Pollutant composition modification of the effect of air pollution on progression of coronary artery calcium: the Multi-Ethnic Study of Atherosclerosis

Joshua P. Keller,

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Timothy V. Larson,

Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA

Elena Austin,

Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA, USA

R. Graham Barr,

Departments of Medicine and Epidemiology, Columbia University, New York, NY, USA

Lianne Sheppard,

Departments of Environmental and Occupational Health Sciences and Biostatistics, University of Washington, Seattle, WA, USA

Sverre Vedal,

Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA, USA

Joel D. Kaufman, and

Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, WA, USA

Adam A. Szpiro

Department of Biostatistics, University of Washington, Seattle, WA, USA

Abstract

Background—Differences in traffic-related air pollution (TRAP) composition may cause heterogeneity in associations between air pollution exposure and cardiovascular health outcomes. Clustering multi-pollutant measurements allows investigation of effect modification by TRAP profiles.

Corresponding Author: Joshua P. Keller, Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205, kellerjp@uw.edu, 443-287-8782.

Conflict of Interest: None declared.

Methods—We measured TRAP components with fixed-site and on-road instruments for two two-week periods in Baltimore, Maryland. We created representative TRAP profiles for cold and warm seasons using predictive *k*-means clustering. We predicted cluster membership for 1005 participants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution with follow-up between 2000 and 2012. We estimated cluster-specific relationships between coronary artery calcification (CAC) progression and long-term exposure to fine particulate matter ($PM_{2.5}$) and oxides of nitrogen (NO_X).

Results—We identified two clusters in the cold season, notable for higher ratios of gases and ultrafine particles, respectively. A 5 μg/m^3 difference in $PM_{2.5}$ was associated with 17.0 (95% Confidence Interval [CI]: 7.2, 26.7) and 42.6 (95% CI: 25.7, 59.4) Agatston units/year CAC progression among participants in clusters 1 and 2, respectively (effect modification *p*=0.006). A 40ppb difference in NO_X was associated with 22.2 (95% CI: 7.7, 36.7) and 41.9 (95% CI: 23.7, 60.2) Agatston units/year CAC progression in clusters 1 and 2, respectively (*p*=0.08). Similar trends occurred using clusters identified from warm season measurements. Clusters correlated highly with baseline pollution level.

Conclusions—Clustering TRAP measurements identified spatial differences in composition. We found evidence of greater CAC progression rates per unit $PM_{2.5}$ exposures among people living in areas characterized by high ratios of ultrafine particle counts relative to NO_X concentrations.

Introduction

The relationship between exposure to traffic-related air pollution and cardiovascular morbidity and mortality has been well-established.[1–6] Traffic-related air pollution (TRAP) is a complex mixture of many different particulate and gaseous components that can vary across fine spatial scales[7] and is of particular interest in urban areas[1]. Our focus in this study is on directly emitted TRAP whose composition and concentration levels vary across an urban area with traffic intensity[8], distance to roadway[7], fuel type (heavy duty diesel vs. gasoline), age and condition of vehicle, and source[9] (tailpipe, brake wear, tire wear, etc.). There is a recognized need to move beyond single-pollutant epidemiological analyses and consider the effects of exposure to mixtures of pollutants.[10–14]

Complex multi-pollutant datasets are often analyzed using dimension reduction techniques, which simplify the complex variability of the data into a smaller set of features. Clustering methods, which provide a promising approach for understanding multi-pollutant health effects,[10] partition multi-pollutant observations into a pre-specified number of groups, or clusters. This provides a categorical division of the data based upon pollutant profile that simplifies the interpretation of simultaneous exposure to multiple pollutants.[15,16] In the popular '*k*-means' algorithm, clusters are selected to minimize the (Euclidean) distance between each observation and the center, or representative exposure vector, of its assigned cluster.[17] For analyses of administrative data, records can be assigned to a cluster based upon city.[16,18] For cohort studies, a classification model can be used to predict cluster membership at subject residences.[15] Cluster membership for each subject or record can then serve as an effect modifier for an association between a single exposure and outcome.[10,15,16] This allows for heterogeneity in the association between a single composite pollutant (e.g.,

NO_X, $PM_{2.5}$) and health outcomes to be identified across groups distinguished by predicted differences in pollution composition at subject locations.

Prior modeling of spatial variation in TRAP across cities has focused on separate land-use regression models for ultrafine particles[19–22] and VOCs[23]. While these models can provide estimates of associations with differences in level of a single TRAP component, the single component approach cannot capture the mixture features identifiable from considering variation in multiple TRAP components simultaneously as can be done through clustering.

The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air) investigated the association between long-term air pollution exposure and progression of subclinical measures of atherosclerosis, including coronary artery calcification (CAC).[24] Using spatiotemporal predictions of individual pollutants, Kaufman et al.[6] found an association between CAC progression and higher levels of NO_X and $PM_{2.5}$ exposure. Recent supplemental monitoring campaigns have obtained an expansive suite of multi-pollutant measurements in MESA Air cities.[25,26] These observations provide the opportunity to explore variation in CAC progression due to differences in TRAP composition.

Methods

Study Population and Exposure Assessment

The MESA and MESA Air cohorts have been described extensively previously.[6,24,27] In Baltimore, MD, 1081 participants were recruited between July 2000 and August 2002. Subjects received CT scans at baseline and at multiple follow up visits through 2012. Scanner type varied by visit, with most baseline scans made using a Aquilion scanner (Toshiba) and later follow-up exams using a Volume Zoom scanner (Siemens). Scans were scored for coronary artery calcium using the Agatston method.[28] Baseline characteristics of the cohort are summarized in Table 1 for the 1005 participants for whom complete covariate and exposure information was available. The study protocol was approved by the institutional review board at the coordinating and local study centers and participants provided written informed consent.

Predictions of participant-specific outdoor ambient exposures of NO_X and $PM_{2.5}$ were made using a standardized set of spatiotemporal prediction models developed for MESA Air.[29] These predictions were resolved to the exact residential addresses of participants at a two-week time scale. Cross-validated R^2 for the models indicated good out-of-sample prediction accuracy (R^2 of 0.92 and 0.84, for NO_X and $PM_{2.5}$, respectively).[29] For modeling CAC, NO_X and $PM_{2.5}$ exposure predictions were aggregated into long-term averages from recruitment through exam visit, based upon participant residential history.

Multipollutant TRAP Data

Pollution measurements were made at 43 intersections within the Baltimore, MD metropolitan area during two-week periods in February 2012 and June 2012, which we refer to as the cold and warm seasons, respectively. Measurements were made in two seasons to capture differences in TRAP composition due to changes in sources and meteorology throughout the year. Most of the intersections were in residential areas and not on major

roads, given the goal to characterize TRAP profiles relevant to subjects' ambient exposures. Measurements of nitrogen dioxide (NO_2), oxides of nitrogen (NO_X), ozone (O_3), and specific volatile organic compounds (VOCs) were made using stationary badge monitors (see Table 2). Carbon monoxide (CO) concentrations and particle number (PN) counts for different size bins were measured using an on-road mobile platform at a collection of locations in and near the intersection by traversing the blocks bordering the intersection of interest. The smallest size bin for PN counts captured particles 25nm to 400nm in diameter. Riley et al.[26] provide a detailed description of the sample collection procedures. The mobile measurements were made during the afternoon commuting period and were adjusted for day-to-day variability by subtracting the fifth percentile of each pollutant and taking the median value of all adjusted observations at each location.[26]

We scaled observations from both the badge and mobile platforms by the measured badge NO_X concentration at the respective location. Standardization by NO_X level allows clusters derived from the data to be informed by relative pollution composition and not solely by overall level. We then standardized these pollution fractions to have mean zero and unit standard deviation.

Clustering

Using the predictive *k*-means method,[15] we clustered the component species fractions separately by season. The predictive *k*-means method identifies cluster centers by simultaneously optimizing the deviation of the multi-pollutant observation from its assigned cluster center and the assignment of each monitor location to a cluster based upon prediction variables, which are typically a function of spatial location. This results in clusters whose membership can be better predicted at subject locations compared to clusters from traditional *k*-means clustering, which does not incorporate prediction variables when identifying cluster centers. The predictive *k*-means method is implemented in the publicly-available R package 'predkmeans', and additional technical details are provided in eAppendix 1.

We used a large collection of geographic variables containing land use information, distance to roadways, emissions inventories, and other values derived from Geographic Information Systems (GIS). These covariates, listed in eTable 1, have been used previously to develop spatial and spatiotemporal prediction models.[29–32] We reduced the values of these geographic variables at monitor locations to a small set of principal component analysis (PCA) scores, which we included as prediction variables for modeling cluster assignment predictive *k*-means.

We chose the number of clusters and the number of PCA scores by ten-fold cross validation (CV). We limited the models to between 2 and 5 clusters and 1 and 6 PCA scores. Models were compared according to their mean squared prediction error (MSPE), which is the sum of the squared distances between the observed pollutant fractions (after standardization) and their predicted cluster centers.[15] This provides an aggregate estimate of the error in using predicted cluster membership relative to the observed value for each measured TRAP component. eAppendix 2 provides additional detail for calculating this CV metric.

We predicted cluster membership at each participant residence using a multinomial logistic regression model (see eAppendix 1). The covariates for prediction were PCA scores evaluated using geographic covariates at participant residence locations and based upon the relationship between variables and scores at monitor locations. We assigned cluster membership for each participant record according to residence at the time of the exam.

CAC Progression Analyses

Following the approach of the primary MESA Air analyses,[6] we estimated the association between CAC progression and pollution exposure ($PM_{2.5}$ or NO_X) via a mixed model. The model includes a cross-sectional component that models baseline CAC, a longitudinal component that accommodates time-varying confounders, and a time-varying component that includes transient factors affecting measurements.[6,33] Variables included are age, sex, race/ethnicity, site, scanner type, adiposity, physical activity level, smoking and second-hand smoke exposure status, employment outside the home, total cholesterol level, high density lipoprotein level, triglyceride level, statin use, an index of neighborhood socioeconomic status,[34] education and income. We excluded all data from participants after a coronary revascularization procedure.

We included cluster membership as an effect modifier for the longitudinal associations between CAC and $PM_{2.5}$ (or NO_X). In total, we fit four separate models, corresponding to the two pollutants of interest and the two groupings of the cohort based upon the cold and warm season clusters. The coefficients from the cluster-pollutant-time interaction provide cluster-specific estimates of the association between pollution exposure ($PM_{2.5}$ or NO_X) and CAC progression. Statistical significance of the effect modification was assessed using a likelihood ratio test.

Sensitivity Analyses

Because we found strong correlation between membership in the identified clusters and baseline exposure levels for participants, we conducted sensitivity analyses that estimated cluster-specific CAC progression associations for alternative cluster definitions. To compare against clusters not derived from the TRAP measurements, we split the cohort into those with baseline pollution exposure (NO_X or $PM_{2.5}$) above and below the cohort-wide median level. As a second sensitivity analysis, we orthogonalized the GIS covariates against baseline NO_X level (using year 2000 annual averages for monitor locations) and then computed new PCA scores from these modified covariates and clustered the pollutant measurements via predictive k-means. This sensitivity analysis was designed to identify clusters that were less correlated with the regional trend in baseline exposure levels.

Results

TRAP Data

The locations of the 43 monitoring sites are provided in Figure 1. Three (different) locations in each of the cold and warm seasons were removed due to instrument error in the processing of badge measurements, leaving forty sites for each season. Figure 2 shows a

heatmap of the correlations between the pollutants in the cold season, after scaling by NO_X. eFigure 1 provides the analogous plot for the warm season data.

Clustering Results

In the cold season, the predictive *k*-means model with 3 clusters and 2 PCA scores performed the best in CV (MSPE=16.82) (eTable 1). However, parameters in this model were not fully identifiable, since one cluster included only two locations but had three coefficients for classification. Therefore, we chose the model with 3 clusters and 1 PCA score for the cold season since it was not overdetermined and had the second-best CV performance (MSPE=16.94). In the warm season, the best CV results were for the model with 3 clusters and 2 PCA scores (MSPE = 15.32).

The cluster centers from the cold season are depicted in Figure 3. Most monitors were assigned to Cluster 1 (21 locations) or Cluster 2 (17 locations). Cluster 1 was characterized by above average ratios of NO_2 and ozone relative to NO_X, while the profile for Cluster 2 showed the opposite trend, with lower fractions of gases but higher ratios of ultrafine (25-400nm) and accumulation mode (0.05-1μm) particle counts relative to NO_X. Locations west and north of downtown were primarily assigned to Cluster 1, while those assigned to Cluster 2 were located closer to downtown (see Figure 1). Cold season Cluster 3, which comprised two locations, had high ratios of almost all gases and particle sizes. A summary of cluster attributes is provided in Table 3.

We predicted that, at baseline, 743 and 262 participants belonged to cold season Clusters 1 and 2, respectively, and none to Cluster 3. Table 1 summarizes baseline characteristics of the study cohort, stratified by cold season and warm season cluster membership. In addition to the clear geographic patterns between clusters, participants in Cluster 1 tended to have higher levels of education and socioeconomic status than those in Cluster 2, while baseline NO_X and $PM_{2.5}$ exposure levels were higher in Cluster 2 (Figure 4).

The centers for the clusters identified in the warm season are depicted in eFigure 2. The first cluster (to which 19 locations were assigned) was characterized by lower ratios of VOCs and higher ratios of NO_2, O_3, and PN_{1-3}. The second cluster (12 locations) had the highest ratios of all VOCs and CO, but low ratios of particle counts. The third cluster (9 locations) was notable for its low PN_{1-3} counts and NO_2 and O_3 concentrations.

Warm season cluster membership also showed a spatial pattern, with monitors located in the urban center primarily assigned to Cluster 1 and locations furthest from downtown generally assigned to Cluster 3 (eFigure 3). Most participants (701 at baseline) were predicted to belong to Cluster 1 in the warm season. A total of 151 and 153 participants were predicted to belong to Clusters 2 and 3, respectively. Participants in Cluster 2 tended to have lower levels of income, education, and neighborhood socioeconomic status (Table 1). The warm season clusters also showed patterns in baseline NO_X and $PM_{2.5}$ concentrations at residences (Figure 4).

In the sensitivity analysis using clusters based upon GIS covariates with correlation with baseline NO_X removed, the best models (according to CV MSPE) had 3 clusters and 2 PCA

scores in the cold season and 2 clusters and 4 PCA scores in the warm season (eTable 1). Similar to cold-season Cluster 1 from the primary analysis, cold season Cluster 1 from the sensitivity analysis had below-average ratios for all TRAP components relative to NO_X except for ultrafine and accumulation mode particle counts (eFigure 4). However, this cluster included several monitors from suburban areas in addition to those located in the downtown area (eFigure 5). Cold season Cluster 2 from the sensitivity analysis had low ratios of ultrafine and accumulation mode particle counts, similar to Cluster 1 from the primary analysis. Cluster 3 had high ratios of almost all TRAP components relative to NO_X, but only included four monitors. In the warm season, Cluster 1 broadly resembled a combination of Cluster 1 and 3 from the primary analysis and had high particle count fractions and low VOC fractions, while Cluster 2 showed the opposite trend (eFigure 6). In both seasons, the baseline NO_X exposures for participants did not show a strong correlation with cluster membership (Figure 4).

CAC Progression Estimates

In a model without effect modification by cluster, a difference of 5 μg/m^3 in $PM_{2.5}$ was associated with 23.0 Agatston units per year CAC progression (95% Confidence Interval [CI]: 14.2, 31.7). When cold season cluster membership was used as an effect modifier, the estimated association with a 5 μg/m^3 difference in $PM_{2.5}$ was 17.0 (95% CI: 7.2, 26.7) units/ year for participants belonging to Cluster 1 and 42.6 (95% CI: 25.7, 59.4) units/year for participants in Cluster 2 (see Table 4). The model with cluster-specific progression terms was significantly different from the city-wide model without cluster interactions (p=0.006). Effect modification by warm season cluster was not statistically significant (p=0.10) but showed a similar trend: 17.1 (95% CI: 7.1, 27.0), 24.7 (95% CI: –0.1, 49.6), and 39.8 (95% CI: 20.0, 59.6) Agatston units/year CAC progression associated with 5 μg/m^3 difference in $PM_{2.5}$ among participants in Clusters 1, 2, and 3, respectively. The median baseline $PM_{2.5}$ concentration was 15.77 μg/m^3. The estimated associations (per 5 μg/m^3 difference in $PM_{2.5}$) were 13.7 (95% CI: 2.1, 25.4) and 33.1 (95% 21.2, 44.9) units/year for participants with baseline exposure below and above this value, respectively.

The estimated associations with NO_X exposure showed a similar trend across clusters (see Table 5). A difference of 40ppb in NO_X exposure was associated with 22.2 (95% CI: 7.7, 36.7) and 41.9 (95% CI: 23.7, 60.2) Agatston units/year CAC progression among participants in cold season Clusters 1 and 2, respectively. These results were not significantly different (p=0.076) from the estimate for all participants pooled together (28.2 units/year, 95% CI: 17.1, 39.3). Estimates for the warm season clusters followed the same pattern as for $PM_{2.5}$: 20.7 (5.9, 35.6), 25.7 (–2.0, 53.4), and 38.1 (16.8, 59.5) for Clusters 1, 2, and 3, respectively. There was no evidence for effect modification by baseline NO_X (p=0.64), although the point estimates had a similar trend to results from the model with effect modification by cold season cluster.

In the sensitivity analysis using covariates that had been orthogonalized with respect to NO_X, there was little difference in the estimated association between pollution exposure and CAC progression between clusters from the warm season (eTable 2). When clustering subjects using the cold season data, the estimates from Cluster 1 vs. 2 were significantly

different (p=0.023 and p=0.017 for $PM_{2.5}$ and NO_X, respectively). The confidence intervals for the progression estimates in Cluster 2 were notably broad and included zero for NO_X and $PM_{2.5}$. In cold season Cluster 1, the estimated association between CAC progression and both $PM_{2.5}$ and NO_X exposure was similar to, but greater than, the estimates from the city-wide analysis without clustering. Similar to the trend observed in the original analysis, the largest associations with CAC progression were observed among people predicted to be exposed to pollution with higher ratios of ultrafine and accumulation mode particles relative to NO_X.

Discussion

We have used a novel collection of near- and on-road multipollutant measurements to assess how spatial differences in pollution composition affect the relationship of $PM_{2.5}$ and NO_X exposure with coronary artery calcification. By reducing dimension of the measurements using clustering and categorizing the cohort by these clusters, we were able to incorporate multi-pollutant information into a longitudinal model for CAC progression.

In our primary analysis, we found significant differences in the association between CAC progression and $PM_{2.5}$ exposure when grouping subjects by predicted cold season TRAP profile. The association was strongest among participants in Cluster 2, which was notable for its locations being primarily in the downtown area and its above average ratios of ultrafine and accumulation mode particle counts relative to NO_X. Freshly emitted vehicle exhaust is one likely source for these higher particle counts in the urban center.[9] When grouping by predicted warm season TRAP profiles, the estimated association between CAC progression and $PM_{2.5}$ exposure was weaker in the cluster with lower ratios of ultrafine particle counts (Cluster 2) and strongest in the cluster with high ratios of ultrafine particle counts (Cluster 3).

Our results suggest that overall $PM_{2.5}$ exposure among individuals whose ambient residential TRAP concentrations have high levels of ultrafine particles relative to NO_X has greater adverse cardiovascular impacts than exposure among participants whose ambient residential TRAP concentrations have different composition. Evidence for increased risk of atherosclerosis due to exposure to ultrafine particles, relative to larger particle sizes, has been found in mice[35]. Ultrafine particles have also been linked to a broader set of inflammation- and endothelial-related outcomes.[36,37] Although a defining feature of the identified TRAP profiles were the relative fractions of ultrafine particle counts, the moderately high correlation between ultrafine and accumulation mode particle (Figure 2 and eFigure 1) mean that the differences between clusters cannot be attributed to ultrafine particles alone. The lower fractions of VOCs and gases in cold season Cluster 2 and warm season Cluster 1 may also play a role in the observed differences.

A striking feature of the clustering results is the strong correlation between cluster membership, geographic location, and overall $PM_{2.5}$ and NO_X levels. The predictive k-means clustering procedure incorporated an aggregation of geographic covariates via the PCA scores, which have a strong gradient from downtown out to the suburbs. These covariates are similar to those used for predicting exposures in the MESA Air cohort,[29]

which may drive some of the correlation between cluster membership and exposure level. This correlation makes it difficult to determine whether the between-cluster differences identified are due to differences in TRAP composition or differences in baseline exposure. We addressed this concern by conducting a sensitivity analysis that removed variability from the covariates that is explainable by baseline NO_X and yielded clusters that did not correlate well with baseline exposure. Using the cold season TRAP measurements, the sensitivity analysis found larger associations between CAC progression and $PM_{2.5}$ exposure among participants predicted to have TRAP profiles with higher fractions of ultrafine particles (Cluster 1). This supports the primary results, which found greater rates of progression among people predicted to have TRAP profiles with higher fractions of ultrafine particles.

Predicted cluster membership correlated with socioeconomic status, in addition to baseline exposure and geographic location. Strong relationships between socioeconomic status and pollution exposure have been previously identified in this cohort.[38] The results from our analysis, however, identify differences that are not fully explainable by socioeconomic differences alone. In our primary analyses, participants in cold-season Cluster 1 and warm season Cluster 1 had higher levels of affluence and weaker, but still non-zero, estimated associations between CAC progression and $PM_{2.5}$ and NO_X exposure. However, participants in warm season Cluster 2 had the lowest income and neighborhood-level socioeconomic index, but their estimated association between $PM_{2.5}$ and NO_X exposure and CAC progression was very similar to the city-wide average. Additionally, participants in warm season Cluster 3 had higher average income and were more affluent, but also had a stronger estimated association between $PM_{2.5}$ and CAC progression than participants in Cluster 2.

We estimated TRAP profiles for the cold and warm season separately and found similar trends in the relative component fractions across seasons. This could be due to similar widespread sources of directly emitted TRAP in each location across seasons, in contrast to residential heating sources or secondary pollutants that vary by season. The trend of effect modification by predicted profile was similar as well, although with different levels of statistical evidence. This is likely due in part to the overlap between the predicted cluster membership.

A limitation of this analysis was the relatively small number of locations at which component measurements were made. Although a sample size of forty locations is far more than the one or two locations per city at which component data is typically available via the Chemical Speciation Network, it is nonetheless a relatively small sample for building a cluster prediction model. Including spatial splines in the prediction model as a mechanism for spatial smoothing is impractical with cluster sizes of 21 and 17 from the cold season data. This leads to the prediction model being derived from one or two PCA scores, which can capture small-scale variability but can also be dominated by larger trends, as was the case in this study. The results may be impacted by differences in the time period of the data. The geographic variables used for cluster prediction were from the period 2000-2006. The cluster profiles were derived from multi-pollutant measurements in 2012 but were used to predict representative exposure profiles for the entire study period. It is possible the TRAP profiles and the relationship between geographic covariates and TRAP within the Baltimore

region changed over time, although major highway patterns and industrial sources were largely stable.

The city-wide and cluster-specific estimates were all greater than the estimated association in the full MESA Air cohort, which includes participants from five other metropolitan areas. In the full cohort, differences of 5μg/m^3 difference in $PM_{2.5}$ and 40ppb in NO_X were associated with 4.1 (95% CI: 1.4, 6.8) and 4.8 (95% CI: 0.9, 8.7) Agatston units/year CAC progression, respectively.[6] While the exposure prediction models for Baltimore did have the best overall predictive accuracy,[29] the difference in progression estimates is not likely due to exposure assessment accuracy alone. Differences in pollution composition between the six MESA Air cities could potentially cause some of this difference, however such differences are masked by between-city differences in cohort members. The defining feature of the MESA cohort is the overrepresentation of different ethnicities. The Baltimore sub-cohort, however, only includes white and black participants, while the other cities also have different racial-ethnic groups, which could be one source of the differences from the Baltimore-only results.

We have presented a novel approach to using multipollutant TRAP measurements within a metropolitan area to assess effect modification by pollution composition in longitudinal relationships between pollution exposure and CAC. Our results found that the same difference in $PM_{2.5}$ or NO_X concentration was associated with faster CAC progression among participants living in areas predicted to have higher ratios of ultrafine particle counts relative to NO_X during the cold season. These results highlight how incorporating multi-pollutant measurements into health effect analyses can yield insight into heterogeneity in the relationships between air pollution exposure and health.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Hoek G, Krishnan RM, Beelen R, et al. Long-term air pollution exposure and cardio- respiratory mortality: a review. Environ Heal. 2013; 12(1):43.doi: 10.1186/1476-069X-12-43

2. Brook RD, Rajagopalan S, Pope CA, et al. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation. 2010; 121(21):2331–78. DOI: 10.1161/CIR.0b013e3181dbece1 [PubMed: 20458016]

3. Beelen R, Raaschou-Nielsen O, Stafoggia M, et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet. 2014; 383(9919):785–795. DOI: 10.1016/S0140-6736(13)62158-3 [PubMed: 24332274]

4. Dockery DW, Pope CA, Xu X, et al. An Association between Air Pollution and Mortality in Six US. Cities N Engl J Med. 1993; 329(24):1753–1759. [PubMed: 8179653]

5. Miller KA, Siscovick DS, Sheppard L, et al. Long-term exposure to air pollution and incidence of cardiovascular events in women. N Engl J Med. 2007; 356(5):447–458. [PubMed: 17267905]

6. Kaufman JD, Adar SD, Barr RG, et al. Association between air pollution and coronary artery calcification within six metropolitan areas in the USA (the Multi-Ethnic Study of Atherosclerosis and Air Pollution): a longitudinal cohort study. Lancet. 2016; 388:696–704. DOI: 10.1016/S0140-6736(16)00378-0 [PubMed: 27233746]

7. Karner AA, Eisinger DS, Niemeier DA. Near-roadway air quality: Synthesizing the findings from real-world data. Environ Sci Technol. 2010; 44(14):5334–5344. DOI: 10.1021/es100008x [PubMed: 20560612]

8. Fischer PH, Hoek G, Van Reeuwijk H, et al. Traffic-related differences in outdoor and indoor concentrations of particles and volatile organic compounds in Amsterdam. Atmos Environ. 2000; 34(22):3713–3722. DOI: 10.1016/S1352-2310(00)00067-4

9. Vu TV, Delgado-saborit JM, Harrison RM. Review: Particle number size distributions from seven major sources and implications for source apportionment studies. Atmos Environ. 2015; 122:114–132. DOI: 10.1016/j.atmosenv.2015.09.027

10. Oakes M, Baxter L, Long TC. Evaluating the application of multipollutant exposure metrics in air pollution health studies. Environ Int. 2014; 69:90–99. DOI: 10.1016/j.envint.2014.03.030 [PubMed: 24815342]

11. Greenbaum D, Shaikh R. First steps toward multipollutant science for air quality decisions. Epidemiology. 2010; 21(2):195–197. DOI: 10.1097/EDE.0b013e3181ccc52a [PubMed: 20160562]

12. Vedal S, Kaufman JD. What does multi-pollutant air pollution research mean? Am J Respir Crit Care Med. 2011; 183(1):3–4. DOI: 10.1164/rccm.201011-1881ED [PubMed: 21193782]

13. Dominici F, Peng R, Barr CD, Bell ML. Protecting human health from air pollution: Shifting from a single-pollutant to a multi-pollutant approach. Epidemiology. 2010; 21(2):187–194. DOI: 10.1097/EDE.0b013e3181cc86e8 [PubMed: 20160561]

14. Health Effects Institute. HEI Strategic Plan for Understanding the Health Effects of Air Pollution 2015-2020. Boston, MA: 2015.

15. Keller JP, Drton M, Larson T, Kaufman JD, Sandler DP, Szpiro AA. Covariate-adaptive clustering of exposures for air pollution epidemiology cohorts. Ann Appl Stat. 2017; 11(1):93–113. DOI: 10.1214/16-AOAS992 [PubMed: 28572869]

16. Kioumourtzoglou M-A, Austin E, Koutrakis P, Dominici F, Schwartz J, Zanobetti A. PM2.5 and survival among older adults: Effect modification by particulate composition. Epidemiology. 2015; 26(3):321–327. DOI: 10.1097/EDE.0000000000000269 [PubMed: 25738903]

17. Hartigan J, Wong M. Algorithm AS 136: A k-means clustering algorithm. Appl Stat. 1979; 28(1):100–108.

18. Austin E, Coull BA, Zanobetti A, Koutrakis P. A framework to spatially cluster air pollution monitoring sites in US based on the PM2.5 composition. Environ Int. 2013; 59:244–54. DOI: 10.1016/j.envint.2013.06.003 [PubMed: 23850585]

19. Hoek G, Beelen R, Kos G, et al. Land use regression model for ultrafine particles in Amsterdam. Environ Sci Technol. 2011; 45(2):622–628. DOI: 10.1021/es1023042 [PubMed: 21158386]

20. Abernethy RC, Allen RW, McKendry IG, Brauer M. A land use regression model for ultrafine particles in Vancouver, Canada. Environ Sci Technol. 2013; 47(10):5217–5225. DOI: 10.1021/es304495s [PubMed: 23550900]

21. Rivera M, Basagaña X, Aguilera I, et al. Spatial distribution of ultrafine particles in urban settings: A land use regression model. Atmos Environ. 2012; 54:657–666. DOI: 10.1016/j.atmosenv.2012.01.058

22. Weichenthal S, Van Ryswyk K, Goldstein A, Shekarrizfard M, Hatzopoulou M. Characterizing the spatial distribution of ambient ultrafine particles in Toronto, Canada: A land use regression model. Environ Pollut. 2016; 208:241–248. DOI: 10.1016/j.envpol.2015.04.011 [PubMed: 25935348]

23. Mukerjee S, Smith LA, Johnson MM, Neas LM, Stallings CA. Spatial analysis and land use regression of VOCs and NO2from school-based urban air monitoring in Detroit/Dearborn, USA. Sci Total Environ. 2009; 407(16):4642–4651. DOI: 10.1016/j.scitotenv.2009.04.030 [PubMed: 19467697]

24. Kaufman JD, Adar SD, Allen RW, et al. Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease: The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Am J Epidemiol. 2012; 176(9):825–37. DOI: 10.1093/aje/kws169 [PubMed: 23043127]

25. Xu W, Riley EA, Austin E, et al. Use of mobile and passive badge air monitoring data for NOX and ozone air pollution spatial exposure prediction models. J Expo Sci Environ Epidemiol. 2017; 27(2):184–192. DOI: 10.1038/jes.2016.9 [PubMed: 27005742]

26. Riley EA, Schaal LN, Sasakura M, et al. Correlations between short-term mobile monitoring and long-term passive sampler measurements of traffic-related air pollution. Atmos Environ. 2016; 132:229–239. DOI: 10.1016/j.atmosenv.2016.03.001

27. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: Objectives and design. Am J Epidemiol. 2002; 156(9):871–881. DOI: 10.1093/aje/kwf113 [PubMed: 12397006]

28. Carr JJ, Nelson JC, Wong ND, et al. Calcified Coronary Artery Plaque Measurement with Cardiac CT in Population-based Studies: Standardized Protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) Study1. Radiology. 2005; 234(1):35–43. DOI: 10.1148/radiol.2341040439 [PubMed: 15618373]

29. Keller JP, Olives C, Kim S-Y, et al. A unified spatiotemporal modeling approach for predicting concentrations of multiple air pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution. Environ Health Perspect. 2015; 123(4):301–309. DOI: 10.1289/ehp.1408145 [PubMed: 25398188]

30. Mercer LD, Szpiro AA, Sheppard L, et al. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). Atmos Environ. 2011; 45(26):4412–4420. DOI: 10.1016/j.atmosenv.2011.05.043

31. Wang M, Keller JP, Adar SD, et al. Development of long-term spatiotemporal models for ambient ozone in six metropolitan regions of the United States: The MESA Air study. Atmos Environ. 2015; 123(2):79–87. DOI: 10.1016/j.atmosenv.2015.10.042

32. Bergen S, Sheppard L, Sampson PD, et al. A National Prediction Model for PM2.5 Component Exposures and Measurement Error-Corrected Heatlh Effect Inference. Environ Health Perspect. 2013; 121(9):1017–1025. DOI: 10.1289/ehp.1206010 [PubMed: 23757600]

33. Gassett AJ, Sheppard L, McClelland RL, et al. Risk Factors for Long-Term Coronary Artery Calcium Progression in the Multi-Ethnic Study of Atherosclerosis. J Am Heart Assoc. 2015; 4(8):e001726.doi: 10.1161/JAHA.114.001726 [PubMed: 26251281]

34. Moore K, Diez Roux AV, Auchincloss A, et al. Home and work neighbourhood environments in relation to body mass index: the Multi-Ethnic Study of Atherosclerosis (MESA). J Epidemiol Community Health. 2013; 67(10):846–53. DOI: 10.1136/jech-2013-202682 [PubMed: 23868527]

35. Araujo JA, Barajas B, Kleinman M, et al. Ambient particulate pollutants in the ultrafine range promote early atherosclerosis and systemic oxidative stress. Circ Res. 2008; 102(5):589–596. DOI: 10.1161/CIRCRESAHA.107.164970 [PubMed: 18202315]

36. Nel A, Xia T, Lutz M, Li N. Toxic Potential of Materials at the Nanolevel. Science (80-). 2007; 311(5726):622–627. DOI: 10.1126/science.1114397

37. Oberdörster G, Oberdörster E, Oberdörster J. Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles. Environ Health Perspect. 2005; 113(7):823–839. DOI: 10.1289/ehp.7339 [PubMed: 16002369]

38. Hajat A, Diez-roux AV, Adar SD, Auchincloss AH, Lovasi GS, Neill MSO. Air Pollution and Individual and Neighborhood Socioeconomic Status: Evidence from the Multi-Ethnic Study of Atherosclerosis (MESA). Environ Health Perspect. 2013; 121(11):1325–1333. [PubMed: 24076625]

What this study adds

This paper presents a novel investigation of how differences in traffic-related air pollution may modify the relationship between long-term particulate matter exposure and cardiovascular health. A distinctive feature is the use of near- and on-road pollutant measurements to inform predictions of pollution profiles using modern statistical clustering methods. This study provides evidence of greater health impact of exposure to pollution with higher proportions of ultrafine particles within metropolitan regions.
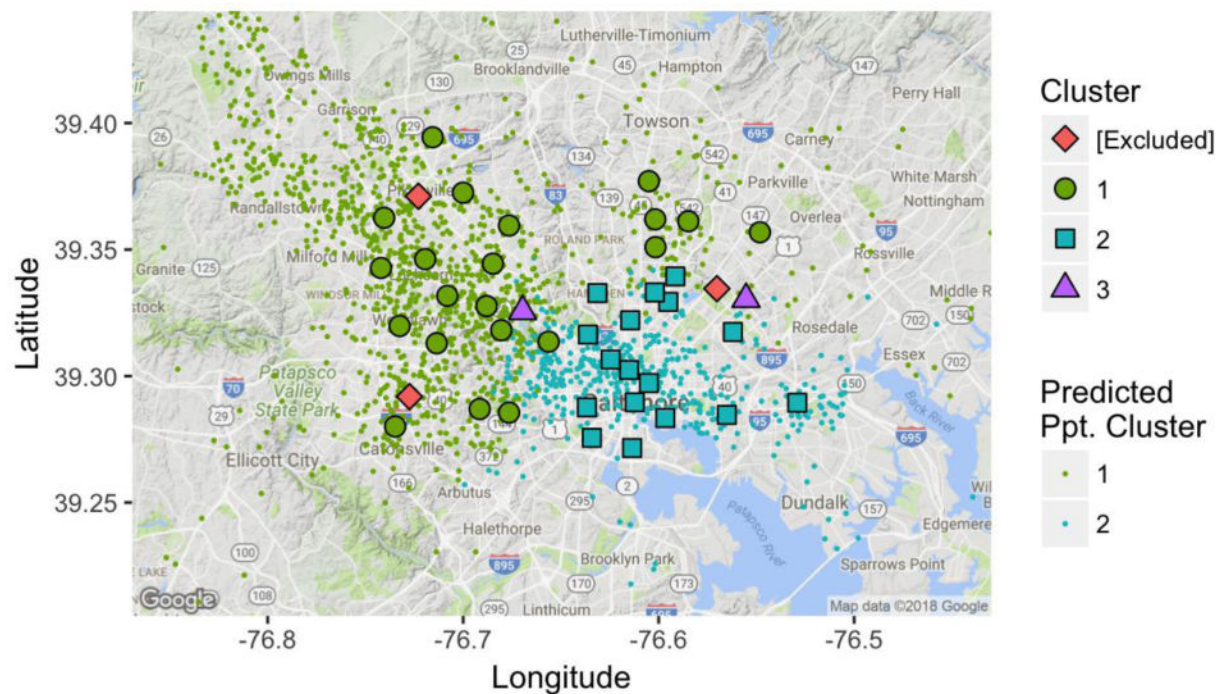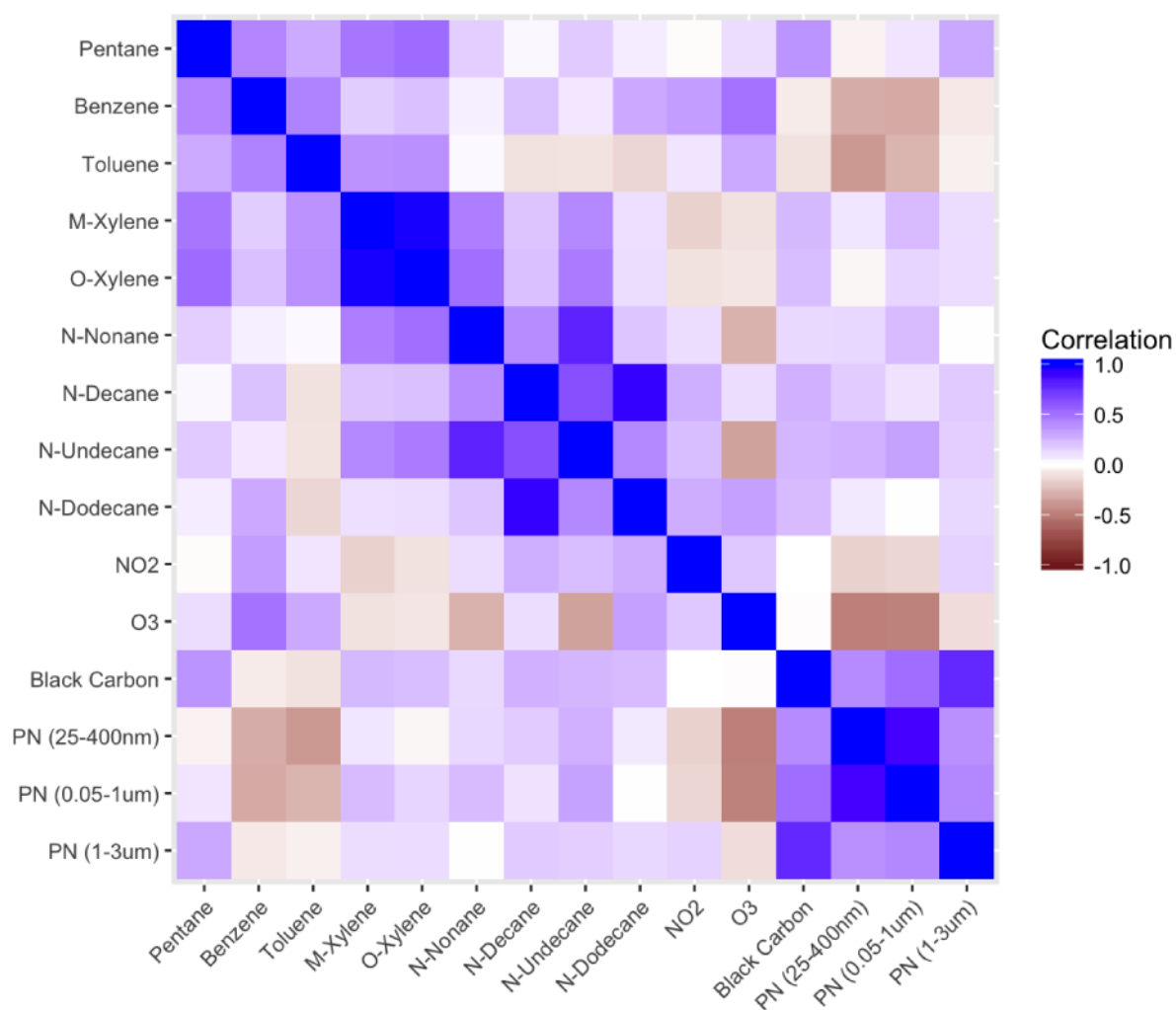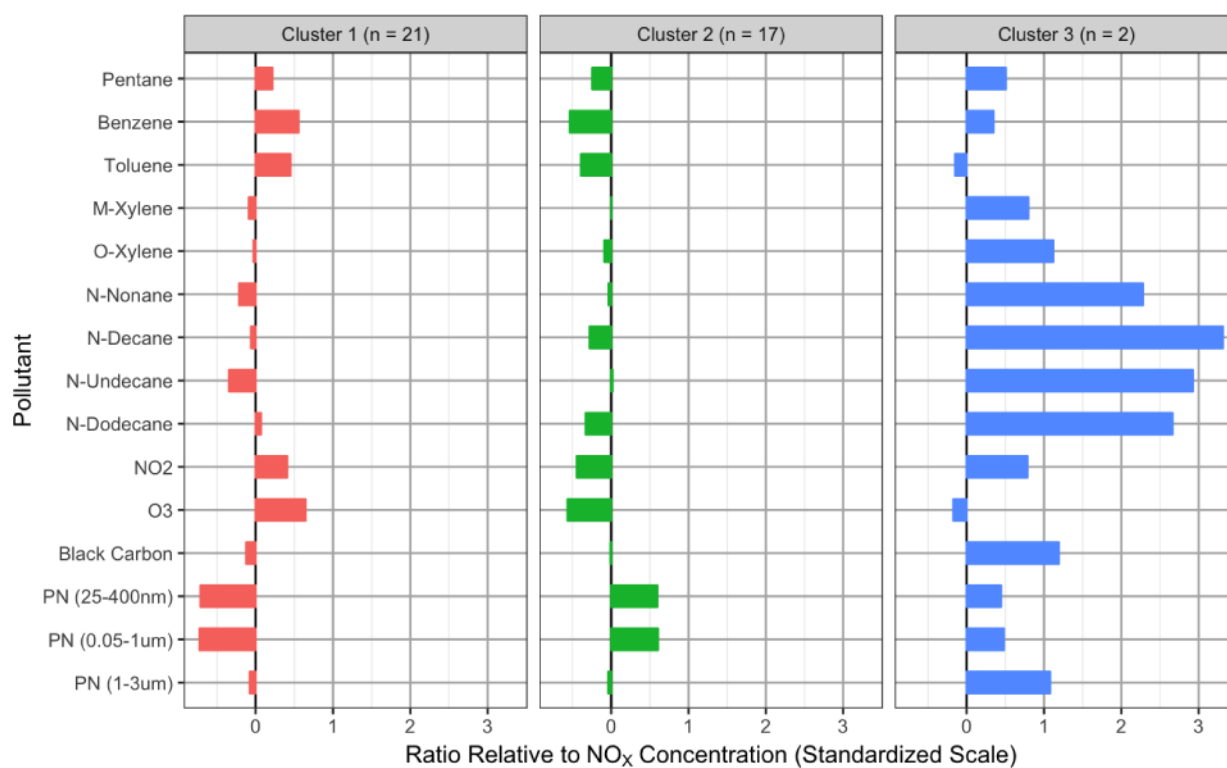
Figure 1.
Monitoring locations, colored by membership in cold season cluster.
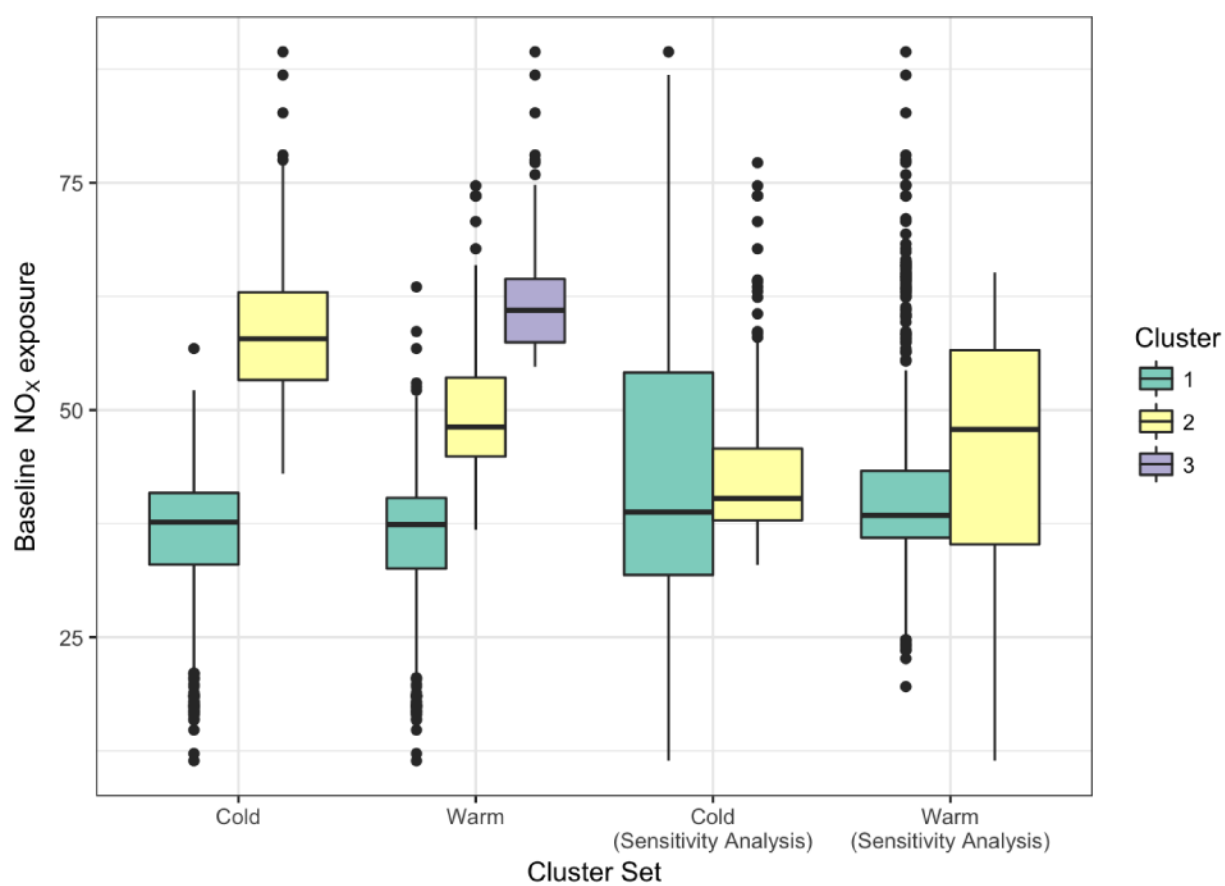
Figure 2.
Heatmap of the correlation between measurements in the cold season.

Figure 3.
Cold season cluster centers.

Figure 4.
Baseline NO$_X$ exposure by cluster membership.

Table 1

Baseline characteristics of the study population together and stratified by predicted membership in cold and warm season clusters.

Characteristics	Full Cohort	Cold Season		Warm Season		
		Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3
Participants (n)	1005	743	262	701	151	153
Age (years)	63 ± 10	63 ± 9.8	64 ± 10	63 ± 9.8	66 ± 11	62 ± 9.6
Male	471 (47)	357 (48)	114 (44)	357 (48)	52 (39)	67 (49)
Ethnicity						
White	498 (50)	373 (50)	125 (48)	379 (54)	33 (22)	86 (56)
Black	507 (50)	370 (50)	137 (52)	322 (46)	118 (78)	67 (44)
Education						
Less than high school	112 (11)	57 (8)	55 (21)	48 (7)	38 (25)	26 (17)
High School	202 (20)	145 (20)	57 (22)	137 (20)	41 (27)	24 (16)
Some college or technical	303 (30)	240 (32)	63 (24)	226 (32)	40 (26)	37 (24)
University or graduate	388 (39)	301 (41)	87 (33)	290 (41)	32 (21)	66 (43)
Smoking Status						
Never	432 (43)	340 (46)	92 (35)	321 (46)	64 (42)	47 (31)
Former	423 (42)	308 (41)	115 (44)	295 (42)	55 (36)	73 (48)
Current	150 (15)	95 (13)	55 (21)	85 (12)	32 (21)	33 (22)
Other socioeconomic						
Income ($1,000s)	53 ± 35	58 ± 34	39 ± 34	59 ± 34	30 ± 20	46 ± 39
Normalized SES Index[a]	0.16 ± 0.73	−0.29 ± 0.90	0.83 ± 0.78	−0.36 ± 0.87	1.09 ± 0.42	0.59 ± 0.91
General Health						
BMI	29 ± 5.8	29 ± 5.7	30 ± 6.0	29 ± 5.7	30 ± 6.6	30 ± 5.6
Cholesterol (mmol/L)	192 ± 36	193 ± 35	192 ± 37	193 ± 35	195 ± 38	188 ± 35
HDL (mmol/L)	51 ± 15	52 ± 15	50 ± 16	51 ± 14	53 ± 19	50 ± 16
Triglycerides (mmol/L)	118 ± 72	117 ± 72	122 ± 72	119 ± 74	111 ± 63	122 ± 73
Agatston Score	186 ± 462	170 ± 455	232 ± 478	176 ± 463	218 ± 442	201 ± 476
Pollutant Exposure						
NO_X (ppb)	42 ± 12	36 ± 7.4	58 ± 7.2	36 ± 7.6	50 ± 8.0	62 ± 5.8
$PM_{2.5}$ (µg/m^3)	15.9 ± 0.80	15.6 ± 0.52	16.8 ± 0.72	15.5 ± 0.51	16.2 ± 0.42	17.2 ± 0.59

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Values given as mean± standard deviation or n (%). BMI: Body-Mass Index; HDL: High-density lipoprotein; SES: Socioeconomic status.

[a]Unitless scale. Lower values indicate greater affluence.

Table 2

Pollutants measured on mobile and fixed monitor platforms.

Platform	Pollutant
Stationary Badges	Ozone (O_3)
	Nitrogen dioxide (NO_2)
	Oxides of Nitrogen (NO_X)
	Pentanes
	Benzene
	Toluene
	m-Xylene
	o-Xylene
	Nonane
	Decane
	Undecane
	Dodecane[a]
Mobile	Carbon monoxide (CO)[b]
	Black Carbon (BC)
	Particle Number Concentrations[c]: PN_{UF}: 25–400 nm PN_1: 0.02–1μm PN_{1-3}: 1–3μm

[a]Cold season only.

[b]Warm season only.

[c]Each size range measured by different instrument.

Table 3

Descriptive summary of clusters.

Season	Clustering Covariates	Cluster	# of Monitors	# of Subjects[a]	Predominant Location(s)	Distinctive High Ratio[b] Features	Distinctive Low Ratio[b] Features
Cold	Original	1	21	743	Northern and western suburban areas	NO_2, O_3, some VOCs	PN_{UF}, PN_1
		2	17	262	Downtown area	PN_{UF}	NO_2, O_3, some VOCs
		3	2	0	Singletons	VOCs, BC, PN_{UF}, PN_1, PN_{1-3}	
	Decorrelated with NO_X	1	19	664	Downtown area, with some suburban sites	PN_{UF}	VOCs, NO_2, O_3, BC, PN_{1-3}
		2	17	341	Western and northern suburbs	Some VOCs	Particle Counts
		3	4	0	Singletons	VOCs and particles	
Warm	Original	1	19	701	Western suburban area	NO_2, O_3, PM_{1-3}	VOCs
		2	12	151	Between downtown and suburbs	VOCs, CO	PN_{UF}, PN_1
		3	9	153	Downtown area	PN_{UF}, PN_1	NO_2, O_3, PM_{1-3}
	Decorrelated with NO_X	1	26	722	Downtown area and western suburbs	NO_2, PN_{UF}, PN_1, PN_{1-3}	VOCs, CO
		2	14	283	Areas just north and west of downtown	VOCs, CO	NO_2, PN_{UF}, PN_1, PN_{1-3}

[a]Counts are numbers of subjects predicted to belong to the cluster at baseline.

[b]Ratios are relative to NO_X concentration.

Table 4

Cluster-specific estimates of the association between CAC progression, in Agatston units per year, and differences of 5 μg/m^3 PM$_{2.5}$.

Source of Clusters	Cluster Name	Estimate (95% Confidence Interval)	Effect Modification *p*-value[a]
None	City-wide	23.0 (14.2, 31.7)	
Cold Season Data[b]	Cluster 1	17.0 (7.2, 26.7)	0.006
	Cluster 2	42.6 (25.7, 59.4)	
Warm Season Data	Cluster 1	17.1 (7.1, 27.0)	0.10
	Cluster 2	24.7 (–0.1, 49.6)	
	Cluster 3	39.8 (20.0, 59.6)	
Baseline PM	Median	13.7 (2.1, 25.4)	0.013
	> Median	33.1 (21.1, 44.9))	

[a]*p*-values are from a likelihood ratio test comparing against the model without cluster-specific progression estimates.

[b]No subjects were predicted to belong to Cluster 3 in the Cold Season.

Table 5

Cluster-specific estimates of the association between CAC progression, in Agatston units per year, and differences of 40 ppb NO_X.

Source of Clusters	Cluster Name	Estimate (95% Confidence Interval)	Effect Modification *p*-value[a]
None	City-wide	28.2 (17.1, 39.3)	
Cold Season Data	Cluster 1	22.2 (7.7, 36.7)	0.076
	Cluster 2	41.9 (23.7, 60.2)	
Warm Season Data	Cluster 1	20.7 (5.9, 35.6)	0.38
	Cluster 2	25.7 (−2.0, 53.4)	
	Cluster 3	38.1 (16.8, 59.5)	
Baseline $PM_{2.5}$	Median	25.1 (5.1, 45.1)	0.64
	> Median	30.6 (17.6, 43.6)	

[a]
p-values are from a likelihood ratio test comparing against the model without cluster-specific progression estimates.