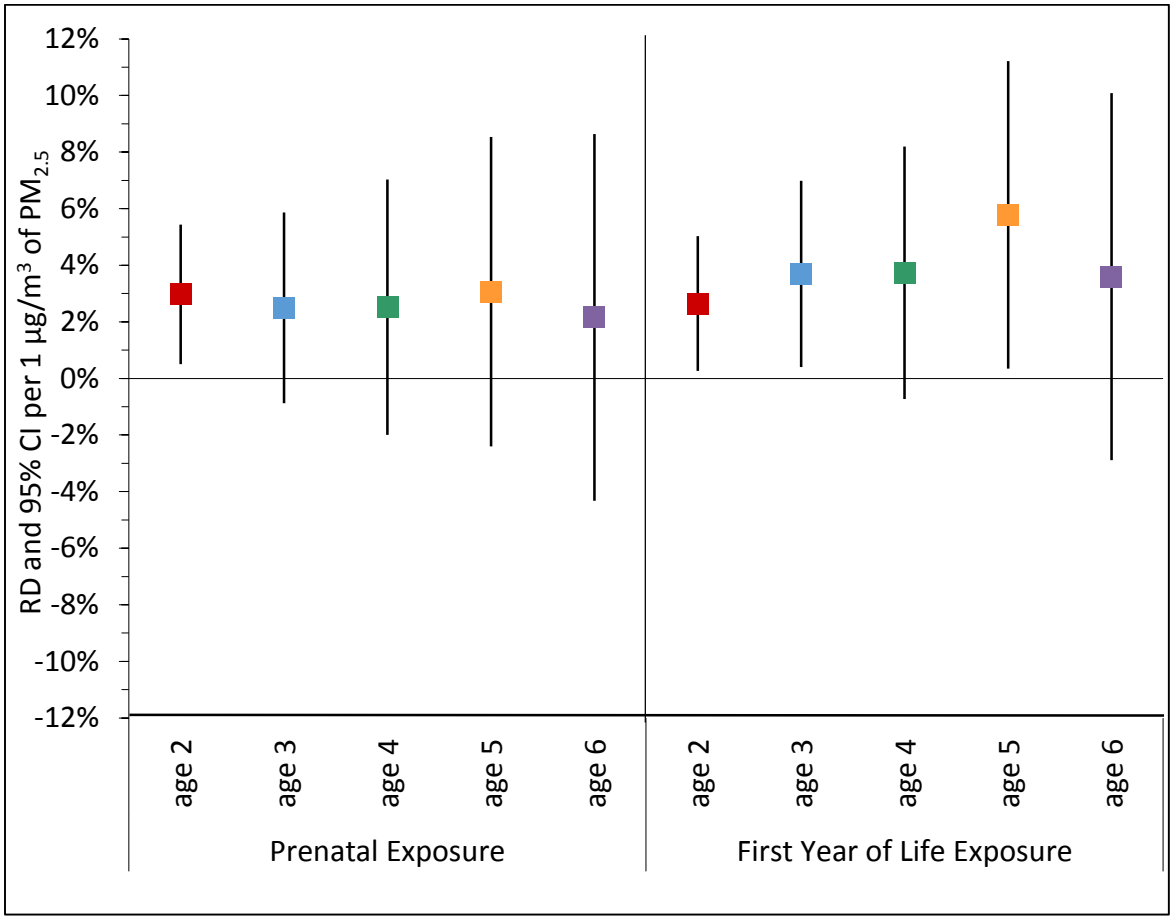
Model 5: child sex, child race, maternal asthma, birth year, median household income, median year structure built, median house value, percent less than high school, percent families in poverty

Model 6: child sex, child race, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤5 miles from metro Atlanta, >5–10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 7 (**FINAL MODEL**): child sex, child race, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 8: child sex, child race, maternal asthma, birth year, median household income, median year structure built, median house value, percent less than high school, percent families in poverty, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Figure 5.5. Adjusted risk differences for prenatal and first year of life traffic $PM_{2.5}$ and asthma incidence from final models



RD = Risk Difference, CI = Confidence Interval. Models control for child sex, child race, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

CHAPTER 6

Prenatal and first year of life exposure to total $PM_{2.5}$ and childhood asthma incidence in a birth cohort

INTRODUCTION

Particulate matter (PM) is a heterogeneous mixture of matter suspended in air created by a variety of sources including motor vehicles, power plants, and naturally occurring processes. The chemical composition of PM varies regionally; fine particulate matter (particulate matter equal to or less than 2.5 micrometers in diameter ($PM_{2.5}$)) in the Southeastern United States is composed of more sulfate and less nitrates than in other parts of the country (Godish 2004, Bell 2012). In the metropolitan area of Atlanta, $PM_{2.5}$ is primarily composed of organic carbon and sulfates, with smaller contributions from elemental carbon, metals, nitrate and ammonium. There is some epidemiologic evidence that the health effects of $PM_{2.5}$ vary depending on source, suggesting that PM toxicity may be determined by composition (Bell et al. 2009, Ito et al. 2011, Gass et al. 2015).

A large body of literature has assessed the acute respiratory effects of exposure to $PM_{2.5}$ from all sources in children (Villeneuve et al. 2007, Silverman et al. 2010, Alhanti et al. 2015, Fan et al. 2015). Several of these studies have found positive associations between $PM_{2.5}$ and asthma-related emergency department visits and hospital admissions. The majority of studies looking instead at long-term respiratory effects of chronic exposure to $PM_{2.5}$ focus only on pollution from traffic and don't capture exposure to total $PM_{2.5}$ from all sources (Clark et al. 2010, Gehring et al. 2010, Carlsten et al. 2011). Results of many of these studies are suggestive of a positive association between traffic-related $PM_{2.5}$ and childhood asthma incidence. While fewer studies have focused on its effects on asthma development, exposure to total $PM_{2.5}$ may be biologically relevant for the development of asthma, and

recent study results show an association between prenatal $PM_{2.5}$ and asthma incidence by age 6 among boys (Leon Hsu et al. 2015).

This study addresses current knowledge gaps by assessing the association between prenatal and first year of life exposure to total $PM_{2.5}$ and development of childhood asthma. This work was completed in a birth cohort of children enrolled in Kaiser Permanente Georgia using daily $PM_{2.5}$ data created by an innovative new method that integrates modeled $PM_{2.5}$ data from all sources and spatially resolved estimates of primary $PM_{2.5}$ from traffic.

METHODS

The Kaiser Air Pollution and Pediatric Asthma study (KAPPA) is a birth cohort of 24,608 children born between 2000 and 2010, residing in metropolitan Atlanta, and enrolled in Kaiser Permanente Georgia. This study has been previously described in Chapter 5.

Prenatal and First Year of life $PM_{2.5}$ Exposure Estimates

Colleagues at Georgia Institute of Technology modeled daily total $PM_{2.5}$ concentrations for 2002 through 2010 at 250 meter grid resolution for the Atlanta metropolitan area. These data were created using a downscaling approach that integrates estimates of $PM_{2.5}$ from all sources at 12 kilometer grid resolution and estimates of primary $PM_{2.5}$ from traffic, which is more spatially heterogeneous, at 250 meter grid resolution to produce estimates of total $PM_{2.5}$ at 250 meter resolution (Bates et al. 2016). Data from three sources were used to obtain these estimates: 2011 RLINE (a research line-source dispersion model for near-surface releases) annual estimates of primary $PM_{2.5}$ from vehicle emissions, 2002-2010 CMAQ (community multiscale air quality model) daily estimates of total $PM_{2.5}$,

and observations from stationary air pollution monitors. RLINE estimates (described in detail in Chapter 5) were provided by Atlanta Regional Commission and calibrated by the methods described in Chapter 5 (Zhai et al. 2015).

The starting dataset for this downscaling process was a 12-km by 12-km fused-CMAQ dataset. CMAQ is an Eulerian emissions-based chemical transport model developed by the U.S. Environmental Protection Agency described in detail in Chapter 3 (Byun DW et al. 1999, Byun et al. 2006). While CMAQ has impressive spatial and temporal coverage, it tends to under-estimate daily variability in $PM_{2.5}$ and the data it creates are impacted by the accuracy of data inputs and model specifications. Estimates from stationary air pollution monitors better estimate day-to-day variability, but are spatially sparse. Observations from 42 ambient $PM_{2.5}$ monitors in Georgia and surrounding states were used to improve the accuracy of daily 12-km CMAQ estimates. This data fusion process was completed by first creating two sets of spatial fields using both the CMAQ data and monitoring data, with temporal variability in one spatial field driven by monitoring observations, and temporal variability in the other field driven by CMAQ. These datasets were then averaged, optimizing temporal variance prediction, using a weighting process based on daily error estimates. This novel fusion approach is described in more detail by two manuscripts (one published in 2015 and one currently submitted for publication) (Friberg et al. 2015, Hao et al. 2015).

This fused-CMAQ dataset was downscaled from 12 km resolution to 250 m resolution using the calibrated RLINE data for the metropolitan Atlanta area. In order to avoid double counting primary $PM_{2.5}$ from traffic (since RLINE estimates $PM_{2.5}$ from traffic, and CMAQ captures $PM_{2.5}$ from all sources including traffic), fused-CMAQ estimates were first adjusted to remove $PM_{2.5}$ from traffic impacts. This was achieved by subtracting the average of RLINE estimates over each 12 km CMAQ grid from each 12 km CMAQ estimate.

Next, the adjusted 12 km CMAQ estimates were downscaled using bilinear interpolation to create spatially smooth $PM_{2.5}$ estimates at 250 meter grid resolution. This downscaling process made the assumption that the 12 km grid estimate applied to the 12 km grid centroid. It then used a weighted average approach to calculate an estimate for each 250 meter grid taking into account the centroid estimates of all surrounding 12 km grids. Lastly, the 2011 RLINE estimates were added to these downscaled CMAQ estimates in order to add back in $PM_{2.5}$ from traffic impacts. The resulting daily $PM_{2.5}$ estimates attempt to capture both spatial and temporal variability from all sources over this time period. Please note, the air quality data for this work are still being developed and may change before the publication of results.

$PM_{2.5}$ exposure during pregnancy and the first year of life was estimated using these daily $PM_{2.5}$ concentrations and mother and child residential history captured in the administrative data from Kaiser Permanente Georgia. The first year of life period was defined as the date of birth to the day before the first birthday. The start of the gestational period was determined by first counting back the number of weeks gestation (using gestational age at birth from the birth certificate) from the date of birth. Then, to account for the obstetric convention of starting the gestational week count on the day of the last menstrual period, the start date of the prenatal period was moved forward 2 weeks to the estimated date of conception. In the absence of a birth certificate (for 24.5% of the cohort), 38 weeks of gestation were assumed. Residential $PM_{2.5}$ was calculated for each day in the exposure windows, the resulting estimates were averages of all daily estimates in each exposure window of interest. Daily $PM_{2.5}$ estimates were available for 2002 to 2010, but prenatal exposure windows for children in the cohort ranged from 1999 to 2010 and first year of life exposure windows ranged from 2000 to 2011. For estimating $PM_{2.5}$ exposure, 2002 data were used for 1999-2001, and 2010 data were used for 2011. Children residing

outside the pollution region at any time during the first year of life period were excluded from analyses on first year of life exposure. Children whose mothers resided outside the pollution region at any time during the prenatal period were excluded from analyses on prenatal exposure. A secondary analysis was completed excluding children with any dates of their exposure windows falling in years for which air pollution data was unavailable.

Analytic Approach

Binomial linear regression was used to estimate the association between prenatal and first year of life $PM_{2.5}$ and cumulative asthma incidence at follow-up ages two through six. These models were implemented using generalized estimating equations with an exchangeable correlation structure to account for correlation between siblings. Robust variance estimation was used to account for potential variance heteroscedasticity and misspecification of the working correlation matrix. Asthma was defined as at least one asthma diagnosis (ICD-9 code 493.XX) and one asthma-related medication dispensing after the first year of life. A list of medications considered asthma-related is provided in Table A2 in Appendix A. Cumulative asthma incidence was assessed so once a child was classified as having asthma they were classified as asthmatic at every subsequent follow-up age. Each model assessed the association between $PM_{2.5}$ and cumulative asthma incidence by each follow-up age among children enrolled in Kaiser Permanente Georgia until that follow-up age (e.g., cumulative asthma incidence by the fifth birthday among children enrolled until their fifth birthday). The following factors were considered as potential confounders: race, sex, ethnicity, maternal age, maternal asthma, maternal education, paternal education, maternal marital status, neighborhood socioeconomic status (SES) and city region. Neighborhood SES was defined using demographic clusters from Georgia Department of Public Health that integrate data from the U.S. Census to characterize census block groups

into 4 major demographic clusters and 18 minor demographic clusters (Figure 5.2) (Demographic Clusters of Georgia 2012, Zhou 2012, Georgia Department of Public Health et al. 2013-2015). The use of variables from the American Community Survey at census tract level to control for neighborhood SES was also investigated (e.g. median household income, median home value) (U.S. Census Bureau 2010). Due to concerns about potential confounding by factors that vary with distance to city center not already controlled for (e.g., healthcare access, agricultural exposures, cockroach allergen exposure), the impact of adjusting for city region was explored. All models were constructed to estimate the association between first year of life $PM_{2.5}$ and childhood asthma incidence including assessments of potential confounding and interaction. The final adjusted model was then used to examine the association between prenatal $PM_{2.5}$ and childhood asthma incidence in order to produce adjusted results comparable to those of first year of life $PM_{2.5}$ exposure. A secondary adjusted analysis of the association of prenatal $PM_{2.5}$ exposure was completed excluding covariates that would not be expected to be related to prenatal exposure, and hence are not confounders of the association between prenatal $PM_{2.5}$ exposure and asthma incidence. Lastly, both exposures were included in the same model to tease out the relative contributions of exposure in each of these developmental windows. More details about this modeling approach and the selection and definition of potential confounders is included in the methods section of Chapter 5.

There are temporal trends in both $PM_{2.5}$ and asthma incidence. $PM_{2.5}$ varies seasonally and annually due to changes in meteorology, emissions, regulation, and population size among other factors. Asthma incidence varies seasonally due to pollen, illness, and weather, and also over time due to changes in doctors' diagnosing habits and potential changes in the true underlying burden of asthma in the population. Temporal trends in both our exposure and outcome necessitate control for these trends to prevent a

spurious association between $PM_{2.5}$ and asthma. In models assessing exposure during the first year of life, we explored the use of both month and year of birth categories, and transformed cubic splines on date of birth to control for these trends. The cubic splines we used were linearly transformed in order to reduce correlation between spline variables. For prenatal exposure models we took a similar approach, but instead used month and year of conception categories and date of conception for the time splines. This adjustment aimed to control for confounding by any factor, known or unknown, which exhibits seasonal or longer-term trends.

It is not uncommon for binomial risk difference models not to converge. A frequent cause of convergence difficulties is that at least one individual in the model has a predicted probability outside of, or very close to, the probability bounds of zero to one. When this occurs, the variance covariance structure does not converge leading to model output that only includes regression coefficients with no variance estimation. In this situation, we used the model to output the predicted probability of the outcome for each child to identify the children with problematic values. If a small number of children were preventing convergence, these children were deleted from the cohort and the model was re-run. In the event that new individuals had the same issue in the new model, the approach was used iteratively until a cohort of children was determined for whom the model converged or until it appeared that this approach will not resolve the convergence issues.

Additional Analyses and Sensitivity Analyses

We completed all additional analyses and sensitivity analyses discussed in the methods section of Chapter 5. These included analyses using asthma case definitions restricted to less-transient asthma phenotypes (i.e. using persistent asthma as an outcome and using events between age 2 and 5 to inform disease classification at age 2), analyses

assessing the impact of data cleaning decisions by excluding certain groups of children (i.e. children with any of the following characteristics: children missing race information, children missing 90 days or more of residence data during exposure windows, children not linked to mothers or for whom their maternal match was deemed less reliable, and children not linked to birth certificates), analyses investigating potential selection bias (e.g., restricting analyses to children followed until age 6), analyses using different case definitions for asthma, and analyses using alternative measures of association (i.e. risk ratios, odds ratios, and hazard ratios).

There is both spatial and temporal variability in our exposure data. In order to achieve purely spatial contrasts of exposure, we completed a sensitivity analysis assigning exposure with 2010 annual averages instead of the daily data from 2002-2010. For this analysis, we first calculated the average of the daily estimates in 2010 for each 250 meter grid. We then calculated first year of life $PM_{2.5}$ exposure for each child using these annual averages for each grid, rather than using the daily pollution data as was done for our main exposure estimates. Unadjusted and adjusted models were completed assessing the association between this exposure estimate and asthma incidence by ages two through six.

Lastly, since our exposure periods extend beyond the years for which we have $PM_{2.5}$ data, we applied data from 2002 to 1999-2001 and data from 2010 to 2011. As an additional sensitivity analysis, we excluded children with exposure periods including years for which pollution data were unavailable to see if this substitution of pollution data from different years impacted estimates of association.

RESULTS

Descriptive Results

Estimated $PM_{2.5}$ exposure in the first year of life was available for 23,100 children. Of these children, 19,951 also had information on prenatal $PM_{2.5}$ exposure. This is a racially diverse cohort, with 34.6% of children identifying with African American race and 12% of children identifying with other non-white racial groups (Table 6.1). As may be expected from an HMO cohort, there is high socioeconomic status in this population; 58.2% of children are born to mothers who attended at least some college, and 62.3% of children have their residence at birth located in major demographic cluster A, the demographic cluster with the highest SES (Table 6.1, Table 6.2). Only 10.5% of children in the cohort live inside metropolitan Atlanta as defined by living inside the I-285 perimeter highway that encircles Atlanta and (on average) these children have the highest prenatal and first year of life $PM_{2.5}$ exposure (Table 6.2). Average $PM_{2.5}$ exposure estimates are lowest among children who live more than 10 miles from the city center (Table 6.2). Overall, children of black race have higher $PM_{2.5}$ exposure than children of white race or children identifying with other racial groups, but this pattern does not hold in all regions of the city (Table 6.3). Cumulative asthma incidence, defined as at least one asthma diagnosis and one asthma-related medication dispensing after the first year of life, increased from 11.1% at age 2 to 32.4% at age 6 (Table 6.4). Risk factors for asthma in this population have been previously described (Chapter 5 text, Table 5.5). $PM_{2.5}$ exposure was fairly similar between children with and without asthma (Table 6.4).

There was a dramatic decrease in $PM_{2.5}$ in this region starting in 2008 and lasting until 2010 (Figure 6.1). This drop largely reflects the implementation of flue gas desulphurization ("scrubbers") at power plants near the metropolitan Atlanta area in 2008 and 2009 which dramatically reduced sulfur dioxide (SO_2) emissions, subsequently reducing secondary sulfate (SO_4) production. Other potential causes of this drop in pollution include the economic recession, which impacted energy use, and meteorology (the summer

of 2009 had more rain and cooler temperatures than usual). This change in $PM_{2.5}$ is reflected in the $PM_{2.5}$ exposure among children in this cohort with children born between 2008 and 2010 having the lowest exposure estimates (Table 6.1, Figure 6.2, Figure 6.3). Estimates of $PM_{2.5}$ exposure during the prenatal period exhibit patterns consistent with $PM_{2.5}$ seasonality; children born in the winter months have the highest estimates and children born in the summer months have the lowest estimates (Figure 6.3). As expected, first year of life $PM_{2.5}$ estimates show no seasonality because they are averaged over a full year (Figure 6.3). The Spearman correlation coefficient between prenatal and first year of life $PM_{2.5}$ exposure is 0.59, which declines to 0.22 when controlling for birth year.

Compared to children lost to follow-up, children retained in the cohort have slightly higher first year of life $PM_{2.5}$ exposure estimates and somewhat higher percentages of asthma at every follow-up age (see Table C1 in Appendix C). However, the difference between $PM_{2.5}$ exposure among children with and without asthma is similar between children retained in the cohort and lost to follow-up.

Association between first year of life $PM_{2.5}$ and asthma incidence

Results from unadjusted and adjusted models are presented in Table 6.5 and plotted in Figure 6.4 and show little evidence for an association between first year of life $PM_{2.5}$ and childhood asthma incidence at ages 2 through 6. All risk differences were calculated for a change of 1 $\mu g/m^3$ of $PM_{2.5}$ which is equivalent to moving from the 30^{th} to the 58^{th} percentile of the exposure distribution. In unadjusted models, and models adjusted for factors other than city region, risk differences for an increase of 1 $\mu g/m^3$ of $PM_{2.5}$ range from -0.0162 to 0.0105 (Table 6.5 Models 0-5). When adding individual-level covariates and temporal control to an unadjusted model, risk differences move in the negative direction, with confidence intervals around some negative risk differences excluding the null value of 0. For

example, an increase of 1 μg/m^3 of $PM_{2.5}$ is associated with a 1.6% decrease in the absolute risk of asthma by the fourth birthday (Model 1 RD (95% CI): -0.016 (-0.031, 0.001)) when controlling for sex, race, ethnicity, maternal asthma, cubic splines on date of birth, maternal age, and maternal education. Adding control for neighborhood SES either with major demographic clusters, minor demographic clusters, or census tract variables from the American Community Survey, moves risk differences in the positive direction with minor clusters having the biggest impact on risk difference estimates. Models controlling for individual-level factors, time trends, neighborhood SES, and city region produced positive risk differences for the association between first year of life $PM_{2.5}$ and childhood asthma with all confidence intervals around these estimates including the null value of zero (Table 6.5 Models 6-8). Model results were similar when controlling for city region by dividing Atlanta into three or four regions. A model with three regions was chosen as our final model for simplicity, and to be consistent with models in Chapter 5. This final adjusted model (Model 7) controls for child sex, child race, maternal asthma, cubic splines on date of birth, minor demographic cluster, and city region and shows little evidence for an association between $PM_{2.5}$ exposure in the first year of life and childhood asthma incidence at ages 2 through 6. The inclusion of product terms between city region and race, and metro Atlanta and demographic cluster in the adjusted model did not affect estimates of the association between $PM_{2.5}$ and childhood asthma (see Table C2 in Appendix C).

Final adjusted models show that the strongest risk factors for childhood asthma in this cohort are male sex, black race, and maternal asthma (Table 6.7). A male child of black race whose mother has asthma is at almost a 30% greater absolute risk of developing asthma by age 6 than a female child of white race whose mother does not have asthma (estimated as 8.6% increase from male sex, 5.9% increase from black race, and 15.2% increase from maternal asthma history assuming no interaction between these factors).

There are temporal trends in asthma incidence and $PM_{2.5}$ exposure in our cohort making consideration of temporal confounding important (Figure 6.3, Figure C1 in Appendix C). All of the adjusted models in Table 6.5 control for time trends using cubic splines with one knot per year. This control was used because it yielded the degree of temporal control desired and resulted in less convergence issues for models of first year of life and prenatal exposure than other time control explored (Table 6.6). Figure C2 in Appendix C shows the impact on trend smoothing from models using cubic splines with one and two knots per year. While compared to models including year of birth instead of the cubic splines on date, results were not meaningfully different (Table 6.6), we retained the splines in the final model because we believe they are the most methodologically appropriate way to control for temporal trends in our exposure and outcome.

Results of stratified models and models containing interaction terms with $PM_{2.5}$ show no indication that the association of first year of life $PM_{2.5}$ exposure with childhood asthma varies by child race, sex, maternal asthma status, or the region of metropolitan Atlanta in which a child was born (see Tables C3-C10 in Appendix C). Confidence intervals around almost all adjusted stratified risk differences include the null value of zero. An exception is the strong association observed between $PM_{2.5}$ and childhood asthma at almost every follow-up age among children born more than 10 miles from the I-285 perimeter around metropolitan Atlanta. Among children born in this area, an increase of 1 $\mu g/m^3$ of $PM_{2.5}$ is associated with a 4 percent increase in risk of asthma at follow-up ages 5 and 6 (see Table C9 in Appendix C, RD (95% CI) age 5 0.038 (0.005, 0.071), age 6 0.044 (0.003, 0.084)). However, there is no evidence of statistical interaction between $PM_{2.5}$ and city region (see Table C10 in Appendix C). There is a suggestion from adjusted estimates that the effect is higher among white children than black children, but there is no evidence of statistical interaction (see Table C3 and Table C4 in Appendix C).

Association between prenatal $PM_{2.5}$ and asthma incidence

Neither unadjusted nor adjusted models show evidence of an association between prenatal $PM_{2.5}$ and childhood asthma (Table 6.8, Figure 6.5). Risk differences of the association indicate less than a 1% change in asthma risk associated with an increase of 1 $\mu g/m^3$ of $PM_{2.5}$, with all confidence intervals including the null value of zero. The adjusted model included all covariates from the final adjusted model examining the impact of first year of life $PM_{2.5}$ exposure: child sex, child race, maternal asthma, minor demographic cluster, city region, and temporal control using cubic splines. For a secondary adjusted analysis of the association between prenatal $PM_{2.5}$ and childhood asthma incidence, we excluded covariates that would not be expected to be related to $PM_{2.5}$ exposure during the prenatal window and subsequently could not be confounders. Specifically, we excluded child sex, and controlled for maternal race instead of child race. While both child sex and child race are strong predictors of the outcome, it is not conceivable that either of these variables could be causing prenatal $PM_{2.5}$ exposure. The results of this analysis were comparable to the results of the main analysis presented in Table 6.8. At every follow-up age the risk difference for the association between prenatal $PM_{2.5}$ and childhood asthma were an absolute 0.07% to 0.2% larger than the risk differences from the main adjusted model, with all confidence intervals still including the null value of zero.

To estimate the separate contributions of exposure during the prenatal and first year of life periods, unadjusted and adjusted models containing both exposures were run (Table 6.9). Compared to estimates from adjusted models with prenatal exposure alone (Table 6.8), all estimates of prenatal exposure are smaller when adjusted for first year of life exposure, with confidence intervals from both types of models containing the null. Compared to estimates from adjusted models with first year of life exposure alone (Table

6.8), estimates of first year of life exposure move in the positive direction at ages 2 through 5, with estimates at ages 2 and 5 no longer including the null (RD (95% CI): 0.012 (0.002, 0.023) and 0.027 (0.004, 0.051) respectively).

Additional analyses and sensitivity analyses

In most analyses using an outcome classification of persistent asthma, there is a stronger association between first year of life $PM_{2.5}$ and childhood asthma than from analyses using other outcome definitions (Table 6.10). For example, from our main models there is an estimated 1.8% increase in the absolute risk of asthma by age 6 associated with an increase of 1 $\mu g/m^3$ of $PM_{2.5}$, but using an outcome of persistent asthma, there is a 2.5% increase in absolute risk of persistent asthma (RD (95% CI): 0.018 (-0.011, 0.046), 0.025 (-0.001, 0.051) respectively). Effect sizes were larger when using events at later ages to inform disease classifications at an earlier age. In an analysis restricted to children enrolled until age 5 (n=8,592), the risk difference for the association between first year of life $PM_{2.5}$ and childhood asthma is 0.003 when classifying a case as a child with at least one asthma diagnosis or asthma-related medication dispensing by age 2 (RD (95% CI): 0.003 (-0.013, 0.018)). When completing the same analysis, but further restricting the case group to children with at least one asthma medication or medication dispensing between ages 2 and 5 the risk difference is 0.011 (RD (95% CI): 0.011 (-0.006, 0.028)). These analyses controlled for all covariates in the final model, but used major demographic cluster to control for SES (instead of minor demographic cluster) in order to ameliorate convergence issues.

The estimate of the association between $PM_{2.5}$ and childhood asthma is dependent on the asthma case definition used in analyses. Table 6.11 shows this variation for the association between first year of life $PM_{2.5}$ and cumulative asthma incidence by age 5. Risk

differences from adjusted analyses range from 0.004 to 0.030 with 6 of the 15 estimated confidence intervals excluding the null value of zero. The strongest effect estimates tend to be from less stringent outcome classifications (e.g., 1 asthma or wheeze diagnosis RD (95% CI): 0.026 (0.002, 0.051)), and many of the weakest effect estimates are from the stricter outcome classifications (e.g., 1 asthma diagnosis and 2 asthma-related medication dispensings 1 of which must be a steroid RD (95% CI): 0.007 (-0.010, 0.024)). The effect estimate from the outcome definition used by KAPPA (1 asthma diagnosis and 1 asthma-related medication dispensing) fell in the middle of the estimated risk differences (RD (95% CI): 0.018 (-0.005, 0.041).

For a purely spatial comparison, we completed an analysis using the 2010 250 meter grid annual averages to assign first year of life $PM_{2.5}$ exposure (Table 6.12). The Spearman correlation coefficient between this first year of life $PM_{2.5}$ estimate, and the first year of life exposure estimate from the daily data was 0.28. At ages 2 and 3, results from this adjusted analysis estimate a greater increase in risk than in the main analysis where daily $PM_{2.5}$ data is used to determine $PM_{2.5}$ exposure. For example, by age 3 this analysis shows a 1.5% increase in absolute risk of asthma associated with an increase of 1 $\mu g/m^3$ of $PM_{2.5}$, compared with a 0.7% increase risk in the main analysis (RD (95% CI): 0.015 (0.002, 0.028) compared to 0.007 (-0.006, 0.020)). Associations at ages 4 to 6 are comparable to those from main analyses.

Five of the completed sensitivity analyses excluded different groups of children, in order to assess the impact of using data from different years for exposure assignment, data cleaning decisions, and the impact of missing data (see Table C11 in Appendix C). All analyses reached the same conclusion as main analyses that overall there is little evidence of an association between first year of life total $PM_{2.5}$ exposure and childhood asthma

incidence in these data. These analyses also indicate that the decisions examined did not have a large impact on the estimated risk differences.

To assess whether the association between $PM_{2.5}$ and asthma is different among children enrolled until the later follow-up ages and those who were lost to follow-up or not born early enough to be followed until the later ages, we completed sensitivity analyses of asthma at earlier ages among the cohort of children enrolled Kaiser Permanente Georgia until age 6. Among the 6,628 children enrolled until age 6, the risk difference for the association between first year of life $PM_{2.5}$ and asthma by age 3 was -0.0083 and the risk difference for the association between first year of life $PM_{2.5}$ and asthma by age 4 was 0.0004 (RD (95% CI): -0.0083 (-0.0310, 0.0144) and 0.0004 (-0.0248, 0.0256) respectively). The age 2 analysis would not converge when limiting the sample to children enrolled until age 6. The sensitivity analyses at ages 3 and 4 both yielded smaller risk differences than from our main models, but both confidence intervals included the null value of 0 as in the results from our main analyses.

Alternative measures of association

Results from analyses using the risk ratio, odds ratio, and hazard ratio as the measure of association of interest are presented in Table 6.13. All estimated risk ratios and odds ratios from adjusted models are greater than one but provide no strong evidence for an association between first year of life $PM_{2.5}$ and childhood asthma incidence. From a graphical assessment and extended Cox models we concluded that all variables met the proportional hazards assumption and that no time-dependent variables were necessary in our final Cox proportional hazards models (see Figure B5 in Appendix B and Figure C3 and Table C12 in Appendix C). Estimated hazard ratios are greater than one in adjusted models. Due to the larger sample size in the Cox proportional hazards models (by including all

children in each analysis regardless of the time at which a child was lost to follow-up), the hazard ratio effect estimates are more precise than the estimates from other analyses, with almost all confidence intervals excluding the null value of one.

DISCUSSION

The results of this study provide little evidence for an association between exposure to total $PM_{2.5}$ in the prenatal and first year of life periods and childhood asthma incidence. This conclusion is from results of models separately examining the prenatal and first year of life $PM_{2.5}$ exposure, and also from models containing exposure in both of these developmental periods. Given the high correlation between these exposures (Spearman correlation coefficient = 0.59), there may not be enough exposure variability to reliably estimate each separate association. Results from models with both exposures (Table 6.9) should therefore be interpreted cautiously. Nevertheless, from models containing both exposures the conclusion is the same as from models with only one exposure that there is little support for an association between total $PM_{2.5}$ and childhood asthma in this cohort.

Many of the considerations discussed in Chapter 5 are also relevant to this study. These include the following strengths of this study: use of comprehensive medical record data, residential history data that include information on residential mobility, data on both individual and neighborhood-level socioeconomic status, and estimates of $PM_{2.5}$ exposure in important developmental windows. Limitations of the study discussed previously include potential limited generalizability, incomplete race and familial linkage data, high loss to follow-up rates, no assurance that maternal matches are biologic, no access to information on some potential confounders (e.g. income, early life environment), not capturing daily movement which may impact $PM_{2.5}$ exposure, measurement error in the outcome, and not being able to control for the effects of other pollutants. There is great interest in whether

the association of $PM_{2.5}$ with childhood asthma differs by atopic and non-atopic asthma phenotypes. We were unable to assess this question in our data because we did not have the detailed clinical information (i.e. data on blood IgE levels and skin prick testing) needed to reliably distinguish these two phenotypes.

The RLINE data used for exposure assignment in Chapter 5 were one of the inputs used to create the daily downscaled CMAQ data used for exposure assignment in this chapter. This means that the limitations of RLINE discussed in Chapter 5, such as the inability to account for changes in the roadway structure, hills, and barriers that may impact small scale meteorology, may also impact the $PM_{2.5}$ data used in these analyses. This is the first epidemiologic study to use daily CMAQ data downscaled to a 250 meter spatial resolution. The methods used to create these data are novel and integrated data from multiple sources to best estimate $PM_{2.5}$ at every point in the metropolitan Atlanta area. These data allowed us to estimate $PM_{2.5}$ exposure during key developmental windows taking into account both spatial and temporal variability in exposure. Nevertheless, since this is the first study to use air pollution data created by this method, there may be limitations of our $PM_{2.5}$ exposure estimates of which we are not yet aware. Since $PM_{2.5}$ estimates were not available for all years of our study, we applied data from neighboring years in order to maximize the sample size in our main analyses. Sensitivity analyses excluding children for whom data from different years were used to calculate $PM_{2.5}$ exposure produced similar risk difference estimates as in our main analyses (see Table C11 in Appendix C).

We chose our final model, adjusting for individual-level factors, demographic factors, city region, and temporal trends aiming to error on the side of over-control for potential confounding. When first adding individual-level factors to the model our risk differences moved in the negative direction from our unadjusted results (Table 6.5, Model 1

vs. Model 0). Maternal age, maternal education, and child ethnicity were dropped from the model because their inclusion had minimal impact on the estimate of the association between first year of life $PM_{2.5}$ and childhood asthma incidence (Table 6.5, Model 2 vs. Model 1). Next, we added demographic control in the form of major demographic clusters, minor demographic clusters, or variables from the American Community Survey, each of which moved risk differences in the positive direction (Table 6.5, Models 3-5). The addition of city region had a large impact on risk differences; estimated associations were larger when including this variable (Table 6.5, Models 6-8). We added this regional control due to concerns that children in the same demographic cluster are not exchangeable since the same demographic cluster could include individuals in the most urban and most rural areas of the study region. There may be key differences between a child living in the lowest SES demographic cluster inside the city center and a child living in the same demographic cluster furthest away from the city center. While we consider Model 7 the most appropriate model for our data, since it is conservative in its confounding control, our conclusions would not change if we had used any of the other models in Table 6.5 as our final model. Confidence intervals around almost all effect estimates from these models contain the null value of zero. Our conclusions would also have been the same had we used the risk ratio or odds ratio as our primary measure of association (Table 6.13). The use of hazard ratios resulted in more precise estimates of effect, likely due to the increased power from including more children in each analysis, with many of the confidence intervals around the hazard ratios excluding the null. While Cox proportional hazards models are a natural fit for our data, and benefited from increased power in this analysis, one limitation of them is the use of exact time of asthma diagnosis which is unlikely to be reliable. In our other analyses, less emphasis is placed on the timing of diagnoses by looking at the outcome of asthma by different birthdays.

We considered it important to control for temporal trends in our analyses, due to the patterns of $PM_{2.5}$ in the region during this time period and the changes in asthma incidence by birth year and season in our data (Figure 6.1, Table 5.5, Figure C1 in Appendix C). We chose the risk difference as our primary measure of association because of an inherent interest in additive effects and because it lends itself well to the assessment of additive interaction which many think is more relevant to public health than multiplicative interaction (Greenland S et al. 2008, Vanderweele et al. 2014). However, risk difference models are prone to convergence difficulties and the addition of temporal control to these models resulted in substantial convergence issues that were not easily resolved. These issues resulted in the inability to estimate particular risk differences of interest. For example, we were unable to estimate the association between $PM_{2.5}$ and childhood asthma among white children by age 2 while controlling for all covariates in other adjusted stratified models. In some analyses we were able to estimate a risk difference, but only with less confounding control than desired. For example, in our analysis of asthma by age 2 restricted to children enrolled until age 5, we controlled for neighborhood SES using major demographic cluster instead of minor demographic cluster in order for the model to converge. Our results in Table 6.5 (comparing Model 4 to Model 3) indicate that had we been able to instead include minor demographic cluster in this model, our risk difference may have been larger than the one produced by our analysis. We ran some of our sensitivity analyses using log binomial models (which estimate risk ratios), and while they were more likely to converge in some scenarios, they did not solve all convergence issues. Our method of dropping individual children from certain models in order to achieve model convergence is unconventional, but helped several models converge. We dropped between one and fifty children from certain models making the assumption that a handful of individuals were not driving overall associations between $PM_{2.5}$ and asthma incidence.

In the literature, there are some available methods to aid in convergence for binomial and log binomial models. The use of a robust Poisson model is an easy to implement method that has helped others (Spiegelman et al. 2005), but it did help our models to converge. The COPY method, which involves creating an expanded dataset with copies of observations from the original dataset, was developed for log binomial models and we are unaware of any use of this method in binomial models with an identity link (Deddens et al. 2008). It is possible to obtain risk difference and risk ratio estimates from logistic models, however, only in the case of a dichotomous outcome and exposure (Austin 2010). This is an active area of current research, with recent work proposing that fitting a marginal structural binomial regression model can circumvent current convergence issues in SAS even when the exposure of interest is continuous (Richardson et al. 2015). This dissertation emphasizes the importance of these avenues of research by highlighting the convergence limitations of traditional approaches even with sample sizes greater than 15,000 individuals and a high disease prevalence.

Analyses using the asthma case definitions from Chapter 4 revealed that the estimated measure of association between first year of life $PM_{2.5}$ exposure and asthma incidence by age 5 differed depending on case definition used (Table 6.11). Use of more sensitive asthma case definitions, that are easier to satisfy, generally resulted in larger effect estimates than the more specific asthma case definitions. It is possible that misclassification of asthma cases as non-cases diluted the effect estimate when using the stricter asthma case definitions. Alternatively, the association between first year of life $PM_{2.5}$ and asthma may differ depending asthma phenotype. The stricter asthma case definitions likely identify children with the more severe asthma phenotypes as having asthma, and classify children with milder phenotypes as non-diseased. Perhaps air pollution has no causal effect on the most severe asthma cases and these cases are caused solely by non-environmental factors.

The results of using different asthma case definitions can be combined with the results of using a more persistent asthma case definition (Table 6.10). When using a definition of persistent asthma, aimed at removing children with transient asthma from the outcome group, effect estimates are larger. If there is a truly an association between $PM_{2.5}$ in the first year of life and asthma incidence it would make sense that risk differences increased when reducing misclassification of non-diseased as diseased (persistent asthma analysis) and risk differences decreased when increasing misclassification of diseased as non-diseased (using an outcome definition focusing on severe asthma). The results of these analyses highlight the importance of choosing an outcome definition *a priori.* Completing sensitivity analyses using alternative disease classifications is useful before making conclusions from any study classifying asthma using medical records, and is not currently a common practice in this field.

There are many places at which selection bias could have impacted our results. A large percentage of children were lost to follow-up and not all children were linked to birth certificates or mothers. We completed sensitivity analyses to assess whether loss to follow-up biased our results and whether there was any difference in results when restricting our cohort to children linked to both mothers or birth certificates (see Table C1 and Table C11 in Appendix C). While there were slight differences in children retained in the cohort and lost to follow-up (children retained in the cohort had more asthma and higher $PM_{2.5}$ exposure during the first year of life than children lost to follow-up), our analyses provided no evidence that the association between $PM_{2.5}$ and childhood asthma would have been different had all children been retained in the cohort. Crude differences in first year of life $PM_{2.5}$ exposure by asthma status were comparable between children retained in the cohort and children lost to follow-up, and effect estimates changed little when excluding children not linked to mothers or children not linked to birth certificates from analyses (see Table C1

and Table C11 in Appendix C). Results from analyses examining asthma by ages 3 and 4 only among children enrolled until age 6 reached similar conclusions as our main analyses. Taken together, these analyses provide no evidence that selection bias altered our study conclusions.

Results of this study provide little evidence for an association between total $PM_{2.5}$ and childhood asthma incidence. While it is possible that there is no true association between $PM_{2.5}$ in early life and asthma incidence, there are other potential explanations for our results. Biologically, particulate matter from different sources may have different impacts on the respiratory and immune systems. It is possible that there is a strong association between $PM_{2.5}$ from one source, for example traffic, and childhood asthma, but that the effect is washed out in this study by examining total $PM_{2.5}$. If future source apportionment work determines that the impact of $PM_{2.5}$ is heterogeneous depending on source, it may not be appropriate to combine $PM_{2.5}$ from different sources in a single analysis. Another consideration is that the relevant exposure window may be shorter than an entire pregnancy or the entire first year of life. Different stages of respiratory and immune system development occur during these periods meaning vulnerability to the effect of environmental insults may vary throughout these windows. A recent study by Leon Hsu and colleagues assessed week-specific effects of $PM_{2.5}$ during pregnancy and saw an association between $PM_{2.5}$ exposure during weeks 16-25 and asthma incidence among boys, but saw no associations with $PM_{2.5}$ exposure in other weeks (Leon Hsu et al. 2015). While the relevant exposure window could be shorter than the windows we examined, it could also be longer. Perhaps, it is cumulative exposure to $PM_{2.5}$ throughout pregnancy and early childhood that is important rather than exposure in any one smaller window. This study moves us closer to understanding whether exposure to total $PM_{2.5}$ during pregnancy and the

first year of life is associated with development of asthma in childhood by describing these relationships in a well-defined cohort of children enrolled in Kaiser Permanente Georgia.

Table 6.1. KAPPA cohort characteristics

	Children with first year of life $PM_{2.5}$ data		Children with prenatal $PM_{2.5}$ data [a]	
Characteristic	n (%)	Mean first year of life $PM_{2.5}$	n (%)	Mean prenatal $PM_{2.5}$
Cohort	23,100	14.63	19,951	14.89
Children with siblings in cohort	7,313 (31.7)	14.77	7,123 (35.7)	15.04
Sex				
Female	11,330 (49.1)	14.63	9,748 (48.9)	14.87
Male	11,770 (51.0)	14.63	10,203 (51.1)	14.92
Race/Ethnicity				
Black	7,995 (34.6)	14.71	7,220 (36.2)	14.98
White	9,034 (39.1)	14.48	8,467 (42.4)	14.76
Other [b]	2,771 (12.0)	14.43	2,643 (13.3)	14.84
Unknown Race	3,300 (14.3)	15.05	1,621 (8.1)	15.29
Hispanic Ethnicity	1,839 (8.0)	13.49	1,759 (8.8)	14.05
Maternal Education				
<12th grade	285 (1.2)	14.84	280 (1.4)	15.24
High School/GED	2,605 (11.3)	14.45	2,524 (12.7)	14.72
Some College or more	13,442 (58.2)	14.64	13,113 (65.7)	14.91
Missing	6,768 (29.3)	14.68	4,034 (20.2)	14.93
Maternal Asthma				
Yes	2,488 (10.8)	14.59	2,419 (12.1)	14.88
No	17,998 (77.9)	14.64	17,532 (87.9)	14.90
Missing	2,614 (11.3)	14.65	0	--
Kaiser Permanente Enrollment Duration [c]				
Enrolled until age 2	17,960 (77.8)	14.70	15,631 (78.4)	14.93
Enrolled until age 3	14,251 (61.7)	14.79	12,434 (62.3)	14.98
Enrolled until age 4	10,999 (47.6)	15.00	9,620 (48.2)	15.27
Enrolled until age 5	8,592 (37.2)	15.34	7,521 (37.7)	15.50
Enrolled until age 6	6,629 (28.7)	15.59	5,806 (29.1)	15.47
Birth Year				
2000	2,456 (10.6)	15.03	2,054 (10.3)	15.04
2001	2,369 (10.3)	15.01	1,977 (9.9)	15.00
2002	2,266 (9.8)	15.21	1,946 (9.8)	15.00
2003	2,185 (9.5)	15.92	1,929 (9.7)	15.33
2004	2,138 (9.3)	16.11	1,871 (9.4)	15.94
2005	2,023 (8.8)	16.25	1,741 (8.7)	16.18
2006	2,198 (9.5)	16.14	1,935 (9.7)	16.28
2007	2,216 (9.6)	14.76	1,919 (9.6)	16.21
2008	2,101 (9.1)	11.68	1,835 (9.2)	14.21
2009	1,585 (6.9)	11.14	1,403 (7.0)	11.30
2010	1,563 (6.8)	11.93	1,341 (6.7)	11.41

[a] A subset of children with first year of life $PM_{2.5}$ data [b] Includes Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander, and children identifying with more than one racial group [c] Reduction in sample size across follow-up reflects the shorter follow-up time available for children born in later years of the study (e.g., a child born in 2010 could be at most 3 years old at the time KAPPA follow-up ended in September 2013) as well as HMO enrollment attrition over time.

Table 6.2. Distribution of prenatal and first year of life $PM_{2.5}$ by demographic clusters and city region

	Children with first year of life $PM_{2.5}$ data		Children with prenatal $PM_{2.5}$ data	
Characteristic	n (%)	Mean first year of life $PM_{2.5}$	n (%)	Mean prenatal $PM_{2.5}$
Major Demographic Cluster				
A	14,401 (62.3)	14.50	12,626 (63.3)	14.77
B	2,293 (9.9)	15.06	1,925 (9.7)	15.37
C	1,130 (4.9)	14.36	974 (4.9)	14.66
D	5,272 (22.8)	14.87	4,423 (22.2)	15.09
Minor Demographic Cluster				
A.1	3,065 (13.3)	14.54	2,747 (13.8)	14.79
A.2	2,243 (9.7)	14.75	1,976 (9.9)	15.00
A.3	9,093 (39.4)	14.42	7,903 (39.6)	14.71
B.1	1,080 (4.7)	15.07	896 (4.5)	15.37
B.3/B.4	1,213 (5.3)	15.06	1,029 (5.2)	15.37
C.1/C.2	856 (3.7)	14.15	735 (3.7)	14.47
C.3/C.4	274 (1.2)	15.02	239 (1.2)	15.25
D.1	2,453 (10.6)	14.69	2,083 (10.4)	14.95
D.3	631 (2.7)	15.14	520 (2.6)	15.34
D.4	1,436 (6.2)	14.96	1,204 (6.0)	15.13
D.5	450 (2.0)	15.00	378 (1.9)	15.23
D.6/D.7	302 (1.3)	15.23	238 (1.2)	15.26
City Region				
Metro Atlanta [a]	2,425 (10.5)	15.20	2,030 (10.2)	15.48
≤10 miles from metro Atlanta	9,894 (42.8)	14.91	8,449 (42.4)	15.13
>10 miles from metro Atlanta	10,781 (46.7)	14.25	9,472 (47.5)	14.56

[a] Metro Atlanta defined as inside the I-285 perimeter of Atlanta (Figure B2). Minor demographic clusters with less than 150 individuals were combined with the closest demographic cluster within its major category for analyses (reflected in table).

Table 6.3. $PM_{2.5}$ exposure and census tract household income and poverty by city region

	All Regions	Metro Atlanta	≤10 miles from metro Atlanta	>10 miles from metro Atlanta
n (row %)				
All Races	23,100	2,425 (10.5)	9,894 (42.8)	10,781 (46.7)
Black	7,995	884 (11.1)	4,901 (61.3)	2,210 (27.6)
White	9,034	851 (9.4)	2,204 (24.4)	5,979 (66.2)
Other Race [a]	2,771	275 (9.9)	1,221 (44.1)	1,275 (46.0)
Unknown Race	3,300	415 (12.6)	1,568 (47.5)	1,317 (39.9)
Mean first year of life $PM_{2.5}$				
All Children	14.63	15.20	14.91	14.25
Black	14.71	15.18	14.85	14.22
White	14.48	15.06	14.95	14.22
Other Race [a]	14.43	15.11	14.72	14.00
Unknown	15.05	15.59	15.21	14.69
Spearman Correlation with first year of life $PM_{2.5}$				
Median household income [b]	ρ = -0.075	ρ = 0.018	ρ = -0.019	ρ = 0.032
Percent families poverty [b]	ρ = 0.057	ρ = -0.041	ρ = 0.023	ρ = -0.029

[a] Includes Asian, American Indian, Alaska Native, Native Hawaiian or other Pacific Islander, and children identifying with more than one racial group. [b] American Community Survey 2010 data (5 year estimates) at census tract level, prepared by Social Explorer

Table 6.4. Asthma incidence and prenatal and first year of life $PM_{2.5}$ exposure by each follow-up age

	Children with first year of life $PM_{2.5}$ data; First year of life $PM_{2.5}$			Children with prenatal $PM_{2.5}$ data; Prenatal $PM_{2.5}$		
Cohort	Asthma n (%)	Mean $PM_{2.5}$ among cases	Mean $PM_{2.5}$ among non-cases	Asthma n (%)	Mean $PM_{2.5}$ among cases	Mean $PM_{2.5}$ among non-cases
Age 2	1,994 (11.1)	14.74	14.70	1,731 (11.1)	14.95	14.92
Age 3	2,627 (18.4)	14.82	14.79	2,316 (18.6)	14.98	14.98
Age 4	2,650 (24.1)	15.03	14.99	2,309 (24.0)	15.27	15.27
Age 5	2,465 (28.7)	15.33	15.35	2,132 (28.4)	15.45	15.51
Age 6	2,149 (32.4)	15.57	15.60	1,854 (31.9)	15.45	15.48

Asthma incidence (at least one asthma diagnosis (ICD-9 493.XX) and one asthma-related medication dispensing after the first year of life) calculated among children enrolled until each follow-up age (see Table 6.1 for number enrolled until each age). For example, the age 5 cohort examines asthma incidence by the 5th birthday among children enrolled until at least their 5th birthday.

Table 6.5. Risk differences for first year of life $PM_{2.5}$ and childhood asthma incidence in models adjusting for different sets of covariates

	Cohort	MODEL 0		MODEL 1		MODEL 2	
		RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Unadjusted models & models with individual-level covariates	Age 2	0.0015 (-0.0010, 0.0039)	0.24	-0.0058 (-0.0135, 0.0020)	0.14	-0.0069 (-0.0142, 0.0004)	0.06
	Age 3	0.0015 (-0.0019, 0.0050)	0.38	-0.0124 (-0.0237, -0.0011)	0.03	-0.0126 (-0.0231, -0.0020)	0.02
	Age 4	0.0031 (-0.0016, 0.0078)	0.20	-0.0162 (-0.0314, -0.0010)	0.04	-0.0159 (-0.0302, -0.0017)	0.03
	Age 5	-0.0034 (-0.0118, 0.0050)	0.43	-0.0056 (-0.0243, 0.0131)	0.56	-0.0065 (-0.0241, 0.0112)	0.47
	Age 6	-0.0106 (-0.0258, 0.0047)	0.18	-0.0071 (-0.0307, 0.0165)	0.55	-0.0087 (-0.0309, 0.0135)	0.44
	Cohort	MODEL 3		MODEL 4		MODEL 5	
		RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Models with individual-level & demographic covariates	Age 2	-0.0018 (-0.0099, 0.0063)	0.66	0.0012 (-0.0071, 0.0095)	0.78	0.0004 (-0.0086, 0.0094)[a]	0.93
	Age 3	-0.0035 (-0.0154, 0.0084)	0.57	-0.0019 (-0.0137, 0.0100)	0.76	-0.0017 (-0.0144, 0.0109)	0.79
	Age 4	-0.0087 (-0.0245, 0.0072)	0.28	-0.0054 (-0.0218, 0.0111)	0.52	-0.0026 (-0.0196, 0.0143)	0.76
	Age 5	0.0032 (-0.0164, 0.0228)	0.75	0.0035 (-0.0170, 0.0239)	0.74	0.0076 (-0.0132, 0.0283)	0.47
	Age 6	0.0022 (-0.0221, 0.0266)	0.86	0.0018 (-0.0236, 0.0272)	0.89	0.0105 (-0.0145, 0.0355)	0.41
	Cohort	MODEL 6		MODEL 7: **FINAL MODEL**		MODEL 8	
		RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Models with individual-level, demographic & spatial covariates	Age 2	0.0081 (-0.0016, 0.0178)	0.10	0.0076 (-0.0019, 0.0170)	0.12	0.0073 (-0.0023, 0.0169)	0.13
	Age 3	0.0073 (-0.0059, 0.0205)	0.28	0.0066 (-0.0064, 0.0196)	0.32	0.0056 (-0.0081, 0.0193)	0.42
	Age 4	0.0049 (-0.0138, 0.0236)	0.61	0.0040 (-0.0143, 0.0223)	0.67	0.0033 (-0.0149, 0.0215)	0.73
	Age 5	0.0194 (-0.0037, 0.0425)	0.10	0.0178 (-0.0049, 0.0405)	0.12	0.0175 (-0.0049, 0.0398)	0.13
	Age 6	0.0205 (-0.0081, 0.0491)	0.16	0.0176 (-0.0106, 0.0457)	0.22	0.0197 (-0.0078, 0.0472)	0.16

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] deleted 1 child in order for model to converge

List of Model Covariates

Model 0: unadjusted model

Model 1: child sex, child race, child ethnicity, maternal asthma, cubic splines on date of birth (1 knot per year in May), maternal age, maternal education

Model 2: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May)

Model 3: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), major demographic cluster

Model 4: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), minor demographic cluster

Model 5: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), median household income, median year structure built, median house value, percent less than high school, percent families in poverty

Model 6: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), minor demographic cluster, city region (metro Atlanta, ≤5 miles from metro Atlanta, >5–10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 7 (**FINAL MODEL**): child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 8: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), median household income, median year structure built, median house value, percent less than high school, percent families in poverty, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 6.6. Risk differences for prenatal and first year of life $PM_{2.5}$ and childhood asthma incidence, assessing impact of temporal control

First year of life $PM_{2.5}$						
Cohort	MODEL A		MODEL B		MODEL C	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0069 (0.0005, 0.0133)	0.03	0.0076 (-0.0019, 0.0170)	0.12	0.0061 (-0.0039, 0.0160)	0.23
Age 3	0.0053 (-0.0029, 0.0135)	0.21	0.0066 (-0.0064, 0.0196)	0.32	0.0041 (-0.0089, 0.0171)	0.53
Age 4	0.0101 (-0.0022, 0.0223)	0.11	0.0040 (-0.0143, 0.0223)	0.67	0.0007 (-0.0186, 0.0200)	0.94
Age 5	0.0153 (-0.0010, 0.0316)	0.07	0.0178 (-0.0049, 0.0405)	0.12	0.0185 (-0.0046, 0.0416)	0.12
Age 6	0.0181 (-0.0069, 0.0432)	0.16	0.0176 (-0.0106, 0.0457)	0.22	0.0181 (-0.0101, 0.0464)	0.21
Prenatal $PM_{2.5}$						
Cohort	MODEL A[a]		MODEL B		MODEL C	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0015 (-0.0021, 0.0052)	0.41	0.0006 (-0.0036, 0.0047)[b]	0.78	0.0031 (-0.0048, 0.0110)[b]	0.44
Age 3	0.0011 (-0.0041, 0.0062)	0.69	0.0002 (-0.0056, 0.0060)	0.94	—	—
Age 4	0.0001 (-0.0065, 0.0068)	0.97	-0.0011 (-0.0085, 0.0062)	0.76	0.0010 (-0.0136, 0.0156)	0.89
Age 5	-0.0029 (-0.0112, 0.0053)	0.49	-0.0029 (-0.0117, 0.0059)	0.52	—	—
Age 6	0.0028 (-0.0081, 0.0136)	0.62	0.0036 (-0.0072, 0.0144)	0.51	0.0163 (-0.0061, 0.0387)[b]	0.15

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, [a] Combined conception year categories for years with less than 5 children in order for models to converge [b] <10 children deleted in order for model to converge

Model A: child sex, child race, maternal asthma, minor demographic cluster, city region, year *[birth year for first year of life exposure, conception year for prenatal exposure]*

Model B: child sex, child race, maternal asthma, minor demographic cluster, city region, cubic splines on date with 1 knot per year in May *[date of birth for first year of life exposure, date of conception for prenatal exposure]*

Model C: child sex, child race, maternal asthma, minor demographic cluster, city region, cubic splines on date with 2 knots per year in April and October *[date of birth for first year of life exposure, date of conception for prenatal exposure]*

Table 6.7. Full adjusted model results for the association between first year of life $PM_{2.5}$ and asthma incidence at ages 2, 4, and 6

	Age 2 n = 17,958 Correlation = 0.16		Age 4 n = 10,998 Correlation = 0.20		Age 6 n = 6,628 Correlation = 0.24	
Parameter	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Intercept	-0.0486 (-0.2069, 0.1098)	0.55	0.2266 (-0.0735, 0.5267)	0.14	0.2079 (-0.2338, 0.6497)	0.36
$PM_{2.5}$	0.0076 (-0.0019, 0.0170)	0.12	0.0040 (-0.0143, 0.0223)	0.67	0.0176 (-0.0106, 0.0457)	0.22
Male sex	0.0437 (0.0350, 0.0525)	<0.01	0.0781 (0.0625, 0.0937)	<0.01	0.0862 (0.0641, 0.1082)	<0.01
Black race	0.0228 (0.0106, 0.0350)	<0.01	0.0283 (0.0075, 0.0492)	<0.01	0.0594 (0.0300, 0.0888)	<0.01
Unknown/other race	0.0028 (-0.0086, 0.0143)	0.63	0.0014 (-0.0202, 0.0229)	0.90	0.0432 (0.0103, 0.0762)	0.01
Maternal asthma	0.0615 (0.0449, 0.0780)	<0.01	0.1040 (0.0780, 0.1300)	<0.01	0.1521 (0.1168, 0.1874)	<0.01
Metro Atlanta	-0.0358 (-0.0556, -0.0160)	<0.01	-0.0492 (-0.0861, -0.0122)	<0.01	-0.0799 (-0.1347, -0.0251)	<0.01
≤10 mi from metro Atlanta	-0.0156 (-0.0279, -0.0032)	0.01	-0.0179 (-0.0393, 0.0034)	0.10	-0.0276 (-0.0581, 0.0028)	0.08
Demographic cluster A.2	0.0095 (-0.0089, 0.0280)	0.31	0.0301 (-0.0029, 0.0631)	0.07	-0.0091 (-0.0565, 0.0384)	0.71
Cluster A.3	0.0157 (0.0012, 0.0301)	0.03	0.0373 (0.0117, 0.0628)	<0.01	0.0000 (-0.0376, 0.0376)	1.00
Cluster B.1	-0.0182 (-0.0389, 0.0026)	0.09	-0.0422 (-0.0829, -0.0015)	0.04	-0.0624 (-0.1268, 0.0020)	0.06
Cluster B.3/B.4	0.0203 (-0.0044, 0.0450)	0.11	0.0588 (0.0132, 0.1044)	0.01	0.0204 (-0.0441, 0.0849)	0.54
Cluster C.1/C.2	0.0164 (-0.0113, 0.0442)	0.25	0.0518 (-0.0016, 0.1053)	0.06	-0.0256 (-0.0965, 0.0454)	0.48
Cluster C.3/C.4	0.0165 (-0.0245, 0.0575)	0.43	0.1220 (0.0427, 0.2013)	<0.01	0.1418 (0.0293, 0.2544)	0.01
Cluster D.1	0.0110 (-0.0082, 0.0302)	0.26	0.0402 (0.0064, 0.0739)	0.02	-0.0027 (-0.0516, 0.0463)	0.92
Cluster D.3	0.0110 (-0.0184, 0.0405)	0.46	0.0503 (-0.0044, 0.1050)	0.07	0.0308 (-0.0482, 0.1098)	0.44
Cluster D.4	0.0046 (-0.0173, 0.0265)	0.68	0.0315 (-0.0091, 0.0721)	0.13	-0.0128 (-0.0714, 0.0458)	0.67
Cluster D.5	0.0333 (-0.0061, 0.0727)	0.10	0.0244 (-0.0406, 0.0894)	0.46	-0.0172 (-0.1017, 0.0673)	0.69
Cluster D.6/D.7	0.0051 (-0.0381, 0.0484)	0.82	0.0451 (-0.0256, 0.1158)	0.21	0.0794 (-0.0226, 0.1813)	0.13

RD = Risk Difference, CI = Confidence Interval, p = p-value, mi = miles. Models also include cubic splines on date of birth with 1 knot per year in May (results not included in table). Reference group: female sex, white race, no or unknown maternal asthma, >10 miles from metro Atlanta, demographic cluster A.1

Table 6.8. Risk differences for prenatal and first year of life $PM_{2.5}$ and asthma incidence, unadjusted and final adjusted models

Cohort (adjusted model n)	First Year of Life $PM_{2.5}$ Unadjusted Model		First Year of Life $PM_{2.5}$ Final Adjusted Model	
	RD (95% CI)	p	RD (95% CI)	p
Age 2 (n=17,958)	0.0015 (-0.0010, 0.0039)	0.24	0.0076 (-0.0019, 0.0170)	0.12
Age 3 (n=14,249)	0.0015 (-0.0019, 0.0050)	0.38	0.0066 (-0.0064, 0.0196)	0.32
Age 4 (n=10,998)	0.0031 (-0.0016, 0.0078)	0.20	0.0040 (-0.0143, 0.0223)	0.67
Age 5 (n=8,591)	-0.0034 (-0.0118, 0.0050)	0.43	0.0178 (-0.0049, 0.0405)	0.12
Age 6 (n=6,628)	-0.0106 (-0.0258, 0.0047)	0.18	0.0176 (-0.0106, 0.0457)	0.22
Cohort (adjusted model n)	Prenatal $PM_{2.5}$ Unadjusted Model		Prenatal $PM_{2.5}$ Final Adjusted Model	
	RD (95% CI)	p	RD (95% CI)	p
Age 2 (n=15,622)	0.0006 (-0.0018, 0.0031)	0.62	0.0006 (-0.0036, 0.0047)[a]	0.78
Age 3 (n=12,432)	-0.0000 (-0.0036, 0.0035)	0.98	0.0002 (-0.0056, 0.0060)	0.94
Age 4 (n=9,619)	0.0006 (-0.0047, 0.0059)	0.82	-0.0011 (-0.0085, 0.0062)	0.76
Age 5 (n=7,520)	-0.0059 (-0.0134, 0.0016)	0.12	-0.0029 (-0.0117, 0.0059)	0.52
Age 6 (n=5,805)	-0.0041 (-0.0137, 0.0054)	0.40	0.0036 (-0.0072, 0.0144)	0.51

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] 7 children deleted in order for model to converge

Final adjusted models control for child sex, child race, maternal asthma, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta), and date of birth/conception (using cubic splines with 1 knot per year; date of birth used in first year of life exposure model and date of conception used in prenatal exposure model)

Table 6.9. Risk differences for prenatal and first year of life $PM_{2.5}$ and asthma incidence, from unadjusted and adjusted models containing both exposure windows

Cohort	UNADJUSTED			
	Prenatal $PM_{2.5}$		First year of life $PM_{2.5}$	
	RD (95% CI)	P	RD (95% CI)	p
Age 2	0.0001 (-0.0032, 0.0034)	0.94	0.0008 (-0.0027, 0.0043)	0.64
Age 3	-0.0004 (-0.0051, 0.0043)	0.86	0.0006 (-0.0044, 0.0056)	0.80
Age 4	-0.0008 (-0.0071, 0.0056)	0.81	0.0025 (-0.0036, 0.0086)	0.42
Age 5	-0.0053 (-0.0132, 0.0026)	0.19	-0.0023 (-0.0118, 0.0072)	0.64
Age 6	-0.0016 (-0.0124, 0.0092)	0.77	-0.0096 (-0.0280, 0.0087)	0.30
Cohort	ADJUSTED			
	Prenatal $PM_{2.5}$		First year of life $PM_{2.5}$	
	RD (95% CI)	P	RD (95% CI)	p
Age 2	0.0002 (-0.0040, 0.0044)[a]	0.92	0.0124 (0.0021, 0.0227)[a]	0.02
Age 3	-0.0001 (-0.0059, 0.0057)	0.97	0.0099 (-0.0045, 0.0244)	0.18
Age 4	-0.0013 (-0.0086, 0.0061)	0.73	0.0055 (-0.0142, 0.0252)	0.59
Age 5	-0.0041 (-0.0129, 0.0048)	0.37	0.0273 (0.0035, 0.0512)	0.02
Age 6	0.0024 (-0.0087, 0.0134)	0.68	0.0173 (-0.0125, 0.0472)	0.26

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a]10 children deleted in order for model to converge

Adjusted models control for child sex, child race, maternal asthma, cubic splines on date of conception with 1 knot per year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 6.10. Risk differences for first year of life $PM_{2.5}$ and persistent asthma

Cohort	Children with persistent asthma [n (%)]	Including children with incident, but not persistent asthma as non-cases		Excluding children with incident, but not persistent asthma	
		RD (95% CI)	p	RD (95% CI)	p
Age 2	1,994 (11.1)	0.0076 (-0.0019, 0.0170)[a]	0.12	0.0076 (-0.0019, 0.0170)[a]	0.12
Age 3	2,196 (15.4)	0.0072 (-0.0045, 0.0189)[b]	0.23	0.0076 (-0.0038, 0.0191)[b]	0.19
Age 4	1,965 (17.9)	0.0020 (-0.0139, 0.0179)	0.81	0.0030 (-0.0138, 0.0198)	0.73
Age 5	1,629 (19.0)	0.0184 (-0.0008, 0.0375)	0.06	0.0206 (-0.0002, 0.0414)	0.05
Age 6	1,350 (20.4)	0.0232 (-0.0004, 0.0469)	0.05	0.0248 (-0.0014, 0.0510)	0.06

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] Identical analysis to incident asthma at age 2, [b]≤20 children deleted in order for model to converge

Persistent asthma defined as a child meeting the incident asthma classification (at least 1 asthma diagnosis (ICD-9 493.XX) and 1 asthma-related medication dispensing) with evidence of asthma in the past year (at least 1 asthma diagnosis or 1 asthma-related medication dispensing)

Models adjust for child sex, child race, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 6.11. First year of life $PM_{2.5}$ and incident asthma by age 5 among children enrolled through age 5 (n=8,592), comparing different outcome definitions

Outcome Definition	n(%) meeting definition	UNADJUSTED		ADJUSTED	
		RD (95% CI)	p	RD (95% CI)	p
1. 1 asthma or wheeze diagnosis	3,082 (35.9)	-0.0023 (-0.0113, 0.0068)	0.62	0.0261 (0.0016, 0.0505)	0.04
2. 1 asthma diagnosis	2,570 (29.9)	-0.0043 (-0.0129, 0.0043)	0.33	0.0235 (0.0004, 0.0466)	0.05
3. 2 asthma diagnoses	1,757 (20.5)	-0.0060 (-0.0137, 0.0017)	0.12	0.0070 (-0.0124, 0.0264)	0.48
4. 3 asthma diagnoses	1,311 (15.3)	-0.0032 (-0.0101, 0.0038)	0.37	0.0044 (-0.0132, 0.0220)	0.63
6. 2 asthma diagnoses OR 1 acute asthma diagnosis	1,856 (21.6)	-0.0063 (-0.0142, 0.0015)	0.11	0.0092 (-0.0108, 0.0292)	0.37
7. 1 asthma diagnosis OR 2 medication dispensings	3,322 (38.7)	-0.0059 (-0.0150, 0.0032)	0.20	0.0301 (0.0054, 0.0548)	0.02
8. 1 asthma diagnosis AND 1 medication dispensing *(KAPPA study definition)*	2,465 (28.7)	-0.0034 (-0.0118, 0.0050)	0.43	0.0178 (-0.0049, 0.0405)	0.12
9. 1 asthma diagnosis AND 2 medication dispensings	2,168 (25.2)	-0.0043 (-0.0124, 0.0038)	0.30	0.0223 (0.0013, 0.0433)	0.04
10. 1 asthma diagnosis OR 2 medication dispensings 1 of which must be a steroid	2,685 (31.3)	-0.0041 (-0.0127, 0.0046)	0.36	0.0279 (0.0045, 0.0512)	0.02
11. 1 asthma diagnosis AND 2 medication dispensings 1 of which must be a steroid	1,388 (16.2)	0.0099 (0.0038, 0.0161)	<0.01	0.0073 (-0.0096, 0.0242)[a]	0.40
12. 1 asthma diagnosis OR 1 controller dispensing	2,715 (31.6)	-0.0046 (-0.0132, 0.0041)	0.30	0.0234 (-0.0001, 0.0469)	0.05
13. 1 asthma diagnosis AND 1 controller dispensing	1,434 (16.7)	0.0080 (0.0017, 0.0143)	0.01	0.0073 (-0.0098, 0.0244)[a]	0.40
14. 1 asthma diagnosis AND (2 reliever dispensings OR 1 controller dispensing)	2,181 (25.4)	-0.0044 (-0.0125, 0.0037)	0.28	0.0222 (0.0012, 0.0433)	0.04
15. Any of the following: a) 1 asthma diagnosis AND 1 medication dispensing in the same year, b) 1 asthma-related ED visit or hospitalization, c) 3 asthma diagnoses	2,450 (28.5)	-0.0030 (-0.0115, 0.0054)	0.48	0.0184 (-0.0044, 0.0412)	0.11

[a]<15 children deleted in order for model to converge. These are the minimum required events for each case definition. Definition numbers align with numbers from Dissertation Chapter 4 (excluded definition 5 (atopic asthma)). Only 1 diagnosis per day counted. ED = emergency department; Asthma diagnosis = ICD-9 code 493.XX; Wheeze diagnosis = ICD-9 code 786.07; Acute asthma diagnosis = a) emergency department or inpatient asthma diagnosis *or* b) asthma diagnosis with status asthmaticus or acute exacerbation (ICD-9 codes 493.01, 493.02, 493.11, 493.12, 493.21, 493.22, 493.91, 493.92); Asthma controller (underlined medications contain a steroid) = Aminophylline, beclomethasone diproprionate, budesonide, budesonide/formoterol fumarate, cromolyn sodium, fluticasone propionate, fluticasone/sameterol, mometasone furoate, montelukast sodium, salmeterol xinafoate, theophylline anhydrous, tiotropium bromide, triamcinolone acetonide; Asthma reliever = Albuterol, albuterol sulfate, ipratropium bromide, ipratropium/albuterol sulfate, levalbuterol, metaproterenol sulfate; Medication dispensing = dispensing of any asthma controller or reliever. Adjusted models control for child sex, child race, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 6.12. Risk differences for first year of life $PM_{2.5}$ and childhood asthma incidence, assigning first year of life $PM_{2.5}$ using 2010 annual averages

	Unadjusted Model		Adjusted Model	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	-0.0057 (-0.0113, -0.0001)	0.05	0.0115 (0.0024, 0.0207)[a]	0.01
Age 3	-0.0068 (-0.0147, 0.0011)	0.09	0.0152 (0.0022, 0.0282)[a]	0.02
Age 4	-0.0071 (-0.0172, 0.0030)	0.17	0.0072 (-0.0098, 0.0242)	0.41
Age 5	-0.0011 (-0.0134, 0.0112)	0.86	0.0174 (-0.0034, 0.0383)	0.10
Age 6	-0.0042 (-0.0187, 0.0104)	0.57	0.0105 (-0.0145, 0.0356)	0.41

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] deleted 1 child from cohort in order for model to converge

Adjusted models control for child sex, child race, maternal asthma, birth year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Table 6.13. First year of life $PM_{2.5}$ and incident asthma, comparing different measures of association

	UNADJUSTED		ADJUSTED	
Cohort	RISK DIFFERENCES			
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0015 (-0.0010, 0.0039)	0.24	0.0076 (-0.0019, 0.0170)	0.12
Age 3	0.0015 (-0.0019, 0.0050)	0.38	0.0066 (-0.0064, 0.0196)	0.32
Age 4	0.0031 (-0.0016, 0.0078)	0.20	0.0040 (-0.0143, 0.0223)	0.67
Age 5	-0.0034 (-0.0118, 0.0050)	0.43	0.0178 (-0.0049, 0.0405)	0.12
Age 6	-0.0106 (-0.0258, 0.0047)	0.18	0.0176 (-0.0106, 0.0457)	0.22
Cohort	RISK RATIOS			
	RR (95% CI)	p	RR (95% CI)	p
Age 2	1.01 (0.99, 1.04)	0.26	1.12 (1.02, 1.23)	0.02
Age 3	1.01 (0.99, 1.03)	0.40	1.06 (0.98, 1.14)	0.16
Age 4	1.01 (0.99, 1.03)	0.23	1.04 (0.96, 1.12)	0.32
Age 5	0.99 (0.96, 1.02)	0.44	1.06 (0.98, 1.15)	0.13
Age 6	0.97 (0.93, 1.02)	0.19	1.03 (0.94, 1.12)	0.50
Cohort	ODDS RATIOS			
	OR (95% CI)	p	OR (95% CI)	p
Age 2	1.01 (0.99, 1.04)	0.25	1.13 (1.02, 1.26)	0.02
Age 3	1.01 (0.99, 1.03)	0.40	1.07 (0.97, 1.18)	0.17
Age 4	1.02 (0.99, 1.04)	0.22	1.05 (0.94, 1.17)	0.37
Age 5	0.98 (0.94, 1.02)	0.43	1.10 (0.98, 1.23)	0.12
Age 6	0.95 (0.89, 1.02)	0.18	1.07 (0.93, 1.22)	0.33
Cohort	HAZARD RATIOS [a]			
	HR (95% CI)	p	HR (95% CI)	p
Age 2	1.02 (0.99, 1.04)	0.15	1.11 (1.01, 1.21)	0.03
Age 3	1.01 (0.99, 1.03)	0.19	1.07 (0.99, 1.16)	0.07
Age 4	1.02 (1.00, 1.03)	0.09	1.07 (1.00, 1.15)	0.05
Age 5	1.02 (1.00, 1.03)	0.07	1.09 (1.02, 1.17)	0.01
Age 6	1.02 (1.00, 1.04)	0.04	1.09 (1.02, 1.17)	0.01

RD = Risk Difference, RR = Risk Ratio, OR = Odds Ratio, HR = Hazard Ratio, all calculated for 1 $\mu g/m^3$
CI = Confidence Interval, p = p-value, [a] Hazard ratios calculated using the 22,987 children in the KAPPA cohort with first year of life $PM_{2.5}$ estimates enrolled in Kaiser Permanente Georgia until at least their first birthday. The outcome of interest was asthma incidence between the first birthday and the birthday of each age cohort. For example, in the age 4 analysis the outcome of interest is asthma between the first and fourth birthdays.

Adjusted models control for child sex, child race, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Figure 6.1. Modeled $PM_{2.5}$ (µg/m³) from all sources for 2002-2010

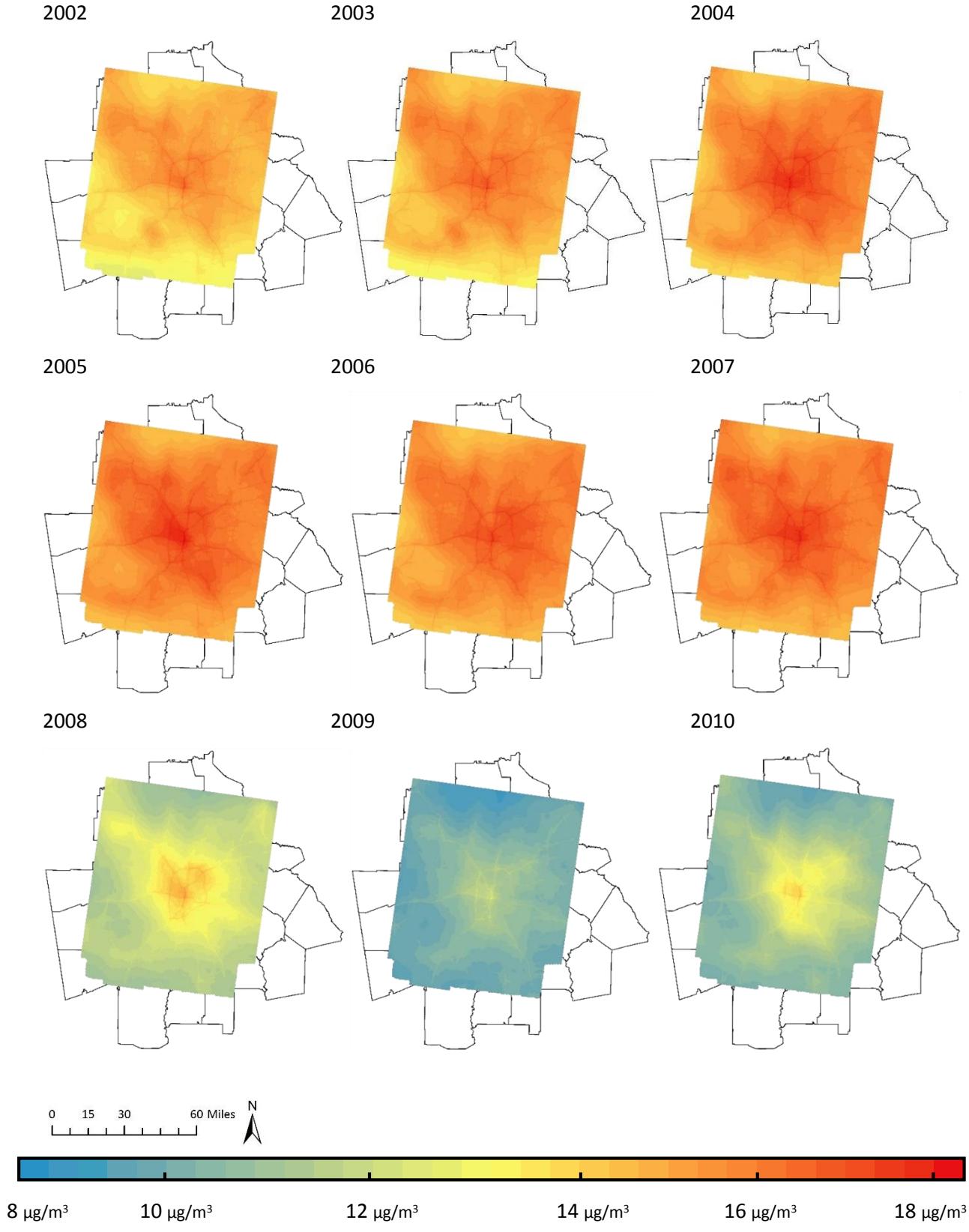
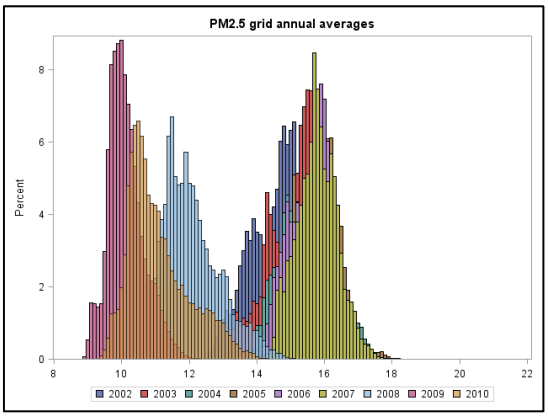
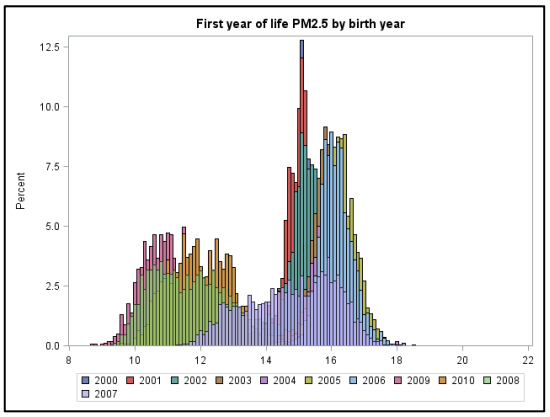


Figure 6.2. A) $PM_{2.5}$ grid annual averages by year B) Estimated first year of life $PM_{2.5}$ by birth year C) Estimated prenatal $PM_{2.5}$ by birth year D) Estimated prenatal $PM_{2.5}$ by year of conception

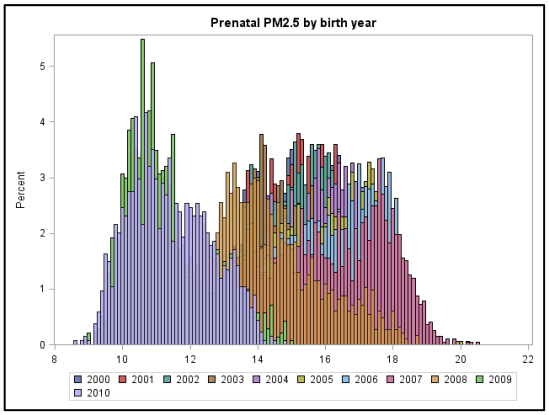
A.



B.



C.



D.

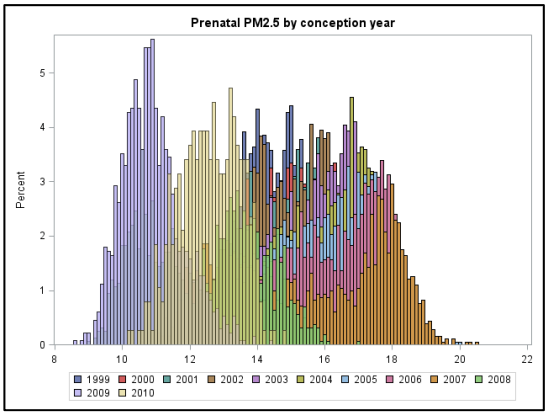


Figure 6.3. Prenatal and first year of life $PM_{2.5}$ exposure by month and year of birth

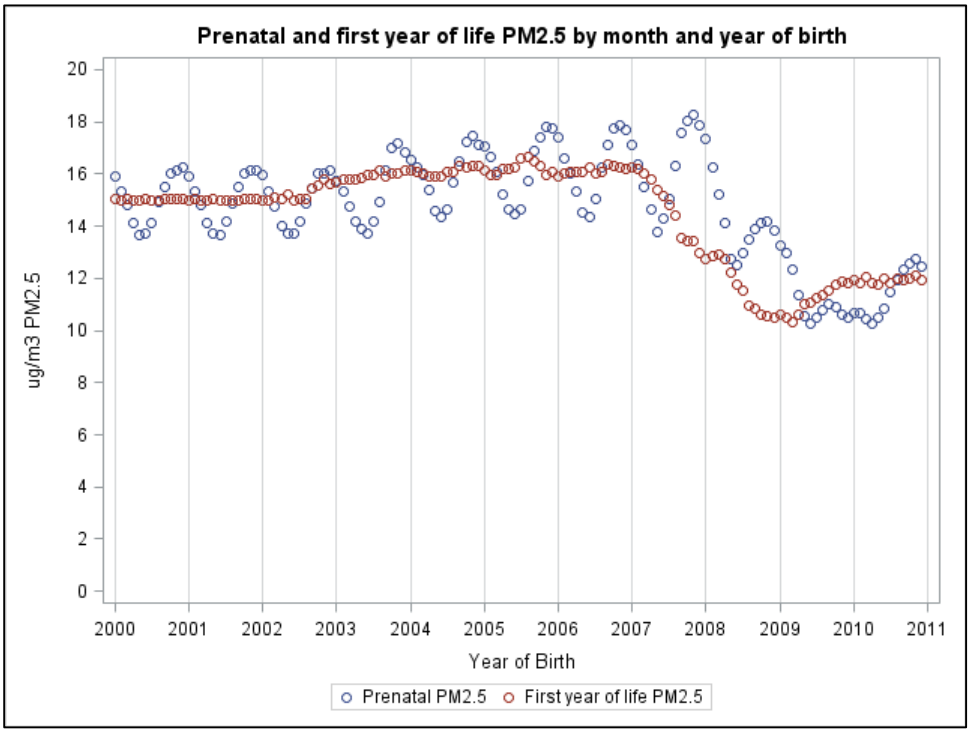
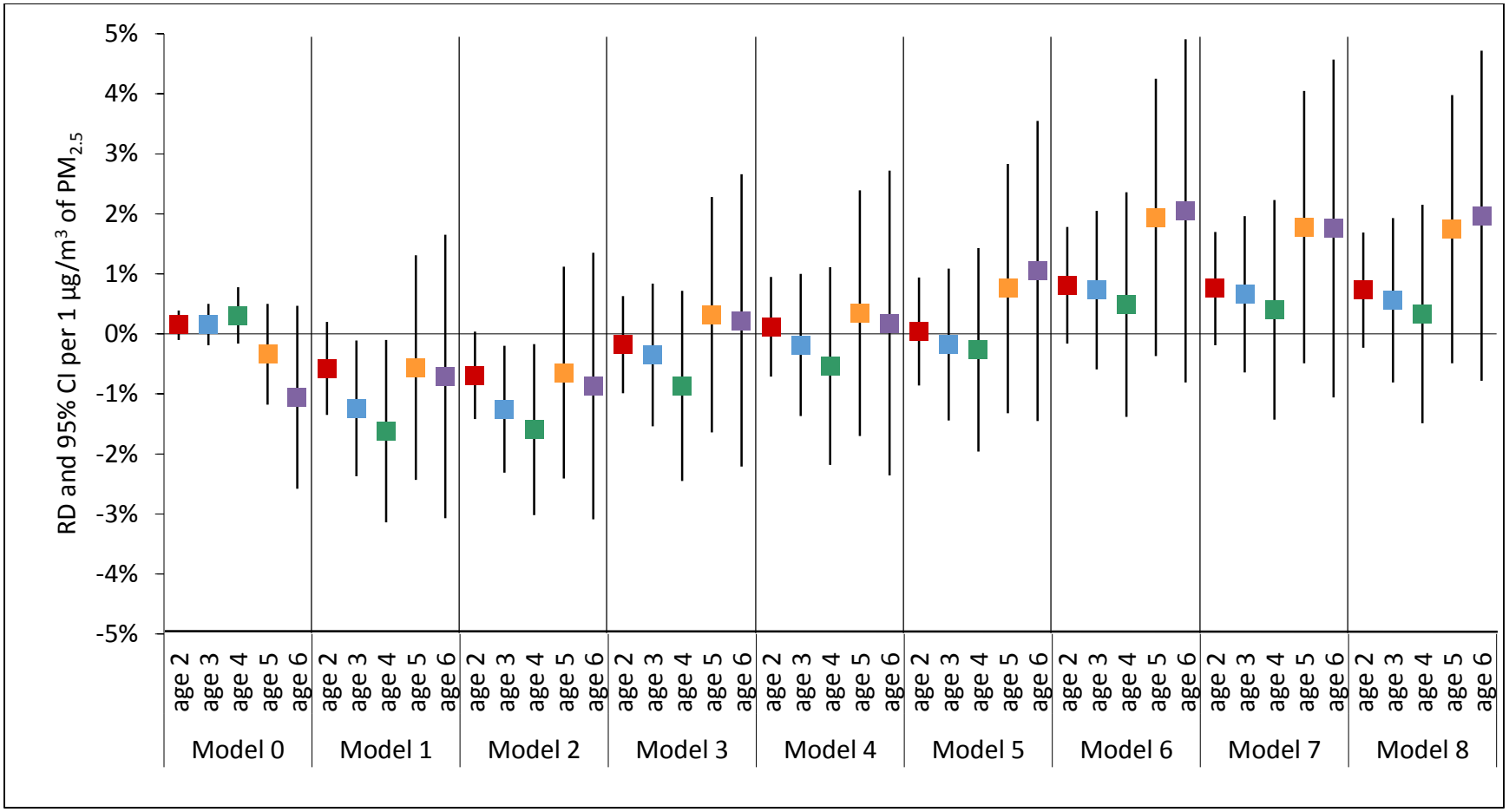


Figure 6.4. Risk differences for first year of life $PM_{2.5}$ and childhood asthma incidence in models adjusting for different sets of covariates



RD = Risk Difference, CI = Confidence Interval

List of Model Covariates

Model 0: unadjusted model

Model 1: child sex, child race, child ethnicity, maternal asthma, cubic splines on date of birth (1 knot per year in May), maternal age, maternal education

Model 2: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May)

Model 3: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), major demographic cluster

Model 4: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), minor demographic cluster

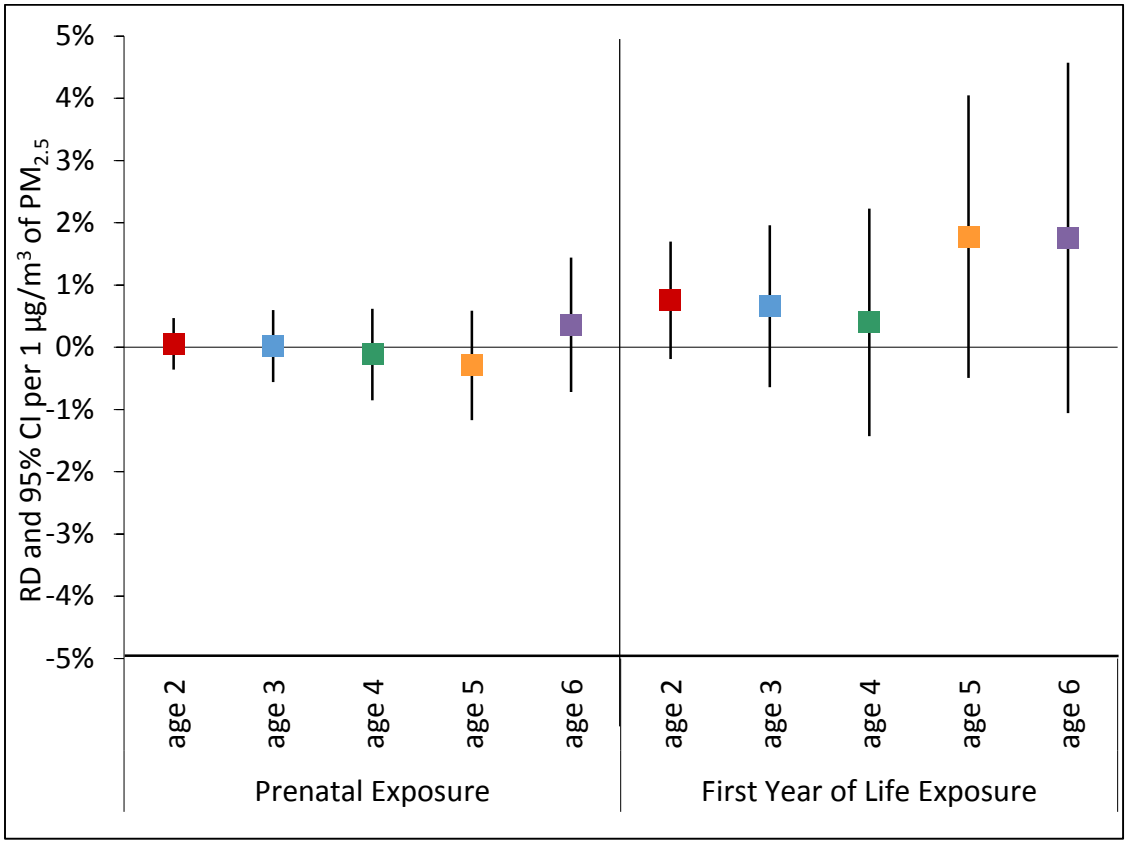
Model 5: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), median household income, median year structure built, median house value, percent less than high school, percent families in poverty

Model 6: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), minor demographic cluster, city region (metro Atlanta, ≤5 miles from metro Atlanta, >5–10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 7 (**FINAL MODEL**): child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Model 8: child sex, child race, maternal asthma, cubic splines on date of birth (1 knot per year in May), median household income, median year structure built, median house value, percent less than high school, percent families in poverty, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

Figure 6.5. Adjusted risk differences for prenatal and first year of life $PM_{2.5}$ and asthma incidence from final models



RD = Risk Difference, CI = Confidence Interval. Models control for child sex, child race, maternal asthma, cubic splines on date of conception with 1 knot per year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta)

CHAPTER 7

Conclusion

In this dissertation, the impact of prenatal and first year of life exposure to $PM_{2.5}$ (particulate matter equal to or less than 2.5 micrometers in diameter) on childhood asthma incidence was explored in the Kaiser Air Pollution and Pediatric Asthma (KAPPA) Study, a historical birth cohort of 24,608 children enrolled in Kaiser Permanente Georgia. The KAPPA study is a well-defined cohort which allowed us to examine the risk of asthma development. A challenge in many epidemiologic studies is deciding how to define the outcome of interest. Asthma is particularly hard to define so we completed a systematic evaluation of case definitions for incident asthma in early life to decide how to define asthma using medical records in this study (Aim 1). We then estimated the effects of exposure to primary $PM_{2.5}$ from traffic emissions and total $PM_{2.5}$, which includes $PM_{2.5}$ from traffic, during pregnancy and the first year of life on cumulative asthma incidence by ages two through six (Aims 2 and 3). In Aim 2 we made a spatial comparison of pollution, estimating exposure using child residence during the first year of life and maternal residence during pregnancy. In Aim 3 we made both a spatial and temporal comparison of pollution.

In Aims 2 and 3 we used air pollution data created by colleagues at Georgia Institute of Technology. In Aim 2, we used 2011 annual average primary $PM_{2.5}$ from traffic emissions at 250 meter grid resolution created by RLINE, a line-source dispersion model (Community Modeling and Analysis System 2015, Zhai et al. 2015). These data vary spatially over the Atlanta region, but have no temporal variation. In Aim 3, we used daily $PM_{2.5}$ data created by a novel downscaling approach that integrates data from CMAQ, RLINE, and stationary air pollution monitors (Bates et al. 2016). These data differ from those used in Aim 3 in that

they capture total $PM_{2.5}$ and have both spatial and temporal variation. As noted in Chapters 5 and 6, the air quality data for this work are still being developed and may change before the publication of results. One such potential change is the use of emissions data at a different spatial and temporal resolution.

Moving to the results of this dissertation, in Aim 1 (Chapter 4), we evaluated 15 case definitions for asthma and found that choice of case definition had a substantial impact on the estimate of asthma incidence in early life; cumulative incidence of asthma by age 3 ranged from 9.0% to 35.2% depending on which definition was used. We also assessed the ability of asthma case definitions in the first three years of life to predict school-age asthma, defined as at least one asthma diagnosis between ages five and eight. In this population, with a prevalence of school-age asthma of 24%, the early-life asthma case definitions were far superior at ruling out school-age asthma (negative predictive values ranged from 79.1% to 86.6%) than they were at predicting school-age asthma (positive predictive values ranged from 43.5% to 71.5%). While positive and negative likelihood ratios indicate that overall predictive ability was limited, several of the case definitions examined performed similarly to clinical asthma prediction tools such as the Asthma Predictive Index and the Environmental and Childhood Asthma severity index (Castro-Rodriguez et al. 2000, Devulapalli et al. 2008).

Before conducting this analysis, we had planned to classify a child as asthmatic if they have one asthma diagnosis or two asthma-related medication dispensings (case definition seven). However, in Aim 1 we saw that compared to other case definitions, definition seven yielded the highest asthma incidence (35.2% by age 3) and correctly predicted asthma status at school age for the smallest percentage of children (71.4%) (Table 4.4). While these results convinced us not to use this case definition, there was not a clear best case definition out of the other 14 examined. We decided to use case definition

eight (one asthma diagnosis and one asthma-related medication dispensing) in our air pollution analyses because a child is classified as asthmatic if a doctor diagnosed their condition as asthma and they have some evidence of respiratory morbidity requiring medication. The inconsistencies between different asthma case definitions highlighted in Aim 1 motivated sensitivity analyses in Aims 2 and 3 which found that choice of case definition can also impact the estimate of the association between $PM_{2.5}$ exposure and asthma incidence (Table 5.10 and Table 6.11).

Aim 2 (Chapter 5) of this dissertation assessed whether exposure to primary $PM_{2.5}$ from traffic during pregnancy and the first year of life was associated with childhood asthma incidence in the KAPPA cohort. In fully adjusted models, an increase of 1 $\mu g/m^3$ of traffic $PM_{2.5}$ in the first year of life was associated with a 2.7% to 5.8% increase in the absolute risk of asthma by ages two through six with some 95% confidence intervals excluding the null (Table 5.8). Effect estimates were smaller for the association between traffic $PM_{2.5}$ exposure during pregnancy and asthma incidence, ranging from a 2.2% to 3.0% increase in the absolute risk of asthma. Estimated associations were greatly impacted by including city region in the model; risk differences from models including this variable were larger than risk differences from unadjusted models and adjusted models not including this variable (Table 5.6). Controlling for city region was motivated by concerns about lack of exchangeability between children living in the most urban and rural parts of the metropolitan Atlanta area, even after controlling for neighborhood socioeconomic status. Including it in the model allows us to compare children living in the same city region and aims to control for unmeasured potential confounders that vary spatially such as agricultural exposures and health care utilization. In our opinion, this addition to the model is important because we are making a spatial comparison of pollution exposure.

In Aim 2, we could not tease apart the effects of traffic $PM_{2.5}$ during pregnancy and the first year of life due to the high correlation between exposure estimates during these two time windows (Spearman correlation coefficient = 0.80). This correlation was higher than expected based on substantial residential mobility observed in this cohort, with 36% of children changing residence at least once during the period from conception to the first birthday, but is partially due to the lack of temporal variability in the $PM_{2.5}$ data used for exposure assignment. Overall, results from Aim 2 provide some evidence of a positive association between exposure to primary $PM_{2.5}$ from traffic during early life and childhood asthma incidence. However, our conclusions are somewhat tempered due to the dependence of results on which variables are included in the model. Including city region in our models, the rationale for which is described above, had a substantial impact on model results shifting all risk differences in the positive direction and resulting in confidence intervals that exclude the null. These results give the suggestion of negative bias if we don't control for city region. While we believe the models containing this variable most appropriately control for confounding, we note that our results are highly dependent on the inclusion of one variable.

In Aim 3 (Chapter 6) we assessed whether exposure to total $PM_{2.5}$ during pregnancy and the first year of life was associated with childhood asthma incidence in the KAPPA cohort. Results show little evidence of an association between exposure to total $PM_{2.5}$ and asthma incidence by ages 2 through 6. In fully adjusted models, an increase of 1 $\mu g/m^3$ total $PM_{2.5}$ during the first year of life was associated with an absolute increase in risk of asthma, with estimates of that increase ranging from 0.4% at age four (RD (95% CI) 0.004 (-0.014, 0.022)) to 1.8% at age five (RD (95% CI) 0.018 (-0.005, 0.041)). Risk differences for an increase of 1 $\mu g/m^3$ total $PM_{2.5}$ during the prenatal period were smaller and ranged from a 0.1% decrease in risk at age four ((RD (95% CI) -0.001 (-0.009, 0.006)) to a 0.4% increase in

risk at age six (RD (95% CI) 0.004 (-0.007, 0.014)). Confidence intervals around all adjusted estimates of both prenatal and first year of life exposure include the null. Due to the spatial and temporal variability in the $PM_{2.5}$ data used in Aim 3, there was a smaller correlation between prenatal and first year of life $PM_{2.5}$ estimates than in Aim 2 (Spearman correlation coefficient in Aim 3 = 0.59). This greater variability better enabled estimation of the independent effects of exposure in each of these periods by including both exposures in the same model. Effect estimates from this model are similar to those from models containing each exposure separately and provide little evidence of an effect of exposure during either of these periods on asthma incidence in childhood (Table 6.9).

There are some key differences between the air pollution data used in Aim 2 (RLINE) and Aim 3 (CMAQ-RLINE fusion). The RLINE data capture only primary $PM_{2.5}$ from traffic emissions and all variation in assigned pollution is spatial (i.e. one $PM_{2.5}$ estimate per grid cell over the study period). The CMAQ-RLINE fusion data capture total $PM_{2.5}$, which includes $PM_{2.5}$ from traffic as well as from other sources (e.g., power plants, biomass burning), and variation in the assigned pollution is both spatial and temporal (i.e. 3,287 daily $PM_{2.5}$ estimates per grid cell over the study period). Additionally, in the CMAQ-RLINE fusion data the spatial structure can change over time, although change was somewhat limited due to the spatial structure of the traffic impacts remaining the same. We completed an analysis to compare spatial and temporal variation in these datasets. To calculate spatial variability in the CMAQ-RLINE fusion data, we averaged estimates over space resulting in an estimate of mean and variance for the entire region for each of the 3,287 days for which data were available. To calculate temporal variability in the data, we took a similar approach, but instead averaged over time, resulting in an estimate of mean and variance for the entire time period for each of the 25,212 grids in the region in which KAPPA children

and mothers lived. For the RLINE data there is no temporal variability, and the spatial variability was examined by calculating the mean and variance of estimates across grids.

These analyses revealed that in the CMAQ-RLINE fusion data there is more temporal variability in $PM_{2.5}$ concentration than spatial variability; the mean of the variance is 2.9 for spatial variation and 50.3 for temporal variation. There was a large drop in $PM_{2.5}$ concentrations in 2008 so we were further interested in whether this shift was driving our estimate of temporal variation. We completed an analysis controlling for seasonality (by calculating 365 day moving averages) for years 2002-2007 and 2009-2010. Using this approach, the mean of the variance for years 2002-2007 is 0.23 and the mean of the variance for years 2009-2010 is 0.69. This drop in variation from the previous analysis, to a level lower than our estimate of spatial variation, implies that the estimate of temporal variability in $PM_{2.5}$ concentrations is largely driven by the decrease in $PM_{2.5}$ in 2008. To compare the RLINE and CMAQ-RLINE fusion data, we used the coefficient of variation as a measure due to the difference in the magnitude of $PM_{2.5}$ concentrations. One caveat to this approach is that if the mean and variance do not vary multiplicatively, it may not capture the true contrast of the variability. For example, if mean and variance vary additively, this approach may not accurately identify the data with more variability. Using this measure, it appears that the RLINE data has more spatial variation than the CMAQ-RLINE fusion data. The coefficient of variation of the 2011 annual average RLINE estimates is 24.7. The coefficient of variation of the annual averages of the CMAQ-RLINE fusion data range from 2.7 to 7.3. This result is consistent with our knowledge that secondary PM is fairly spatially homogenous and that the majority of PM is secondary, for example sulfate that is produced in secondary reactions of emissions from power plants. In comparison, primary PM from traffic emissions has more spatial heterogeneity which is determined by local roadway

emissions. Subsequently, it makes sense that the RLINE data would have more spatial heterogeneity than the CMAQ-RLINE fusion data.

Comparing the results from Aims 2 and 3, we see evidence of a stronger association between primary $PM_{2.5}$ from traffic and childhood asthma (Aim 2) than between total $PM_{2.5}$ and childhood asthma (Aim 3). The correlation between our estimate of first year of life exposure to total $PM_{2.5}$ and traffic $PM_{2.5}$ is low (Spearman correlation coefficient = 0.26). However, the correlation increases to 0.57 when removing some of the temporal variability by controlling for birth year. Correlation between prenatal estimates using the two types of pollution data is much lower (overall correlation = 0.20, correlation controlling for year of conception = 0.29). This low correlation is partially due to seasonality in prenatal estimates of total $PM_{2.5}$ (Figure 6.3) and lack of seasonality in prenatal estimates of traffic $PM_{2.5}$ which are always averaged over a full year. In Aim 2 the variation in exposure is purely spatial while in Aim 3 the variation in exposure has both spatial and temporal components. In order to facilitate a better comparison between results from these two aims, we completed an analysis in Aim 3 that made a purely spatial comparison by using only 2010 $PM_{2.5}$ annual averages for exposure assignment (Table 6.12). We found that an increase of 1 $\mu g/m^3$ of total $PM_{2.5}$ is associated with a 0.7% to 1.7% increase in the absolute risk of asthma, with most confidence intervals including the null. Similar to the main results from Aim 3, these estimates show less evidence of an association than those from Aim 2.

Considering the primary sources of bias of confounding, misclassification, and selection bias, I believe that confounding is of the most concern in this dissertation. We completed extensive analyses in Chapters 4, 5 and 6 to assess potential selection bias and observed no convincing evidence that it impacted our results. To help mitigate concerns about the impact of disease misclassification, in Chapter 4 we examined different ways to define asthma in our cohort. In Chapters 5 and 6 we completed sensitivity analyses to assess

whether the estimate of the association between $PM_{2.5}$ and asthma changed if we used different asthma case definitions. There is no indication that misclassification of this dichotomous outcome would be dependent on exposure status or other factors so any misclassification would be expected to bias our results toward the null. We believe that our exposure of interest, residential ambient $PM_{2.5}$, is fairly well estimated with the available air pollution data. Since our exposure, $PM_{2.5}$, is continuous we cannot assume that nondifferential misclassification would bias results toward the null. However, it is difficult to conceptualize a scenario in which this type of misclassification would bias our estimates away from the null. If personal exposure was our primary exposure of interest, which it is not, we would be more concerned about the potential influence of exposure misclassification on our results.

In comparison, we have less ability to determine how uncontrolled potential confounding may be impacting our results. We are able to control for key variables that could be potential confounders and our control for some variables is superior to that in previous studies. For example, we assessed the impact of both individual-level SES and neighborhood-level (contextual) SES, each of which may be important and are not always both considered in epidemiologic studies of air pollution and asthma. Our concern about potential confounding stems from analyses in Chapter 5 where including city region in models of the association between traffic $PM_{2.5}$ and asthma incidence has a large impact on results (Table 5.6). The fact that our results are sensitive to control for this variable is noteworthy particularly since this is a variable that has not been adjusted for in similar analyses from previous cohort studies. While we believe that conditioning on city region, making comparisons between children of the same city region, results in the most valid estimates of the association between $PM_{2.5}$ and asthma, there is no way to confirm this. The large impact of controlling for city region on model results tempers our conclusions

somewhat and raises concerns about other possible unmeasured confounders that could have similarly large impacts on estimated associations. In Chapter 6 control for city region also impacts results, but results are similar enough between models that inclusion does not meaningfully impact conclusions.

While the primary measure of association for this dissertation is the risk difference, we also assessed associations using the risk ratio, odds ratio, and hazard ratio. I will briefly discuss some differences between the risk difference, risk ratio, and hazard ratio. The risk difference and risk ratio both estimate changes in risk and can only be calculated with denominator data, like those available from the KAPPA study. The hazard ratio estimates changes in hazard which is arguably a less intuitive measure than risk. Out of these measures, only the risk difference assesses additive effects (all other measures assess multiplicative effects). Subsequently, the risk difference is well suited to assess additive interaction which is thought to be more relevant to public health than multiplicative interaction (Greenland S et al. 2008, Vanderweele et al. 2014). The interpretation of risk differences is arguably the most straightforward out of any of these measures since risk differences represent absolute change in the probability of disease, rather than relative change expressed in ratio measures. A limitation of the risk difference is that the binomial models which are used to calculate risk differences often encounter convergence problems. Convergence problems can sometimes be easily solved, for example by changing category groupings of a covariate or changing iteration limits in SAS. However, they can also be more serious and in some cases can prevent an analysis from producing results. From using the risk difference as our primary measure, we encountered both easily solvable and unresolvable convergence issues in this dissertation.

Risk ratios are calculated from log binomial models which, like the risk difference, are also prone to convergence difficulties, but perhaps less frequently. One nuance of the

risk ratio is that it decreases as baseline prevalence of an outcome increases. For example, if the risk of an outcome among unexposed individuals is 20%, a risk ratio of 2.0 would occur if the risk of the outcome among exposed individuals is 40%. On the other hand, if the risk of the outcome among the unexposed is 60%, it is not possible to have a risk ratio of 2.0 while keeping risk in the exposed group within the bounds of 0 to 1. The highest possible risk ratio in the population would be 1.7 and would arise from the exposed group having a risk of the outcome of 100%. This attribute of the risk ratio is apparent when comparing the risk ratio results between different follow-up ages in this dissertation. In general, as age increases and the baseline prevalence of asthma increases, the risk ratio for the association between $PM_{2.5}$ and asthma gets closer to the null. In all analyses using the risk ratio as the measure of association the risk ratio at age 2 is further away from the null than the risk ratio at age 6 (Table 5.11, Table 6.13). One strength of the risk ratio is that in a scenario with independent and nondifferential disease misclassification with perfect specificity, there is no bias of the risk ratio (Greenland S et al. 2008). This is unique to the risk ratio; in the same scenario the risk difference would be biased downward.

Our analyses using the hazard ratio as the measure of association of interest were completed using Cox proportional hazards models. These analyses had more power than our binomial and log binomial models due to the use of censoring to include children who were lost to follow-up in analyses (rather than excluding these children). As mentioned in Chapter 6, a limitation of Cox proportional hazards models is that unlike our other analyses they use the exact time of asthma diagnosis which is unlikely to be reliable. There was more consistency in hazard ratios between different follow-up ages than there was in risk differences or risk ratios at different follow-up ages. This is partially due to the fact that the exact same children are in each analysis, unlike in other analyses where some children are lost to follow-up. So for example, most of the data contributing to the age 6 analysis are

identical to the data in the age 5 analysis, leading to very similar results. The implicit proportional hazards assumption, that a single hazard ratio is valid for the entire time period in each analysis, likely also plays a role in the consistency between hazard ratios across follow-up ages.

This dissertation contributes to our knowledge of the use of medical records to define asthma in early life, and the associations between $PM_{2.5}$ exposure, overall and specifically from traffic sources, in key developmental windows and childhood asthma incidence. The KAPPA study is different from previous studies in its use of the risk difference as a measure of association, the use of RLINE data, and the use of CMAQ data downscaled to a 250 meter grid resolution. This study leverages work by colleagues at Georgia Institute of Technology to create innovative spatially and temporally resolved estimates of ambient $PM_{2.5}$. Our results indicate that discrepant results between studies could be due to the use of different asthma case definitions and that future studies may want to explore controlling for city region as a way to increase exchangeability between children if making a spatial comparison in a city with a similar pattern of pollution and urban sprawl. The high correlation between prenatal and first year of life $PM_{2.5}$ exposure, despite the residential mobility of our cohort and the use of temporally varying pollution data, suggests that the results of studies looking at exposure in only one of these windows may be driven by the effect of exposure during another period. This body of work adds to a growing body of research showing some evidence for a positive association between $PM_{2.5}$ exposure, particularly from traffic sources, in early life and childhood asthma. Future research will help us better understand these relationships; the knowledge of which could ultimately lead to cleaner air and reduced risk of asthma.

APPENDIX A

Additional detail on data sources

Table A1. Demographic Cluster Descriptions Provided by Georgia Department of Public Health, Office of Health Indicators for Planning (OHIP)

Citation: (Millard et al. 2012)

A.1 Georgia's wealthiest cluster is primarily populated by "new money" executives and professionals living in tract mansions of metropolitan suburbs and exurbs. Predominantly White with an above average index for Asians, this highly educated cluster is composed of married couples in their middle adulthood ages (45-64) with young and adolescent children.
A.2 This well-educated, suburban cluster, dominated by professionals and managers, has the second highest level of affluence in the state. Mostly White with a high percentage in their middle or late adulthood (55+), they have adolescent and grown children.
A.3 Found in the metro suburbs, this mixed-ethnicity with majority of Whites and high index for African-Americans, more youthful cluster is populated by married couples in their late 20's through early 40's with young children. The majority has some college degree or are college graduates. Most are employed in sales and other white-collar jobs, while some are high-earning blue-collar families. This cluster has a median household income well above the state average.
B.1 This cluster is characterized by its high concentration of White and Asian non-family households renting in upscale apartments. With easy access to major highways, this cluster is the home for young managers and professionals in their late 20's through early 40's, predominately with college degrees and beyond. They live a modern urban lifestyle in the most densely populated urban neighborhoods before they establish families and move to suburban areas.
B.2 This small cluster is populated by military personnel in their early and young adulthood ages (18-34) with some college degrees. Majority of the population are White. They live in rented apartments and condos in urban areas. Their median income is around the state average.
B.3 This is a mixed-ethnicity cluster with a high index of Asian and Multiracial non-family households living in middle-range value apartments in urban/suburban areas. Although many have some college degrees or are college graduates, their median income is below the state average due to their recent entry into the workforce.
B.4 This mixed-ethnicity cluster mainly represents the college populations in Georgia (populations living in group quarters). They are mostly between 18-24 years of age and have incomes lower than the state average.
C.1 This is a White, middle-class rural cluster dominated by married families of people aged 55 years and over. They are mainly home owners, but the value of their housing is lower than in some of the urban and suburban clusters. Many in this cluster are high school graduates. Found predominantly in N/NE rural counties of Georgia, this cluster is highly represented in farming, production, and construction.
C.2 This rural cluster is dominated by married families of people in their middle adulthood ages with young and adolescent children. Found widespread in rural counties of Georgia, the cluster is White with some African-American population. Many people are in construction and production jobs; their incomes are average compared to the state.
C.3 Found in relatively populated areas in rural counties, this mixed-ethnicity cluster with high index of African-Americans is populated by older people living in old houses. With mixed levels of education, people in this cluster mainly work in lower paying service, sales and managerial jobs earning below state average incomes.
C.4 This rural cluster is composed of married and single parent families of predominantly White population with or without children. Most have high school diploma or less; they mainly work in farming, production, and construction earning well below the state average income.

Table A1. Demographic Cluster Descriptions Provided by Georgia Department of Public Health, Office of Health Indicators for Planning (OHIP) *(Continued)*

D.1 An urban cluster, this mixed-race group has a high representation of single-parent families with or without children. Most have a high school diploma or less; this group mainly works in the service industry earning lower than state average income. They live in rented apartments or old houses of low housing values.
D.2 This is a small cluster composed of military personnel in their early and young adulthood ages (18-34) with some college degrees. A mixed-ethnicity group with majority of Whites, this cluster is populated by married and single families with young children. The percentage of population 18 years of age and younger is higher than any other cluster in the state.
D.3 This is the oldest urban cluster with high proportion of 55 years of age and older. Primarily African-American with a high index for non-Hispanic Whites, this cluster is characterized by single family or non-family households living in their own old houses in urban/suburban areas. They work in low-paying service and sales jobs earning incomes lower than the state average.
D.4 This cluster is composed predominantly of African-Americans with a high percentage of single family households with or without children. It is relatively young among urban clusters with a high percentage of population between 18-34 years of age. They are primarily renters, have high school or less than high school educations and work in service industry--making 30% below the state average in income.
D.5 This is a mixed-ethnicity cluster with a high index of Hispanics and Multiracial groups. Most have high school diploma or less; they mainly work in low-paying blue collar jobs in production and construction industries. The cluster's housing is half owner-occupied and half renter-occupied with a high percentage of vacant housing.
D.6 This cluster is predominantly populated by African-Americans with high percentage of population in their 60's and over. Most have a high school diploma or less; they mainly work in service industries. Their median income is second lowest in the state.
D.7 This cluster is predominantly composed of very young African-Americans with more females than males. The cluster has the highest percentage of population less than 18 years of age in nonmilitary clusters in the state, of whom most live in female-headed households. Most have a high school diploma or less; they work in low-paying jobs and live in rental units. The median household income in this cluster is the lowest in the state.

Table A2. Asthma-related medications

List and classifications of asthma medications	
Maintenance Medications	Rescue Medications
Combinations (ICS and LABA)	*Combinations*
Budesonide/Formoterol	Ipratropium/Albuterol
Fluticasone/Salmeterol	*Anticholinergics*
Inhaled corticosteroids (ICS)	Ipratropium
Beclomethasone	Tiotropium
Budesonide	*Short acting beta-agonists*
Flunisolide	Albuterol
Flunisolide/Menthol	Albuterol Sulfate
Fluticasone	Bitolterol
Mometasone	Isoetharine
Triamcinolone	Isoproterenol
Methylxanthines	Levalbuterol
Aerolate	Metaproterenol
Aminophyllin/ephed/potiod/pb	Pirbuterol
Aminophylline	Terbutaline
Aminophylline/ephed/amobarb	
Aminophylline/ephed/phenobarb	
Aminophylline/Ephedrine	
Aminophylline/Phenobarb	
Aminophylline/Quinine	
Dyphylline	
Mast cell stabilizers	
Cromolyn	
Nedocromil	
Long acting beta-agonists (LABA)	
Arformotherol	
Formoterol	
Salmeterol	
Leukotriene antagonist (LRA)	
Montelukast	
Zafirlukast	
Zileuton	
Theop/Isoproterenol/epd/ki/pb	
Theophyll/caff/aa13/cinn/hc135	
Theophyll/ephed hcl/phenobarb	
Theophyll/ephed/potiodide/pb	
Theophylline	
Theophylline-ephed-butaba	
Theophylline-ephed-phenob	
Theophylline-ephedrine	
Dyphylline-ephedrein-phen	
Guaifen/Dyphyllin/ephed/pb	
Guaifen/theop anhyd/pephed	
Guaifenesin/Dyphylline	
Guaifenesin/Oxtriphylline	
Guaifenesin/Theophylline	
Oxtriphylline	
Oxtriphylline-Guaifenesin	

Table A2. Asthma-related medications *(Continued)*

List of medications used to search Kaiser electronic medical records (Includes combinations and specifics which may be omitted list of asthma medications above)	
GENERIC_NM_GROUP	GNN_GENERIC_NM
AEROLATE	AEROLATE
ALBUTEROL	ALBUTEROL
ALBUTEROL	ALBUTEROL SULFATE
AMINOPHYLLIN/EPHED/POT IOD/PB	AMINOPHYLLIN/EPHED/POT IOD/PB
AMINOPHYLLINE	AMINOPHYLLINE
AMINOPHYLLINE/EPHED/AMOBARB	AMINOPHYLLINE/EPHED/AMOBARB
AMINOPHYLLINE/EPHED/PHENOBARB	AMINOPHYLLINE/EPHED/PHENOBARB
AMINOPHYLLINE/EPHEDRINE	AMINOPHYLLINE-EPHEDRINE-A
AMINOPHYLLINE/EPHEDRINE	AMINOPHYLLINE-EPHEDRINE-P
AMINOPHYLLINE/EPHEDRINE	AMINOPHYLLINE-GG
AMINOPHYLLINE/PHENOBARB	AMINOPHYLLINE WITH PHENOBARBITAL
AMINOPHYLLINE/QUININE	AMINOPHYLLINE/QUININE SULFATE
ARFORMOTEROL	ARFORMOTEROL TARTRATE
BECLOMETHASONE	BECLOMETHASONE DIPROPIONATE
BITOLTEROL	BITOLTEROL MESYLATE
BUDESONIDE	BUDESONIDE
BUDESONIDE/FORMOTEROL	BUDESONIDE/FORMOTEROL FUMARATE
CROMOLYN	CROMOLYN SODIUM
DYPHYLLINE	DYPHYLLINE
DYPHYLLINE-EPHEDRINE-PHEN	DYPHYLLINE-EPHEDRINE-PHEN
FLUNISOLIDE	FLUNISOLIDE
FLUNISOLIDE/MENTHOL	FLUNISOLIDE/MENTHOL
FLUTICASONE	FLUTICASONE PROPIONATE
FLUTICASONE/SALMETEROL	FLUTICASONE/SALMETEROL
FORMOTEROL	FORMOTEROL FUMARATE
GUAIFEN/DYPHYLLIN/EPHED/PB	GUAIFEN/DYPHYLLIN/EPHED/PB
GUAIFEN/THEOP ANHYD/P-EPHED	GUAIFEN/THEOP ANHYD/P-EPHED
GUAIFENESIN/DYPHYLLINE	GUAIFENESIN/DYPHYLLINE
GUAIFENESIN/OXTRIPHYLLINE	GUAIFENESIN/OXTRIPHYLLINE
GUAIFENESIN/THEOPHYLLINE	GUAIFENESIN/THEOPHYLLINE
IPRATROPIUM	IPRATROPIUM BROMIDE
IPRATROPIUM/ALBUTEROL	IPRATROPIUM/ALBUTEROL SULFATE
ISOETHARINE	ISOETHARINE HCL
ISOPROTERENOL	ISOPROTERENOL HCL
ISOPROTERENOL	ISOPROTERENOL SULFATE
LEVALBUTEROL	LEVALBUTEROL HCL
LEVALBUTEROL	LEVALBUTEROL TARTRATE
METAPROTERENOL	METAPROTERENOL SULFATE
MOMETASONE	MOMETASONE FUROATE
MONTELUKAST	MONTELUKAST SODIUM
NEDOCROMIL	NEDOCROMIL SODIUM
OXTRIPHYLLINE	OXTRIPHYLLINE
OXTRIPHYLLINE-GUAIFENESIN	OXTRIPHYLLINE-GUAIFENESIN
PIRBUTEROL	PIRBUTEROL ACETATE
SALMETEROL	SALMETEROL XINAFOATE
TERBUTALINE	TERBUTALINE SULFATE

Table A2. Asthma-related medications *(Continued)*

THEOP/ISOPROTERENOL/EPD/KI/PB	THEOP/ISOPROTERENOL/EPD/KI/PB
THEOPHYLL/CAFF/AA13/CINN/HC135	THEOPHYLL/CAFF/AA13/CINN/HC135
THEOPHYLL/EPHED HCL/PHENOBARB	THEOPHYLL/EPHED HCL/PHENOBARB
THEOPHYLL/EPHED/BUTABARBITAL	THEOPHYLL/EPHED/BUTABARBITAL
THEOPHYLL/EPHED/POT IODIDE/PB	THEOPHYLL/EPHED/POT IODIDE/PB
THEOPHYLLINE	THEOPHYLLINE
THEOPHYLLINE	THEOPHYLLINE ANHYDROUS
THEOPHYLLINE	THEOPHYLLINE SODIUM GLYCI
THEOPHYLLINE	THEOPHYLLINE SODIUM GLYCINATE-GG
THEOPHYLLINE	THEOPHYLLINE TIMED RELEASE
THEOPHYLLINE	THEOPHYLLINE-ALCOHOL,SUGAR,DYE FREE
THEOPHYLLINE-EPHED-BUTABA	THEOPHYLLINE-EPHED-BUTABA
THEOPHYLLINE-EPHED-PHENOB	THEOPHYLLINE-EPHED-PHENOB
THEOPHYLLINE-EPHEDRINE	THEOPHYLLINE-EPHEDRINE
THEOPHYLLINE-EPHEDRINE-GG	THEOPHYLLINE-EPHEDRINE-GG
THEOPHYLLINE-EPHEDRINE-PB	THEOPHYLLINE-EPHEDRINE-PB
THEOPHYLLINE-GUAIFENESIN	THEOPHYLLINE-GUAIFENESIN
THEOPHYLLINE-IODINATED GL	THEOPHYLLINE-IODINATED GL
THEOPHYLLINE-KI	THEOPHYLLINE-KI
THEOPHYLLINE-PSE-GG	THEOPHYLLINE-PSE-GG
THEOPHYLLINE/DIETARY SUP.CMB9	THEOPHYLLINE/DIETARY SUP.CMB9
THEOPHYLLINE/EPHED/HYDROXYZINE	THEOPHYLLINE/EPHED/HYDROXYZINE
THEOPHYLLINE/POTASSIUM IODIDE	THEOPHYLLINE/POTASSIUM IODIDE
TIOTROPIUM	TIOTROPIUM BROMIDE
TRIAMCINOLONE	TRIAMCINOLONE ACETONIDE
ZAFIRLUKAST	ZAFIRLUKAST
ZILEUTON	ZILEUTON

APPENDIX B

Additional analyses of prenatal and first year of life exposure to primary $PM_{2.5}$ from traffic and childhood asthma incidence in a birth cohort

Table B1. Loss to follow-up in the KAPPA cohort

Age at follow-up	KAPPA children eligible for each follow-up age [n]	Children followed to each age [n (% of eligible children)]
Age 2	23,100	17,960 (77.8%)
Age 3	22,627	14,251 (63.0%)
Age 4	21,117	10,999 (52.1%)
Age 5	19,362	8,592 (44.4%)
Age 6	17,251	6,629 (38.4%)

To be eligible for each follow-up age, a child has to be born early enough to be that age in September 2013 when the KAPPA cohort was defined. For example, a child born in 2010 could be at most 3 years old at the time KAPPA follow-up ended in September 2013.

Table B2. Comparison of covariates for children retained in cohort and lost to follow-up

Cohort	Children retained in cohort until follow-up age				Children eligible for follow-up age, but lost to follow-up before age			
Mean traffic first year of life $PM_{2.5}$ [$\mu g/m^3$]								
Age 2	1.17				1.17			
Age 3	1.17				1.17			
Age 4	1.17				1.16			
Age 5	1.17				1.16			
Age 6	1.17				1.16			
Incident asthma at previous follow-up age [a] [%]								
Age 3	11.5				9.4			
Age 4	18.8				16.9			
Age 5	24.7				23.0			
Age 6	28.8				28.1			
Mean traffic $PM_{2.5}$ stratified by asthma at previous age [a] [$\mu g/m^3$]								
	Asthma		No Asthma		Asthma		No Asthma	
Age 3	1.15		1.17		1.16		1.17	
Age 4	1.16		1.17		1.15		1.16	
Age 5	1.16		1.17		1.13		1.16	
Age 6	1.17		1.17		1.15		1.15	
Child race [%]								
	Black	White	Other	Unknown	Black	White	Other	Unknown
Age 2	36.9	39.7	11.5	11.9	26.6	37.0	13.7	22.8
Age 3	38.7	39.6	11.2	10.6	27.7	38.0	13.4	20.9
Age 4	40.1	39.7	10.9	9.4	28.2	37.8	12.7	21.4
Age 5	41.8	39.2	10.4	8.6	28.3	38.2	12.0	21.5
Age 6	43.2	39.0	10.0	7.9	28.8	38.6	11.0	21.6
Major demographic cluster [%]								
	A	B	C	D	A	B	C	D
Age 2	62.7	9.6	4.7	23.0	61.2	10.9	5.6	22.4
Age 3	63.0	9.2	4.7	23.2	61.4	10.9	5.3	22.3
Age 4	62.9	9.0	4.6	23.5	62.1	10.3	5.1	22.5
Age 5	62.7	8.9	4.7	23.7	62.4	9.7	5.2	22.7
Age 6	62.7	8.9	4.4	24.0	62.3	9.3	5.4	23.0
City region [%]								
	Metro Atlanta	≤10 mi from metro Atlanta	>10 mi from metro Atlanta		Metro Atlanta	≤10 mi from metro Atlanta	>10 mi from metro Atlanta	
Age 2	10.5	43.3	46.2		10.4	41.2	48.5	
Age 3	10.4	43.9	45.7		10.4	41.2	48.3	
Age 4	10.3	45.1	44.6		10.2	41.2	48.6	
Age 5	10.3	46.1	43.6		9.8	41.4	48.8	
Age 6	10.4	47.3	42.3		9.6	42.1	48.3	

To be eligible for each follow-up age a child has to be born early enough to be that age in September 2013 when the KAPPA cohort was defined. For example, a child born in 2010 could be at most 3 years old at the time KAPPA follow-up ended in September 2013. [a] Incident asthma at previous follow-up age only calculated among children in previous cohort (can't calculate for age 2 since no asthma classifications at age 1).

Table B3. Additional model results for the association between first year of life traffic $PM_{2.5}$ exposure and asthma incidence

Cohort	MODEL A1		MODEL A2		MODEL A3	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Age 2	-0.0234 (-0.0380, -0.0087)[a]	<0.01	0.0278 (0.0038, 0.0518)	0.02	0.0251 (0.0014, 0.0488)	0.04
Age 3	-0.0299 (-0.0553, -0.0045)	0.02	0.0369 (0.0040, 0.0698)	0.03	0.0334 (0.0007, 0.0661)[a]	0.05
Age 4	-0.0331 (-0.0657, -0.0004)	0.05	0.0355 (-0.0091, 0.0801)	0.12	0.0363 (-0.0071, 0.0797)	0.10
Age 5	-0.0194 (-0.0586, 0.0199)	0.33	0.0586 (0.0045, 0.1128)	0.04	0.0560 (0.0030, 0.1090)	0.04
Age 6	-0.0313 (-0.0774, 0.0149)	0.18	0.0349 (-0.0299, 0.0997)	0.29	0.0361 (-0.0275, 0.0996)	0.27
Cohort	MODEL A4		MODEL A5			
	RD (95% CI)	p	RD (95% CI)	p		
Age 2	0.0295 (0.0050, 0.0540)[a]	0.02	0.0267 (0.0025, 0.0510)	0.03		
Age 3	0.0380 (0.0047, 0.0712)[a]	0.03	0.0363 (0.0032, 0.0693)[a]	0.03		
Age 4	0.0398 (-0.0048, 0.0843)	0.08	0.0369 (-0.0079, 0.0817)	0.11		
Age 5	0.0600 (0.0057, 0.1143)	0.03	0.0570 (0.0025, 0.1115)	0.04		
Age 6	0.0376 (-0.0272, 0.1023)	0.26	—	—		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, even if dropped birth year, [a] birth year dropped in order for model to converge

List of Model Covariates

Model A1: child sex, child race, child ethnicity, maternal asthma, birth year, maternal age, maternal education, paternal education, maternal marital status *[Model 1 with paternal education, maternal marital status]*

Model A2: child sex, child race, maternal asthma, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta) *[Model 7 (final model) without birth year]*

Model A3: child sex, child race, maternal asthma, minor demographic cluster, birth year, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta) *[Model 7 (final model) without accounting for correlation between siblings, correlation structure (necessary to implement robust variance estimation) determined using child study id instead of family id]*

Model A4: child sex, child race, maternal asthma, minor demographic cluster, birth year, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta), interaction between city region and race *[Model 7 (final model) with interaction term between city region and race] [Score statistic p-values for interaction terms: age 2 – did not converge, age 3 – 0.77, age 4 – 0.40, age 5 – 0.21, age 6 – 0.49]*

Model A5: child sex, child race, maternal asthma, minor demographic cluster, birth year, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta), interaction between metro Atlanta and minor demographic cluster *[Model 7 (final model) with interaction term between metro Atlanta and minor demographic cluster. Did not include interaction terms with all city regions because of model convergence issues] [Score statistic p-values for interaction terms: age 2 – did not converge, age 3 – 0.98, age 4 – 0.83, age 5 – 0.81]*

Table B4. First year of life traffic $PM_{2.5}$ exposure and asthma incidence, stratified by race

Cohort	Unadjusted models stratified by race			
	Black		White	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0004 (-0.0311, 0.0319)	0.98	-0.0262 (-0.0516, -0.0007)	0.04
Age 3	0.0031 (-0.0395, 0.0456)	0.89	-0.0381 (-0.0738, -0.0024)	0.04
Age 4	0.0070 (-0.0465, 0.0605)	0.80	-0.0652 (-0.1102, -0.0202)	<0.01
Age 5	0.0367 (-0.0275, 0.1009)	0.26	-0.0594 (-0.1137, -0.0052)	0.03
Age 6	0.0110 (-0.0638, 0.0858)	0.77	-0.0735 (-0.1373, -0.0098)	0.02
Cohort	Adjusted models stratified by race			
	Black		White	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0481 (0.0076, 0.0886)	0.02	—	—
Age 3	0.0277 (-0.0302, 0.0856)	0.35	—	—
Age 4	0.0361 (-0.0364, 0.1085)	0.33	0.0728 (0.0036, 0.1419)[a]	0.04
Age 5	0.0695 (-0.0182, 0.1572)	0.12	0.0880 (0.0066, 0.1694)	0.03
Age 6	0.0329 (-0.0696, 0.1353)	0.53	0.0545 (-0.0440, 0.1529)	0.28

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, even if dropped birth year, [a] birth year dropped in order for model to converge

Adjusted models include all covariates in the final adjusted model: child sex, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). The following minor demographic clusters were combined in order to aid in model convergence: A.2 and A.3; D.1 and D.3; D.5, D.6 and D.7.

Table B5. First year of life traffic $PM_{2.5}$ exposure and asthma incidence, assessing interaction between traffic $PM_{2.5}$ and race

Parameter	AGE 2		AGE 3 [a]		AGE 4	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Traffic $PM_{2.5}$	0.0288 (-0.0026, 0.0601)	0.07	0.0403 (-0.0038, 0.0844)	0.07	0.0275 (-0.0302, 0.0853)	0.35
Black race	0.0118 (-0.0361, 0.0596)	0.63	0.0105 (-0.0545, 0.0755)	0.75	0.0000 (-0.0839, 0.0840)	1.00
Unknown/other race	0.0291 (-0.0195, 0.0776)	0.24	0.0391 (-0.0255, 0.1037)	0.24	0.0004 (-0.0964, 0.0971)	0.99
Traffic $PM_{2.5}$*black race	0.0102 (-0.0304, 0.0508)	0.62	0.0146 (-0.0400, 0.0693)	0.60	0.0259 (-0.0451, 0.0968)	0.48
Traffic $PM_{2.5}$*unknown/other race	-0.0214 (-0.0596, 0.0168)	0.27	-0.0348 (-0.0837, 0.0142)	0.16	0.0014 (-0.0753, 0.0780)	0.97
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
Traffic $PM_{2.5}$	0.0412 (-0.0285, 0.1108)	0.25	0.0214 (-0.0627, 0.1056)	0.62		
Black race	0.0002 (-0.0994, 0.0997)	1.00	0.0255 (-0.0936, 0.1446)	0.67		
Unknown/other race	0.0313 (-0.0894, 0.1520)	0.61	0.0261 (-0.1193, 0.1716)	0.72		
Traffic $PM_{2.5}$*black race	0.0469 (-0.0382, 0.1319)	0.28	0.0294 (-0.0718, 0.1306)	0.57		
Traffic $PM_{2.5}$*unknown/other race	-0.0082 (-0.1043, 0.0878)	0.87	0.0129 (-0.1030, 0.1288)	0.83		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] Birth year dropped in order for model to converge

Table only includes model output relevant to interaction of interest. Models also control for child sex, birth year, maternal asthma, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Reference group: white race. *Score statistic p-values for interaction terms: age 2 – did not converge, age 3 – 0.32, age 4 – 0.77, age 5 – 0.50, age 6 – 0.86*

Table B6. First year of life traffic $PM_{2.5}$ exposure and asthma incidence, stratified by sex

Cohort	Unadjusted models stratified by sex			
	Male		Female	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	-0.0355 (-0.0615, -0.0094)	<0.01	-0.0094 (-0.0300, 0.0111)	0.37
Age 3	-0.0341 (-0.0701, 0.0020)	0.06	-0.0326 (-0.0619, -0.0032)	0.03
Age 4	-0.0471 (-0.0927, -0.0014)	0.04	-0.0304 (-0.0693, 0.0085)	0.13
Age 5	-0.0199 (-0.0763, 0.0364)	0.49	-0.0170 (-0.0634, 0.0294)	0.47
Age 6	-0.0546 (-0.1199, 0.0107)	0.10	0.0002 (-0.0560, 0.0564)	0.99
Cohort	Adjusted models stratified by sex			
	Male		Female	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0410 (0.0029, 0.0791)	0.04	0.0159 (-0.0132, 0.0450)	0.28
Age 3	0.0544 (0.0016, 0.1073)	0.04	0.0202 (-0.0194, 0.0598)[a]	0.32
Age 4	0.0375 (-0.0301, 0.1051)	0.28	0.0269 (-0.0280, 0.0817)	0.34
Age 5	0.0678 (-0.0139, 0.1494)	0.10	0.0390 (-0.0295, 0.1075)	0.27
Age 6	-0.0148 (-0.1100, 0.0804)	0.76	0.0631 (-0.0208, 0.1471)	0.14

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, even if dropped birth year, [a] birth year dropped in order for model to converge

Adjusted models include all covariates in the final adjusted model: child race, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). The following minor demographic clusters were combined in order to aid in model convergence: A.2 and A.3; D.1 and D.3; D.5, D.6 and D.7.

Table B7. First year of life traffic $PM_{2.5}$ exposure and asthma incidence, assessing interaction between traffic $PM_{2.5}$ and sex

Parameter	AGE 2 [a]		AGE 3 [a]		AGE 4	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Traffic $PM_{2.5}$	0.0393 (0.0141, 0.0645)	<0.01	0.0374 (0.0012, 0.0736)	0.04	0.0449 (-0.0052, 0.0951)	0.08
Male sex	0.0756 (0.0374, 0.1138)	<0.01	0.0702 (0.0178, 0.1225)	<0.01	0.0982 (0.0285, 0.1679)	<0.01
Traffic $PM_{2.5}$*male sex	-0.0271 (-0.0586, 0.0043)	0.09	-0.0011 (-0.0440, 0.0419)	0.96	-0.0169 (-0.0743, 0.0404)	0.56
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
Traffic $PM_{2.5}$	0.0654 (0.0048, 0.1259)	0.03	0.0727 (-0.0003, 0.1457)	0.05		
Male sex	0.1066 (0.0229, 0.1902)	0.01	0.1775 (0.0788, 0.2762)	<0.01		
Traffic $PM_{2.5}$*male sex	-0.0166 (-0.0864, 0.0531)	0.64	-0.0780 (-0.1599, 0.0040)	0.06		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] Birth year dropped in order for model to converge

Table only includes model output relevant to interaction of interest. Models also control for child race, birth year, maternal asthma, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Reference group: female sex.
Score statistic p-values for interaction terms: age 2 –did not converge, age 3 –0.96 , age 4 –0.58 , age 5 –0.66 , age 6 –0.08

Table B8. First year of life traffic $PM_{2.5}$ exposure and asthma incidence, stratified by maternal asthma

Cohort	Unadjusted models stratified by maternal asthma			
	Maternal Asthma		No Maternal Asthma	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	-0.0174 (-0.0758, 0.0410)	0.56	-0.0225 (-0.0415, -0.0035)	0.02
Age 3	-0.0498 (-0.1282, 0.0286)	0.21	-0.0266 (-0.0534, 0.0002)	0.05
Age 4	-0.0821 (-0.1792, 0.0149)	0.10	-0.0271 (-0.0615, 0.0072)	0.12
Age 5	-0.1032 (-0.2215, 0.0151)	0.09	-0.0019 (-0.0432, 0.0395)	0.93
Age 6	-0.0490 (-0.1849, 0.0869)	0.48	-0.0207 (-0.0699, 0.0284)	0.41
Cohort	Adjusted models stratified by maternal asthma			
	Maternal Asthma		No Maternal Asthma	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	—	—	0.0270 (0.0021, 0.0519)	0.03
Age 3	0.0181 (-0.0906, 0.1267)	0.74	0.0405 (0.0038, 0.0772)	0.03
Age 4	0.0005 (-0.1319, 0.1329)	0.99	0.0380 (-0.0134, 0.0893)	0.15
Age 5	-0.0118 (-0.1777, 0.1541)	0.89	0.0552 (-0.0067, 0.1171)	0.08
Age 6	-0.0016 (-0.1932, 0.1900)	0.99	0.0331 (-0.0404, 0.1067)	0.38

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, even if dropped birth year

Adjusted models include all covariates in the final adjusted model: child sex, child race, birth year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta).

Table B9. First year of life traffic $PM_{2.5}$ exposure and asthma incidence, assessing interaction between traffic $PM_{2.5}$ and maternal asthma

Parameter	AGE 2		AGE 3 [a]		AGE 4	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Traffic $PM_{2.5}$	0.0271 (0.0030, 0.0512)	0.03	0.0393 (0.0061, 0.0724)	0.02	0.0436 (-0.0020, 0.0891)	0.06
Maternal asthma	0.0711 (0.0007, 0.1415)	0.05	0.1201 (0.0229, 0.2173)	0.02	0.1822 (0.0622, 0.3022)	<0.01
Traffic $PM_{2.5}$*maternal asthma	-0.0088 (-0.0671, 0.0495)	0.77	-0.0292 (-0.1098, 0.0514)	0.48	-0.0672 (-0.1668, 0.0323)	0.19
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
Traffic $PM_{2.5}$	0.0684 (0.0130, 0.1238)	0.02	0.0401 (-0.0262, 0.1064)	0.24		
Maternal asthma	0.2622 (0.1158, 0.4086)	<0.01	0.1947 (0.0263, 0.3632)	0.02		
Traffic $PM_{2.5}$*maternal asthma	-0.1072 (-0.2302, 0.0158)	0.09	-0.0368 (-0.1783, 0.1047)	0.61		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value [a] Birth year dropped in order for model to converge

Table only includes model output relevant to interaction of interest. Models also control for child sex, child race, birth year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Reference group: no maternal asthma or unknown maternal asthma status. *Score statistic p-values for interaction terms: age 2 –did not converge, age 3 –0.52, age 4 –0.22, age 5 –0.09, age 6 –0.62*

Table B10. First year of life traffic $PM_{2.5}$ exposure and asthma incidence, stratified by city region

Cohort	Unadjusted models stratified by city region					
	Metro Atlanta		≤10 miles from metro Atlanta		>10 miles from metro Atlanta	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	P
Age 2	-0.0257 (-0.0892, 0.0379)	0.43	-0.0013 (-0.0341, 0.0315)	0.94	0.0267 (-0.0076, 0.0610)	0.13
Age 3	-0.1208 (-0.2020, -0.0396)	<0.01	-0.0020 (-0.0478, 0.0437)	0.93	0.0343 (-0.0127, 0.0814)	0.15
Age 4	-0.1247 (-0.2424, -0.0070)	0.04	-0.0020 (-0.0598, 0.0559)	0.95	0.0055 (-0.0543, 0.0652)	0.86
Age 5	-0.1568 (-0.2964, -0.0172)	0.03	0.0014 (-0.0696, 0.0724)	0.97	0.0726 (-0.0001, 0.1452)	0.05
Age 6	-0.2260 (-0.4113, -0.0406)	0.02	-0.0198 (-0.1048, 0.0653)	0.65	0.0850 (-0.0010, 0.1710)	0.05
Cohort	Final model stratified by city region					
	Metro Atlanta		≤10 miles from metro Atlanta		>10 miles from metro Atlanta	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	P
Age 2	0.0004 (-0.0629, 0.0636)[a]	0.99	0.0103 (-0.0267, 0.0473)	0.59	—	—
Age 3	—	—	0.0513 (-0.0005, 0.1031)	0.05	0.0469 (-0.0069, 0.1007)	0.09
Age 4	-0.0160 (-0.1228, 0.0908)[a]	0.77	0.0462 (-0.0211, 0.1136)	0.18	0.0286 (-0.0410, 0.0982)	0.42
Age 5	0.0193 (-0.1181, 0.1566)	0.78	0.0309 (-0.0521, 0.1138)	0.47	0.0828 (-0.0010, 0.1667)	0.05
Age 6	-0.1166 (-0.2591, 0.0260)	0.11	-0.0082 (-0.1067, 0.0903)	0.87	0.1014 (0.0034, 0.1993)	0.04

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, even if dropped birth year, [a] birth year dropped in order for model to converge

Adjusted models include all covariates in the final adjusted model: child sex, child race, maternal asthma, birth year, and minor demographic cluster. The following minor demographic clusters were combined in order to aid in convergence: A.2 and A.3; D.1 and D.3; D.5, D.6 and D.7.

Table B11. First year of life traffic $PM_{2.5}$ exposure and childhood asthma incidence, assessing interaction between traffic $PM_{2.5}$ and city region

Parameter	AGE 2		AGE 3		AGE 4	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Traffic $PM_{2.5}$	0.0451 (0.0091, 0.0811)	0.01	Did not converge, even after dropping birth year		0.0435 (-0.0211, 0.1080)	0.19
Metro Atlanta	0.0135 (-0.0855, 0.1124)	0.79			0.0304 (-0.1434, 0.2042)	0.73
≤10 mi from metro Atlanta	0.0156 (-0.0397, 0.0709)	0.58			-0.0269 (-0.1257, 0.0719)	0.59
Traffic $PM_{2.5}$*metro Atlanta	-0.0412 (-0.1097, 0.0272)	0.24			-0.0607 (-0.1812, 0.0598)	0.32
Traffic $PM_{2.5}$*≤10 mi from metro	-0.0300 (-0.0780, 0.0181)	0.22			0.0006 (-0.0852, 0.0864)	0.99
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
Traffic $PM_{2.5}$	0.0866 (0.0097, 0.1635)	0.03	0.0875 (-0.0035, 0.1784)	0.06		
Metro Atlanta	0.0688 (-0.1447, 0.2824)	0.53	0.1826 (-0.0860, 0.4512)	0.18		
≤10 mi from metro Atlanta	0.0072 (-0.1134, 0.1279)	0.91	0.0454 (-0.0981, 0.1888)	0.54		
Traffic $PM_{2.5}$*metro Atlanta	-0.1067 (-0.2553, 0.0419)	0.16	-0.1929 (-0.3786, -0.0072)	0.04		
Traffic $PM_{2.5}$*≤10 mi from metro	-0.0399 (-0.1448, 0.0651)	0.46	-0.0702 (-0.1952, 0.0548)	0.27		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value.

Table only includes model output relevant to interaction of interest. Models also control for child sex, child race, birth year, maternal asthma, and minor demographic cluster. Reference group: >10 miles from metro Atlanta. *Score statistic p-values for interaction terms: age 2 – did not converge, age 4 – 0.68, age 5 – 0.48, age 6 – 0.16*

Table B12. Sensitivity analyses of the association between first year of life traffic $PM_{2.5}$ exposure and asthma incidence

Cohort	Final adjusted model (for comparison)		Excluding children with unknown race		Excluding children missing ≥90 days of first year residence data	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0265 (0.0027, 0.0503)	0.03	—	—	0.0270 (0.0031, 0.0510)[a]	0.03
Age 3	0.0369 (0.0040, 0.0698)[a]	0.03	0.0350 (-0.0013, 0.0713)[a]	0.06	0.0334 (-0.0013, 0.0680)[a]	0.06
Age 4	0.0373 (-0.0073, 0.0819)	0.10	0.0370 (-0.0112, 0.0852)	0.13	0.0341 (-0.0120, 0.0803)	0.15
Age 5	0.0578 (0.0035, 0.1122)	0.04	0.0606 (0.0024, 0.1188)	0.04	0.0586 (0.0027, 0.1146)	0.04
Age 6	0.0359 (-0.0289, 0.1008)	0.28	0.0231 (-0.0457, 0.0918)	0.51	0.0296 (-0.0372, 0.0963)	0.39
Cohort	Excluding children with no maternal matches and unreliable maternal matches [b]		Excluding children not linked to birth certificates			
	RD (95% CI)	p	RD (95% CI)	p		
Age 2	—	—	0.0297 (0.0024, 0.0569)[a]	0.03		
Age 3	0.0359 (-0.0002, 0.0720)	0.05	0.0277 (-0.0114, 0.0668)[a]	0.17		
Age 4	0.0338 (-0.0146, 0.0823)	0.17	0.0325 (-0.0200, 0.0850)	0.22		
Age 5	0.0476 (-0.0110, 0.1062)	0.11	0.0474 (-0.0167, 0.1115)	0.15		
Age 6	0.0214 (-0.0479, 0.0906)	0.55	0.0220 (-0.0540, 0.0980)	0.58		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, even if dropped birth year, [a] birth year dropped in order for model to converge, [b] Maternal matches were considered unreliable if they were completed using incomplete medical record information and not confirmed by a birth certificate, or if birth certificates and medical records included discrepant maternal information

Models adjust for all covariates in the final adjusted model: child sex, maternal asthma, birth year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta).

Table B13. Extended Cox proportional hazards regression model results

Likelihood ratio tests				
Time dependent variables included in model	-2 log likelihood (with covariates)	χ^2	degrees of freedom	p
None (reduced model)	88102.453	—	—	—
1 sex*time variable	88094.387	8.066	1	<0.01
2 race*time variables	88087.208	15.245	2	<0.01
11 minor cluster*time variables	88089.154	13.299	11	0.27
10 birth year*time variables	88087.901	14.552	10	0.15
Hazard ratios for association between traffic $PM_{2.5}$ and asthma				
Model	Hazard Ratio (95% CI)			p
Adjusted	1.277 (1.081, 1.508)			<0.01
Adjusted + sex*time	1.276 (1.081, 1.507)			<0.01
Adjusted + race*time	1.274 (1.079, 1.505)			<0.01
Adjusted + sex*time + race*time	1.274 (1.079, 1.504)			<0.01

CI = Confidence Interval, p = p-value. Models adjust for child sex, child race, maternal asthma, birth year, minor demographic clusters, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Completed for the outcome of asthma incidence between the first and sixth birthdays (age 6 analysis).

Figure B1. Overlap between KAPPA residences during pregnancy and the first year of life and the RLINE $PM_{2.5}$ data region

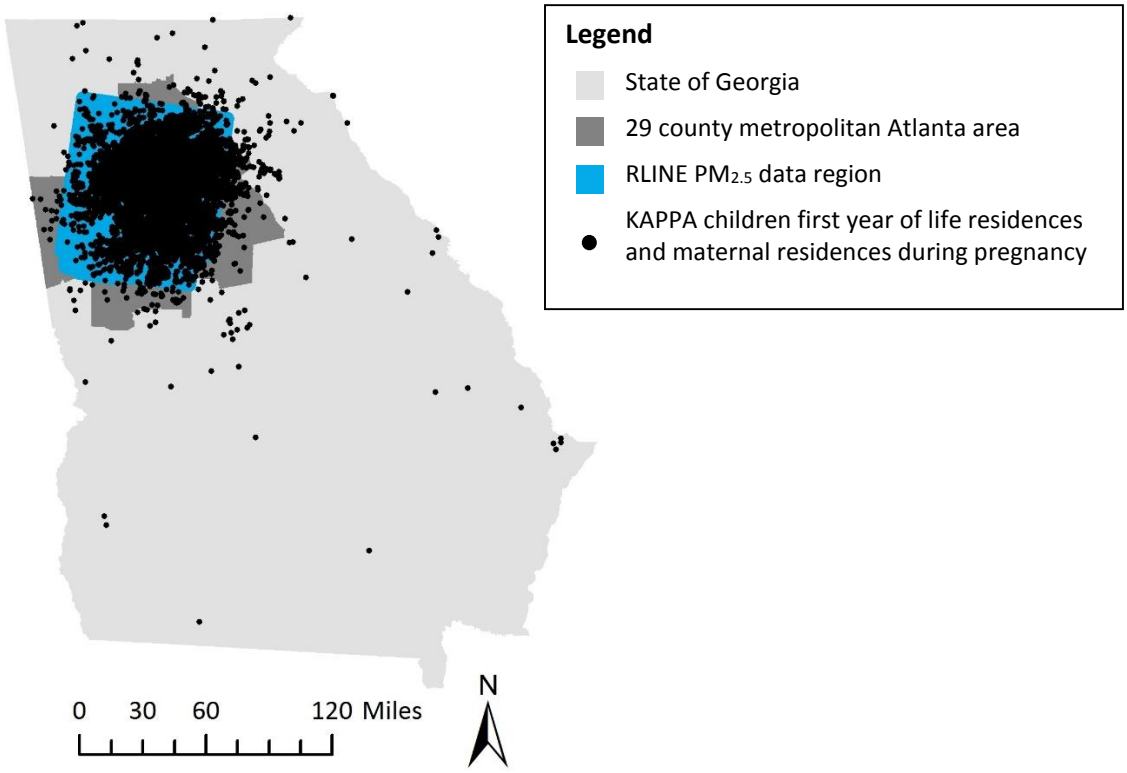


Figure B2. 29 county metropolitan Atlanta area divided into regions

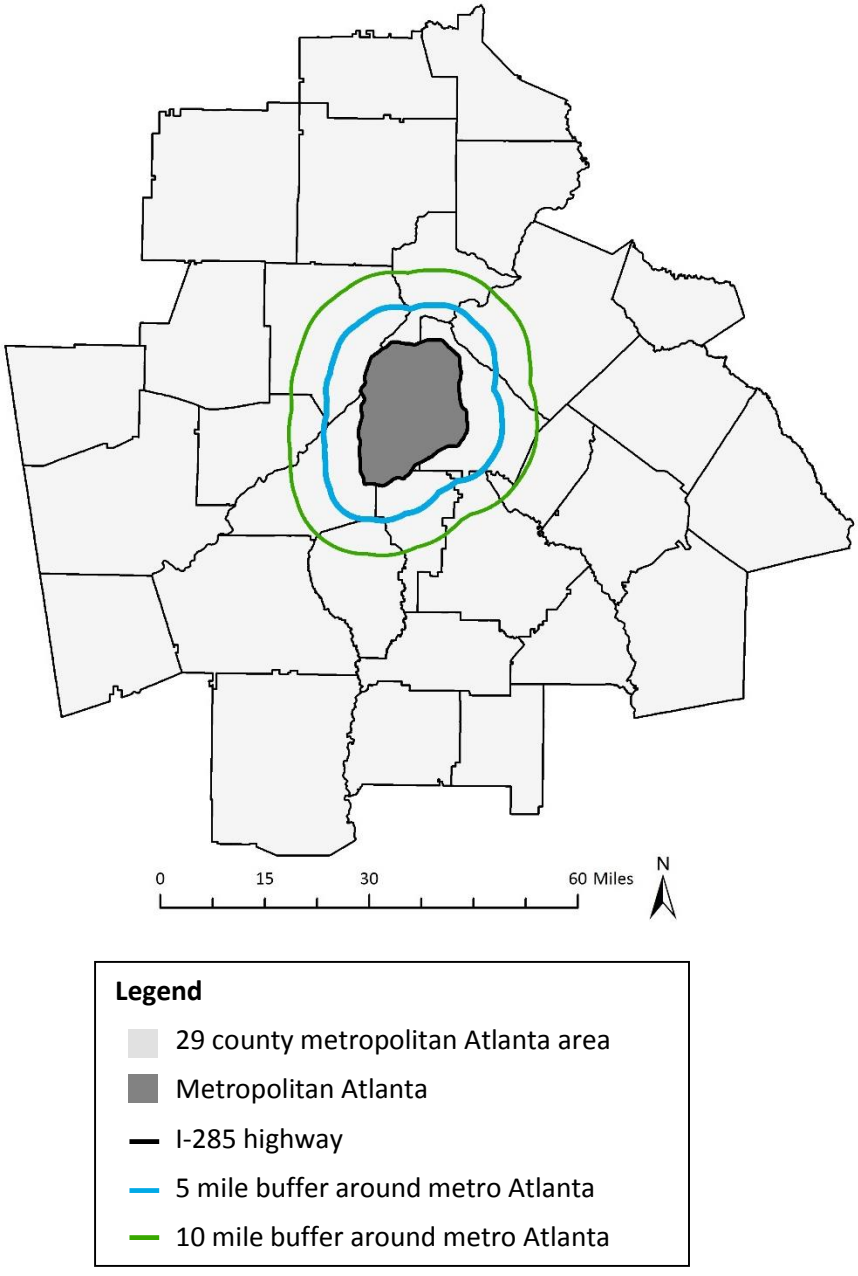


Figure B3. KAPPA residences at birth by race, restricted to the RLINE $PM_{2.5}$ data region

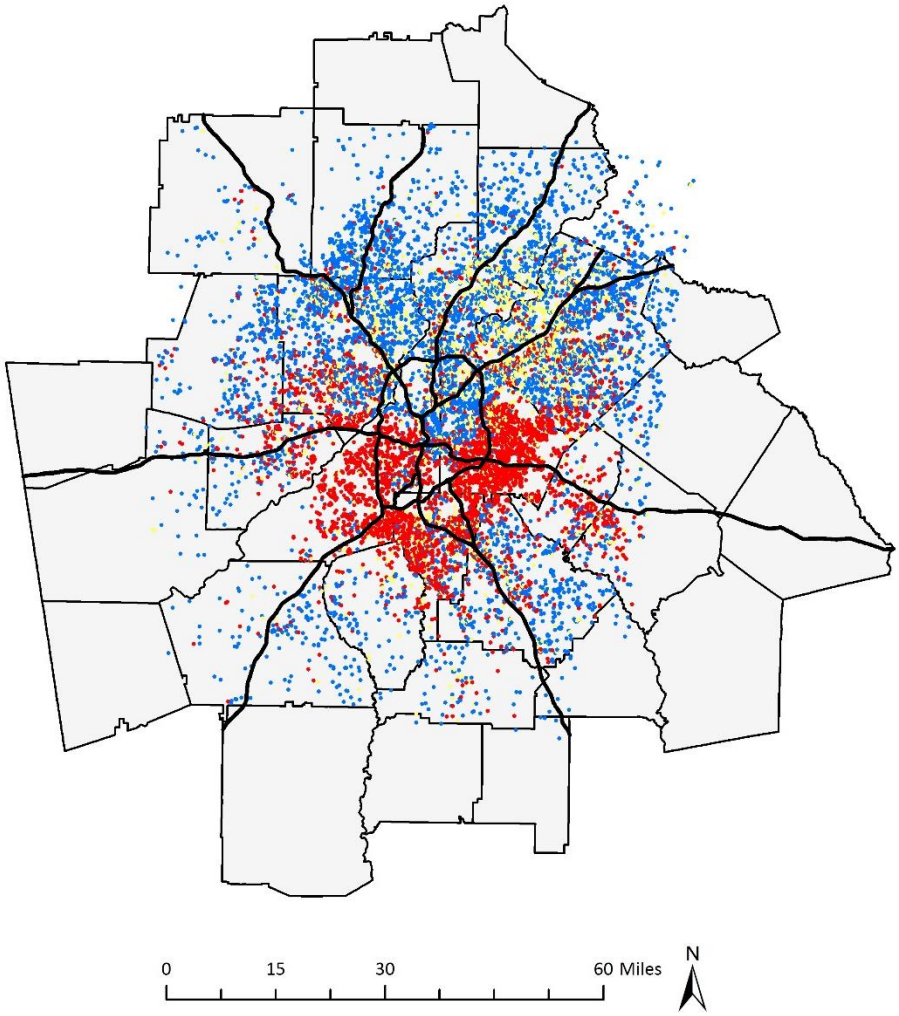


Figure B4. Asthma incidence survival curve for the KAPPA cohort

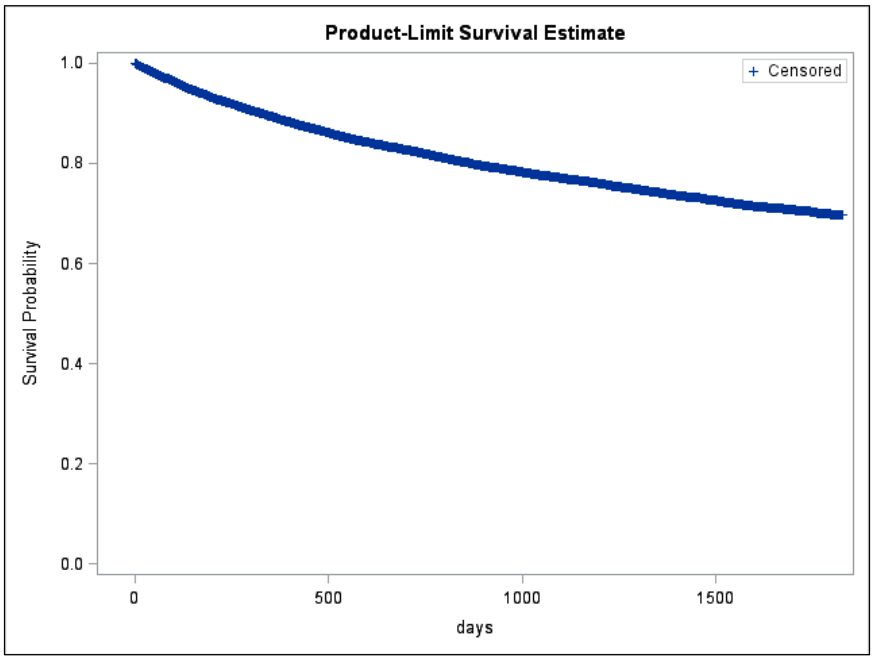


Figure B5 (Part 1 of 3). Graphical assessment of the proportional hazards assumption

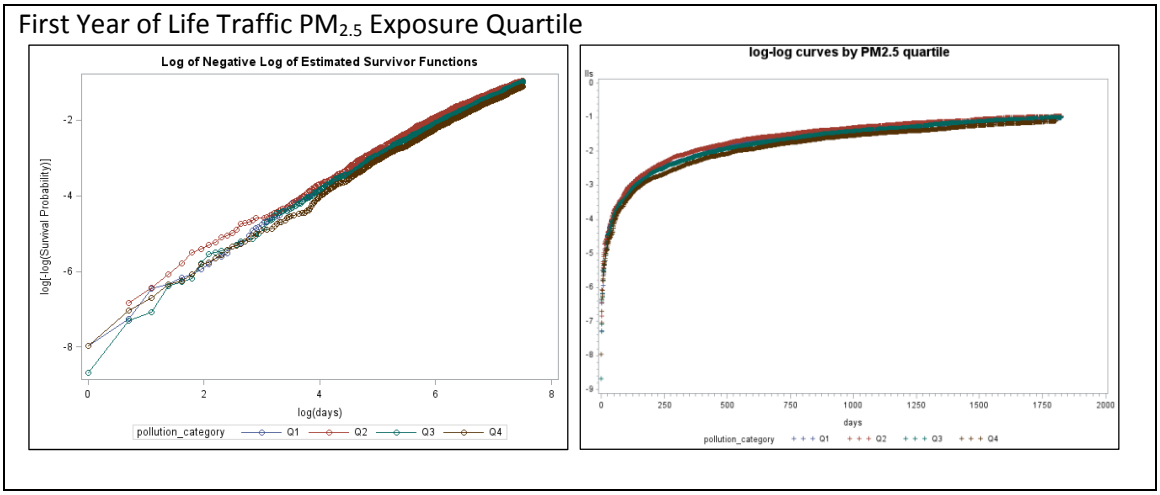


Figure B5 (Part 2 of 3). Graphical assessment of the proportional hazards assumption

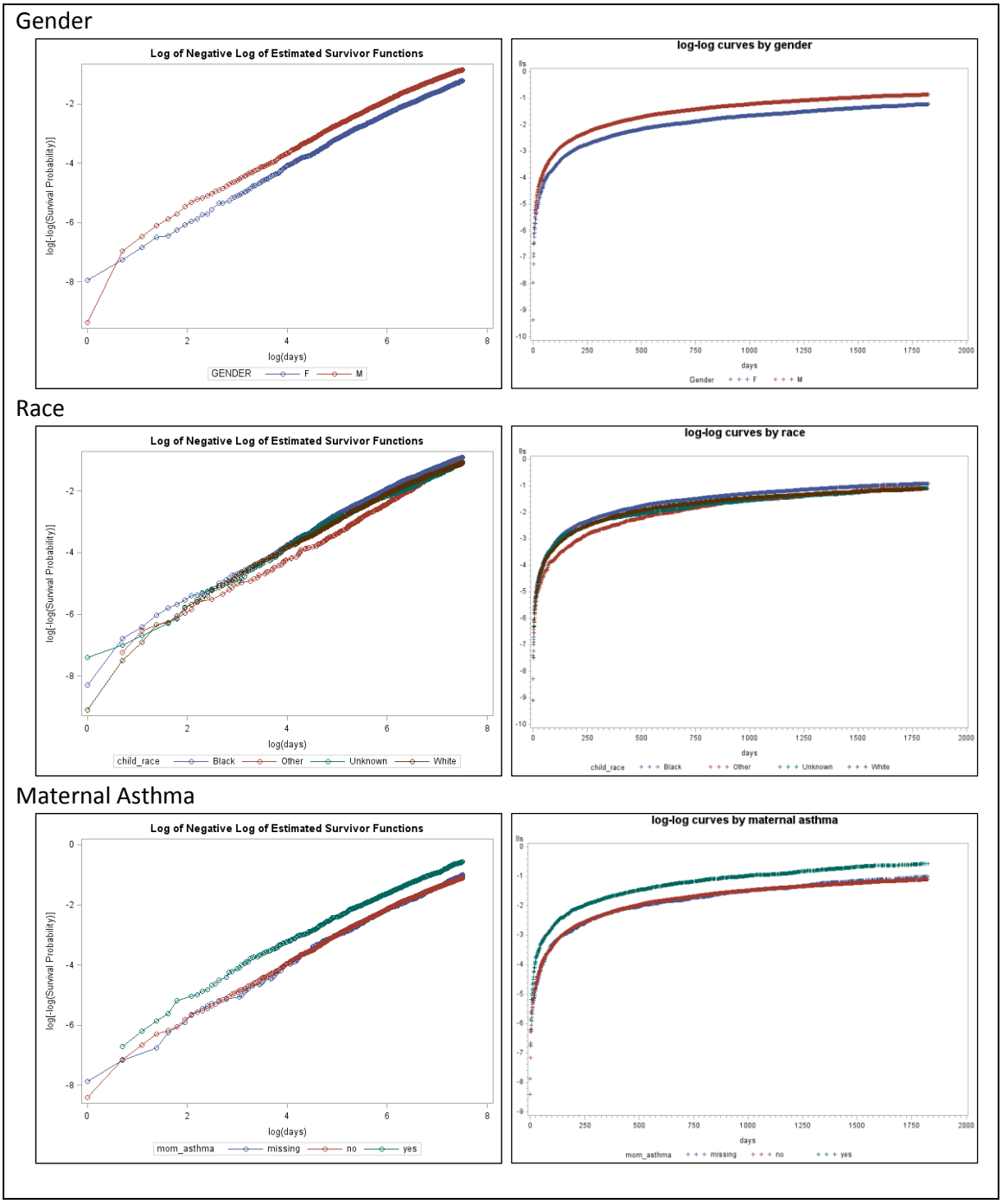
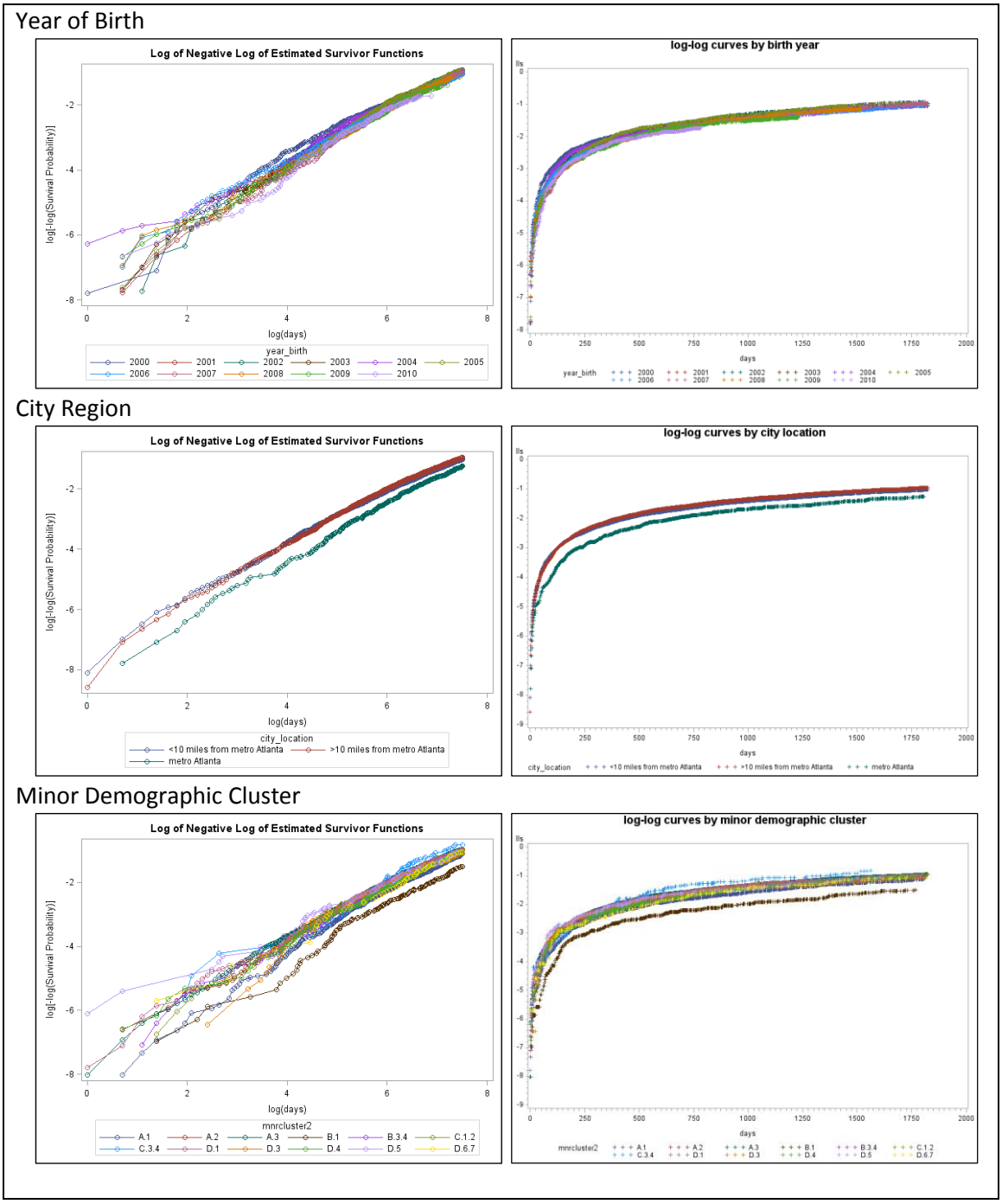


Figure B5 (Part 3 of 3). Graphical assessment of the proportional hazards assumption



APPENDIX C

Additional analyses of prenatal and first year of life exposure to total $PM_{2.5}$ and childhood asthma incidence in a birth cohort

Table C1. Comparison of covariates for children retained in cohort and lost to follow-up

Cohort	Children retained in cohort until follow-up age				Children eligible for follow-up age, but lost to follow-up before age			
Mean first year of life $PM_{2.5}$ [$\mu g/m^3$]								
Age 2	14.70				14.40			
Age 3	14.79				14.52			
Age 4	15.00				14.78			
Age 5	15.34				15.20			
Age 6	15.59				15.62			
Incident asthma at previous follow-up age [a] [%]								
Age 3	11.5				9.4			
Age 4	18.8				16.9			
Age 5	24.7				23.0			
Age 6	28.8				28.1			
Mean $PM_{2.5}$ stratified by asthma at previous age [a] [$\mu g/m^3$]								
	Asthma		No Asthma		Asthma		No Asthma	
Age 3	14.83		14.79		14.65		14.59	
Age 4	15.02		14.99		14.91		14.94	
Age 5	15.33		15.35		15.21		15.29	
Age 6	15.57		15.60		15.59		15.60	
Child race [%]								
	Black	White	Other	Unknown	Black	White	Other	Unknown
Age 2	36.9	39.7	11.5	11.9	26.6	37.0	13.7	22.8
Age 3	38.7	39.6	11.2	10.6	27.7	38.0	13.4	20.9
Age 4	40.1	39.7	10.9	9.4	28.2	37.8	12.7	21.4
Age 5	41.8	39.2	10.4	8.6	28.3	38.2	12.0	21.5
Age 6	43.2	39.0	10.0	7.9	28.8	38.6	11.0	21.6
Major demographic cluster [%]								
	A	B	C	D	A	B	C	D
Age 2	62.7	9.6	4.7	23.0	61.2	10.9	5.6	22.4
Age 3	63.0	9.2	4.7	23.2	61.4	10.9	5.3	22.3
Age 4	62.9	9.0	4.6	23.5	62.1	10.3	5.1	22.5
Age 5	62.7	8.9	4.7	23.7	62.4	9.7	5.2	22.7
Age 6	62.7	8.9	4.4	24.0	62.3	9.3	5.4	23.0
City region [%]								
	Metro Atlanta	≤10 mi from metro Atlanta	>10 mi from metro Atlanta		Metro Atlanta	≤10 mi from metro Atlanta	>10 mi from metro Atlanta	
Age 2	10.5	43.3	46.2		10.4	41.2	48.5	
Age 3	10.4	43.9	45.7		10.4	41.2	48.3	
Age 4	10.3	45.1	44.6		10.2	41.2	48.6	
Age 5	10.3	46.1	43.6		9.8	41.4	48.8	
Age 6	10.4	47.3	42.3		9.6	42.1	48.3	

To be eligible for each follow-up age a child has to be born early enough to be that age in September 2013 when the KAPPA cohort was defined. For example, a child born in 2010 could be at most 3 years old at the time KAPPA follow-up ended in September 2013. [a] Incident asthma at previous follow-up age only calculated among children in previous cohort (can't calculate for age 2 since no asthma classifications at age 1)

Table C2. Additional model results for the association between first year of life $PM_{2.5}$ exposure and asthma incidence

Cohort	MODEL A1		MODEL A2		MODEL A3	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Age 2	-0.0074 (-0.0155, 0.0008)[a]	0.08	0.0026 (0.0003, 0.0049)	0.02	0.0077 (-0.0017, 0.0171)	0.11
Age 3	-0.0123 (-0.0245, -0.0000)	0.05	0.0031 (-0.0001, 0.0062)	0.06	0.0072 (-0.0056, 0.0200)	0.27
Age 4	-0.0207 (-0.0370, -0.0044)	0.01	0.0050 (0.0005, 0.0096)	0.03	0.0044 (-0.0135, 0.0222)	0.63
Age 5	-0.0077 (-0.0280, 0.0125)	0.45	-0.0008 (-0.0092, 0.0075)	0.85	0.0176 (-0.0046, 0.0397)	0.12
Age 6	-0.0097 (-0.0356, 0.0162)	0.46	-0.0041 (-0.0203, 0.0120)	0.62	0.0185 (-0.0090, 0.0461)	0.19
Cohort	MODEL A4		MODEL A5			
	RD (95% CI)	p	RD (95% CI)	p		
Age 2	0.0083 (-0.0012, 0.0179)[a]	0.09	0.0079 (-0.0015, 0.0174)	0.10		
Age 3	0.0067 (-0.0063, 0.0196)[a]	0.31	0.0061 (-0.0068, 0.0190)	0.36		
Age 4	0.0048 (-0.0136, 0.0231)	0.61	0.0030 (-0.0153, 0.0214)	0.75		
Age 5	0.0187 (-0.0041, 0.0415)	0.11	0.0174 (-0.0054, 0.0402)	0.13		
Age 6	0.0183 (-0.0098, 0.0465)	0.20	—	—		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, [a]≤17 children deleted in order for model to converge

List of Model Covariates

Model A1: child sex, child race, child ethnicity, maternal asthma, cubic splines on date of birth with 1 knot per year, maternal age, maternal education, paternal education, maternal marital status *[Model 1 with paternal education, maternal marital status]*

Model A2: child sex, child race, maternal asthma, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta) *[Model 7 (final model) without cubic splines on date of birth, no temporal control]*

Model A3: child sex, child race, maternal asthma, minor demographic cluster, cubic splines on date of birth with 1 knot per year, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta) *[Model 7 (final model) without accounting for correlation between siblings, correlation structure (necessary to implement robust variance estimation) determined using child study id instead of family id]*

Model A4: child sex, child race, maternal asthma, minor demographic cluster, cubic splines on date of birth with 1 knot per year, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta), interaction between city region and race *[Model 7 (final model) with interaction term between city region and race] [Most Wald test p-values for interaction terms >0.05, score statistics did not converge]*

Model A5: child sex, child race, maternal asthma, minor demographic cluster, cubic splines on date of birth with 1 knot per year, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta), interaction between metro Atlanta and minor demographic cluster *[Model 7 (final model) with interaction term between metro Atlanta and minor demographic cluster. Did not include interaction terms with all city regions because of model convergence issues] [All Wald test p-values for interaction terms >0.05, score statistics did not converge]*

Table C3. First year of life $PM_{2.5}$ exposure and asthma incidence, stratified by race

Cohort	Unadjusted models stratified by race			
	Black		White	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0003 (-0.0040, 0.0046)	0.89	0.0020 (-0.0016, 0.0057)	0.27
Age 3	0.0030 (-0.0027, 0.0088)	0.30	0.0021 (-0.0032, 0.0074)	0.44
Age 4	0.0041 (-0.0036, 0.0118)	0.29	0.0021 (-0.0052, 0.0094)	0.57
Age 5	-0.0091 (-0.0225, 0.0044)	0.19	-0.0050 (-0.0181, 0.0081)	0.45
Age 6	0.0005 (-0.0236, 0.0245)	0.97	-0.0206 (-0.0441, 0.0028)	0.09
Cohort	Adjusted models stratified by race			
	Black		White	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0173 (0.0004, 0.0343)	0.05	—	—
Age 3	0.0028 (-0.0212, 0.0269)	0.82	0.0120 (-0.0066, 0.0306)[a]	0.21
Age 4	0.0001 (-0.0307, 0.0309)	1.00	0.0118 (-0.0139, 0.0374)	0.37
Age 5	0.0017 (-0.0360, 0.0394)	0.93	0.0224 (-0.0101, 0.0549)	0.18
Age 6	-0.0154 (-0.0598, 0.0291)	0.50	0.0418 (-0.0000, 0.0835)[a]	0.05

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, [a]<50 children deleted in order for model to converge

Adjusted models include all covariates in the final adjusted model: child sex, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). The following minor demographic clusters were combined in order to aid in convergence: A.2 and A.3; D.1 and D.3; D.5, D.6 and D.7.

Table C4. First year of life $PM_{2.5}$ exposure and asthma incidence, assessing interaction between $PM_{2.5}$ and race

Parameter	AGE 2[a]		AGE 3		AGE 4	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
$PM_{2.5}$	0.0078 (-0.0015, 0.0172)	0.10	0.0073 (-0.0058, 0.0205)	0.27	0.0038 (-0.0149, 0.0226)	0.69
Black race	0.0409 (-0.0380, 0.1199)	0.31	0.0223 (-0.0834, 0.1280)	0.68	0.0092 (-0.1408, 0.1593)	0.90
Unknown/other race	-0.0010 (-0.0790, 0.0770)	0.98	0.0725 (-0.0454, 0.1903)	0.23	0.0354 (-0.1479, 0.2187)	0.70
$PM_{2.5}$*black race	-0.0012 (-0.0066, 0.0041)	0.65	0.0003 (-0.0068, 0.0074)	0.93	0.0013 (-0.0087, 0.0112)	0.80
$PM_{2.5}$*unknown/other race	0.0002 (-0.0051, 0.0056)	0.93	-0.0051 (-0.0129, 0.0026)	0.19	-0.0023 (-0.0144, 0.0099)	0.71
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
$PM_{2.5}$	0.0192 (-0.0053, 0.0436)	0.13	0.0153 (-0.0187, 0.0494)	0.38		
Black race	0.1658 (-0.1153, 0.4470)	0.25	-0.0878 (-0.6086, 0.4330)	0.74		
Unknown/other race	-0.1552 (-0.4765, 0.1660)	0.34	0.2820 (-0.3761, 0.9401)	0.40		
$PM_{2.5}$*black race	-0.0073 (-0.0256, 0.0110)	0.43	0.0094 (-0.0240, 0.0428)	0.58		
$PM_{2.5}$*unknown/other race	0.0115 (-0.0095, 0.0324)	0.28	-0.0153 (-0.0574, 0.0268)	0.48		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] 1 child deleted from cohort in order for model to converge

Table only includes model output relevant to interaction of interest. Models also control for child sex, cubic splines on date of birth with 1 knot per year, maternal asthma, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Reference group: white race. *Score statistics did not converge for any of the models.*

Table C5. First year of life $PM_{2.5}$ exposure and asthma incidence, stratified by sex

Cohort	Unadjusted models stratified by sex			
	Male		Female	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0017 (-0.0019, 0.0054)	0.35	0.0010 (-0.0021, 0.0041)	0.53
Age 3	0.0034 (-0.0018, 0.0085)	0.20	-0.0007 (-0.0053, 0.0039)	0.77
Age 4	0.0045 (-0.0024, 0.0114)	0.20	0.0011 (-0.0052, 0.0074)	0.73
Age 5	-0.0068 (-0.0192, 0.0056)	0.28	-0.0017 (-0.0127, 0.0093)	0.76
Age 6	-0.0121 (-0.0348, 0.0106)	0.30	-0.0095 (-0.0297, 0.0106)	0.36
Cohort	Adjusted models stratified by sex			
	Male		Female	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0192 (0.0040, 0.0344)	0.01	0.0003 (-0.0109, 0.0115)[a]	0.96
Age 3	0.0193 (-0.0021, 0.0407)	0.08	-0.0010 (-0.0145, 0.0124)[a]	0.88
Age 4	0.0163 (-0.0111, 0.0438)	0.24	-0.0067 (-0.0299, 0.0165)	0.57
Age 5	0.0219 (-0.0124, 0.0562)	0.21	0.0097 (-0.0195, 0.0389)	0.52
Age 6	-0.0098 (-0.0518, 0.0321)	0.65	0.0402 (0.0040, 0.0765)	0.03

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a]<40 children deleted in order for model to converge

Adjusted models include all covariates in the final adjusted model: child race, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). The following minor demographic clusters were combined in order to aid in convergence: A.2 and A.3; D.1 and D.3; D.5, D.6 and D.7.

Table C6. First year of life $PM_{2.5}$ exposure and asthma incidence, assessing interaction between $PM_{2.5}$ and sex

Parameter	AGE 2		AGE 3		AGE 4	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
$PM_{2.5}$	0.0078 (-0.0015, 0.0172)	0.10	0.0058 (-0.0076, 0.0192)	0.39	0.0037 (-0.0150, 0.0225)	0.70
Male sex	0.0542 (-0.0118, 0.1203)	0.11	0.0460 (-0.0487, 0.1407)	0.34	0.0702 (-0.0644, 0.2047)	0.31
$PM_{2.5}$*male sex	-0.0007 (-0.0052, 0.0038)	0.75	0.0016 (-0.0047, 0.0080)	0.62	0.0005 (-0.0084, 0.0095)	0.91
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
$PM_{2.5}$	0.0207 (-0.0029, 0.0442)	0.09	0.0202 (-0.0104, 0.0508)	0.20		
Male sex	0.1841 (-0.0626, 0.4307)	0.14	0.1730 (-0.2869, 0.6330)	0.46		
$PM_{2.5}$*male sex	-0.0063 (-0.0223, 0.0097)	0.44	-0.0056 (-0.0350, 0.0239)	0.71		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value

Table only includes model output relevant to interaction of interest. Models also control for child race, cubic splines on date of birth with 1 knot per year, maternal asthma, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Reference group: female sex. *Score statistics did not converge for any of the models.*

Table C7. First year of life $PM_{2.5}$ exposure and asthma incidence, stratified by maternal asthma

Cohort	Unadjusted models stratified by maternal asthma			
	Maternal Asthma		No Maternal Asthma	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0079 (0.0001, 0.0157)	0.05	0.0002 (-0.0025, 0.0029)	0.89
Age 3	0.0050 (-0.0061, 0.0160)	0.38	-0.0003 (-0.0042, 0.0036)	0.87
Age 4	0.0051 (-0.0092, 0.0194)	0.48	0.0023 (-0.0030, 0.0076)	0.39
Age 5	-0.0078 (-0.0347, 0.0192)	0.57	-0.0039 (-0.0133, 0.0054)	0.41
Age 6	-0.0216 (-0.0687, 0.0254)	0.37	-0.0086 (-0.0256, 0.0084)	0.32
Cohort	Adjusted models stratified by maternal asthma			
	Maternal Asthma		No Maternal Asthma	
	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0352 (0.0074, 0.0629)[a]	0.01	0.0082 (-0.0022, 0.0187)	0.12
Age 3	0.0058 (-0.0401, 0.0517)[a]	0.80	0.0069 (-0.0079, 0.0217)	0.36
Age 4	0.0143 (-0.0412, 0.0697)	0.61	0.0009 (-0.0199, 0.0217)	0.93
Age 5	0.0266 (-0.0426, 0.0957)[a]	0.45	0.0161 (-0.0093, 0.0415)	0.22
Age 6	0.0093 (-0.0762, 0.0948)[a]	0.83	0.0162 (-0.0155, 0.0478)	0.32

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a]<10 children deleted in order for model to converge

Adjusted models include all covariates in the final adjusted model: child sex, child race, cubic splines on date of birth with 1 knot per year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). The following minor demographic clusters were combined in order to aid in convergence: A.2 and A.3; D.1 and D.3; D.5, D.6 and D.7.

Table C8. First year of life $PM_{2.5}$ exposure and asthma incidence, assessing interaction between $PM_{2.5}$ and maternal asthma

Parameter	AGE 2		AGE 3		AGE 4	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
$PM_{2.5}$	0.0072 (-0.0023, 0.0168)	0.14	0.0063 (-0.0067, 0.0194)	0.34	0.0040 (-0.0143, 0.0224)	0.67
Maternal asthma	-0.0330 (-0.1511, 0.0852)	0.58	0.0474 (-0.1239, 0.2188)	0.59	0.1065 (-0.1218, 0.3347)	0.36
$PM_{2.5}$*maternal asthma	0.0065 (-0.0016, 0.0145)	0.11	0.0027 (-0.0088, 0.0141)	0.65	-0.0002 (-0.0153, 0.0150)	0.98
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
$PM_{2.5}$	0.0182 (-0.0046, 0.0410)	0.12	0.0188 (-0.0097, 0.0473)	0.20		
Maternal asthma	0.2042 (-0.2252, 0.6336)	0.35	0.3455 (-0.4285, 1.1195)	0.38		
$PM_{2.5}$*maternal asthma	-0.0044 (-0.0322, 0.0235)	0.76	-0.0124 (-0.0620, 0.0372)	0.62		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value

Table only includes model output relevant to interaction of interest. Models also control for child sex, child race, cubic splines on date of birth with 1 knot per year, minor demographic cluster, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Reference group: no maternal asthma or unknown maternal asthma status. *Score statistics did not converge for any of the models.*

Table C9. First year of life $PM_{2.5}$ exposure and asthma incidence, stratified by city region

Cohort	Unadjusted models stratified by city region					
	Metro Atlanta		≤10 miles from metro Atlanta		>10 miles from metro Atlanta	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	P
Age 2	0.0090 (0.0029, 0.0151)	<0.01	0.0011 (-0.0029, 0.0051)	0.58	0.0022 (-0.0013, 0.0057)	0.22
Age 3	0.0105 (0.0009, 0.0200)	0.03	0.0044 (-0.0012, 0.0100)	0.12	0.0003 (-0.0048, 0.0053)	0.92
Age 4	0.0143 (0.0016, 0.0269)	0.03	0.0054 (-0.0022, 0.0130)	0.17	0.0014 (-0.0056, 0.0083)	0.70
Age 5	0.0034 (-0.0230, 0.0297)	0.80	0.0003 (-0.0133, 0.0139)	0.97	-0.0028 (-0.0153, 0.0097)	0.66
Age 6	-0.0150 (-0.0595, 0.0296)	0.51	-0.0042 (-0.0285, 0.0202)	0.74	-0.0022 (-0.0276, 0.0232)	0.86
Cohort	Final model stratified by city region					
	Metro Atlanta		≤10 miles from metro Atlanta		>10 miles from metro Atlanta	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	P
Age 2	—	—	0.0059 (-0.0098, 0.0216)	0.46	0.0183 (0.0039, 0.0328)[a]	0.01
Age 3	—	—	0.0051 (-0.0180, 0.0281)	0.67	0.0258 (0.0054, 0.0462)	0.01
Age 4	-0.0261 (-0.0753, 0.0230)[a]	0.30	-0.0033 (-0.0336, 0.0270)	0.83	0.0199 (-0.0077, 0.0474)	0.16
Age 5	0.0181 (-0.0424, 0.0786)[a]	0.56	-0.0054 (-0.0430, 0.0321)	0.78	0.0380 (0.0046, 0.0713)	0.03
Age 6	-0.0041 (-0.0841, 0.0759)[a]	0.92	-0.0160 (-0.0616, 0.0296)	0.49	0.0437 (0.0031, 0.0843)	0.03

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, "—" = did not converge, [a]<50 children deleted in order for model to converge

Adjusted models include all covariates in the final adjusted model: child sex, child race, maternal asthma, cubic splines on date of birth with 1 knot per year, and minor demographic cluster. The following minor demographic clusters were combined in order to aid in convergence: A.2 and A.3; D.1 and D.3; D.5, D.6 and D.7.

Table C10. First year of life $PM_{2.5}$ exposure and childhood asthma incidence, assessing interaction between $PM_{2.5}$ and city region

Parameter	AGE 2 [a]		AGE 3 [a]		AGE 4 [a]	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
$PM_{2.5}$	0.0080 (-0.0017, 0.0177)	0.10	0.0057 (-0.0079, 0.0193)	0.41	0.0023 (-0.0161, 0.0207)	0.81
Metro Atlanta	-0.1374 (-0.2396, -0.0353)	<0.01	-0.1491 (-0.2604, -0.0378)	<0.01	-0.1945 (-0.3436, -0.0454)	0.01
≤10 mi from metro Atlanta	-0.0149 (-0.0920, 0.0622)	0.71	-0.0773 (-0.1881, 0.0335)	0.17	-0.0662 (-0.2187, 0.0862)	0.39
$PM_{2.5}$*metro Atlanta	0.0067 (-0.0001, 0.0134)	0.05	0.0067 (-0.0006, 0.0139)	0.07	0.0096 (-0.0003, 0.0194)	0.06
$PM_{2.5}$*≤10 mi from metro	-0.0000 (-0.0052, 0.0051)	0.99	0.0041 (-0.0033, 0.0114)	0.28	0.0033 (-0.0068, 0.0133)	0.52
Parameter	AGE 5		AGE 6			
	RD (95% CI)	p	RD (95% CI)	p		
$PM_{2.5}$	0.0141 (-0.0096, 0.0378)	0.24	0.0112 (-0.0221, 0.0446)	0.51		
Metro Atlanta	-0.3315 (-0.7149, 0.0519)	0.09	-0.3059 (-1.0987, 0.4870)	0.45		
≤10 mi from metro Atlanta	-0.1470 (-0.4235, 0.1296)	0.30	-0.1911 (-0.7332, 0.3510)	0.49		
$PM_{2.5}$*metro Atlanta	0.0167 (-0.0076, 0.0409)	0.18	0.0143 (-0.0352, 0.0639)	0.57		
$PM_{2.5}$*≤10 mi from metro	0.0077 (-0.0103, 0.0256)	0.40	0.0106 (-0.0243, 0.0454)	0.55		

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] ≤36 children deleted in order for model to converge

Table only includes model output relevant to interaction of interest. Models also control for child sex, child race, cubic splines on date of birth with 1 knot per year, maternal asthma, and minor demographic cluster. Reference group: >10 miles from metro Atlanta. *Score statistics did not converge for any of the models.*

Table C11. Sensitivity analyses of the association between first year of life $PM_{2.5}$ exposure and asthma incidence

Cohort	Final adjusted model (for comparison)		Excluding children with unknown race		Excluding children missing ≥90 days of first year residence data	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0076 (-0.0019, 0.0170)	0.12	0.0118 (0.0020, 0.0216)[b]	0.02	0.0057 (-0.0037, 0.0151)[b]	0.24
Age 3	0.0066 (-0.0064, 0.0196)	0.32	0.0068 (-0.0070, 0.0207)[b]	0.33	0.0039 (-0.0095, 0.0174)	0.57
Age 4	0.0040 (-0.0143, 0.0223)	0.67	0.0041 (-0.0153, 0.0235)	0.68	-0.0001 (-0.0189, 0.0188)	1.00
Age 5	0.0178 (-0.0049, 0.0405)	0.12	0.0189 (-0.0051, 0.0430)	0.12	0.0146 (-0.0088, 0.0379)	0.22
Age 6	0.0176 (-0.0106, 0.0457)	0.22	0.0142 (-0.0155, 0.0439)	0.35	0.0104 (-0.0187, 0.0395)	0.48
Cohort	Excluding children with no maternal matches and unreliable maternal matches [a]		Excluding children with pollution data assigned from a different year		Excluding children not linked to birth certificates	
	RD (95% CI)	p	RD (95% CI)	p	RD (95% CI)	p
Age 2	0.0118 (0.0019, 0.0217)[b]	0.02	0.0095 (-0.0014, 0.0205)	0.09	0.0117 (0.0011, 0.0223)[b]	0.03
Age 3	0.0069 (-0.0064, 0.0202)	0.31	0.0086 (-0.0055, 0.0227)[b]	0.23	0.0069 (-0.0075, 0.0213)[b]	0.35
Age 4	0.0027 (-0.0170, 0.0225)	0.79	0.0022 (-0.0181, 0.0225)	0.83	0.0006 (-0.0204, 0.0215)	0.96
Age 5	0.0186 (-0.0056, 0.0428)	0.13	0.0105 (-0.0157, 0.0367)	0.43	0.0180 (-0.0079, 0.0440)	0.17
Age 6	0.0158 (-0.0140, 0.0457)	0.30	0.0041 (-0.0297, 0.0378)	0.81	0.0159 (-0.0164, 0.0482)	0.34

RD = Risk Difference for 1 $\mu g/m^3$, CI = Confidence Interval, p = p-value, [a] Maternal matches were considered unreliable if they were completed using incomplete medical record information and not confirmed by a birth certificate, or if birth certificates and medical records included discrepant maternal information, [b]≤26 children deleted in order for model to converge

Models adjust for all covariates in the final adjusted model: child sex, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic cluster, city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta).

Table C12. Extended Cox proportional hazards regression model results

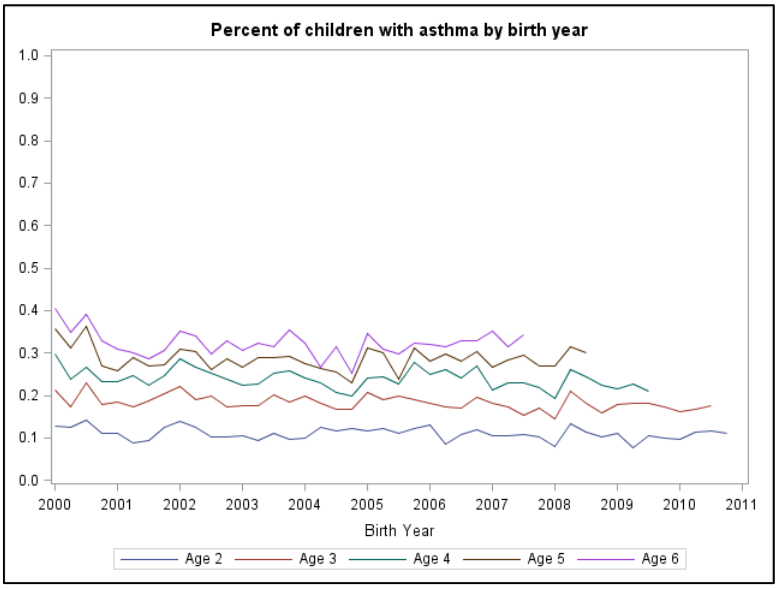
Likelihood ratio tests				
Time dependent variables included in model	-2 log likelihood (with covariates)	χ^2	degrees of freedom	p
None (reduced model)	88102.102	—	—	—
1 sex*time variable	88093.990	8.112	1	<0.01
2 race*time variables	88086.661	15.441	2	<0.01
11 minor cluster*time variables	88088.771	13.331	11	0.27
Hazard ratios for association between $PM_{2.5}$ and asthma				
Model	Hazard Ratio (95% CI)			p
Adjusted	1.090 (1.020, 1.165)			0.01
Adjusted + sex*time	1.091 (1.021, 1.166)			0.01
Adjusted + race*time	1.090 (1.020, 1.165)			0.01
Adjusted + sex*time + race*time	1.091 (1.021, 1.166)			0.01

CI = Confidence Interval, p = p-value. Models adjust for child sex, child race, maternal asthma, cubic splines on date of birth with 1 knot per year, minor demographic clusters, and city region (metro Atlanta, ≤10 miles from metro Atlanta, >10 miles from metro Atlanta). Completed for the outcome of asthma incidence between the first and sixth birthdays (age 6 analysis).

Figure C1. Percent of children classified as asthmatic by A) birth year B) conception year

Four points per year plotted one for each season. Order within each year: winter, spring, summer, fall

A.



B.

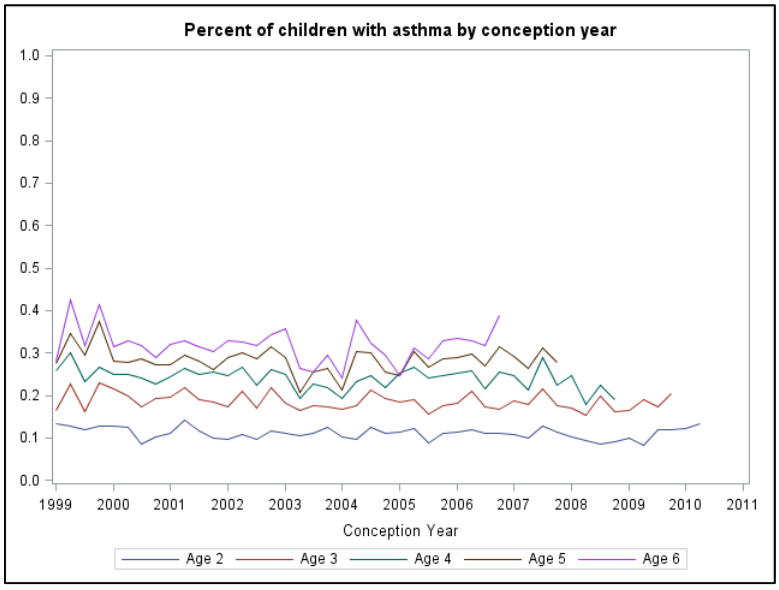
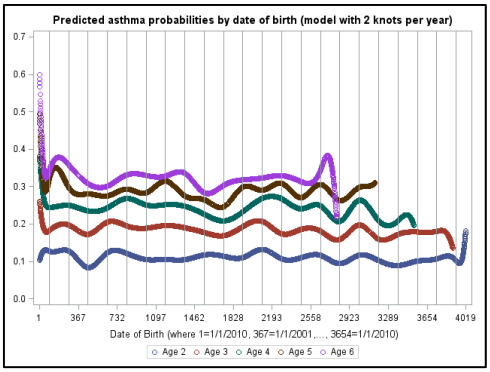
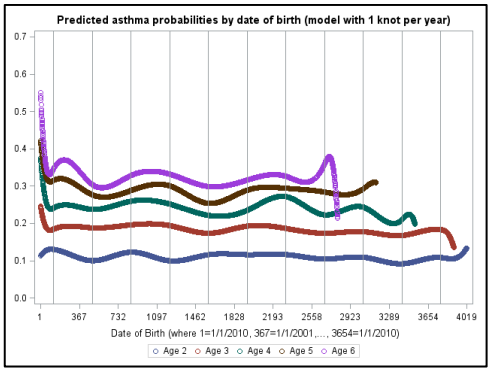


Figure C2. Plots of predicted probability of asthma incidence by time, modeled using different types of time control with cubic splines: A) splines on date of birth with 2 knots per year in April and October B) splines on date of birth with 1 knot per year in May C) splines on date of conception with 2 knots per year in April and October D) splines on date of conception with 1 knot per year in May

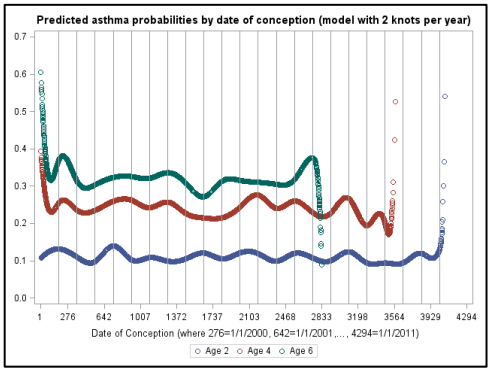
A.



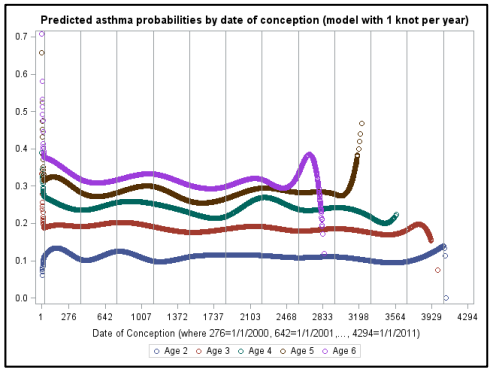
B.



C.

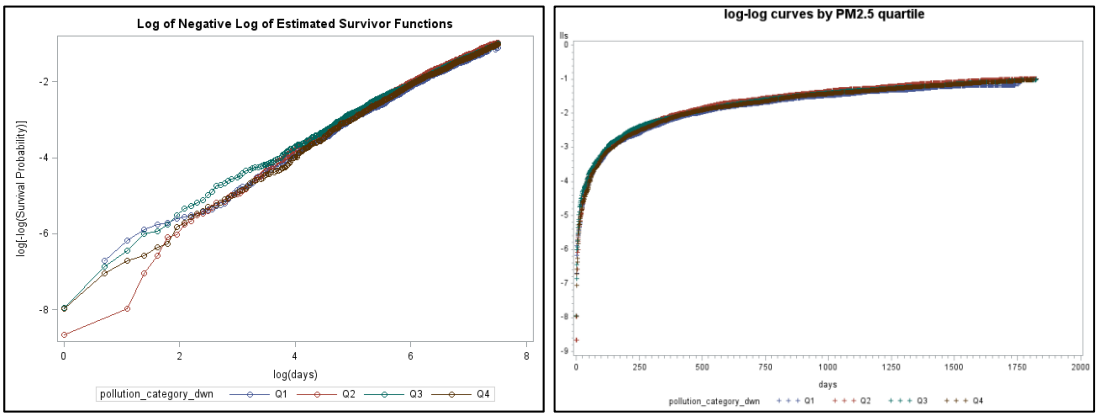


D.



Vertical lines indicate knot locations. Age 3 and 5 models did not converge when using cubic splines for date of conception with knots in April and October.

Figure C3. Graphical assessment of the proportional hazards assumption for first year of life $PM_{2.5}$ exposure quartile



CHAPTER 8

References

Akinbami L, Moorman J, Bailey C, et al. (2012). Trends in Asthma Prevalence, Health Care Use, and Mortality in the United States, 2001–2010. NCHS Data Brief. U.S. Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics.

Akinbami LJ, LaFleur BJ and Schoendorf KC (2002). "Racial and income disparities in childhood asthma in the United States." *Ambul Pediatr* 2(5): 382-387.

Akinbami LJ, Moorman JE, Garbe PL and Sondik EJ (2009). "Status of childhood asthma in the United States, 1980–2007." *Pediatrics* 123(Supplement 3): S131-S145.

Alhanti BA, Chang HH, Winquist A, Mulholland JA, Darrow LA and Sarnat SE (2015). "Ambient air pollution and emergency department visits for asthma: a multi-city assessment of effect modification by age." *J Expo Sci Environ Epidemiol*.

Almqvist C, Worm M and Leynaert B (2008). "Impact of gender on asthma in childhood and adolescence: a GA2LEN review." *Allergy* 63(1): 47-57.

American Academy of Pediatrics (1999). Handbook of Pediatric Environmental Health, American Academy of Pediatrics.

Anandan C, Nurmatov U, Van Schayck OCP and Sheikh A (2010). "Is the prevalence of asthma declining? Systematic review of epidemiological studies." *Allergy* 65(2): 152-167.

Anderson HR, Favarato G and Atkinson R (2013). "Long-term exposure to air pollution and the incidence of asthma: meta-analysis of cohort studies." *Air Quality, Atmosphere & Health* 6(1): 47-56.

Annesi-Maesano I, Moreau D, Caillaud D, et al. (2007). "Residential proximity fine particles related to allergic sensitisation and asthma in primary school children." *Respir Med* 101(8): 1721-1729.

Arbes Jr SJ, Gergen PJ, Vaughn B and Zeldin DC (2007). "Asthma cases attributable to atopy: Results from the Third National Health and Nutrition Examination Survey." *Journal of Allergy and Clinical Immunology* 120(5): 1139-1145.

Austin PC (2010). "Absolute risk reductions, relative risks, relative risk reductions, and numbers needed to treat can be obtained from a logistic regression model." *J Clin Epidemiol* 63(1): 2-6.

Barraza-Villarreal A, Escamilla-Nunez MC, Hernandez-Cadena L, et al. (2011). "Elemental carbon exposure and lung function in school children from Mexico City." *Eur Respir J* 38(3): 548-552.

Bates JT, Flak AL, Darrow LA, et al. (2016). "Comparison of statistical and data fusion techniques used to obtain spatially resolved air pollutant concentrations." *Manuscript in preparation*.

Baxi SN and Phipatanakul W (2010). "The role of allergen exposure and avoidance in asthma." *Adolesc Med State Art Rev* 21(1): 57-71, viii-ix.

Bell ML (2012). Assessment of the Health Impacts of Particulate Matter Characteristics. Research Report 161. Boston, MA, Health Effects Institute.

Bell ML, Ebisu K, Peng RD, Samet JM and Dominici F (2009). "Hospital admissions and chemical composition of fine particle air pollution." *Am J Respir Crit Care Med* 179(12): 1115-1120.

Bisgaard H, Jensen SM and Bonnelykke K (2012). "Interaction between asthma and lung function growth in early life." *Am J Respir Crit Care Med* 185(11): 1183-1189.

Black MH, Zhou H, Takayanagi M, Jacobsen SJ and Koebnick C (2013). "Increased asthma risk and asthma-related health care complications associated with childhood obesity." *Am J Epidemiol* 178(7): 1120-1128.

Bowatte G, Lodge C, Lowe AJ, et al. (2015). "The influence of childhood traffic-related air pollution exposure on asthma, allergy and sensitization: a systematic review and a meta-analysis of birth cohort studies." *Allergy* 70(3): 245-256.

Bråbäck L and Forsberg B (2009). "Does traffic exhaust contribute to the development of asthma and allergic sensitization in children: findings from recent cohort studies." *Environ Health* 8: 17.

Bracken MB, Belanger K, Cookson WO, Triche E, Christiani DC and Leaderer BP (2002). "Genetic and perinatal risk factors for asthma onset and severity: a review and theoretical analysis." *Epidemiol Rev* 24(2): 176-189.

Brand PL (2011). "The Asthma Predictive Index: not a useful tool in clinical practice." *J Allergy Clin Immunol* 127(1): 293-294.

Brand PL, Baraldi E, Bisgaard H, et al. (2008). "Definition, assessment and treatment of wheezing disorders in preschool children: an evidence-based approach." *The European respiratory journal: official journal of the European Society for Clinical Respiratory Physiology* 32(4): 1096-1110.

Brauer M, Hoek G, Smit HA, et al. (2007). "Air pollution and development of asthma, allergy and infections in a birth cohort." *The European respiratory journal: official journal of the European Society for Clinical Respiratory Physiology* 29(5): 879-888.

Brooks C, Pearce N and Douwes J (2013). "The hygiene hypothesis in allergy and asthma: an update." *Curr Opin Allergy Clin Immunol* 13(1): 70-77.

Buescher PA and Jones-Vessey K (1999). "Using Medicaid data to estimate state- and county-level prevalence of asthma among low-income children." *Matern Child Health J* 3(4): 211-216.

Burra TA, Moineddin R, Agha MM and Glazier RH (2009). "Social disadvantage, air pollution, and asthma physician visits in Toronto, Canada." *Environ Res* 109(5): 567-574.

Bush A (2007). "Diagnosis of asthma in children under five." *Prim Care Respir J* 16(1): 7-15.

Byun D and Schere K (2006). "Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system." *Applied Mechanics Review* 59: 51-77.

Byun DW, Ching JKS, U. S. Environmental Protection Agency and Office of Research and Development (1999). Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. EPA/600/R-99/030. Washington D.C.

Cabana MD, Lara M and Shannon J (2007). "Racial and ethnic disparities in the quality of asthma care." *Chest* 132(5 Suppl): 810S-817S.

Carlsten C, Dybuncio A, Becker A, Chan-Yeung M and Brauer M (2011). "Traffic-related air pollution and incident asthma in a high-risk birth cohort." *Occup Environ Med* 68(4): 291-295.

Castro-Rodriguez J, Holberg CJ, Wright AL and Martinez FD (2000). "A clinical index to define risk of asthma in young children with recurrent wheezing." *Am J Respir Crit Care Med* 162(4): 1403-1406.

Castro-Rodriguez JA (2010). "The Asthma Predictive Index: a very useful tool for predicting asthma in young children." *J Allergy Clin Immunol* 126(2): 212-216.

Castro-Rodriguez JA (2011). "The Asthma Predictive Index: early diagnosis of asthma." *Curr Opin Allergy Clin Immunol* 11(3): 157-161.

Castro-Rodriguez JA, Cifuentes L and Rodriguez-Martinez CE (2011). "The asthma predictive index remains a useful tool to predict asthma in young children with recurrent wheeze in clinical practice." *J Allergy Clin Immunol* 127(4): 1082-1083.

Caudri D, Wijga A, A Schipper CM, et al. (2009). "Predicting the long-term prognosis of children with symptoms suggestive of asthma at preschool age." *J Allergy Clin Immunol* 124(5): 903-910.

Celedón JC, Fuhlbrigge A, Rifas-Shiman S, Weiss ST and Finkelstein JA (2004). "Antibiotic use in the first year of life and asthma in early childhood." *Clin Exp Allergy* 34(7): 1011-1016.

Cho CE and Norman M (2013). "Cesarean section and development of the immune system in the offspring." *American Journal of Obstetrics and Gynecology* 208(4): 249-254.

Clark NA, Demers PA, Karr CJ, et al. (2010). "Effect of early life exposure to air pollution on development of childhood asthma." *Environmental health perspectives* 118(2).

Cockcroft DW and Swystun VA (1996). "Asthma control versus asthma severity." *Journal of Allergy and Clinical Immunology* 98(6, Part 1): 1016-1018.

Coker TR, Kaplan RM and Chung PJ (2012). "The association of health insurance and disease impairment with reported asthma prevalence in U.S. children." *Health Serv Res* 47(1 Pt 2): 431-445.

Community Modeling and Analysis System. (2015). "R-LINE: A Research LINE-source dispersion model for near-surface releases." Retrieved September 28, 2015, from https://www.cmascenter.org/r-line/.

Crocker D, Brown C, Moolenaar R, et al. (2009). "Racial and ethnic disparities in asthma medication usage and health-care utilization: data from the National Asthma Survey." *Chest* 136(4): 1063-1071.

Darrow LA, Klein M, Strickland MJ, Mulholland JA and Tolbert PE (2011). "Ambient air pollution and birth weight in full-term infants in Atlanta, 1994-2004." *Environ Health Perspect* 119(5): 731-737.

Dawood FS, Kamimoto L, D'Mello TA, et al. (2011). "Children with asthma hospitalized with seasonal or pandemic influenza, 2003-2009." *Pediatrics* 128(1): e27-32.

Deddens JA and Petersen MR (2008). "Approaches for estimating prevalence ratios." *Occup Environ Med* 65(7): 481, 501-486.

Demographic Clusters of Georgia (2012). Georgia Department of Public Health: Office of Health Indicators for Planning (OHIP).

Devulapalli CS, Carlsen KC, Haland G, et al. (2008). "Severity of obstructive airways disease by age 2 years predicts asthma at 10 years of age." *Thorax* 63(1): 8-13.

Dockery DW, Pope CA, Xu X, et al. (1993). "An association between air pollution and mortality in six U.S. cities." *New England Journal of Medicine* 329(24): 1753-1759.

Dombkowski KJ, Wasilevich EA and Lyon-Callo SK (2005). "Pediatric asthma surveillance using Medicaid claims." *Public Health Rep* 120(5): 515-524.

Dominici F, Peng RD, Barr CD and Bell ML (2010). "Protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach." *Epidemiology* 21(2): 187-194.

Douwes J, Gibson P, Pekkanen J and Pearce N (2002). "Non-eosinophilic asthma: importance and possible mechanisms." *Thorax* 57(7): 643-648.

Drews AC, Pizzichini MM, Pizzichini E, et al. (2009). "Neutrophilic airway inflammation is a main feature of induced sputum in nonatopic asthmatic children." *Allergy* 64(11): 1597-1601.

Duijts L (2012). "Fetal and infant origins of asthma." *European journal of epidemiology* 27(1): 5-14.

Durrani SR, Viswanathan RK and Busse WW (2011). "What effect does asthma treatment have on airway remodeling? Current perspectives." *J Allergy Clin Immunol* 128(3): 439-448.

Eigen H (2008). "Differential diagnosis and treatment of wheezing and asthma in young children." *Clin Pediatr* 47(8): 735-743.

Fan J, Li S, Fan C, Bai Z and Yang K (2015). "The impact of PM2.5 on asthma emergency department visits: a systematic review and meta-analysis." *Environ Sci Pollut Res Int*.

Finkelstein JA, Barton MB, Donahue JG, Algatt-Bergstrom P, Markson LE and Platt R (2000). "Comparing asthma care for Medicaid and non-Medicaid children in a health maintenance organization." *Arch Pediatr Adolesc Med* 154(6): 563-568.

Firoozi F, Lemiere C, Ducharme FM, et al. (2010). "Effect of maternal moderate to severe asthma on perinatal outcomes." *Respir Med* 104(9): 1278-1287.

Fouzas S and Brand PL (2013). "Predicting persistence of asthma in preschool wheezers: crystal balls or muddy waters?" *Paediatr Respir Rev* 14(1): 46-52.

Friberg MD, Zhai X, Holmes HA, et al. (2015). "Fusion of observational data and chemical transport model simulations for estimating spatiotemporally-resolved ambient air quality concentrations for use in health studies." *Submitted for publication*.

Gaffin JM and Phipatanakul W (2009). "The role of indoor allergens in the development of asthma." *Curr Opin Allergy Clin Immunol* 9(2): 128-135.

Gallagher EJ (1998). "Clinical utility of likelihood ratios." *Ann Emerg Med* 31(3): 391-397.

Garshick E, Laden F, Hart JE, Davis ME, Eisen EA and Smith TJ (2012). "Lung cancer and elemental carbon exposure in trucking industry workers." *Environ Health Perspect* 120(9): 1301-1306.

Gasana J, Dillikar D, Mendy A, Forno E and Ramos Vieira E (2012). "Motor vehicle air pollution and asthma in children: A meta-analysis." *Environmental research* 117(0): 36-45.

Gass K, Balachandran S, Chang HH, Russell AG and Strickland MJ (2015). "Ensemble-based source apportionment of fine particulate matter and emergency department visits for pediatric asthma." *Am J Epidemiol* 181(7): 504-512.

Gdalevich M, Mimouni D and Mimouni M (2001). "Breast-feeding and the risk of bronchial asthma in childhood: a systematic review with meta-analysis of prospective studies." *J Pediatr* 139(2): 261-266.

Gehring U, Wijga AH, Brauer M, et al. (2010). "Traffic-related air pollution and the development of asthma and allergies during the first 8 years of life." *Am J Respir Crit Care Med* 181(6): 596-603.

Georgia Department of Public Health and Office of Health Indicators for Planning (OHIP). (2013-2015). "Online Analytical Statistical Information System: Demographic Clusters of Georgia: Accessing the Georgia Deparmtent of Public Health's Data Warehouse." Retrieved 11/15/2015, from https://oasis.state.ga.us/gis/demographiccluster/DemoClusters2011.htm.

Gergen PJ, Arbes SJ, Jr., Calatroni A, Mitchell HE and Zeldin DC (2009). "Total IgE levels and asthma prevalence in the US population: results from the National Health and Nutrition Examination Survey 2005-2006." *J Allergy Clin Immunol* 124(3): 447-453.

Gershwin ME and Albertson TE (2012). "The practical understanding and treatment of asthma." *Clin Rev Allergy Immunol* 43: 1-2.

Getahun D, Strickland D, Zeiger RS, et al. (2010). "Effect of chorioamnionitis on early childhood asthma." *Arch Pediatr Adolesc Med* 164(2): 187-192.

Gibson PG, Henry RL and Thomas P (2000). "Noninvasive assessment of airway inflammation in children: induced sputum, exhaled nitric oxide, and breath condensate." *The European respiratory journal: official journal of the European Society for Clinical Respiratory Physiology* 16(5): 1008-1015.

Global Initiative for Asthma (GINA) (2009). Global Strategy for the Diagnosis and Management of Asthma in Children 5 Years and Younger. Available from: http://www.ginasthma.org/.

Glymour MM and Greenland S (2008). Causal Diagrams. Modern Epidemiology. Editors: Rothman KJ, Greenland S and Lash TL. Philadelphia, Lippincott Williams & Wilkins**:** 199-200.

Godish T (2004). Chapter 2: Atmospheric Pollution and Pollutants. Air Quality. United States, Lewis Publishers**:** 23-69.

Godish T (2004). Chapter 5: Health Effects. Air Quality. United States, Lewis Publishers**:** 141-182.

Gold DR and Wright R (2005). "Population disparities in asthma." *Annual Review of Public Health* 26(1): 89-113.

Gold J, Reyes-Gastelum D, Turner J and Davies HD (2014). "A quality improvement study using fishbone analysis and an electronic medical records intervention to improve care for children with asthma." *Am J Med Qual* 29(1): 70-77.

Gordian ME, Haneuse S and Wakefield J (2006). "An investigation of the association between traffic exposure and the diagnosis of asthma in children." *J Expo Sci Environ Epidemiol* 16(1): 49-55.

Gowers AM, Cullinan P, Ayres JG, et al. (2012). "Does outdoor air pollution induce new cases of asthma? Biological plausibility and evidence; a review." *Respirology* 17(6): 887-898.

Goyal NK, Fiks AG and Lorch SA (2011). "Association of late-preterm birth with asthma in young children: practice-based study." *Pediatrics* 128(4): e830-838.

Grana J, Preston S, McDermott PD and Hanchak NA (1997). "The use of administrative data to risk-stratify asthmatic patients." *Am J Med Qual* 12(2): 113-119.

Greenland S, Lash TL and Rothman KJ (2008). Concepts of Interaction. Modern Epidemiology. Editors: Rothman KJ, Greenland S and Lash TL. Philadelphia, Lippincott Williams & Wilkins**:** 71-83.

Greenland S, Rothman KJ and Lash TL (2008). Measures of Effect and Measures of Association. Modern Epidemiology. Editors: Rothman KJ, Greenland S and Lash TL. Philadelphia, Lippincott Williams & Wilkins**:** 60-61.

Greenland S, Rothman KJ and Lash TL (2008). Validity in Epidemiologic Studies. Modern Epidemiology. Editors: Rothman KJ, Greenland S and Lash TL. Philadelphia, Lippincott Williams & Wilkins**:** 128-147.

Greenland S and Finkle WD (1995). "A critical look at methods for handling missing covariates in epidemiologic regression analyses." *Am J Epidemiol* 142(12): 1255-1264.

Grol MH, Gerritsen J, Vonk JM, et al. (1999). "Risk factors for growth and decline of lung function in asthmatic individuals up to age 42 years. A 30-year follow-up study." *American journal of respiratory and critical care medicine* 160(6): 1830-1837.

Gruchalla RS, Pongracic J, Plaut M, et al. (2005). "Inner City Asthma Study: Relationships among sensitivity, allergen exposure, and asthma morbidity." *Journal of Allergy and Clinical Immunology* 115(3): 478-485.

Gruzieva O, Bergstrom A, Hulchiy O, et al. (2013). "Exposure to air pollution from traffic and childhood asthma until 12 years of age." *Epidemiology* 24(1): 54-61.

Hafkamp-de Groen E, Lingsma HF, Caudri D, et al. (2012). "Predicting asthma in preschool children with asthma symptoms: study rationale and design." *BMC Pulm Med* 12: 65.

Hao H, Chang HH, Holmes HA, et al. (2015). "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." *Environ Health Perspect*.

Harrison RM (1999). Measurements of Concentrations of Air Pollutants. Air Pollution and Health. S. T. Holgate, J. M. Samet, H. S. Koren and R. L. Maynard. London, Academic Press**:** 63-81.

Hedlin G, Konradsen J and Bush A (2012). "An update on paediatric asthma." *Eur Respir Rev* 21(125): 175-185.

HEI Panel on the Health Effects of Traffic-Related Air Pollution (2010). Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. HEI Special Report 17. Boston, Mass, Health Effects Institute.

Hirsch T, Weiland SK, von Mutius E, et al. (1999). "Inner city air pollution and respiratory health and atopy in children." *Eur Respir J* 14(3): 669-677.

Holt PG, Upham JW and Sly PD (2005). "Contemporaneous maturation of immunologic and respiratory functions during early childhood: Implications for development of asthma prevention strategies." *Journal of Allergy and Clinical Immunology* 116(1): 16-24.

Hovland V, Riiser A, Mowinckel P, Carlsen KH and Carlsen KC (2012). "The significance of early recurrent wheeze for asthma outcomes in late childhood." *The European respiratory journal: official journal of the European Society for Clinical Respiratory Physiology*.

Iniguez C, Ballester F, Estarlich M, et al. (2009). "Estimation of personal NO2 exposure in a cohort of pregnant women." *Sci Total Environ* 407(23): 6093-6099.

Ito K, Mathes R, Ross Z, Nadas A, Thurston G and Matte T (2011). "Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City." *Environ Health Perspect* 119(4): 467-473.

Jerrett M, Shankardass K, Berhane K, et al. (2008). "Traffic-related air pollution and asthma onset in children: a prospective cohort study with individual exposure measurement." *Environmental health perspectives* 116(10): 1433-1438.

Juhn Y, Kung A, Voigt R and Johnson S (2011). "Characterisation of children's asthma status by ICD-9 code and criteria-based medical record review." *Prim Care Respir J* 20(1): 79-83.

Kajekar R (2007). "Environmental factors and developmental outcomes in the lung." *Pharmacol Ther* 114(2): 129-145.

Kattan M, Gergen PJ, Eggleston P, Visness CM and Mitchell HE (2007). "Health effects of indoor nitrogen dioxide and passive smoking on urban asthmatic children." *J Allergy Clin Immunol* 120(3): 618-624.

Kelly FJ (2003). "Oxidative stress: its role in air pollution and adverse health effects." *Occupational and Environmental Medicine* 60(8): 612-616.

Kim H, Kieckhefer GM, Greek AA, Joesch JM and Baydar N (2009). "Health care utilization by children with asthma." *Prev Chronic Dis* 6(1): A12.

King ME, Mannino DM and Holguin F (2004). "Risk factors for asthma incidence. A review of recent prospective evidence." *Panminerva Med* 46(2): 97-110.

Krämer U, Sugiri D, Ranft U, et al. (2009). "Eczema, respiratory allergies, and traffic-related air pollution in birth cohorts from small-town areas." *Journal of Dermatological Science* 56(2): 99-105.

Kurukulaaratchy RJ, Matthews S, Holgate ST and Arshad SH (2003). "Predicting persistent disease among children who wheeze during early life." *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* 22(5): 767-771.

Kusel MM, de Klerk NH, Kebadze T, et al. (2007). "Early-life respiratory viral infections, atopic sensitization, and risk of subsequent development of persistent asthma." *J Allergy Clin Immunol* 119(5): 1105-1110.

Lafata JE, Xi H and Divine G (2002). "Risk factors for emergency department use among children with asthma using primary care in a managed care environment." *Ambul Pediatr* 2(4): 268-275.

Lamb AK, Ervice J, Lorenzen K, Prentice B and White S (2011). "Reducing asthma disparities by addressing environmental inequities: a case study of regional asthma management and prevention's advocacy efforts." *Fam Community Health* 34 Suppl 1: S54-62.

Laurent O, Pedrono G, Filleul L, et al. (2009). "Influence of socioeconomic deprivation on the relation between air pollution and beta-agonist sales for asthma." *Chest* 135(3): 717-723.

Lee YJ, Kim KW, Choi BS, Sohn MH and Kim KE (2013). "Clinical characteristics of eosinophilic and noneosinophilic asthma in children." *Acta Paediatr* 102(1): 53-57.

Leon Hsu HH, Mathilda Chiu YH, Coull BA, et al. (2015). "Prenatal Particulate Air Pollution and Asthma Onset in Urban Children. Identifying Sensitive Windows and Sex Differences." *Am J Respir Crit Care Med* 192(9): 1052-1059.

Li DK, Chen H and Odouli R (2011). "Maternal exposure to magnetic fields during pregnancy in relation to the risk of asthma in offspring." *Arch Pediatr Adolesc Med* 165(10): 945-950.

Lieu TA, Quesenberry CP, Sorel ME, Mendoza GR and Leong AB (1998). "Computer-based models to identify high-risk children with asthma." *Am J Respir Crit Care Med* 157(4 Pt 1): 1173-1180.

Litonjua AA, Carey VJ, Weiss ST and Gold DR (1999). "Race, socioeconomic factors, and area of residence are associated with asthma prevalence." *Pediatric Pulmonology* 28(6): 394-401.

Maher JE, Mullooly JP, Drew L and DeStefano F (2004). "Infant vaccinations and childhood asthma among full-term infants." *Pharmacoepidemiol Drug Saf* 13(1): 1-9.

McConnell R, Berhane K, Yao L, et al. (2006). "Traffic, susceptibility, and childhood asthma." *Environmental health perspectives* 114(5): 766-772.

McDaniel M, Paxson C and Waldfogel J (2006). "Racial disparities in childhood asthma in the United States: evidence from the National Health Interview Survey, 1997 to 2003." *Pediatrics* 117(5): e868-877.

McHugh ML (2012). "Interrater reliability: the kappa statistic." *Biochemia Medica: Casopis Hrvatskoga Drustva Medicinskih Biokemicara* 22(3): 276-282.

McLeish S and Turner SW (2007). "Gene-environment interactions in asthma." *Arch Dis Child* 92(11): 1032-1035.

Melen E, Nyberg F, Lindgren CM, et al. (2008). "Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease." *Environ Health Perspect* 116(8): 1077-1084.

Millard F and Zhou K. (2012). "Demographic Clusters of Georgia." Retrieved July 30, 2013, from http://oasis.state.ga.us/GADemoProfile/documents/DemoClusters2011Description.pdf.

Miller JE, Gaboda D and Davis D (2001). "Early childhood chronic illness: comparability of maternal reports and medical records." *Vital Health Stat 2*(131): 1-10.

Milton DK, Solomon GM, Rosiello RA and Herrick RF (1998). "Risk and incidence of asthma attributable to occupational exposure among HMO members." *Am J Ind Med* 33(1): 1-10.

Moore K and Persaud T (2003). The Developing Human: Clinically Oriented Embryology. Philadelphia, PA, Saunders.

Moorman JE, Akinbami LJ, Bailey CM, et al. (2012). "National Surveillance of Asthma: United States, 2001-2010. National Center for Health Statistics." *Vital Health Stat* 3(35).

National Asthma Education and Prevention Program. Expert Panel Report 3 (2007). Full Report 2007: Guidelines for the Diagnosis and Management of Asthma. NIH Publication Number 10-4051. Bethesda, MD, U.S. Department of Health and Human Services. National Institutes of Health. National Heart Lung and Blood Institute.

National Asthma Education and Prevention Program. Expert Panel Report 3 (2007). Summary Report 2007: Guidelines for the Diagnosis and Management of Asthma. NIH Publication Number 08-5846. Bethesda, MD, U.S. Department of Health and Human Services. National Institutes of Health. National Heart Lung and Blood Institute.

Neidell MJ (2004). "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma." *J Health Econ* 23(6): 1209-1236.

Nelson DA, Johnson CC, Divine GW, Strauchman C, Joseph CLM and Ownby DR (1997). "Ethnic differences in the prevalence of asthma in middle class children." *Annals of Allergy, Asthma & Immunology* 78(1): 21-26.

Nethery E, Brauer M and Janssen P (2009). "Time-activity patterns of pregnant women and changes during the course of pregnancy." *J Expo Sci Environ Epidemiol* 19(3): 317-324.

Noal RB, Menezes AM, Macedo SE and Dumith SC (2011). "Childhood body mass index and risk of asthma in adolescence: a systematic review." *Obes Rev* 12(2): 93-104.

Oftedal B, Nystad W, Brunekreef B and Nafstad P (2009). "Long-term traffic-related exposures and asthma onset in schoolchildren in Oslo, Norway." *Environmental health perspectives* 117(5): 839-844.

Osborne ML, Vollmer WM, Johnson RE and Buist AS (1995). "Use of an automated prescription database to identify individuals with asthma." *Journal of Clinical Epidemiology* 48(11): 1393-1397.

Papoutsakis C, Priftis KN, Drakouli M, et al. (2013). "Childhood overweight/obesity and asthma: is there a link? a systematic review of recent epidemiologic evidence." *Journal of the Academy of Nutrition and Dietetics* 113(1): 77-105.

PASS Statistical Software. Version 8.0.7. "Kaysville, Utah."

Patel MM, Quinn JW, Jung KH, et al. (2011). "Traffic density and stationary sources of air pollution associated with wheeze, asthma, and immunoglobulin E from birth to age 5 years among New York City children." *Environmental research* 111(8): 1222-1229.

Pearce N, Pekkanen J and Beasley R (1999). "How much asthma is really attributable to atopy?" *Thorax* 54(3): 268-272.

Pearlman DN (2012). "Considerations and challenges for eliminating asthma disparities in children." *Expert Rev Respir Med* 6(5): 509-512.

Pearlman DN, Zierler S, Meersman S, Kim HK, Viner-Brown SI and Caron C (2006). "Race disparities in childhood asthma: does where you live matter?" *J Natl Med Assoc* 98(2): 239-247.

Peden DB (2000). "Development of atopy and asthma: candidate environmental influences and important periods of exposure." *Environ Health Perspect* 108 Suppl 3: 475-482.

Peled R, Reuveni H, Pliskin JS, Benenson I, Hatna E and Tal A (2006). "Defining localities of inadequate treatment for childhood asthma: a GIS approach." *Int J Health Geogr* 5: 3.

Pinkerton KE and Joad JP (2000). "The mammalian respiratory system and critical windows of exposure for children's health." *Environmental health perspectives* 108 Suppl 3: 457-462.

Pizzichini E, Pizzichini MM, Efthimiadis A, et al. (1996). "Indices of airway inflammation in induced sputum: reproducibility and validity of cell and fluid-phase measurements." *Am J Respir Crit Care Med* 154(2 Pt 1): 308-317.

Quinto KB, Zuraw BL, Poon KY, Chen W, Schatz M and Christiansen SC (2011). "The association of obesity and asthma severity and control in children." *J Allergy Clin Immunol* 128(5): 964-969.

Ramsey CD and Celedon JC (2005). "The hygiene hypothesis and asthma." *Curr Opin Pulm Med* 11(1): 14-20.

Rauh VA, Chew GR and Garfinkel RS (2002). "Deteriorated housing contributes to high cockroach allergen levels in inner-city households." *Environmental health perspectives* 110 Suppl 2: 323-327.

Reichman NE and Nepomnyaschy L (2008). "Maternal pre-pregnancy obesity and diagnosis of asthma in offspring at age 3 years." *Matern Child Health J* 12(6): 725-733.

Richardson DB, Kinlaw AC, MacLehose RF and Cole SR (2015). "Standardized binomial models for risk or prevalence ratios and differences." *Int J Epidemiol*.

Roberts EM (2002). "Racial and ethnic disparities in childhood asthma diagnosis: the role of clinical findings." *J Natl Med Assoc* 94(4): 215-223.

Roberts G (2009). "Predicting the long-term outcome of preschool wheeze: are we there yet?" *J Allergy Clin Immunol* 124(5): 911-912.

Romagnani S (2004). "The increased prevalence of allergy and the hygiene hypothesis: missing immune deviation, reduced immune suppression, or both?" *Immunology* 112(3): 352-363.

Rosenstreich DL, Eggleston P, Kattan M, et al. (1997). "The role of cockroach allergy and exposure to cockroach allergen in causing morbidity among inner-city children with asthma." *N Engl J Med* 336(19): 1356-1363.

Saadeh FB, Clark MA, Rogers ML, et al. (2013). "Pregnant and moving: understanding residential mobility during pregnancy and in the first year of life using a prospective birth cohort." *Matern Child Health J* 17(2): 330-343.

Salam MT, Islam T and Gilliland FD (2008). "Recent evidence for adverse effects of residential proximity to traffic sources on asthma." *Curr Opin Pulm Med* 14(1): 3-8.

Schatz M, Nakahiro R, Jones CH, Roth RM, Joshua A and Petitti D (2004). "Asthma population management: development and validation of a practical 3-level risk stratification scheme." *Am J Manag Care* 10(1): 25-32.

Schatz M, Zeiger RS, Vollmer WM, et al. (2006). "Development and validation of a medication intensity scale derived from computerized pharmacy data that predicts emergency hospital utilization for persistent asthma." *Am J Manag Care* 12(8): 478-484.

Selgrade MK, Lemanske RF, Jr., Gilmour MI, et al. (2006). "Induction of asthma and the environment: what we know and need to know." *Environmental health perspectives* 114(4): 615-619.

Shankardass K, McConnell R, Jerrett M, Milam J, Richardson J and Berhane K (2009). "Parental stress increases the effect of traffic-related air pollution on childhood asthma incidence." *Proc Natl Acad Sci U S A* 106(30): 12406-12411.

Shields AE, Comstock C, Finkelstein JA and Weiss KB (2003). "Comparing asthma care provided to Medicaid-enrolled children in a Primary Care Case Manager plan and a staff model HMO." *Ambul Pediatr* 3(5): 253-262.

Shields AE, Finkelstein JA, Comstock C and Weiss KB (2002). "Process of care for Medicaid-enrolled children with asthma: served by community health centers and other providers." *Med Care* 40(4): 303-314.

Shima M, Nitta Y, Ando M and Adachi M (2002). "Effects of air pollution on the prevalence and incidence of asthma in children." *Arch Environ Health* 57(6): 529-535.

Silverman RA and Ito K (2010). "Age-related association of fine particles and ozone with severe acute asthma in New York City." *J Allergy Clin Immunol* 125(2): 367-373.e365.

Simon PA, Zeng Z, Wold CM, Haddock W and Fielding JE (2003). "Prevalence of childhood asthma and associated morbidity in Los Angeles County: impacts of race/ethnicity and income." *J Asthma* 40(5): 535-543.

Sinclair AH and Tolsma DD (2006). "Gender differences in asthma experience and disease care in a managed care organization." *Journal of Asthma* 43(5): 363-367.

Siwik JP, Johnson CC, Havstad SL, Peterson EL, Ownby DR and Zoratti EM (2005). "Airway hyperresponsiveness to methacholine at age 6 to 8 years in nonasthmatic patients is not related to increased health-care utilization for asthma in the ensuing 5 years: a longitudinal study of a birth cohort." *Chest* 128(4): 2420-2426.

Sly PD, Boner AL, Björksten B, et al. (2008). "Early identification of atopy in the prediction of persistent asthma in children." *The Lancet* 372(9643): 1100-1106.

Smith LA, Hatcher-Ross JL, Wertheimer R and Kahn RS (2005). "Rethinking race/ethnicity, income, and childhood asthma: racial/ethnic disparities concentrated among the very poor." *Public Health Rep* 120(2): 109-116.

Snyder MG, Venkatram A, Heist DK, Perry SG, Petersen WB and Isakov V (2013). "RLINE: A line source dispersion model for near-surface releases." *Atmospheric Environment* 77: 748-756.

Southeastern Center for Air Pollution and Epidemiology (SCAPE). (2010). "Website." Retrieved November 28, 2015, from http://www.scape.gatech.edu/.

Spiegelman D and Hertzmark E (2005). "Easy SAS calculations for risk or prevalence ratios and differences." *Am J Epidemiol* 162(3): 199-200.

SPSS (2001). The SPSS TwoStep Cluster Component: A scalable component enabling more efficient customer segmentation. TSCWP-0101. USA.

Stewart LJ (2008). "Pediatric asthma." *Prim Care* 35(1): 25-40, vi.

Strachan DP (1989). "Hay fever, hygiene, and household size." *BMJ* 299(6710): 1259-1260.

Strachan DP and Cook DG (1998). "Health effects of passive smoking. 6. Parental smoking and childhood asthma: longitudinal and case-control studies." *Thorax* 53(3): 204-212.

Strunk RC, Weiss ST, Yates KP, Tonascia J, Zeiger RS and Szefler SJ (2006). "Mild to moderate asthma affects lung growth in children and adolescents." *J Allergy Clin Immunol* 118(5): 1040-1047.

Szefler SJ, Wenzel S, Brown R, et al. (2012). "Asthma outcomes: biomarkers." *J Allergy Clin Immunol* 129(3 Suppl): S9-23.

Szidat S, Ruff M, Perron N, et al. (2009). "Fossil and non-fossil sources of organic carbon (OC) and elemental carbon (EC) in Göteborg, Sweden." *Atmospheric Chemistry and Physics* 9(5): 1521-1535.

Tamari M, Tanaka S and Hirota T (2013). "Genome-wide association studies of allergic diseases." *Allergol Int* 62(1): 21-28.

Taylor DR (2011). "Using biomarkers in the assessment of airways disease." *Journal of Allergy and Clinical Immunology* 128(5): 927-934.

Toren K, Brisman J and Jarvholm B (1993). "Asthma and asthma-like symptoms in adults assessed by questionnaires. A literature review." *Chest* 104(2): 600-608.

Tzivian L (2011). "Outdoor air pollution and asthma in children." *J Asthma* 48(5): 470-481.

U.S. Census Bureau (2010). Social Explorer Tables: ACS 2010 (5-year estimates). Prepared by Social Explorer. Web.

U.S. Centers for Disease Control and Prevention and National Center of Environmental Health. (2011). "Behavioral Risk Factor Surveillance System (BRFSS) Prevalence Data: 1999-2010 Tables and Graphs." from http://www.cdc.gov/asthma/brfss/archive.htm.

U.S. Environmental Protection Agency. (2013). "Sulfur Dioxide - Health." Retrieved July 9, 2013, from http://www.epa.gov/airquality/sulfurdioxide/health.html.

U.S. Environmental Protection Agency. (2015). "Meteorological Processors and Accessory Programs." Technology Transfer Network Support Center for Regulatory Atmospheric Modeling Retrieved September 29, 2015, from http://www3.epa.gov/scram001/metobsdata_procaccprogs.htm.

UCLA Institute for Digital Research and Education (2013). "SPSS Data Analysis Examples: Discriminant Function Analysis. Available from: http://www.ats.ucla.edu/stat/spss/dae/discrim.htm."

Vanderweele TJ and Knol MJ (2014). "A tutorial on interaction." *Epidemiologic Methods* 3: 33-72.

Verstraeten T, Jumaan AO, Mullooly JP, et al. (2003). "A retrospective cohort study of the association of varicella vaccine failure with asthma, steroid use, age at vaccination, and measles-mumps-rubella vaccination." *Pediatrics* 112(2): e98-103.

Villeneuve PJ, Chen L, Rowe BH and Coates F (2007). "Outdoor air pollution and emergency department visits for asthma among children and adults: a case-crossover study in northern Alberta, Canada." *Environ Health* 6: 40.

Vollmer WM, Heumann MA, Breen VR, et al. (2005). "Incidence of work-related asthma in members of a health maintenance organization." *J Occup Environ Med* 47(12): 1292-1297.

Wang J, Visness CM, Calatroni A, Gergen PJ, Mitchell HE and Sampson HA (2009). "Effect of environmental allergen sensitization on asthma morbidity in inner-city asthmatic children." *Clinical & Experimental Allergy* 39(9): 1381-1389.

Wang L and Pinkerton KE (2008). "Detrimental effects of tobacco smoke exposure during development on postnatal lung function and asthma." *Birth Defects Research Part C* 84(1): 54-60.

Wells EM, Bonfield TL, Dearborn DG and Jackson LW (2013). "The relationship of blood lead with immunoglobulin E, eosinophils, and asthma among children: NHANES 2005-2006." *Int J Hyg Environ Health*.

Williams DR, Sternthal M and Wright RJ (2009). "Social determinants: taking the social context of asthma seriously." *Pediatrics* 123(Supplement 3): S174-S184.

Williams TJ (2004). "The eosinophil enigma." *J Clin Invest* 113(4): 507-509.

Wright AL (2002). "Epidemiology of asthma and recurrent wheeze in childhood." *Clin Rev Allergy Immunol* 22(1): 33-44.

Yang IA, Fong KM, Zimmerman PV, Holgate ST and Holloway JW (2009). "Genetic susceptibility to the respiratory effects of air pollution." *Postgrad Med J* 85(1006): 428-436.

Yoo KH, Johnson SK, Voigt RG, Campeau LJ, Yawn BP and Juhn YJ (2007). "Characterization of asthma status by parent report and medical record review." *J Allergy Clin Immunol* 120(6): 1468-1469.

Yoshikawa T and Kanazawa H (2011). "Phenotypic differences between asymptomatic airway hyperresponsiveness and remission of asthma." *Respir Med* 105(1): 24-30.

Yost GS, Buckpitt AR, Roth RA and McLemore TL (1989). "Mechanisms of lung injury by systemically administered chemicals." *Toxicol Appl Pharmacol* 101(2): 179-195.

Young S, Arnott J, Le Souef PN and Landau LI (1994). "Flow limitation during tidal expiration in symptom-free infants and the subsequent development of asthma." *J Pediatr* 124(5, Part 1): 681-688.

Yunginger JW, Reed CE, O'Connell EJ, Melton LJ, 3rd, O'Fallon WM and Silverstein MD (1992). "A community-based study of the epidemiology of asthma. Incidence rates, 1964-1983." *Am Rev Respir Dis* 146(4): 888-894.

Zeiger RS, Hay JW, Contreras R, et al. (2008). "Asthma costs and utilization in a managed care organization." *J Allergy Clin Immunol* 121(4): 885-892 e885.

Zeiger RS, Schatz M, Li Q, Zhang F, Purdum AS and Chen W (2010). "Step-up care improves impairment in uncontrolled asthma: an administrative data study." *Am J Manag Care* 16(12): 897-906.

Zhai X, Sampath P, Mulholland JA, et al. (2015). "Comparison and calibration of Research-Line (RLINE) model results with measurement-based an CMAQ-based source impacts." *Poster presented at Community Modeling and Analysis System (CMAS) Conference*.

Zhou C, Baiz N, Zhang T, Banerjee S and Annesi-Maesano I (2013). "Modifiable exposures to air pollutants related to asthma phenotypes in the first year of life in children of the EDEN mother-child cohort study." *BMC Public Health* 13(1): 506.

Zhou Y (2012). Creating Demographic Clusters of Georgia, 2011, Georgia Department of Public Health, Office of Health Indicators for Planning.

Zmirou D, Gauvin S, Pin I, et al. (2004). "Traffic related air pollution and incidence of childhood asthma: results of the Vesta case-control study." *J Epidemiol Community Health* 58(1): 18-23.