Project 3 (Framingham):
Identifying the cognitive and vascular effects of air pollution sources and mixtures in the Framingham Offspring and Third Generation Cohorts

Principal Investigator: Murray Mittleman
Co-Principal Investigator: Joel Schwartz
ABSTRACT - Project 3

a. EPA-RC2009-STAR-C1
b. Title: Identifying the Cognitive and Vascular Effects of Air Pollution Sources and Mixtures in the Framingham Offspring and Third Generation Cohorts.
c. Investigators: PI: Murray Mittleman, Co-PI: Joel Schwartz
d. Institution: Harvard School of Public Health
e. Project Period and Location: 2010 - 2015, Boston, MA
f. Project Cost: $1 Million
g. Project Summary:

1) Objectives and Hypotheses: Long- and short-term exposures to ambient air pollution are associated with adverse acute and chronic cardiovascular and perhaps cognitive function, but these effects are poorly understood. Using data from the Framingham Offspring and Third Generation Cohorts, well-characterized populations that have not been previously investigated in association with ambient environmental exposures, we will: (1) determine whether long-term exposures to ambient pollutants and mixtures are associated with cognitive impairment and cognitive interference; (2) test whether short-term and long-term exposures to pollutants, mixtures and sources are associated with acute and chronic vascular and endothelial function; and (3) consider whether markers of biological susceptibility and vulnerability differentially influence these associations, allowing us to identify subpopulations at increased risk for harmful effects of air pollution.

2) Experimental Approach: Exposures will be assigned using data from the Harvard School of Public Health Boston Supersite, a network of New England regional sites, rotating monitors, and satellite aerosol optical depth data. This information will be used to obtain predictors for linear regressions with covariate-adjusted models for cognitive outcomes (MMSE, CERAD Word List Memory and Victoria Stroop test) as well as vascular measures (blood pressure, brachial artery diameter, flow mediated dilation, digital pulse amplitude). Cross-product terms will be used to test susceptibilities and vulnerabilities.

3) Expected Results: We will estimate health risks associated with short- and long-term exposure to individual air pollutants, sources and air pollution mixtures within the Framingham Offspring and Third Generation populations. We will address which individual and area-level factors, measuring vulnerability, susceptibility, and individual air pollutants, sources and mixtures, are the major determinants in explaining spatial and temporal variability of the health risks. Also, we will be able to add to data addressing the effects of pollutants on the life course examined in Project 2: the Normative Aging Study by investigating cognitive performance and vascular function in populations of middle- and older-aged men and women.

Supplemental Keywords: Air pollution, ambient particles, multi-pollutant mixtures, cognitive function, vascular function, inflammation, susceptibility, vulnerability
1. OBJECTIVES/HYPOTHESES

There is mounting evidence that long- and short-term exposures to ambient air pollution are associated with both acute and chronic adverse cardiovascular effects. Recent data suggest that long-term exposure may also adversely affect cognitive function, but these effects are poorly understood. An important potential explanation for heterogeneity of findings is that variation in specific pollutants and mixtures may have differential effects that have not been well-studied. It is also likely that pollutants from different sources have variation in their effects. Furthermore, individuals may differ in either biologic susceptibility to the effects of air pollution, or levels of vulnerability to exposure based on psychosocial characteristics. Our objectives are therefore to determine the chronic effects of exposure to air pollutants and mixtures from specific sources on cognitive function, and the acute and chronic effects of such exposures on vascular and endothelial function among middle-aged and older adults enrolled in the Framingham Offspring and Third Generation cohorts.

Our Specific Aims are to test the following 3 hypotheses:

**Hypothesis 1:** Long-term exposures to ambient pollutants and mixtures have adverse chronic effects on **cognitive function** among middle-aged and older adults. Furthermore, different individual air pollutants, sources, and pollutant mixtures vary in their effects, after accounting for age, on:

- Cognitive decline as measured by Mini-Mental State Exam (MMSE) and the CERAD Word List Memory test
- Cognitive interference as measured by Victoria Stroop test

**Hypothesis 2:** Both short-term and long-term exposures to pollutants, mixtures and specific sources have adverse acute and chronic effects on **vascular and endothelial function** among middle aged and older adults as measured by:

- Decreased brachial artery diameter and flow mediated dilatation (FMD)
- Decreased digital pulse amplitude response to hyperemia (PAT)
- Increased systolic, diastolic and pulse pressure

**Hypothesis 3:** The adverse effects of short-term and long-term exposure to ambient individual pollutants, sources, and pollutant mixtures on **cognitive, vascular and endothelial function** as described in Hypotheses 1 and 2 are differentially impacted by measures of **risk modifiers associated with susceptibility** and **vulnerability.**
1.2. Background and Introductory Information

1.2.1 Potential Mechanisms of Neurotoxic Effects of Exposure to Particulate Air Pollution

Emerging evidence suggests that particulate air pollution induces harmful effects on neurocognitive function. Two major hypotheses have been suggested to explain these observed associations. First, recent evidence from animal and human models suggests that ambient air pollutants can adversely affect endothelial function, damage microvasculature, and eventually cause progression of atherosclerosis.\(^1,2\) Given the strong associations of cognitive decline with these clinical conditions and neuropathogenic processes, long-term exposures to ambient air pollution may affect neurocognitive function through such processes. The second hypothesis is that air pollution induces cognitive effects more directly. Studies in animal models have found that particles can translocate from the nose via the olfactory nerve into the brain with exposures shown not only to the olfactory bulb, but also to the striatum, frontal cortex, and cerebellum.\(^3,4\)

In human volunteers, exposure to diesel exhaust led to changes in electroencephalographic patterns indicative of cortical stress.\(^5\) Calderon-Garciduenas and coworkers have compared the health effects of air pollution in Mexico City and a control, much less-polluted Mexican city. MRI evaluation of the brains of children living in more polluted locations revealed greater prefrontal lesions, and highly-exposed children and young adults showed upregulated inflammatory markers (cyclooxygenase-2, IL1-\(\beta\), CD14). Also, dogs from Mexico City had greater rates of prefrontal lesions, neuroinflammation, gliosis, and particle deposition.\(^6\)

Despite such evidence, research on health effects of air pollution mixtures has produced different results dependent on the individual air pollutants tested and socio-demographic factors present.\(^7\) Therefore, our studies are designed to identify differential adverse effects of individual pollutants, sources, and pollutant mixtures on neurocognitive function. Furthermore, we will be able to identify sub-populations at greater risk of these adverse outcomes based on individual markers of susceptibility and vulnerability.

1.2.2 Evidence that Particulate Air Pollution Affects Vascular and Endothelial Function

Numerous health studies have shown acute\(^8,9\) and chronic\(^10\) particulate air pollution exposure to be associated with early death, particularly from cardiovascular disease (CVD). A classic marker of cardiovascular health is elevated blood pressure (BP), which is directly associated with cardiovascular events. Air pollution has been found positively associated with markers of blood pressure (systolic, diastolic and pulse pressure) within the Boston metropolitan area\(^11,12\) and elsewhere,\(^13-16\) but results have been inconsistent.\(^17,18\) Recently, we observed that exposure to Concentrated Ambient Particles (CAPs) and O\(_3\) increased diastolic BP by 2.0 mmHg\(^19\) but no other statistically significant blood pressure changes were observed. The contribution of components within mixtures from multiple sources in affecting blood pressure is poorly understood.

Newer cardiovascular measures may provide more comprehensive information about cardiovascular function. For example, brachial flow-mediated dilation (FMD) is a noninvasive physiological parameter that can measure endothelial function and has been validated as a predictor of cardiovascular events in older adults.\(^20\) Literature on the effects of ambient pollutants on FMD is limited and a few small controlled experiments have examined FMD association with air pollution, with mixed results. We conducted a study in the Boston area and observed changes in flow-mediated and nitroglycerine mediated vascular reactivity among diabetics.\(^21\) However, in young healthy adults who inhaled 150 \(\mu g/m^3\) of concentrated ambient fine particles and ozone (120 ppb) for 2 hours, decreased basal brachial artery diameter in
comparison with control inhalation of filtered air, without affecting FMD or responses to nitroglycerine. Discrepant findings may have arisen because of differing sources, composition and mixtures of pollutants. Varying susceptibilities of study subjects and differences in time windows studied (a 24-hour average to a moving average of 6 days) may have also contributed to different outcomes.

While FMD assesses a large conduit artery, it does not provide information regarding microvascular circulation, which may be an important dimension of chronic ischemic heart disease. Digital peripheral arterial tonometry (PAT) is a noninvasive measure of peripheral vascular function. Augmentation of pulse amplitude in the finger with hyperemia is a response to ischemia reflecting both change in flow and digital microvessel dilation. In prior clinical studies, impairment of pulse amplitude hyperemic response was associated with the presence of coronary artery endothelial dysfunction. Brauner et al. reported that filtering re-circulated indoor air for 48 hours improved microvascular function as measured by PAT in healthy elderly using a double-blind crossover design. A limitation of this study is that there was that only indoor air quality was examined and associations observed may be due to indoor sources such as cooking or candle burning rather than vehicle emissions. Our work within the Framingham Offspring and Third Generation participants would be the first large population-based study to investigate both macro- and microvascular effects of pollution with detailed assessment of composition, mixtures and sources.

1.2.3 Susceptibility and Vulnerability Factors in Health Effects of Particulate Air Pollution

In addition to studying direct effects of pollutants, sources and mixtures on vascular and neurocognitive function, we plan to address factors of susceptibility which may differentially influence these outcomes.

Aging and Air Pollution: Elderly individuals are at increased risk for CVD, and those with CVD may be particularly susceptible to effects of air pollution. Higher risks of particulate matter for deaths in persons aged >65 yrs were found in Philadelphia, and healthy elderly subjects exposed to moderate pollution for 2 hr while at rest exhibited altered heart rate variability (HRV). Older individuals may have experienced larger cumulative doses of particulate air pollution than inferred from current levels, because pollution levels have decreased dramatically in recent years, and they also have age-related weakening of defense mechanisms. Whether susceptibility to air pollution at older age varies by pollutant types, sources and mixtures has been inadequately explored, but may have important regulatory implications.

Obesity and Diabetes: There is evidence that those with obesity or diabetes are more susceptible to the effects of air pollution, putting an increasing number of people at risk. In research funded by our current Center we assessed obesity and diabetes as susceptibility markers Progress Report (NAS Project 2, Section 1.3.7). We found that diabetics had double the risk of a PM$_{10}$-associated cardiovascular admission compared with non-diabetics, in a study of four US cities. We also observed that diabetic subjects had a 2.0-fold higher mortality risk associated with PM$_{10}$ exposure, in a case-crossover study. A study of 9 Italian cities likewise found stronger effects of PM$_{10}$ on mortality in diabetics than non-diabetics. A study of the effect of PM$_{2.5}$ on Heart Rate Variability (HRV) reductions found that obese individuals had twice the standard deviation of the normal-to-normal interval (SDNN) reduction of non-obese individuals and had more PM$_{2.5}$ mediated HR increases, supporting the hypothesis that obesity is a susceptibility factor for the acute cardiovascular effects of fine particles.
Gender Differences in Responses to Particulate Air Pollution: Women and men have different risk factors for CVD and aging and it is also possible that they respond to air pollution differently. Women were found particularly susceptible to ambient ozone-related death, an effect observed not at high exposure levels, but rather in cities with lower ozone concentrations. In addition to biological differences between men and women, it is also possible that differences in exposure precision influence observed and expected results.

Socioeconomic Position: Air pollution and environmental risks remain unevenly distributed. A study in Southern Ontario found that mortality rates varied by neighborhood of residence in a cohort of people whose lung function was tested, and that income and air pollution levels were important correlates of mortality. Furthermore, a recent study found that a high level of perceived neighborhood problems (traffic, noise, trash, odors, and fire) was associated with poorer quality of life among asthmatics. When asthma severity and socio-demographics were taken into account, people reporting scores of ≥8 on a 0-25 scale for serious problems (top quartile of seriousness) in their neighborhoods had significantly poorer Quality of Life (QOL) scores (mean difference=5.91; SE=1.63), poorer physical functioning (mean difference = −3.04; SE=1.27), and almost a 5-fold increase in depressive symptoms (odds ratio=4.79; 95% CI=2.41, 9.52). Despite numerous theories and methods suggested for studying these effects, the gaps in knowledge persist. Better understanding of social factors that influence health effects of air pollution is essential to develop targeted approaches to mitigate exposures among individuals who are most exposed and at risk.

1.3 Preliminary Results
Below are Framingham Offspring and Third Generation characteristics and results from previous investigations in Framingham populations relevant to our study approach.

1.3.1 Framingham Offspring and Third Generation Recruitment
The Offspring Study was initiated in 1971. A sample of 5,124 men and women, consisting of the offspring of the Original Cohort and their spouses, was recruited. Offspring Exam 8 began in March 2005. Beginning in 2002, 4,095 Third Generation participants, who had at least 1 parent in the Offspring Cohort, were enrolled. The second exam cycle for the Third Generation study began in 2007.

Table 1 Age-Sex Distributions at Entry for Offspring and Third Generation Participants

| Offspring: AGE-SEX DISTRIBUTION AT ENTRY (1971) |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Age              | <10              | 10-19            | 20-29            | 30-39            | 40-49            | 50-59            | 60-70            | Totals           |
| Men              | 0                | 126              | 543              | 789              | 694              | 293              | 38               | 2,483            |
| Women            | 6                | 113              | 692              | 835              | 740              | 242              | 13               | 2,641            |
| Totals           | 6                | 239              | 1,235            | 1,624            | 1,434            | 535              | 51               | 5,124            |

| Third Generation: AGE-SEX DISTRIBUTION AT ENTRY (2005) |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Age              | 19               | 20-29            | 30-39            | 40-49            | 50-59            | 60-69            | 70-79            | Totals           |
| Men              | 4                | 220              | 656              | 737              | 276              | 19               | 1                | 1,913            |
| Women            | 3                | 262              | 759              | 848              | 293              | 16               | 1                | 2,182            |
| Totals           | 7                | 482              | 1,415            | 1,585            | 569              | 35               | 2                | 4,095            |

Source: FHS website, [http://www.framinghamheartstudy.org/participants/](http://www.framinghamheartstudy.org/participants/)
1.3.2 Flow Mediated Dilation Results from Framingham Offspring
Benjamin et al. examined the clinical correlates of endothelial function cross-sectionally as measured by Flow Mediated Dilation (FMD) within the Framingham Offspring Study population at the Seventh Examination Cycle (1998-2001). Mean FMD% was 3.3±3.0% in women and 2.4±2.4% in men. Stepwise multivariable predictors of increasing FMD% were female gender, heart rate, and having the walk test first. Predictors of decreasing FMD% were advancing age, SBP, BMI, lipid-lowering medication, and smoking within the past 6 hours.38

1.3.3 Peripheral Arterial Tonometry Results from Framingham Third Generation
Hamburg et al. examined the relationship between digital vascular function to cardiovascular risk factors among Framingham Third Generation participants during the first examination cycle (2002-2005).39 Using a fingertip peripheral arterial tonometry (PAT) device digital pulse amplitude was measured in Framingham Third Generation Cohort participants (n=1,957; mean age, 40 +/- 9 years; 49% women) at baseline and in 30-second intervals for 4 minutes during reactive hyperemia induced by 5-minute forearm cuff occlusion. Augmentation of pulse amplitude in the finger associated with hyperemia is a response to restriction in the blood supply reflective of changes in microvessel dilation. Among Framingham Third Generation participants, digital pulse amplitude hyperemic response was higher in women than in men and with advancing age and was inversely related to multiple risk factors, particularly diabetes mellitus, body mass index, higher cholesterol concentrations, and smoking. In stepwise linear models male sex, body mass index, ratio of total to high-density lipoprotein cholesterol, diabetes mellitus, smoking and lipid-lowering treatment were inversely related to PAT ratio, whereas increasing age was positively related to PAT ratio.39

1.3.4 Vascular Effects of Air Pollution Within the Study Region
Our group has previously demonstrated associations between blood pressure and air pollution in cardiac rehabilitation patients.11 In a panel study of diabetic patients, we have shown that particulate matter was associated with acute (6 to 24-hour moving average) increases in arterial blood pressure. Strongest associations were seen for the organic carbon fraction and diastolic blood pressure. A 1.9 μg/m³ increase (inter-quartile range) in the 12-hour mean of organic carbon was associated with a 1.7 mmHg (95% CI 0.6 to 2.7 mmHg) increase in diastolic blood pressure and a 2.1 mmHg (95% CI 0 to 4.3 mmHg) increase in systolic blood pressure (Hoffmann et al, unpublished data). In the same population, we also measured baseline brachial artery diameter and flow-mediated dilation and dilation in response to nitroglycerine administration. Preliminary analyses indicate brachial artery constriction at baseline in response to higher levels of PM2.5, OC, and SO4 with maximal effect at 2 to 4-day moving averages. The figure shows that for each pollutant, an IQR increment in pollutant level was associated with an approximate 0.08 mm smaller brachial artery diameter.

Figure 1: Brachial Artery Diameter Change per IQR Pollutant.
1.4 Project 3 Integration within Center Framework
In this project we will evaluate the impact of pollutants, mixtures and sources on chronic neurocognitive effects and acute and chronic vascular and endothelial function in middle age and older adults. The study team will work closely with members of the Exposure Core to ensure that each exposure metric is optimally assigned and will also interact with members of the Biostatistics Core in the planning and execution of data analyses. Project 3 will be informed by the toxicology studies of Project 1, the studies of early life vascular changes in Project 4 (VIVA) and the studies of oxidative stress, vascular and neurocognitive function in Project 2 (Normative Aging Study). This project will inform the population-based studies of Project 5 (National Study). Our integrated program will provide a comprehensive approach to evaluating health effects of individual pollutants, sources, and pollutant mixtures across life stages.

2. APPROACH/ACTIVITIES

2.1. Framingham Populations
To determine the neurocognitive and vascular effects of particulate air pollution we will study participants from the Offspring and Third Generation cohorts of the Framingham Heart Study. The Framingham Study began in 1948, and consisted of 5,209 men and women in the Original Cohort. In 1971, 5,124 adult children (and their spouses) of the Original Cohort enrolled into the Offspring Cohort and were examined every 4-8 years. Third Generation enrollment is described in Section 1.3.1. Data have been collected through Exam 2.

2.2 Health Data
All participants in both the Framingham Offspring and Third Generation cohorts underwent routine physical examination, anthropometry, and risk factor assessment. Fasting morning plasma glucose, total and HDL cholesterol, triglycerides, waist circumference, weight and height were measured. Diabetes was defined as fasting plasma glucose $\geq 6.99$ mmol/L (126 mg/dL) or treatment with insulin or hypoglycemic agents. Participants were considered current smokers if they smoked at least one cigarette per day for the year prior to examination. Moderate to heavy alcohol use was assessed through physician-prompted questions and were defined as consumption of $>14$ drinks per week in men, or $>7$ drinks per week in women. Among Offspring, HR and BP were measured by automated device (Dinamap, Critikon, Inc.). Among Third Generation participants, HR was recorded during baseline flow velocity measurement. Women were considered to be menopausal if their menses had stopped for $\geq 1$ year. CVD was defined as recognized myocardial infarction, coronary insufficiency, angina pectoris, stroke, transient ischemic attack, or intermittent claudication.

2.3 Cognitive Outcomes
For these analyses we will examine cognitive impairment by utilizing the Mini-Mental State exam in the Framingham Offspring and the Consortium to Establish a Registry for Alzheimer’s disease (CERAD) Word Recall test.

2.3.1 Mini-Mental State Exam: The Mini-Mental State Exam (MMSE) has been collected at each visit since Cycle 5 (1991–1995) for Framingham Offspring participants. THE MMSE is a brief test that assesses cognitive abilities in several domains, including orientation to place and time, memory, attention, language, and ability to copy a design. It is used primarily as a screening instrument for dementia and is widely used in epidemiologic studies.
2.3.2 CERAD Word List Memory: The CERAD Word List Memory test was conducted at Exam 2 for Framingham Third Generation participants. This test assesses mild cognitive impairment, and is part of the CERAD consortium battery, developed to evaluate individuals with mild and early-stage cognitive impairment. Its 10-word list is one of the more sensitive tests commonly used, and is scored by recording the number of words recalled in each of the trials. A single cutoff score for the delayed recall trial is used to determine whether cognitive impairment exists.

2.3.3 Victoria Stroop: The Victoria Stroop test was conducted at Exam 2 (2007-present) for Framingham Third Generation participants. It measures verbal inhibitory capacity, and includes 3 tasks with 24 items each. The first contains colored dots and the participant is asked to read the colors out loudly from left to right. The second task includes common words printed in the colors of the dots in trial 1 and the participant is asked to read the colors of the words as quickly as possible. The final task is similar, but involves the names of the printed colors, and the text never matching the color in which it is printed. The participant is asked to disregard the text and read only the color of the ink. The summary score is the number of naming errors on the third task.

2.4 Endothelial Outcomes

2.4.1 Blood Pressure: Blood pressure (BP) has been collected at all visits for both Framingham Offspring and Third Generation participants. In the Offspring cohort, heart rate and BP were measured by automated device (Dinamap, Critikon, Inc.). In Third Generation participants, heart rate was recorded during baseline flow velocity measurement. Left-arm BP was measured to the nearest 2 mmHg with a mercury column sphygmomanometer after the subject was seated quietly for 5 min. Two readings by the physician were averaged to calculate systolic and diastolic BP.

2.4.2 Brachial Artery Measurements: Fasting brachial artery tracings were measured by one of three experienced sonographers following rigorous standardized written protocols and were acquired for Framingham Offspring and Third Generation participants at Exams 7 and 1 respectively. Baseline diameter, flow mediated diameter (FMD)% (percent change in diameter from baseline), baseline mean flow velocity and mean hyperemic flow velocity were determined. Participants were scanned using a Toshiba SSH-140A ultrasound system with a 7.5 MHz linear-array transducer (Offspring) and an Agilent Sonos 1000 system with an 11 L MHz transducer (Third Generation). Brachial studies were measured with commercially available software (Brachial Analyzer v. 3.2.3, Medical Imaging Applications). Investigators measured brachial artery diameter at baseline and 1min after reactive hyperemia induced by 5-min forearm cuff occlusion. Doppler flow was assessed at baseline and during reactive hyperemia with a 3.75 MHz (Offspring) or 3.6-MHz (Third Generation) carrier frequency and with correction for insonation angle. Mean baseline and hyperemic flow velocities were analyzed from digitized audio data with semi-automated signal averaging (Cardiovascular Engineering). Decreased FMD and reactive hyperemia are associated with endothelial dysfunction, because endothelium-dependent vasodilation (by mediators such as nitric oxide) is induced by flow or shear stress.

2.4.3 Peripheral Arterial Tonometry (PAT): Impaired digital pulse amplitude hyperemic response is associated with coronary artery endothelial dysfunction. In Framingham Offspring and Third Generation participants, fasting digital pulse amplitude was measured with a PAT device placed on the tip of each index finger (Endo-PAT2000, Itamar Medical, Israel). The device is a pneumatic plethysmograph that applies uniform pressure to the fingertip surface, allowing measurement of pulse volume changes in the finger. Throughout the study, the inflation pressure of the digital device was electronically set to 10 mmHg below DBP or 70 mmHg, whichever was lower. Baseline pulse amplitude was measured from each fingertip for 2 min 20 sec. Arterial flow
was interrupted for 5 min by forearm cuff (Hokanson model AG101, D.E. Hokanson, WA) at occlusion pressure 200 mmHg or 60 mmHg plus SBP, whichever was higher.

Pulse amplitude was recorded electronically in both fingers and analyzed by a computerized, automated algorithm (Itamar) that provided the average for each 30-sec interval after cuff deflation up to 4 min. Figure 2 compares the hyperemic response of individuals in the first and third tertile of responsiveness. For each 30-sec interval, pulse amplitude response to hyperemia was calculated from the hyperemic fingertip as the ratio of postdeflation pulse amplitude to baseline pulse amplitude ($X_h/X_b$, with $X=$ pulse amplitude, $h=$hyperemic finger, $t=$time interval, 0=baseline). We divide the result by the ratio obtained from the contralateral, control hand ($X_c/X_0$); ($c=$control finger, $t=$time interval, 0=baseline) to obtain the PAT ratio for analysis.

**Figure 2: Pulse Amplitude and PAT Ratio in Hyperemic and Control Fingers**

2.5 Markers of Susceptibility and Vulnerability
The susceptibility markers that we will investigate include gender, aging, obesity and diabetes. Descriptive data on gender and age is available from questionnaires. Health, obesity and diabetes information will be assessed via study visit/exam data (Section 2.2). We will investigate socioeconomic status as a potential marker for vulnerability. Questions about education and income are included in the Framingham Heart Study questionnaires. In addition, we will utilize geo-coded addresses of participants to identify census block group information. To assess depression as a vulnerability marker, we will use the Center for Epidemiologic Studies Depression scale (CES-D), a short self-report scale that measures depressive symptoms and was collected at all study visits. Its items are symptoms associated with depression, used in previously-validated longer scales.

2.6 Environmental Exposures
Here we present a succinct overview of the exposure assessment utilized in our study. The full details of the methodologies for utilizing these sources of data are described in our Exposure and Biostatistics Cores. Depending on the specific hypothesis to be tested, one of six different exposure metrics for epidemiologic analysis will be employed: (a) Short-term Exposures to individual Pollutants (b) Short-term Exposures to individual Sources (c) Short-term Exposures to pollutant Mixtures (d) Long-term Exposures to individual Pollutants, (e) Long-term to Individual Sources, and (f) Long-term exposures to pollutant Mixtures.
2.6.1 Short-Term Exposures and Acute Effects
We will use short-term exposures to examine acute effects on vascular function and on serum biomarkers of inflammation, oxidative stress, and endothelial function. Short-term exposures will be averages of one day, several days, or several weeks before each participant’s study visit, with the length of time depending on the specific health outcome examined. Short-term exposure measurements will include:

(a) Individual Pollutants: Daily concentrations of O₃, PM₂.₅ mass, particle number, elements, ions, BC, EC and OC, VOCs, PM₁₀ and PM₁₀₋₂.₅ will be acquired from the Boston Supersite and the Boston city sites. We will use GIS-based spatio-temporal modeling to estimate 24-hr BC and O₃ exposures at each participant address, and get spatially and temporally resolved estimates of 24-hr PM₂.₅ exposure by calibrating satellite data.

(b) Sources: We will estimate the daily contributions to PM using source apportionment methods. We will estimate contributions from traffic, soil, coal combustion, and sea salt among others, from 1995 onward. These contributions will be estimated using the PMF EPA source apportionment method and for this we will use data from the Boston Supersite only.

(c) Multi-Pollutant Mixtures: We will group days according to their multi-pollutant mixture profiles based on the daily composition of all measured individual pollutant levels, using clustering methods. We will use mixture type as an effect modifier when assessing the effects of 24-hour PM₂.₅ exposure.

2.6.2 Long Term Exposures
For long-term effects, we will use a 1-year exposure window. A follow-up analysis of the Six City Study looked at year-to-year changes in particle concentrations to examine the lag between change in exposure and change in mortality rate.⁴⁵ Schwartz and coworkers, using a penalized spline with up to 18 degrees of freedom (essentially, a polynomial with 18 terms to capture any deviation from linearity), showed that the association was essentially linear down to 8 µg/m³, where the data become sparse, and that the effects of reduced particle exposure on mortality appear to be mostly seen within two years. This conclusion is also supported by natural experiments. Pope and coworkers reported that mortality fell in the Utah valley in the year a strike closed a steel mill, and returned to its previous level the next year, when mill operations resumed.⁴⁶ Another cohort study examined over 66,000 nurses living in the Northeast and upper Midwest.⁴⁷ Using a spatial model that estimated monthly PM₂.₅ concentrations at the addresses of each nurse, they found that a 10 µg/m³ increase in PM₂.₅ at the nurse’s address was associated with a 26% increase in risk of dying in that year. This increase was predominantly seen within a year of a change in exposure.

(a) Individual Pollutants: We will create annual averages for each participant using our spatio-temporal models for 24-hr BC and O₃, and using our spatially and temporally resolved estimates of 24-hr PM₂.₅ exposure. Additionally, using the new 2010-2014 monitoring data, we will build pollutant-specific spatial surfaces for each season during 1995-2009, calibrating the surfaces with the 1995-2009 central site measurements, and validating the estimates with the 20 monitor sites currently in place. From these estimates, we will compute annual averages for the year prior to each participant’s study visit, for all measured pollutants.

(b) Sources: Starting with the year 2000, we will run source apportionment models for all sites, separately by year. Aggregating the source contribution estimates from all years, we will then stratify the data by season and spatial smooth the seasonal source contribution estimates for each
season separately, calibrating the levels of pollution using the central site monitor (see the Biostatistics Core for specification of the proposed modeling framework). We will estimate annual mean source contributions by geo-code.

(c) Multi-Pollutant Mixtures: Using the individual pollutant-specific spatial surfaces described in Section 2.6.2 (a), we will calculate annual averages of each pollutant species for each zip code and cluster the zip codes according to multiple pollutant averages. We will use mixture type for each zip code as an effect modifier for the individual-specific annual PM$_{2.5}$.

2.7. Data Analysis
Detailed statistical methods to be used are provided in the Biostatistics Core. Briefly, our principal outcomes will be continuous measures, with effect modification by numerous outcome-specific parameters examined to test the hypothesis that PM component effects are modified by factors influencing susceptibility to cognitive, endothelial and vascular functions. Covariate selection will be outcome-specific and based on biological factors and other study findings. For all outcomes except when we are assessing the modifying effect of each variable, variables will include age, sex, BMI, season, temperature, and day of the week.

2.7.1 Univariate Analyses
Univariate analyses will be performed for all variables. Expected ranges for all of the variables will be defined a priori and out-of-range values or outlier values will be checked for errors. In addition to data cleaning, this initial analysis will serve the purpose of describing the study characteristics, identifying skewed variables that need transformation.

2.7.2 Linear and Hierarchical Models
Many of the outcomes we will examine are continuous measures. We will use linear models for outcomes with only one measurement per person (Flow-Mediated Dilation). Hierarchical mixed models will be used to account for outcomes with repeated measures on each subject (i.e., blood pressure, MMSE). These regressions combine fixed effects, which are the predictors in traditional regression models, with random effects, such as subject-specific intercepts or slopes, and/or autoregressive errors. Individual covariates will be examined as predictors of outcome, but it is unlikely that they will explain all inter-individual differences. This will necessitate the inclusion of individual intercepts. Specifically, let $Y_{ij}$ be the response in the $i$th subject on day $j$, $age_i$, $temperature_j$, denote for simplicity the set of covariates of interest, including time varying covariates (temperature, took medication today) and covariates that only vary by subject. Then we will consider models of the form:

$$Y_{ij} = \beta_0 + \beta_1 age_i + \beta_2 temperature_j + \ldots + [pollution]_{ij} + b_i + \epsilon_{ij}$$

Here $b_i$ represents a subject specific intercept. To examine effect modification by subject characteristics related to individual susceptibility and vulnerability (such as having diabetes, being obese or depressed) we will use an interaction term to fit separate pollution slopes for each subgroup. Thus, we extend the model above by adding the main effect of the modifier and the pollution x modifier interaction term(s).

2.7.3. Approach to Testing Associations with Pollutants, Mixtures and Sources
To test characteristics of each type of pollutant we will use the following general framework (Section 4.2.1 of Biostatistics Core). The choice of exposure metric above will relate to a specific choice of the pollution term in the model. Briefly, the form of the covariate for each type of scientific hypothesis listed above is:
(a) Short-Term Exposures to Individual Pollutants: For this exposure hypothesis, the term \( [pollution_{ij}] \) represents the daily (or multi-day) concentration of a given pollutant. For the majority of pollutants, we will use measurements taken from our HSPH Boston Supersite. Black carbon, for which we have spatially resolved data, will be modeled with a two-tiered approach, using both daily measures recorded at our Supersite as a first analysis and daily predicted exposure at a given location estimated from an exposure model as a second level of analysis. \( \text{O}_3 \) models will be explored to develop a similar approach. \( \text{PM}_{2.5} \) measurements will utilize AOD converted to mass and regression calibration of daily measurements.

(b) Short-Term Exposures to Individual Sources: As outlined briefly above (Section 2.6) and more comprehensively in the Exposure Core (Section 3.2.1), estimates of source-specific exposures will be constructed using source apportionment methods. Factor analysis will be used to produce daily source contribution for identified sources including crustal sources, coal-fired sources, fuel oil combustion, salts, and traffic pollution.

(c) Short-Term Exposures to Individual Mixtures: Rather than attempting to estimate independent component effects, which in many settings are effectively unobservable due to highly collinear components, we seek to develop methods that characterize profiles of mixtures for which pollution is associated with a health outcome. As outlined in Section 2.3 of the Exposure Core and Section 4.4.2 of the Biostatistics Core, let \( K_j = k, k = 1, ..., C \) denote a variable reflecting cluster membership in one of \( C \) classes for day \( j \) in a study. Again framing it in the linear mixed model framework, we will consider the model:

\[
Y_{ij} = \beta_0 + \gamma_1 c_{j1} + ... + \gamma_{C-1} c_{jC-1} + \beta_{1} c_{j1} pm2.5_j + \beta_{2} c_{j2} pm2.5_j + ... + \beta_{C} c_{jC} pm2.5_j + confounders + b_i + \epsilon_{ij}
\]

where \( c_{jk} \) is an indicator variable equal to 1 if day \( j \) belongs to mixture class \( k \) and 0 otherwise. This model assumes that both the overall level of the health outcome and the association between \( \text{PM}_{2.5} \) mass and the outcome varies among clusters of days defined by their pollution mixture. Interest will focus on testing both whether the \( \text{PM}_{2.5} \) mass associations vary by cluster (i.e., \( H_0 : \beta_1 = \beta_2 = ... = \beta_C \)), and, if so, whether the health outcomes overall vary by mixture (i.e., \( H_0 : \gamma_1 = \gamma_2 = ... = \gamma_{C-1} = 0 \)).

(d) Long-Term Exposures to Individual Pollutants: For hypotheses that concern the health effects associated with long-term (greater than 6 months) exposure to individual pollutants, the “pollution” term in our model represents an estimated address-specific long-term average concentration of a given pollutant. For \( \text{PM}_{2.5}, \text{BC}, \) and \( \text{O}_3 \), we will average daily exposure estimates obtained from the data sources described in (a) above. For all other pollutants, we will apply the long-term spatial models described in the Biostatistics Core (Section 4.4.3) to the long-term monitoring data to be collected during 2010-2014 as part of this Center (see the Exposure Core for a description of this proposed monitoring effort). We propose to use simpler models to characterize the long-term spatial variability for these pollutants (\( \text{PM}_{2.5}, \text{PM}_{10}, \text{PM}_{10-2.5}, \text{BC}, \text{OC}, \text{EC}, \text{elements}, \text{O}_3, \text{NO}_2 \) and \( \text{VOCs} \)) because of the paucity of spatial data and the lack of well-established parameters that can be used to predict spatial patterns. For instance, for \( \text{BC} \) and \( \text{O}_3 \) models, traffic or population density could be included as predictors, whereas this may not be the case for \( \text{PM}_{10-2.5} \) or elements that are associated with a number of sources. The long-term spatial models that will be used to predict geo-coded seasonal average levels for individual pollutants are described in the Biostatistics.
Core. These models will include two components: 1) a spatial component that will use data from the 40 HSPH multi-pollutant spatial sites collected during 2010/11-2014; and 2) a temporal component that will use Supersite data. The developed models will also be employed to predict retrospective exposures for the period of 2001-2009. For these calculations we will use the same spatial structure developed for the period of 2010-2014 and Supersite data to account for yearly pollution trends during 2001-2009. These predictions will be validated using measurements from the spatial networks collected during 2001-2009.

(e) Long-Term Exposures to Sources: We will estimate long-term exposures to sources using a two-stage process. First, we will conduct separate source apportionment analyses for all of the spatial sites (see the Exposure Core for a description of the available ambient monitoring data on PM composition) in the Region using the EPA Positive Matrix Factorization (PMF) method. In order to allow factor loadings to vary over time, this analysis will be performed separately by year. Putting together the resulting daily source contribution estimates from all years, we will then stratify these daily estimates by season. At the second stage, for each source in each season, we will spatially smooth these estimated source contributions using the long-term spatial models that were proposed above. Although we do not directly observe measures of source contributions, we will check the ability of these long-term spatial models to predict source-specific exposures using tracer elements of each source. We will do this by first constructing estimates of the concentrations of tracer elements, for locations for which we have data during 1999-2009, and comparing them to the observed values. This strategy represents an out-of-sample validation step for the tracer elements.

(f) Long-Term Exposures to Individual Mixtures: We will investigate whether health effects associated with long-term exposure to individual pollutants differ across groups of geographical locations that have a similar air pollution mixture profile. Specifically, consider a data structure in which multivariate exposure to several individual pollutants is available in many geographical locations. We will apply the clustering methods outlined in Biostatistics Core (Section 4.4.2) and in the Exposure Core to the area-level exposure data to define groups of geographical locations having similar profiles of long-term air pollution. We will fit the following regression model

\[
[pollution]_{ij} = \gamma c_{i1} + \ldots + \gamma_{C-i} c_{C-i} + \beta_1 c_{i1} pm2.5_{ij} + \ldots + \beta_c c_{C} pm2.5_{ij}
\]

where \( c_{ki} \) denotes a variable that indicates whether or not subject \( i \) resides in an area that belongs to cluster \( k \). We will use a three-step procedure. First, we will apply the long-term spatial models described in the Biostatistics Core (Section 4.4.3) to the long-term monitoring data proposed to be collected during 2010-2014 as part of this Center (see the Exposure and Engineering Cores Core for a description of this proposed monitoring effort). Second, we will use the resulting model fits to compute estimated average concentrations for each modeled air pollution component at the zip code level. Third, we will apply the clustering methods outlined above to zip code level concentration averages.

2.7.4. Covariate Selection
Covariate selection will be determined \textit{a priori} and will be outcome-specific, based on biological factors and other study findings. For all outcomes, variables will include age, sex, BMI, season, temperature, and day of the week. Other covariates (e.g. medication use) will be included if it that covariate has potential to be a confounder based on the structure of the association with the exposure and outcome of interest.
2.7.5 Statistical Methods for Specific Aims

**Aim 1: Cognitive Function**
To evaluate the long-term effects on the longitudinal MMSE, we will use a nonparametric regression model, to assess declines in cognitive function as a function of age, and whether this decline is associated with pollution exposures. Specifically, we will consider models of the form:

\[ Y_{ij} = \beta_0 + \sum_{i} (age_{ij}) + b_i + \epsilon_{ij} \]

where \( Y_{ij} \) is the MMSE outcome for subject \( i \) on day \( j \), \( e_i = 0, 1, \) or \( 2 \) is a categorical exposure indicator defined by the tertile of a given pollutant, and \( age_{ij} \) is the age of subject \( i \) on day \( j \). This model accommodates the well-known nonlinear trend of MMSE score as a function of age, due to the fact that there typically exists an age threshold at which subjects begin to experience cognitive decline, and allows this trend to vary by low, medium, and high levels of chronic exposure. The subject-specific random effects account for correlation among repeated assessments taken on the same subject. Furthermore, we will also model air pollution level as a continuous predictor, allowing us to estimate the rate of cognitive decline by contrasting any exposure difference of interest. Cognitive tests performed in the Third Generation participants (CERAD Word List Memory and Victoria Stroop) are available at only one time and will therefore be examined at the time of visit 2 using a similar approach to explore non-linear trends by age.

**Aim 2: Vascular Function**
Outcomes reflecting vascular function will be analyzed using linear regression as described above (Section 2.7.2) to examine cross-sectional associations. Blood pressure (DBP, SBP, pulse pressure and MAP) will be analyzed longitudinally with mixed models to account for individual trajectories over 3 study visits. Flow mediated diameter was examined only once and will therefore be examine cross-sectionally as a continuous outcome for associations between baseline brachial artery diameter, FMD (mm), and FMD (%). Peripheral arterial tonometry will be assessed using the PAT ratio. Specifically, to estimate the effects of exposures on blood pressure, models will follow the form described in Section 2.7.2,

\[ Y_{ij} = \beta_0 + \beta_1 age_{ij} + \beta_2 temperature_{ij} + ... + [pollution]_{ij} + b_i + \epsilon_{ij} \]

where \( b_i \) represents a subject specific intercept. Covariates to be included in these models include: body mass index (BMI), hypertension medication usage, alcohol intake, and diabetes mellitus diagnosis.

**Aim 3: Susceptibility and Vulnerability**
Aim 3 will be an extension of Aims 1 and 2, and will utilize cross product terms for the categorical markers of susceptibility (diabetes diagnosis, obesity) and will consider both categorical and linear associations with age. We will investigate factors that affect susceptibility in the above models by including interaction terms for pollution and an indicator of potential modifier of interest.

2.7.6 Power Calculations
Benjamin et al. examined the clinical correlates and heritability of endothelial function as measured by FMD within the Framingham Offspring Study population. The final results included 2,883 participants in the Offspring Cohort. We will also examine the 3,813 Third Generation participants with FMD measures. We expect that our cross-sectional analyses will include approximately 6,000 individuals from the Offspring and Third Generation cohorts. For
example, with a sample size of approximately 6,000 subjects we will have >80% power to detect a difference in systolic and diastolic blood pressures of 1.5 and 0.8 mmHg respectively, for an interquartile range difference in black carbon, assuming a 2-sided alpha of 0.05. Likewise, we will have >80% power to detect a difference in baseline FMD of .01 mm. For cognitive outcomes we will have fewer observations available, as we will analyze associations in the Offspring and Third Generation cohorts separately. We conservatively estimate that we will have CERAD Word List Recall data on 3,000 of the 3,813 Third Generation participants with FMD measurements. Based on our pilot data from the NAS, we estimate that we will have 80% power to detect a one-unit change in total word recall score, a composite of the 3 recall trials. These power estimates are based on simplifying assumptions linear regression. Thus our actual analyses will have substantially more power and will be able to detect even smaller effects.

3. EXPECTED RESULTS, BENEFITS, OUTPUTS AND OUTCOMES

We will estimate health risks associated with short and long-term exposure to individual air pollutants, sources and air pollution mixtures within the Framingham Offspring and Third Generation populations. We will address which individual and area-level factors, measuring vulnerability, susceptibility, and individual air pollutants, sources and mixtures, are major determinants in explaining spatial and temporal variability of the health risks. Also, we will be able to add to data addressing the effects of pollutants on the life course examined in Project 2: the Normative Aging Study by investigating cognitive performance and vascular function in populations of middle- and older-aged men and women. Understanding which sources and mixtures are more harmful can aid decision-makers in developing targeted air quality regulations and effective management of air quality. Identifying which populations are more susceptible, due to variation in effect by race, pre-existing health conditions, or gender, will help inform research on biological mechanisms. Thus our findings may provide evidence of air pollution health impacts encompassing both susceptibility, through biological characteristics, and vulnerability, through non-biological characteristics. The Center will communicate the research findings through scientific publications, presentations, the Center Website, and progress reports.

This project will incorporate rich data from the Framingham Offspring and Third Generation Studies on vascular and cognitive function. The integration of this data with our Center Exposure and Statistics cores will create new opportunities to test additional novel hypotheses for which external funds beyond this proposal will be sought. These hypotheses include but are not limited to evaluating the effects of pollutants mixtures and sources on health effects such as: 1) Cerebrovascular function as assessed by Magnetic Resonance Imaging (MRI); 2) Adipokines (including Tumor Necrosis Factor-alpha (TNF-α), Plasminogen activator inhibitor-1 (PAI-1), and Interleukin-6 (IL6)); 3) Carotid intima–media thickness (cIMT); and 4) Biomarkers including matrix-metalloproteases (MMPs)

4. GENERAL PROJECT INFORMATION

4.1 Investigators: Our team includes experts in the analyses of cardiovascular and air pollution data. Dr. Mittleman (PI) and Dr. Schwartz (co-PI) have established a track record of collaborating on projects involving analysis of large epidemiologic studies. Dr. Mittleman (PI) is an Associate Professor of Medicine and Epidemiology at Harvard Medical School and School of Public Health and Director of the BIDMC Cardiovascular Epidemiology Research Unit. He has extensive expertise in conducting cardiovascular epidemiologic studies. For the past decade he has been actively involved in studies of the effects of air pollution on cardiovascular outcomes.
Dr. Schwartz is one of the world’s leading environmental epidemiologists. He has numerous publications on cognitive decline as a function of environmental exposures and was senior author of a major review of conceptual and methodological issues in analyzing confounding and interaction by socioeconomic and psychosocial factors in studies of air pollution. He has had a leading role in identifying differences in the effects of particles, ozone, and temperature by gender, socio-economic position, pre-existing medical conditions, diet, age, race, and particle composition.

Through many years of collaboration within the EPA Center, and through the NIEHS Program Project ES009825, Dr. Gold has worked with Drs. Mittleman and Schwartz in evaluation of multiple pollutant effects on vascular function, pulmonary and systemic inflammation and oxidative stress in vulnerable populations. Currently, they collaborate on a study that she leads on Boston pollution effects on vascular function in people with diabetes or coronary artery disease. In the Framingham EPA Center Project, she will be a co-Investigator in evaluation of the effects of pollutants, sources and mixtures on vascular responses and inflammation, and will contribute her experience in measurement and interpretation of these outcomes. We will also collaborate with Framingham investigators including Drs. Emelia Benjamin, Dan Levy, Naomi Hamburg, and Sudha Shesadri.

4.2 Facilities: Analyses for this project will take place at Harvard School of Public Health and Beth Israel Deaconess Medical Center. Extensive computing resources are available at BIDMC, including Dell Optiplex PCs, MS Office applications, SAS, Stata, Arc-GIS and R statistical software, and a password-protected LAN. Our group uses a password-protected shared research server that is maintained by the IT department. All networked drives are mirrored and backed up daily. We have access to the Harvard Medical School research computing cluster, which utilizes the Linux operating system, LSF queuing and cluster software from Platform Computing, providing an open and powerful environment for computational research. Hardware includes 162 compute nodes for computational work and an infrastructure built on 7 Xeon servers. Other resources include full digital access to the Harvard Libraries. Our group has successfully used geocoding software and has established access to databases of area-based measures of socioeconomic position. ArcGIS 9.1 (ESRI, Inc) offers an integrated collection of graphical information systems software products that enables us to geocode addresses and to assign area-based measures of socioeconomic position to individuals. Investigators participating in Project 3 will meet weekly. The Steering Committee will meet monthly at Harvard to monitor research progress, review research initiatives within and outside of the Center, and manage operations. Steering Committee meetings will be held in the Exposure, Epidemiology, and Risk conference room at Harvard, which is fully set up for telephone and video-conferencing, with up to four outside participants so that full committee attendance can be assured.
References


