1-25-2007

MORTALITY IN THE MEDICARE POPULATION AND CHRONIC EXPOSURE TO FINE PARTICULATE AIR POLLUTION

Scott L. Zeger
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Francesca Dominici
Johns Hopkins Bloomberg School of Public Health, fdominic@jhsph.edu

Aidan McDermott
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Jonathan M. Samet
Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology

Mortality in the Medicare Population and Chronic Exposure to Fine Particulate Air Pollution

Scott L. Zeger [1]
Francesca Dominici [1]
Aidan McDermott [1]
Jonathan M Samet [2]

Abbreviations:

CMS:	Center for Medicare and Medicaid Services
MCAPS:	Medicare Air Pollution Cohort
NAAQS:	National Ambient Air Quality Standard

[1] The Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, Baltimore, MD, USA

[2] The Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Baltimore, MD, USA

For reprints, please contact:

Jonathan M. Samet, MD, MS
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe Street, Rm. W6041
Baltimore, Maryland 21205 USA
jsamet@jhsph.edu

Running head: Mortality and air pollution in a Medicare dataset

Abstract word count: 197
Text word count: 3,595

Abstract

Prospective cohort studies have provided evidence on longer-term mortality risks of fine particulate matter ($PM_{2.5}$), but due to their complexity and costs, only a few have been conducted.

By linking monitoring data to the U.S. Medicare system by county of residence, we developed a retrospective cohort study, the Medicare Air Pollution Cohort Study (MCAPS), comprising over 20 million enrollees in the 250 largest counties during 2000-2002. We estimated log-linear regression models having as outcome the age-specific mortality rate for each county and as the main predictor, the average $PM_{2.5}$ level for the study period 2000. Area-level covariates were used to adjust for socio-economic status and smoking. We reported results under several degrees of adjustment for spatial confounding and with stratification into by eastern, central and western counties.

We estimated that a 10 μg/m^3 increase in $PM_{2.5}$ is associated with a 7.6% increase in mortality (95% CI: 4.4 to 10.8%). We found a stronger association in the eastern counties than nationally, with no evidence of an association in western counties. When adjusted for spatial confounding, the estimated log-relative risks drop by 50%. We demonstrated the feasibility of using Medicare data to establish cohorts for follow-up for effects of air pollution.

Particulate matter (PM) air pollution is a global public health problem (1). In developing countries, levels of airborne particles still reach concentrations at which serious health consequences are well-documented; in developed countries, recent epidemiologic evidence shows continued adverse effects, even though particle levels have declined in the last two decades (2-6). Increased mortality associated with higher levels of PM air pollution has been of particular concern, giving an imperative for stronger protective regulations (7).

Evidence on PM and health comes from studies of acute and chronic adverse effects (6). The London Fog of 1952 provides dramatic evidence of the unacceptable short-term risk of extremely high levels of PM air pollution (8-10); multi-site time-series studies of daily mortality show that far lower levels of particles are still associated with short-term risk (5)(11-13). Cohort studies provide complementary evidence on the longer-term risks of PM air pollution, indicating the extent to which exposure reduces life expectancy. The design of these studies involves follow-up of cohorts for mortality over periods of years to decades and an assessment of mortality risk in association with estimated long-term exposure to air pollution (2-4;14-17). Because of the complexity and costs of such studies, only a small number have been conducted. The most rigorously executed, including the Harvard Six Cities Study and the American Cancer Society's (ACS) Cancer Prevention Study II, have provided generally consistent evidence for an association of long-term exposure to particulate matter air pollution with increased all-cause and cardio-respiratory mortality (2,4,14,15). Results from these studies have been used in risk assessments conducted for setting the U.S. National

Ambient Air Quality Standard (NAAQS) for PM and for estimating the global burden of disease attributable to air pollution (18,19).

Additional prospective cohort studies are necessary, however, to confirm associations between long-term exposure to PM and mortality, to broaden the populations studied, and to refine estimates by regions across which particle composition varies. Toward this end, we have used data from the U.S. Medicare system, which covers nearly all persons 65 years of age and older in the United States. We linked Medicare mortality data to $PM_{2.5}$ (particulate matter less than 2.5 μm in aerodynamic diameter) air pollution monitoring data to create a new retrospective cohort study, the Medicare Air Pollution Cohort Study (MCAPS), consisting of 20 million persons from 250 counties and representing about 50% of the US population of elderly living in urban settings. In this paper, we report on the relationship between longer-term exposure to $PM_{2.5}$ and mortality risk over the period 2000 to 2002 in the MCAPS.

Materials and Methods

MCAPS is a retrospective cohort study of participants 65 years and older enrolled in the U.S. Medicare system during the three-year period 2000-2002. To create the cohort, we used the Medicare enrollment file for the study period, which provides a listing of all enrollees, along with demographic information (age, race, and gender) and county of residence.

We obtained the date of death from the Center for Medicare and Medicaid Services (CMS); it is provided to CMS by the Social Security Agency. To validate the mortality

data, we compared annual age- and gender-adjusted mortality rates from CMS with the corresponding rates calculated from National Center of Health Statistics (NCHS) data. Figure 1 displays a scatter-plot of age-gender adjusted all-cause mortality rates (log scale) from MCAPS and NCHS data for the 250 cities for the year 2000. The correlation coefficient is 0.998 on the original scale and 0.986 on the log scale, indicating a high level of agreement between the two sources of mortality data for the one-year period.

For this paper, the outcome measure considered is the three-year (2000-2002) county mortality rate for each age stratum (65-74, 75-84, and 85 years and older). Initial analyses were also stratified by gender, but the estimated effects for men and women were very similar: consequently, the reported analyses are not stratified by gender.

The $PM_{2.5}$ data for the 250 largest U.S. counties with the most complete data were obtained from the U.S. EPA AirData database (http://www.epa.gov/oar/data/), which included more than 1,000 monitors for the period 2000-2002. Mean annual $PM_{2.5}$ values were compiled for 2000-2002 for every county with more than 10 months of data per year; months with four or more observations were considered to have sufficient data. Because of little variation in $PM_{2.5}$ annual averages within a county across the three years, we used the average county $PM_{2.5}$ for 2000-2002 as the chronic exposure estimate. As 1999 was the initial year of the EPA monitoring program and coverage was limited, 1999 data were omitted.

We estimated the region-specific relative risks of chronic $PM_{2.5}$ exposure for the eastern half of the U.S. with 167 urban locations (eastern U.S.), the mid-section with 51 counties (central U.S.), and the western U.S. with 32 counties extending from Washington State to southern California (Figure 2). We conducted the analyses for all 250 counties and separately within each geographical region.

In estimating the effect on mortality of PM or other air pollutants, cohort studies have accounted for potential confounding by 1) individual-level lifestyle factors, such as age, gender, smoking and 2) area-level characteristics such as socio-economic factors (SES). The MCAPS provides individual-level age, gender, and race data but not variables on lifestyle factors. To account for SES at the county level, variables from the 2000 US Census were used. After preliminary analysis of the full suite of available indicators, we selected two education variables -- percentage of the population with a high school diploma and the percentage with a higher education degree (associate or above) -- and two household income measures -- percentage of households living below the poverty level and median household income --, as well as percent unemployed. Area-level differences in cigarette smoking could potentially confound the $PM_{2.5}$ mortality association, although previous cohort studies found little effect in adjusting for self-reported smoking status (20). As the MCAPS database does not have individual- or county-level smoking information, we used mortality from chronic obstructive pulmonary disease (COPD) as a surrogate for past smoking, since most deaths from COPD in the United States result from smoking (21). We used data from the NCHS to calculate the standardized mortality ratio (SMR) for COPD for the period 1993-2002, adjusted for age, race, and gender. We then included the COPD SMR in the

regression model as a surrogate indicator of the long-term smoking pattern of county residents.

We investigated potential confounding by unmeasured factors that vary smoothly across locations. To account for these difficult-to-measure factors, we included in the regression model a smooth function $s(u_i, \lambda)$ of county location u_i (longitude and latitude) with degrees of freedom λ. Smaller λ corresponds to a rougher $s(u_i, \lambda)$ and less control for spatial confounding, while larger λ corresponds to a smoother $s(u_i, \lambda)$ and more control for spatial confounding; $\lambda = 0$, corresponds to no adjustment for spatial confounding.

In summary, within each age stratum, we estimated the following generalized additive model (GAM) (22) :

$$\log E[Y_i] = \log N_i + \beta_0 + \beta_{PM} PM_i + \beta_Z Z_i + s(u_i, \lambda).$$

Here β_{PM} is the log relative risk of mortality associated with a one μg/m^3 difference in average $PM_{2.5}$ comparing counties that are otherwise similar with respect to SES and COPD SMR and a smooth function of location. We report results by age stratum and with aggregation across strata. To obtain the aggregated value, we fit a single log-linear regression with a common PM effect across the strata by allowing the adjustment for Z_i and for the smooth function to be done separately for each stratum.

We estimated the log-relative risk β_{PM} as we varied λ. As λ increased, control for spatial confounding increased and mortality was contrasted in cities with higher and lower $PM_{2.5}$ concentrations that are nearer to each other in space.

We smoothed in two dimensions in the eastern U.S. and the central U.S., and with one dimension, latitude, in the western region. If there are no important spatial confounders and minimal measurement error, then the PM coefficient will not change with $\tilde{\lambda}$. We estimated the PM effect with $\lambda = 0$ using all 250 cities and compared our findings with those previously published for three different levels of covariate adjustment: age-gender only, SES covariates, and COPD SMR. All analyses were carried out with the statistical programs R (http://www.r-project.org/) and SAS (http://www.sas.com/technologies/analytics/statistics).

Results

Table 1 presents the total number of counties, the study population, person-years of follow-up, the number of deaths, and the crude death rates for the eastern, central, and western U.S. The study population and the numbers of deaths exceed 20 million and 2.5 million, respectively. The study includes more than 50 million person-years of follow-up between 2000 and 2002.

The MCAPS data provide strong evidence that mortality is higher in counties with higher $PM_{2.5}$ when there is no adjustment for potential confounding by unmeasured spatial variables ($\lambda = 0$). Table 2 summarizes the estimated log relative risk of death associated

with a 10 μg/m^3 increase in $PM_{2.5}$ level obtained using all 250 and setting $\lambda = 0$. We estimate that a county with 10 μg/m^3 per cubic meter higher long-term average $PM_{2.5}$ has 7.6 percent higher (95% CI: 4.4 to 10.8% higher) mortality than a county with comparable age and gender distribution. Control for SES indicators and COPD SMR results in little change in the estimates. The analyses were done using Medicare and NCHS mortality rates and results were qualitatively similar.

Table 3 presents the estimated log relative risks of death separately for the eastern U.S., the central U.S., and the western U.S. without controlling for spatial confounding ($\lambda = 0$), and for the three levels of adjustment for area-level measured covariates. For the 167 eastern U.S. counties and the 52 central US counties, we found that the log-relative risks estimates were higher than the national estimate. Estimates are smaller but still statistically significant with adjustment for SES and COPD. For the 32 western counties, we found little evidence for an association of $PM_{2.5}$ with mortality.

Table 4 summarizes separately for the three geographical regions the percent of PM variation remaining for estimation of the log relative risk for several values of the smoothing parameter ℵ. Even for the largest values of ℵ considered, at least 15% of the variation in PM remains to estimate the log-relative risk.

Figures 2, 3, and 4 show by region the estimated PM log relative risks after controlling for age, gender, SES indicators, and COPD SMR for a range of values of λ. For the eastern U.S. (Figure 2), the estimated log-relative risks become progressively smaller as λ increases; an increase in λ removes more of the spatially smooth variation in PM

2.5 from this predictor, and thus protects against spatial confounders that also vary smoothly. Figures A1 and A2 (see supplemental material) show overlapping hills of mortality and $PM_{2.5}$ in the eastern U.S. These coincident hills are the major source of the positive association estimated by the log-linear model with $\lambda \leq 8$. Setting this smooth variation aside and considering instead the variation at more local spatial scales, the PM effect on mortality is roughly half as large; however, it remains statistically significant until $\lambda \geq 40$. For the central U.S. (Figure 3), the estimated log-relative risk also diminishes as λ increases and loses statistical significance for $\lambda \geq 18$. Finally, for the 32 western U.S. locations (Figure 4), there is little evidence of an association between chronic $PM_{2.5}$ and mortality for any of the three age strata and for any value of λ. The annual average $PM_{2.5}$ levels vary among the western counties with a standard deviation of 4.6 $\mu g/m^3$, nearly twice as great as observed in the eastern U.S. Annual mortality rates also vary substantially, but air pollution levels and mortality rates, aggregated to the county level, are not correlated for the 32 western locations.

Discussion

This paper presents results from the Medicare Cohort Air Pollution Study (MCAPS), the largest cohort study of air pollution effects on morbidity and mortality to date, with 2.5 million deaths during more than 50 million person-years of follow-up. In comparison, the ACS Study (20) had 20,765 deaths in the sub-cohort used in the analyses of air pollution and mortality. Given the availability of 250 locations, we have stratified the analyses geographically, choosing strata that broadly reflected differing source mixes and background disease patterns. This stratification also controls for potential confounders that vary on broad geographic scales.

The MCAPS basic analysis ($\lambda = 0$) gives results similar to those previously published from the Six Cities (14) and the ACS (2) studies. The relative risk estimate is 15.4% per 10 μg/m^3 increase in $PM_{2.5}$ (95% CI: 10.2 to 20.6%) for the youngest age group in MCAPS, compared to the Six Cities and ACS studies values of 15.3 and 12.4 %, respectively. While the MCAPS data lack individual-level risk factor information, the MCAPS results changed little with inclusion of county-level SES indicators and the COPD SMR in the log-linear regression model (Table 3).

In MCAPS, we found evidence for effect modification by age - a drop in relative risk with increasing age (Table 2) - and by geographic location. This decline may reflect the many competing causes of death for which the hazard of death increases with age. If the hazard is increased for only a subset of the competing causes by $PM_{2.5}$, then the relative risk will drop with increasing age.

There was a strong pattern of effect modification by geographic location with evidence for the positive association between $PM_{2.5}$ and mortality coming entirely from the central and eastern United States. The positive PM-mortality association estimated from the 176 counties in the eastern U.S. could be caused by the overlapping hills of mortality and $PM_{2.5}$ that peak near the Ohio River Valley. The elevated relative risk estimates in the central region are substantially determined from comparisons of Texas counties, where $PM_{2.5}$ concentrations and mortality rates are higher, to more northern counties where both are relatively lower.

No positive association was found between county-level $PM_{2.5}$ concentration and mortality rates for the 32 urban counties in the western U.S. In the MCAPS cohort, the lack of association for the West is largely because the Los Angeles area counties have higher $PM_{2.5}$ levels than other western counties, but not higher adjusted mortality rates.

Several previous studies have addressed PM exposure and mortality in California, based on gradients of exposure within the state. Abbey et al. (23) reported a follow-up of the Adventist Health Study, a cohort study of more than 6,000 nonsmoking residents of three air basins: San Francisco, Los Angeles, and San Diego. They found a non-significant increase in all-cause mortality of roughly 5% per 10 $\mu g/m^3$ increase in PM_{10} in males and no effect in females. They also reported a statistically significant association in respiratory deaths with fraction of days above 100 $\mu g/m^3$ of PM_{10} for both genders. Enstrom (24) tracked mortality from 1973 through 2002 in 50,000 California participants in the first national cohort study carried out by the ACS. Using $PM_{2.5}$ data for 11 counties in 1979-1983, he found no association across the full follow-up period and evidence of a small effect during the first decade of follow-up. Misclassification arising from the limited exposure data available may have biased this study towards the null. Most recently, Jerrett et al. (3) investigated the PM-mortality association in a subset of the second ACS cohort living in Los Angeles. They estimated an 11% increase in mortality per 10 $\mu g/m^3$ of $PM_{2.5}$ (95% CI: -1 to 25%), using an indicator of chronic PM exposure based on a model that incorporated measured PM, traffic patterns, and residential proximity to freeways. The many methodological differences between the

MCAPS cohort and the other three cohorts providing findings for California residents prevent any reconciliation of the differing findings.

By increasing the degrees of freedom λ for the nuisance spatial function, we examined the change in relative risk estimates, as greater weight is given to the comparison of counties closer to one another. Absent confounding by unmeasured variables and any effect on measurement error, the $PM_{2.5}$ relative risk of death should be similar regardless of the value of λ. That is, on the assumption that $PM_{2.5}$ exposure causes death, the evidence for association should exist when comparing distant locations (λ smaller) or neighboring locations (λ greater).

Using this analytic approach, we find that the relative risk estimates for the eastern and central U.S. counties decrease by roughly 50% when more distant city comparisons are down-weighted in the analysis. Nevertheless, both positive and statistically significant associations remain for a wide range of smoothing parameters. There are competing explanations for the finding that the relative risks estimates decrease as λ is increased. Confounding by unobserved variables is one possibility. Assuming that the hill in mortality in the east is caused by $PM_{2.5}$ (Figures A1 and A2, see supplemental material), the reduced association at smaller spatial scales may reflect unmeasured confounders that vary locally, but not regionally. If the PM-mortality association is not causal, the mortality hill might be caused by factors other than $PM_{2.5}$ that co-vary smoothly in space as $PM_{2.5}$.

These potential confounders include “ecologic” effects that result from using aggregate rather than individual exposure and risk factor data. As detailed in Appendix A, the observed mortality rate for a county may be a biased estimate of the true rate because only county-level exposure and covariate data and not individual-level data are available. The appendix decomposes the bias into five terms, coming from: 1) differences between the baseline-risk-weighted personal $PM_{2.5}$ exposure (x) and the aggregated values used in our models; 2) differences between the baseline-risk-weighted averages of personal characteristics (e.g., smoking and SES) (z) and their aggregated values used in our models; 3-4) baseline-risk-weighted variance within counties in $PM_{2.5}$ exposure (x) and personal characteristics (z); and 5) covariance within counties in the personal $PM_{2.5}$ exposure and personal characteristics. If the county-specific bias comprising these five terms varies across counties with $PM_{2.5}$ exposure, then the relative risk estimates will be biased.

Within the data reported here, we cannot evaluate the potential for ecologic bias from these terms. However, Krewski et al. (20) have compared relative risk estimates from both the American Cancer Society and Six Cities studies with and without control for individual-level characteristics including smoking, exercise, education, and occupational exposures. They found little change in the PM relative risk estimates, suggesting that terms 2, 4, and 5 are negligible for those personal characteristics measured in these two studies.

Terms 1 and 3 are the effects of using ambient rather than personal level exposure data as in most previous cohort studies of chronic air pollution exposure. While there is likely some difference between average personal PM2.5 exposure and the average ambient values used in this, and all other cohort studies, we do not anticipate that this difference co-varies with PM2.5 levels across the country, a spatial pattern of variation that would be necessary to produce bias in relative risk estimates. In its classification of exposure, the MCAPS design is comparable to both the Six Cities and ACS studies, which also assigned exposures based on centrally-sited monitors. In some other cohort studies, exposures were been estimated at the individual level using models and residence location (3,17). This approach can assign estimates to most or even all study participants, potentially reducing the effects of exposure measurement error.

An alternate explanation for the dependence of the relative risk estimate on λ is that $PM_{2.5}$ is not the causal factor but a surrogate for one of its constituents or for another pollutant. If the actual toxic exposure is a component of $PM_{2.5}$ that varies smoothly and is highly correlated with $PM_{2.5}$ at larger spatial scales, but has little or no local variation, then the local $PM_{2.5}$ variation would not be associated with mortality. Our results point toward a component for which $PM_{2.5}$ is a better surrogate on larger spatial scales than on more local scales.

A strength of these and some other cohort studies is their use of mortality rates that are adjusted not only for age, gender, and race, but also for personal characteristics that might confound the PM-mortality association. However, in the American Cancer Society

Study, when the sensitivity of pollution relative rates to potential confounding variables was studied in detail (20), little evidence of confounding by individual-level information was observed, once age, gender, and race were account for. The analyses in these studies have not included a smooth spatial term, so that the approach is equivalent to setting $\lambda = 0$ in our model formulation. Krewski et al. (20) and Burnett et al. (25) have previously accounted for spatial autocorrelation in estimating chronic effects of $PM_{2.5}$ and PM_{10} by including a loess smoother in the model. They used a frailty extension of the Cox proportional hazards model which assumes the existence of a latent spatial process that explains the correlation among deaths rates from neighboring regions. However, they assume that the latent process is orthogonal to the spatial field of $PM_{2.5}$ concentrations, allowing $PM_{2.5}$ to explain as much as possible of the mortality spatial pattern and then using a latent process to represent residual spatial autocorrelation. In our approach, we assume that the latent spatial process is a fixed effect that may not be orthogonal to the $PM_{2.5}$ field. We remove the spatial variation that can be captured by a smooth function (e.g., loess or spline) with the chosen degrees of freedom and then estimate the association of mortality with the remaining variation in $PM_{2.5}$.

In spite of these methodological complexities, we have shown that a cohort can be established using Medicare participants and routine monitoring data for investigating air pollution and mortality on longer-term time frames. In our initial analyses of the MCAPS data, we confirmed the association between $PM_{2.5}$ and mortality found in other studies but find substantial and unexplained geographic heterogeneity in the effect of $PM_{2.5}$ across the United States.

Figure Legends

Figure A1: Scatter-plot of Medicare versus NCHS age-gender adjusted all-cause mortality rates (log scale) for the 250 cities for the year 2000.

Figure A2: Smoothed $PM_{2.5}$ surface using GAM with a two dimensional loess smoother of latitude and longitude employing 12 degrees of freedom while controlling for proportion with high-school education, proportion with degree, proportion living in poverty, proportion unemployed, median income, and the standardized mortality ratio for COPD. $PM_{2.5}$ residuals are plotted on the pollution surface, red indicating where the actual $PM_{2.5}$ is above the predicted surface and green when below.

Figure 1: Smoothed mortality surface using GAM with a two dimensional loess smoother of latitude and longitude employing 12 degrees of freedom while controlling for proportion with high-school education, proportion with degree, proportion living in poverty, proportion unemployed, median income, and the standardized mortality ratio for COPD. Standardized mortality ratios (100 * actual mortality/predicted mortality) are shown. The color red is used to indicate when the number of deaths exceeds the predicted number of deaths and the color green when the predicted number of deaths exceeds the number of deaths at a location.

Figure 2: Point estimates and 95% confidence intervals of the mortality log-relative risks associated with average exposure to $PM_{2.5}$ over the period 2000-2002 for different values of the smoothing parameter λ for the 167 counties located in the eastern U.S.

Figure 3: Point estimates and 95% confidence intervals of the mortality log-relative risks associated with average exposure to $PM_{2.5}$ over the period 2000-2002 for different values of the smoothing parameter λ for the 51 counties located in the central U.S.

Figure 4: Point estimates and 95% confidence intervals of the mortality log-relative risks associated with average exposure to $PM_{2.5}$ over the period 2000-2002 for different values of the smoothing parameter λ for the 32 counties located in the western U.S.

Table 1. Number of counties, person-years of follow-up, deaths and crude death rates stratified by region for the MCAPS data; US Medicare population (as described), 2000-2002.

	Eastern U.S.	Central U.S.	Western U.S.	Total

Number of counties	167	51	32	250
Persons (millions)	13.1	3.6	3.7	20.3
Person-years (millions)	33.5	8.9	9.5	51.9
Deaths (millions)	1.76	0.46	0.46	2.68
Crude rate (deaths/1,000 person years)	52.5	51.7	48.8	51.7

Table 2. Percent increase and 95% confidence interval in mortality per 10 μg/m^3 $PM_{2.5}$ estimated from the log-linear regression using all 250 counties with no spatial adjustment ($\lambda = 0$) for MCAPS data for three levels of adjustment for demographic and socioeconomic variables; US Medicare population (as described), for the period 2000-2002..

	Ages							
Adjustment	All		65-74		75-84		85+	
	% Inc	95% CI	% Inc	95% CI	% Inc	95% CI	% Inc	95% CI
Age	7.7	5.4,10.0	15.6	11.7, 19.6	8.1	4.2, 12.1	-0.5	4.4, 3.5
Age+SES	8.6	6.8,10.4	14.5	11.4, 17.6	9.1	6.0, 12.1	2.3	0.7, 5.4
Age+SES+COPD	7.7	6.2, 9.3	12.2	9.5, 14.8	8.6	6.0, 11.2	2.6	0.0, 5.2

Table 3. Percent increase and 95% confidence interval in mortality per 10 μg/m^3 $PM_{2.5}$ estimated from the log-linear regression separately for the 167 counties in the eastern U.S, the 51 counties in the central U.S. and the 32 counties in the western U.S. (32) with no spatial adjustment ($\lambda = 0$) for MCAPS data for three levels of adjustment for demographic and socioeconomic variables; for the period 2000-2002..

	All		Eastern U.S. (n=167)		Central U.S. (n=51)		Western U.S. (n=32)	
	% Inc	95% CI	% Inc	95% CI	% Inc	95% CI	% Inc	95% CI
Age	7.7	5.4,10.0	12.5	9.1,15.9	19.6	11.5, 27.7	2.9	-0.6 ,6.4
Age+SES	8.6	6.8,10.4	10.6	8.0,13.2	16.8	10.3, 23.2	0.6	-3.4, 4.6
Age+SES+COPD	7.7	6.2, 9.3	5.9	4.0, 7.7	17.8	11.9, 23.7	-2.5	-6.2, 1.1

Table 4. Residual variance and percent of total for mean $PM_{2.5}$ (2000-2002) as a function of degrees of freedom () in loess smoother; US Medicare population (as described), for the period 2000-2002..

Eastern U.S.	df (λ)	0	3.0	8.0	15.0	20.3	35.2
	Variance	4.98	2.59	1.37	1.13	1.05	0.78
	% Variance	100	52.1	27.6	22.8	21.1	15.7
Central U.S.	df (λ)	0	3.3	8.0	10.2	15.0	20.9
	Variance	3.67	2.06	1.29	1.20	0.90	0.61
	% Variance	100	56.1	35.2	32.7	24.6	16.5
Western U.S.	df (λ)	0	1.0	2.0	4.0	5.9	10.2
	Variance	20.8	11.09	8.90	6.98	5.94	4.05
	% Variance	100	53.3	42.8	33.6	28.6	19.5

Appendix A

This appendix derives expressions for the expected rate μ_c of mortality for county c from a relative risk model for the risk of death λ_{ci} for person I in county c. Let x_{ci} and z_{ci} be the long-term average personal exposure and a set of personal characteristics (e.g., smoking, exercise level, occupational exposures, and SES) for person i. Let x_c be the measured ambient $PM_{2.5}$ exposure and let z_c be the county average of the personal characteristics, $z_c = 1/n_c \sum_i z_{ci}$. Finally, let u_c be unmeasured county characteristics such as availability of health services, local cultural beliefs and other factors that might confound the $PM_{2.5}$ –mortality association.

The individual level risk model can be written

$$\lambda_{ci} = \lambda_{0ci} \exp\{\beta_x x_{ci} + \beta_z z_{ci} + \beta_u u_c\}$$
$$\lambda_{ci} = \lambda_{0ci} \exp\{\beta_x (x_{ci} - x_c + x_c) + \beta_z (z_{ci} - z_c + z_c) + \beta_u u_c\}$$

Let y_{ci} =1 if person I from county c dies during m_{ci} years of follow-up and define $y_c = \sum_i y_{ci}$ to be the number of deaths in the county with M_c total years of follow-up.

Then the expected number of deaths in county c is given by

$$\mu_c = E(y_c) = \sum_i m_{ci} \lambda_{ci} = \sum_i m_{ci} \lambda_{0ci} \exp\{\beta_x x_{ci} + \beta_z z_{ci} + \beta_u u_c\}$$
$$\mu_c = \exp\{\beta_x x_c + \beta_z z_c + \beta_u u_c\} \times b_c$$

where b_c is the potential "bias-causing" factors that results from using aggregated rather than individual-level exposure and personal-characteristic data.

The bias term can be approximated by

$$b_c = \sum_i m_{ci}\lambda_{0ci}\exp\{\beta_x x_{ci} - x_c) + \beta_z(z_{ci} - z_c)\}$$

$$b_c \approx \beta_x(\bar{x}_c^w - x_c) + \beta_z(\bar{z}_c^w - z_c) + \frac{\beta_x^2}{2}Var_c^w(x) + \frac{\beta_z^2}{2}Var_c^w(z) + Cov_c^w(x,z).$$

Here $\bar{x}_c^w$ and $\bar{z}_c^w$ are the cumulative baseline risk weighted mean PM exposure and mean personal characteristics for county c; $Var_c^w(x)$ and $Var_c^w(z)$ are the cumulative baseline-risk weighted variances; and $Cov_c^w(x,z)$ is the analogous covariance between the exposure and covariate values for individuals within county c.

Acknowledgements: Funding for Drs. Zeger, Dominici, McDermott, and Samet was provided by the U.S. Environmental Protection Agency (RD-83054801 and RD-83241701). Funding for Drs. Zeger, Dominici, McDermott, and Samet was also provided by the National Institute for Environmental Health Sciences (ES012054-03). Although the research described in this article has been funded wholly or in part by the United States Environmental Protection Agency through grant agreements (RD-83054801 and RD-83241701) to Johns Hopkins University, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

Reference List

(1) Cohen AJ, Anderson HR, Ostro B, Pandey KD, Krzyzanowski M, Kunzli N et al. Urban air pollution. In: Ezzati M, Lopez AD, Rodgers A, Murray CJ, editors. Comparative quantification of health risks. Global and regional burden of disease attributable to selected major risk factors. Geneva: World Health Organization, 2004: 1353-1434.

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

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

(4) Laden F, Schwartz J, Speizer FE, Dockery DW. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *Am J Respir Crit Care Med* 2006; 173(6):667-672.

(5) Dominici F, Peng RD, Bell ML, Pham L, McDermott A, Zeger SL et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA* 2006; 295(10):1127-1134.

(6) Pope CA, III, Dockery DW. Health effects of fine particulate air pollution: lines that connect. *J Air Waste Manag Assoc* 2006; 56(6):709-742.

(7) World Health Organization. Air Quality Guidelines for Europe, 3rd edition

(8) Logan WPD. Mortality in the London fog incident, 1952. *Lancet* 1953; 1:336-338.

(9) Bell ML, Davis DL. Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. *Environ Health Perspect* 2001; 109 Suppl 3:389-394.

(10) Bell ML, Davis DL, Fletcher T. A retrospective assessment of mortality from the London smog episode of 1952: the role of influenza and pollution. *Environ Health Perspect* 2004; 112(1):6-8.

(11) Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S et al. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from the APHEA project. *Br Med J* 1997; 314:1658-1663.

(12) Samoli E, Schwartz J, Wojtyniak B, Touloumi G, Spix C, Balducci F et al. Investigating regional differences in short-term effects of air pollution on daily mortality in the APHEA project: a sensitivity analysis for controlling long-term trends and seasonality. *Environ Health Perspect* 2001; 109:349-353.

(13) Lee JT, Kim H, Hong YC, Kwon HJ, Schwartz J, Christiani DC. Air pollution and daily mortality in seven major cities of Korea, 1991-1997. *Environ Res* 2000; 84(3):247-254.

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

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

(16) Krewski D, Burnett RT, Goldberg MS, Hoover K, Siemiatycki J, Abrahamowicz M et al. Validation of the Harvard Six Cities Study of particulate air pollution and mortality. *N Engl J Med* 2004; 350(2):198-199.

(17) Hoek G, Brunekreef B, Goldbohm S, Fischer P, van den Brandt PA. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *Lancet* 2002; 360(9341):1203-1209.

(18) US Environmental Protection Agency (EPA), Clean Air Scientific Advisory Committee (CASAC). Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. OAQPS Staff paper. 2005. Research Triangle Park, NC, USEPA.

(19) Cohen AJ, Ross AH, Ostro B, Pandey KD, Krzyzanowski M, Kunzli N et al. The global burden of disease due to outdoor air pollution. *J Toxicol Environ Health A* 2005; 68(13-14):1301-1307.

(20) Krewski D, Burnett RT, Goldberg MS, Hoover K, Siemiatycki J, Abrahamowicz M et al. Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of particulate air pollution and mortality. Investigators' reports parts I and II. 2000. Cambridge, MA, Health Effects Institute.

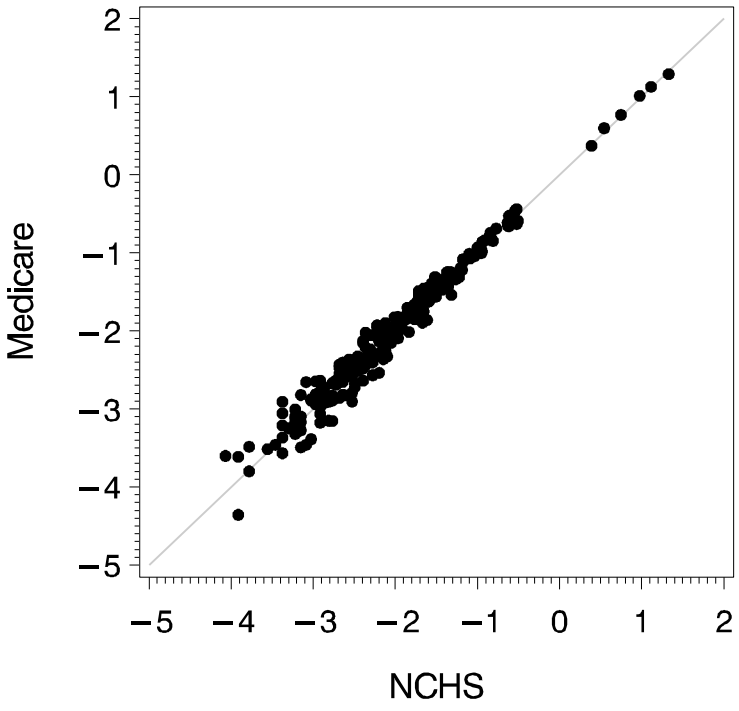
(21) US Department of Health and Human Services (USDHHS). The health effects of active smoking: A report of the Surgeon General. 2004. Washington, D.C., U.S. Government Printing Office.

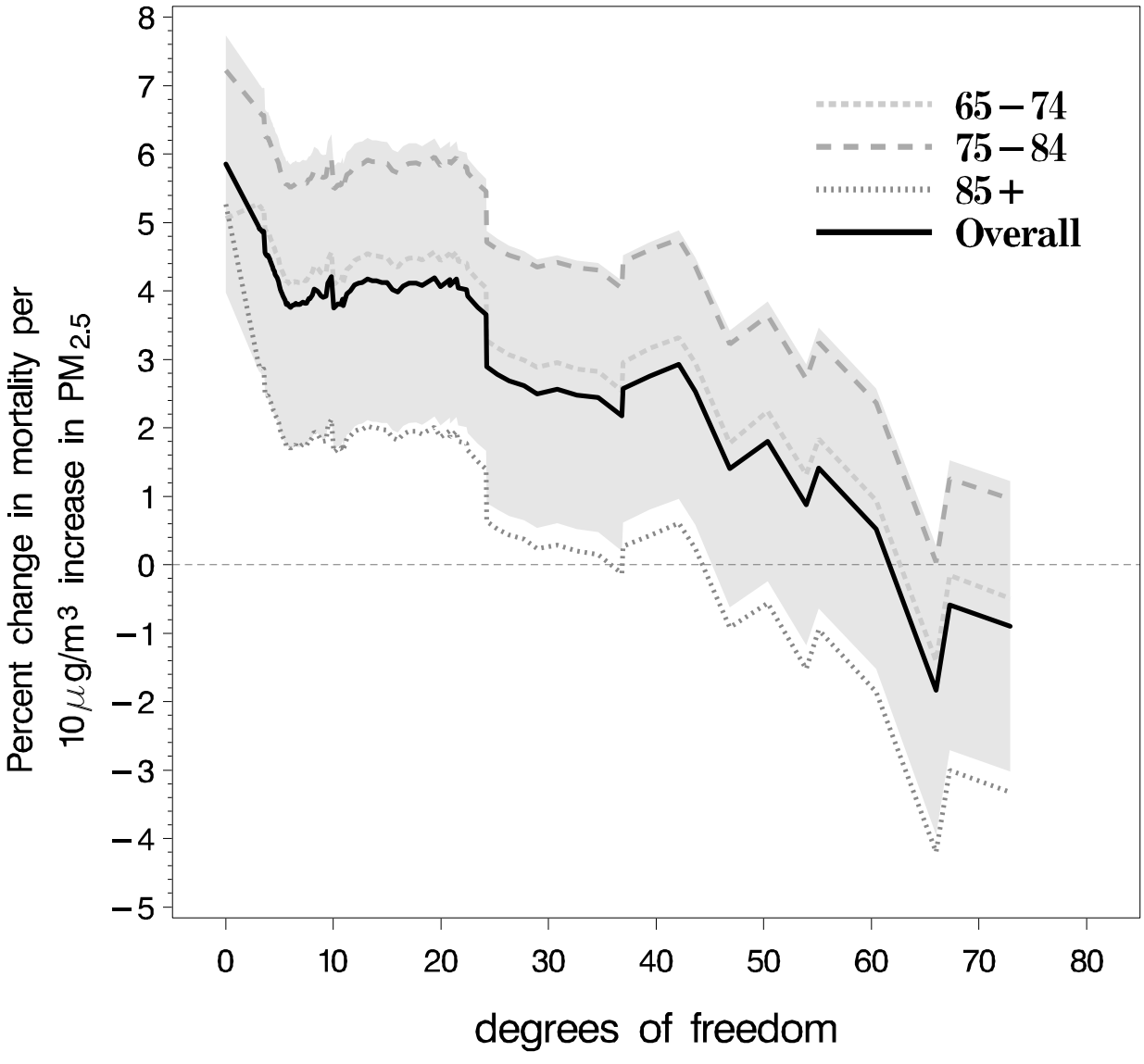
(22) Hastie TJ, Tibshirani RJ. Generalized additive models. New York: Chapman and Hall, 1990.

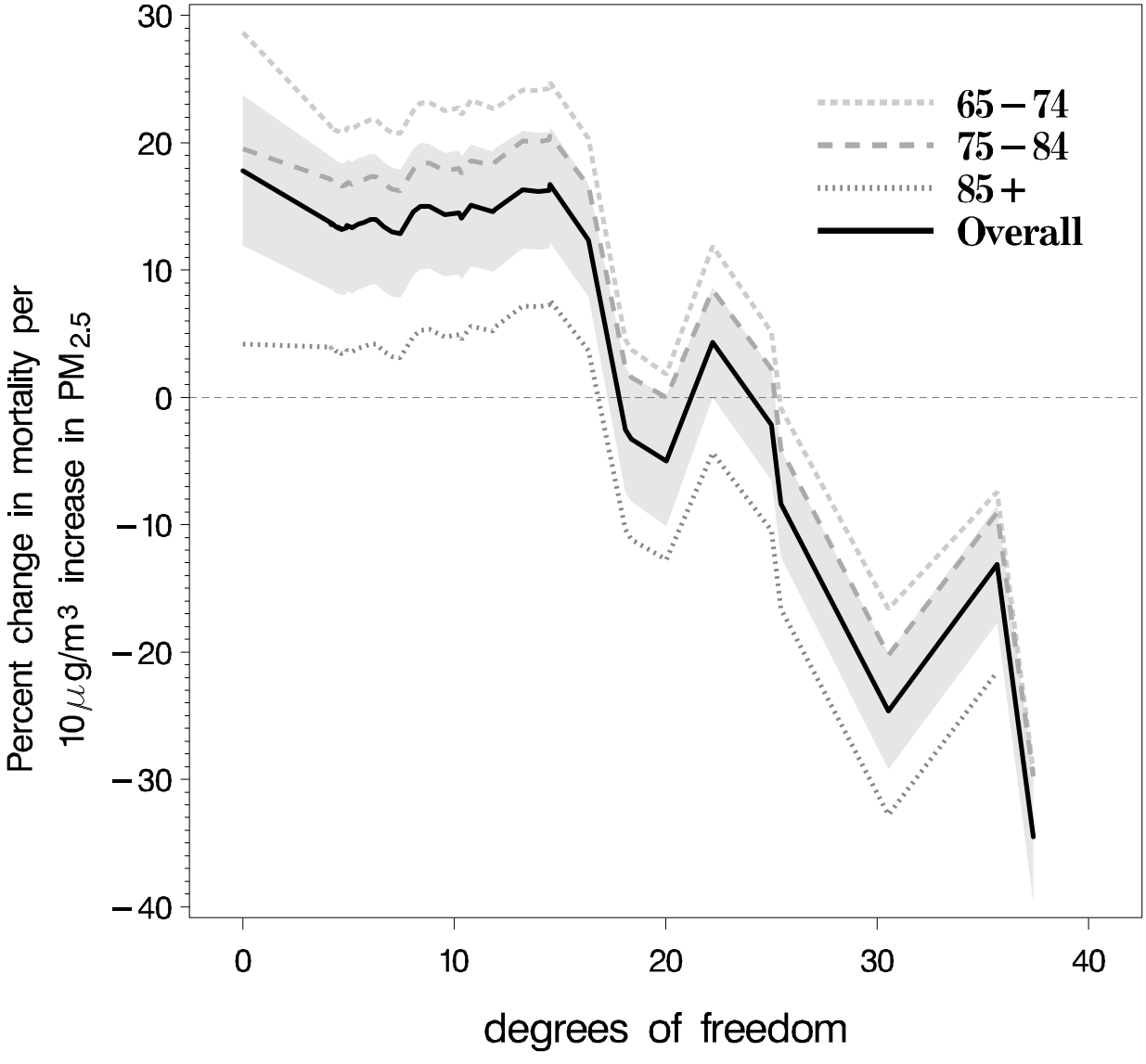
(23) Abbey DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Lawrence BW et al. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. *Am J Respir Crit Care Med* 1999; 159(2):373-382.

(24) Enstrom JE. Fine particulate air pollution and total mortality among elderly Californians, 1973-2002. *Inhal Toxicol* 2005; 17(14):803-816.

(25) Burnett R, Ma R, Jerrett M, Goldberg MS, Cakmak S, Pope CA, III et al. The spatial association between community air pollution and mortality: a new method of analyzing correlated geographic cohort data. *Environ Health Perspect* 2001; 109 Suppl 3:375-380.


Medicare
NCHS
2
1
0
−1
−2
−3
−4
−5
−5
−4
−3
−2
−1
0
1
2




65 – 74
75 – 84
85 +
Overall
Percent change in mortality per
10 μg/m3 increase in PM2.5
30
20
10
0
−10
−20
−30
−40
0
10
20
30
40
degrees of freedom