ADDITIONAL MATERIALS AVAILABLE ON THE HEI WEBSITE

Research Report 200

Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Phase 1

Dominici et al.

Additional Materials. Di Q, et al. 2017. Air pollution and mortality in the Medicare population. N Engl J Med 376:2513–2522; doi:10.1056/NEJMoa1702747 and Supplementary Appendix.

The Additional Materials were not formatted or edited by HEI. This document was part of the HEI Low-Exposure Epidemiology Studies Review Panel's review process.

Correspondence may be addressed to Dr. Francesca Dominici, Harvard T.H. Chan School of Public Health, 677 Huntington Ave., Boston, MA 02115; e-mail: *fdominic@hsph.harvard.edu*.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83467701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 JUNE 29, 2017 VOL. 376 NO. 26

Air Pollution and Mortality in the Medicare Population

Qian Di, M.S., Yan Wang, M.S., Antonella Zanobetti, Ph.D., Yun Wang, Ph.D., Petros Koutrakis, Ph.D., Christine Choirat, Ph.D., Francesca Dominici, Ph.D., and Joel D. Schwartz, Ph.D.

ABSTRACT

BACKGROUND

Studies have shown that long-term exposure to air pollution increases mortality. However, evidence is limited for air-pollution levels below the most recent National Ambient Air Quality Standards. Previous studies involved predominantly urban populations and did not have the statistical power to estimate the health effects in underrepresented groups.

METHODS

We constructed an open cohort of all Medicare beneficiaries (60,925,443 persons) in the continental United States from the years 2000 through 2012, with 460,310,521 person-years of follow-up. Annual averages of fine particulate matter (particles with a mass median aerodynamic diameter of less than 2.5 μm [$PM_{2.5}$]) and ozone were estimated according to the ZIP Code of residence for each enrollee with the use of previously validated prediction models. We estimated the risk of death associated with exposure to increases of 10 μg per cubic meter for $PM_{2.5}$ and 10 parts per billion (ppb) for ozone using a two-pollutant Cox proportional-hazards model that controlled for demographic characteristics, Medicaid eligibility, and area-level covariates.

RESULTS

Increases of 10 μg per cubic meter in $PM_{2.5}$ and of 10 ppb in ozone were associated with increases in all-cause mortality of 7.3% (95% confidence interval [CI], 7.1 to 7.5) and 1.1% (95% CI, 1.0 to 1.2), respectively. When the analysis was restricted to person-years with exposure to $PM_{2.5}$ of less than 12 μg per cubic meter and ozone of less than 50 ppb, the same increases in $PM_{2.5}$ and ozone were associated with increases in the risk of death of 13.6% (95% CI, 13.1 to 14.1) and 1.0% (95% CI, 0.9 to 1.1), respectively. For $PM_{2.5}$, the risk of death among men, blacks, and people with Medicaid eligibility was higher than that in the rest of the population.

CONCLUSIONS

In the entire Medicare population, there was significant evidence of adverse effects related to exposure to $PM_{2.5}$ and ozone at concentrations below current national standards. This effect was most pronounced among self-identified racial minorities and people with low income. (Supported by the Health Effects Institute and others.)

From the Departments of Environmental Health (Q.D., Yan Wang, A.Z., P.K., J.D.S.) and Biostatistics (Yun Wang, C.C., F.D.), Harvard T.H. Chan School of Public Health, Boston. Address reprint requests to Dr. Dominici at Harvard T.H. Chan School of Public Health, Biostatistics Department, Bldg. 2, 4th Flr., 655 Huntington Ave., Boston, MA 02115, or at fdominic@hsph.harvard.edu.

N Engl J Med 2017;376:2513-22.
DOI: 10.1056/NEJMoa1702747

A Quick Take is available at NEJM.org

The adverse health effects associated with long-term exposure to air pollution are well documented.[1,2] Studies suggest that fine particles (particles with a mass median aerodynamic diameter of less than 2.5 μm [$PM_{2.5}$]) are a public health concern,[3] with exposure linked to decreased life expectancy.[4-6] Long-term exposure to ozone has also been associated with reduced survival in several recent studies, although evidence is sparse.[4,7-9]

Studies with large cohorts have investigated the relationship between long-term exposures to $PM_{2.5}$ and ozone and mortality[4,9-13]; others have estimated the health effects of fine particles at low concentrations (e.g., below 12 μg per cubic meter for $PM_{2.5}$).[14-18] However, most of these studies have included populations whose socioeconomic status is higher than the national average and who reside in well-monitored urban areas. Consequently, these studies provide limited information on the health effects of long-term exposure to low levels of air pollution in smaller cities and rural areas or among minorities or persons with low socioeconomic status.

To address these gaps in knowledge, we conducted a nationwide cohort study involving all Medicare beneficiaries from 2000 through 2012, a population of 61 million, with 460 million person-years of follow-up. We used a survival analysis to estimate the risk of death from any cause associated with long-term exposure (yearly average) to $PM_{2.5}$ concentrations lower than the current annual National Ambient Air Quality Standard (NAAQS) of 12 μg per cubic meter and to ozone concentrations below 50 parts per billion (ppb). Subgroup analyses were conducted to identify populations with a higher or lower level of pollution-associated risk of death from any cause.

Methods

Mortality Data

We obtained the Medicare beneficiary denominator file from the Centers for Medicare and Medicaid Services, which contains information on all persons in the United States covered by Medicare and more than 96% of the population 65 years of age or older. We constructed an open cohort consisting of all beneficiaries in this age group in the continental United States from 2000 through 2012, with all-cause mortality as the outcome. For each beneficiary, we extracted the date of death (up to December 31, 2012), age at year of Medicare entry, year of entry, sex, race, ZIP Code of residence, and Medicaid eligibility (a proxy for low socioeconomic status). Persons who were alive on January 1 of the year following their enrollment in Medicare were entered into the open cohort for the survival analysis. Follow-up periods were defined according to calendar years.

Assessment of Exposure to Air Pollution

Ambient levels of ozone and $PM_{2.5}$ were estimated and validated on the basis of previously published prediction models.[19,20] Briefly, we used an artificial neural network that incorporated satellite-based measurements, simulation outputs from a chemical transport model, land-use terms, meteorologic data, and other data to predict daily concentrations of $PM_{2.5}$ and ozone at unmonitored locations. We fit the neural network with monitoring data from the Environmental Protection Agency (EPA) Air Quality System (AQS) (in which there are 1928 monitoring stations for $PM_{2.5}$ and 1877 monitoring stations for ozone). We then predicted daily $PM_{2.5}$ and ozone concentrations for nationwide grids that were 1 km by 1 km. Cross-validation indicated that predictions were good across the entire study area. The coefficients of determination (R^2) for $PM_{2.5}$ and ozone were 0.83 and 0.80, respectively; the mean square errors between the target and forecasting values for $PM_{2.5}$ and ozone were 1.29 μg per cubic meter and 2.91 ppb, respectively. Data on daily air temperature and relative humidity were retrieved from North American Regional Reanalysis with grids that were approximately 32 km by 32 km; data were averaged annually.[21]

For each calendar year during which a person was at risk of death, we assigned to that person a value for the annual average $PM_{2.5}$ concentration, a value for average ozone level during the warm season (April 1 through September 30), and values for annual average temperature and humidity according to the ZIP Code of the person's residence. The warm-season ozone concentration was used to compare our results with those of previous studies.[10] In this study, "ozone concentration" refers to the average concentration during the warm season, unless specified otherwise.

As part of a sensitivity analysis, we also obtained data on $PM_{2.5}$ and ozone concentrations from the EPA AQS and matched that data with

each person in our study on the basis of the nearest monitoring site within a distance of 50 km. (Details are provided in Section 1 in the Supplementary Appendix, available with the full text of this article at NEJM.org.)

STATISTICAL ANALYSIS

We fit a two-pollutant Cox proportional-hazards model with a generalized estimating equation to account for the correlation between ZIP Codes.[22] In this way, the risk of death from any cause associated with long-term exposure to $PM_{2.5}$ was always adjusted for long-term exposure to ozone, and the risk of death from any cause associated with long-term exposure to ozone was always adjusted for long-term exposure to $PM_{2.5}$, unless noted otherwise. We also conducted single-pollutant analyses for comparability. We allowed baseline mortality rates to differ according to sex, race, Medicaid eligibility, and 5-year categories of age at study entry. To adjust for potential confounding, we also obtained 15 ZIP-Code or county-level variables from various sources and a regional dummy variable to account for compositional differences in $PM_{2.5}$ across the United States (Table 1, and Section 1 in the Supplementary Appendix). We conducted this same statistical analysis but restricted it to person-years with $PM_{2.5}$ exposures lower than 12 μg per cubic meter and ozone exposures lower than 50 ppb (low-exposure analysis) (Table 1, and Section 1 in the Supplementary Appendix).

To identify populations at a higher or lower pollution-associated risk of death from any cause, we refit the same two-pollutant Cox model for some subgroups (e.g., male vs. female, white vs. black, and Medicaid eligible vs. Medicaid ineligible). To estimate the concentration-response function of air pollution and mortality, we fit a log-linear model with a thin-plate spline of both $PM_{2.5}$ and ozone and controlled for all the individual and ecologic variables used in our main analysis model (Section 7 in the Supplementary Appendix). To examine the robustness of our results, we conducted sensitivity analyses and compared the extent to which estimates of risk changed with respect to differences in confounding adjustment and estimation approaches (Sections S2 through S4 in the Supplementary Appendix).

Data on some important individual-level covariates were not available for the Medicare cohort, including data on smoking status, body-mass index (BMI), and income. We obtained data from the Medicare Current Beneficiary Survey (MCBS), a representative subsample of Medicare enrollees (133,964 records and 57,154 enrollees for the period 2000 through 2012), with individual-level data on smoking, BMI, income, and many other variables collected by means of telephone survey. Using MCBS data, we investigated how the lack of adjustment for these risk factors could have affected our calculated risk estimates in the Medicare cohort (Section 5 in the Supplementary Appendix). The computations in this article were run on the Odyssey cluster, which is supported by the FAS Division of Science, Research Computing Group, and on the Research Computing Environment, which is supported by the Institute for Quantitative Social Science in the Faculty of Arts and Sciences, both at Harvard University. We used R software, version 3.3.2 (R Project for Statistical Computing), and SAS software, version 9.4 (SAS Institute).

RESULTS

COHORT ANALYSES

The full cohort included 60,925,443 persons living in 39,716 different ZIP Codes with 460,310,521 person-years of follow-up. The median follow-up was 7 years. The total number of deaths was 22,567,924. There were 11,908,888 deaths and 247,682,367 person-years of follow-up when the $PM_{2.5}$ concentration was below 12 μg per cubic meter and 17,470,128 deaths and 353,831,836 person-years of follow-up when the ozone concentration was below 50 ppb. These data provided excellent power to estimate the risk of death at air-pollution levels below the current annual NAAQS for $PM_{2.5}$ and at low concentrations for ozone (Table 1).

Annual average $PM_{2.5}$ concentrations across the continental United States during the study period ranged from 6.21 to 15.64 μg per cubic meter (5th and 95th percentiles, respectively), and the warm-season average ozone concentrations ranged from 36.27 to 55.86 ppb (5th and 95th percentiles, respectively). The highest $PM_{2.5}$ concentrations were in California and the eastern and southeastern United States. The Mountain region and California had the highest ozone concentrations; the eastern states had lower ozone concentrations (Fig. 1).

Table 1. Cohort Characteristics and Ecologic and Meteorologic Variables.

Characteristic or Variable	Entire Cohort	Ozone Concentration		$PM_{2.5}$ Concentration	
		≥50 ppb*	<50 ppb	≥12 $\mu g/m^3$	<12 $\mu g/m^3$
Population					
Persons (no.)	60,925,443	14,405,094	46,520,349	28,145,493	32,779,950
Deaths (no.)	22,567,924	5,097,796	17,470,128	10,659,036	11,908,888
Total person-yr†	460,310,521	106,478,685	353,831,836	212,628,154	247,682,367
Median yr of follow-up	7	7	7	7	7
Average air-pollutant concentrations‡					
Ozone (ppb)	46.3	52.8	44.4	48.0	45.3
$PM_{2.5}$ ($\mu g/m^3$)	11.0	10.9	11.0	13.3	9.6
Individual covariates‡					
Male sex (%)	44.0	44.3	43.8	43.1	44.7
Race or ethnic group (%)§					
White	85.4	86.6	85.1	82.0	88.4
Black	8.7	7.2	9.2	12.0	5.9
Asian	1.8	1.8	1.8	2.1	1.6
Hispanic	1.9	2.0	1.9	1.9	1.9
Native American	0.3	0.6	0.3	0.1	0.6
Eligible for Medicaid (%)	16.5	15.3	16.8	17.8	15.3
Average age at study entry (yr)	70.1	69.7	70.2	70.1	70.0
Ecologic variables‡					
BMI	28.2	27.9	28.4	28.0	28.4
Ever smoked (%)	46.0	44.9	46.2	45.8	46.0
Population including all people 65 yr of age or older (%)					
Hispanic	9.5	13.4	8.4	8.4	10.0
Black	8.8	7.2	9.3	13.3	6.3
Median household income (1000s of $)	47.4	51.0	46.4	47.3	47.4
Median value of housing (1000s of $)	160.5	175.8	156.3	161.7	159.8
Below poverty level (%)	12.2	11.4	12.4	12.5	12.0
Did not complete high school (%)	32.3	30.7	32.7	35.3	30.6
Owner-occupied housing (%)	71.5	71.3	71.6	68.6	73.2
Population density (persons/km^2)	3.2	0.7	3.8	4.8	2.2
Low-density lipoprotein level measured (%)	92.2	92.0	92.2	92.2	92.2
Glycated hemoglobin level measured (%)	94.8	94.6	94.8	94.8	94.8
≥1 Ambulatory visits (%)¶	91.7	92.2	91.6	91.7	91.7
Meteorologic variables‡					
Average temperature (°C)	14.0	14.9	13.8	14.5	13.7
Relative humidity (%)	71.1	60.8	73.9	73.7	69.6

* Summary statistics were calculated separately for persons residing in ZIP Codes where average ozone levels were below or above 50 ppb and where $PM_{2.5}$ levels were below or above 12 μg per cubic meter. The value 12 μg per cubic meter was chosen as the current annual National Ambient Air Quality Standard (NAAQS) (e.g., the "safe" level) for $PM_{2.5}$. BMI denotes body-mass index (the weight in kilograms divided by the square of the height in meters) and ppb parts per billion.

† The number for total person-years of follow-up indicates the sum of individual units of time that the persons in the study population were at risk of death from 2000 through 2012.

‡ The average values for air pollution levels and for ecologic and meteorologic variables were computed by averaging values over all ZIP Codes from 2000 through 2012.

§ Data on race and ethnic group were obtained from Medicare beneficiary files.

¶ The variable for ambulatory visits refers to the average annual percentage of Medicare enrollees who had at least one ambulatory visit to a primary care physician.

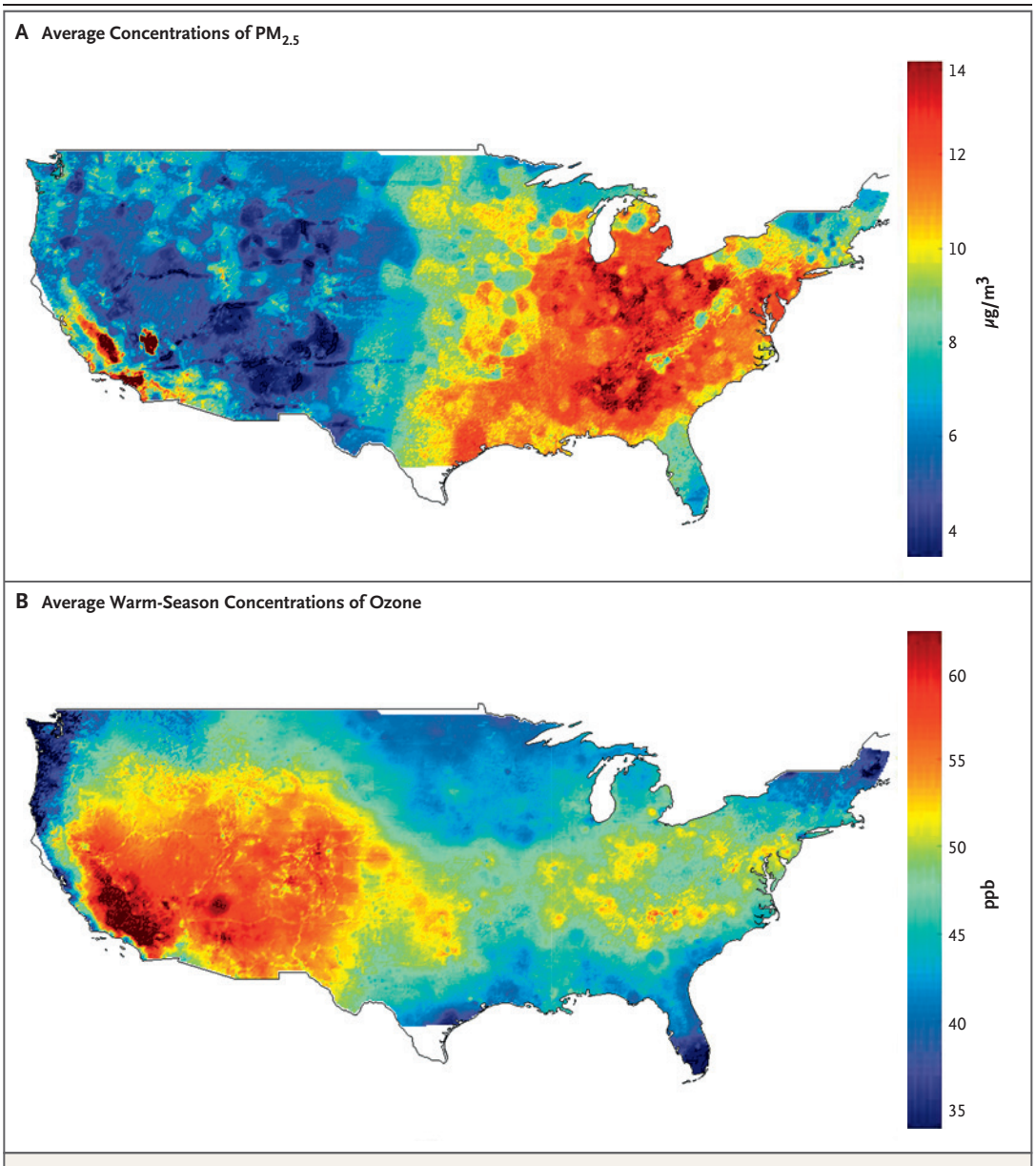


Figure 1. Average $PM_{2.5}$ and Ozone Concentrations in the Continental United States, 2000 through 2012.

Panel A shows the average concentrations of fine particulate matter (particles with a mass median aerodynamic diameter of less than 2.5 μm [$PM_{2.5}$]) in micrograms per cubic meter, as estimated on the basis of all daily predictions during the study period. Panel B shows the concentration of ozone levels in parts per billion as averaged from April 1 through September 30 throughout the study period.

In a two-pollutant analysis, each increase of 10 μg per cubic meter in annual exposure to $PM_{2.5}$ (estimated independently of ozone) and each increase of 10 ppb in warm-season exposure to ozone (estimated independently of $PM_{2.5}$) was associated with an increase in all-cause mortality of 7.3% (95% confidence interval [CI], 7.1 to 7.5) and 1.1% (95% CI, 1.0 to 1.2), respectively. Estimates of risk based on predictive, ZIP-Code–specific assessments of exposure were slightly higher than those provided by the nearest data-monitoring site (Table 2). When we restricted the $PM_{2.5}$ and ozone analyses to location-years with low concentrations, we continued to see significant associations between exposure and mortality (Table 2). Analysis of the MCBS

Table 2. Risk of Death Associated with an Increase of 10 μg per Cubic Meter in $PM_{2.5}$ or an Increase of 10 ppb in Ozone Concentration.*

Model	$PM_{2.5}$	Ozone
	hazard ratio (95% CI)	
Two-pollutant analysis		
Main analysis	1.073 (1.071–1.075)	1.011 (1.010–1.012)
Low-exposure analysis	1.136 (1.131–1.141)	1.010 (1.009–1.011)
Analysis based on data from nearest monitoring site (nearest-monitor analysis)†	1.061 (1.059–1.063)	1.001 (1.000–1.002)
Single-pollutant analysis‡	1.084 (1.081–1.086)	1.023 (1.022–1.024)

* Hazard ratios and 95% confidence intervals were calculated on the basis of an increase of 10 μg per cubic meter in exposure to $PM_{2.5}$ and an increase of 10 ppb in exposure to ozone.
† Daily average monitoring data on $PM_{2.5}$ and ozone were obtained from the Environmental Protection Agency Air Quality System. Daily ozone concentrations were averaged from April 1 through September 30 for the computation of warm-season averages. Data on $PM_{2.5}$ and ozone levels were obtained from the nearest monitoring site within 50 km. If there was more than one monitoring site within 50 km, the nearest site was chosen. Persons who lived more than 50 km from a monitoring site were excluded.
‡ For the single-pollutant analysis, model specifications were the same as those used in the main analysis, except that ozone was not included in the model when the main effect of $PM_{2.5}$ was estimated and $PM_{2.5}$ was not included in the model when the main effect of ozone was estimated.

subsample provided strong evidence that smoking and income are not likely to be confounders because they do not have a significant association with $PM_{2.5}$ or ozone (Section 5 in the Supplementary Appendix).

SUBGROUP ANALYSES

Subgroup analyses revealed that men; black, Asian, and Hispanic persons; and persons who were eligible for Medicaid (i.e., those who had low socioeconomic status) had a higher estimated risk of death from any cause in association with $PM_{2.5}$ exposure than the general population. The risk of death associated with ozone exposure was higher among white, Medicaid-eligible persons and was significantly below 1 in some racial subgroups (Fig. 2). Among black persons, the effect estimate for $PM_{2.5}$ was three times as high as that for the overall population (Table S3 in the Supplementary Appendix). Overall, the risk of death associated with ozone exposure was smaller and somewhat less robust than that associated with $PM_{2.5}$ exposure. We also detected a small but significant interaction between ozone exposure and $PM_{2.5}$ exposure (Table S8 in the Supplementary Appendix). Our thin-plate–spline fit indicated a relationship between $PM_{2.5}$, ozone, and all-cause mortality that was almost linear, with no signal of threshold down to 5 μg per cubic meter and 30 ppb, respectively (Fig. 3, and Fig. S8 in the Supplementary Appendix).

DISCUSSION

This study involving an open cohort of all persons receiving Medicare, including those from small cities and rural areas, showed that long-term exposures to $PM_{2.5}$ and ozone were associated with an increased risk of death, even at levels below the current annual NAAQS for $PM_{2.5}$. Furthermore, the study showed that black men and persons eligible to receive Medicaid had a much higher risk of death associated with exposure to air pollution than other subgroups. These findings suggest that lowering the annual NAAQS may produce important public health benefits overall, especially among self-identified racial minorities and people with low income.

The strengths of this study include the assessment of exposure with high spatial and temporal resolution, the use of a cohort of almost 61 million Medicare beneficiaries across the entire continental United States followed for up to 13 consecutive years, and the ability to perform subgroup analyses of the health effects of air pollution on groups of disadvantaged persons. However, Medicare claims do not include extensive individual-level data on behavioral risk fac-

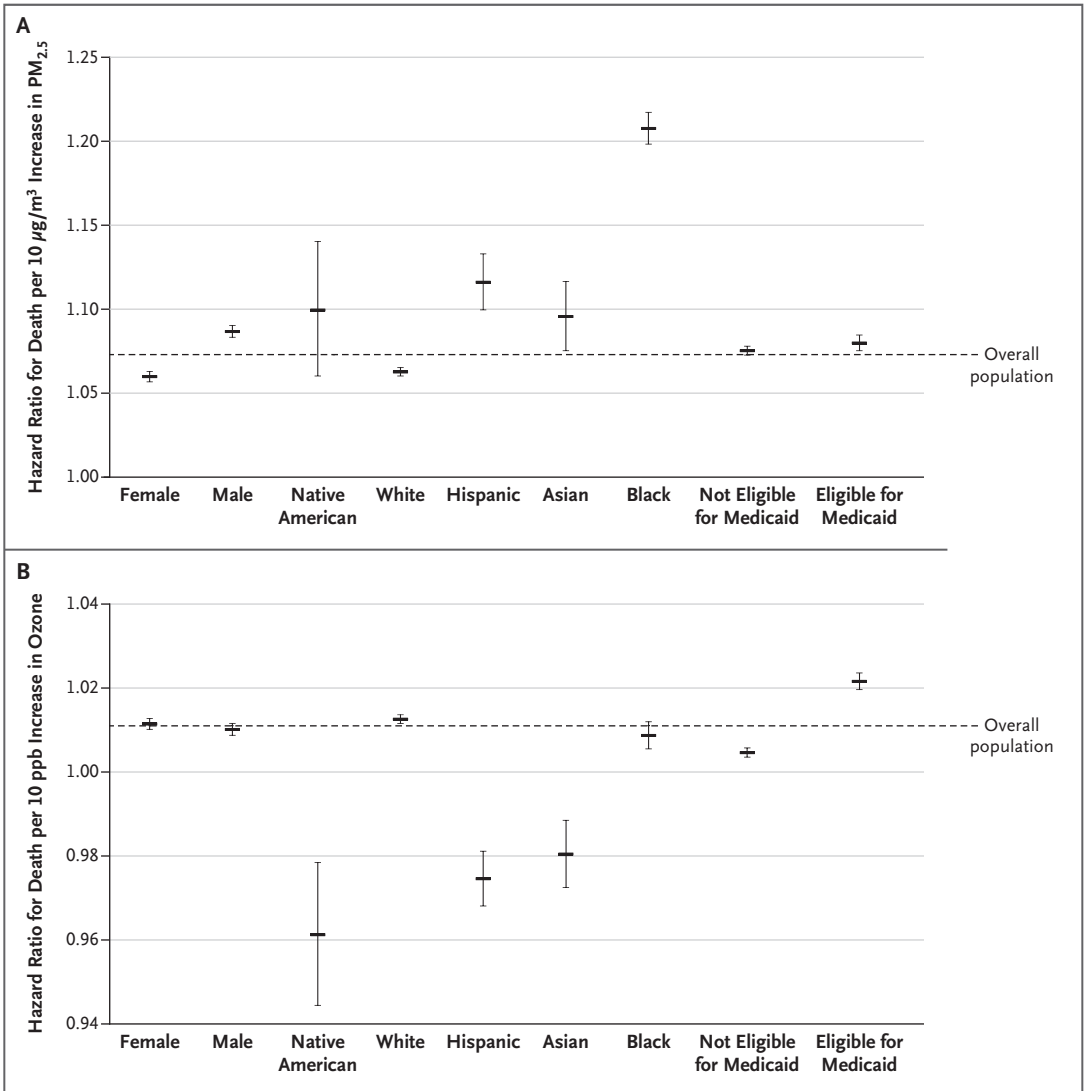


Figure 2. Risk of Death Associated with an Increase of 10 μg per Cubic Meter in $PM_{2.5}$ Concentrations and an Increase of 10 ppb in Ozone Exposure, According to Study Subgroups.

Hazard ratios and 95% confidence intervals are shown for an increase of 10 μg per cubic meter in $PM_{2.5}$ and an increase of 10 parts per billion (ppb) in ozone. Subgroup analyses were conducted by first restricting the population (e.g., considering only male enrollees). The same two-pollutant analysis (the main analysis) was then applied to each subgroup. Numeric results are presented in Tables S3 and S4 in the Supplementary Appendix. Dashed lines indicate the estimated hazard ratio for the overall population.

tors, such as smoking and income, which could be important confounders. Still, our analysis of the MCBS subsample (Table S6 in the Supplementary Appendix) increased our level of confidence that the inability to adjust for these individual-level risk factors in the Medicare cohort did not lead to biased results (Section 5 in the Supplementary Appendix). In another study, we analyzed a similar Medicare subsample with detailed individual-level data on smoking, BMI, and many other potential confounders linked to Medicare claims.[23] In that analysis, we found that for mortality and hospitalization, the risks of exposure to $PM_{2.5}$ were not sensitive to the additional control of individual-level variables that were not available in the whole Medicare population.

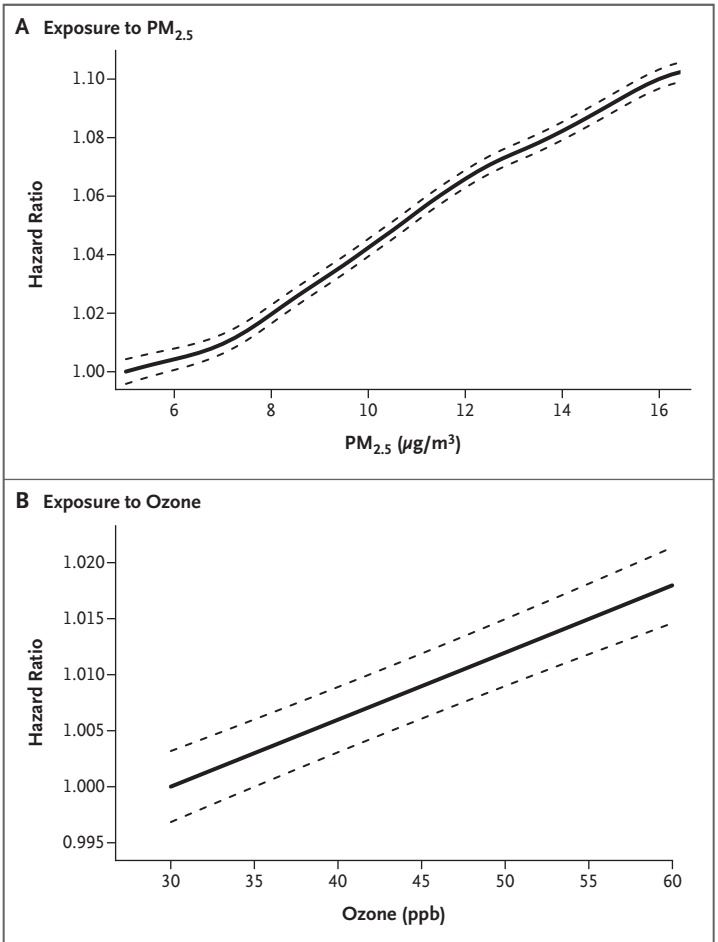


Figure 3. Concentration–Response Function of the Joint Effects of Exposure to $PM_{2.5}$ and Ozone on All-Cause Mortality.

A log-linear model with a thin-plate spline was fit for both $PM_{2.5}$ and ozone, and the shape of the concentration-response surface was estimated (Fig. S8 in the Supplementary Appendix). The concentration–response curve in Panel A was plotted for an ozone concentration equal to 45 ppb. The concentration–response curve in Panel B was plotted for a $PM_{2.5}$ concentration equal to 10 μg per cubic meter. These estimated curves were plotted at the 5th and 95th percentiles of the concentrations of $PM_{2.5}$ and ozone, respectively. The complete concentration–response three-dimensional surface is plotted in Fig. S8 in the Supplementary Appendix.

We also found that our results were robust when we excluded individual and ecologic covariates from the main analysis (Fig. S2 and Table S2 in the Supplementary Appendix), when we stratified age at entry into 3-year and 4-year categories rather than the 5 years used in the main analysis (Fig. S3 in the Supplementary Appendix), when we varied the estimation procedure (by means of a generalized estimating equation as opposed to mixed effects) (Tables S3 and S4 in the Supplementary Appendix), and when we used different types of statistical software (R, version 3.3.2, vs. SAS, version 9.4). Finally, we found that our results were consistent with others published in the literature (Section 6 in the Supplementary Appendix).[5,17,24-28]

There was a significant association between $PM_{2.5}$ exposure and mortality when the analysis was restricted to concentrations below 12 μg per cubic meter, with a steeper slope below that level. This association indicated that the health-benefit-per-unit decrease in the concentration of $PM_{2.5}$ is larger for $PM_{2.5}$ concentrations that are below the current annual NAAQS than the health benefit of decreases in $PM_{2.5}$ concentrations that are above that level. Similar, steeper concentration-response curves at low concentrations have been observed in previous studies.[29] Moreover, we found no evidence of a threshold value — the concentration at which $PM_{2.5}$ exposure does not affect mortality — at concentrations as low as approximately 5 μg per cubic meter (Fig. 3); this finding is similar to those of other studies.[18,30]

The current ozone standard for daily exposure is 70 ppb; there is no annual or seasonal standard. Our results strengthen the argument for establishing seasonal or annual standards. Moreover, whereas time-series studies have shown the short-term effects of ozone exposure, our results indicate that there are larger effect sizes for longer-term ozone exposure, including in locations where ozone concentrations never exceed 70 ppb. Unlike the American Cancer Society Cancer Prevention Study II,[9,10] our study reported a linear connection between ozone concentration and mortality. This finding is probably the result of the interaction between $PM_{2.5}$ and ozone (Section 7 in the Supplementary Appendix). The significant, linear relationship between seasonal ozone levels and all-cause mortality indicates that current risk assessments,[31-33] which incorporate only the acute effects of ozone exposure on deaths each day from respiratory mortality, may be substantially underestimating the contribution of ozone exposure to the total burden of disease.

The enormous sample size in this study, which includes the entire Medicare cohort, allowed for unprecedented accuracy in the estimation of risks among racial minorities and disadvantaged subgroups. The estimate of effect size for $PM_{2.5}$ expo-

sure was greatest among male, black, and Medicaid-eligible persons. We also estimated risks in subgroups of persons who were eligible for Medicaid and in whites and blacks alone to ascertain whether the effect modifications according to race and Medicaid status were independent. We found that black persons who were not eligible for Medicaid (e.g., because of higher income) continued to have an increased risk of death from exposure to $PM_{2.5}$ (Fig. S4 in the Supplementary Appendix). In addition, we found that there was a difference in the health effects of $PM_{2.5}$ exposure between urban and rural populations, a finding that may be due to compositional differences in the particulates (Table S3 Supplementary Appendix).

Although the Medicare cohort includes only the population of persons 65 years of age or older, two thirds of all deaths in the United States occur in people in that age group. Although our exposure models had excellent out-of-sample predictive power on held-out monitors, they do have limitations. Error in exposure assessment remains an issue in this type of analysis and could attenuate effect estimates for air pollution.[34]

The overall association between air pollution and human health has been well documented since the publication of the landmark Harvard Six Cities Study in 1993.[25] With air pollution declining, it is critical to estimate the health effects of low levels of air pollution — below the current NAAQS — to determine whether these levels are adequate to minimize the risk of death. Since the Clean Air Act requires the EPA to set air-quality standards that protect sensitive populations, it is also important to focus more effort on estimating effect sizes in potentially sensitive populations in order to inform regulatory policy going forward.

The views expressed in this article are those of the authors and do not necessarily represent the official views of the funding agencies. Furthermore, these agencies do not endorse the purchase of any commercial products or services related to this publication.

Supported by grants from the Health Effects Institute (4953-RFA14-3/16-4), the National Institutes of Health (R01 ES024332-01A1, ES-000002, ES024012, R01ES026217), the National Cancer Institute (R35CA197449), and the Environmental Protection Agency (83587201-0 and RD-83479801).

No potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Stacey C. Tobin, Ph.D., for editorial assistance on an earlier version of the manuscript, Sarah L. Duncan and William J. Horka for their support with the Research Computing Environment, and Ista Zahn at the Institute for Quantitative Social Science, Harvard University, for SAS programming support.

REFERENCES

1. Ambient (outdoor) air quality and health. Fact sheet no. 313. Updated March 2014. Geneva: World Health Organization, 2015.
2. Brook RD, Rajagopalan S, Pope CA III, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation 2010; 121:2331-78.
3. Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380:2224-60.
4. Crouse DL, Peters PA, Hystad P, et al. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect 2015;123:1180-6.
5. Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. Estimating causal effects of long-term PM2.5 exposure on mortality in New Jersey. Environ Health Perspect 2016;124:1182-8.
6. 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;383:785-95.
7. Atkinson RW, Butland BK, Dimitroulopoulou C, et al. Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies. BMJ Open 2016;6(2):e009493.
8. Hao Y, Balluz L, Strosnider H, Wen XJ, Li C, Qualters JR. Ozone, fine particulate matter, and chronic lower respiratory disease mortality in the United States. Am J Respir Crit Care Med 2015;192:337-41.
9. Turner MC, Jerrett M, Pope CA III, et al. Long-term ozone exposure and mortality in a large prospective study. Am J Respir Crit Care Med 2016;193:1134-42.
10. Jerrett M, Burnett RT, Pope CA III, et al. Long-term ozone exposure and mortality. N Engl J Med 2009;360:1085-95.
11. Krewski D, Jerrett M, Burnett RT, et al. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. Res Rep Health Eff Inst 2009;140:5-114, discussion 115-136.
12. Carey IM, Atkinson RW, Kent AJ, van Staa T, Cook DG, Anderson HR. Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. Am J Respir Crit Care Med 2013; 187:1226-33.
13. Ostro B, Hu J, Goldberg D, et al. Associations of mortality with long-term exposures to fine and ultrafine particles, species and sources: results from the California Teachers Study Cohort. Environ Health Perspect 2015;123:549-56.
14. Crouse DL, Peters PA, van Donkelaar A, et al. Risk of nonaccidental and cardiovascular mortality in relation to long-term exposure to low concentrations of fine particulate matter: a Canadian national-level cohort study. Environ Health Perspect 2012;120:708-14.
15. Wang Y, Shi L, Lee M, et al. Long-term exposure to PM2.5 and mortality among older adults in the southeastern US. Epidemiology 2017;28:207-14.
16. Thurston GD, Ahn J, Cromar KR, et al. Ambient particulate matter air pollution exposure and mortality in the NIH-AARP Diet and Health cohort. Environ Health Perspect 2016;124:484-90.
17. Pinault L, Tjepkema M, Crouse DL, et al.

Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian Community Health Survey cohort. Environ Health 2016;15:18.
18. Shi L, Zanobetti A, Kloog I, et al. Low-concentration PM2.5 and mortality: estimating acute and chronic effects in a population-based study. Environ Health Perspect 2016;124:46-52.
19. Di Q, Kloog I, Koutrakis P, Lyapustin A, Wang Y, Schwartz J. Assessing PM2.5 exposures with high spatiotemporal resolution across the continental United States. Environ Sci Technol 2016;50:4712-21.
20. Di Q, Rowland S, Koutrakis P, Schwartz J. A hybrid model for spatially and temporally resolved ozone exposures in the continental United States. J Air Waste Manag Assoc 2017;67:39-52.
21. Kalnay E, Kanamitsu M, Kistler R, et al. The NCEP/NCAR 40-Year Reanalysis Project. Bull Am Meteorol Soc 1996;77:437-71.
22. Lee EW, Wei L, Amato DA, Leurgans S. Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In: Klein JP, Goel PK, eds. Survival analysis: state of the art. Berlin: Springer, 1992:237-47.
23. Makar M, Antonelli J, Di Q, Cutler D, Schwartz J, Dominici F. Estimating the causal effect of low levels of fine particulate matter on hospitalization. Epidemiology, May 25, 2016 (http://journals.lww.com/epidem/Abstract/publishahead/Estimating_the_Causal_Effect_of_Low_Levels_of_Fine.98844.aspx).
24. Kioumourtzoglou MA, Schwartz J, James P, Dominici F, Zanobetti A. PM2.5 and mortality in 207 US cities: modification by temperature and city characteristics. Epidemiology 2016;27:221-7.
25. Dockery DW, Pope CA III, Xu X, et al. An association between air pollution and mortality in six U.S. cities. N Engl J Med 1993;329:1753-9.
26. Lepeule J, Laden F, Dockery D, Schwartz J. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. Environ Health Perspect 2012;120:965-70.
27. Pope CA III, Burnett RT, Thurston GD, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. Circulation 2004;109:71-7.
28. Eftim SE, Samet JM, Janes H, McDermott A, Dominici F. Fine particulate matter and mortality: a comparison of the six cities and American Cancer Society cohorts with a Medicare cohort. Epidemiology 2008;19:209-16.
29. Pope CA III, Burnett RT, Krewski D, et al. Cardiovascular mortality and exposure to airborne fine particulate matter and cigarette smoke: shape of the exposure-response relationship. Circulation 2009;120:941-8.
30. Schwartz J, Coull B, Laden F, Ryan L. The effect of dose and timing of dose on the association between airborne particles and survival. Environ Health Perspect 2008;116:64-9.
31. Smith RL, Xu B, Switzer P. Reassessing the relationship between ozone and short-term mortality in U.S. urban communities. Inhal Toxicol 2009;21:Suppl 2:37-61.
32. Zanobetti A, Schwartz J. Mortality displacement in the association of ozone with mortality: an analysis of 48 cities in the United States. Am J Respir Crit Care Med 2008;177:184-9.
33. Regulatory impact analysis of the final revisions to the National Ambient Air Quality Standards for ground-level ozone. Research Triangle Park, NC: Environmental Protection Agency, 2015 (https://www.epa.gov/naaqs/regulatory-impact-analysis-final-revisions-national-ambient-air-quality-standards-ground-level).
34. Spiegelman D. Evaluating public health interventions. 4. The Nurses' Health Study and methods for eliminating bias attributable to measurement error and misclassification. Am J Public Health 2016;106:1563-6.

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Di Q, Wang Y, Zanobetti A, et al. Air pollution and mortality in the Medicare population. N Engl J Med 2017;376:2513-22. DOI: 10.1056/NEJMoa1702747

Supplementary Material for

Air Pollution and Mortality in the Medicare Population

Qian Di, M.S.,[†] Yan Wang, M.S.,[†] Antonella Zanobetti, Ph.D.,[†] Yun Wang, Ph.D.,[§] Petros Koutrakis, Ph.D.,[†] Christine Choirat, Ph.D.,[§] Francesca Dominici, Ph.D.,[§*] Joel D Schwartz, Ph.D.,[†]

[†] Department of Environmental Health, Harvard T.H. Chan School of Public Heath, Boston, MA, 02115, USA

[§] Department of Biostatistics, Harvard T.H. Chan School of Public Heath, Boston, MA, 02115, USA

* Corresponding author: Francesca Dominici, Phone: 410-258-5886, Email: fdominic@hsph.harvard.edu

SECTION 1: Details Regarding Confounding Adjustment by Individual-level and Area-level Covariates

SECTION 2: Sensitivity Analysis with Respect to Variables Included in Confounding Adjustment

SECTION 3: Sensitivity Analysis with Respect to the Categorization of Age at Entry

SECTION 4: Sensitivity Analysis with Respect to the Estimation Approach and Statistical Software

SECTION 5: Sensitivity Analysis with Respect to Lack of Adjustment for Individual-Level Behavioral Risk Factors

SECTION 6: Comparison of Our Results with Others in the Literature

SECTION 7: Concentration-Response Function

1. Details Regarding Confounding Adjustment by Individual-level and Area-level Covariates

In our main analysis, we adjusted for **20** covariates. These includes **four** individual level covariates; **two** county-level variables from the Behavioral Risk Factor Surveillance System (BRFSS); **eight** ZIP code-level variables from U.S. Census; **three** hospital service area-level variables from Dartmouth Atlas of Health Care; **two** meteorological variables; and **one** dummy variable indicating geographical regions. Except for the individual covariates, **16** area-level covariates included in our main analysis were denoted as ecological variables. Details regarding the definition of each of these variables, and how they were linked to the mortality data are described below.

Individual-level variables: We considered a 5-year category of age at entry (65 to 69, 70 to 74, 75 to 79, 80 to 84, 85 to 89, 90 to 94, 95 to 99, and above 100), race (White, Black, Asian, Hispanic, Native American, and other), sex (male or female) and a dummy variable for eligibility for Medicaid. Our sensitivity analyses of the Medicare Current Beneficiary Survey (MCBS) provided strong evidence that Medicaid eligibility is an excellent proxy for individual-level income in our population (see Section 4 for details).

ZIP code-level variables: We acquired data at the ZIP Code Tabulation Areas (ZCTA)-level from the 2000 U.S. Census, the 2010 U.S. Census, and from the American Community Survey (ACS) for each year from 2005 to 2012. Not all variables were available for all years, and we linearly interpolated them between two available years. We matched data from ZCTA to ZIP code and manually resolved some minor differences between ZCTA and ZIP codes. ZIP code-level variables from the Census included: percentage Hispanic, percentage Black, median household income, median value of owner-occupied housing, percentage above age 65 living below the poverty level, percentage above age of 65 with less than high school education, percentage of owner-occupied housing units, and population density.

County-level variables: We acquired county-level body mass index (BMI) and percentage of ever smokers from BRFSS, for each year from 2000 to 2012.[1] We assigned the same values of these county-level variables to all ZIP codes that fell within the county boundary.

Hospital service area-level variables: We acquired hospital service area-level variables from the Dartmouth Atlas of Health Care, for all available years.[2] We considered the following variables: percentage of Medicare enrollees having: 1) a blood lipid (LDL-C) test, 2) a hemoglobin A1c test, and 3) at least one ambulatory visit to a primary care clinician. We used existing crosswalk files provided by the Dartmouth Atlas of Health Care to match data from hospital service area to ZIP code.

Gridded weather and air pollution variables: We acquired daily 32 km × 32 km gridded temperature and humidity data from the North American Regional Reanalysis data.[3] We also acquired daily 1 km × 1 km gridded air pollution levels ($PM_{2.5}$ and ozone) from previously developed and validated air pollution prediction models.[4,5] We obtained ZIP code-level variables by taking inverse-distance averages of the four nearest grid cells to the ZIP code's centroid and then computed the annual averages for temperature, humidity, and $PM_{2.5}$, and the warm-season (from April 1 to September 30) average for ozone.

Monitor level air pollution variables: We acquired air pollution monitoring data from the U.S. EPA Air Quality System (1,928 monitors for $PM_{2.5}$ and 1,877 monitors for ozone).[6] We first obtained $PM_{2.5}$ annual average and daily 8-hour maximal ozone. We computed warm-season ozone for each monitoring site by averaging the daily ozone measurements from April 1 to September 30. To join monitoring data to each residential ZIP code, we identified the nearest monitoring site within 50 km of the ZIP code (based on centroid point) and assigned air pollutant measurements to that ZIP code. If there was more than one monitoring site, we chose the nearest

one; if there were no monitoring site within 50 km, we treated the monitored exposure level as missing and excluded that ZIP code from the analysis.

Regional dummy variable: To categorize ZIP codes into regions, we first simulated concentrations of five major chemical components of $PM_{2.5}$: sulfate, nitrate, organic carbon, elemental carbon, and ammonium, using GEOS-Chem, a 3D global chemical transport model.[7] Long-term averaged concentrations of the five $PM_{2.5}$ components were linearly interpolated to each ZIP code. Then we calculated the percentage of each $PM_{2.5}$ component with respect to the total $PM_{2.5}$ mass. We used k-mean clustering to classify all ZIP codes into 10 geographical regions based on the percentage of these five $PM_{2.5}$ components. ZIP codes that share a similar chemical profile of $PM_{2.5}$ were assigned to the same geographical region (Figure S1).

Table S1 summarizes the Pearson correlation coefficients between each pollutant and ecological covariate.

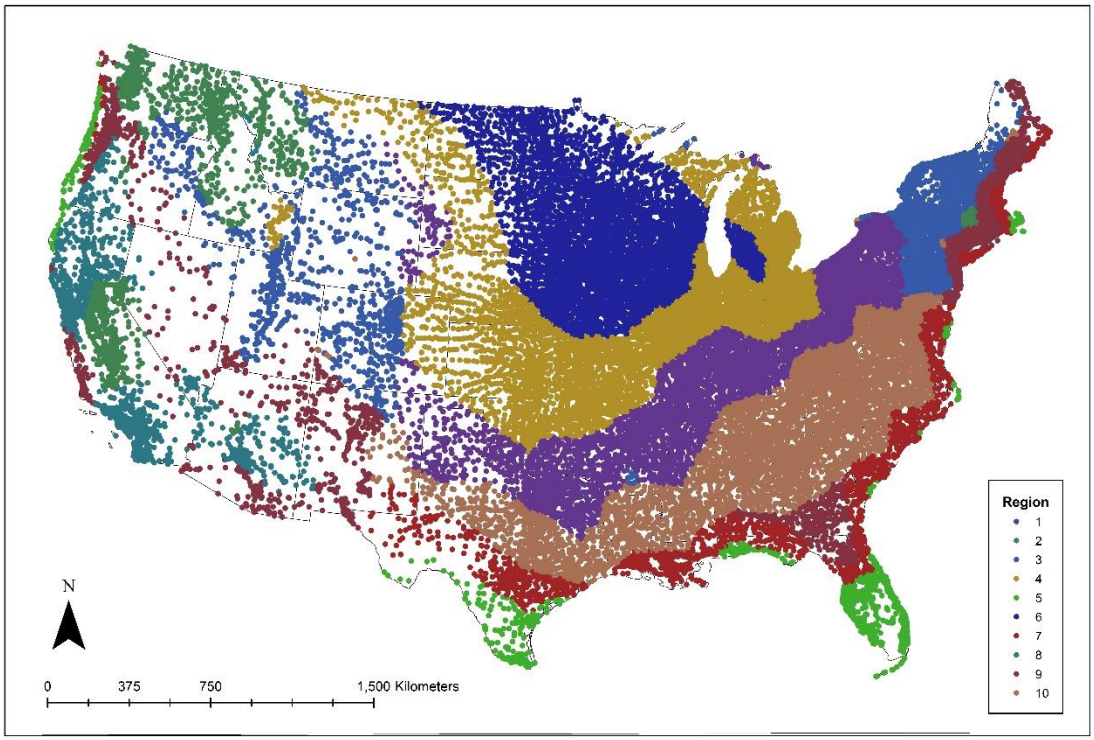


Figure S1. Regional Dummy Variable

ZIP codes with the same color belong to the same region and share the same value of the regional dummy variable.

Table S1. Pearson Correlation between Air Pollution, Ecological Variables, and Meteorological Variables*

Variable	$PM_{2.5}$	Ozone
Air Pollutants		
$PM_{2.5}$ ($\mu g/m^3$)	1.000	0.239
Ozone (ppb)	0.239	1.000
County-level variables --- from BRFSS		
BMI (kg/m^2)	-0.149	0.022
Ever Smoker (%)	-0.055	-0.096
ZIP code-level variables --- from US Census		
Hispanic Population (%)	-0.018	0.050
Black Population (%)	0.207	-0.042
Median household income (US dollars)	-0.049	-0.029
Median value of housing (US dollars)	-0.042	-0.099
% below poverty level	0.020	-0.010
% below high school education	0.219	0.089
% of owner occupied housing	-0.143	0.076
Population density (person/ km^2)	0.007	-0.020
Hospital service area level variables --- from Dartmouth Atlas of Health Care		
% with LDL-C test	-0.085	-0.033
% with hemoglobin A1c test	-0.070	-0.045
% with ≥1 ambulatory visit	0.055	0.031
Meteorological variables		

Temperature (°C)	0.131	0.110
Relative humidity (%)	0.286	-0.508

* Pearson correlation was computed for every pair of variables for each year (from 2000 to 2012) across the 40,177 ZIP codes.

2. Sensitivity Analysis with Respect to Variables Included in Confounding Adjustment

We conducted a sensitivity analysis to assess the robustness of our results to different sets of variables included in the Cox proportional hazards model for the confounding adjustment. Starting from the main analysis that included 20 variables, we considered several alternative models; each of these models exclude a different set of variables (e.g., excluding individual covariates or excluding meteorological variables, etc.). We compared models fit at various levels of adjustment and the estimated hazard ratios. Various levels of adjustment allowed us to evaluate the impact of ecological and individual covariates and judge the direction of potential biases of omitting individual covariates.

Table S2 displays the AIC and -2*log likelihood values corresponding to Cox models that exclude different subsets of covariates. Both AIC values and likelihood ratio tests indicate that the main analysis provides better fit to the data than all sensitivity analyses that exclude some of these variables.

Figure S2 shows the estimated HR and 95% confidence intervals under the different model specifications for confounding adjustment. Results are presented for both $PM_{2.5}$ (also adjusted by ozone) and ozone (also adjusted by $PM_{2.5}$). The vertical line is placed at the estimated HR from the main analysis (which includes all 20 variables). Risk estimates change moderately after excluding regional dummy variables or U.S. Census ZIP code-level variables, but are very robust to the omission of other sets of variables. This is expected as regional dummy variables explain a large amount of spatial variation in the air pollution exposure and mortality rates.

Table S2. -2* log Likelihood and AIC Values at Different Levels of Adjustment for Confounding

Name[†]	AIC	-2*log likelihood	df	p[‡]
Main analysis	615404575	615404523	26	
Main analysis excluding sex	645833387	645833335	26	N/A
Main analysis excluding Medicaid eligibility	639509036	639508984	26	N/A
Main analysis excluding race	637761029	637760977	26	N/A
Main analysis excluding regional dummy	615413233	615413199	17	<0.001
Main analysis excluding meteorological variables	615405208	615405160	24	<0.001
Main analysis excluding BRFSS	615409354	615409306	24	<0.001
Main analysis excluding Dartmouth	615406339	615406293	23	<0.001
Main analysis excluding U.S. Census	615480804	615480768	18	<0.001

[†] For different levels of adjustment, we started from our main analysis and omitted sex, Medicaid eligibility, race, regional dummy variables, meteorological variables, BRFSS county-level variables, ecological variables from Dartmouth Atlas of Health Care, or ZIP code-level variables from the U.S. Census.

[‡] p-values were based on likelihood ratio test.

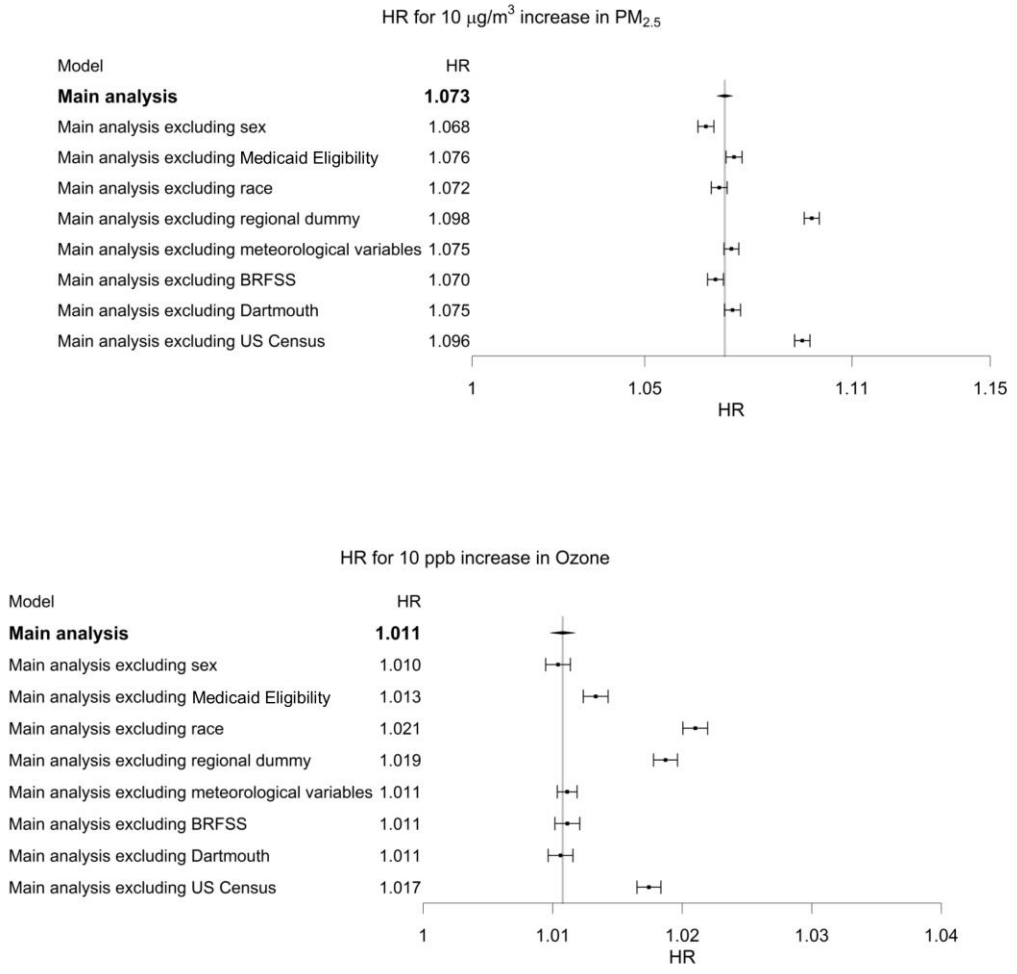


Figure S2. Estimated Risk of Death Associated with $PM_{2.5}$ and Ozone Exposure at Different Levels of Adjustment.

Vertical lines are placed at the estimated HR obtained from the main analysis.

3. Sensitivity Analysis with Respect to the Categorization of Age at Entry

In our main analysis, we considered age at entry in the Medicare cohort categorized into 5-year intervals: 65 to 69, 70 to 74, 75 to 79, 80 to 84, 85 to 89, 90 to 94, 95 to 99, and above 100. We conducted a sensitivity analysis where we re-fit our models using 4-year and 3-year intervals. For $PM_{2.5}$, when we consider finer age groups at entry the mortality risk estimates were lower (1.07 to 1.05) but still significant. For ozone, this finer stratification had little impact on the risk estimate (Figure S3).

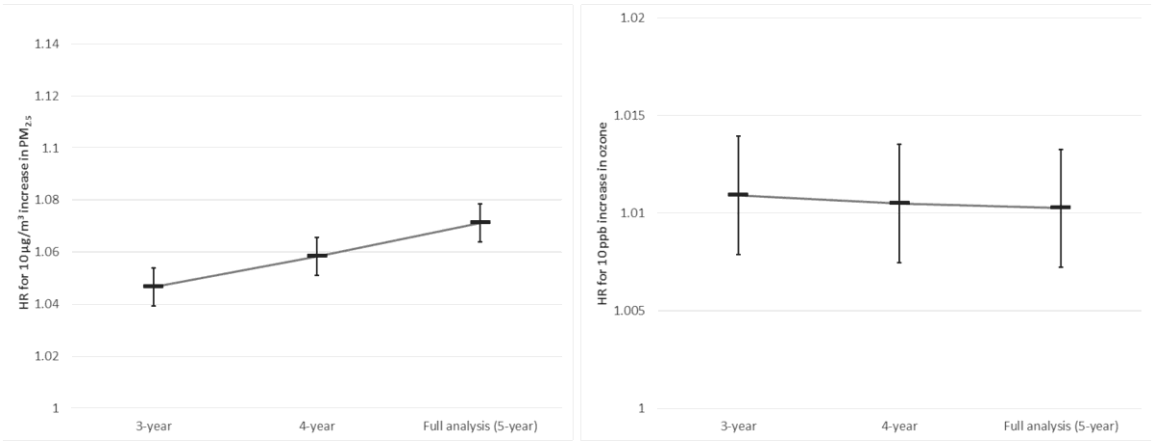


Figure S3. Risk of Death Associated with $PM_{2.5}$ and Ozone Exposure at Different Age Groups at Entry

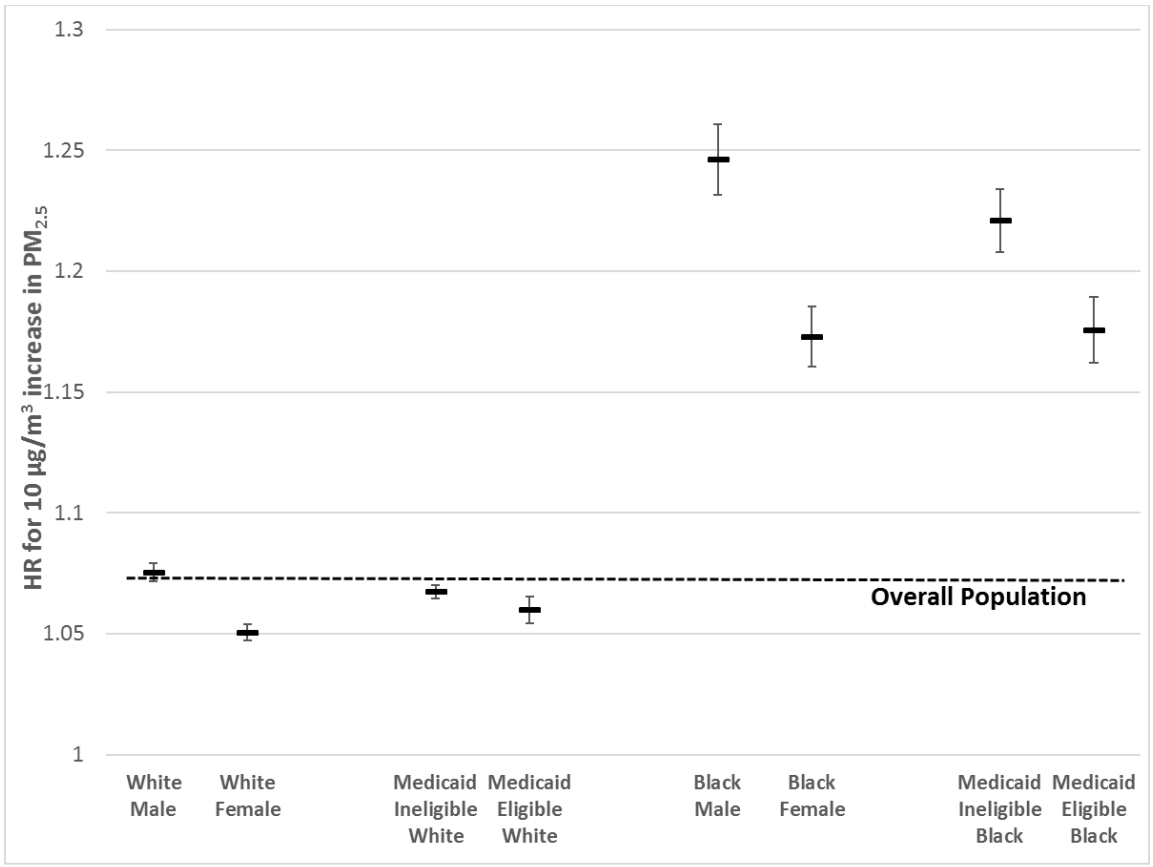
The main analysis stratified by 5-year category of age at entry. As a sensitivity analysis examining the impact of categorization of entry age, we modified the main analysis by stratifying by 3-year and 4-year categories of entry age. Two figures visualize risk estimates associated with each 10 $\mu g/m^3$ increase in $PM_{2.5}$ and 10 ppb increase in ozone. Running Cox model with exceeding numbers of strata (e.g., stratifying by 1-year category of entry age) on the whole data set was computationally infeasible.

4. Sensitivity Analysis with Respect to the Estimation Approach and Statistical Software

We joined data and ran our main analysis using a Cox proportional hazard model with Generalized Estimating Equation (GEE) to account for correlated measures. We fit this model in SAS 9.4 to the whole data set (460.3 million records and 60.9 million subjects) as well to population subgroups (Figure S4). To assess the robustness of our risk estimates to both the estimation approach and statistical software, we repeated the main analysis by fitting a Cox proportional hazards model with the same model specification as our main analysis (20 covariates), but with a random intercept at the ZIP code level instead of accounting for correlation using GEE. We joined data in R version 3.3.2 and implemented a mixed-effect Cox model with the *coxme* package version 2.2-5.[8] We compared the risk estimates from the mixed-effect Cox model with those from GEE.

Running a mixed-effect Cox proportional hazards model on the whole study population was computationally infeasible. Instead, we randomly divided all subjects into 50 groups with equal probability. We conducted our analysis in each group separately and used meta-analysis to obtain summary results. We pooled the point estimates (i.e., beta coefficients from Cox models) and corresponding standard errors from the 50 groups using a fixed-effect meta-analysis.

Table S3 and Table S4 summarize risk estimates obtained under the two statistical approaches. We found that the two sets of estimates are almost identical, thus increasing our level of confidence regarding reproducibility of results with respect to the assumptions of correlations, statistical estimation, and software. All of our computer programs are hosted on the GitHub social coding platform and are accessible upon request.


1.3
1.25
1.2
1.15
1.1
1.05
1
HR for 10 µg/m³ increase in PM2.5
Overall Population
White Male
White Female
Medicaid Ineligible White
Medicaid Eligible White
Black Male
Black Female
Medicaid Ineligible Black
Medicaid Eligible Black

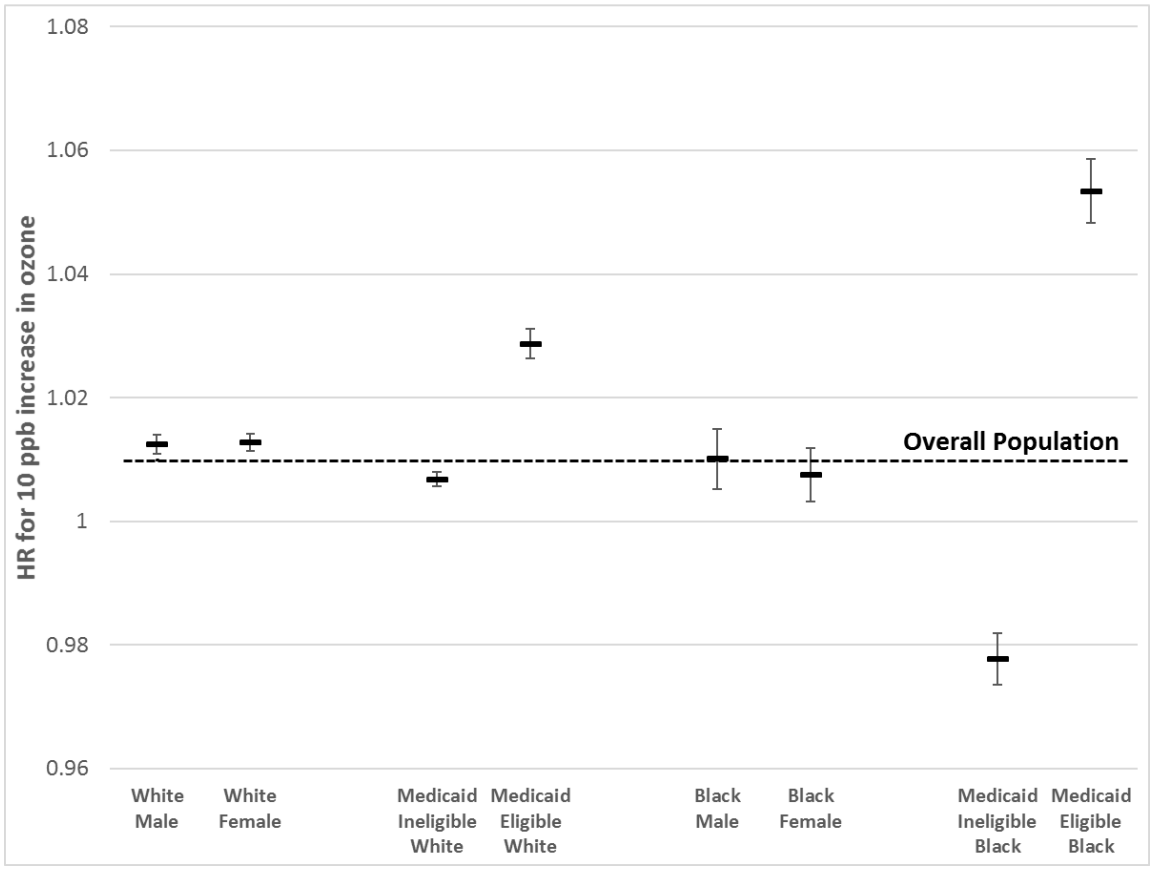

1.08
1.06
1.04
1.02
1
0.98
0.96
HR for 10 ppb increase in ozone
Overall Population
White Male
White Female
Medicaid Ineligible White
Medicaid Eligible White
Black Male
Black Female
Medicaid Ineligible Black
Medicaid Eligible Black

Figure S4. Risk of Death Associated with $PM_{2.5}$ and Ozone Exposure in Subgroups

Hazard ratios (HRs) are presented for each 10 $\mu g/m^3$ increase in $PM_{2.5}$ and 10 ppb increase in ozone. Subgroup analyses were conducted by first restricting the population (e.g., considering only male enrollees) and then the same model specification as the main analysis was applied to each subgroup. Numeric results are presented in Table S3 and Table S4. Dashed line indicates the HR for the overall population.

Table S3. Risk of Death Associated with 10 μg/m^3 Increase in $PM_{2.5}$

Analysis	$PM_{2.5}$ (GEE)	$PM_{2.5}$ (coxme)	Effect Modification[§]
Main Analysis	1.073 (1.071, 1.075)	1.081 (1.078, 1.083)	
Low-Exposure Analysis[*]	1.136 (1.131, 1.141)	1.134 (1.129, 1.139)	
Single-Pollutant Analysis[†]	1.084 (1.081, 1.086)	1.089 (1.087, 1.091)	
Nearest-Monitor Analysis[‡]	1.061 (1.059, 1.063)	1.072 (1.069, 1.074)	
By Sex			
Male	1.087 (1.083, 1.090)	1.089 (1.086, 1.093)	Ref
Female	1.060 (1.057, 1.063)	1.062 (1.058, 1.065)	<0.001
By Medicaid Eligibility			
Non-eligible	1.075 (1.073, 1.078)	1.079 (1.076, 1.082)	Ref
Eligible	1.080 (1.075, 1.085)	1.089 (1.084, 1.094)	0.092
By Race			
White	1.063 (1.060, 1.065)	1.068 (1.065, 1.070)	Ref
Black	1.208 (1.199, 1.217)	1.216 (1.206, 1.225)	<0.001
Asian	1.096 (1.075, 1.117)	1.140 (1.116, 1.164)	0.002
Hispanic	1.116 (1.100, 1.133)	1.127 (1.109, 1.144)	<0.001
Native Americans	1.100 (1.060, 1.140)	1.145 (1.090, 1.203)	0.067
By Age Groups			
<75	1.147 (1.142, 1.152)	1.187 (1.183, 1.192)	Ref
75 to 84	1.029 (1.025, 1.032)	1.071 (1.067, 1.074)	<0.001
≥85	0.998 (0.994, 1.002)	1.024 (1.020, 1.027)	<0.001
By Population Density			
Population Density (low)	1.067 (1.063, 1.072)	1.065 (1.061, 1.069)	Ref
Population Density (medium-low)	1.105 (1.100, 1.111)	1.131 (1.126, 1.136)	<0.001
Population Density (medium-high)	1.098 (1.093, 1.104)	1.117 (1.112, 1.123)	<0.001
Population Density (high)	1.080 (1.074, 1.085)	1.144 (1.139, 1.150)	<0.001
Among White			
By Sex			
White Male	1.075 (1.072, 1.079)	1.077 (1.073, 1.080)	Ref
White Female	1.051 (1.047, 1.054)	1.051 (1.047, 1.054)	<0.001
By Medicaid Eligibility			
Non-eligible White	1.067 (1.065, 1.070)	1.070 (1.067, 1.073)	Ref
Eligible White	1.060 (1.055, 1.065)	1.063 (1.057, 1.068)	0.015
Among Black			
By Sex			
Black Male	1.246 (1.232, 1.261)	1.249 (1.234, 1.264)	Ref
Black Female	1.173 (1.161, 1.185)	1.178 (1.165, 1.190)	<0.001
By Medicaid Eligibility			
Non-eligible Black	1.221 (1.208, 1.234)	1.226 (1.212, 1.239)	Ref

Eligible Black	1.176 (1.162, 1.189)	1.182 (1.168, 1.196)	<0.001
Among Male			
By Medicaid Eligibility			
Non-eligible Male	1.097 (1.093, 1.101)	1.096 (1.092, 1.100)	Ref
Eligible Male	1.076 (1.068, 1.084)	1.080 (1.071, 1.088)	<0.001
Among Female			
By Medicaid Eligibility			
Non-eligible Female	1.052 (1.049, 1.056)	1.054 (1.050, 1.057)	Ref
Eligible Female	1.083 (1.077, 1.088)	1.081 (1.075, 1.087)	<0.001

[*] Low-exposure analysis used the same model specifications as the main analysis, with $PM_{2.5}$ concentrations constrained to below 12 μg/m^3.

[†] For the single-pollutant analysis, model specifications were the same as in the main analysis, except that ozone was not included when assessing the main effect of $PM_{2.5}$.

[‡] Daily $PM_{2.5}$ monitoring data were retrieved from the U.S. EPA Air Quality System (AQS) and averaged for the whole year. Subjects were assigned to the $PM_{2.5}$ levels from the nearest monitoring site within 50 kilometers. If there was more than one monitoring site, the nearest one was chosen. Subjects who lived more than 50 kilometers away from any monitoring site were excluded.

[§] To determine the risk estimates of $PM_{2.5}$, for example in male vs. females, are statistically different (H_0: $\beta_{male} = \beta_{female}$), we have: $Z = \frac{\beta_{male} - \beta_{female}}{\sqrt{se(\beta_{male})^2 + se(\beta_{female})^2}}$

Table S4. Risk of Death Associated with 10 ppb Increase in Ozone

Analysis	Ozone (GEE)	Ozone (coxme)	Effect Modification[§]
Main Analysis	1.011 (1.010, 1.012)	1.009 (1.008, 1.010)	
Low-Exposure Analysis[*]	1.010 (1.009, 1.011)	1.008 (1.006, 1.009)	
Single-Pollutant Analysis[†]	1.023 (1.022, 1.024)	1.022 (1.021, 1.023)	
Nearest-Monitor Analysis[‡]	1.001 (1.000, 1.002)	1.000 (0.999, 1.001)	
By Sex			
Male	1.010 (1.009, 1.012)	1.008 (1.007, 1.010)	Ref
Female	1.011 (1.010, 1.013)	1.010 (1.008, 1.011)	0.181
By Medicaid Eligibility			
Non-eligible White	1.005 (1.004, 1.006)	1.004 (1.003, 1.005)	Ref
Eligible White	1.022 (1.020, 1.024)	1.019 (1.017, 1.021)	<0.001
By Race			
White	1.013 (1.012, 1.014)	1.011 (1.010, 1.012)	Ref
Black	1.009 (1.005, 1.012)	1.006 (1.003, 1.009)	0.026
Asian	0.980 (0.972, 0.988)	0.967 (0.958, 0.976)	<0.001
Hispanic	0.975 (0.968, 0.981)	0.971 (0.964, 0.977)	<0.001
Native Americans	0.961 (0.944, 0.978)	0.951 (0.928, 0.975)	<0.001
By Age Groups			
<75	1.012 (1.010, 1.014)	1.007 (1.005, 1.009)	Ref
75 to 84	1.004 (1.002, 1.005)	1.017 (1.016, 1.019)	<0.001
≥85	1.015 (1.013, 1.016)	1.024 (1.023, 1.026)	0.061
By population Density			
Population Density (low)	1.029 (1.027, 1.031)	1.038 (1.036, 1.040)	Ref
Population Density (medium-low)	1.006 (1.004, 1.008)	1.007 (1.004, 1.009)	<0.001
Population Density (medium-high)	0.997 (0.995, 0.999)	1.001 (0.999, 1.003)	<0.001
Population Density (high)	0.983 (0.981, 0.985)	0.986 (0.984, 0.988)	<0.001
Among White			
By Sex			
White Male	1.012 (1.011, 1.014)	1.011 (1.009, 1.013)	Ref
White Female	1.013 (1.011, 1.014)	1.011 (1.010, 1.013)	0.795
By Medicaid Eligibility			
Non-eligible White	1.007 (1.006, 1.008)	1.006 (1.005, 1.007)	Ref
Eligible White	1.029 (1.026, 1.031)	1.026 (1.024, 1.029)	<0.001
Among Black			
By Sex			
Black Male	1.010 (1.005, 1.015)	1.007 (1.002, 1.012)	Ref
Black Female	1.008 (1.003, 1.012)	1.006 (1.002, 1.011)	0.443
By Medicaid Eligibility			

Non-eligible Black	0.978 (0.974, 0.982)	0.977 (0.973, 0.981)	Ref
Eligible Black	1.053 (1.048, 1.059)	1.049 (1.044, 1.054)	<0.001
Among Male			
By Medicaid Eligibility			
Non-eligible Male	1.006 (1.004, 1.007)	1.005 (1.003, 1.006)	Ref
Eligible Male	1.018 (1.015, 1.021)	1.016 (1.012, 1.019)	<0.001
Among Female			
By Medicaid Eligibility			
Non-eligible Female	1.004 (1.002, 1.005)	1.003 (1.002, 1.005)	Ref
Eligible Female	1.024 (1.021, 1.026)	1.021 (1.018, 1.023)	<0.001

* Low-exposure analysis used the same model specifications as the main analysis, with ozone concentrations constrained to below 50 ppb.

† For the single-pollutant analysis, model specifications were the same as the main analysis, except that $PM_{2.5}$ was not included when assessing the main effect of ozone.

‡ Daily ozone monitoring data were retrieved from the U.S. EPA Air Quality System (AQS). Daily ozone concentrations were averaged from April 1 to September 30 to compute the warm-season average. Subjects were assigned to ozone levels from the nearest monitoring site within 50 kilometers. If there was more than one monitoring site, the nearest one was chosen. Subjects who lived more than 50 kilometers away from any monitoring site were excluded.

§ The method to assess effect modification is the same as Table S3.

5. Sensitivity Analysis with Respect to Lack of Adjustment for Individual-Level Behavioral Risk Factors

Individual-level behavioral risk factors such as Body Mass Index (BMI), smoking, and income level could confound the association between long-term exposure to air pollution and mortality. A potential limitation of our study is that claims from the Medicare cohort do not provide this information. To assess whether our results could be affected by confounding bias due to the omission of these variables from the main analysis, we gathered and analyzed an additional data source called the Medicare Current Beneficiary Survey (MCBS). MCBS is a phone survey of a nationally representative sample of Medicare beneficiaries for the whole continental U.S. MCBS data provides very extensive information on behavioral risk factors (e.g., smoking, BMI, and income) and more than 150 individual confounders. The number of annual surveyed MCBS enrollees ranges from 9,224 to 11,227 (mean annual enrollees: 10305, with a total of 133,964 records and 57,154 enrollees) for the period 2000 to 2012. Among 57,154 enrollees, 10,346 were surveyed for two consecutive years and 33,232 were surveyed for three consecutive years. Table S5 summarizes descriptive statistics of the key individual-level behavioral risk factors for the MCBS population (N=57,154 MCBS enrollees who live in 6,690 ZIP codes). Figure S5 displays the geographic distribution of the ZIP codes of residence of the MCBS enrollees.

We have demonstrated, repeatedly and in peer reviewed publications, that in order for an individual-level variable Z (e.g., smoking) to confound the relationship between X (e.g., air pollution) and Y (e.g., mortality), the variable Z must be a strong predictor of X conditional to all the other covariates that are included in the survival model.[9-11] Therefore, we fit a mixed-effect model to the MCBS data, with the dependent variable the exposure to air pollution ($PM_{2.5}$ or ozone, averaged across time and assessed at the residential ZIP code for each individual) and with the independent variables: individual-level smoking, individual-level BMI, and individual-level income plus all 20 individual- and area-level variables included in the main analysis. We used

sampling weight provided by the MCBS. We included random intercept by person and by ZIP code to account for between-person variation and geographic difference.

Table S6 displays the results of this analysis. Except for BMI, all the other individual-level risk factors were not strong predictors of air pollution exposure conditionally to the other variables included in the main analysis. For BMI, we detected a significant association (p values = 0.002 and 0.026 for $PM_{2.5}$ and ozone, respectively). However, the beta coefficients were very small, indicating that BMI and the other variables have negligible effects on air pollution. For example, for an interquartile change in BMI (from 23.39 to 29.99), the expected $PM_{2.5}$ exposure decreased negligibly, by only (29.99-23.39)*2.43E-03 = 0.016 $\mu g/m^3$.

In a second analysis, we acquired another data set, called MCBS-Medicare. In this data set, the health interviews from MCBS were linked at the individual level to claims data from Medicare. In previous work under review at *Epidemiology*,[12] we constructed a new cohort of 32,119 Medicare beneficiaries residing in 5,138 ZIP codes that were interviewed as part of the MCBS between 2002 and 2010 with the same air pollution exposure as considered here. We considered four outcomes: death, all-cause hospitalizations, hospitalizations for circulatory diseases, and hospitalizations for respiratory diseases. We fit survival models with 123 potential confounders (from Medicare claims, MCBS, census) and assessed the sensitivity of the estimated air pollution health effects to the exclusion of 73 individual level risk factors (including smoking, BMI, and income). We found that our results were robust to the lack of adjustment for these variables.

Note that in our main analysis, we included in the model an individual-level variable called "eligibility to Medicaid". We found that this variable is an excellent surrogate for individual-level income. Using MCBS, which has information on both individual-level Medicaid eligibility and individual-level income, we found that the area under the ROC curve was 0.91. Results were adjusted by age, sex, and race. Therefore, we feel confident that our main analysis allows an adequate adjustment by individual level income via Medicaid eligibility.

Table S5. Descriptive Information of MCBS Subjects (N=57,154 MCBS enrollees who live in 6,690 ZIP codes)

	25% percentile	Mean	75% percentile
Age (years)	68.0	72.3	82.0
Ozone (ppb)*	41.9	46.0	50.4
$PM_{2.5}$ ($\mu g/m^3$)*	9.5	11.5	13.4
Temperature (°C)	11.0	15.2	18.5
Relative Humidity (%)	70.05	71.7	78.6
Income (U.S. dollars)	12000	30874	36000
BMI (kg/m^2)	23.39	27.03	29.99
	Percentage		
% with Smoking History	58.91%		
% Current Smokers	13.67%		

* $PM_{2.5}$ and ozone concentrations were estimated from the prediction model. $PM_{2.5}$ values were averaged across the whole year. Ozone values were averaged from April 1 to September 30.

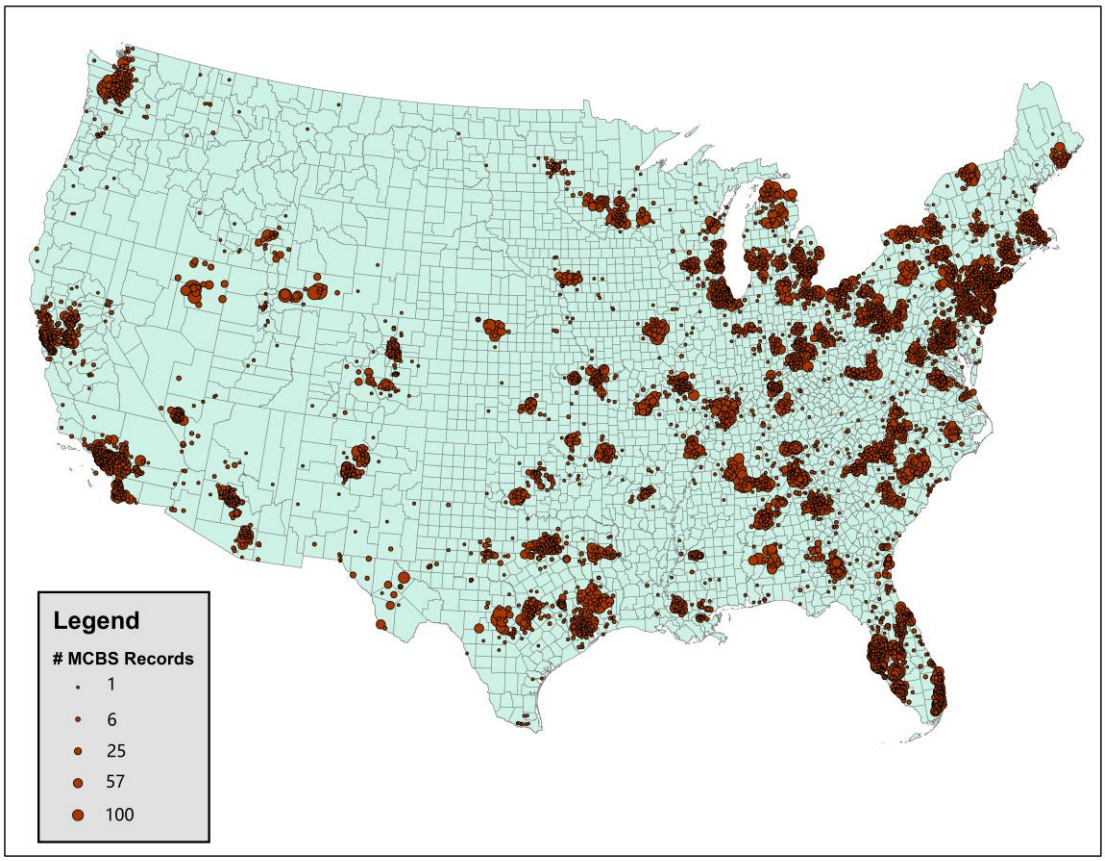


Figure S5. Geographic Distribution of MCBS Enrollees

The map shows the residential ZIP codes of MCBS enrollees. The diameter of each circle is proportional to the number of MCBS enrollees that have a place of residence in that ZIP code. The 57,154 MCBS enrollees live in 6,690 ZIP codes. Each ZIP code has from 1 to 374 MCBS enrollees (median: 2; mean: 9) and from 1 to 860 MCBS records (median: 4; mean: 20).

Table S6. Sensitivity Analysis: Association Between Individual-Level Behavioral Risk Factors in MCBS and Exposure to Air Pollution*

	$PM_{2.5}$ ($\mu g/m^3$)		**Ozone (ppb)**	
Individual covariates	Beta	p-value	Beta	p-value
BMI (kg/m^2)	-2.43E-03	0.002	3.32E-03	0.026
Indicator for being current smoker	1.25E-02	0.384	-3.89E-02	0.162
Indicator for history of smoking	6.23E-03	0.522	-1.89E-02	0.298
Income (dollars)	-9.56E-08	0.119	-1.39E-07	0.276

* We fit mixed-effect models with annual $PM_{2.5}$ (or ozone) as dependent variables. The independent variables were BMI (or indicator for being current smoker, indicator for history of smoking, or income) plus all 20 covariates included in the main analysis. For models with $PM_{2.5}$ as the response variable, we controlled for ozone, and vice versa. This model had random intercepts by person and by ZIP code.

6. Comparison of Our Results with Others in the Literature

To systematically compare our results with others in the existing literature, we gathered risk estimates of $PM_{2.5}$ and standard errors from recent studies. We pooled the existing results across studies using random-effect models for meta-analysis. The risk estimate of $PM_{2.5}$ from our main analysis (HR = 1.073 (1.071, 1.075)) is very close to the pooled estimate (HR = 1.11 (1.08, 1.15), random-effect model) (Figure S6).

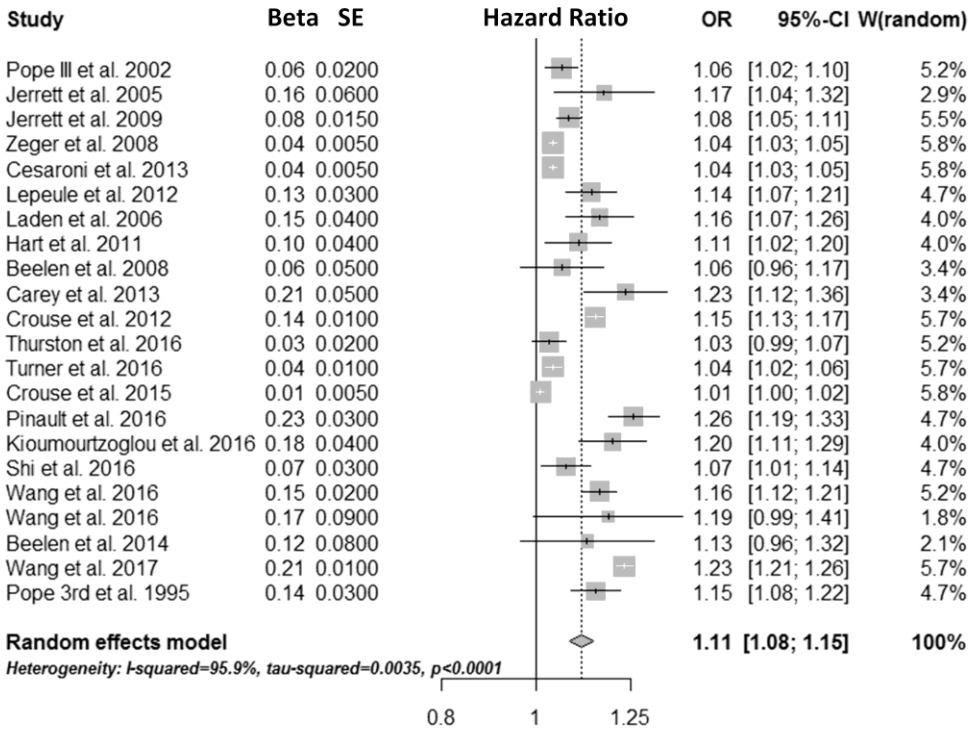


Figure S6. Forest Plot of Recent Studies on $PM_{2.5}$

The dotted line is placed at the summary HR from the random-effect model. I-square indicates that risk estimates from previous studies demonstrates a high degree of heterogeneity.

7. Concentration-Response Function

Previous studies found an almost linear concentration-response relationship between $PM_{2.5}$ and mortality,[13] and a unconfirmed threshold effect for ozone.[14,15] We examined the potential nonlinear effects of both ozone and $PM_{2.5}$ on mortality by fitting a Cox proportional hazards model with separate penalized splines for $PM_{2.5}$ and ozone. We adjusted for the same variables as in the main analysis.

Due to computational issues, running a Cox model with two separate penalized splines on the whole data set was not possible. Alternatively, we randomly divided all subjects into 50 groups with equal probability and obtained concentration-response functions separately for each of the 50 groups. To combine concentration-response curves across groups, we applied the meta-smoothing approach that has been used and modified in previous studies.[16-18] In each group, the estimated HR and its point-wise standard error were computed for each 1 $\mu g/m^3$ increment in $PM_{2.5}$ or 1 ppb increment in ozone. These group-level effect estimates ($\hat{\beta}_{ij}$=log HR) in each group *i* and for exposure level *j*, and corresponding standard error $se(\hat{\beta}_{ij})$ were combined by regressing the $\hat{\beta}_{ij}$ against indicator variables for each exposure level, with inverse variance weights. We assumed:

$$\hat{\beta}_{ij} \sim N(\beta_1 d_1 + \beta_2 d_2 + \cdots + \beta_j d_j, V_{ij})$$

where d_j is indicator variable for exposure level j and V_{ij} is the estimated variance in group i at exposure level j.

Figure S7 shows the estimated concentration-response relationships. The narrow confidence interval in most of the range reflects the large sample size in this study. The concentration-response curve for $PM_{2.5}$ is roughly linear. The concentration-response curve for ozone seems to indicate a threshold around 40 ppb, which is consistent with previous studies.[14]

Threshold Analysis

To assess as whether there is evidence of a threshold for the concentration response for ozone, we conducted a threshold analysis. The threshold model is the same as that of our main analysis except that it sets the ozone concentration to zero below the threshold value and the concentration minus the threshold value otherwise.

For ozone, we found a threshold estimate of 40 ppb based on minimizing the AIC values (Table S7). This result is consistent with visual interpretation of the concentration-response curves (Figure S7). The beta coefficient for the linear component of the threshold model is larger than the beta coefficient from the linear model.

Concentration-Response Three-Dimensional Surface

We found significant evidence for an interaction between $PM_{2.5}$ and ozone (Table S8). To further investigate the potential non-linear effect and interaction between $PM_{2.5}$ and ozone, we fit a log-linear model with a thin plate spline on both $PM_{2.5}$ and ozone and controlling for the same 20 covariates as we did in the main analysis. We incorporated a dummy variable for follow-up year to allow the baseline hazard rate to change for each follow-up year. Running a log-linear model on the whole data set also raised computational issues. Similar to obtaining concentration-response curves, we randomly divided the data into 50 splits, plotted concentration-response surfaces separately, and combined them together (Figure S8, Figure 3).

Unlike the ozone concentration-response curve, the concentration-response surface displays no threshold effect (Figure S8, Figure 3). Higher ozone is linearly associated with increased mortality at all $PM_{2.5}$ concentrations. This distinction may be due to interaction between $PM_{2.5}$ and ozone, which may change how ozone affects mortality below 40 ppb. The interaction between $PM_{2.5}$ and ozone deserves more attention and further investigation.

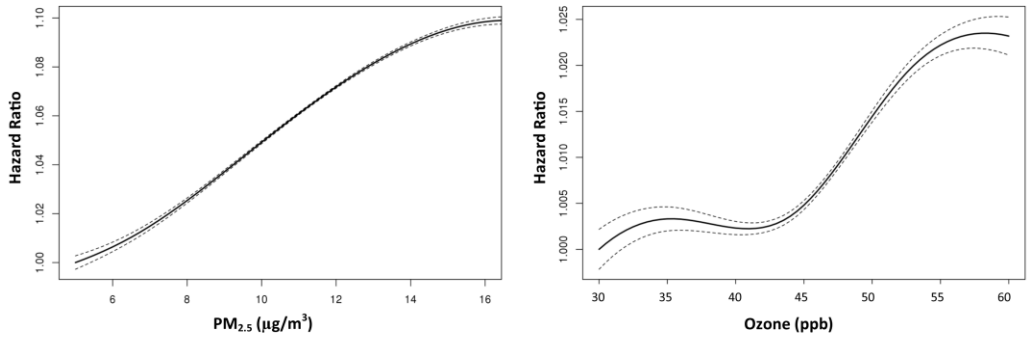


Figure S7. Concentration-Response Curves of $PM_{2.5}$ and Ozone on Mortality

We fit a Cox proportional hazards model with two penalized splines on $PM_{2.5}$ and ozone respectively, and adjusted for individual covariates (sex, race, Medicaid eligibility, age group at entry); meteorological variables (temperature and humidity), a dummy variable for region, and ecological variables (BMI, percentage of ever smoker, percentage of Hispanic population, percentage of Black population, median household income, median value of housing, percentage above age 65 living below the poverty level, percentage above age of 65 with less than high school education, percentage of owner-occupied housing units, population density, percentage of Medicare enrollees having a blood lipid (LDL-C) test, a hemoglobin A1c test, and at least one ambulatory visit to a primary care clinician).

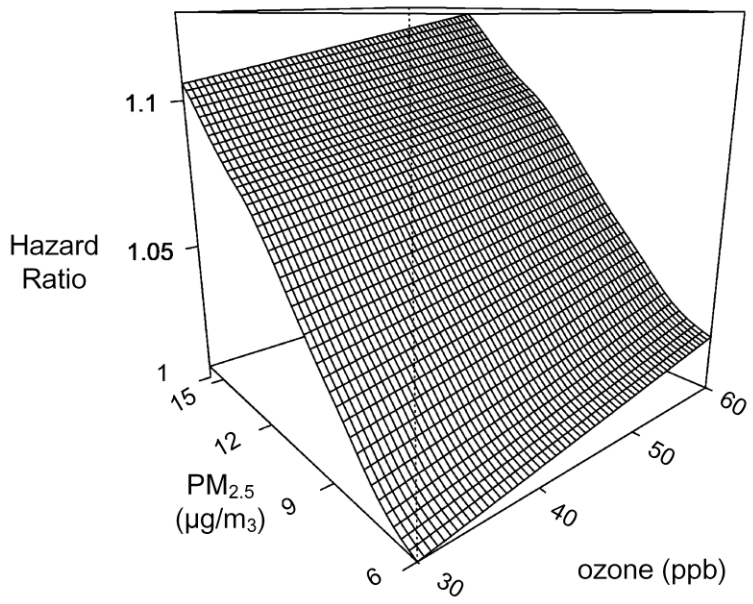


Figure S8. Concentration-Response Surface of $PM_{2.5}$ and Ozone on Mortality

We fit a log-linear model with a thin plate spline on both $PM_{2.5}$ and ozone, and adjusted for individual covariates (sex, race, Medicaid eligibility, and age group at entry); meteorological variables (temperature and humidity); a dummy variable for region; and ecological variables (BMI, percentage of ever smoker, percentage of Hispanic population, percentage of Black population, median household income, median value of housing, percentage above age 65 living below the poverty level, percentage above age of 65 with less than high school education, percentage of owner-occupied housing units, population density, percentage of Medicare enrollees having a blood lipid (LDL-C) test, percentage of Medicare enrollees having a hemoglobin A1c test, and percentage of Medicare

enrollees having at least one ambulatory visit to a primary care clinician). Then, we exported the dose-response surface.

Table S7. Threshold Analysis on Ozone

Threshold value (ppb)	**-2*log likelihood**	**AIC**	**beta**	**se**
0	615404523	615404575	0.001071	0.000048
30	615404562	615404614	0.001047	0.000049
35	615404577	615404629	0.001075	0.000051
36	615404563	615404615	0.001107	0.000052
37	615404542	615404594	0.001151	0.000053
38	615404515	615404567	0.001206	0.000054
39	615404490	615404542	0.001265	0.000055
40[†]	**615404475**	**615404527**	**0.001318**	**0.000057**
41	615404481	615404533	0.001354	0.000059
42	615404501	615404553	0.001378	0.000061
43	615404532	615404584	0.001394	0.000063
44	615404579	615404631	0.001390	0.000067
45	615404637	615404689	0.001367	0.000070
50	615404932	615404984	0.000924	0.000101

[†]Threshold analysis with threshold value at 40 ppb yields lower AIC values than threshold analyses with other threshold values.

Table S8. Interaction between $PM_{2.5}$ and Ozone

Variable	Beta	se	p
$PM_{2.5}$ (μg/m^3)	2.263E-02	4.770E-04	<0.0001
Ozone (ppb)	5.033E-03	1.262E-04	<0.0001
$PM_{2.5}$*ozone	-3.233E-04	9.640E-06	<0.0001

Starting from the two-pollutant main analysis, we added the interaction term with $PM_{2.5}$ and ozone to it and fit the model on the whole data set. Beta coefficient for $PM_{2.5}$ stands for its hypothetical risk estimate when ozone level is 0 ppb; beta coefficient for ozone stands for its hypothetical risk estimate when $PM_{2.5}$ level is 0 μg/m^3.

References

1. Health UDo, Services H. Centers for Disease Control and Prevention (CDC) Behavioral Risk Factor Surveillance System Survey Questionnaire. Atlanta 2004.
2. Wennberg JE, Cooper MM. The Dartmouth atlas of health care: American Hospital Publishing Chicago, IL; 1996.
3. Kalnay E, Kanamitsu M, Kistler R, et al. The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological Society 1996;77:437-71.
4. Di Q, Kloog I, Koutrakis P, Lyapustin A, Wang Y, Schwartz J. Assessing PM2. 5 Exposures with High Spatiotemporal Resolution across the Continental United States. Environmental science & technology 2016;50:4712-21.
5. Di Q, Rowland S, Koutrakis P, Schwartz J. A hybrid model for spatially and temporally resolved ozone exposures in the continental United States. Journal of the Air & Waste Management Association 2017;67:39-52.
6. Agency UEP. Air Quality System Data Mart (daily summary of PM2.5 FRM/FEM Mass (88101), daily summary of ozone (44201)). 2016. Available at http://www.epa.gov/ttn/airs/aqsdatamart. Accessed June, 2016.
7. Bey I, Jacob DJ, Yantosca RM, et al. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. Journal of Geophysical Research: Atmospheres 2001;106:23073-95.
8. Therneau T. Coxme and the Laplace Approximation. 2015.
9. Dominici F, Wang C, Crainiceanu C, Parmigiani G. Model selection and health effect estimation in environmental epidemiology. Epidemiology 2008;19:558-60.
10. Wang C, Dominici F, Parmigiani G, Zigler CM. Accounting for uncertainty in confounder and effect modifier selection when estimating average causal effects in generalized linear models. Biometrics 2015;71:654-65.
11. Wang C, Parmigiani G, Dominici F. Bayesian effect estimation accounting for adjustment uncertainty. Biometrics 2012;68:661-71.
12. Makar M, Antonelli J, Di Q, Cutler D, Schwartz J, Dominici F. Estimating the causal effect of lowering particulate matter levels below the United States standards on hospitalization and death: observational study using an open cohort. Epidemiology 2016.
13. Schwartz J, Coull B, Laden F, Ryan L. The Effect of Dose and Timing of Dose on the Association between Airborne Particles and Survival. . Environmental health perspectives 2008;116:64–9.
14. Turner MC, Jerrett M, Pope III CA, et al. Long-Term Ozone Exposure and Mortality in a Large Prospective Study. American journal of respiratory and critical care medicine 2015.
15. Jerrett M, Burnett RT, Pope III CA, et al. Long-term ozone exposure and mortality. New England Journal of Medicine 2009;360:1085-95.
16. Schwartz J, Zanobetti A. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. Epidemiology 2000;11:666-72.
17. Schwartz J, Ballester F, Saez M, et al. The concentration-response relation between air pollution and daily deaths. Environmental health perspectives 2001;109:1001-6.

18. Schwartz J, Laden F, Zanobetti A. The concentration-response relation between PM(2.5) and daily deaths. Environmental health perspectives 2002;110:1025-9.