Associations between Long-Term Exposure to Chemical Constituents of Fine Particulate Matter (PM$_{2.5}$) and Mortality in Medicare Enrollees in the Eastern United States

Yeonseung Chung,[1] Francesca Dominici,[2] Yun Wang,[2] Brent A. Coull,[2] and Michelle L. Bell[3]

[1]Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea; [2] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA; [3]School of Forestry and Environmental Studies, Yale University, New Haven, Connecticut, USA

BACKGROUND: Several epidemiological studies have reported that long-term exposure to fine particulate matter (PM$_{2.5}$) is associated with higher mortality. Evidence regarding contributions of PM$_{2.5}$ constituents is inconclusive.

OBJECTIVES: We assembled a data set of 12.5 million Medicare enrollees ($\geq$ 65 years of age) to determine which PM$_{2.5}$ constituents are *a)* associated with mortality controlling for previous-year PM$_{2.5}$ total mass (main effect); and *b)* elevated in locations exhibiting stronger associations between previous-year PM$_{2.5}$ and mortality (effect modification).

METHODS: For 518 PM$_{2.5}$ monitoring locations (eastern United States, 2000–2006), we calculated monthly mortality rates, monthly long-term (previous 1-year average) PM$_{2.5}$, and 7-year averages (2000–2006) of major PM$_{2.5}$ constituents [elemental carbon (EC), organic carbon matter (OCM), sulfate (SO$_4^{2-}$), silicon (Si), nitrate (NO$_3^-$), and sodium (Na)] and community-level variables. We applied a Bayesian hierarchical model to estimate location-specific mortality rates associated with previous-year PM$_{2.5}$ (model level 1) and identify constituents that contributed to the spatial variability of mortality, and constituents that modified associations between previous-year PM$_{2.5}$ and mortality (model level 2), controlling for community-level confounders.

RESULTS: One–standard deviation (SD) increases in 7-year average EC, Si, and NO$_3^-$ concentrations were associated with 1.3% [95% posterior interval (PI): 0.3, 2.2], 1.4% (95% PI: 0.6, 2.4), and 1.2% (95% PI: 0.4, 2.1) increases in monthly mortality, controlling for previous-year PM$_{2.5}$. Associations between previous-year PM$_{2.5}$ and mortality were stronger in combination with 1-SD increases in SO$_4^{2-}$ and Na.

CONCLUSIONS: Long-term exposures to PM$_{2.5}$ and several constituents were associated with mortality in the elderly population of the eastern United States. Moreover, some constituents increased the association between long-term exposure to PM$_{2.5}$ and mortality. These results provide new evidence that chemical composition can partly explain the differential toxicity of PM$_{2.5}$.

CITATION: Chung Y, Dominici F, Wang Y, Coull BA, Bell ML. 2015. Associations between long-term exposure to chemical constituents of fine particulate matter (PM$_{2.5}$) and mortality in Medicare enrollees in the eastern United States. Environ Health Perspect 123:467–474; http://dx.doi.org/10.1289/ehp.1307549

Introduction

Regulatory control of particulate matter (PM) could be dramatically improved with robust quantification of the evidence regarding the toxicity of various constituents of the PM mixture and of their sources [U.S. Environmental Protection Agency (EPA) 2004]. Currently, PM is regulated based on the total mass concentration without regard to its chemical composition, but scientific evidence on which types of constituents are most harmful could result in more effective regulations. However, the knowledge regarding differential toxicities of PM constituents has been identified as a crucial research gap (National Research Council 2004).

Responding to the need for such evidence, for the last decade, data have been accumulated for the constituents of fine particulate matter ($\leq$ 2.5 µm in aerodynamic diameter; PM$_{2.5}$) nationwide in the United States and have provided opportunities for studying the association between morbidity/mortality risk and PM$_{2.5}$ constituents. Using

these data, numerous epidemiological studies have reported evidence of health effects associated with PM$_{2.5}$ constituents, both in short-term (a few days previous) and long-term (a few years previous) exposure time frames. Studies focusing on the short-term health effects include those by Bell et al. (2014), Cao et al. (2012), Ito et al. (2011), Kim et al. (2012), Levy et al. (2012), Ostro et al. (2009), Peng et al. (2009), and Zhou et al. (2011). Fewer studies have investigated the long-term health effects of different PM$_{2.5}$ constituents, including those by Dockery et al. (1993), Ostro et al. (2010), and Pope et al. (1995, 2002). However, studies have not reported consistent findings regarding associations with specific constituents. Such discrepancies may derive from different aspects of the study design (e.g., population, confounding control, time frame, and statistical analysis), and the U.S. EPA (2009) has called for further research.

To fill this research gap, we investigated the differential toxicity of long-term

PM$_{2.5}$ exposure according to its chemical composition, based on a large-scale national database including approximately 12.5 million Medicare enrollees ($\geq$ 65 years of age). Combining several sources of data, we constructed a monthly multi-site time-series data set for 518 PM$_{2.5}$ monitoring locations in the eastern region of the United States during 2000–2006. The data include, for each monitoring location, monthly mortality rates, monthly values of the average PM$_{2.5}$ concentration over the previous 12 months, 7-year average concentrations of PM$_{2.5}$ constituents, and community-level confounding variables on socioeconomic status (SES) and racial composition. Using a Bayesian hierarchical (BH) regression model, we estimated spatially varying (SV) mortality rates associated with previous-year PM$_{2.5}$ and identified chemical constituents that explained the spatial variability of the mortality rates, controlling for PM$_{2.5}$ and community-level characteristics.

Methods

Data description. **PM$_{2.5}$ total mass.** We obtained daily (24-hr average) concentrations of PM$_{2.5}$ at 518 monitors in the eastern United States (Figure 1) for 2000–2006 from the U.S. EPA Air Quality System (AQS) database (U.S. EPA 2014). Using the daily PM$_{2.5}$ data, we calculated monthly long-term exposure to PM$_{2.5}$ as described in detail by

Greven et al. (2011). In brief, for the first day of every month and at each of the 518 monitor locations, we calculated the previous 1-year average of daily $PM_{2.5}$ concentrations (x_{ij}) for ith monitor at jth month, with $i = 1,\ldots, n$ and $j = 1,\ldots, n_i$. Because not all monitors had valid measurements for the entire study period, the number of monthly $PM_{2.5}$ values at a given monitoring location (n_i) varied from 33 to 70.

$PM_{2.5}$ chemical constituents. Although the U.S. EPA measures > 50 $PM_{2.5}$ chemical constituents, we focused on 6 identified in previous research as contributing substantially to $PM_{2.5}$ total mass (Bell et al. 2007): elemental carbon (EC), organic carbon matter (OCM), sulfate (SO_4^{2-}), silicon (Si), nitrate (NO_3^-), and sodium (Na). We obtained 7-year averages of the 6 chemical constituents at 174 monitors in the eastern region for 2000–2006 from the U.S. EPA AQS database, as described in detail elsewhere (Bell et al. 2007).

One challenge in our study is that the $PM_{2.5}$ constituents are measured at monitors (174 monitors) that are not collocated with the $PM_{2.5}$ monitors (518 monitors) (Figure 1). We assumed that the levels of constituents are spatially homogenous within a 6-mile radius, and therefore linked $PM_{2.5}$ monitors to $PM_{2.5}$ constituent monitors within 6 miles. Of the 518 $PM_{2.5}$ monitors, 241 had $PM_{2.5}$

constituent monitors within 6 miles, and we assigned 7-year averages of each $PM_{2.5}$ constituent of the closest constituent monitor to each of 241 $PM_{2.5}$ monitors. For the remaining 277 monitoring locations, we treated the levels of $PM_{2.5}$ chemical constituents as missing and applied a statistical approach to impute the missing data, as described in "Statistical methods." For either measured or imputed values, we let $z_i = (z_{i1},\ldots z_{i6})'$ denote the 7-year average concentrations of the six chemical constituents for ith $PM_{2.5}$ monitor location.

Mortality count and total number of people at risk. Mortality counts and the total number of people at risk were obtained at ZIP-code level from billing claims of Medicare enrollees who are fee-for-service Medicare beneficiaries (≥ 65 years of age) (Greven et al. 2011). For each of the 518 $PM_{2.5}$ monitor locations, we calculated monthly numbers of deaths and people at risk among the Medicare enrollees residing in each ZIP code with a centroid < 6 miles from a $PM_{2.5}$ monitor location. Depending on the location, 6-mile buffers around the monitors included the centroids of at least 3 and up to 20 different ZIP codes, and the data were aggregated over 3–20 ZIP codes. We let Y_{ij} and N_{ij} denote the number of deaths and the people at risk for ith monitor

location at jth month. For the whole study period across all 518 locations, the total size of the study population was 12.5 million enrollees, with the total number of deaths equal to 2.2 million approximately residing in 4,974 ZIP codes. For the 241 locations with $PM_{2.5}$ constituent data available, 1.2 million deaths occurred among 7.5 million enrollees approximately living in 3,425 ZIP codes.

Community-level confounders. We obtained ZIP code–level data on community-level confounding variables including SES and racial composition from the U.S. Census 2000 (U.S. Census Bureau 2000). We averaged values over all ZIP codes with centroids within 6 miles of each $PM_{2.5}$ monitor and assigned the averaged value to each monitor. We let $w_i = (w_{i1},\ldots w_{i5})'$ denote the five community-level confounders: median family income, proportion of people with high-school diploma or equivalent, proportion of residents in urban environment, proportion of white residents, and proportion of black residents.

Statistical methods. We analyzed the linked data using a BH Poisson regression model. The first level, a Poisson regression model with SV random effects, was used to estimate the association between month-to-month variation in mortality rate and month-to-month variation in long-term (previous 1-year average) $PM_{2.5}$:

$$Y_{ij} \sim \text{Poisson}(\lambda_{ij}),\ i = 1,\ldots, n \text{ and } j = 1,\ldots, n_i,$$
$$\text{Log}(\lambda_{ij}) = \log(N_{ij}) + \alpha_{i0} + \alpha_{i1}x_{ij}^*, \quad [1]$$

where Y_{ij} and N_{ij} are the number of deaths and the size of the population at risk for the ith monitoring location and jth month, x_{ij}^* is the previous-year average $PM_{2.5}$ centered around the location-specific average (i.e., $x_{ij}^* = x_{ij} - \bar{x}_i$), and α_{i0} and α_{i1} are the location-specific (SV) random intercepts and slopes. The parameter α_{i0} represents the SV baseline mortality rate when the previous-year average $PM_{2.5}$ is equal to its location-specific average (i.e., $x_{ij}^* = 0$). The parameter α_{i1} represents the SV association between month-to-month variation in mortality rate and month-to-month variation in previous-year $PM_{2.5}$.

The second level of the BH model regresses the location-specific 7-year averages of $PM_{2.5}$ constituents and community-level confounders on the SV intercept and slope, α_{i0} and α_{i1}:

$$\alpha_{i0} = \beta_0 + \Sigma_{k=1}^6 \beta_k z_{ik}^* + \Sigma_{l=1}^5 \beta_{6+l} w_{il}^*$$
$$+ \beta_{12}\bar{x}_i + \varepsilon_{i0} \quad [2]$$

$$\alpha_{i1} = \gamma_0 + \Sigma_{k=1}^6 \gamma_k z_{ik}^* + \Sigma_{l=1}^5 \gamma_{6+l} w_{il}^*$$
$$+ \varepsilon_{i1}, \quad [3]$$

where z_{ik}^* is the level of the kth chemical constituent and w_{il}^* is lth community-level confounder at ith location, and ε_{i0} and ε_{i1} are
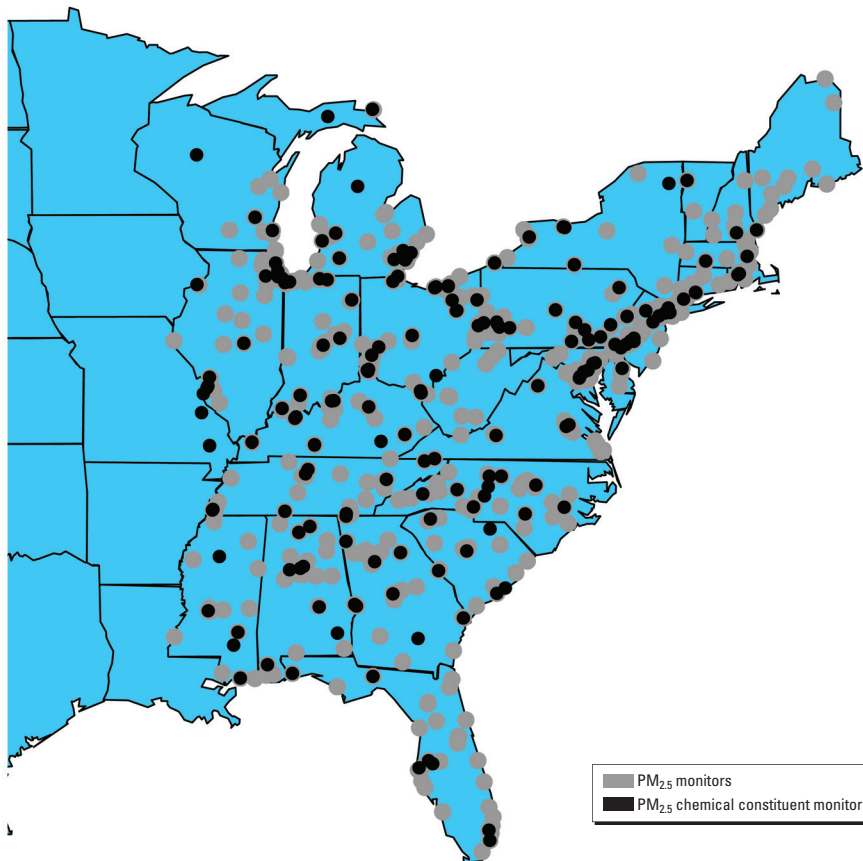

Figure 1. Map of 518 $PM_{2.5}$ monitor locations and 174 $PM_{2.5}$ chemical constituent monitor locations in the eastern region of the United States.

Legend:
- PM_{2.5} monitors
- PM_{2.5} chemical constituent monitor

random errors. We centered and scaled all explanatory variables to simplify interpretation and reduce multicollinearity. Note that $\bar{x}_i$ is included in the SV intercept in Equation 2 to control for total PM$_{2.5}$ concentration when estimating the effects of constituents on the SV mortality rate (Mostofsky et al. 2012).

To account for potential residual spatial correlation in the second level, we assumed the error terms could be spatially correlated using a standard approach (Gelfand et al. 2003) (see Supplemental Material, "Accounting for residual spatial correlation"). We fit our BH model using a Monte Carlo Markov chain (MCMC) method [see Supplemental Material, "Two-stage estimation and the Markov Chain Monte Carlo (MCMC) algorithm"]. All computations were conducted using R statistical software (R Core Team 2013).

There are four sets of parameters of interest. From Equation 1, we obtained a) the SV (i.e., monitor-specific) baseline mortality rates when the previous-year PM$_{2.5}$ was equal to its monitor-specific overall average (SV intercepts α_{i0} for each location i, expressed as deaths/month/1,000 persons); and b) the SV association between month-to-month variation in mortality rate and month-to-month variation in previous-year PM$_{2.5}$ (SV slopes α_{i1} for each location i, expressed as the percentage increase in the mortality rate associated with a 1-µg/m^3 increase in previous-year PM$_{2.5}$). From Equations 2 and 3, we obtained c) the association between the SV intercepts and the monitor-specific 7-year averages of PM$_{2.5}$ constituents, adjusted by community-level confounders and previous-year PM$_{2.5}$ (the β_k coefficients from Equation 2, expressed as the percentage increase in the mortality rate associated with a 1-SD increase in the 7-year average concentration of each constituent), and d) the association between the SV slopes and the monitor-specific 7-year averages of PM$_{2.5}$ constituents, adjusted by community-level confounders (the γ_k coefficients from Equation 3, expressed as the percentage increase in the mortality rate ratio for previous year PM$_{2.5}$ associated with a 1-SD increase in the 7-year average concentration of each constituent).

To find the best fit for Equation 2, we conducted an extensive sensitivity analysis. We considered the following eight models: no explanatory variable, constituents only, community-level confounders only, or both constituents and community-level confounders as explanatory variables, all with and without spatially correlated errors. Among the eight options, we chose the best fit based on the Deviance Information Criteria (DIC) (Spiegelhalter et al. 2002).

There were 277 PM$_{2.5}$ monitoring locations with missing values for the constituents. Separately for each constituent, we fit a Bayesian spatial Gaussian process (GP) model based on the observed data (i.e., 241 locations) and estimated a spatial correlation using the spBayes R package (Finley et al. 2007) and imputed the missing values based on the posterior predictive sample means for the 277 PM$_{2.5}$ monitors (see Supplemental Material, "Bayesian spatial Gaussian process (GP) for missing imputation"). Before using the imputed constituent levels in the analysis, we confirmed that the Bayesian spatial GP modeling was appropriate for imputation via a cross-validation (CV) study (see Supplemental Material, "Cross validation study").

We conducted the analysis for the complete-case data ($n = 241$ monitoring locations with the data available for both PM$_{2.5}$ total mass and the chemical constituents) and for the all-sites data ($n = 518$ monitoring locations using imputed values for the 277 locations without measurements for PM$_{2.5}$ chemical constituents). Also, we analyzed the data for the entire elderly population (≥ 65 years) and stratified by two age groups (65–74 vs. ≥ 75 years).

Results

Table 1 reports summary statistics for each variable for the complete-case data ($n = 241$) and for the all-sites data ($n = 518$). Figure 2A displays maps of 7-year averages of PM$_{2.5}$ exposure levels (micrograms per cubic meter) and Figure 2B presents maps of 7-year averages of mortality rates (deaths/month/1,000 persons) for 518 monitoring locations. Figure 3 shows maps of 7-year averages of each chemical constituent (micrograms per cubic meter) for the 241 locations with available data. SO$_4^{2-}$ and NO$_3^-$ levels seem to exhibit strong spatial correlations; OCM, Si, and Na levels moderate spatial correlations; and EC levels weak spatial correlations, with high values only in a few locations. Estimated spatial correlations obtained from the spatial GP model between pairs of monitors with a distance of about 40 miles are 0.05, 0.20, 0.21, 0.19, 0.21, and 0.20 for EC, OCM, SO$_4^{2-}$, Si, NO$_3^-$, and Na, respectively. All five community-level confounders are also spatially mapped (see Supplemental Material, Figure S1) over the 518 locations.

The correlations among the 7-year averages of PM$_{2.5}$ total mass, the PM$_{2.5}$ chemical constituents, and community-level confounders are summarized in Supplemental Material, Table S1, for the complete-case data ($n = 241$). We observed that PM$_{2.5}$ is correlated positively with OCM (0.43), SO$_4^{2-}$ (0.61), and the proportion of black residents (0.32) and inversely with Na (–0.41). The highest positive correlations among the constituents were observed between EC and OCM (0.44), SO$_4^{2-}$ and OCM (0.41), and Si and OCM (0.43). Among the community-level confounders, strong positive correlations were observed as 0.62 between median family income and the proportion of people with high school diploma or equivalent, and 0.50 between the proportions of white residents and high school graduates, whereas the strongest negative correlation was –0.84 between the proportions of white and black residents. Between the constituents and the community-level confounders, the highest correlations were observed for OCM at 0.39 and for Si at 0.34 with the proportion of black residents.

Before the BH regression modeling, we conducted CV studies for our imputation method for the missing constituent levels. The sample correlation coefficients between the observed and predicted values for the test data are 0.64–0.94 for all constituents averaged over 5 CV data sets (see Supplemental

Table 1. Summary statistics for each variable (mean ± SD).

Variable	Complete-case data ($n = 241$)	All-sites data ($n = 518$)
Population size at risk (n/month)	16901.00 ± 17307.31	14538.92 ± 15453.71
Mortality count (n/month)	76.01 ± 74.04	64.93 ± 66.7
Long-term PM$_{2.5}$ exposure level (µg/m^3)	14.56 ± 1.88	13.7 ± 2.13
PM$_{2.5}$ chemical constituents (µg/m^3)		
Elemental carbon (EC)	0.71 ± 0.33	0.68 ± 0.24[a]
Organic carbon matter (OCM)	4.1 ± 1.06	4.05 ± 0.90[a]
Sulfate (SO$_4^{2-}$)	4.22 ± 0.81	4.14 ± 0.80[a]
Silicon (Si)	0.09 ± 0.03	0.09 ± 0.03[a]
Nitrate (NO$_3^-$)	1.86 ± 0.86	1.68 ± 0.85[a]
Sodium (Na)	0.16 ± 0.08	0.17 ± 0.07[a]
Community-level confounders		
Family income ($)	38247.39 ± 11149.87	40305.00 ± 12461.21
Percent high school graduate	0.78 ± 0.06	0.79 ± 0.08
Percent urban	0.93 ± 0.16	0.86 ± 0.23
Percent white	0.64 ± 0.20	0.70 ± 0.20
Percent black	0.25 ± 0.18	0.19 ± 0.17

For population size at risk, mortality count and long-term (previous 1-year average) PM$_{2.5}$, location-specific monthly values are averaged across locations for the whole study period (2000–2006). For chemical constituents, location-specific 7-year averages are averaged across locations. For community-level confounders, location-specific values are averaged across locations.
[a]Numbers were calculated including the imputed PM$_{2.5}$ constituent levels for the 277 PM$_{2.5}$ monitoring locations with missing constituent levels.

Material, Table S2). The root mean square error (RMSE) for prediction for each constituent and the average RMSE over five CV data sets is about half of the sample standard deviation for all constituents (see Supplemental Material, Table S3). Scatter plots for the observed versus predicted data show that the points generally follow the reference line (meaning observed values = predicted values) (see Supplemental Material, Figure S2). Based on the CV study results, we concluded that the Bayesian spatial GP method was appropriate for imputing the missing constituents in our study.

We analyzed the complete-case data (n = 241) and the all-sites data (n = 518), separately. We fit Equations 1, 2, and 3 and Supplemental Material, Equation S1 with eight different options and chose the best fit based on the smallest DIC (see Supplemental Material, Table S4). We obtained the smallest DIC for the model including both chemical constituents and community-level confounders as explanatory variables and with spatially independent errors both for the SV intercept and slope model in complete-case data as well as in the all-sites data.

Figure 4 displays results from the first level of the BH model for complete-case data (left panels) and all-sites data (right panels), respectively. Both sets of data showed similar results. Figure 4A shows the estimated monthly mortality rate when the previous-year $PM_{2.5}$ is equal to its monitor-specific overall average ranges from 3.37 to 6.15 (deaths/month/1,000 persons) over the study region from all-sites data analysis. Also, Figure 4B shows that the estimated association of mortality rate with a 1-$\mu g/m^3$ increase in the previous-year $PM_{2.5}$ is from –1.0 to 4.6 (percent increase in mortality rate).

Figure 5 reports the results from the second level of the BH model for complete-case data (left-solid bars) and all-sites data (right-dashed bars), respectively. Results were similar between complete-case data and all-sites data, but the all-sites estimates were somewhat smaller and their confidence intervals are narrower. In both analyses, we observed that adjusting for the community-level confounders and $PM_{2.5}$ total mass, EC, Si, and NO_3^- were positively associated with mortality rate (the SV intercept, β_k), whereas SO_4^{2-} was inversely related to mortality (Figure 5A). Meanwhile, SV slope estimates (γ_k) indicated that the percentage increase in mortality rate with a 1-unit increase in average previous-year $PM_{2.5}$ was greater than expected when combined with a 1-SD increase in SO_4^{2-} and Na (Figure 5B).

Finally, we analyzed the data stratified by two age groups (65–74 vs. ≥ 75 years). For the SV intercept model (see Supplemental Material, Figure S3), results for the younger age group were similar to those for the all-age analysis (i.e., main effect estimates for EC, Si, and NO_3^- were significant), whereas for the SV slope model (see Supplemental Material, Figure S4), both age groups (65–74 and ≥ 75 years) had results similar to those for the all-age analysis (≥ 65 years).

Discussion

We investigated *a*) whether month-to-month changes in mortality rates were associated with month-to-month changes in the previous-year average exposure to $PM_{2.5}$, and *b*) whether 7-year average levels of $PM_{2.5}$ chemical constituents modified this association.

For the SV intercept, EC, Si, and NO_3^- were positively associated with mortality rates after adjusting for $PM_{2.5}$ total mass and the community-level confounders. For EC, our results are consistent with those of a previous cohort study of female public school professionals in California (Ostro et al. 2010).

Evidence for the toxicity of Si was also found in other studies. Ostro et al. (2010) reported that long-term exposure to Si was positively associated with pulmonary mortality. Si may serve as a surrogate for toxic constituents found in mineral dust associated with traffic (Reff et al. 2009). A previous study reported that long-term exposure to traffic or traffic-related pollutants is associated with cardiopulmonary mortality (Jerrett et al. 2005). Few studies examined associations for NO_3^- with mortality in a long-term framework. Ostro et al. (2010) reported that long-term exposure to NO_3^- was significant for mortality in a single-pollutant model but not in a multipollutant model. Additionally, SO_4^{2-} was found to be inversely associated with mortality rates, which is inconsistent with previous studies where positive associations were found (Dockery et al. 1993; Ostro et al. 2010; Pope et al. 1995, 2002). Although the observed positive associations were from single-pollutant approaches, our study used a multipollutant analysis that also included adjustment for the $PM_{2.5}$ total mass and community-level confounders. Therefore, the inverse associations that we observed should be interpreted with caution, because they may be an artifact of multicollinearity resulting from correlations between SO_4^{2-} and other constituents, $PM_{2.5}$ total mass, and community-level confounders.

SO_4^{2-} and Na were significant modifiers of monitor-specific associations between previous-year average $PM_{2.5}$ and mortality rates. Previous long-term exposure studies for $PM_{2.5}$ constituents have reported significant positive associations of SO_4^{2-} with all-cause mortality (Dockery et al. 1993; Pope et al. 2002) or with cardiopulmonary mortality (Ostro et al. 2010; Pope et al. 2002). Na was also a significant modifier that strengthened the association between
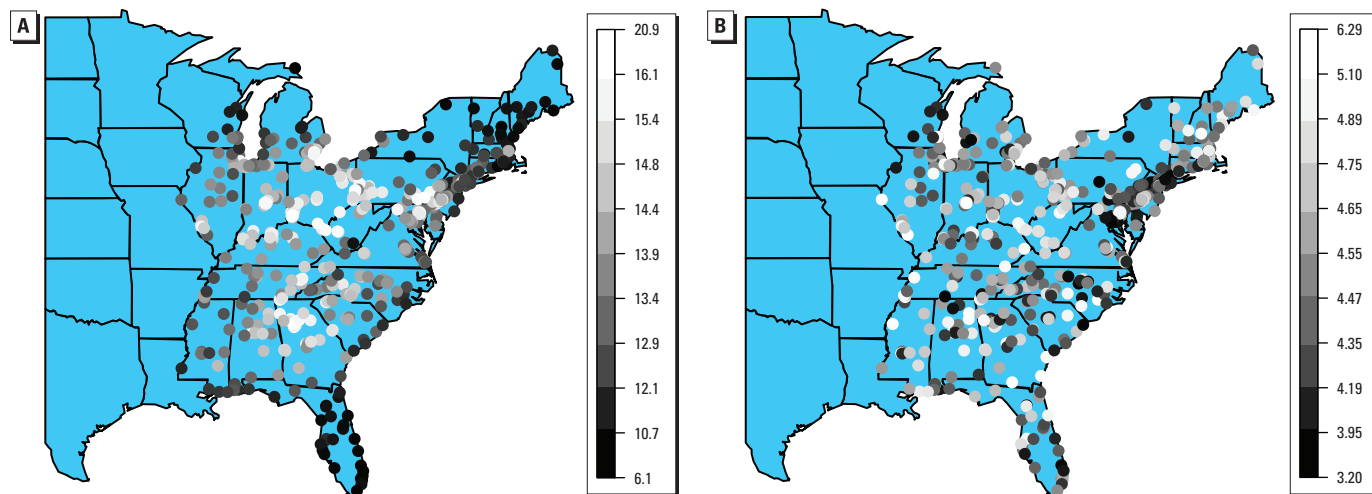

Figure 2. (*A*) Map of 7-year (2000–2006) averages of monthly long-term (previous 1-year average) $PM_{2.5}$ exposure ($\mu g/m^3$) for all $PM_{2.5}$ monitor locations (n = 518). (*B*) Map of 7-year (2000–2006) averages of monthly mortality rate (deaths/month/1,000 persons) for all $PM_{2.5}$ monitor locations (n = 518).

long-term PM$_{2.5}$ and mortality. Few studies have estimated associations between Na and health outcomes, with some showing evidence of associations with mortality (Krall et al. 2013) or hospital admission (Zanobetti et al. 2009).

Several possible mechanisms have been proposed in human subject studies linking constituents to biomarkers: systemic inflammation and oxidative stress associated with EC (Neophytou et al. 2013), altered DNA methylation related to Si (Hou et al. 2014),
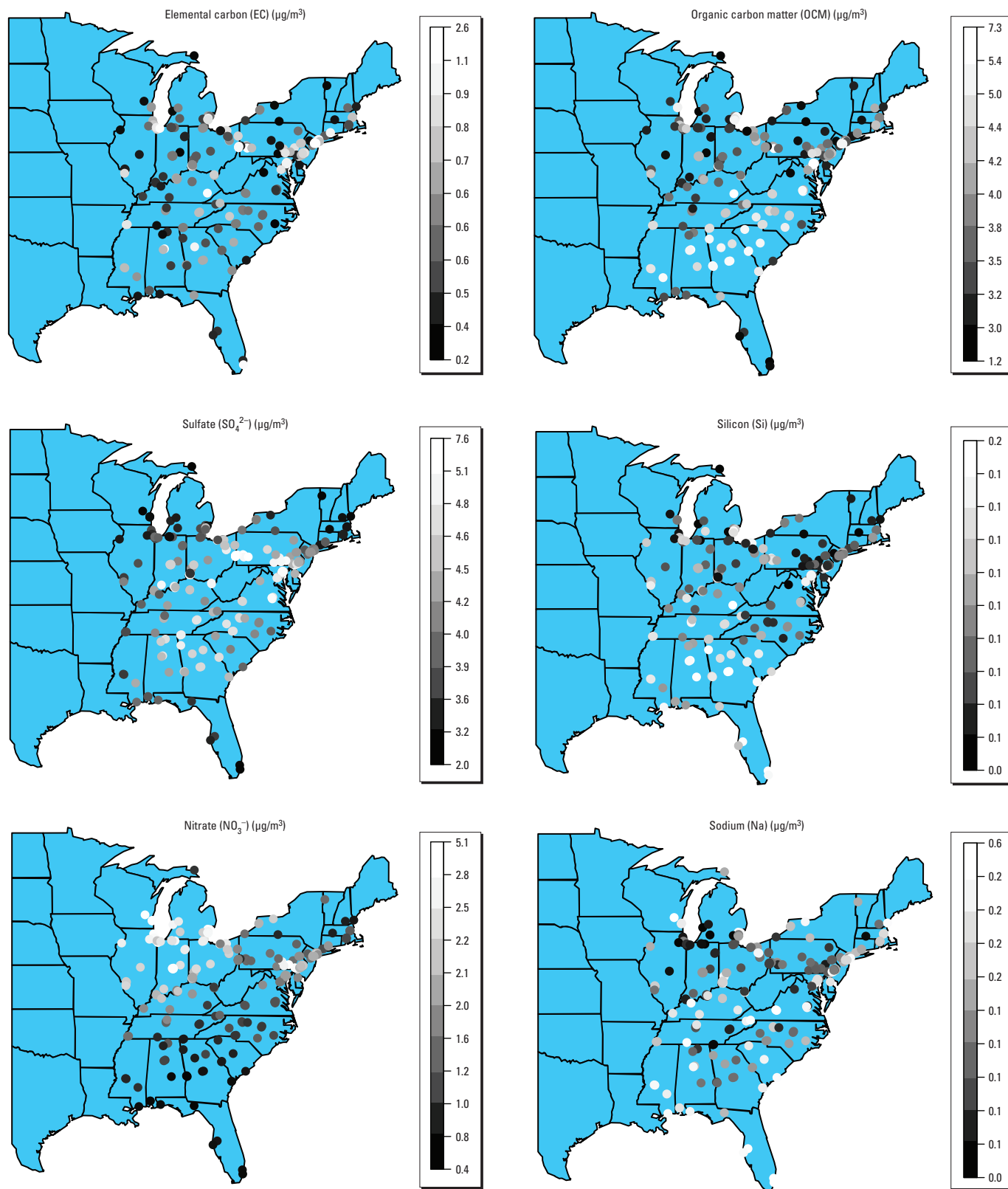

Figure 3. Maps of 7-year (2000–2006) averages of each of the six PM$_{2.5}$ chemical constituents (µg/m³) for 241 monitor locations with available data.

and inflammation related to NO_3^- and SO_4^{2-} (Wu et al. 2012). However, because of limited evidences from experimental/toxicological studies, the biological pathway through which short-term exposure to $PM_{2.5}$ and its components affects health is still an area of active investigation, and the mechanisms for long-term exposure are less understood.

The U.S. EPA measures > 50 different chemical constituents. Analyzing all available constituents would present problems of multiple comparisons. We selected the six constituents that were previously shown to be the largest contributors to $PM_{2.5}$ total mass and/or co-vary with $PM_{2.5}$ total mass (Bell et al. 2007). Also, the reliability of a community-level average of $PM_{2.5}$ constituent exposure varies by constituent. For the six constituents investigated in the present study, the average correlation of monitors in close proximity (< 5 km) ranges from 0.60 to 0.93 and for larger distances (20–50 km) ranges from 0.46 to 0.88 (Bell et al. 2011). The spatial heterogeneity of many other constituents may be larger, limiting the interpretation of community-level exposures. However, we recognize that other constituents have also been found to be associated with human health. In particular, associations of health outcomes with $PM_{2.5}$ metal constituents that were not included in our analysis, such as aluminum, calcium, chromium, lead, manganese, nickel, titanium, vanadium, and zinc, have been reported in previous studies (Bell et al. 2014; Cavallari et al. 2008; Hsu et al. 2011; Lippmann et al. 2006; Wu et al. 2012).

One limitation of the available air pollution data is that monitors that measure $PM_{2.5}$ total mass and monitors that measure the $PM_{2.5}$ chemical constituents are misaligned (Figure 1). We addressed this limitation by assuming that ambient levels of $PM_{2.5}$ constituents were homogeneous within a 6-mile radius. This spatial homogeneity assumption for air pollutants builds on previous research (Bell et al. 2011) where 6 miles (about 10 km) in radius is a reasonable buffer size for the homogeneity assumption. In Bell et al. (2011), the estimated spatial correlations between pairs of monitors with distances of 5–10 km are 0.67, 0.85, 0.95, 0.62, 0.95, and 0.59 for EC, OCM, SO_4^{2-}, Si, NO_3^-, and Na, respectively. However, spatial variability varies by constituent—for example, with more heterogeneity for Si or Na than for SO_4^{2-} or NO_3^-—and different buffer sizes may be applied for different constituents when aligning various sources of data.

Another limitation for air pollution data is that monitors that measure $PM_{2.5}$ total mass are much denser than monitors
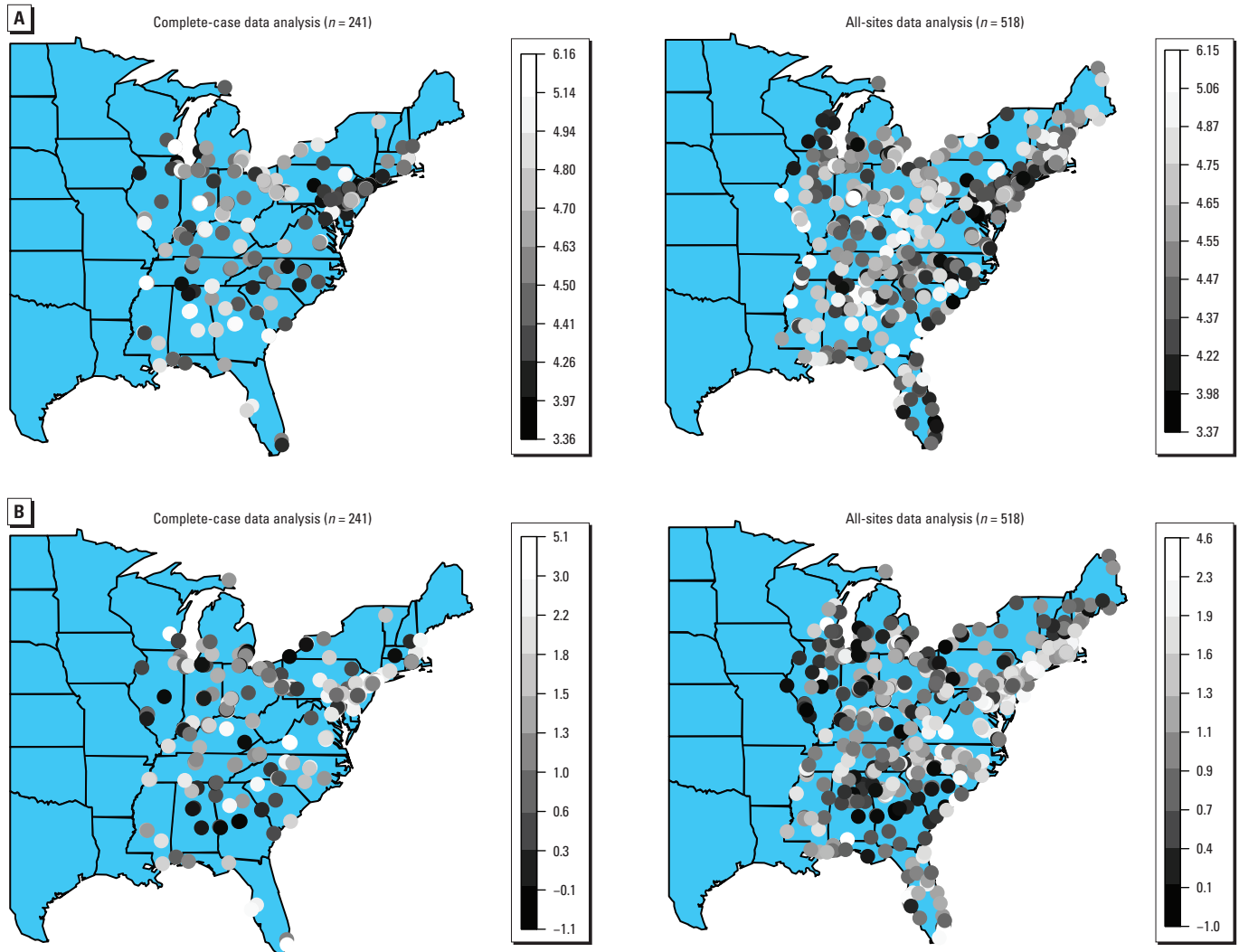

Figure 4. (*A*) Maps of the estimates (posterior means) of the SV intercept from the complete-case data analysis (*n* = 241, left) and the all-sites data analysis (*n* = 518, right). The values represent the monthly mortality rate (deaths/month/1,000 persons) when previous-year $PM_{2.5}$ is at location-specific average. (*B*) Maps of the estimates (posterior means) of the SV slope from the complete-case data analysis (*n* = 241, left) and the all-sites data analysis (*n* = 518, right). The values represent the percent increase in the monthly mortality rate associated with a 1-μg/m³ increase in previous-year $PM_{2.5}$.

that measure PM$_{2.5}$ chemical constituents (Figure 1). When we aligned the two kinds of monitors, missing data occurred for almost half of the PM$_{2.5}$ monitors. To avoid simply removing the observations with missing values and reducing the sample size to half, we adopted a Bayesian spatial GP modeling and conducted a single value imputation for the missing data separately for each constituent. We compared the results between the complete-case data analysis and the all-sites data analysis with the imputed values. Although imputation did not change our primary conclusions, results based on the imputed data should be interpreted with caution. Specifically, using a single-value imputation does not incorporate uncertainty for prediction, and measurement error can occur for the explanatory variables in regression modeling (Gryparis et al. 2009).

In our study, the PM data are the ambient levels, which we use to approximate the actual human exposure. The ambient level of a given pollutant is not a perfect surrogate of personal exposure to that pollutant, which can induce exposure measurement error into the analysis with variations in error by constituent. In a multipollutant analysis such as the present study, this type of error may induce upward bias in regression coefficient estimates, resulting in anticonservative inference on health effects. However, several authors have shown that this type of bias barely occurs in situations in which the amount of error or the correlations among pollutants in analysis are extremely large (Schwartz and Coull 2003). Therefore, it is unlikely that differences between ambient levels and personal exposures explain the observed associations in our study.

Our analysis is based on multi-site time-series data where long-term exposure was estimated by calculating previous 1-year average of daily exposure values at each temporal point (i.e., first day of each month). However, results may be sensitive to different choices of time frames. Kim et al. (2012) reported that different lag values should be selected for the short-term effects of PM$_{2.5}$ constituents depending on health outcomes. Shorter or longer time frames than a year could be considered for examining long-term health effects of PM$_{2.5}$.

Our study focused on the eastern region of the United States, and our findings may not be generalizable to other areas because the characteristics of PM mixtures and populations are quite different across the United States (Bell et al. 2007), and effect modification by the chemical composition of PM$_{2.5}$ may vary among regions. Also, we focused on the elderly population, which may be more susceptible to effects of exposure than other age groups. In our study, slight differences in results were found between two age groups (65–74 and ≥ 75 years).

To our knowledge, this is the first large-scale study (covering the eastern United States) to investigate the association between long-term exposure to PM$_{2.5}$ and mortality rate and effect modification by the chemical constituents of PM$_{2.5}$. Unlike previous studies of PM$_{2.5}$ constituents, we used a BH regression approach, where PM$_{2.5}$ constituents were modeled as potential modifiers of the main effect of PM$_{2.5}$ on health outcomes. Despite limitations, our findings add new evidence regarding the differential toxicity of PM$_{2.5}$ constituents and their potential influence on the long-term health effects of PM$_{2.5}$.

References

Bell ML, Dominici F, Ebisu K, Zeger SL, Samet JM. 2007. Spatial and temporal variation in PM$_{2.5}$ chemical composition in the United States for health effects studies. Environ Health Perspect 115:989–995; doi:10.1289/ehp.9621.

Bell ML, Ebisu K, Leaderer BP, Gent JF, Lee HJ, Koutrakis P, et al. 2014. Associations of PM$_{2.5}$ constituents and sources with hospital admissions: analysis of four counties in Connecticut and Massachusetts (USA) for persons ≥ 65 years of age. Environ Health Perspect 122:138–144; doi:10.1289/ehp.1306656.

Bell ML, Ebisu K, Peng RD. 2011. Community-level spatial heterogeneity of chemical constituent levels of fine particulates and implications for epidemiological research. J Expo Sci Environ Epidemiol 21:372–384.

Cao J, Xu H, Xu Q, Chen B, Kan H. 2012. Fine particulate matter constituents and cardiopulmonary mortality in a heavily polluted Chinese city. Environ Health Perspect 120:373–378; doi:10.1289/ehp.1103671.

Cavallari JM, Eisen EA, Fang SC, Schwartz J, Hauser R, Herrick RF, et al. 2008. PM$_{2.5}$ metal exposures and nocturnal heart rate variability: a panel study of boilermaker construction workers. Environ Health 7:36; doi:10.1186/1476-069X-7-36.

Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, et al. 1993. An association between air pollution and mortality in six U.S. cities. New Engl J Med 329(24):1753–1759.

Finley AO, Banerjee S, Carlin BP. 2007. spBayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. J Stat Softw 19(4):1–24.

Gelfand AE, Kim HJ, Sirmans CF, Banerjee S. 2003. Spatial modeling with spatially varying coefficient processes. J Am Stat Assoc 98:387–396.

Greven S, Dominici F, Zeger S. 2011. An approach to the estimation of chronic air pollution effects using spatio-temporal information. J Am Stat Assoc 106:396–406.

Gryparis A, Paciorek CJ, Zeka A, Schwartz J, Coull BA. 2009. Measurement error caused by spatial misalignment in environmental epidemiology. Biostatistics 10:258–274.

Hou L, Zhang X, Zheng Y, Wang S, Dou C, Guo L, et al. 2014. Altered methylation in tandem repeat element and elemental component levels in inhalable air particles. Environ Mol Mutagen 55:256–265.

Hsu SO, Ito K, Lippmann M. 2011. Effects of thoracic and fine PM and their components on heart rate and pulmonary function in COPD patients. J Exp Sci Environ Epidemiol 21:464–472.

Ito K, Mathes R, Ross Z, Nádas A, Thurston G, Matte T. 2011. Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. Environ Health Perspect 119:467–473; doi:10.1289/ehp.1002667.

Jerrett M, Burnett RT, Ma R, Pope CA III, Krewski D, Newbold KB, et al. 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16(6):727–736.

Kim SY, Peel JL, Hannigan MP, Dutton SJ, Sheppard L, Clark ML, et al. 2012. The temporal lag structure of short-term associations of fine particulate matter chemical constituents and cardiovascular and respiratory hospitalizations. Environ Health Perspect 120:1094–1099; doi:10.1289/ehp.1104721.

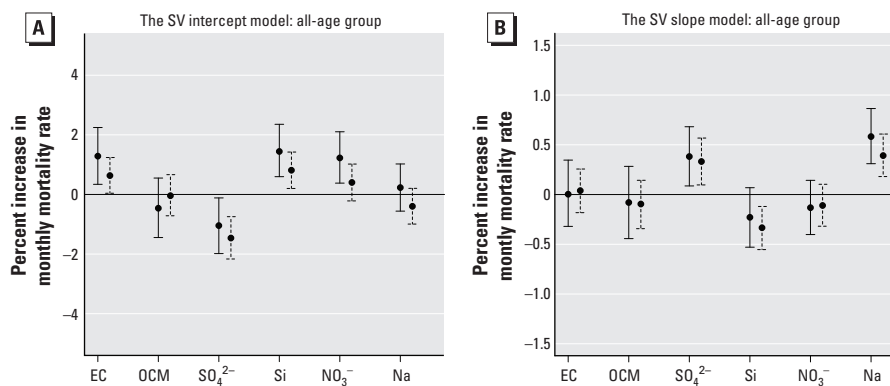Krall JR, Anderson B, Dominici F, Bell ML. Peng RD. 2013. Short-term exposure to particulate matter

Figure 5. (*A*) Posterior estimates with 95% posterior intervals for the β_k regression coefficients in the second-level SV intercept model. Solid error bars are for the complete-case data (*n* = 241) and dashed error bars are for the all-sites data (*n* = 518). Values correspond to the estimated percentage increase in monthly mortality rate associated with a 1-SD increase in each constituent, adjusted for previous-year average of PM$_{2.5}$ total mass and for community-level covariates. (*B*) Posterior estimates with 95% posterior intervals for the γ_k regression coefficients in the second-level SV slope model. Solid error bars are for the complete-case data (*n* = 241) and dashed error bars are for the all-sites data (*n* = 518). Values correspond to the estimated percentage increase in the association between previous-year average of PM$_{2.5}$ and mortality when combined with a 1-SD increase in each constituent, adjusted for community level covariates.

constituents and mortality in a national study of U.S. urban communities. Environ Health Perspect 121:1148–1153; doi:10.1289/ehp.1206185.

Levy JI, Diez D, Dou Y, Barr CD, Dominici F. 2012. A meta-analysis and multisite time-series analysis of the differential toxicity of major fine particulate matter constituents. Am J Epidemiol 175(11):1091–1099.

Lippmann M, Ito K, Hwang JS, Maciejczyk P, Chen LC. 2006. Cardiovascular effects of nickel in ambient air. Environ Health Perspect 114:1662–1669; doi:10.1289/ehp.9150.

Mostofsky E, Schwartz J, Coull BA, Koutrakis P, Wellenius GA, Suh HH, et al. 2012. Modeling an association between particle constituents of air pollution and health outcomes. Am J Epidemiol 176(4):317–326.

National Research Council. 2004. Research Priorities for Airborne Particulate Matter. IV. Continuing Research Progress. Washington, DC:National Academies Press.

Neophytou AM, Hart JE, Cavallari HM, Smith TJ, Dockery DW, Coull BA, et al. 2013. Traffic-related exposures and biomarkers of systemic inflammation, endothelial activation and oxidative stress: a panel study in the US trucking industry. Environ Health 12:105; doi:10.1186/1476-069X-12-105.

Ostro B, Lipsett M, Reynolds P, Goldberg D, Hertz A, Garcia C, et al. 2010. Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study. Environ Health Perspect 118:363–369; doi:10.1289/ehp.0901181.

Ostro B, Roth L, Malig B, Marty M. 2009. The effects of fine particle constituents on respiratory hospital admissions in children. Environ Health Perspect 117:475–480; doi:10.1289/ehp.11848.

Peng RD, Bell ML, Geyh AS, McDermott A, Zeger SL, Samet JM, et al. 2009. Emergency admissions of cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. Environ Health Perspect 117:957–963; doi:10.1289/ehp.0800185.

Pope CA III, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287(9):1132–1141.

Pope CA III, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, et al. 1995. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. Am J Respir Crit Care Med 151(3 pt 1):669–674.

R Core Team. 2013. R: A language and environment for statistical computing. Vienna, Austria:R Foundation for Statistical Computing. Available: http://www.R-project.org/ [accessed 10 October 2014].

Reff A, Bhave PV, Simon H, Pace TG, Pouliot GA, Mobley JD, et al. 2009. Emissions inventory of $PM_{2.5}$ trace elements across the United States. Environ Sci Technol 43:5790–5796.

Schwartz J, Coull BA. 2003. Control for confounding in the presence of measurement error in hierarchical models. Biostatistics 4:539–553.

Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A. 2002. Bayesian Measures of Model Complexity and Fit. J R Stat Soc Series B Stat Methodol 64(4):583–639.

U.S. Census Bureau. 2000. Census 2000 Gateway. Available: http://www.census.gov/main/www/cen2000.html [accessed 22 March 2013].

U.S. EPA (U.S. Environmental Protection Agency). 2004. Advisory on Plans for Health Effects Analysis in the Analytical Plan for EPA's Second Prospective Analysis—Benefits and Costs of the Clean Air Act, 1990–2020. EPA-SAB-COUNCIL-ADV-04-002. Washington, DC:EPA Science Advisory Board. Available: http://yosemite.epa.gov/sab%5CSABPRODUCT.NSF/08E1155AD24F871C85256E5400433D5D/$File/council_adv_04002.pdf [accessed 28 March 2014].

U.S. EPA (U.S. Environmental Protection Agency). 2009. Integrated Science Assessment for Particulate Matter. EPA/600/R-08/139F. Washington, DC:U.S. EPA.

U.S. EPA (U.S. Environmental Protection Agency). 2014. Technology Transfer Network (TTN). Air Quality System (AQS). Available: http://www.epa.gov/ttn/airs/airsaqs/ [accessed 1 July 2013].

Wu S, Deng F, Wei H, Huang J, Wang H, Shima M, et al. 2012. Chemical constituents of ambient particulate air pollution and biomarkers of inflammation, coagulation and homocysteine in healthy adults: a prospective panel study. Part Fibre Toxicol 9:49; doi:10.1186/1743-8977-9-49.

Zanobetti A, Franklin M, Koutrakis P, Schwartz J. 2009. Fine particulate air pollution and its components in association with cause-specific emergency admissions. Environ Health 8:58; doi:10.1186/1476-069X-8-58.

Zhou J, Ito K, Lall R, Lippmann M, Thurston G. 2011. Time-series analysis of mortality effects of fine particulate matter constituents in Detroit and Seattle. Environ Health Perspect 119:461–466; doi:10.1289/ehp.1002613.