APPROACH TO ESTIMATING PARTICIPANT POLLUTANT
EXPOSURES IN THE MULTI-ETHNIC STUDY OF
ATHEROSCLEROSIS AND AIR POLLUTION (MESA AIR)

Martin A. Cohen1*, Sara D. Adar1, Ryan W. Allen1,2, Edward Avol3, Cynthia L. Curl1,
Timothy Gould4, David Hardie1, Anne Ho,1 Patrick Kinney5, Timothy V. Larson4, Paul
Sampson6, Lianne Sheppard7, Karen D. Stukovsky7, Susan S. Swan1,8, L-J Sally Liu1,9,
Joel D. Kaufman1

1Department of Environmental and Occupational Health Sciences, University of
Washington, Seattle, WA, USA
2Current address: Faculty of Health Sciences, Simon Fraser University, Burnaby, BC,
Canada
3Department of Preventive Medicine, University of Southern California, Los Angeles,
CA, USA
4Department of Civil and Environmental Engineering, University of Washington, Seattle,
WA, USA
5Department of Environmental Health Sciences, Columbia University, New York, NY,
USA
6Department of Statistics, University of Washington, Seattle, WA, USA
7Department of Biostatistics, University of Washington, Seattle, WA, USA
8Current address: MDS Pharma Services, Inc., Bothell, WA, USA
9Current address: Institute of Social and Preventive Medicine, University of Basel, Basel,
Switzerland

Submitted to Environmental Science & Technology
November, 2008

*Corresponding Author: Martin Cohen, mcohen@u.washington.edu, Phone: 206-616-
1905
Abstract

Most published epidemiology studies of long-term air pollution health effects have relied on central site monitoring data to understand the health impacts of regional-scale differences in exposure. Few cohort studies have had sufficient data to characterize localized variations in pollution, despite the fact that large gradients can exist over small spatial scales. Similarly, previous data have generally been limited to measurements of particle mass or several of the criteria gases. The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air) is an innovative investigation undertaken to link subclinical and clinical cardiovascular health effects with individual-level estimates of personal exposure to ambient-origin pollution. This project improves on prior work by implementing an extensive exposure assessment program to characterize long-term average concentrations of ambient-generated PM$_{2.5}$, specific PM$_{2.5}$ chemical components, and co-pollutants, with particular emphasis on capturing concentration gradients within cities.

This paper describes exposure assessment in MESA Air, including questionnaires, community sampling, home monitoring, and personal sampling. Summary statistics describing the performance of the sampling methods are presented along with descriptive statistics of the air pollution concentrations by city.
Introduction

Several epidemiologic studies have reported associations between long-term exposure to fine particulate matter (PM2.5) and cardiovascular morbidity and mortality [1]. One mechanism through which chronic exposures are hypothesized to affect cardiac health is through acceleration of atherosclerosis. This hypothesis is supported by evidence from animal toxicological investigations [2-4] and cross-sectional epidemiologic studies of prevalent atherosclerosis. [5-8]. Since no published studies have yet evaluated air pollution and the progression of atherosclerosis, the Multi-Ethnic Study of Atherosclerosis and Air Pollution (henceforth “MESA Air”) was initiated in 2004 to address this question. The primary aims of the ten-year project are to examine the impacts of chronic exposures to PM2.5 of ambient origin on the progression of atherosclerosis and the incidence of cardiovascular disease. Ambient-origin PM2.5 is specifically defined in this study as both outdoor PM2.5 and the fraction of outdoor PM2.5 that has infiltrated into indoor environments. Secondary aims of MESA Air and its ancillary studies are to evaluate associations between specific gases and particulate components and cardiovascular disease.

In order to meet these aims, MESA Air has developed an exposure assessment methodology that will allow for participant-specific estimates of long-term exposure to PM2.5 of ambient origin during the time period of interest. The concept that ambient exposure should be considered a key metric in air pollution epidemiology is not new [9], and ambient exposures are known to be highly correlated with ambient concentrations [10] [11]. However, in comparison with ambient concentrations, ambient exposures have shown larger health effect estimates with smaller confidence intervals [12].
Previous chronic air pollution studies have predominantly relied on limited regulatory monitoring data and have assigned one ambient concentration as an assumed exposure to numerous participants. We designed an extensive monitoring campaign to better capture within-city variations in pollution, allowing for the resolution of concentrations at a relatively small spatial scale (i.e. 10’s of meters rather than kilometers). Within-city differences are an important focus of our exposure assessment since recent evidence suggests that local concentration gradients are linked to differential cardiovascular outcomes [1, 13, 14].

Another key feature of MESA Air is the inclusion of participant-specific residential infiltration efficiencies and time-location patterns. These unique data allow us to estimate exposures to ambient-origin air pollution rather than using ambient concentrations as surrogates for ambient [11, 15]. This is a potentially important refinement from past research, since individuals spend the vast majority of their time indoors [16], ambient PM$_{2.5}$ concentrations are attenuated as they move indoors, and attenuations differ between homes and over time within homes [17-19].

Since it was not feasible to measure exposures for all 6,200 MESA Air participants, a modeling approach was selected to characterize each participant’s long-term exposure to ambient-origin air pollution. This paper describes the general exposure assessment approach of MESA Air, with an emphasis on the air pollution measurements collected, chemical analyses employed, and quality assurance/quality control (QA/QC) procedures utilized. Brief information on the cohort and the spatiotemporal modeling methods are presented, with more detailed information on the epidemiologic design including the statistical methodology to be utilized for assessing health effects in
Kaufman et al. (in preparation). A description of the ambient concentration models is presented in Szpiro et al. [20].

**Study Population and Geographic Areas**

MESA Air is an ancillary study to the Multi-Ethnic Study of Atherosclerosis (MESA) described by Bild and colleagues [21]. The primary purpose of MESA was to characterize subclinical cardiovascular disease and its associated risk factors among African American, Caucasian, Chinese, and Hispanic persons aged 45-84 and free of clinical cardiovascular disease at recruitment. MESA recruited a population-based sample of men and women from 6 U.S. metropolitan areas: Baltimore City and Baltimore County, MD; Chicago, IL; Forsyth County (Winston-Salem), NC; Los Angeles County, CA (centered in the San Gabriel Valley); Manhattan and the Bronx, NY; and St. Paul, MN. All MESA participants were eligible for inclusion in MESA Air along with new participants recruited from three additional areas: coastal Los Angeles County, CA; Riverside County, CA; and Rockland County, NY. These additional participants were recruited to enhance exposure gradients among the cohort. Since the new recruits in coastal Los Angeles were within close proximity to the other Los Angeles MESA participants (generally less than 15 km), these two areas were considered as one for our exposure monitoring efforts. This resulted in a total of eight areas of interest for our exposure characterization campaigns. Further details about subject recruitment and demographics of the MESA Air cohort can be found in Kaufman et al. (in preparation).

**General Approach to Exposure Assessment**
Exposure assessment within MESA Air is a multi-step process, illustrated in Figure 1. To characterize the long-term concentrations of ambient pollution, extensive pollution measurements were made at the homes and in the communities of a subset of participants in each of the MESA Air areas. These data, along with local geographic, meteorological, and emission information, inform a hierarchical spatiotemporal model that predicts long-term average concentrations outside of each participant’s home [20]. Similarly, measurements of indoor and outdoor particulate sulfur are integrated with housing characteristic data collected from each participant to predict home-specific infiltration efficiencies ($F_{inf}$) and personalized estimates of indoor concentrations of ambient-generated PM$_{2.5}$ [22]. The ambient concentration to which individual $i$ within area $k$ during time period $v$ is exposed ($A^A_{kiv}$), is a function of fraction of time spent outdoors ($f^o$), the particle infiltration efficiency ($F_{inf}$), and the ambient concentration ($C^A_{kiv}$):

$$A^A_{kiv} = [f^o + (1 - f^o) F_{inf}] A^A_{kiv}$$

The total personal exposure can be written as:

$$E^P_{kiv} = E^A_{kiv} + E^N_{kiv} = [f^o + (1 - f^o) F_{inf}] C^A_{kiv} + E^N_{kiv}$$

where $E^N$, or exposure to non-ambient sources, is the sum of exposure to indoor-generated PM$_{2.5}$ ($E^A$) and exposure to personal activity PM$_{2.5}$. $E^I$ is exposure to the fraction of the indoor particles generated indoors, expressed as the difference between total indoor concentration $C^I$ and infiltrated concentration:

$$E^I = [1-f^o] [C^I - (F_{inf} x C^A)]$$

2/13/09  5  Revision 6.0
Exposure to personal activity PM$_{2.5}$ is the difference between the measured $E^p$ and the sum of calculated $E^A$ and $E^d$. [23]

Further detail pertaining to the modeling of the long-term average ambient concentration at the subject’s home address ($C_{kiv}^d$) can be found in Szpiro et al [20].

Information regarding the data collection, measured parameters, analytic methods, and quality control procedures are presented below.

**Pollutants of Interest**

The pollutant of primary interest in MESA Air is PM$_{2.5}$. Due to the influence of traffic-generated pollution on within-city spatial gradients and the potential importance of those gradients on health [24], we also measured traffic pollution indicators including light absorbing carbon (LAC), total oxides of nitrogen (NO$_x$), and nitrogen dioxide (NO$_2$). In addition, we characterized the composition of PM$_{2.5}$ by analyzing for elemental composition, organic (OC) and elemental carbon (EC), and specific organic compounds. We also measured other gaseous pollutants, ozone (O$_3$) and sulfur dioxide (SO$_2$), in select situations.

**Data Sources and Collection Methods**

*Community-Based Monitoring*

Environmental Protection Agency (EPA) Air Quality System (AQS) Stations

For regulatory compliance purposes, the EPA contracts with local agencies to maintain multiple monitoring stations in and around all of the MESA Air areas except Rockland County, NY. These stations include at least one site per area where speciated PM$_{2.5}$ data are collected. Data from the EPA monitoring stations are collected on various time scales, including continuous measurements which are reported hourly as well as 24-
hr average concentrations measured every one, three, or six days. These data provide
temporally rich but spatially limited information regarding ambient pollutants.

**MESA Air Fixed Sites**

In order to expand our characterization of temporal variability to more spatial
locations, we operated between 1 and 5 fixed monitoring stations in each study area.

These “fixed sites” remained at set locations and collected two-week integrated samples
of PM$_{2.5}$, LAC, NO$_x$, NO$_2$, EC, OC, elemental and organic composition of PM$_{2.5}$, and
SO$_2$ (Table 1) over numerous years while MESA Air participants were being observed
for sub-clinical and clinical disease. The locations for these sites included libraries,
schools, or other buildings that were in participant-dense areas underrepresented by the
AQS network. One monitor was also situated within 100 m of an interstate or state
highway in each area to identify and characterize any differences in temporal trends near
major roadways, where AQS PM monitors are generally not located. One fixed site per
area was also co-located with a local EPA speciation site (except in Rockland, NY) to
calibrate our measurements with the Federal Reference Method (FRM) measurements.

**Community Saturation Monitoring**

In each area, we collected approximately 100 simultaneous two-week samples of
NO$_2$, NO$_x$, and SO$_2$ (Table 1) to create a spatially rich dataset that would identify
important geographic predictors of within-city variability. Since traffic is a major source
of small-scale spatial variation, the saturation samples focused on roadway concentration
gradients but also captured other local sources.

Saturation sampling occurred during three sampling periods (December through
February; May through August; and in October, November, March, or April) to capture
seasonal differences in spatial concentration patterns. All samples were collected using passive Ogawa samplers attached to utility poles approximately 3 m above ground. Based on our preliminary spatial models and past investigations [25-27], we chose a factorial sampling design to sample near and far from major roadways and in areas of high and low population density. To maximize variability while minimizing technician driving time, we selected a roadway gradient design for most locations with six samples collected along a trajectory perpendicular to a major roadway (i.e., interstate or state highways or major arterials) as defined by the U.S. Census Bureau’s Census Feature Class Code (CFCC) A1, A2, or A3. In each direction, one sampler was situated between 0 and 50, 50 and 100, and 100-350 meters, respectively, from the major roadway’s edge. Care was taken to avoid intersections with a large fraction of “stop and go” traffic and samplers were deployed both in North-South and in East-West orientations. To ensure adequate capture of non-roadway conditions, gradient locations were sited in areas of low, medium, and high population density (defined by tertiles of the general population in our region of interest) and in areas of varied land-use. We also collected 10% of all samples at random locations and added new locations based on important predictors identified during the first round of sampling.

**Home and Participant Monitoring**

**MESA Air Questionnaire**

To extrapolate our measured data to every participant, all subjects completed the *MESA Air Questionnaire* upon entry into the study. The questionnaire included questions about residential predictors of $F_{\text{inf}}$ including building characteristics, heating and air conditioning, and window opening as well as information on time-location patterns such
as secondary residences (e.g. winter homes), employment or volunteer positions, time spent indoors and outdoors, and time spent in traffic. All questionnaires and diaries used in this study were written in English and translated into Spanish and Chinese.

Questionnaires are re-administered during follow-up phone calls approximately every nine months if a participant moves or has major life changes that might affect their time-location information (such as retirement or the death of a spouse). Questionnaires will also be re-administered to all participants during the follow-up clinical exam in 2010-2012 when participants will be examined for the progression of subclinical atherosclerosis.

**Home and Personal Pollution Measurements**

Between 25 and 95 homes per area were selected for home outdoor monitoring to add broader spatial coverage to the existing AQS monitoring network. In addition to selecting homes for good geographic coverage of our areas (implemented by forcing data in each of 4 zones per area), we also aimed to capture more extreme within-city variation in concentrations than the AQS. Therefore, we over-sampled homes “near” (<50 meters) major roads (25% of samples) and homes “far” (> 300 meters) from roads (25% of samples).

Outdoor sampling was typically conducted in a participant’s backyard, away from all structures. When it was not possible to sample in a backyard, samplers were placed approximately 1 meter out a window that was sealed with weather-stripping. All outdoor monitoring at participant homes included analysis for PM$_{2.5}$, LAC, NO$_x$, NO$_2$, and SO$_2$ and a subset included EC, OC, elemental and organic composition (Table 1). Homes
were visited on a rolling basis with as many as 12 homes in each area monitored concurrently during any given two-week period.

A portion of the homes (between 25 and 80 per area) selected for outdoor monitoring also received indoor pollution measurements. This sampling allowed us to obtain information to characterize the $F_{inf}$ rates of outdoor pollution into different non-smoking homes in each area. Indoor samplers were set up in the participant’s main activity room away from potential pollutant sources or ventilation systems. These samplers collected the same pollutants as were collected outdoors (Table 1). The ratio of indoor elemental sulfur concentration to outdoor elemental sulfur concentration was used to estimate the infiltration rate of PM$_{2.5}$ [22] [11] and O$_3$ was measured to characterize the $F_{inf}$ of reactive gases.

At all homes with indoor measurements, technicians administered an Infiltration Questionnaire to the participants. This questionnaire characterized the predictors of $F_{inf}$ also assessed on the MESA Air Questionnaire, but during the nominal two-week period when paired indoor-outdoor monitoring occurred. Questions were also asked about indoor sources of sulfur that might bias our approach to estimate $F_{inf}$. These data inform city-specific models of $F_{inf}$ that will be used to predict $F_{inf}$ for all MESA Air homes.

A subset (between 4 and 16 per area) of participants whose homes underwent indoor-outdoor monitoring was also selected for personal exposure monitoring. Personal monitoring was conducted to better understand potential sources of measurement error in our study, and assess the modeled estimates of ambient-generated PM$_{2.5}$ and traffic-related gases. Participants carried active samplers to collect PM$_{2.5}$ mass and elements, including sulfur, and passive samplers to measure NO$_2$, NOx, and SO$_2$. By including
sulfur, the ratio of personal sulfur exposure to ambient sulfur concentration can be used as an estimate of the ratio of ambient component of personal PM$_{2.5}$ to the ambient PM$_{2.5}$ concentration. This latter ratio is the “attenuation factor” [28], or the “ambient exposure factor”, $\alpha$ [29], and is equal to $f^{\alpha} + (1-f^{\alpha})F_{\text{inf.}}$, described previously.

Due to the higher level of alert dictated by the US Department of Homeland Security, sampling among participants who rode public transportation in New York City was limited to passive measurements only. All samples were collected during the same 2-week periods that indoor and outdoor samples were collected at the participant’s home (see Table 1). Only participants living in non-smoking households participated in personal monitoring.

In addition to the MESA Air Questionnaire and the Infiltration Questionnaire, personal monitoring participants completed a Time-Location Diary (TLD) to understand the impact of activities and locations (e.g. commuting) on personal exposures to ambient-generated pollution and to provide a comparison (non-contemporaneous) with the time-activity information in the MESA Air Questionnaire. On the TLD, participants recorded their hourly presence in seven microenvironments: home indoor or outdoor, work indoor or outdoor, motor vehicle, or other indoor or outdoor. In addition, participants indicated if they conducted or experienced any of the following: exposure to environmental tobacco smoke (active or passive), cooking, dusting or vacuuming, close proximity to a fire, proximity to a major road while not in a vehicle or building, heavy traffic while driving, or other sources of smoke.
All home and participant-based sampling was designed to be conducted twice per location/person with repeat samples targeted to capture two of three distinct seasons: summer, winter, or a transitional season (spring or fall).

**Measurement Methods**

PM$_{2.5}$ mass concentrations were measured gravimetrically; the Teflon filters used to determine PM$_{2.5}$ mass concentration were pre- and post-weighed at the UW in a temperature and humidity controlled environment [30] using standard filter weighing procedures [31]. The Teflon filters were also used to determine the amount of LAC, a surrogate for EC, via reflectance measurements. The relationship between LAC and EC is developed empirically in each area. Following post-sample gravimetric analysis and reflectometry, the filters from homes selected for paired indoor-outdoor or personal monitoring were analyzed for 48 elements by X-Ray Fluorescence (XRF). EC and OC were also determined using pre-fired quartz fiber filters at these locations.

Ogawa passive samplers were used to measure NO$_2$, NO$_x$, SO$_2$, and O$_3$. Ion chromatography was used to analyze the sample extracts for nitrite, nitrate, and sulfate for the quantification of NO$_2$, O$_3$, and SO$_2$, respectively. Ultraviolet spectroscopy was used for analysis of NO$_x$, and the mass of NO$_2$ was subtracted from the mass of NO$_x$ to estimate the net mass of NO. Ambient concentrations of each pollutant were calculated using the equations provided by Ogawa & Co.[32]

A more in-depth description of the measurement methods can be found in the online supporting information.

**Geographic Data**
Geographic parameters such as proximity to roadway, population density, land use, and elevation were calculated or assigned to each participant home and monitoring location using ArcGIS 9.2 (ESRI Corporation, Redlands, CA). Meteorological dispersion models (e.g., CalRoads) were also implemented in order to better predict pollutant gradients around roadways. These geographic data and dispersion modeling results will be used to inform our spatiotemporal concentration prediction models [20].

All participant home locations were geocoded using Dynamap 2000 TeleAtlas (Menlo Park, CA). Only those homes with a minimum match score of 80% were geocoded in an automated fashion. All others were geocoded interactively with a minimum match score of 90 to ensure high accuracy. In addition, at all monitored homes our technicians collected curb-side GPS measurements using a handheld Etrex global positioning system (GPS) (Garmin International Inc, Olathe, KS) at the centerline of the property and at selected sampling locations to evaluate spatial accuracy and reproducibility.

Geographic data originated from a wide variety of sources. For example, block group information and population density data were obtained from the US Bureau of the Census (http://www.census.gov/) while land use was obtained from the United States Geological Survey (http://edc.usgs.gov/geodata/). Pollution emission data were characterized using the Toxic Release Inventory database (http://www.epa.gov/tri/) and traffic volumes were characterized by overlaying traffic demand models developed between 2000 and 2006 by regional traffic authorities on the Dynamap 2000 TeleAtlas road network.

**Quality Control/Quality Assurance**
Both field blanks and co-located duplicate samples were deployed at a rate of 10% at the fixed and home sites for all sampling media. Blanks and duplicate samples were deployed during community saturation sampling at rates of 10% and 15%, respectively. To reduce participant burden, no duplicate personal PM$_{2.5}$ samples were collected. Duplicate Ogawa samples were collected in personal sampling at a rate of 15%. As a result of artifact issues associated with quartz filter sampling [33], we used both standard blanks (10% of samples) and dynamic blanks (20% of samples); the latter are back up filters placed downstream of the sample filter to collect any off-gassed organic carbon compounds. A more in-depth description of the measurement methods can be found in the on-line supplement.

**Preliminary Results**

Data Summary and Design Goals

Air pollution monitoring began in July, 2005 and is planned to continue through August, 2009. Sampling goals for the home, personal, and community saturation sampling were met by August, 2008, with 620 unique homes and 793 unique community locations monitored. Of the 620 homes with outdoor air samples, 443 were also sampled for indoor air, and 88 of those homes with paired indoor-outdoor samplers had additional personal monitoring. Fixed site locations, which will continue operating through August of 2009, totaled 27 unique locations. Counts of the sampling locations by area are presented in the on-line supporting information. Overall, we were generally successful in meeting our design goals for siting of our samples with good variation in proximity to roadways, population density, and land use. Details are presented in the on-line supporting information.
Quality assurance demonstrated high data quality with all pre-defined data quality objectives (DQOs) met. Precision, as assessed by the relative percent difference of duplicate samples divided by the square-root of 2, was less than 10% for all pollutants. Comparisons of co-located MESA Air and EPA AQS stations also showed good overall correlation between methods, with $R^2$ ranging between 0.56 and 0.93, with some scatter explainable by the small number of days included in the 2-week averages from the AQS sites. Further detail is presented in the on-line supporting information.

**Preliminary Findings**

Figure 2 presents preliminary monitoring results for PM$_{2.5}$ from outdoor home samples collected in each of the MESA Air areas. In order to reduce the impact of temporal variations on the 2-week averages, we have normalized these concentrations by dividing by the area-average concentration during the 2-week time period and multiplying by the long-term average (2005-2008) concentration for that area. Area-wide short- and long-term averages were calculated using AQS data.

These data demonstrate that there is intra- and inter-urban spatial variation in concentrations. Riverside and Los Angeles communities consistently have the highest PM$_{2.5}$ concentrations while concentrations were consistently lowest in Rockland and St-Paul, resulting in a between-area range of approximately 15 µg/m$^3$. Within-area variability of PM$_{2.5}$ was generally on the order of 5 and 10 µg/m$^3$ (25-75$^{th}$ percentile differences) with smaller variation documented in Baltimore and Winston-Salem and larger variation found in Riverside and Los Angeles. Szpiro et al. have made more in-depth investigation into the structure of the spatio-temporal variability of the data.
Substantial variation in $F_{inf}$ was also observed between and within-areas. Figure 3 presents preliminary indoor-outdoor particulate sulfur ratios from a subset of homes that were sampled simultaneously for indoor and outdoor air pollution. These data are presented independent of season due to the limited homes with analyzed speciation data at the time of publication. These ratios range on average from 55% in St Paul, MN to approximately 85% in New York and Los Angeles. The indoor-outdoor ratios were also quite variable with the 10th to 90th percentiles differing by as much as 60%, showing the within-city variability.

**Limitations**

Despite its considerable strengths, some limitations of this exposure assessment approach should be noted. First, one emphasis of this project was to characterize residential exposures to ambient-origin pollution. Although information was collected on commuting patterns and work exposures, we do not have comparable monitoring data or infiltration information to accurately quantify concentrations in non-residential scenarios. Additionally, MESA Air was not intended to fully characterize indoor sources of exposure such as cooking, smoking or occupational emissions. The exposure assessment methodology employed assumes that average non-ambient personal exposure neither varies across communities nor with ambient exposure. However, this assumption is consistent with previous research [28] [34], and we will also be able to explore its validity by comparing questionnaire-derived data on non-ambient pollutant sources across communities.

An additional limitation of this project is that our PM$_{2.5}$ samples may not capture volatile or semi-volatile components; therefore our measured concentrations may in some
cases be underestimates [35, 36]. This may be especially problematic in and around Los Angeles due to the high concentrations of particulate nitrate in that region [36].

Nevertheless, the FRM also underestimates the concentration of nitrate and other semi-volatile species, so our PM$_{2.5}$ measurements are expected to be comparable to those measured using the FRM [36]. This issue also has implications for our estimation of $F_{\text{inf}}$ since the $F_{\text{inf}}$ of sulfur accurately represents $F_{\text{inf}}$ of non-volatile PM$_{2.5}$ (assuming that there are no indoor-generated sources of sulfur) but may overestimate the $F_{\text{inf}}$ of total PM$_{2.5}$ in regions with relatively high nitrate concentrations [37].

**Discussion**

Links between ambient fine particulate air pollution and adverse cardiovascular health effects are increasingly well established[1], though epidemiological studies to date have used very simple exposure estimation methods. MESA Air was funded by EPA to reduce uncertainty in the relationship between PM$_{2.5}$ and cardiovascular disease in part through improved exposure assessment. Unlike many past chronic epidemiology studies that have been limited to assigning one exposure estimate to all participants within a city, MESA Air aims to predict unique exposure estimates for ambient-generated PM$_{2.5}$ for every study participant using a statistical modeling-based approach. This approach characterizes and incorporates spatial and temporal variations in ambient concentrations, differences in residential $F_{\text{inf}}$, and the general movement of study participants between various indoor and outdoor microenvironments. To our knowledge, this is the first epidemiological study of long-term air pollution effects to directly incorporate many of these potentially important sources of exposure heterogeneity into an exposure assessment. Based on past research by Ebelt and colleagues, [12] we anticipate that these
improvements in our exposures assessment will result in less biased and more precise
effect estimates.

In most urban areas, a major source of within-city variability in air pollution
concentrations is traffic, and evidence is accumulating that these within-city
concentration contrasts may play an important role on cardiovascular mortality and
morbidity [13, 14, 24, 38]. Due to a lack of routine pollution monitoring near roads it has
become common for chronic health studies to use residential proximity to major
roadways as a surrogate for traffic-generated air pollution exposures. Although this
approach is supported by data showing elevated concentrations within approximately
100-300 meters of the roadway’s edge [39], this approach is limited by uncertainties in
the proximity-exposure relationship (due to topography, wind direction, or other factors)
[40]. The extensive data collected in MESA Air should allow us to better characterize
small-scale spatial gradients and determine empirically the relationships between traffic-
generated air pollutants and geographic predictors in each study area.

Another limitation of previous air pollution epidemiology studies is that they have
relied on outdoor monitoring data as a surrogate for personal exposure to ambient
pollution. This approach implicitly assumes that the attenuation of ambient
concentrations by buildings is the same for all participants. However, previous studies
have demonstrated differences in residential $F_{inf}$ between homes (both between and
within cities) and over time within homes [18, 19, 41]. Failure to account for this
exposure heterogeneity may lead to uncertainty and/or bias in health effect estimates
[41]. The ability of MESA Air to account for these differences in such a large cohort of
individuals in multiple locations is another unique feature of this investigation.
In summary, MESA Air aims to characterize important sources of between and within-city heterogeneity in exposure to ambient-generated PM$_{2.5}$ and co-pollutants. Using data collected as detailed in this manuscript, in combination with a statistical modeling-based approach, MESA Air will calculate exposure estimates for each study participant that will incorporate personalized estimates of ambient concentrations outside their home, individual-level $F_{inf}$ estimates for their homes, and weights for the duration of time spent inside their homes. Together, these efforts should provide more accurate estimates of ambient-source exposure than have previously been available on cohorts of this size, with an expected result of improving our ability to assess relationships between chronic exposures to air pollution and cardiovascular disease.

**Acknowledgements**

Although the research described in this manuscript has been funded wholly or in part by the United States Environmental Protection Agency through RD831697 to the University of Washington, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. The work was also partially supported by National Institutes of Health grant ES013195. The authors greatly appreciate the dedication of the air pollution monitoring technicians, the field center and exposure assessment center staff supporting these activities, and the local air quality agencies who have given us access to their sites. We would also like to thank all of the study participants who graciously participated in our air pollution monitoring.

**Supporting Information**
Supporting information is available on-line that presents more details of the measurement methods and quality control/quality assurance procedures and outcomes.
References


32. Ogawa and Company, U., Inc.: Pompano Beach, FL 1998., NO, NO$_{2}$, NO$_{x}$, and SO$_{2}$ sampling protocol using the Ogawa sampler.


### Table 1 Sampling sites and parameters measured

<table>
<thead>
<tr>
<th></th>
<th>Fixed</th>
<th>Community Saturation</th>
<th>Home Outdoor</th>
<th>Home Indoor&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Personal&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal Sample Duration</td>
<td>2 weeks</td>
<td>2 weeks</td>
<td>2 weeks</td>
<td>2 weeks</td>
<td>2-3 days/2 weeks&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Number of Sites Per Area</td>
<td>1-5</td>
<td>44-152</td>
<td>25-80</td>
<td>25-80</td>
<td>4-16</td>
</tr>
<tr>
<td>Number of Sites Per Area Per 2 Week Session</td>
<td>1-5</td>
<td>0 or 100</td>
<td>3-5</td>
<td>2-3</td>
<td>1-2</td>
</tr>
<tr>
<td>Number of Repeat 2 Week Sampling Sessions at Each Sampling Sites</td>
<td>Sequential, ~26/yr</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Questionnaire&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Infiltration Questionnaire</th>
<th>Time Location Diary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pollutant Measurements (Method)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM&lt;sub&gt;2.5&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Harvard PEM - Gravimetric)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light-Absorbing Carbon&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Harvard PEM - Reflectometry)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elements&lt;sup&gt;f&lt;/sup&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(X-ray fluorescence)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC/OC&lt;sup&gt;f&lt;/sup&gt;</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(IMPROVE, Thermal Optical Reflectance)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Organic Compounds&lt;sup&gt;f&lt;/sup&gt;</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(Gas chromatography mass spectroscopy)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO&lt;sub&gt;x&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Ogawa – UV Spectrometry)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Ogawa – Ion Chromatography)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SO&lt;sub&gt;2&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Ogawa – Ion Chromatography)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O&lt;sub&gt;3&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(Ogawa – Ion Chromatography)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>All home indoor measurements are concurrent with home outdoor measurements.

<sup>b</sup>All personal measurements are concurrent with home outdoor and indoor measurements.

<sup>c</sup>PM<sub>2.5</sub> is collected with 2 or 3 day consecutive samples which are then composited over the ~2 week period, NO<sub>x</sub>/NO<sub>2</sub> samples are nominally 2 weeks.

<sup>d</sup>All participants, including those not selected for home monitoring, complete the full MESA Air questionnaire and the follow-up questionnaire.

<sup>e</sup>Measured on PM<sub>2.5</sub> Teflon filters.

<sup>f</sup>Measured on PM<sub>2.5</sub> Quartz fiber filters.
Figure 1. Conceptual Diagram of MESA Air Exposure Assignment. The wiggly arrows indicate multiple sources of input.
Figure 2. Normalized outdoor PM$_{2.5}$ (µg/m$^3$) concentrations at participants’ homes.
Figure 3. Indoor to outdoor particulate sulfur concentration ratios by area. Winston-Salem data not presented because samples had yet to be analyzed.