Association between Air Pollution Exposure and Self-Report of Recent Respiratory Infection
A MESA Air study

Carly Marie Miller

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Public Health

University of Washington
2014

Committee:
Joel D Kaufman
Sverre Vedal
Anjum Hajat

Program Authorized to Offer Degree:
Public Health - Epidemiology

University of Washington

Abstract

Association between Air Pollution Exposure and Self-Report of Recent Respiratory Infection
A MESA Air study

Carly Marie Miller

Chair of the Supervisory Committee:
Joel Kaufman
Department of Epidemiology

Background: Ambient air pollution has been linked to multiple adverse health outcomes. Most of these studies have use large-scale models to estimate air pollution exposures and have identified health outcomes based on hospital diagnosis codes. The Multi-Ethnic Study of Atherosclerosis (MESA) and MESA Air provide residence-level air pollution data. We investigated the association between these data and information collected prospectively on recent MESA participant infection.

Methods & Results: MESA, a prospective cohort study, followed over 6800 participants with 5 study visits over 12 years. Data collected included participant self-report of recent respiratory infection. Individual ambient air pollution exposures to fine particulate matter ($PM_{2.5}$), oxides of nitrogen (NO_x and NO_2), and black carbon (BC) were estimated. Cross-sectional and longitudinal associations were examined with generalized linear models for each pollutant adjusted for confounders: study site, season of exam, age, sex, race/ethnicity, smoking, and socio-economic status. Across all study sites, air pollution exposure estimates decreased. Report of recent respiratory infection was associated with season, the highest prevalence in winter. Nearly 18% of participants reported infection during the 2 weeks prior to their visit. Significant associations were seen between increased prevalence of infection and elevated exposures to $PM_{2.5}$, NO_x, and NO_2.

Discussion: In a large population-based cohort study, we found robust evidence that the prevalence of respiratory infection is increased in participants exposed to higher levels of ambient air pollution.

Table of Contents

List of Figures……...v

List of Tables……vi

Acknowledgments……...……vii

Introduction & Literature Review…….……1

Ambient air pollution & historical review of pollution research……..…….....1

Air pollution and cardiovascular disease……..……...4

Air pollution and respiratory disease……..……..…5

Covariates to consider with air pollution exposure…….9

MESA and MESA Air……...…….....10

Methods……...11

Study subjects…….…….……..……11

Outcome and risk factor data…….…….……...12

Other variables and covariates…….…….……13

Air pollution exposure estimation…….…….……....15

Exclusion criteria…….……..…….....15

Statistical analysis…….……..……...16

Directed Acyclic Graph…….……18

Results……...…….....21

Predicted pollutant exposures…….…….…...21

Subject characteristics…….…….….25

Prevalence of health outcomes…….…….…26

Association between long-term air pollution exposure and self-reported infection……..….30

Discussion…….……..…….35

Conclusion……...……37

References……..…..……38

List of Figures

Figure 1: US map of counties by NAAQS compliance……3

Figure 2: Directed Acyclic Graph (DAG) detailing model with main effects and covariates……..……20

Figure 3: Annual pollutant levels by study visit…….……23

Figure 4: Annual pollutant levels by study site……...……24

Figure 5: Prevalence of infection by month (season) of exam……27

List of Tables

Table 1: Mean annual pollutant predictions, by study site……22

Table 2: Mean two-week levels of pollutant predictions, by study site……22

Table 3: Correlation between paired pollutants, using annual predictions……24

Table 4: Correlation between pollutant annual and two-week predictions……24

Table 5: Participant characteristics at baseline examination……25

Table 6: Percent and counts of health outcomes by visit……27

Table 7: Prevalence ratios and 95% confidence intervals of association between covariates and self-reported infection……29

Table 8: Differences in inflammatory markers by report of infection……30

Table 9: Prevalence ratios of association between annual pollutant predictions and self-reported infection, by study visit……31

Table 10. Prevalence ratios for the association between annual pollutant predictions and prevalence of any infection……32

Table 11: Prevalence ratios for the association between two-week pollutant predictions and prevalence of any infection……33

Table 12: Prevalence ratios for the association between annual pollutant predictions and prevalence of fever, cold/flu, sinus infection, and bronchitis/pneumonia……34

Table 13: prevalence ratios for the association between annual pollutant prediction and prevalence of any infection with other variables for individual SES: education level, employment status, home type……34

Acknowledgements

I would like to thank my thesis advisors, Dr. Joel Kaufman, Dr. Sverre Vedal, and Dr. Anjum Hajat, as well as the MESA and MESA Air staff for their guidance and patience throughout this process. This work would not have been possible without the many people who contributed to the design and development of the Multi-Ethnic Study of Atherosclerosis (MESA) and the ancillary study of air pollution, MESA Air. MESA is registered in the ClinicalTrials.gov database under unique registry identifier NCT00005487, and a full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org. This research was supported by awards RD831697 and 83386401 from the EPA; N01-HC-95159 through N01-HC-95165, N01-HC-95169, R01-HL-069003, and R01-HL-081352 from the NHLBI; and K2ES013195, P50ES015915, and P30ES07033 from the NIEHS. The funding agencies were not involved in the design of this study and this thesis does not reflect their vies. I do not have any relevant financial disclosure related to this topic.

Introduction & Literature Review

I. Ambient air pollution and historical review of air pollution research

The Clean Air Act of 1963, signed into law by President Lyndon B Johnson, brought concerns of exposure to ambient air pollution to the nation's collective consciousness. This legislation authorized federal and state authorities to enact air pollution standards and required the Environmental Protection Agency (EPA) to develop and enforce air pollution directives. This original legislation established a research program that has since been extensively amended and modified and now includes comprehensive federal and state regulations.[1] The National Ambient Air Quality Standards (NAAQS) was established with the Clean Air Act Extension of 1970 and requires the EPA to maintain national standards for six criteria pollutants which currently include sulfur dioxide (SO_2), carbon monoxide (CO), particulate matter, nitrogen dioxide (NO_2), ozone (O_3), and lead. The EPA continues to monitor air pollution and air quality standards and mandates metropolitan areas to release daily information regarding their adherence to these standards.[1]

Even before the Clean Air Act air pollution and its health effects had been a subject of much scientific research and discussion. Bascom noted the multitude of adverse health effects associated with ambient air pollution in her 1996 summary paper "Health Effects of Outdoor Air Pollution." This paper discusses the varied composition of ambient air pollution and acute health effects associated with exposure including: excess cardiorespiratory mortality, increased health care utilization, asthma exacerbations, increased respiratory illness, increased respiratory symptoms, decreased lung function, increased airways reactivity, lung inflammation, and altered host defense.[2] This article provides a helpful summary of the research findings up to this point and discusses research methods being used to investigate associations. Reviewing papers from the early years of air pollution research, it is evident how far we have come. Early reviews focused on evidence of detrimental health effects based on events such as the London Fog in 1952. These, and other events,

led to the development of air quality standards as well as the legal action to enforce them. Research methods have refined over time and more studies now consider chronic air pollution exposure and the association with long-term effects on human health.

Ambient air pollution is a diverse mixture that includes byproducts of anthropogenic processes, namely combustion from industry (especially power generation) and motor vehicle use. The resulting pollutants that can be readily measured include particulate matter, oxides of nitrogen (including nitrogen dioxide), and black carbon. As described above, the EPA frequently evaluates ambient levels of such pollutants and regulates their environmental concentrations. They propose regulations of exposure levels for the general population as well as exposure levels for "sensitive" populations, including the elderly, children, asthmatics, and individuals with cardiorespiratory disease. Current EPA regulations of the annual average concentrations from the National Ambient Air Quality Standards (NAAQS) are:

NO_2 = 53 parts per billion (ppb)

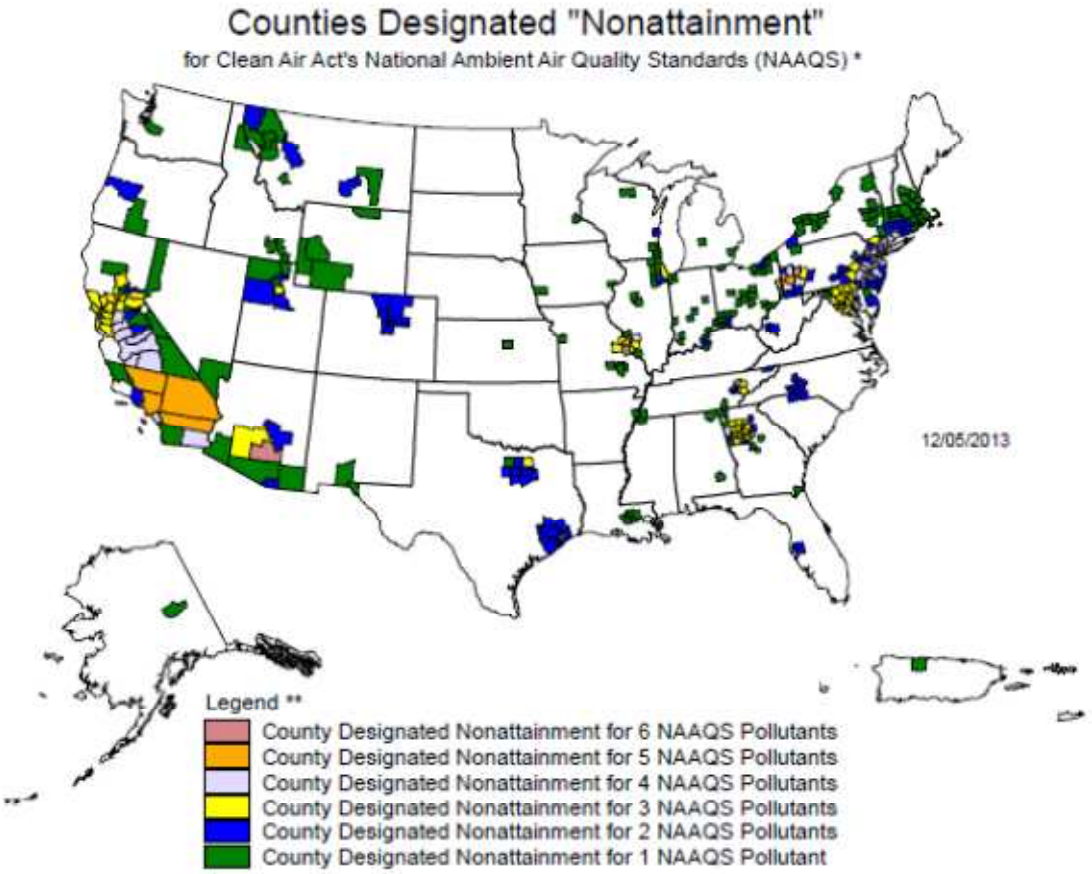
$PM_{2.5}$ = 12 micrograms per meter cubed ($\mu g/m^3$).

$PM_{2.5}$ refers to particulate matter with an aerodynamic diameter < 2.5 $\mu g/m^3$.

The EPA also regulates lead, SO_2, O_3, and CO levels as criteria air pollutants.[3]

A scientific statement issued by the American Heart Association in 2010 lists these NAAQS values alongside current US averages.[4] According to their research, the national average range of $PM_{2.5}$ is 5-50 $\mu g/m^3$ with a mean of 13.4 $\mu g/m^3$. The national range of NO_2 as .5-50 ppb and black carbon as 0.1-3.0 $\mu g/m^3$. Figure 1 provides a map of all US counties that do not meet the NAAQS for at least one pollutant.

Figure 1. US Map of counties by NAAQS compliance.[5]



Guam - Piti and Tanguisson Counties are designated nonattainment for the SO2 NAAQS

* The National Ambient Air Quality Standards (NAAQS) are health standards for Carbon Monoxide, Lead (1978 and 2008), Nitrogen Dioxide, 8-hour Ozone (1997 and 2008), Particulate Matter (PM-10 and PM-2.5 (1997 and 2006)), and Sulfur Dioxide.(1971 and 2010)

** Included in the counts are counties designated for NAAQS and revised NAAQS pollutants. 1-hour Ozone is excluded. Partial counties, those with part of the county designated nonattainment and part attainment, are shown as full counties on the map.

Particulate matter is arguably the most well studied component of air pollution with regard to health outcomes, with $PM_{2.5}$ often considered the most critical component. This smaller size of particulate matter is theorized to be more important from a health perspective, and the most potentially damaging, as the particle size is within the respirable range, that is, particles able to reach the air-exchange areas of human lungs.[6] Such respirable particles may elicit an inflammatory

reaction within an individual's airways, and perhaps a systemic reaction, and are believed to be responsible for many of the pathologic changes associated with increased pollutant exposure.[7]

The scientific statement by the American Heart Association explains evidence for the association between these respirable particles and the systemic inflammatory reaction they elicit.[4] Research data have led scientists to believe that an inflammatory reaction caused by exposure to pollutants is responsible for the pathologic effects. The authors write that exposure to ambient particulate matter can be associated with increased levels of inflammatory biomarkers, such as C reactive protein (CRP), fibrinogen and interleukin-6 (IL-6). Elevated biomarkers may point to a systemic response to ambient air pollution, a response that extends outside of the air-exchange interface.[4]

II. Air Pollution and Cardiovascular Disease

Much research has been conducted evaluating the effects of ambient air pollution on cardiovascular disease. Many of the groundbreaking studies in this field, including the Harvard Six Cities study, evaluated the association between metropolitan ambient air pollution levels and mortality rates from cardiovascular causes.[8] These studies found an association between elevated ambient air pollution levels and cardiovascular mortality and spurred further research in the field. The American Heart Association, in their 2010 scientific statement, neatly summarized much of the relevant research evaluating the association between heart disease and exposure to particulate matter. They conclude:

> exposure to PM <2.5μg in diameter ($PM_{2.5}$) over a few hours to weeks can trigger cardiovascular disease-related mortality and nonfatal events; longer-term exposure (eg, a few years) increases the risk for cardiovascular mortality to an even greater extent than exposures over a few days and reduces life expectancy within more highly exposed segments of the population by several months to a few years….overall evidence is consistent with a casual relationship between $PM_{2.5}$ exposure and cardiovascular morbidity and mortality [4].

This scientific statement updated the public and scientific community's awareness concerning the relationship between short-term and long-term risks associated with ambient air pollution and cardiovascular mortality. The statement reviews the results of much research and concludes that most studies show an elevated risk of cardiovascular morbidity and mortality associated with increasing exposure to fine particulate matter and nitrogen dioxide.[4] Some studies evaluate these events and retrospectively associate them with levels of ambient air pollution components. For example, the Harvard Six Cities Study evaluated health outcomes in relation to air pollution levels and concluded that cities with greater pollution levels had increased adjusted daily mortality rates compared to cities with lower pollution levels.[4,8] These studies can provide ecological epidemiologic evidence, but do not provide much more due to their crude pollution measures, causality cannot be assessed. The pollution measures used in these studies generally associated one level of exposure to all participants within a geographic area using fixed site monitors. These studies were integral in determining the acute morbidity and mortality associated with high levels of air pollution; they also serve as important examples of the issues associated with assessing pollution exposure at an individual level. Other studies involve more precise measurements of air pollution concentrations by using multiple monitors throughout the area to assess the temporal-spatial distribution of air pollutants.

III. Air Pollution and Pulmonary Disease

Studies evaluating the effects of air pollution exposure on pulmonary disease have typically focused on populations with preexisting pulmonary disease: those with asthma, chronic obstructive pulmonary disease (COPD), and chronic bronchitis. Many studies have evaluated changes in health care utilization by these populations during times of short-term fluctuations of air pollution levels in metropolitan areas. In 1993 in Seattle, Schwartz and colleagues evaluated the association between

daily levels of PM_{10} and emergency room visits for asthma in adults. (Like $PM_{2.5}$, PM_{10} refers to particulate matter with an aerodynamic diameter < 10 $\mu g/m^3$). They determined that the mean PM_{10} levels four days prior were good predictors for ER visits in asthmatics and that increased levels of PM_{10} were associated with an increase in ER visits.[9] Additionally, Peel and colleagues assessed ambient air pollution levels and acute upper respiratory infections in the Atlanta metropolitan area using ICD-9 diagnoses of asthma, COPD, upper respiratory tract infections (URI), and pneumonia.[10] They retrospectively measured air pollution levels and determined O_3, NO_2, PM_{10} were associated with an increased number of emergency department visits for URI. Further, they determined that a $2\mu g/m^3$ increase in $PM_{2.5}$ was associated with a 3% increase in pneumonia cases for all ages.[10]

The authors of the Air Pollution and Health: A European Approach (APHEA) studies also evaluated the short-term effects of air pollution on hospital admission and mortality related to respiratory disease. Using fixed site monitors in five European cities, they determined that a significant increase in daily hospital admissions for respiratory disease was seen with elevated levels of ozone (O_3).[11] Chauhan and colleagues conducted a review of ambient air pollution and respiratory effects and commented specifically on NO_2 exposure in the APHEA study.[12] They determined that acute increases in NO_2 exposures were associated with a 2.6% increase in asthma admissions as well as a 1.3% increase in daily all-cause mortality.[12] Neupane and colleagues conducted a case-control study to determine the association between long term exposure to ambient air pollution and the risk of community acquired pneumonia among adults >65 years of age.[13] They determined that increases in particulate matter exposure were associated with an increased risk of community-acquired pneumonia.[13]

In contrast to the prior studies discussed here, Brauer, Hoek, and colleagues studied a longitudinal cohort study in the Netherlands to determine the association between ambient air pollution exposure and acute infections and asthma. This study used a birth cohort of approximately

4000 infants and followed them, initially for two years.[14] The first iteration of this study allowed parents to complete a questionnaire asking them to report symptoms of asthma, allergy, and acute respiratory infection in newborns. The study team used air pollution models and land use regression to generate an exposure for each child's residence, which allowed them to capture variability in pollutant exposure. They determined a general positive association between traffic related air pollutants and wheezing, asthma, and respiratory infection in the cohort.[14] The follow-up study included data up to age 4 on the birth cohort and had similar findings. Specifically, they determined that an increase in elemental carbon exposure was associated with significant increases in ear, nose, and throat infections and flu/serious colds.[15]

Other studies have examined the relationship between ambient air pollution exposure and acute respiratory infection, with children as participants. MacIntyre and colleagues performed a meta-analysis using data from ten European birth cohorts to examine this association. They studies all examined the association between traffic related air pollutants and respiratory infections (pneumonia, croup, and otitis media) in childhood using parent report of physician diagnosis as the outcome measure. The authors concluded that a positive association was evident, with the strongest evidence associated with pneumonia diagnoses and higher levels of ambient pollutants.[16] Similarly, Mehta et al. conducted a similar study to elucidate the global attributable burden of mortality from acute lower respiratory infections (ALRI) in children. The study group calculated a summary estimate of a 1.12 increased risk of ALRI per 10 $\mu g/m^3$ increase in annual $PM_{2.5}$.[17] Respiratory infection is the leading cause of morbidity and mortality worldwide, perhaps explaining the abundance of research in this field focusing on children.[17] Globally, indoor air pollution resulting from indoor cook stoves is thought to cause much of the illnesses, particularly in lower income settings.

Despite the robust research conducted with children, few studies have examined the relationship between adult lung function, acute infection, and ambient air pollution. The Swiss Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) was conducted as a cross-sectional study evaluating the long-term relationship between lung function measurements and air pollution levels in 8 Swiss communities.[18] The researchers determined that NO_2, SO_2, and PM_{10} were associated with a lower FEV_1 and FVC in all study groups. However, this study did not consider acute infections associated with these pollutant exposures. Additionally, this study was published in 1997 and did not employ the sophisticated pollutant measurement methods that are now available to improve the existing variability among municipal areas.

Ciencewicki and Jaspers summarized the current evidence suggesting the association between infections and specific ambient air pollutants.[19] The authors describe the epidemiological and experimental effects of ambient pollutants on respiratory infections, particularly on the over 200 viruses known to cause respiratory infections. They describe a study by Jaakkola et al comparing the incidence of URIs in Polish children. The authors compare the prevalence of URIs over the course of a year in children living in polluted cities to those living in reference cities.[20] They concluded that the children living in more polluted cities had a 2.0 increased prevalence of respiratory infections, including the common cold and otitis media. This study, published in 1991, used rudimentary methods to assess air pollutant exposure yet provides a strong association in an ecological framework. Another study, by Wong et al, describes an ecological study retrospectively examining air pollution levels and hospital admissions for respiratory disease in Hong Kong. This study specifically examined short-term NO_2, O_3, and PM_{10} exposures and found a significant association between higher exposures of these pollutants and hospital visits for pneumonia and influenza.[21]

Although these many different studies take varied approaches and focus on different aspects of ambient air pollution, they generally all conclude that increases in ambient air pollution exposure are associated with poorer health outcomes. However, none of the studies reviewed describe a prospective approach to gathering data on respiratory infections in adults coupled with individual-level air pollutant exposures.

IV. Covariates to Consider with Air Pollution Exposure

Studies investigating the relationship between adverse events and air pollution levels must consider many variables when evaluating their findings.

Data indicate that air pollution components vary within cities and geographic areas. These pollutants also vary with time and have different peak levels throughout the day. For example, NO_2 and black carbon tend to peak during rush hours as their main source is motor vehicle exhaust.[4] Particulate matter levels tend to have a longer atmospheric lifetime than the other pollutants and is affected more by meteorological conditions such as solar radiation. Particulate matter tends to be highest in the mid-afternoon hours. Weather patterns as well as seasonal changes can also impact the concentrations of air pollutants. Because pollutant levels can vary by geographic location within metropolitan areas, neighborhood variables are important to consider. Neighborhoods can be an effective surrogate measure for contextual socioeconomic status and has been shown to influence an individual's health outcomes [22]. Commonly, lower income housing may be closer to roadways, industrial areas, or other exposures that may increase the occupants' exposure to these pollutants.[23] Neighborhood socioeconomic status is also being evaluated as a potential covariate for the relationship between air pollution exposure and adverse health outcomes. A study by Clark and colleagues used a land use regression model to evaluate nitrogen dioxide concentrations and socioeconomic status throughout the United States.[24] Using Census 2000 data and air pollution data

from 2006, they determined inequalities with nitrogen dioxide dispersal. For example, average NO_2 concentrations were 4.6 ppb higher for nonwhites than whites, 1.2 ppb higher for those below the poverty line versus those above, and 3.4 ppb higher for low income whites compared to higher-income whites.[24]

V. MESA and MESA Air

The Multi-Ethnic Study of Atherosclerosis (MESA) is a longitudinal epidemiologic study of subclinical cardiovascular disease and the development of clinical cardiovascular disease in a diverse population-based sample. MESA Air, an ancillary study, uses data from the MESA sample designed to investigate effects of individual-level exposures to ambient $PM_{2.5}$ and traffic related pollutants on subclinical and clinical cardiovascular disease in the MESA cohort. The MESA Air study augments the MESA data with novel methods for ambient air pollution concentration measurements and predictions.[25] Such data enables fluctuations of air pollutant concentrations and within-city gradients to be more specifically studied.

The MESA sample has served as the substrate for more than 700 peer-reviewed publications. The main study and its many ancillary studies have served to elucidate health effects and the course of cardiovascular disease in unprecedented ways. Research involving MESA Air has focused on the relationship between air pollution predictions and cardiovascular disease outcomes such as aortic atherosclerosis[26], left ventricular mass[27], and heart rate variability[26, 28]. Other studies have utilized MESA Air and MESA Neighborhood data to examine the relationship between neighborhood socioeconomic status, residential segregation, and air pollution exposure.[29] The MESA Air studies have also been invaluable data for advancing air pollution exposure assessment[25, 30].

The purpose of this study is to investigate the association between air pollution levels and recent respiratory infection using data from MESA and MESA Air. We hypothesize that there will

be positive associations between these outcomes and higher concentrations of ambient air pollution. Ambient air pollution levels will be evaluated on a long-term and short-term basis. We will consider effect modification by age, gender, obesity, diabetes, race/ethnicity, and socioeconomic status. The potential for confounding by socioeconomic factors and the role of social and physical neighborhood environmental factors to serve as either confounders or effect modifiers will also be considered.

We will determine:

1) Whether estimated annual outdoor residential concentrations of fine particulate matter ($PM_{2.5}$), oxides of nitrogen, (NO_x, NO_2) and black carbon (BC) are associated with self-reported recent infection,
2) Whether estimated concentrations of $PM_{2.5}$, NO_x, NO_2 and BC in the 2-week period preceding participant interview, are associated with report of recent infection; and
3) Whether reported recent infection is associated with markers of inflammation including C reactive protein, fibrinogen, and IL-6.

Methods:

MESA Air, MESA, and air pollution data were made available by the MESA data coordinating center at the University of Washington.

Study subjects

The Multi-Ethnic Study of Atherosclerosis (MESA) is a cohort study in six U.S. metropolitan areas involving four racial-ethnic groups (non-Hispanic white, African American, Hispanic, and Asian) designed to assess the prevalence, correlates, and progression of subclinical cardiovascular disease.

These participants were free of clinical cardiovascular disease at their baseline examination conducted between 2000-2002. Detailed MESA eligibility criteria are available online (http://www.mesa-nhlbi.org). Subject recruitment occurred in six U.S. metropolitan areas: Los Angeles County, CA; Chicago, IL; Forsyth County, NC; Northern Manhattan and the Bronx, NY; Baltimore City and Baltimore County, MD; and St. Paul, MN. The original MESA cohort consisted of 6,814 US adults.

MESA participants were followed for twelve years with 5 evaluations during that time. Evaluations were conducted during the following time frames: Visit 1/baseline: July 2000 – August 2002, Visit 2: September 2002 – February 2004, Visit 3: March 2004-September 2005, Visit 4: September 2005 – May 2007, and Visit 5: April 2010- February 2012.

Demographic, medical history, physical examination, and laboratory data were recorded at each participant visit. Their residential addresses were collected and assigned geographic coordinates using mapping software and geocoding.[13] Of the 6814 MESA cohort members at baseline, 6,256 (91.8%) consented to geocoding. This process enabled predictions to be made regarding ambient air pollution exposure at each individual's location. The protocol was approved by institutional review boards at all participating centers, and all participants gave informed consent.

Outcome and risk factor data

The participants were assessed with physical examination, non-invasive testing, and an extensive questionnaire. Questionnaire data included demographic information, interim medical history and medication use, and behavioral data.

Participants were asked to complete questionnaires at each of the MESA study visits. These questionnaires all included questions based on recent infections, i.e. infections within the past two weeks. The specific questions of interest include: pneumonia, bronchitis, sinus infection, cold/flu,

and fever. The answers were not verified with medical records and the severity of the infection is unknown.

Outcomes we will consider include report of recent infection as listed on the MESA questionnaires. At each visit, prompts on the questionnaire were as follows:

- Fever in the past 2 weeks?
- Sinus infection in the past 2 weeks?
- Cold/flu in the past 2 weeks?
- Bronchitis in the past 2 weeks?
- Pneumonia in the past 2 weeks?

Dental infection and urinary tract infection was also assessed; however, we did not analyze these responses. The included questions were used because the focus of our study is to examine the association between ambient air pollution exposure and respiratory infections. In addition to the five individual infections listed we also developed a dichotomous variable "any infection" that referred to the respondents' answers to the above questions. If any of the above questions were positive, the infection variable was coded as "1". Otherwise, it was "0". There were instances in which responses were irregularly omitted by the participants. We included these respondents if they answered at least 3 of the above questions. Otherwise, respondents were not included in the final analyses, which included 35 participants (0.13%).

Other Variables / Covariates

Participant age, race/ethnicity, sex, and metropolitan area were included as covariates. Occupational status, education level attained, country of birth, household income, socioeconomic status of participant's neighborhood, health history data including body mass index (BMI), diabetes status, cigarette smoking history, and lung disease including emphysema and asthma were also included. Some of the data were collected at each exam, therefore variables are time varying. These

variables include: age, BMI, smoking status, diabetes status, household income, number of people living in household, current job status, marital status, and medication use. Education level was classified as: less than high school, high school, some college, Bachelor's degree, and Graduate or Professional degree. BMI was classified as normal (<25 kg/m^2), overweight (25-<30 kg/m^2), and obese (30+ kg/m^2). Household income was divided into quartiles: <$20,000, $20,000 - <40,000, $40,000 - <75,000, $75,000+ and treated as a categorical variable with these values. Current job status is categorized as: homemaker, employed, unemployed, and retired. Marital status is classified as married or unmarried. Cigarette status is evaluated as never smoked, former smoker, or current smoker. Diabetes is classified as treated diabetic, untreated diabetic, impaired fasting glucose, and not diabetic, this was further collapsed into diabetic (which included treated, untreated and impaired) and not diabetic. Emphysema and asthma were evaluated as binary variables defined as presence or absence of disease. The season of exam is also included as a covariate and is defined as Winter (months 1-3), Spring (4-6), Summer (7-9), and Autumn (10-12).

Individual socioeconomic status was assessed with several different variables. Data were available on participant annual income, current job, type of household (own, mortgage, rent), and number of people in the household. Neighborhood socioeconomic status was also used. This variable is derived from census tract level socioeconomic data from the Census 2000, American Community Survey 2005-2009, and American Community Survey 2007-2011. The resulting variable is a weighted variable that takes into account census measures regarding the racial and ethnic compositions, education and employment levels as well as housing and crowing variables of the neighborhood. They values were then standardized with a z-score to make them more interpretable in a larger data analysis.

Air pollution exposure estimation:

Previous air pollution research evaluated various clinical outcomes with pollution levels using group-level data for pollution exposure.[3,4,11] The model used by MESA researchers included deployment of cohort-oriented air pollution monitors and advanced geostatistical approaches to derive fine-scale participant level exposure levels.[25] Each field center employed multiple monitors throughout the area to collect concentrations of air pollutants including $PM_{2.5}$ ($\mu g/m^3$), NO_x (ppb), NO_2 (in ppb), and BC ($\mu g/m^3$). These monitors included fixed sites such as libraries, schools, or other buildings that were in areas underrepresented by existing EPA monitors. The locations of these monitors were based on distances from roadways, major traffic gradients, and individual monitoring with the goal of capturing within-city variations in pollution.[25] The data from preexisting EPA monitors added to the precision of MESA predictions. This cumulative data enabled MESA researchers to develop a hierarchical spatiotemporal model to predict long-term average air pollution concentrations for each participant. Additionally, 100 two-week samples of the pollutants were collected in each study area to assist the researchers in their attempt to model within-city concentration variability.[25] The exception to this method is the predicted concentrations of black carbon. Temporal-spatial predictions were not available for black carbon over the time period used in this study. Instead, BC data are spatially predicted. Therefore, the only variability in BC is among locations (i.e. they are constant if a participant did not move over the study period). This is a standard practice for exposures where temporal resolution cannot be further defined. More details are available in articles by Cohen et al and Szpiro et al.[25,31]

Exclusion Criteria

MESA participants who did not consent or who reported baseline addresses outside of one of the six MESA metropolitan areas were not included in the present analysis. Sixty-one participants

did not have available pollution data for any of the visits and these were excluded from further analysis. Patients were invariably lost to follow-up, including death. The pollution data specific to their address was only included if the patient participated in follow-up visits. Additionally, participants were excluded from the analysis if they did not answer the questions on the questionnaire related to our outcome of interest: report of recent infection. Thirty-four participants did not answer any questions related to self-report of infection, and these observations were dropped from the data set. (i.e. the participant participated in the exam and questionnaire based on other available data, but did not respond to all 5 questions related to infection). No participants were missing data on 4 of the 5 questions. Only 1 participant left three of the options blank, and was excluded from the study.

Finally, participants who reported a diagnosis of cancer were excluded from the study (n = 532). Visits were also excluded if the participant noted taking oral steroids at the time of the visit (n = 366). This resulted in a total of 6125 participants to evaluate, and a total number of visits at 26663. Air pollution data was available for all included participants.

Statistical Analysis:

Analyses were performed using STATA version 12 (Stata Corp, College Station, TX). Descriptive variables were evaluated based on baseline data per study site. Air pollution levels were compared between study sites and within study sites as visits progressed. This was an important consideration due to the changing air pollution standards over the time period of data collection. We expected air pollution levels to decrease across all sites and across all pollutants as the study progressed. Prevalence of reported infection was evaluated by season of visit. Differences between covariates were assessed using chi-square tests for categorical variables and one-way analysis of variance (ANOVA) for continuous variables. Univariate regression models using a general linear method

examined associations between reported infection and known risk factors as well as selected covariates. Separate models were developed for each pollutant: $PM_{2.5}$, NO_x, NO_2, and BC. We used a generalized linear model with a logistic binomial regression in all multivariate analyses and statistical significance was defined by p-values <.05. Air pollutant predictions were entered as continuous variables and were not transformed. Based on the distribution of our pollution data (shown below), we calculated estimates for the following increments in exposure: 4 $\mu g/m^3$ for $PM_{2.5}$, 25 ppb for NO_x, 9 ppb for NO_2, and 0.3 $\mu g/m^3$ for BC. Since there were several observations per participants, we accounted for within-participant clustering when calculating standard errors. In multivariate models for reported infection with air pollutant exposure as the exposure variable, we added covariates based on the DAG that were selected *a priori*. This was done in an effort to maintain consistency between models. These analyses were conducted for the four pollutants considered here: $PM_{2.5}$, NO_x, NO_2, BC. That is, the analyses were conducted separately for each pollutant. Covariates were added in a step-wise fashion as shown:

Model 1 includes age, race/ethnicity, gender, season, site, smoking status

Model 2…controlling for model 1 + income

Model 3…controlling for model 1 + neighborhood SES

Model 4….controlling for model 1 + income + neighborhood SES

A restricted variable was generated for both NO_x and NO_2 and limited to the 99% of their range in order for multivariate models to converge.

Other modeling approaches were undertaken in order to compare our results to past research. We conducted city-specific analyses to explore heterogeneity in associations across MESA sites. For baseline data, lab values including IL-6, fibrinogen, and CRP were available for most of the participants. The mean values of these markers were compared between those reporting recent infection at the time of the first study visit and those not. CRP was logistically transformed for this

analysis. Analyses of variance were used to detect the significance of differences between these two groups.

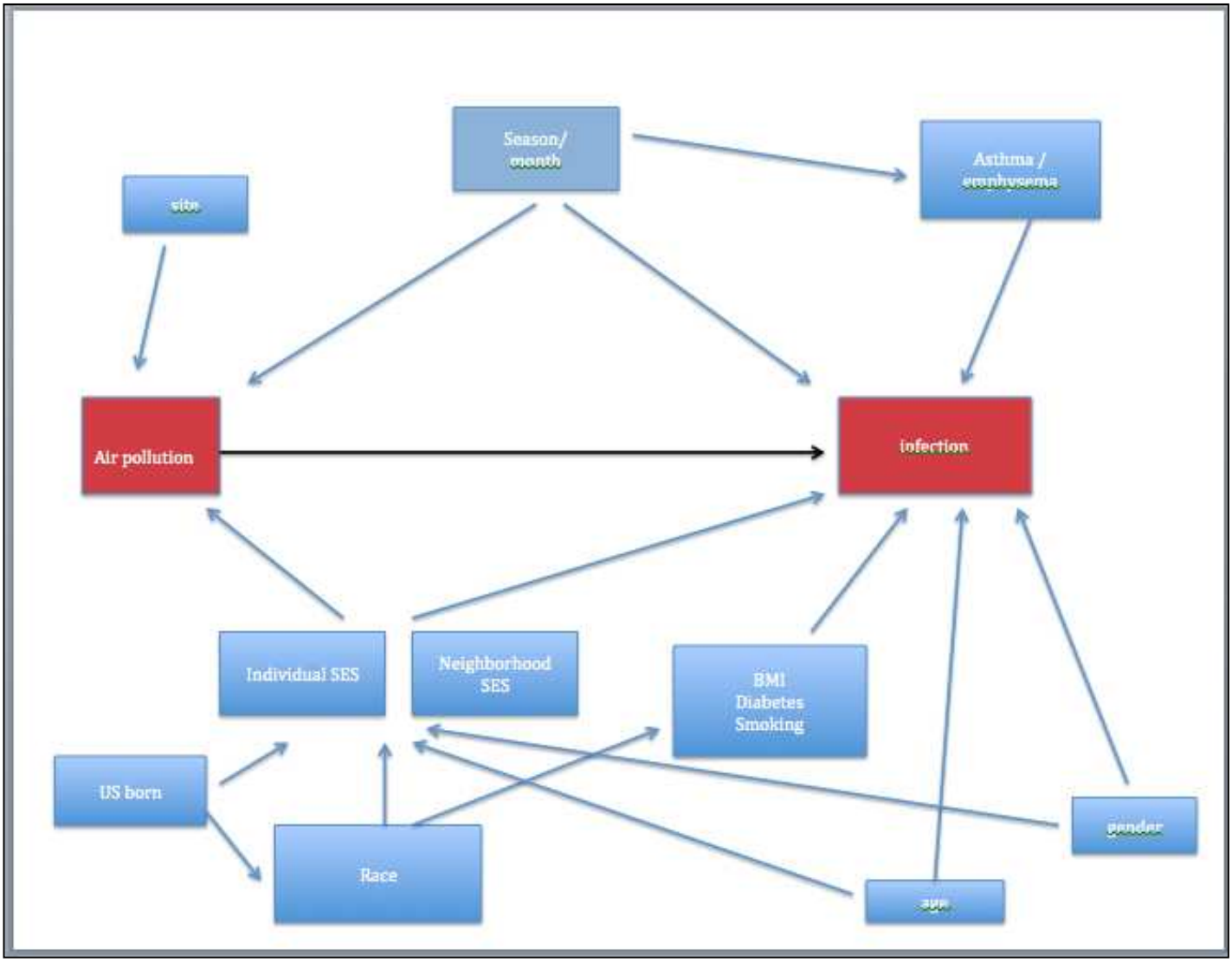
Directed Acyclic Graph

In order to decide which variables from the available MESA data to include in our model, a directed acyclic graph (DAG) was formulated (figure 1). The predictor of interest is "pollutant" so the model is consistent for all four pollutants studied here. Based on this diagram, study site is a variable that may directly influence exposure of air pollutants. Topography, population density, and other metropolitan characteristics likely have an impact on the concentration of pollutants to which study participants are exposed. Season and month during which participants were examined in likely related to both air pollution concentrations as well as to the likelihood of infection. Our study questions participants with regard to "recent" infection, i.e. within the past two weeks. Many illnesses, particularly those that are respiratory related, are well correlated with season. Respiratory viruses tend to occur more frequently in winter months, and therefore participants seen during these time periods are more likely to report recent respiratory infections. It is likely that weather patterns and other seasonal variations contribute to pollution levels as well. Weather patterns and temperature fluctuations can intensify or dilute air pollutant exposures. Participant comorbidities, particularly respiratory problems such as asthma, emphysema, and COPD, may be related to the prevalence of infection within these populations. Self-report of recent infection may also be dependent upon the sex of participants. Generally, females are more likely to report health problems. Additionally, participant sex may be related to individual socioeconomic status (SES). As our population is older, and aging over the course of the longitudinal study, there may be an increasing relationship between sex and SES. We assume age to also be related to self-report of recent infection as well as SES. Older individuals may suffer from less robust immune responses, have greater number of comorbidities, and may be more prone to infections. Older individuals may

also be more likely to report recent infections. We hypothesize that other comorbidities and health risk factors such as obesity (BMI>30), Diabetes Mellitus (DM), and cigarette smoking may also be related to an increased prevalence of recent infection. Additionally, age may be related to individual as well as neighborhood SES. As people retire from the work force, their SES may decline as they subsist on their retirement savings. In concert, this may force older people to move to less desirable neighborhoods, therefore incurring a lower neighborhood SES. In addition to age and gender, we theorize that race/ethnicity and country of birth may also be related to SES in this model. The relationship between race/ethnicity and both individual and neighborhood SES is well documented and well researched. Additionally, we hypothesize that foreign born participants may travel home to be with their families, thus exposing them to different and potentially more frequent infectious agents, possibly resulting in more frequent infections. We assume that individual and neighborhood SES will be related to both air pollutant exposure as well as report of recent infection. There are several variables in the MESA data set that can be used as a surrogate for individual SES: household income, level of education, employment status, number of people in the household, and type of living situation (rent, mortgage, own, other). Based on these *a priori* considerations we used a categorical value for household income as the variable representing individual SES for the main analyses. In exploratory analyses, the other available surrogate variables were also used, to determine if they provide a more powerful model. The variable for neighborhood SES was formulated by MESA researchers, as described above. The relationship between neighborhood SES (NSES) and health outcomes is well described. We know that lower SES neighborhoods tend to be in more polluted areas due to point-source emissions as well as closer to roadways. The stress pathway is also commonly cited as a mechanism by which NSES is associated with health.[32]

Finally, as this is a longitudinal analysis, many of our variables are time varying. We expected many of the variables to change with time as our participant's age. We used age in our model to represent this change in time.

Figure 2. Directed Acyclic Graph (DAG) detailing model with main effects and covariates



Results:

Predicted Pollutant Exposures:

From 1999 through 2012, there was a substantial decrease across study pollutants within all 6 communities, with regard to both annual pollutant predictions as well as two-week exposure predictions (Table 1, 2, Figures 3, 4). Table 1 details the trends in pollutants by study site over the study period, based on annual pollutant exposure predictions. Table 2 details these trends using the two-week predictions. As shown in Figures 3 and 4 all pollutants decreased in concentration over the study period (with the exception of BC because it is not measured as time varying). At the beginning of the study period, participants at the Los Angeles site had the overall highest levels of pollutant exposures, based on annual predictions. While all study sites exhibited a decrease in pollutant concentrations, LA experienced a 47% decrease in $PM_{2.5}$ levels over the study period. At the conclusion of the study, predicted $PM_{2.5}$ levels were highest among MESA participants at the New York site.

Overall, the one-year mean $PM_{2.5}$ estimates across the MESA study regions were 14.6 $\mu g/m^3$ [interquartile range (IQR 12.4-16.1, standard deviation (sd) 3.4]; 2 week mean $PM_{2.5}$ estimates were 14.2 $\mu g/m^3$ [IQR 10.8-16.8, sd 4.9]. The one-year mean NO_x level was 41.3 ppb (IQR 21.6-58.8, sd 24.2); the two-week mean NO_x level was 39.4 ppb (IQR 18.5-52.1, sd 29.3). The 1-year mean NO_2 level was 19.4 ppb (IQR 12.1-26.9, sd 8.9); the 2 week NO_2 estimates were 19.0 ppb (IQR 11.4 - 25.8, sd 9.5). The 1-year mean BC level was .82 $\mu g/m^3$ (IQR .51-1.2, sd .37); the 2 week BC mean was .82 $\mu g/m^3$ (IQR .5-1.2, sd .37). Correlations between all pairs of pollutants were significant, with both annual and two-week predictions. These correlations are shown in Table 3 and range from an r of 0.6 for BC and $PM_{2.5}$ to an r of 0.94 for NO_2 and NO_x. Table 4 provides the correlation between annual average concentrations and 2-week concentrations of each pollutant, which are also relatively high.

Table 1. Mean annual pollutant exposure predictions, by study site

	7/00-8/02	9/02-2/04	3/04-9/05	9/05-5/07	4/10-2/12
Winston					
$PM_{2.5}$ ($\mu g/m^3$)	15.4	13.8	14.2	14.3	10.4
NO_x (ppb)	22.4	17.2	17.1	14.9	9.5
NO_2 (ppb)	10.4	8.9	8.6	8.0	5.5
BC ($\mu g/m^3$)	0.50	0.50	0.50	0.50	0.50
NYC					
$PM_{2.5}$ ($\mu g/m^3$)	16.4	15.5	15.1	15.4	12.3
NO_x (ppb)	81.7	75.5	73.1	66.9	52.8
NO_2 (ppb)	32.8	31.7	30.7	29.9	26.0
BC (ppb)	1.29	1.29	1.29	1.29	1.27
Baltimore					
$PM_{2.5}$ ($\mu g/m^3$)	15.2	14.9	14.8	14.0	10.2
NO_x (ppb)	42.3	37.7	33.0	28.6	19.3
NO_2 (ppb)	17.8	18.2	16,6	14.5	11.0
BC ($\mu g/m^3$)	0.69	0.68	0.67	0.67	0.66
St. Paul					
$PM_{2.5}$ ($\mu g/m^3$)	12.1	10.7	9.8	9.9	9.7
NO_x (ppb)	24.0	22.5	22.6	21.8	16.5
NO_2 (ppb)	13.4	12.2	12.7	12.3	9.5
BC ($\mu g/m^3$)	0.45	0.45	0.45	0.45	0.44
Chicago					
$PM_{2.5}$ ($\mu g/m^3$)	16.2	15.3	13.7	14.3	11.7
NO_x (ppb)	42.9	43.0	37.6	33.8	22.0
NO_2 (ppb)	20.6	20.6	19.4	18.3	14.0
BC ($\mu g/m^3$)	0.65	0.65	0.65	0.65	0.65
LA					
$PM_{2.5}$ ($\mu g/m^3$)	22.3	20.9	19.1	16.7	11.8
NO_x (ppb)	78.7	67.6	58.3	55.0	34.1
NO_2 (ppb)	31.3	29.5	26.9	25.2	21.3
BC ($\mu g/m^3$)	1.23		1.22	1.22	1.22

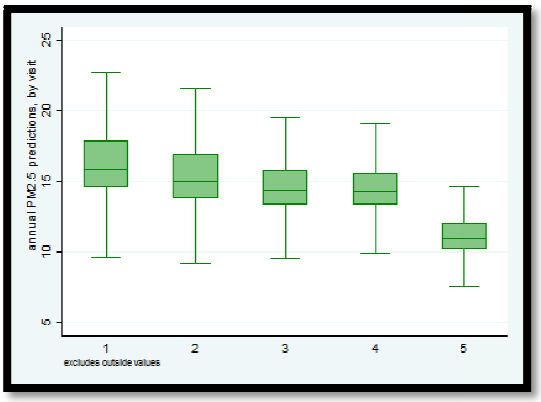
Table 2. Mean two-week levels of pollutant exposure predictions, by study site

	7/00-8/02	9/02-2/04	3/04-9/05	9/05-5/07	4/10-2/12
Winston					
$PM_{2.5}$ ($\mu g/m^3$)	15.0	13.9	14.8	13.7	10.6
NO_x (ppb)	21.1	15.9	13.0	15.3	7,7
NO_2 (ppb)	10.1	8.5	7.3	8.2	4.6
BC ($\mu g/m^3$)	0.5	0.5	0.5	0.5	0.5
NYC					
$PM_{2.5}$ ($\mu g/m^3$)	16.1	15.0	14.8	14.1	12.8
NO_x (ppb)	81	77.9	69.6	63.8	51.2
NO_2 (ppb)	32.8	31.4	30.5	29.0	25.8
BC ($\mu g/m^3$)	1.3	1.3	1.3	1.3	1.3
Baltimore					
$PM_{2.5}$ ($\mu g/m^3$)	14.8	14.3	15.2	12.9	9.9
NO_x (ppb)	44.1	37.6	28.6	29.6	19.1
NO_2 (ppb)	18.2	18.3	15.5	14.1	10.6
BC ($\mu g/m^3$)	0.7	0.7	0.7	0.7	0.7

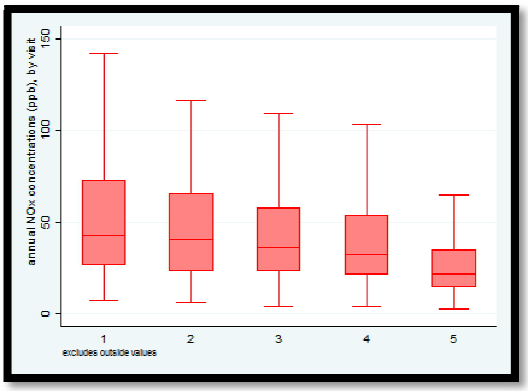
St. Paul					
$PM_{2.5}$ ($\mu g/m^3$)	11.8	10.7	9.5	9.7	8.3
NO_x (ppb)	22.9	24.9	20.6	22.9	13.7
NO_2 (ppb)	13.0	13.2	12.2	12.3	8.0
BC ($\mu g/m^3$)	0.5	0.5	0.4	0.4	0.4
Chicago					
$PM_{2.5}$ ($\mu g/m^3$)	15.9	15.3	13.8	13.1	11.4
NO_x (ppb)	42	43	35.1	34.7	19.5
NO_2 (ppb)	20.8	21.1	18.9	17.9	13.3
BC($\mu g/m^3$)	0.7	0.7	0.7	0.7	0.7
LA					
$PM_{2.5}$ ($\mu g/m^3$)	22.2	20.3	16.6	16.3	11.4
NO_x (ppb)	70.8	67.4	38,1	57.8	32.1
NO_2 (ppb)	30.2	29.7	23.4	26.0	20.5
BC ($\mu g/m^3$)	1.2	1.2	1.2	1.2	1.2

Figure 3. Annual Pollutant levels by study visit, at all sites

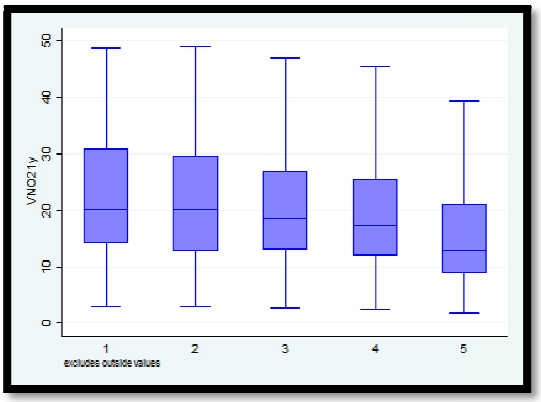
$PM_{2.5}$ per study visit



NO_x per study visit



NO_2 per study visit



BC per study visit

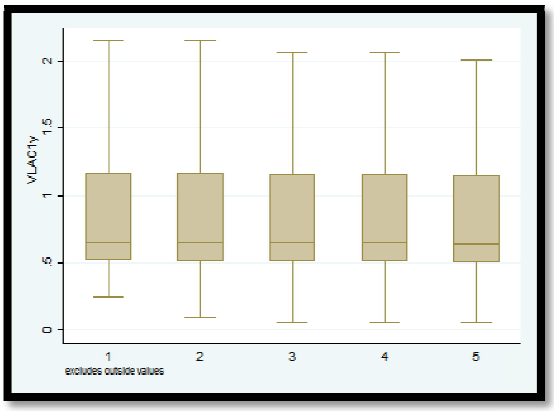
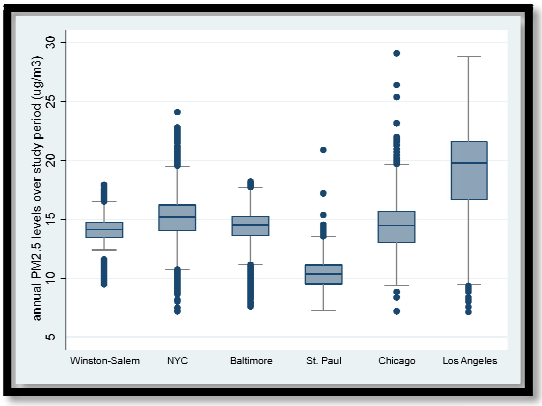
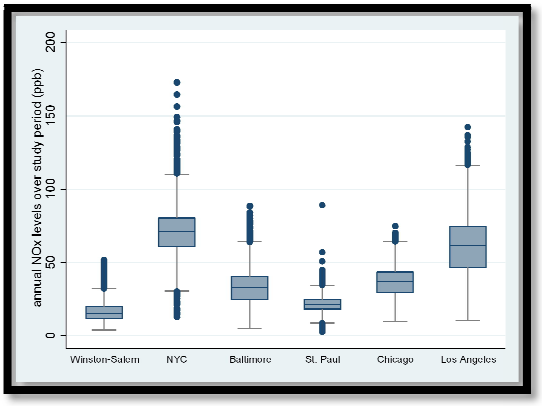


Figure 4: Annual Pollutant Levels by study site, all time points

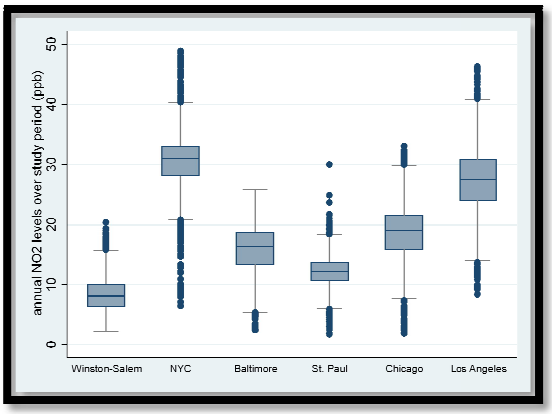
$PM_{2.5}$



NO_x



NO_2



BC

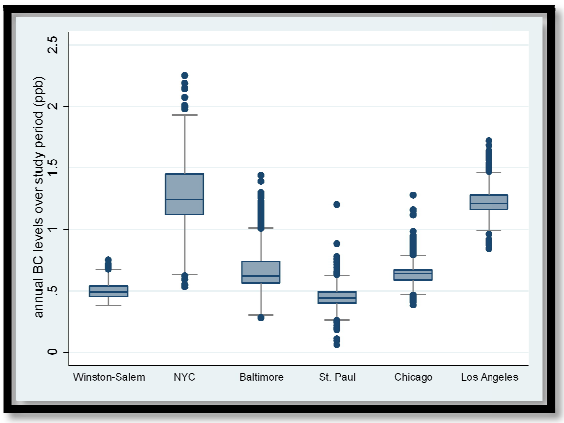


Table 3. Correlation between paired pollutants, using annual exposure predictions

	NO_x	NO_2	BC
$PM_{2.5}$	.71	.68	.60
NO_x		.94	.86
NO_2			.88

Table 4. Correlation between pollutant annual and two-week predictions

	$PM_{2.5}$	NO_x	NO_2	BC
$PM_{2.5}$	0.65			
NO_x		0.78		
NO_2			0.93	
BC				1.0

Subject characteristics:

Baseline data were available on 6804 MESA participants. Mean age of study participants was 62 years, 47.5 percent of participants were male. Ethnic distributions were sampled according to MESA protocol with 36.3% non-Hispanic white, 28.2% African American, 22.7% Hispanic, and 12.7% Asian. Additional characteristics of the study sample are shown in Table 5. The prevalence of current smoking was 13.3% and over half of the participants report never smoking (51%). Nearly 75% of the population was classified as overweight or obese based on BMI at the beginning of the study period. Twenty-six percent of the population was diabetic (including elevated fasting glucose) at the beginning of the study period. Less than 10% reported a diagnosis of asthma and less than 2% reported a diagnosis of emphysema or COPD. Nearly half of the participants reported employment at the beginning of the study period. Over 60% of the population reported being married at the beginning of the study period. About one-third of participants have graduated from college or university, with half of those obtaining graduate or professional education. As described in the methods, income categories were based on quartiles, as evident in the distribution shown here. Nearly 75% of the participants in this study lived alone or with one other individual.

Table 5. Participant characteristics (total N=6125) at the baseline examination 2000-2002

Participant Characteristics	
Participants at Visit 1	6125
Age (yrs)	61.6 (sd 10.2)
Race (%)	
White	36.3
Asian	12.7
Black	28.2
Hispanic	22.7
Male (%)	47.5
BMI (%)	
Normal (<25 kg/m^2)	26.0
Overweight (25-<30 kg/m^2)	35.0
Obese (30+ kg/m^2)	39.0
Asthma (%)	9.5
Emphysema (%)	1.4
Smoking status (%)	

Never	51
Former	35.7
Current	13.3
Diabetes Mellitus (%)	26.4
Non steroidal anti-inflammatory drugs(NSAIDS) (%)	17.2
Income (%)	
<$20,000	24.3
$20,000-<40,000	26.3
$40,000-<75,000	27
$75,000+	22.4
Education (%)	
Less than high school	18.7
High school	18.2
Some college / tech school	28.3
Bachelor's degree	17.2
Graduate/Professional degree	17.7
Current Job (%)	
Homemaker	11.5
Employed	49.7
Unemployed	2.3
Retired	36.5
Marital Status (%)	
Married	60.8
Unmarried	39.2
Number in Household (%)	
1	30.3
2	42.8
3	11.2
4	9.2
5 or more	6.5
Housing Type (%)	
Rent	29.5
Mortgage	39.6
Own	26.7
Other	04.2
Foreign born (%)	30.7
Il6 (mean) pg/ml	1.5 (1.2)
CRP (geometric mean) mg/L	1.9
Fibrinogen (mean) mg/dL	345.3 (73.8)

Prevalence of Health Outcomes:

Overall, there were 4782 participant infections reported over the entire study period. Some participants reported more than one infection at each visit, for a total of 5905 different respiratory related infections. Four hundred and seventy-seven fevers were reported over the study period. Cold and flu were reported 3,247 times by study participants. Participants reported 1800 sinus

infections over this time period. Three hundred and fifty eight cases of bronchitis and 23 cases of pneumonia were reported. Table 6 depicts the counts and prevalence of reported infections based on visit number.

Table 6. Percent and Counts of health outcomes by visit

	Visit 1 n=6120	Visit 2 n=5616	Visit 3 n=5373	Visit 4 n=5252	Visit 5 n=4288	Total N=26649
Age (mean)	61.6	63.1	64.4	65.9	69.4	64.6
Fever % (n)	2.1 (126)	2.1 (116)	1.7 (91)	1.7 (90)	1.3 (54)	1.8 (477)
Cold/flu % (n)	13.6 (835)	13.4 (755)	12.9 (691)	11.7 (617)	8.1 (349)	12.2 (3247)
Sinus % (n)	8.3 (506)	6.1 (345)	6.6 (352)	7.2 (376)	5.2 (221)	6.8 (1800)
Bronchitis % (n)	1.7 (105)	1.1 (63)	1.4 (76)	1.2 (64)	1.2 (50)	1.3 (358)
Pneumonia % (n)	0.05 (3)	0.1 (8)	0.04 (2)	0.1 (5)	0.1 (5)	0.1 (23)
Any infection % (n)	20.8 (1274)	18.7 (1048)	18.0 (968)	17.5 (920)	13.3 (572)	17.9 (4782)

Trends in infection data were evaluated by month and season of study. Figure 5 depicts trends of infection by study month and study season. All data from all visits are included here, presented by study site. The winter season (months Jan, Feb, March) has a higher number of infections.

Figure 5. Prevalence of Infection by month (season) of exam

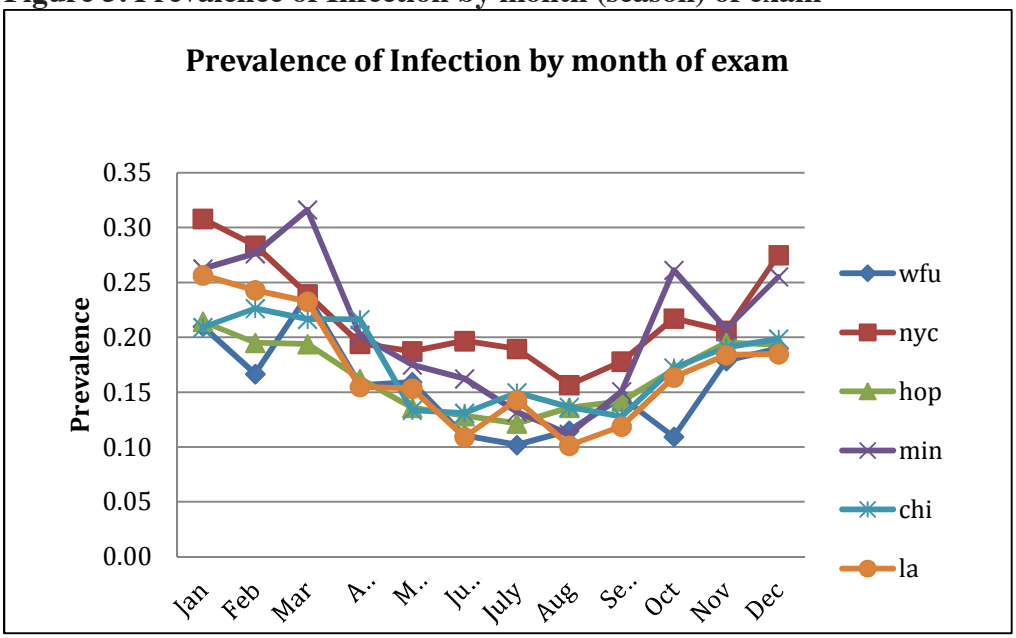


Table 7 provides the bivariate associations between covariates and the main outcome, report of any infection. Data from all five visit periods are represented here. For time-varying variables, data collected during the most recent exam are used. Based on these measures, Blacks and Hispanics have higher prevalence of infections compared to whites and men have lower prevalence of infection compared to women. When examining the relationship between prevalence of infection and study site, New York has the highest prevalence among sites, followed by St. Paul, MN. Winston-Salem, NC has the lowest prevalence. Foreign-born participants have a higher prevalence of reported infection than to US-born participants. The winter season has the highest prevalence of infection, followed by fall, spring, and summer in descending order. Both overweight and obese participants had higher prevalence of infection than did normal weight individuals, with obese participants having the highest prevalence. Participants with asthma as well as participants with emphysema or COPD had a higher prevalence than participants without these respiratory disorders. Diabetic participants had a higher prevalence of infection than non-diabetics. Former smokers had a lower prevalence of infection than did non-smokers, however, current smokers had the highest prevalence among these groups. The highest income category experienced the lowest prevalence of infection, followed in order by the other income categories, with the lowest income category reporting the highest prevalence of infection. Similarly, the highest level of education category experienced the lowest prevalence of infection, followed in order by the other level of education categories. The lowest level of education category experienced the highest prevalence of infection. Married participants had a lower prevalence of infection than non-married participants. No discernible trend in prevalence is seen when examining number of people living in a participant's household. The lowest prevalence rates were seen in the larger households, but this association was not significant.

Table 7. Prevalence ratios and 95% Confidence intervals of association between covariates and self-reported infection

	Prevalence ratios	CI	p-value
Age*	.92	.89-.95	<.001
Race			
White	1 (referent)		
Asian	0.97	0.86-1.09	.58
Black	1.26	1.16-1.37	<.001
Hispanic	1.41	1.29-1.53	<.001
Male	.79	.75-.83	<.001
Site			
Winston-Salem	1 (referent)		
NYC	1.58	1.37-1.81	<.001
Baltimore	1.13	0.97-1.31	.12
St. Paul	1.46	1.27-1.68	<.001
Chicago	1.16	1,01-1.35	.04
Los Angeles	1.13	0.98-1.30	.1
Country of Birth			
US	1 (referent)		
Foreign	1.2	1.02-1.21	.011
Season			
Winter	1 (referent)		
Spring	0.61	0.56-0.66	<.001
Summer	0.49	0.45-0.54	<.001
Fall	0.78	0.72-0.85	<.001
BMI			
Normal (<25 kg/m^2)	1 (referent)		
Overweight (25-<30 kg/m^2)	1.13	1.04-1.22	.002
Obese (30+ kg/m^2)	1.23	1.13-1.33	<.001
Asthma			
No	1 (referent)		
Yes	1.79	1.62-1.96	<.001
Emphysema			
No	1 (referent)		
Yes	1.79	1.53-2.11	<.001
Diabetes			
No	1 (referent)		
Yes	1.12	1.06-1.18	<.001
Smoking Status			
Never	1 (referent)		
Former	0.92	.85-1.0	.06
Current	1.35	1.2-1.53	<.001
Income			
<$20,000	1 (referent)		
$20,000-<40,000	.93	.83-1.03	.19
$40,000-<$75,000	.83	.74-.93	.001
$75,000+	.64	.57-.72	<.001
Education level			
Less than high school	1 (referent)		

High school	0.85	0.77-0.94	.002
Some college / tech	0.85	0.77-0.94	.001
Bachelor's degree	0.76	0.69-0.85	<.001
Grad/Prof degree	0.73	0.66-0.81	<.001
Marital status			
Unmarried	1 (referent)		
Married	.82	.74-.91	<.001
Number in Household			
1	1 (referent)		
2	0.86	0.79-0.92	<.001
3	1.0	0.90-1.11	.997
4	0.96	0.85-1.08	.48
5+	1.08	0.95-1.24	.24

* Association reflects that over a 10 year increase in age

Table 8 displays the differences in mean inflammatory marker values among participants reporting infection and those not reporting infection. These inflammatory markers (fibrinogen, IL-6, and CRP) were only measured at the first study period. With all three makers, participants reporting recent infection had a significantly increased level of inflammatory marker.

Table 8. Differences in Inflammatory Markers by Infection (only available for first study visit)

	Infection, 95% CI		p value
	Yes	No	
Fibrinogen (mg/dL)	360.8 (356.5, 365.0)	345.0 (343.0, 347.0)	<.001
IL-6 (pg/ml)	1.68 (1.61,1.76)	1.5 (1.46, 1.53)	<.001
CRP (mg/L)*	.84 (.78, .9)	.58 (.55, .61)	<.001

* log transformed

Association between long-term Air Pollution Exposure and self-reported infections

Regulations and environmental trends have contributed to a gradual decrease in ambient pollution concentrations in the US, a trend that is discernible over the study period. Given this, we wanted to investigate the cross-sectional association between reported infection and pollution level, which allows us to control for this gradual improvement in air quality, which may otherwise confound our results. These results are shown below in Table 9. Although there is no significant trend to these results, we do have a general increased prevalence of reported infection associated with increased levels of pollution. The strongest association is seen with NO_x as the visit specific models reach statistical significance for all time points except visit 3.

Table 9. Prevalence Ratios between reported infection and level of pollutant, by study visit

Prevalence Ratio (95% CI)	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
$PM_{2.5}$ (IQR = 4μg/m^3)	1.11 (.97, 1.26)	1.09 (.91, 1.31)	.9 (.76, 1.07)	1.25 (1.03,1.51)*	1.1 (.86,1.41)
NO_x (IQR = 25 ppb)	1.17 (1.05,1.32)*	1.19 (1.05,1.35)*	1.08 (.95, 1.22)	1.25 (1.09,1.45)*	1.35 (1.04, 1.69)*
NO_2 (IQR = 9 ppb)	1.1 (.97,1.29)*	1.19 (1.03,1.39)*	1.03 (.89, 1.19)	1.23 (1.05, 1.45)*	1.19 (.97, 1.45)
BC (IQR = 0.3 μg/m^3)	1.1 (1.01,1.19)*	1.08 (.98,1.19)	1.08 (.98, 1.19)	1.12 (1.0, 1.25)*	1.07 (.95, 1.21)

**** controlling for season, site, age**

Table 10 provides the prevalence ratios for the multivariate models developed for each pollutant. The annual average air pollution concentrations were used in these models. The IQR for each pollutant for the entire study period was used to interpret the prevalence ratio. Season, smoking status, income, neighborhood SES, age, race, gender, and study site were adjusted for in the models. Model 4 is the primary model used in this analysis. The subsequent table (table 10) examines specific infections (e.g. fever) and adjusts for the same covariates as this model. Based on this model, the prevalence of any infection is higher for persons exposed to increased levels of air pollution. For example, a 1 IQR increase in $PM_{2.5}$ is associated with a 14% higher prevalence of infection (95% CI: 1.08, 1.19). We see even larger effects for NO_x and NO_2 where a 1 IQR increase is associated with a 16% and 17% higher prevalence respectively. All three of these prevalence ratios reach statistical significant. With black carbon, a 1 IQR increase is associated with an 4% higher prevalence of infection, however this does not reach statistical significance.

Table 10. Prevalence Ratios for the association between annual average air pollution concentrations and prevalence of any infection

Prevalence Ratio (95% CI)	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]
$PM_{2.5}$ (IQR = 4μg/m^3)	1.16 (1.11-1.22)*	1.13 (1.08-1.19)*	1.16 (1.11-1.22)*	1.14 (1.08, 1.19)*
NO_x (IQR = 25 ppb)	1.19 (1.13,1.25)*	1.16 (1.08-1.22)*	1.19 (1.13,1.25)*	1.16 (1.08-1.22)*
NO_2 (IQR = 9 ppb)	1.21 (1.13-1.29)*	1.17(1.08,1.24)*	1.21 (1.13-1.29)*	1.17(1.08,1.24)*
BC (IQR = 0.3 μg/m^3)	1.07 (1.02-1.13)*	1.04 (0.99-1.1)	1.06 (1.00-1.12)*	1.04 (.98,1.1)

[a] Model 1 includes age, race/ethnicity, gender, season, site, smoking status
[b] Model 2…controlling for model 1 + income
[c] Model 3…controlling for model 2 + neighborhood SES
[d] Model 4….controlling for model 1 + income + neighborhood SES
* indicates significance at an alpha<.05

Table 11 provides the prevalence ratios for the multivariate models developed for each pollutant. The two-week average air pollution concentrations were used in these models. The same IQR used previously for each pollutant was used in this model. The same variables were used in the step-wise models and model 4 is the main model used in this analysis. Based on this model, the prevalence of any infection is higher for persons exposed to increased levels of air pollution. For example, a 1 IQR increase in $PM_{2.5}$ is associated with a 2% higher prevalence of infection (95% CI: 1.0, 1.04). Larger effects are seen with NO_x and NO_2 where a 1 IQR increase is associated with a 13% and 10% higher prevalence respectively. With black carbon, a 1 IQR increase is associated with an 5% higher prevalence of infection. Only the models describing NO_x and NO_2 reach statistical significance.

Table 11. Prevalence Ratios for the association between two-week average air pollution concentrations and prevalence of any infection

Prevalence Ratio (95% CI)	Model 1[a]	Model 2[b]	Model 3[c]	Model 4[d]
$PM_{2.5}$ (IQR = 4μg/m^3)	1.03 (1.01,1.05)*	1.02 (1.0,1.04)	1.03 (1.01,1.05)*	1.02 (1.0,1.04)
$NO_x^{\#}$ (IQR = 25 ppb)	1.16 (1.11,1.22)*	1.13 (1.08,1.19)*	1.16 (1.11,1.22)*	1.13 (1.08,1.19)*
$NO_2^{\#}$ (IQR = 9 ppb)	1.13 (1.08,1.2)*	1.1 (1.06,1.17)*	1.1 (1.08,1.2)*	1.1(1.06,1.17)*
BC (IQR = 0.3 μg/m^3)	1.08 (1.02,1.14)*	1.05 (.99,1.11)	1.07 (1.01,1.1)*	1.05 (.99,1.1)

[a] Model 1 includes age, race/ethnicity, gender, season, site, smoking status
[b] Model 2…controlling for model 1 + income
[c] Model 3…controlling for model 2 + neighborhood SES
[d] Model 4….controlling for model 1 + income + neighborhood SES
* indicates significance at an alpha<.05

Table 12 provides the prevalence ratios for the multivariate model describing the association between air pollutants and specific infections. For each of these, the results using model 4, as described above, are displayed. Based on these models, the prevalence of specific infections are higher for persons exposed to increased levels of pollution. For $PM_{2.5}$ and NO_x, the strongest association is seen with participants reporting recent bronchitis or pneumonia with a 15% and 27% increase per IQR, respectively. With NO_2 the strongest association is seen with those reporting recent cold/flu, a 19% increase per IQR. The association with black carbon was highest with participants reporting recent fevers at 7%.

Table 12. Prevalence Ratios for the association between annual average air pollution concentrations and prevalence of fever, cold/flu, sinus infections, and bronchitis/pneumonia

Prevalence Ratio (95% CI)	Fever	Cold/Flu	Sinus Infection	Bronchitis/Pneumonia
$PM_{2.5}$ (IQR = 4 $\mu g/m^3$)	1.18(.98,1.41)*	1.15(1.08,1.22)*	1.12(1.02,1.24)*	1.15(.97,1.36)
NO_x (IQR = 25 ppb)	1.22(1.02,1.49)*	1.16(1.1,1.24)*	1.19(1.07,1.35)*	1.27(1.03,1.57)*
NO_2 (IQR = 9 ppb)	1.16 (.94, 1.45)	1.19 (1.1, 1.29)*	1.12 (.98, 1.29)	1.12 (.88,1.43)
BC (IQR = 0.3 $\mu g/m^3$)	1.07 (.89, 1.28)	1.05 (.98, 1.13)	1.04 (.92, 1.18)	1.0 (.81, 1.24)

** all models are adjusted for age, race/ethnicity, gender, season, site, smoking status, income, and neighborhood SES
* indicates significance at an alpha<.05

Table 13 provides prevalence ratios from a sensitivity analysis evaluating other variables that may be used in our model to represent individual level SES. These models used the complete covariates listed above for model 4 and annual pollutant exposures. Income was substituted for highest education level, current job, and home type. Education and current job result in increased prevalence of infection per IQR increase in pollutant. Controlling for home type provided no real change from the original model.

Table 13. Sensitivity analysis substituting variables to indicate individual SES

Prevalence Ratio (95% CI)	Model 4	Education#	Current job#	Home type#
$PM_{2.5}$ (IQR = 4 $\mu g/m^3$)	1.02 (1.0,1.04)	1.15 (1.1, 1.21)*	1.15 (1.1, 1.21)*	1.01 (.93, 1.1)
NO_x (IQR = 25 ppb)	1.13 (1.08,1.19)*	1.18 (1.1,1.25)*	1.18 (1.1,1.25)*	1.08 (1.0, 1.16)
NO_2 (IQR = 9 ppb)	1.1(1.06,1.17)*	1.19 (1.11, 1.28)*	1.19 (1.11, 1.28)*	1.06 (.96, 1.15)
BC (IQR = 0.3 $\mu g/m^3$)	1.05 (.99,1.1)	1.05 (.99, 1.11)	1.06 (1.0,1.1)*	1.05 (.98, 1.1)

variable was substituted for personal income in each model

Discussion

In a large population-based cohort study of adults without preexisting cardiovascular disease, we found robust evidence that the prevalence of respiratory infection is increased in participants exposed to higher levels of ambient air pollution. Our associations are evident across all four pollutants used in this analysis with the strongest associations observed with increasing levels of the traffic-related air pollutants NO_x and NO_2. These associations were evident between these pollutants and all outcomes: any infection, fever, cold/flu, sinus infection, and bronchitis/pneumonia. Additionally, as the study period progressed, ambient pollution levels declined across all sites. Similarly, the prevalence of infection declined as our study progressed. However, our cross-sectional analysis accounts for these secular trends and yet the association between pollution levels and reported infections remains.

Overall, every model showed an association between increased pollutant levels and increased reporting of respiratory infections. Increased prevalence of infection was associated with a correlated set of pollutants that included nitrogen dioxide, oxides of nitrogen, fine particulate matter, and elemental carbon. As our pollutant levels are highly correlated, we were unable to discern a difference as to the independent effects of each pollutant on respiratory associated infections.

Our main effects models include: one model using annual pollutant predictions, one model using two-week pollutant predictions. Results from table 4 indicated high correlations between these annual and two-week predictions. Similarly, our main effects models show similar prevalence ratios. Results concerning prevalence of infection associated with a 1 IQR increase in NO_x or NO_2 are very similar between the annual and two-week exposure models, as are the BC. The $PM_{2.5}$ models show a more attenuated prevalence ratio when using the two-week exposure predictions. This difference may be due to residual confounding in our two-week models including ambient temperature, wind, and humidity.

The greatest strength of this study is the availability of air pollutant estimates that enable participant-level exposure predictions to be made. As opposed to other studies that use city-level data, our data enables much more specific predictions to be generated and gives a better understanding of within-city ambient pollution variability. Additionally, this study benefited from data collected from a long-term, longitudinal follow-up of a large cohort, with exposure and outcome data collected in a consistent manner throughout the study period. Such robust data collection, particularly standardized data collection, was a great asset to this analysis.

Previous studies have evaluated the association between ambient air pollution and respiratory infections as determined by ICD-9 coding during hospital admissions or emergency room visits. Our study provided further support that increasing levels of pollution are associated with an increased prevalence of respiratory infections. Our study allows us to capture respiratory infections that may be less serious or severe, yet remain important to consider when evaluating the effects of pollution. While our study ideally captures these participants as well as less serious infections, we are only capturing an abbreviated period of their health history.

A limitation of this study is its reliance on participant self-report of infection. This measure is not validated and no sort of medical record or other data are available to assess validity. However, at the time of the first visit, inflammatory markers were collected. Based on the ANOVA of inflammatory marker levels and report of recent infection (at the first visit), all three markers were significantly higher in the group reporting recent infection at this first visit. Such an association provides some evidence that our measure of self-report may have some validity.

The mechanisms by which ambient air pollutant exposure could lead to respiratory infection has been hypothesized and include many possibilities including systemic inflammation,

immunologic mediators, and epigenic mechanisms. Further research regarding the roles of specific pollutants in infection is warranted.

Conclusion

We have shown that exposure to ambient air pollution is correlated with significant increase in prevalence of respiratory infections among healthy adults. This evidence adds to the literature regarding the association of ambient air pollution with adverse heath outcomes. The exposure prediction methods used provide valuable strength to our models and results.

References

1. The Clean Air Act in a Nutshell: How it Works. http://www.epa.gov/air/caa/pdfs/CAA_Nutshell.pdf.
2. Bascom, R., Health effects of outdoor air pollution. *American Journal of Respiratory and Critical Care Medicine* **1996,** *153* (1), 3-50.
3. National Ambient Air Quality Standards (NAAQS). http://www.epa.gov/air/criteria.html (accessed May 12).
4. Brook, R. D.; Rajagopalan, S.; Pope, C. A., 3rd; Brook, J. R.; Bhatnagar, A.; Diez-Roux, A. V.; Holguin, F.; Hong, Y.; Luepker, R. V.; Mittleman, M. A.; Peters, A.; Siscovick, D.; Smith, S. C., Jr.; Whitsel, L.; Kaufman, J. D., Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **2010,** *121* (21), 2331-78.
5. Counties Designated "Nonattainment" for Clean Air Act's National Ambient Air Quality Standards (NAAQS). http://www.epa.gov/airquality/greenbk/map/mapnpoll.pdf (accessed May 27).
6. Brunekreef, B.; Holgate, S. T., Air pollution and health. *Lancet* **2002,** *360* (9341), 1233-42.
7. Kelly, F. J.; Fussell, J. C., Air pollution and airway disease. *Clin Exp Allergy* **2011,** *41* (8), 1059-71.
8. Dockery, D. W.; Pope, C. A., 3rd; Xu, X.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. G., Jr.; Speizer, F. E., An association between air pollution and mortality in six U.S. cities. *N Engl J Med* **1993,** *329* (24), 1753-9.
9. Schwartz, J.; Slater, D.; Larson, T. V.; Pierson, W. E.; Koenig, J. Q., Particulate air pollution and hospital emergency room visits for asthma in Seattle. *Am Rev Respir Dis* **1993,** *147* (4), 826-31.
10. Peel, J. L.; Tolbert, P. E.; Klein, M.; Metzger, K. B.; Flanders, W. D.; Todd, K.; Mulholland, J. A.; Ryan, P. B.; Frumkin, H., Ambient air pollution and respiratory emergency department visits. *Epidemiology* **2005,** *16* (2), 164-74.
11. Spix, C.; Anderson, H. R.; Schwartz, J.; Vigotti, M. A.; LeTertre, A.; Vonk, J. M.; Touloumi, G.; Balducci, F.; Piekarski, T.; Bacharova, L.; Tobias, A.; Ponka, A.; Katsouyanni, K., Short-term effects of air pollution on hospital admissions of respiratory diseases in Europe: a quantitative summary of APHEA study results. Air Pollution and Health: a European Approach. *Arch Environ Health* **1998,** *53* (1), 54-64.
12. Chauhan, A. J.; Johnston, S. L., Air pollution and infection in respiratory illness. *Br Med Bull* **2003,** *68*, 95-112.
13. Neupane, B.; Jerrett, M.; Burnett, R. T.; Marrie, T.; Arain, A.; Loeb, M., Long-term exposure to ambient air pollution and risk of hospitalization with community-acquired pneumonia in older adults. *Am J Respir Crit Care Med* **2010,** *181* (1), 47-53.
14. Brauer, M.; Hoek, G.; Van Vliet, P.; Meliefste, K.; Fischer, P. H.; Wijga, A.; Koopman, L. P.; Neijens, H. J.; Gerritsen, J.; Kerkhof, M.; Heinrich, J.; Bellander, T.; Brunekreef, B., Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children. *Am J Respir Crit Care Med* **2002,** *166* (8), 1092-8.
15. Brauer, M.; Hoek, G.; Smit, H. A.; de Jongste, J. C.; Gerritsen, J.; Postma, D. S.; Kerkhof, M.; Brunekreef, B., Air pollution and development of asthma, allergy and infections in a birth cohort. *Eur Respir J* **2007,** *29* (5), 879-88.
16. MacIntyre, E. A.; Gehring, U.; Molter, A.; Fuertes, E.; Klumper, C.; Kramer, U.; Quass, U.; Hoffmann, B.; Gascon, M.; Brunekreef, B.; Koppelman, G. H.; Beelen, R.; Hoek, G.; Birk, M.; de Jongste, J. C.; Smit, H. A.; Cyrys, J.; Gruzieva, O.; Korek, M.; Bergstrom, A.; Agius, R. M.; de Vocht, F.; Simpson, A.; Porta, D.; Forastiere, F.; Badaloni, C.; Cesaroni, G.; Esplugues, A.;

Fernandez-Somoano, A.; Lerxundi, A.; Sunyer, J.; Cirach, M.; Nieuwenhuijsen, M. J.; Pershagen, G.; Heinrich, J., Air pollution and respiratory infections during early childhood: an analysis of 10 European birth cohorts within the ESCAPE Project. *Environ Health Perspect* **2014,** *122* (1), 107-13.
17. Mehta, S.; Shin, H.; Burnett, R.; North, T.; Cohen, A. J., Ambient particulate air pollution and acute lower respiratory infections: a systematic review and implications for estimating the global burden of disease. *Air Qual Atmos Health* **2013,** *6* (1), 69-83.
18. Ackermann-Liebrich, U.; Leuenberger, P.; Schwartz, J.; Schindler, C.; Monn, C.; Bolognini, G.; Bongard, J. P.; Brandli, O.; Domenighetti, G.; Elsasser, S.; Grize, L.; Karrer, W.; Keller, R.; Keller-Wossidlo, H.; Kunzli, N.; Martin, B. W.; Medici, T. C.; Perruchoud, A. P.; Schoni, M. H.; Tschopp, J. M.; Villiger, B.; Wuthrich, B.; Zellweger, J. P.; Zemp, E., Lung function and long term exposure to air pollutants in Switzerland. Study on Air Pollution and Lung Diseases in Adults (SAPALDIA) Team. *Am J Respir Crit Care Med* **1997,** *155* (1), 122-9.
19. Ciencewicki, J.; Jaspers, I., Air pollution and respiratory viral infection. *Inhal Toxicol* **2007,** *19* (14), 1135-46.
20. Jaakkola, J. J.; Paunio, M.; Virtanen, M.; Heinonen, O. P., Low-level air pollution and upper respiratory infections in children. *Am J Public Health* **1991,** *81* (8), 1060-3.
21. Wong, T. W.; Lau, T. S.; Yu, T. S.; Neller, A.; Wong, S. L.; Tam, W.; Pang, S. W., Air pollution and hospital admissions for respiratory and cardiovascular diseases in Hong Kong. *Occup Environ Med* **1999,** *56* (10), 679-83.
22. O'Neill, M. S.; Jerrett, M.; Kawachi, I.; Levy, J. I.; Cohen, A. J.; Gouveia, N.; Wilkinson, P.; Fletcher, T.; Cifuentes, L.; Schwartz, J., Health, wealth, and air pollution: advancing theory and methods. *Environ Health Perspect* **2003,** *111* (16), 1861-70.
23. Mohai, P.; Lantz, P. M.; Morenoff, J.; House, J. S.; Mero, R. P., Racial and socioeconomic disparities in residential proximity to polluting industrial facilities: evidence from the Americans' Changing Lives Study. *Am J Public Health* **2009,** *99 Suppl 3*, S649-56.
24. Clark, L. P.; Millet, D. B.; Marshall, J. D., National Patterns in Environmental Injustice and Inequality: Outdoor NO2 Air Pollution in the United States. *PLoS One* **2014,** *9* (4), e94431.
25. Cohen, M. A.; Adar, S. D.; Allen, R. W.; Avol, E.; Curl, C. L.; Gould, T.; Hardie, D.; Ho, A.; Kinney, P.; Larson, T. V.; Sampson, P.; Sheppard, L.; Stukovsky, K. D.; Swan, S. S.; Liu, L. J.; Kaufman, J. D., Approach to estimating participant pollutant exposures in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Environ Sci Technol* **2009,** *43* (13), 4687-93.
26. Gill, E. A.; Curl, C. L.; Adar, S. D.; Allen, R. W.; Auchincloss, A. H.; O'Neill, M. S.; Park, S. K.; Van Hee, V. C.; Diez Roux, A. V.; Kaufman, J. D., Air pollution and cardiovascular disease in the Multi-Ethnic Study of Atherosclerosis. *Prog Cardiovasc Dis* **2011,** *53* (5), 353-60.
27. Van Hee, V. C.; Adar, S. D.; Szpiro, A. A.; Barr, R. G.; Bluemke, D. A.; Diez Roux, A. V.; Gill, E. A.; Sheppard, L.; Kaufman, J. D., Exposure to traffic and left ventricular mass and function: the Multi-Ethnic Study of Atherosclerosis. *Am J Respir Crit Care Med* **2009,** *179* (9), 827-34.
28. Park, S. K.; Auchincloss, A. H.; O'Neill, M. S.; Prineas, R.; Correa, J. C.; Keeler, J.; Barr, R. G.; Kaufman, J. D.; Diez Roux, A. V., Particulate air pollution, metabolic syndrome, and heart rate variability: the multi-ethnic study of atherosclerosis (MESA). *Environ Health Perspect* **2010,** *118* (10), 1406-11.
29. Hajat, A.; Diez-Roux, A. V.; Adar, S. D.; Auchincloss, A. H.; Lovasi, G. S.; O'Neill, M. S.; Sheppard, L.; Kaufman, J. D., Air pollution and individual and neighborhood socioeconomic status: evidence from the Multi-Ethnic Study of Atherosclerosis (MESA). *Environ Health Perspect* **2013,** *121* (11-12), 1325-33.
30. Szpiro, A. A.; Sampson, P. D.; Sheppard, L.; Lumley, T.; Adar, S. D.; Kaufman, J., Predicting Intra-Urban Variation in Air Pollution Concentrations with Complex Spatio-Temporal

Dependencies. *Environmetrics* **2009,** *21* (6), 606-631; Keller, J. P.; Olives, C.; Kim, S.-Y.; Sheppard, L.; Sampson, P. D.; Spziro, A. A.; Oron, A. P.; Linstrom, J.; Vedal, S.; Kaufman, J. D., A Unified Spatiotemporal Modeling Approach for Prediction of Multiple Air Pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution. *Environmental Health Perspectives* **2014,** *In Press*.
31. Szpiro, A. A.; Sheppard, L.; Adar, S. D.; Kaufman, J. D., Estimating acute air pollution health effects from cohort study data. *Biometrics* **2014,** *70* (1), 164-74.
32. Steptoe, A.; Feldman, P. J., Neighborhood problems as sources of chronic stress: development of a measure of neighborhood problems, and associations with socioeconomic status and health. *Ann Behav Med* **2001,** *23* (3), 177-85.