The Association Between Air Pollution and Human Health

The Association between Air Pollution and Human Health

Qian Di

A Dissertation Submitted to the Faculty of

The Harvard T.H. Chan School of Public Health

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Science

in the Departments of Environmental Health and Epidemiology

Harvard University

Boston, Massachusetts.

May, 2018

Dissertation Advisor: Dr. Joel Schwartz Qian Di

The Association between Air Pollution and Human Health

Abstract

Air pollution is a risk factor of multiple adverse health outcomes. I and my colleagues investigated the health effect of long-term and short-term exposures to air pollution in the entire Medicare population. We were interested in the effect of air pollution below the current National Ambient Air Quality Standards (NAAQS) set by the U.S. Environmental Protection Agency (EPA). We also identified subgroups that were particularly vulnerable to air pollution.

We first developed air pollution models to predict daily and annual exposures to fine particulate matter (particles with aerodynamic diameter less than 2.5 μm, or $PM_{2.5}$) and ozone among Medicare population with high accuracy ($R^2 = 0.84$ for $PM_{2.5}$ and $R^2 = 0.76$ for ozone). For long-term exposures to air pollution, we found that increases of 10 $\mu g/m^3$ 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 - 7.5) and 1.1% (95% CI, 1.0 - 1.2), respectively. The association was still significant when restricting to $PM_{2.5}$ below 12 $\mu g/m^3$ (the annual standard set by EPA) and ozone below 50 ppb. The risk of mortality was higher among male, African Americans and people of low income than the general population. For short-term exposures to air pollution, increase of 10 $\mu g/m^3$ in daily $PM_{2.5}$ and 10 ppb in daily ozone were significantly associated with increases in daily all-cause mortality of 1.05% (95% CI, 0.95 - 1.15) and 0.51% (95% CI, 0.41 - 0.61), respectively.

When restricting to the analysis below 25 μg/m^3 and 60 ppb, lower than the current daily standards set by EPA, the significant association remained.

In conclusion, exposures to air pollution are associated with mortality both in the long-term and short-term, even below the current national standards, providing strong evidence that these national air quality standards need to be reevaluated. Some subgroups, such as African Americans and people of low income, are more vulnerable to air pollution than the general population.

Table of Contents

List of Figures with Captions

List of Tables with Captions

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Joel Schwartz, who has been an invaluable mentor, for his academic guidance and wisdom during my doctoral study. His foresight, wisdom, and experience have been an inestimable help to my research. I would like to thank Dr. Francesca Dominici, for her patient guidance, incredible wisdom, long-lasting logistics support.

I would also like to thank my research committee member, Dr. Petros Koutrakis, for sparing time to review my work and give suggestions to my research. I am greatly appreciated to Dr. Christine Choirat, Dr. Yun Wang, Dr. Antonella Zanobetti, Dr. Lingzhen Dai, Dr. Yan Wang, Dr. Itai Kloog, Dr. Alexei Lyapustin, Dr. Yujie Wang, Dr. James Kelly for their help on my research. This work could not be completed without their help. I also appreciated Kathy Brenner and Dr. Stacey Tobin for their professional edition on my work.

Thanks to my parents, Di, Hao and Wen, Xiaoling, for their enduring encouragement and support.

Finally, I want to thank Dr. Zhang, Haoyu, who has been always standing by me. It is blessed to have her in my life.

Introduction

Air pollution is a serious public health issue and a leading cause of death worldwide, especially among developing countries. Long-term exposure to fine particulate matter (particles with aerodynamic diameter less than 2.5 μm, or $PM_{2.5}$) was found to be associated with multiple adverse health outcomes, including morbidity,[1,2] mortality,[3] cardiovascular disease,[4] respiratory disease,[5] myocardial infarction,[6] an increase in hospital admission,[1,7,8] and others.[9] Long-term exposure to ozone was also associated with adverse health outcomes.[10-13]

To control air pollution, the Clean Air Act requires the U.S. Environmental Protection Agency (EPA) to enact National Ambient Air Quality Standards (NAAQS). Currently, the daily $PM_{2.5}$ and ozone standards are set at 35 $\mu g/m^3$ and 70 ppb, respectively; the annual standards for $PM_{2.5}$ is 12 $\mu g/m^3$, and there is no long-term standard for ozone.

Several studies used large cohorts to investigate the health effect of long-term exposures to $PM_{2.5}$ and ozone.[11,13-17] Those cohorts are generally convenience samples, which often have higher socioeconomic status (SES) than the national average, and live in well-monitored urban areas, which often have higher air pollution levels than smaller cities and rural areas. Some recent studies used national exposure models to estimate the health effect at lower concentrations but again estimated health effects in convenience samples.[18-21] A critical gap of knowledge is the evidence on health effects of air pollution based on representative samples, at unmonitored areas and areas with air pollution levels below the current NAAQS. Also, we lack a rigorous characterization of which sub-populations (based on their SES and/or race and gender) might be most affected.

Therefore, to assess exposures to air pollution at unmonitored areas, my colleagues and I first developed a neural network-based approach to predict daily $PM_{2.5}$ levels in the continental United

States from 2000 to 2012 with high spatial resolution. This prediction approach integrates satellite-based measurement that quantifies light attenuation due to $PM_{2.5}$, land-use data, simulation outputs from a chemical transport model, meteorological variables, and other datasets to predict daily $PM_{2.5}$ levels, even at places without monitoring stations. Prediction results makes it possible to assess both long-term (at annual level) and short-term (at daily level) exposures to air pollution with high accuracy across the U.S. Related contents are presented in Chapter 1 and have been published on the *Environmental Science & Technology*.[22]

With the long-term exposures to air pollution available, my colleagues and I investigated the adverse health effect of long-term exposures to air pollution among the entire Medicare population using survival analysis. We ascertained exposures to air pollution based on the residential ZIP Code of each participant, controlled for individual-level covariate and ecological variables, and used all-cause mortality as the health outcome. We also assessed the association between air pollution and mortality (1) in several subgroups, defined by socioeconomic status, sex, and race/ethnicity; (2) at locations with air pollution levels below the current annual air quality standards. Details are specified in Chapter 2 and have been published on the *New England Journal of Medicine*[23].

We examined the health effect of short-term exposures to air pollution in a similar way in Chapter 3. We focused on all death cases in the Medicare population and analyzed the association with short-term exposure to air pollution using case-crossover study design. Subgroup analyses were conducted at different subpopulation and at places with daily air pollution levels much lower than the current daily air quality standards to identify the subpopulation that are more vulnerable to air pollution and evaluate whether the current daily air quality standards are stringent enough. Related contents have been published on the *Journal of the American Medical Association* [24].

References

1. Lippmann M, Ito K, Nadas A, Burnett R. Association of particulate matter components with daily mortality and morbidity in urban populations. Research report (Health Effects Institute) 2000:5-72, discussion 3-82.

2. Sarnat JA, Marmur A, Klein M, et al. Fine particle sources and cardiorespiratory morbidity: an application of chemical mass balance and factor analytical source-apportionment methods. Environmental Health Perspectives 2008;116:459.

3. Schwartz J, Dockery DW, Neas LM. Is daily mortality associated specifically with fine particles? Journal of the Air & Waste Management Association (1995) 1996;46:927-39.

4. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. Jama 2006;295:1127-34.

5. Peng RD, Bell ML, Geyh AS, et al. Emergency Admissions for Cardiovascular and Respiratory Diseases and the Chemical Composition of Fine Particle Air Pollution. Environmental Health Perspectives 2009;117:957-63.

6. Peters A, Dockery DW, Muller JE, Mittleman MA. Increased Particulate Air Pollution and the Triggering of Myocardial Infarction. Circulation 2001;103:2810-5.

7. Schwartz J. Air pollution and hospital admissions for cardiovascular disease in Tucson. Epidemiology 1997:371-7.

8. Schwartz J, Morris R. Air pollution and hospital admissions for cardiovascular disease in Detroit, Michigan. American journal of epidemiology 1995;142:23-35.

9. Pope CA, Dockery DW. Health Effects of Fine Particulate Air Pollution: Lines that Connect. Journal of the Air & Waste Management Association 2006;56:709-42.

10. Crouse DL, Peters PA, Hystad P, et al. Ambient PM2. 5, O3, and NO2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environmental health perspectives 2015;123:1180.

11. 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.

12. 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. American journal of respiratory and critical care medicine 2015;192:337-41.

13. Jerrett M, Burnett RT, Pope III CA, et al. Long-term ozone exposure and mortality. New England Journal of Medicine 2009;360:1085-95.

14. 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.

15. 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. American journal of respiratory and critical care medicine 2013;187:1226-33.

16. 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. Environmental Health Perspectives (Online) 2015;123:549.

17. Crouse D, Peters P, Hystad P, et al. Ambient $PM_{2.5}$, O3, and NO2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). . 2015;Environ Health Perspect 1180–6.

18. 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. Environmental health perspectives 2012;120:708.

19. Thurston GD, Ahn J, Cromar KR, et al. Ambient particulate matter air pollution exposure and mortality in the NIH-AARP diet and health cohort. Environmental health perspectives 2016;124:484.

20. 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. Environmental Health 2016;15:1.

21. Shi L, Zanobetti A, Kloog I, et al. Low-concentration PM2. 5 and mortality: Estimating acute and chronic effects in a population-based study. Environmental health perspectives 2016;124:46.

22. 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.

23. Di Q, Wang Y, Zanobetti A, et al. Air pollution and mortality in the Medicare population. New England Journal of Medicine 2017;376:2513-22.

24. Di Q, Dai L, Wang Y, et al. Association of short-term exposure to air pollution with mortality in older adults. Jama 2017;318:2446-56.

Chapter 1. Assessing $PM_{2.5}$ Exposures with High Spatiotemporal Resolution across the Continental United States[1]

Qian Di[1], Itai Kloog[1,2], Petros Koutrakis[1], Alexei Lyapustin[3], Yujie Wang[3], Joel Schwartz[1]

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

[2]Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer Sheva, P.O.B. 653, Israel

[3]GEST/UMBC, NASA Goddard Space Flight Center, Baltimore, MD, USA

[1] Reproduced with permission from Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y. and Schwartz, J., 2016. Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. *Environmental science & technology*, *50*(9), pp.4712-4721.

Abstract

A number of models have been developed to estimate $PM_{2.5}$ exposure, including satellite-based aerosol optical depth (AOD) models, land-use regression or chemical transport model simulation, all with both strengths and weaknesses. Variables like normalized difference vegetation index (NDVI), surface reflectance, absorbing aerosol index and meteoroidal fields, are also informative about $PM_{2.5}$ concentrations. Our objective is to establish a hybrid model which incorporates multiple approaches and input variables to improve model performance. To account for complex atmospheric mechanisms, we used a neural network for its capacity to model nonlinearity and interactions. We used convolutional layers, which aggregate neighboring information, into a neural network to account for spatial and temporal autocorrelation. We trained the neural network for the continental United States from 2000 to 2012 and tested it with left out monitors. Ten-fold cross-validation revealed a good model performance with total R^2 of 0.84 on the left out monitors. Regional R^2 could be even higher for the Eastern and Central United States. Model performance was still good at low $PM_{2.5}$ concentrations. Then, we used the trained neural network to make daily prediction of $PM_{2.5}$ at 1 km×1 km grid cells. This model allows epidemiologists to access $PM_{2.5}$ exposure in both the short-term and the long-term.

Keywords. $PM_{2.5}$; Aerosol optical depth; GEOS-Chem; Absorbing aerosol index; Land-use regression; Convolutional neural network

Introduction

Fine particulate matter ($PM_{2.5}$) is a major concern in public health.[1-6] An adverse health effect is associated with $PM_{2.5}$ exposure in the short term[7,8] and the long term.[9,10] $PM_{2.5}$ is found to be associated with morbidity,[11,12] mortality,[6] cardiovascular disease,[4] respiratory disease,[13] myocardial infarction,[14] an increase in hospital admission[11,15,16] and others.[17]

Accurate exposure assessment of $PM_{2.5}$ is a prerequisite of to investigate its adverse health effect. Early studies estimated $PM_{2.5}$ at the nearest monitoring station.[18] However, nearest monitors cannot capture all variability in $PM_{2.5}$ concentrations and nondifferential misclassification occurs.[19]

Various approaches have been developed to achieve better exposure assessment. Spatial interpolation, including nearest-neighbor interpolation and Kriging interpolation, was used to smooth $PM_{2.5}$ concentration and estimate local exposure. Nonetheless, interpolation adds no additional information to the model. Local emission like highways between two monitor sites is not captured by simple interpolation. Land-use regression (LUR) uses land-use terms, such as road density, percentage of urban and others, as proxies for $PM_{2.5}$ concentration.[20,21] Although LUR could achieve a high spatial resolution, it has limited temporal resolution since land-use terms are usually time-invariant.[22] Recent improvements in land-use regression enable incorporation some level of time-variant factors,[23,24] but land-use terms are still inadequate in modeling short-term variations and often limited by short temporal coverage.[25]

Satellite-based aerosol optical depth (AOD) measurements have been widely used to estimate $PM_{2.5}$ in various models for its large spatial coverage and repeated daily observations.[26] AOD measures the light extinction due to aerosol in the whole atmospheric column.[27] To obtain ground-

level $PM_{2.5}$ concentration, vertical distribution of aerosol is needed. Recent studies proposed different calibration methods.[26,28-32] Most studies focused on quantifying relationship between AOD and $PM_{2.5}$ or predict a long-term average of $PM_{2.5}$, while epidemiological studies also need short-term $PM_{2.5}$ assessment. Some studies combined AOD and land-use regression and used mixed effect model to achieve improvements on model performance.[33-35] However, the drawback of AOD is missing data, which is caused by bright surfaces or cloud contamination, especially in winter.[36] Also, AOD measurements may also have abnormally large values caused by forest fires.[37] For grid cells with missing or abnormal values, the AOD-$PM_{2.5}$ relationship may be problematic, especially for daily $PM_{2.5}$ assessment. The relationship between column aerosol concentration and ground-level concentration can be influenced by multiple factors such as meteorological fields, chemical profile of aerosol, and others.[38,39] The absorbing aerosol index (AAI) provides information about aerosol type and is informative to $PM_{2.5}$ modeling.[40,41]

Chemical transport models (CTMs), like GEOS-Chem,[42] CMAQ,[43] and CHIMERE,[44] simulate the formation, dispersion and deposition of fine particles based on emission inventories and known atmospheric chemical reaction. CTM is another way to assess $PM_{2.5}$ concentration. Due to the complexity of reactions and atmospheric meteorological processes, simulated concentration often deviates from the real world. CTM outputs are often used after calibration.[45,46] CTM provides an aerosol vertical profile, which has been used as scaling factor in AOD calibration.[29,47] Due to the limit of computation, CTM usually has coarse spatial resolution. In a previous study, we have proposed a hybrid model which uses land-use regression to downscale CTM outputs. [48]

Existing approaches have both strengths and weaknesses, and often they complement to each other. In this paper, we incorporated multiple variables into a neural network-based hybrid model, including satellite-based AOD data, AAI, CTM outputs, land-use terms, and meteorological

variables. We validated the model with 10-fold cross-validation and predicted daily $PM_{2.5}$ at 1 km×1 km resolution in the continental United States for the years 2000-2012. Prediction with such a high temporal and spatial resolution allows epidemiological studies to estimate health effect of $PM_{2.5}$ with greater reliability.

Materials

Study Domain

The study domain is the continental United States, including 48 contiguous states and Washington, D.C (Figure S1-1). The study period is from January 1st, 2000 to December 31st, 2012, a total of 4,749 days.

Monitoring Data

Monitoring data for $PM_{2.5}$ were collected by EPA Air Quality System (AQS). In total, there were 1,986 monitor stations available in this period and 1,928 of them were located in the study area. Not every monitoring site has data available throughout the study period. Monitoring sites were densely distributed along coastal areas and the Eastern part, while there were a few monitors in the Mountain Region (Figure S1-1). We calibrated our hybrid model to the daily average of monitored $PM_{2.5}$.

AOD Data

The Moderate Resolution Imaging Spectroradiometer (MODIS) is an instrument aboard the Earth Observing System (EOS) satellite.[49,50] Several algorithms have been developed to retrieve AOD data from MODIS measurement,[51] including a recent algorithm called MAIAC, which retrieves AOD with a spatial resolution of 1 km×1 km. [52-54] We used MAIAC AOD data from Aqua satellite from 2003 to 2012 and Terra satellite from 2001 to 2012. The MAIAC algorithm arranges data at

600 km×600 km tile, which includes 360,000 1 km×1 km grid cells. In total 33 tiles and 11,880,000 grid cells were used in this study, which is also the grid cell we made predictions at. Grid cells over water bodies were excluded from the study.

AOD data has some portion of missing values, especially in the winter. Missing values are caused by bright surfaces (e.g. snow coverage) and cloud contamination.[36] In addition, AOD data may have abnormally large values due to extreme events like forest fires.[37] Usually AOD data with values above 1.5 are excluded from modeling, which also creates missing values.[55] Our previous study calibrated column aerosol mass from CTM outputs to satellite-based AOD and predicted AOD values when satellite-based AOD are missing.[56] For AOD data used in this study, we filled in the missing values using this method as preprocessing (Section 3, Supplementary Material).

Surface Reflectance

Surface characteristics and errors in AOD data products have been well documented by previous studies.[57] The MAIAC algorithm was designed to retrieve AOD over various surfaces, but surface brightness can still affect data quality.[54] We used MODIS surface reflectance data (MOD09A1) to control for that.[58] MOD13A1 has a spatial resolution of 500 m×500 m and a temporal resolution of 8 days. We used surface reflectance from Band 3 and linearly interpolated values for days without measurements.

Chemical Transport Model Outputs

We used GEOS-Chem, a chemical transport model, to simulate ground-level $PM_{2.5}$ concentration. GEOS-Chem is a global 3-dimensional chemical transport model, which uses meteorological inputs and emission inventories to simulate atmospheric components. The details of GEOS-Chem is articulated somewhere else.[42] We performed a nested grid simulation (Version 9.0.2) for North

America at 0.500°×0.667° from 2005 to 2012, with boundary conditions exported from a 2.0°×2.5° global simulation. Since meteorological inputs at 0.500°×0.667° are not available from 2000 to 2004, we used 2.0°×2.5° outputs instead. Based on previous studies and pilot testing, total $PM_{2.5}$ was defined as the sum of nitrate, sulfate, elemental carbon, organic carbon, ammonium, sea salt aerosol, dust aerosol and others (Table S2).[59]

In additional to providing ground-level $PM_{2.5}$ estimation, GEOS-Chem also simulates vertical distribution of aerosol, which could be used for calibrating AOD. Previous studies used GEOS-Chem to compute the percentage of ground-level aerosol in the total column aerosol. This percentage was used in AOD calibration as a scaling factor.[29,60] Both studies utilized GEOS-Chem to provide both direct estimation for ground-level $PM_{2.5}$ and a scaling factor to calibrate AOD.

Meteorological Data

Meteorological fields were obtained from NCEP North American Regional Reanalysis data, which assimilates various data sources like land-surface, ship, radiosonde, pibal, aircraft, satellite and others.[61] Meteorological data are daily estimate at 0.3° grid cells (about 32 km). In total 16 meteorological variables were used in this study. They include air temperature, accumulated total precipitation, downward shortwave radiation flux, accumulated total evaporation, planetary boundary layer height, low cloud area fraction, precipitable water for the entire atmosphere, pressure, specific humidity at 2m, visibility, wind speed, medium cloud area fraction, high cloud area fraction, and albedo. Wind speed was computed as the vector sum of u-wind (east-west component of the wind) at 10m and v-wind (north-south component) at 10m.

Aerosol Index Data

Absorbing aerosol index (AAI) indicates the presence of absorbing aerosols in the atmosphere. Major sources of absorbing aerosol include biomass burning and desert dust; other minor sources could be volcanic ash.[62] AAI is informative for estimating absorbing aerosols, such as organic carbon and soil dust.[63,64] We used AAI Level 3 data products from the Ozone Monitoring Instrument (OMI), where two algorithms are used in retrieval. One is a near-UV algorithm, which retrieves UV aerosol index (OMI data product OMAERUVd);[62,64] and the other one uses multiwavelength aerosol algorithm, whose outputs include aerosol indexes at visible and UV range (OMI data product OMAEROe).[65] Both algorithms have pros and cons, which have been discussed previously.[66] Both data products are complementary, and thus we used both. OMI AAI data is available after October 2004. OMAERUVd data product has a spatial resolution of 1°; OMAEROe data product has a spatial resolution of 0.25°.

Land-use terms

Land-use terms serve as proxies for emissions and are used to capture variations at a small a spatial scale, which may not modeled by GEOS-Chem. The detailed process of obtaining land-terms like elevation, road density, NEI (National Emissions Inventory) emission inventory, population density, percentage of urban, and NDVI has been reported somewhere else.[67] For vegetation coverage, we used percentage of vegetation from NCEP North American Regional Reanalysis data and MODIS MOD13A2, a NDVI data product.[68] MOD13A2 has a spatial resolution of 1 km×1 km and a temporal resolution of 16 days. We linearly interpolated NDVI values for days without measurements.

Regional and Monthly Dummy

Previous studies found the relationship between AOD and $PM_{2.5}$ have regional and daily variation due to the difference in meteorology and aerosol composition.[38,69] Atmospheric mechanism is complex, and relationships between other variables could also differ temporally and spatially. To account for that, we used monthly and regional dummy variables. Regional dummy variable comes from major climate types in the United States (Figure S1-3).[70] Since the AOD-$PM_{2.5}$ relationship can change from day to day, daily dummy variables would be ideal. However, training a neural network with 365 indicator variables in addition to the other variables would be computationally intensive, and we used monthly dummy variables as a compromise.

Methods

We trained a neural network with the above variables to $PM_{2.5}$ monitoring data from the AQS network. The relationships between input variables and $PM_{2.5}$ could be highly nonlinear with complex interactions. Neural networks have the potential to model any type of nonlinearity.[71,72] The details of the neural network, such as its structure and training method were articulated in the supplementary material. All input variables covered the entire study area, but some of them were not available in early years or had higher proportions of missing values. Missing values were especially common in Terra and Aqua AOD data. To deal with the missing values problem and different temporal coverages, we adopted the following steps. We used a calibration method to fill in the missing values in Aqua AOD data from 2003 to 2012 and Terra AOD data from 2001 to 2012 based on the association of GEOS-Chem outputs and land-use terms with non-missing AOD.[56] For the other variables with a low fraction of missing values, we interpolated at grid cells with missing values. Regarding temporal coverage, GEOS-Chem outputs, land-use terms, MODIS outputs, and meteorological variables were available throughout the study period. OMI data, Aqua

AOD, and Terra AOD were unavailable in earlier years. For years with one or more unavailable variables, we fitted the model with the remaining available variables.

Most previous studies used only *in situ* variables for modeling. However, information from a neighboring cell can be informative as well. For example, nearby road density, forest coverage and other land-use variables as well as nearby $PM_{2.5}$ measurements either influence or correlate with local $PM_{2.5}$ measurements. They are informative for modeling and can improve model performance. We accounted for spatial correlation by using convolutional layers in the neural network.[73] A convolutional layer is computed by applying a convolution kernel on an input layer. Values from neighboring cells are combined through the use of the kernel function. The kernel takes the form a function (e.g. weighted average with Gaussian weights based on distance) that produces a scalar estimate from the multidimensional inputs. A convolution layer aggregates nearby information and can simulate some form of autocorrelation. We included convolutional layers for land-use terms and nearby $PM_{2.5}$ measurements as additional predictor variables to account for spatial autocorrelation. Multiple convolution layers were incorporated to allow the neural network to model even more complex autocorrelation or possible interaction with other variables (Supplementary material). In addition to nearby grid cells, observations from nearby days for the same grid cell can be also informative. To incorporate this, we first fitted a neural network and obtained an initial prediction for $PM_{2.5}$. We then computed temporal convolution layers and fitted the neural network again with them (Figure S1-5).

To validate model results and avoid overfitting, we used 10-fold cross-validation, in which all monitoring sites were randomly divided into 10%-90% splits. The model was trained with 90% of data and predicted $PM_{2.5}$ at the remaining 10%. The same process was repeated for other splits. Assembling predicted $PM_{2.5}$ at ten 10% testing sets yielded predicted $PM_{2.5}$ for all the monitors.

We computed correlation between predicted $PM_{2.5}$ and monitored $PM_{2.5}$. Spatial and temporal R^2s were also calculated. Details of calculating R^2 have been specified in the supplementary material.

The trained neural network was then used to make daily $PM_{2.5}$ predictions for each grid cell (1 km×1 km) for each day.

All programming was implemented in Matlab (version 2014a, The MathWorks, Inc.).

Results

To determine input variables, we compared models with different combinations of input variables based on cross-validated total R^2. Model comparison indicated that (1) a hybrid model performed better than any subset models (Figure S1-6); (2) scaling factor was better to be incorporated as a separate input layer (Figure S1-7); (3) convolutional layers for land-use variables and predicted $PM_{2.5}$ both improved model performance (Figure S1-6 and Figure S1-8). Hence, input variables for the final model were GEOS-Chem outputs, Aqua and Terra AOD, scaling factor, OMI AAIs, meteorological variables, NDVI, surface reflectance, land-use terms, convolutional layers, and regional/monthly dummy variables.

Table 1-1 presents model performance after conducting 10-fold cross-validation. Total R^2 between fitted and monitored $PM_{2.5}$ ranged from 0.74 to 0.88 and spatial R^2 was from 0.78 to 0.88. By season, the model usually performed better in summer, followed by autumn, spring, and winter (Table S3). By region, regions in the Eastern United States had the best model performance, followed by the Central United States. The Pacific and Mountain regions had a less satisfying model performance. We also found R^2 remained high before 2008 and dropped after 2010 for sub-regions and the whole study area (Table S4). We will discuss possible reasons later. Region name and division are from U.S. census division (Table S1, Figure S1-2). In terms of spatial pattern, we

found an east-west gradient with model performing better in the Eastern and Central United States but less satisfying in the western coast and the Mountain Region (Figure 1-1). Besides, some areas in the Mountain Region (e.g. Great Basin and Colorado Plateau) with large variability in elevation and surface type have relative low R^2 all the year round. Even in the Eastern United States, where model performance is high in general, areas along Appalachian Mountains and around Ozark Plateau have less satisfying model performance.

Figure 1-2 shows the spatial distribution of total $PM_{2.5}$ in the study area. The Eastern United States generally had higher $PM_{2.5}$ levels than the Western part. The area around Illinois and Ohio, areas around New York City and Philadelphia, and parts of the Southeastern United States witnessed the heaviest $PM_{2.5}$ pollutions in the study area, especially in summer. The San Joaquin Valley, Salt Lake City, and Denver stood out in the Western United States for their high $PM_{2.5}$ levels. Regarding the temporal trend, the national average dropped from 9.2 $\mu g/m^3$ in 2003 to 7.5 $\mu g/m^3$ 2012 (Figure 1-3). By regions, the declining trend was predominantly in the Eastern United States, with largest reduction occurring in East South Central Region (5.8 $\mu g/m^3$).

One additional way to validate our exposure estimates is to see if they can reproduce the spatial autocorrelation in $PM_{2.5}$ concentrations. To do this, we calculated the correlation among all pairs of $PM_{2.5}$ monitors in the EPA network, and plotted them as a function of distance. We compared that to the same plot, but using our predicted $PM_{2.5}$ concentrations instead (Figure 1-4). The results show essential identical trends and substantial overlap between the correlations of actual vs modeled $PM_{2.5}$ with distance.

Table 1-1 Cross-validated R^2 for the Whole Study Area

Year	Total R^2	RMSE	Spatial R^2	RMSE	Temporal R^2	RMSE	Bias	Slope
2000	0.86	3.35	0.85	1.52	0.85	3.07	0.22	1.01
2001	0.84	3.58	0.86	1.40	0.83	3.35	0.22	1.01
2002	0.88	2.99	0.88	1.24	0.88	2.75	0.25	1.00
2003	0.88	2.80	0.87	1.21	0.88	2.57	0.23	1.00
2004	0.88	2.69	0.79	1.50	0.88	2.45	0.22	1.00
2005	0.88	2.94	0.84	1.45	0.89	2.66	0.27	1.00
2006	0.86	2.77	0.80	1.34	0.86	2.50	0.25	1.00
2007	0.87	2.95	0.83	1.31	0.87	2.72	0.21	1.00
2008	0.85	2.64	0.79	1.26	0.86	2.40	0.19	1.00
2009	0.82	2.73	0.81	1.09	0.82	2.54	0.21	1.00
2010	0.81	2.85	0.84	1.21	0.81	2.60	0.51	0.98
2011	0.81	2.83	0.81	1.11	0.81	2.60	0.38	0.99
2012	0.74	3.15	0.78	1.16	0.74	2.92	0.32	1.00
Mean	0.84	2.94	0.83	1.29	0.84	2.70	0.27	1.00

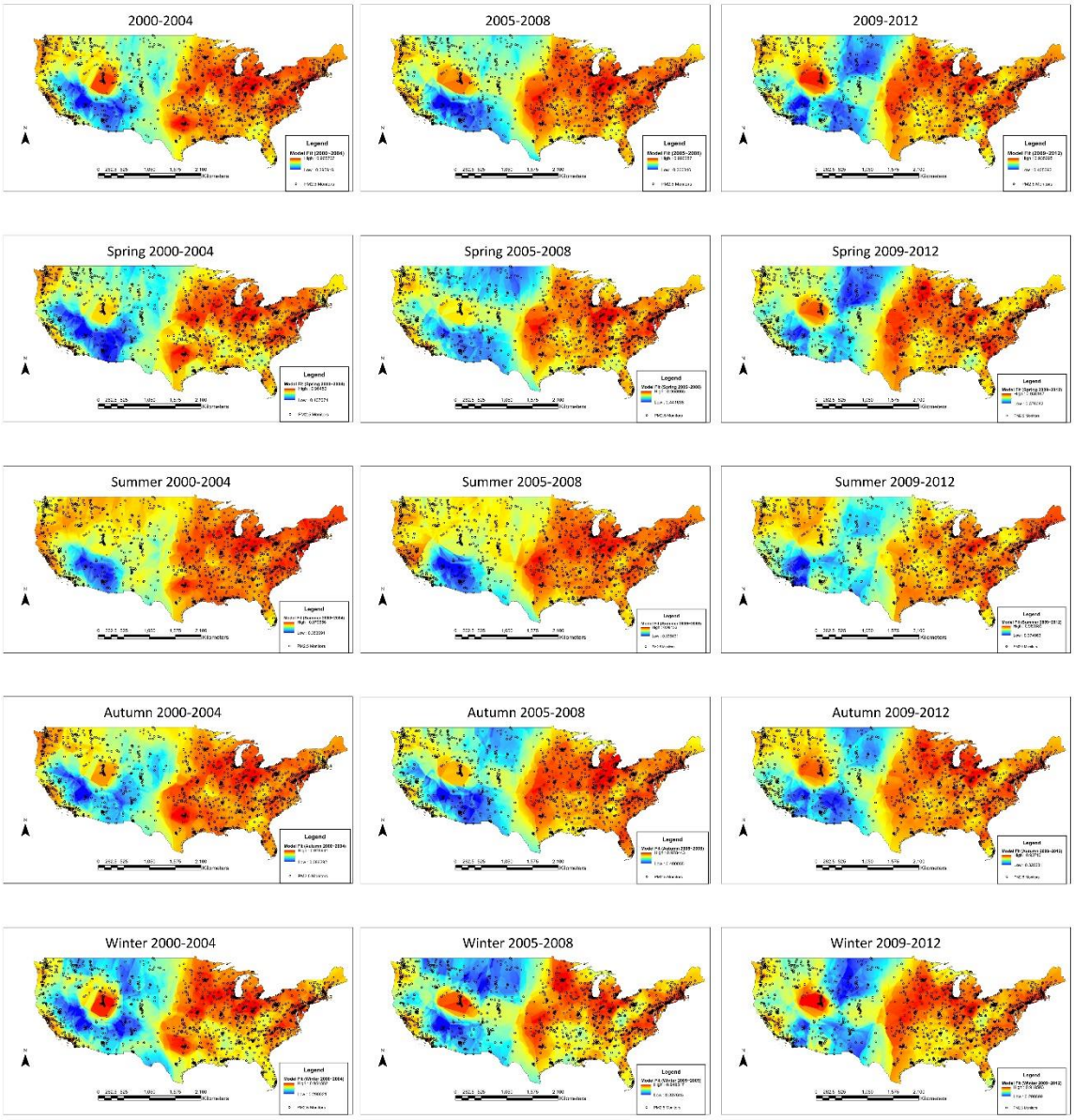


Figure 1-1. Model Performance in the Continental United States

We calculated total R^2 between monitored and predicted $PM_{2.5}$ for each monitoring site and interpolated R^2 to places without monitors using Kriging interpolation. Spring was defined as March to May; summer was defined as June to August; autumn was defined as September to

November; winter was from December to February of the next year (same below). The red color stands for high R^2, and the blue color stands for low R^2.

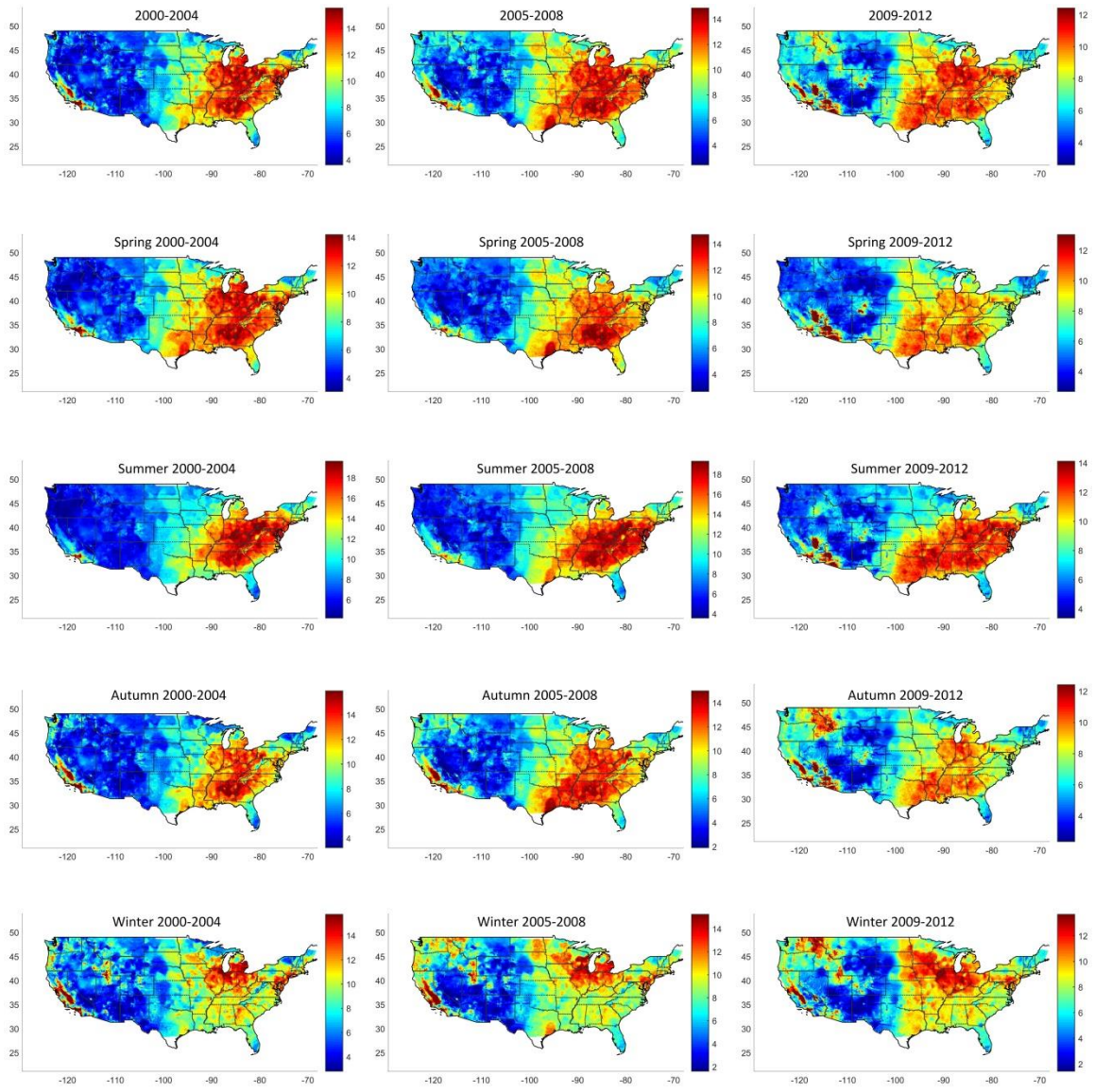


Figure 1-2. Spatial Distribution of Predicted $PM_{2.5}$

Trained neural network predicted daily total $PM_{2.5}$ concentration at 1 km×1 km grid cell in the study area. The red color stands for high concentrations, and the blue color stands for low concentrations.

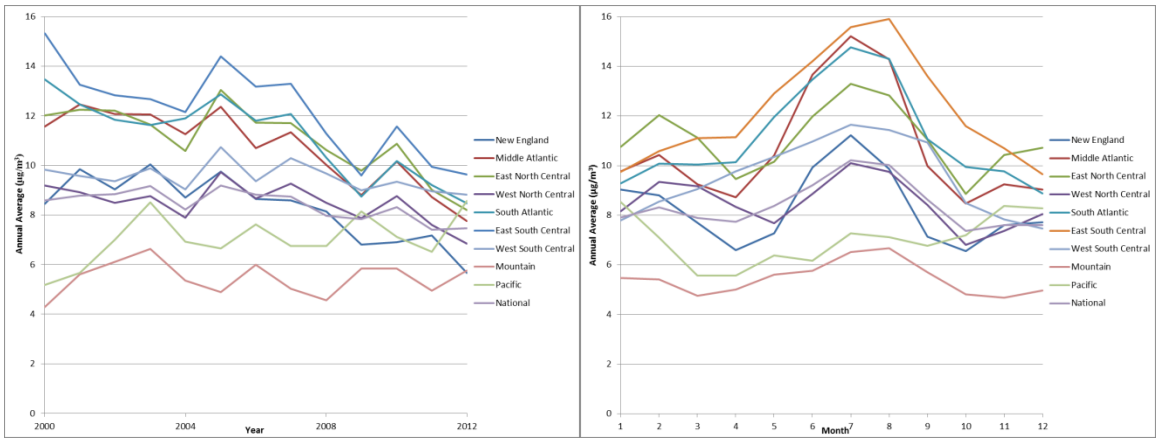


Figure 1-3. Annual Means by Month of Year and By Region

Annual averages were computed by averaging all predicted $PM_{2.5}$ values at 1 km×1 km grid cells in that region or in that month.

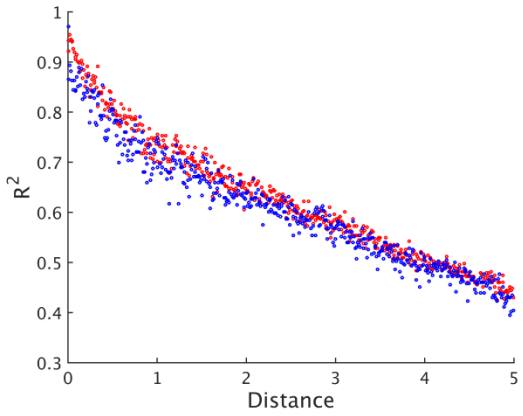


Figure 1-4. Relationship between Correlation and Distance between any Two Monitor Sites

For 1,928 monitoring sites in the study area, we computed the correlation of $PM_{2.5}$ measurements and distance (in degree) between any two monitoring site pairs and plotted the between-site correlation versus between-site distance (red dots). We repeated the same process for predicted $PM_{2.5}$ and plotted the correlation of predicted $PM_{2.5}$ and monitored $PM_{2.5}$ between two site pairs versus distance (blue dots). This figure is for year 2012.

Discussion

Our hybrid model incorporated existing $PM_{2.5}$ models as well as multiple variables and achieved high out-of-sample R^2, averaging 0.84 (0.74~0.88 by year) over the study period. The model performed better in some eastern regions, with an average out of sample R^2 of 0.86~0.89 by region. To our best knowledge, our model performance surpasses existing similar studies. Meanwhile, we predicted $PM_{2.5}$ daily concentrations at nationwide 1 km×1 km grid cells from 2000 to 2012. As discussed below, this level of resolution and coverage is an improvement over current $PM_{2.5}$ models and could be beneficial to epidemiological studies. Epidemiologists could identify long-term and short-term exposure of $PM_{2.5}$ in the whole continental United States at individual level, which helps study adverse health effect of $PM_{2.5}$ with higher accuracy.

There are several advantages and innovations in our approach. First of all, our model covered the whole United States with a spatial resolution of 1 km×1 km and a temporal resolution of 1 day and achieved high R^2. As far as we know, if taking coverage, resolution and model performance into consideration, our model performs better than existing models. As mentioned in the Introduction, most $PM_{2.5}$ modeling work that used AOD data focused on the AOD-$PM_{2.5}$ relationship, instead of making predictions. For studies with a similar research goal as ours, some of them have done AOD calibration at global scale, but their estimation was long-term average[29] or annual average, with some degree of bias (slope=0.68) and modest R^2 (R^2=0.65).[47] A previous study calibrated AOD to daily monitored $PM_{2.5}$ in the Northeastern United States using a mixed model and achieved R^2 around 0.725~0.904.[31] A similar study used the similar method for the Southeastern United States and achieved R^2 around 0.63 to 0.85.[32] Compared with both regional models, our hybrid nationwide model performs slightly better in the Northeastern United States and much better in the Southeastern United States (Table S1-6). One reason is that aerosol formation in the

Southeastern United States is affected by biogenic isoprene emission from trees;[74] while isoprene emission from trees in the Northeastern United States is less of a concern. Secondary organic aerosol that results from isoprene has different absorption than other $PM_{2.5}$ components,[75] which is not well captured by AOD. AAI provides some information about absorption profile, which helps our hybrid model perform much better in the Southeast and almost the same or a little better in the Northeast.

Second, our hybrid model integrated most variables that are known to be informative to $PM_{2.5}$ modeling and improved model performance. This study reminds the importance of hybrid framework and also proposes a possible neural network-based approach to implement that. Atmospheric mechanism is complex, and a single variable can only capture an incomplete picture. For example, AOD measures the light extinction due to aerosol in the whole atmosphere column. Different aerosols vary in terms of aerosol absorption, which can affect AOD. More complexly, even the same aerosol type could have various absorptions under different meteorological conditions and emission features.[39] This discovery suggests that when modeling $PM_{2.5}$ with AOD data, AAI (proxy for aerosol type), meteorological fields, and emission profiles are also necessary. There could be many unknown mechanisms intertwining with other variables. Multiple variables are not redundant but complementary, which can recover the original picture of atmospheric process and improve model performance to the best.

Third, we used a convolutional layer in the neural network for $PM_{2.5}$ modeling, which is an innovation of our study. Primarily used in computer science, a convolutional kernel is placed over nearby pixels to produce a convolutional layer. Similarly, we used convolutional layers in exposure assessment to aggregate variable values from nearby grid cells or monitoring sites. Previous studies incorporated nearby information by using nearby monitoring measurements, nearby road

density, or others, which were all prespecified. Our hybrid model takes multiple convolutional layers, which stand for various ways of aggregating nearby information, and lets learning algorithm decide their relative importance in the model. This approach is versatile and is able to model different neighboring influences, as well as potential interactions with other variables.

Last but not least, we used AOD data with missing values been filled by some calibration model. No further processing is required to deal with missing AOD data, which could have been lengthy and cumbersome in previous studies.

For the east-west gradient in model performance (Figure 1-1), previous studies also reported that correlation between MODIS AOD and ground-measured $PM_{2.5}$ is better in the eastern part but poor in the western part, and they attributed poor model performance to relative low $PM_{2.5}$ level and variability of terrain.[31,41] This study lends support to both statements. We quantified the relationship between model performance and elevation at each monitoring site and found a negative correlation despite of much noise (Figure S10). Similarly, a positive association exists between $PM_{2.5}$ level and model performance (Figure S10), which implies that the drop of model performance after 2010 is probably caused by substantive reduction in $PM_{2.5}$ level after 2010. This is also the reason why the Mountain region, with low $PM_{2.5}$ level, has poor model performance. A lower level of $PM_{2.5}$ means a lower signal-to-noise ratio, and model performance drops as model uncertainty keeps constant. Besides, the reduction of sulfate is mainly responsible for decreasing the $PM_{2.5}$ level. Sulfate is better modeled in GEOS-Chem than other major components like nitrate and ammonium,[76] so dropping sulfate causes unsatisfying model performance. For the same reason, we saw a less satisfying model performance in California despite its high $PM_{2.5}$ level, for the reason that California has a high amount of nitrate originated from vehicle exhaust compared

with the Eastern United States. This argument suggests that it would be informative to include sulfate in $PM_{2.5}$ modeling work in the future.

Our model performance is still good even at low $PM_{2.5}$ levels. To prove that, we fitted a spline regression of prediction $PM_{2.5}$ to measured $PM_{2.5}$. Linearity between measured and predicted $PM_{2.5}$ holds when $PM_{2.5}$ level is below 70 $\mu g/m^3$ and become less obvious above 80 $\mu g/m^3$ due to insufficient measurements (Figure 1-5). Bias at high concentration is less of our concern, since there are few days with the $PM_{2.5}$ level above 80 $\mu g/m^3$ in the study area. If constraining to monitored $PM_{2.5}$ below 35 $\mu g/m^3$, the EPA daily standard for $PM_{2.5}$, our hybrid model performed even better. Mean R^2 increased to 0.85; slope is close 1; and intercept is close to 0 (Table S1-5). Good model performance at low $PM_{2.5}$ concentrations enables epidemiologists to estimate the adverse effect of $PM_{2.5}$ even below EPA daily standard.

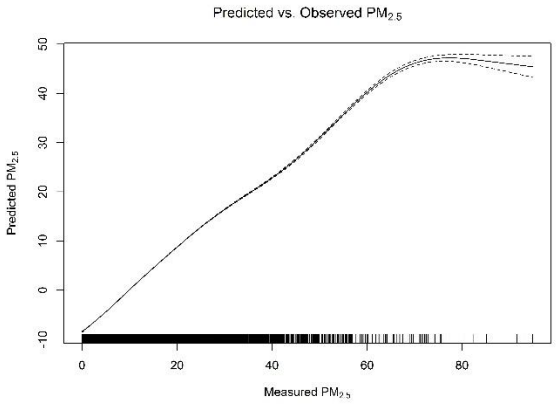


Figure 1-5. Relationship between Measured $PM_{2.5}$ and Predicted $PM_{2.5}$

We fit a penalized spline between measured $PM_{2.5}$ and predicted $PM_{2.5}$ without specifying degree of freedom. This figure is for year 2009.

Figure 1-2 visualizes the spatial distribution of annual and seasonal average of $PM_{2.5}$. There is also an east-west gradient of the $PM_{2.5}$ level. The Eastern and Central United States suffered relatively heavy $PM_{2.5}$ pollutions, except for the Appalachian Mountains, the Florida Peninsula, and some remote areas in the Northeast. The Southeastern United States, especially Alabama and Georgia, witnessed a high $PM_{2.5}$ level in summer and less noticeably in spring and autumn, which results from isoprene emission from trees. Isoprene emission from trees increases with temperature[74,77] and peaks in hot summer.[78] The Western United States had relatively low $PM_{2.5}$ levels, but the San Joaquin Valley, Salt Lake City, and Denver stood out for their abnormally high $PM_{2.5}$ level, which was also featured by clear seasonality and a high $PM_{2.5}$ level in winter. This is caused by temperature inversion in winter which prevents atmospheric convection and trapped air pollution near the surface. For temporal trend, the Eastern and Central United States witnessed a decreasing trend in the $PM_{2.5}$ level (Figure 1-3), which is caused by the reduction of sulfur dioxide from power plant emission. For seasonal cycle, $PM_{2.5}$ level peaks in summer in the Eastern and Central United States due to long-term transported sulfate from power plants and isoprene-related organic carbon. The winter peaks are probably caused by increased fuel burning for heat and local temperature inversion that prevents pollution dispersion.

Exposure assessments are essential for epidemiological studies. The traditional method of exposure assessment relies on nearest monitors, which constraints the number of available participants and introduces measurement errors. Besides, monitoring data from some monitors are intermittent. Our $PM_{2.5}$ predictions have temporal resolution of 1 day and spatial resolution of 1 km×1 km, which lifts the above limitations. Besides, our hybrid model performs still well at low concentrations. Linearity between predicted and monitored $PM_{2.5}$ still holds at low concentrations, without any signal of bias (Figure 1-5). Cross-validated R^2 indicates a good fit when daily $PM_{2.5}$

level is below 35 $\mu g/m^3$ (Table S1-5), which enable epidemiologists to assess the adverse effect of $PM_{2.5}$ even below the EPA standard. In the long term, there is little discrepancy between long-term averages of predicted and monitored $PM_{2.5}$, with difference below 1 $\mu g/m^3$ (Figure S1-9).

Some limitations remain. Our model requires quite a lot of variables, which limits the application in other countries. This data-intensive approach could be difficulty in other regions where public data is sparse. For regions with less data available, we might have to make a tradeoff between model performance and resolution. For example, instead of daily prediction $PM_{2.5}$ at 1 km×1 km, we may model annual average of $PM_{2.5}$ or at coarse spatial resolution. Besides, the chemical profile of $PM_{2.5}$ is not available in this framework. Previous epidemiological studies suggest various toxicities of $PM_{2.5}$ chemical components,[79,80] which is worthy of further investigation.

Supporting Information

Maps of the study area, details on US census division, details on GEOS-Chem, details on neural network and convolutional layers, details on calculating R^2, detailed results for model comparison, cross-validated R^2 by region and by season, and model performance at low concentrations.

Author Contributions

The manuscript was written by Qian Di, edited and approved by all authors. All authors have given approval to the final version of the manuscript.

Acknowledgement

This publication was made possible by USEPA grant R01 ES024332-01A1, RD83479801, and NIEHS grant ES000002. Its contents are solely the responsibility of the grantee and do not

necessarily represent the official views of the USEPA. Further, USEPA do not endorse the purchase of any commercial products or services mentioned in the publication.

References

1. 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. The lancet 2013;380:2224-60.

2. Slama R, Morgenstern V, Cyrys J, et al. Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model. Environmental Health Perspectives 2007:1283-92.

3. Franklin M, Zeka A, Schwartz J. Association between PM2. 5 and all-cause and specific-cause mortality in 27 US communities. Journal of Exposure Science and Environmental Epidemiology 2007;17:279-87.

4. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. Jama 2006;295:1127-34.

5. Gent JF, Triche EW, Holford TR, et al. Association of low-level ozone and fine particles with respiratory symptoms in children with asthma. Jama 2003;290:1859-67.

6. Schwartz J, Dockery DW, Neas LM. Is daily mortality associated specifically with fine particles? Journal of the Air & Waste Management Association (1995) 1996;46:927-39.

7. Halonen JI, Lanki T, Yli-Tuomi T, Kulmala M, Tiittanen P, Pekkanen J. Urban air pollution and asthma and COPD hospital emergency room visits. Thorax 2008.

8. Zanobetti A, Schwartz J, others. The effect of fine and coarse particulate air pollution on mortality: a national analysis. Environ Health Perspect 2009;117:898–903.

9. Boldo E, Medina S, Le Tertre A, et al. Apheis: Health impact assessment of long-term exposure to PM2. 5 in 23 European cities. European journal of epidemiology 2006;21:449-58.

10. Schwartz J. Harvesting and long term exposure effects in the relation between air pollution and mortality. American journal of epidemiology 2000;151:440-8.

11. Lippmann M, Ito K, Nadas A, Burnett R. Association of particulate matter components with daily mortality and morbidity in urban populations. Research report (Health Effects Institute) 2000:5-72, discussion 3-82.

12. Sarnat JA, Marmur A, Klein M, et al. Fine particle sources and cardiorespiratory morbidity: an application of chemical mass balance and factor analytical source-apportionment methods. Environmental Health Perspectives 2008;116:459.

13. Peng RD, Bell ML, Geyh AS, et al. Emergency Admissions for Cardiovascular and Respiratory Diseases and the Chemical Composition of Fine Particle Air Pollution. Environmental Health Perspectives 2009;117:957-63.

14. Peters A, Dockery DW, Muller JE, Mittleman MA. Increased Particulate Air Pollution and the Triggering of Myocardial Infarction. Circulation 2001;103:2810-5.

15. Schwartz J. Air pollution and hospital admissions for cardiovascular disease in Tucson. Epidemiology 1997:371-7.

16. Schwartz J, Morris R. Air pollution and hospital admissions for cardiovascular disease in Detroit, Michigan. American journal of epidemiology 1995;142:23-35.

17. Pope CA, Dockery DW. Health Effects of Fine Particulate Air Pollution: Lines that Connect. Journal of the Air & Waste Management Association 2006;56:709-42.

18. 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. American journal of respiratory and critical care medicine 2006;173:667–72.

19. Pinto JP, Lefohn AS, Shadwick DS. Spatial variability of PM2. 5 in urban areas in the United States. Journal of the Air & Waste Management Association 2004;54:440-9.

20. Beckerman BS, Jerrett M, Martin RV, van Donkelaar A, Ross Z, Burnett RT. Application of the deletion/substitution/addition algorithm to selecting land use regression models for interpolating air pollution measurements in California. Atmospheric Environment 2013;77:172-7.

21. Vienneau D, De Hoogh K, Beelen R, Fischer P, Hoek G, Briggs D. Comparison of land-use regression models between Great Britain and the Netherlands. Atmospheric Environment 2010;44:688-96.

22. Aguilera I, Sunyer J, Fernández-Patier R, et al. Estimation of Outdoor NO_x, NO_2, and BTEX Exposure in a Cohort of Pregnant Women Using Land Use Regression Modeling. Environmental Science & Technology 2008;42:815-21.

23. Ghassoun Y, Ruths M, Löwner M-O, Weber S. Intra-urban variation of ultrafine particles as evaluated by process related land use and pollutant driven regression modelling. Science of the Total Environment 2015;536:150-60.

24. Patton AP, Zamore W, Naumova EN, Levy JI, Brugge D, Durant JL. Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the Boston area. Environmental science & technology 2015;49:6051-60.

25. Hoek G, Beelen R, de Hoogh K, et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmospheric Environment 2008;42:7561-78.

26. Streets DG, Canty T, Carmichael GR, et al. Emissions estimation from satellite retrievals: A review of current capability. Atmospheric Environment 2013;77:1011-42.

27. Morain SA, Budge AM. Environmental Tracking for Public Health Surveillance: CRC Press; 2012.

28. Chu DA, Kaufman Y, Zibordi G, et al. Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS). Journal of Geophysical Research: Atmospheres (1984–2012) 2003;108.

29. van Donkelaar A, Martin RV, Brauer M, et al. Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. Environmental Health Perspectives 2010;118:847-55.

30. Van Donkelaar A, Martin RV, Park RJ. Estimating ground-level PM2. 5 using aerosol optical depth determined from satellite remote sensing. Journal of Geophysical Research: Atmospheres (1984–2012) 2006;111.

31. Engel-Cox JA, Holloman CH, Coutant BW, Hoff RM. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. Atmospheric Environment 2004;38:2495-509.

32. Wang J, Christopher SA. Intercomparison between satellite-derived aerosol optical thickness and PM2. 5 mass: implications for air quality studies. Geophysical research letters 2003;30.

33. Kloog I, Chudnovsky AA, Just AC, et al. A new hybrid spatio-temporal model for estimating daily multi-year $PM_{2.5}$ concentrations across northeastern USA using high resolution aerosol optical depth data. Atmospheric Environment 2014;95:581-90.

34. Kloog I, Koutrakis P, Coull BA, Lee HJ, Schwartz J. Assessing temporally and spatially resolved $PM_{2.5}$ exposures for epidemiological studies using satellite aerosol optical depth measurements. Atmospheric Environment 2011;45:6267-75.

35. Lee M, Kloog I, Chudnovsky A, et al. Spatiotemporal prediction of fine particulate matter using high-resolution satellite images in the Southeastern US 2003–2011. Journal of Exposure Science and Environmental Epidemiology 2015.

36. Liu Y, Paciorek CJ, Koutrakis P. Estimating Regional Spatial and Temporal Variability of $PM_{2.5}$ Concentrations Using Satellite Data, Meteorology, and Land Use Information. Environmental Health Perspectives 2009;117:886-92.

37. Li F, Ramanathan V. Winter to summer monsoon variation of aerosol optical depth over the tropical Indian Ocean. Journal of Geophysical Research: Atmospheres (1984–2012) 2002;107:AAC 2-1-AAC 2-13.

38. Schaap M, Apituley A, Timmermans R, Koelemeijer R, Leeuw Gd. Exploring the relation between aerosol optical depth and $PM_{2.5}$ at Cabauw, the Netherlands. Atmospheric Chemistry and Physics 2009;9:909-25.

39. Dubovik O, Holben B, Eck TF, et al. Variability of absorption and optical properties of key aerosol types observed in worldwide locations. Journal of the atmospheric sciences 2002;59:590-608.

40. Hu R-M, Sokhi R, Fisher B. New algorithms and their application for satellite remote sensing of surface PM2. 5 and aerosol absorption. Journal of Aerosol Science 2009;40:394-402.

41. Al-Saadi J, Szykman J, Pierce RB, et al. Improving national air quality forecasts with satellite aerosol observations. Bulletin of the American Meteorological Society 2005;86:1249-61.

42. Bey I, Jacob DJ, Yantosca RM, et al. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. Journal of Geophysical Research 2001;106:23073.

43. Byun DW, Ching J. Science algorithms of the EPA Models-3 community multiscale air quality (CMAQ) modeling system. 1999.

44. Bessagnet B, Hodzic A, Vautard R, et al. Aerosol modeling with CHIMERE—preliminary evaluation at the continental scale. Atmospheric Environment 2004;38:2803-17.

45. Jun M, Stein ML. Statistical comparison of observed and CMAQ modeled daily sulfate levels. Atmospheric Environment 2004;38:4427-36.

46. Cordero L, Malakar N, Wu Y, Gross B, Moshary F, Ku M. Assessing satellite based $PM_{2.5}$ estimates against CMAQ model forecasts. In: Comeron A, Kassianov EI, Schäfer K, Stein K, Gonglewski JD, editors.2013 2013-10-17. p. 88900U.

47. van Donkelaar A, Martin RV, Brauer M, Boys BL. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter. Environmental health perspectives 2015;123:135.

48. Di Q, Schwartz J, Koutrakis P. A Hybrid Prediction Model for $PM_{2.5}$ Mass and Components Using a Chemical Transport Model and Land Use Regression. Atmospheric Environment 2015.

49. King MD, Kaufman YJ, Menzel WP, Tanre D. Remote sensing of cloud, aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer (MODIS). Geoscience and Remote Sensing, IEEE Transactions on 1992;30:2-27.

50. Salomonson VV, Barnes W, Maymon PW, Montgomery HE, Ostrow H. MODIS: Advanced facility instrument for studies of the Earth as a system. Geoscience and Remote Sensing, IEEE Transactions on 1989;27:145-53.

51. Remer LA, Kaufman YJ, Tanré D, et al. The MODIS Aerosol Algorithm, Products, and Validation. Journal of the Atmospheric Sciences 2005;62:947-73.

52. Chudnovsky A, Lyapustin A, Wang Y, Schwartz J, Koutrakis P. Analyses of high resolution aerosol data from MODIS satellite: a MAIAC retrieval, southern New England, US. First International Conference on Remote Sensing and Geoinformation of Environment; 2013: International Society for Optics and Photonics. p. 87951E-E-7.

53. Lyapustin A, Martonchik J, Wang Y, Laszlo I, Korkin S. Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables. Journal of Geophysical Research: Atmospheres (1984–2012) 2011;116.

54. Lyapustin A, Wang Y, Laszlo I, et al. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. Journal of Geophysical Research: Atmospheres (1984–2012) 2011;116.

55. van Donkelaar A, Martin RV, Park RJ. Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing. Journal of Geophysical Research 2006;111.

56. Di Q, Schwartz J. Using Chemical Transport Model to Fill in the Missingness of Satellite-Based AOD. Atmospheric Environment 2015.

57. Drury E, Jacob DJ, Wang J, Spurr RJ, Chance K. Improved algorithm for MODIS satellite retrievals of aerosol optical depths over western North America. Journal of Geophysical Research: Atmospheres (1984–2012) 2008;113.

58. Vermote E. MOD09A1 MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V006. In: DAAC NELP, ed.2015.

59. Isakov V, Touma J, Touma J, et al. Estimating Fine Particulate Matter Component Concentrations and Size Distributions Using Satellite-Retrieved Fractional Aerosol Optical Depth: Part 2—A Case Study. Journal of the Air & Waste Management Association 2007;57:1360-9.

60. Liu Y. Mapping annual mean ground-level $PM_{2.5}$ concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. Journal of Geophysical Research 2004;109.

61. 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.

62. Herman J, Bhartia P, Torres O, Hsu C, Seftor C, Celarier E. Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data. J Geophys Res 1997;102:911-16.

63. Tegen I, Werner M, Harrison S, Kohfeld K. Relative importance of climate and land use in determining present and future global soil dust emission. Geophysical Research Letters 2004;31.

64. Torres O, Bhartia P, Herman J, Ahmad Z, Gleason J. Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis. Journal of Geophysical Research: Atmospheres (1984–2012) 1998;103:17099-110.

65. Deborah Stein-Zweers aPV. OMI/Aura Multi-wavelength Aerosol Optical Depth and Single Scattering Albedo Daily L3 Global 0.25x0.25 deg Lat/Lon Grid,version 003. In: Center NGSF, ed.2012.

66. Torres O, Tanskanen A, Veihelmann B, et al. Aerosols and surface UV products from Ozone Monitoring Instrument observations: An overview. Journal of Geophysical Research: Atmospheres (1984–2012) 2007;112.

67. Kloog I, Nordio F, Coull BA, Schwartz J. Incorporating Local Land Use Regression And Satellite Aerosol Optical Depth In A Hybrid Model Of Spatiotemporal PM $_{2.5}$ Exposures In The Mid-Atlantic States. Environmental Science & Technology 2012;46:11913-21.

68. Didan K. MOD13A2 MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V006. In: DAAC NELP, ed.2015.

69. Lee H, Liu Y, Coull B, Schwartz J, Koutrakis P. A novel calibration approach of MODIS AOD data to predict PM2. 5 concentrations. Atmos Chem Phys 2011;11:7991-8002.

70. Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. World map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 2006;15:259-63.

71. Bishop CM. Neural networks for pattern recognition: Oxford university press; 1995.

72. Haykin S, Network N. A comprehensive foundation. Neural Networks 2004;2.

73. LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks 1995;3361.

74. Sharkey TD, Wiberley AE, Donohue AR. Isoprene emission from plants: why and how. Annals of Botany 2008;101:5-18.

75. Claeys M, Graham B, Vas G, et al. Formation of secondary organic aerosols through photooxidation of isoprene. Science 2004;303:1173-6.

76. Park RJ, Jacob DJ, Field BD, Yantosca RM, Chin M. Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy. Journal of Geophysical Research: Atmospheres (1984–2012) 2004;109.

77. Sharkey TD, Singsaas EL, Vanderveer PJ, Geron C. Field measurements of isoprene emission from trees in response to temperature and light. Tree Physiology 1996;16:649-54.

78. Geron C, Harley P, Guenther A. Isoprene emission capacity for US tree species. Atmospheric Environment 2001;35:3341-52.

79. Franklin M, Koutrakis P, Schwartz P. The role of particle composition on the association between $PM_{2.5}$ and mortality. Epidemiology 2008;19:680-9.

80. Dai L, Zanobetti A, Koutrakis P, Schwartz JD. Associations of fine particulate matter species with mortality in the United States: a multicity time-series analysis. Environmental health perspectives 2014;122:837.

Supplementary Material

1. Study Area and Sub-regions

Figure S1-1 displays the continental United States and $PM_{2.5}$ monitoring site. The whole study area was divided into sub-region to calculate regional model. Sub-region was defined by US Census Division (Figure S1-2). States in each region are presented in Table S1-1.

2. GEOS-Chem Outputs

GEOS-Chem simulates different chemical components of $PM_{2.5}$ rather than the total mass. Previous studies and GEOS-Chem manual have proposed several combinations of output fields to simulate $PM_{2.5}$.[1,2] By referring to previous studies and doing testing, we decided $PM_{2.5}$ to be the sum of the output fields specified in Table S1-2. Logically, some of the output fields (e.g. PAN, PMN, PPM, and R4N2) are not components of $PM_{2.5}$ and should be excluded from estimating $PM_{2.5}$; but practically we found that including them improved correlation. Since our task is achieving better model performance, we finally decided to include those output fields in estimating $PM_{2.5}$.

3. Filling in Missing AOD Values

Missing values in AOD was filled in by calibrating column aerosol mass from GEOS-Chem to satellite-based AOD using a neural network. We used meteorological variables, land-use terms and absorbing aerosol index as additional input layers of the neural network. The calibration work was done at daily basis. Since all input variables have no or almost no missing values, we can use the trained neural network to predict AOD for the whole study area, even at place with missing AOD values. Details of this work were specified somewhere else.[3]

4. Regional Dummy Variable

The process of creating regional dummy variables from major climate types is displayed in Figure S1-3. Data on climate types were obtained from somewhere else.[4]

5. Convolution Layers

A convolutional layer is computed by taking a convolutional kernel on an input layer over the entire area. The convolutional kernel function has constant parameters and can take many forms, such as mean, median or weighted averages. In a convolution layer, to obtain value of each cell, a kernel function is applied to neighboring cells and produces a scalar summary. Applying different kernel functions yields different convolutional layers.

Assuming we need to compute a convolutional layer for population density at monitoring sites. We chose the kernel function to be weighted average with weight proportional to the inverse of distance. For every site, we took distance-inversed weighted average of nearby population density. That is, for any monitor *k*, we have:

$$Y_k = \frac{\sum_{i=1}^{n} w_{ik} y_i}{\sum_{i=1}^{n} w_{ik}}$$

where Y_k is the value of the convolutional layer at monitor *k*, y_i is population density at grid cell *i*; $w_{ik} \propto 1/d_{ik}^2$; d_{ik} is the distance between grid cell *i* and monitoring site *k*.

By changing other form of convolutional kernels, such as $w_{ik} \propto 1/d_{ik}^3$, $w_{ik} \propto 1/d_{ik}$, we could obtain other convolutional layers.

The above example computes a convolutional layer over spatial neighborhood. It is also possible to have a convolutional layer over nearby days. For instance, a temporal convolutional layer

takes moving average of predicted $PM_{2.5}$ over the nearest *D* days. The weight on each day is proportional to the day difference:

$$Y_k = \frac{\sum_{i=-D}^{D} w_i y_k}{\sum_{i=-D}^{D} w_i}$$

A convolutional layer is some form of aggregation of neighboring information for a variable. Different convolutional layers capture different levels of aggregation. Aggregation of nearby information makes it possible to capture autocorrelation. Multiple convolutional layers were put into a neural network as separate input layers. The neural network adjusts weights on each convolutional layer in learning process, in which various combinations of convolutional layers are explored until the optimized one is found. Combination of convolutional layers can simulate complex autocorrelation and possible interaction with other variables. More introductions about convolutional layers were specified somewhere else.[5]

6. Neural Network

Figure S1-4 presents the conceptual model of a neural network, in which interconnected "neurons" approximate some function. Input layer, hidden layer, and output layer all consist of neurons. Each neuron in the input layer represents one predictor, which could be AOD, OMI AAI, one land-use term, and others; neuron in the output layer represents predicted $PM_{2.5}$, which is trained to $PM_{2.5}$ monitoring data from AQS. Neural network works in the following way: each neuron takes weighted average from neurons in the previous layer and transforms the average by some nonlinear function (e.g. log-sigmoid or hyperbolic tangent sigmoid function). Nonlinear function accounts for possible interactions and nonlinearities. The same process applies to the neurons in the next layer until it reaches to the output layer. Initial weight for each neuron is

randomly assigned and updated in learning process. In the learning process, a learning algorithm updates weights based on some rule to minimize the difference between output and target value. In this study, we used Levenberg-Marquardt optimization algorithm.[6] More hidden layers result in better capacity to model nonlinearity but more computation time. One hidden layer was suggest to avoid overfitting.[7] Our study involves complex and nonlinear atmospheric process. We used two hidden layer to have a compromise between model fit and computation time. Previous study suggested the number of neurons to be in the range of $(2\sqrt{n} + u \sim 2n + 1)$, in which n is number of input variables and u is the number of output variables.[8] We put 15 neurons in each hidden layer. To improve efficiency, we normalized input variables within the range of (-1, 1) as preprocessing.

Convolutional layers were incorporated into a neural network just as ordinary input layers. Some convolutional layers are for predicted $PM_{2.5}$, which requires us to fit a neural network in two stages. First, we fitted a neural network with AOD, OMI AAI, CTM outputs, $PM_{2.5}$ vertical profile, meteorological variables, land-use terms and their convolutional layers, convolutional layers for nearby $PM_{2.5}$ measurements, NDVI, surface reflectance, monthly/regional dummy variables. The target variable was AQS monitored $PM_{2.5}$. After training, we obtained intermediated $PM_{2.5}$ prediction from the neural network and computed both temporal and spatial convolutional layers for predicted $PM_{2.5}$. We fitted the neural network again with above variables and convolutional layers for predicted $PM_{2.5}$. The neural network yielded predicted $PM_{2.5}$ with better model performance, which is the final predicted $PM_{2.5}$. Correspondingly, we predicted $PM_{2.5}$ at 1 km×1 km grid cells in a two-stage way. The two-stage modeling approach is displayed in Figure S1-5.

7. Calculation R^2

We used total R^2, spatial R^2 and temporal R^2 to quantify model performance. R^2s between cross-validated predicted $PM_{2.5}$ and monitored $PM_{2.5}$ were calculated based on description specified somewhere else.[9] We fitted a regression of predicted $PM_{2.5}$ on monitored $PM_{2.5}$ and calculated R^2 for this regression model, which was total R^2. We fitted a regression of annual mean of predicted $PM_{2.5}$ on annual mean of monitored $PM_{2.5}$ and calculated R^2 for this regression model, which was spatial R^2. We fitted a regression of Delta Predicted on Delta Monitored and computed R^2 for this model, which was temporal R^2. Delta Predicted was the difference between predicted $PM_{2.5}$ and annual mean of predicted $PM_{2.5}$ at that location; Delta Monitor was defined in a similar way, the difference between monitored $PM_{2.5}$ and annual mean of monitored $PM_{2.5}$ at that location.

If restricting monitoring records to a time period, to a certain geographic area, or below certain concentrations, we can compute R^2 by season (Table S1-3), by region (Table S1-4) or at low concentrations (Table S1-5).

8. Model Comparison

a. Hybrid model

Model with full set of variables performed better than models with the subset. To illustrate that, for each year, we fitted a model with GEOS-Chem, meteorological variables, and land-use terms as input variables. This was the benchmark model. We fitted the following three hybrid models by adding the following variables to the benchmark model: (1) convolution terms of land-use variables (Hybrid Model 1); (2) convolution terms of land-use variables, Aqua-AOD, and OMI AAI (Hybrid Model 2); (3) convolution terms of land-use variables, Aqua-AOD, Terra-AOD and OMI AAI to the benchmark model (Hybrid Model 3). Compared with the Hybrid Model 3, the

first two hybrid models were subset models. We calculated the cross-validated total R^2. Figure S1-6 visualizes the improvement of model performance and indicates those improvements are of statistical significance. Hybrid Model 3, the one with full list of variables, improved the model performance more than models with subsets of variables. Cross-validated total R^2 were reported to eliminate overfitting.

b. Scaling factor

Previous study used GEOS-Chem output to calculate the percentage of ground aerosol concentration in total column concentration as scaling factor and multiplied AOD by this scaling factor for calibration.[10] We argue that the potential relationship between AOD and a scaling factor can be rather complex. We put the scaling factor input into neural network as a separate layer and let the neural network simulate possible complex relationships. To justify which approach is superior, we first fitted a benchmark model with Aqua/Terra-AOD, meteorological variables and land-use terms for each year. Then we included scaling factor either by multiplying with AOD (denoted as Model 1) or as a separate input layer (denoted as Model 2) and calculated the cross-validated total R^2s. Figure S1-7 indicates that including scaling factor as a separate layer significantly improves model fit even more.

c. Convolutional Layers

Convolutional layers account for temporal and spatial autocorrelation and improve model fit. To demonstrate that, we fitted a benchmark model with GEOS-Chem outputs, AOD, OMI AAI, meteorological variables and land-use terms. On the basis of that, we added either convolutional layers of nearby monitored $PM_{2.5}$ and improved model with statistical significance. If further adding convolution terms of nearby predicted $PM_{2.5}$, model fit can improve again with statistical significance (Figure S1-8).

9. Model Diagnosis

We evaluated the model performance at low concentrations (daily average<35 $\mu g/m^3$). Table S1-5 presents the cross-validated results when monitored $PM_{2.5}$ below 35 $\mu g/m^3$. Compared with Table 1-1, there was a slight increase in R^2 and mean annual R^2 increased 0.01 to 0.85, which indicates good model performance at low concentrations at well. Figure S1-9 visualizes the difference between long-term average of predicted $PM_{2.5}$ and monitored $PM_{2.5}$. The difference is within ± 1 $\mu g/m^3$ for most areas of the United States. Marine, Colorado, and Wisconsin witnessed slight overestimation of long-term averaged $PM_{2.5}$ (<1 $\mu g/m^3$); while Montana has slight underestimation.

Besides, we also investigated the association between model performance with elevation and average $PM_{2.5}$ level (Figure S1-10). We computed the overall model performance from 2000~2012 for each site and its correlation with *in situ* elevation and average $PM_{2.5}$ in the same period. There is a negative correlation between model performance and elevation, with $R^2 = -0.2470$ and p-value<0.0001. Average $PM_{2.5}$ positively associates with model performance, with $R^2 = 0.2697$ and p-value<0.0001.

Table S1-1. Census Bureau Division

Division	Name	States
Division 1	New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Division 2	Middle Atlantic	New Jersey, New York, Pennsylvania
Division 3	East North Central	Indiana, Illinois, Michigan, Ohio, Wisconsin
Division 4	West North Central	Iowa, Nebraska, Kansas, North Dakota, Minnesota, South Dakota, Missouri
Division 5	South Atlantic	Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia
Division 6	East South Central	Alabama, Kentucky, Mississippi, Tennessee
Division 7	West South Central	Arkansas, Louisiana, Oklahoma, Texas
Division 8	Mountain	Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming
Division 9	Pacific	Alaska, California, Hawaii, Oregon, Washington

Note: contents are based on http://www2.census.gov/geo/pdfs/maps-data/maps

Table S1-2. GEOS-Chem Outputs Used to Calculate Total $PM_{2.5}$

Name	Full Name	g/mol
PAN	Peroxyacetyl Nitrate	121
PMN	Peroxymethacroyl Nitrate	147
PPN	Lumped Peroxypropionyl Nitrate	135
R4N2	Lumped Alkyl Nitrate	119
SO4	Sulfate	96
SO4s	Sulfate on surface of sea-salt aerosol	96
NH4	Ammonium	18
NIT	Inorganic nitrates	62
NITs	Inorganic nitrates on surface of sea-salt aerosol	62
BCPI	Hydrophilic black carbon aerosol	12
OCPI	Hydrophilic organic carbon aerosol	12
BCPO	Hydrophobic black carbon aerosol	12
OCPO	Hydrophobic organic carbon aerosol	12
DST1	Dust aerosol, Reff = 0.7 microns	29
DST2	Dust aerosol, Reff = 1.4 microns	29
DST3	Dust aerosol, Reff = 2.4 microns	29
SALA	Accumulation mode sea salt aerosol (Reff = 0.01 – 0.5 microns)	31.4
BrNO3	Bromine nitrate	142
MPN	Methyl peroxy nitrate	93
ISOPN	Isoprene nitrate	147
PROPNN	Propanone nitrate	109
MMN	Nitrate from MACR + MVK	149

Table S1-3. Cross-validated R^2 for Seasons

Year	Spring R^2	RMSE	Summer R^2	RMSE	Autumn R^2	RMSE	Winter R^2	RMSE	Annual R^2	RMSE
2000	0.85	2.60	0.89	2.81	0.88	3.48	0.79	4.62	0.86	3.35
2001	0.85	2.85	0.92	2.70	0.78	4.25	0.82	3.48	0.84	3.58
2002	0.87	2.12	0.90	3.22	0.89	2.79	0.83	3.50	0.88	2.99
2003	0.89	2.34	0.92	2.54	0.87	2.89	0.82	3.50	0.88	2.80
2004	0.87	2.30	0.93	2.18	0.89	2.64	0.83	3.48	0.88	2.69
2005	0.87	2.34	0.92	2.58	0.89	2.88	0.81	3.47	0.88	2.94
2006	0.88	2.17	0.91	2.46	0.86	2.58	0.80	3.72	0.86	2.77
2007	0.86	2.87	0.89	2.79	0.86	2.83	0.83	3.24	0.87	2.95
2008	0.83	2.13	0.88	2.68	0.85	2.54	0.80	3.45	0.85	2.64
2009	0.85	1.96	0.87	1.92	0.78	2.89	0.81	3.52	0.82	2.73
2010	0.80	2.60	0.83	2.50	0.78	2.72	0.80	3.53	0.81	2.85
2011	0.81	2.24	0.84	2.59	0.79	2.56	0.77	3.65	0.81	2.83
2012	0.72	2.60	0.72	2.93	0.75	3.70	0.74	3.29	0.74	3.15
Mean	0.84	2.39	0.88	2.61	0.84	2.98	0.80	3.57	0.84	2.94

Table S1-4. Cross-validated R^2 for Sub-regions

Year	New England	Middle Atlantic	East North Central	West North Central	South Atlantic	East South Central	West South Central	Mountain	Pacific	National
2000	0.82	0.87	0.89	0.88	0.85	0.83	0.86	0.75	0.84	0.86
2001	0.87	0.89	0.91	0.57	0.86	0.87	0.86	0.79	0.83	0.84
2002	0.90	0.91	0.92	0.92	0.90	0.88	0.87	0.79	0.82	0.88
2003	0.89	0.89	0.92	0.91	0.91	0.88	0.81	0.73	0.84	0.88
2004	0.90	0.90	0.92	0.91	0.91	0.85	0.88	0.80	0.83	0.88
2005	0.89	0.88	0.93	0.92	0.90	0.88	0.87	0.75	0.83	0.88
2006	0.92	0.89	0.90	0.88	0.90	0.88	0.78	0.69	0.79	0.86
2007	0.91	0.89	0.90	0.91	0.87	0.88	0.83	0.78	0.83	0.87
2008	0.88	0.88	0.89	0.90	0.87	0.87	0.83	0.75	0.80	0.85
2009	0.86	0.82	0.89	0.78	0.82	0.85	0.83	0.78	0.79	0.82
2010	0.88	0.83	0.87	0.85	0.81	0.80	0.77	0.76	0.70	0.81
2011	0.82	0.79	0.85	0.86	0.84	0.86	0.76	0.76	0.78	0.81
2012	0.78	0.75	0.80	0.79	0.78	0.81	0.73	0.70	0.71	0.74
Mean	0.87	0.86	0.89	0.85	0.86	0.86	0.82	0.76	0.80	0.84

Table S1-5. Cross-validated R^2 for $PM_{2.5}$<35 µg/m^3

Year	Total R^2	MSE	Spatial R^2	MSE	Temporal R^2	MSE	Bias	Slope
2000	0.86	2.71	0.85	1.30	0.85	2.44	0.69	0.96
2001	0.87	2.59	0.87	1.17	0.86	2.37	0.66	0.96
2002	0.87	2.43	0.86	1.12	0.87	2.21	0.66	0.95
2003	0.87	2.36	0.86	1.12	0.88	2.14	0.58	0.96
2004	0.88	2.28	0.81	1.27	0.89	2.05	0.50	0.97
2005	0.88	2.49	0.83	1.32	0.88	2.23	0.56	0.97
2006	0.87	2.38	0.82	1.21	0.87	2.13	0.49	0.97
2007	0.87	2.44	0.84	1.16	0.87	2.21	0.57	0.96
2008	0.86	2.24	0.78	1.16	0.87	2.01	0.44	0.97
2009	0.84	2.22	0.81	1.01	0.84	2.04	0.42	0.97
2010	0.82	2.53	0.82	1.13	0.82	2.27	0.62	0.96
2011	0.82	2.50	0.81	1.06	0.82	2.27	0.61	0.96
2012	0.76	2.56	0.74	1.08	0.76	2.33	0.64	0.96
Mean	0.85	2.44	0.82	1.16	0.85	2.21	0.57	0.96

Table S1-6. Model Performance Compared with Existing Regional Studies

Northeastern United States	Hybrid Model	Existing Study [a]	Existing Study [b]	Southeastern United States	Hybrid Model	Existing Study [c]
2000	0.82	0.73		**2000**	0.85	N/A
2001	0.87	0.84		**2001**	0.86	N/A
2002	0.90	0.85		**2002**	0.90	N/A
2003	0.89	0.85	0.89	**2003**	0.91	0.72~0.83
2004	0.90	0.85	0.89	**2004**	0.91	0.74~0.80
2005	0.89	0.84	0.88	**2005**	0.90	0.75~0.83
2006	0.92	0.86	0.89	**2006**	0.90	0.74~0.84
2007	0.91	0.90	0.90	**2007**	0.87	0.70~0.85
2008	0.88	0.81	0.88	**2008**	0.87	0.69~0.78
2009	0.86	N/A	0.86	**2009**	0.82	0.66~0.78
2010	0.88	N/A	0.90	**2010**	0.81	0.65~0.80
2011	0.82	N/A	0.82	**2011**	0.84	0.63~0.79
2012	0.78	N/A		**2012**	0.78	N/A

a. Results cited from Table 1, Kloog et al. (2011);[9] b. Results cited from Table 1, Kloog et al. (2014);[11] c. Results cited from Table 2, Lee et al. (2015).[12]

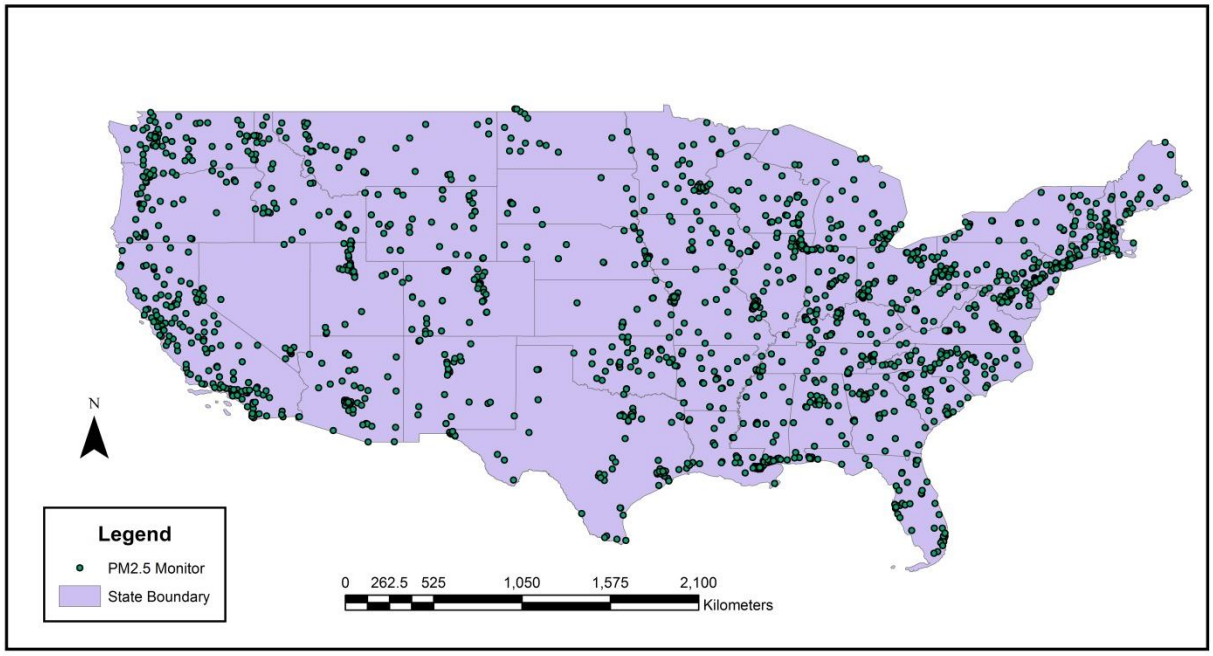


Figure S1-1. Map of the Study Area with $PM_{2.5}$ Monitors

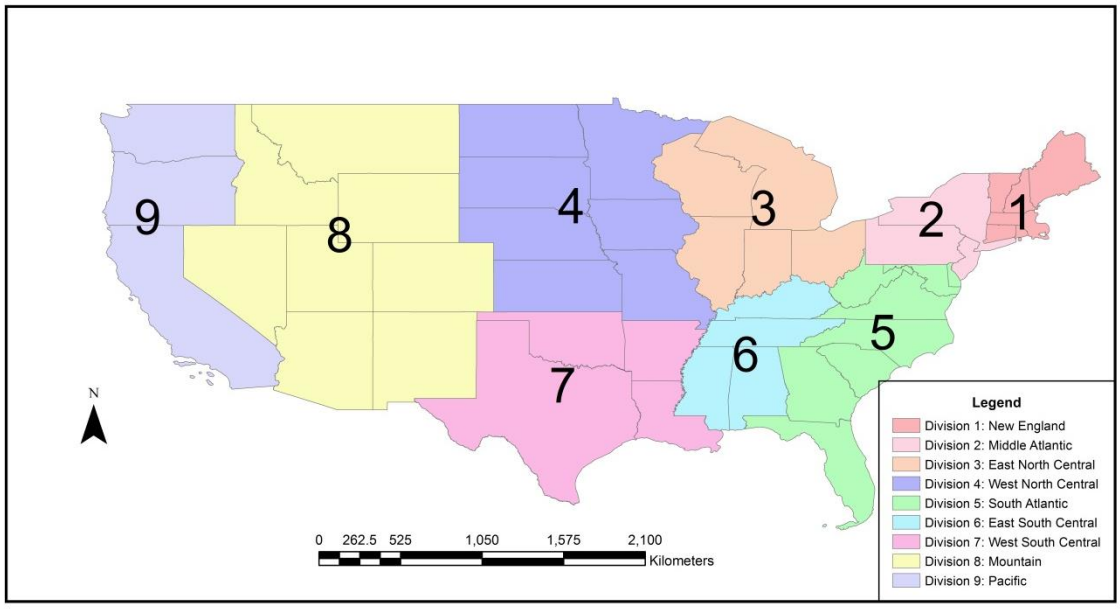


Figure S1-2. U.S. Census Division

Hawaii, Alaska and other overseas territories are not included in the study area and not shown in the figure.

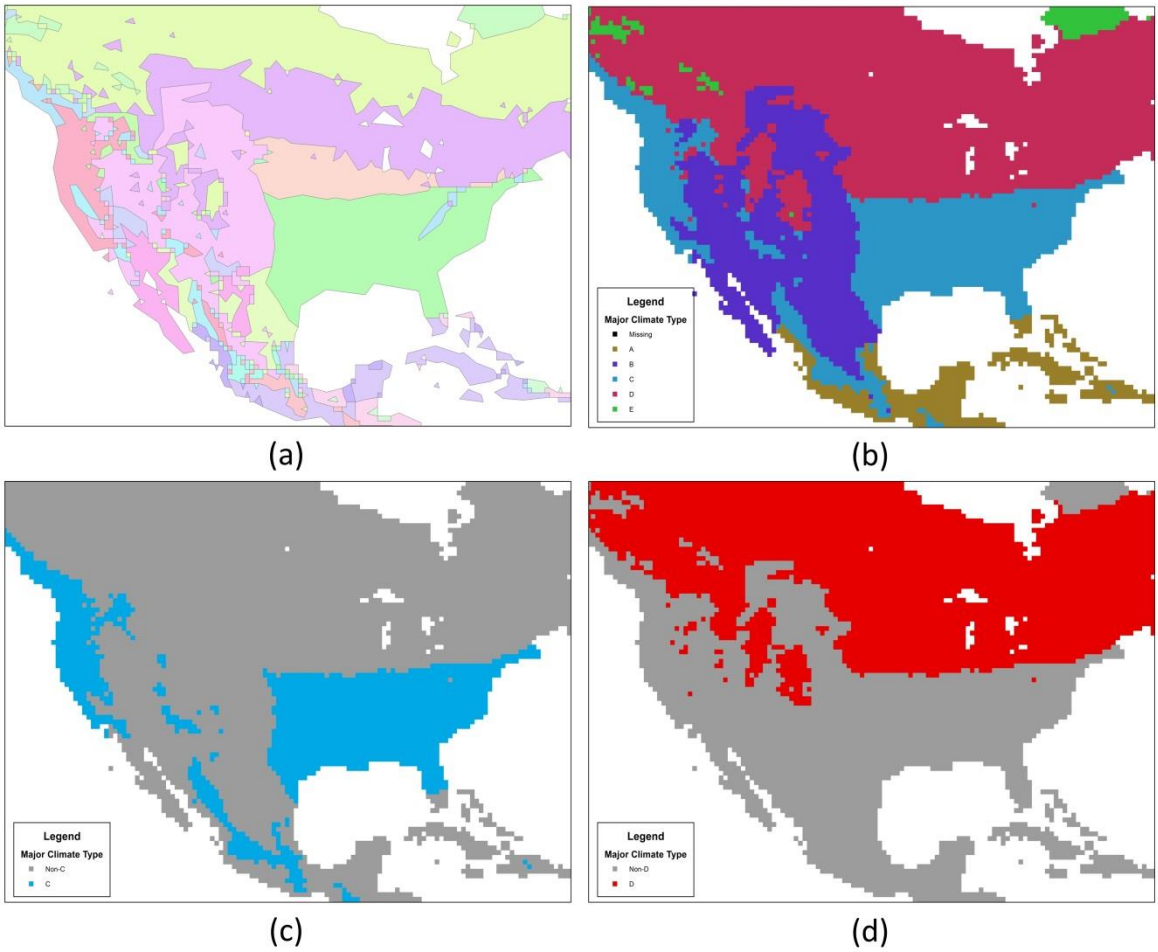


Figure S1-3. Regional Dummy from Major Climate Types

The original climate classification is presented in (Figure (a)). We aggregated minor climate types to five major types: A (equatorial climates), B (arid climates), C (warm temperate climates), D (snow climates) and E (Polar climates) in (Figure (b)).[4] For each major climate type, we created a regional dummy. Take major climate type C as an example, we assigned regional dummy to be 1 for all grid cells covered by climate type C (blue areas in Figure (c)) and 0 otherwise (grey areas in Figure (c)). The same also held for climate type D (Figure (d)) and other types.

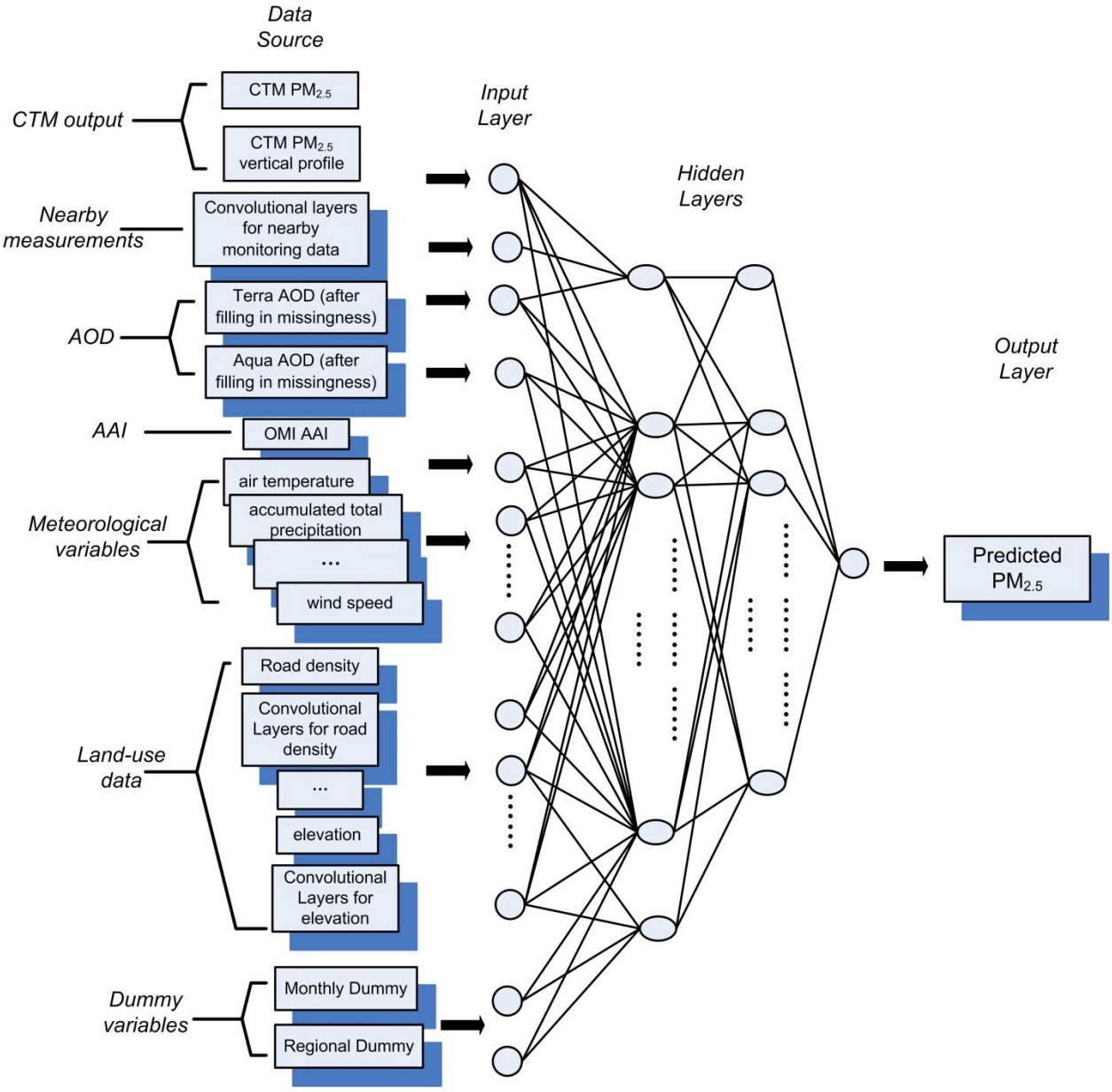


Figure S1-4. Conceptual Model of Neural Model

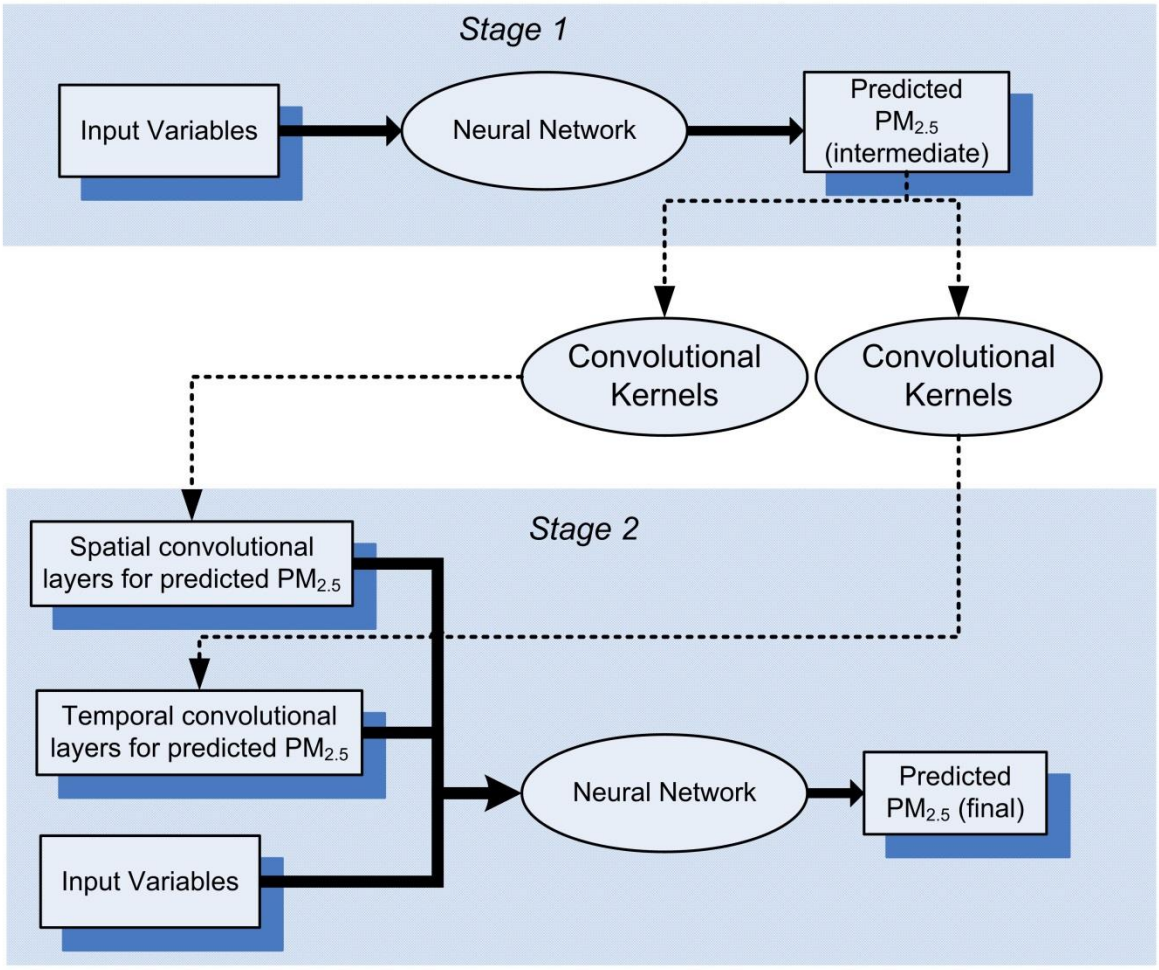


Figure S1-5. Flowchart of Two-stage Modeling Approach

Note: We used *Input Variables* to stand for all input variables, which include AOD, OMI AAI, CTM outputs, $PM_{2.5}$ vertical profile, meteorological variables, land-use terms and their convolutional layers, convolutional layers for nearby $PM_{2.5}$ measurements, NDVI, surface reflectance, and monthly/regional dummy variables.

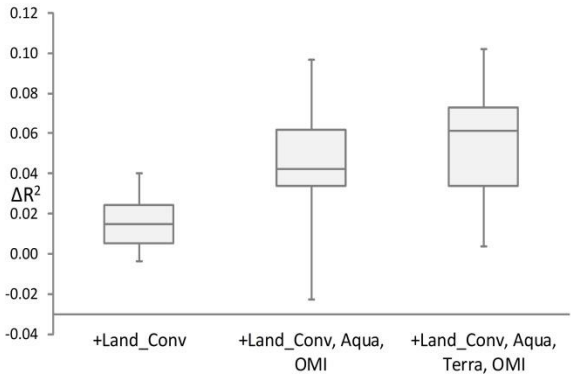


Figure S1-6. Improvement of Model Fit after Adding Extra Variables

After adding convolution terms of land-use variables to the benchmark model (Hybrid Model 1), model fit improved ($p<0.0001$; one-sided paired Wilcoxon signed rank test, the same below). If further adding Aqua-AOD and OMI AAI (Hybrid Model 2), model fit improved ($p=0.0133$). Further adding Terra-AOD also improved model fit ($p=0.0054$).

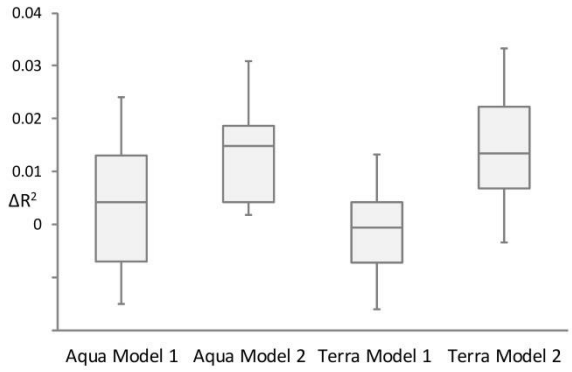


Figure S1-7. Improvement of Model Fit after Adding Scaling Factor

Input layers of benchmark model: AOD, meteorological variables and land-use terms; input layers of Model 1: AOD*scaling factor, meteorological variables and land-use terms; input layers of Model 2: AOD, scaling factor, meteorological variables and land-use terms. Y-axis is the improvement of R^2 compared with benchmark model. For Model 1, R^2 increased by 0.0034 (p=0.0020, one-sided paired Wilcoxon signed rank test, the same below) and -0.0008 (p=0.6890) for Aqua and Terra respectively; for Model 2, R^2 increased by 0.0139 (p=0.0010) and 0.0149 (p=0.0021) for Aqua and Terra respectively. Model 2 demonstrated further improvement than Model 1 (p=0.0020 for Aqua and p=0.0002 for Terra).

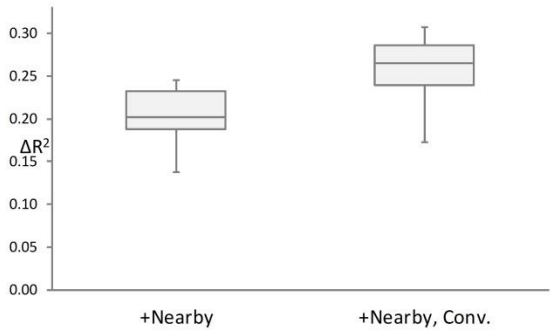


Figure S1-8. Improvement of Model Fit after Adding Convolutional Layers

Input layers of benchmark model: GEOS-Chem outputs, Aqua-AOD, Terra-AOD, OMI AAI, meteorological variables and land-use terms. After adding nearby monitoring data as a convolutional layer into the model, R^2 increased by 0.2043 (p=0.0001). If further adding convolutional layer of predicted $PM_{2.5}$, R^2 increased by 0.0517 (p=0.0001).

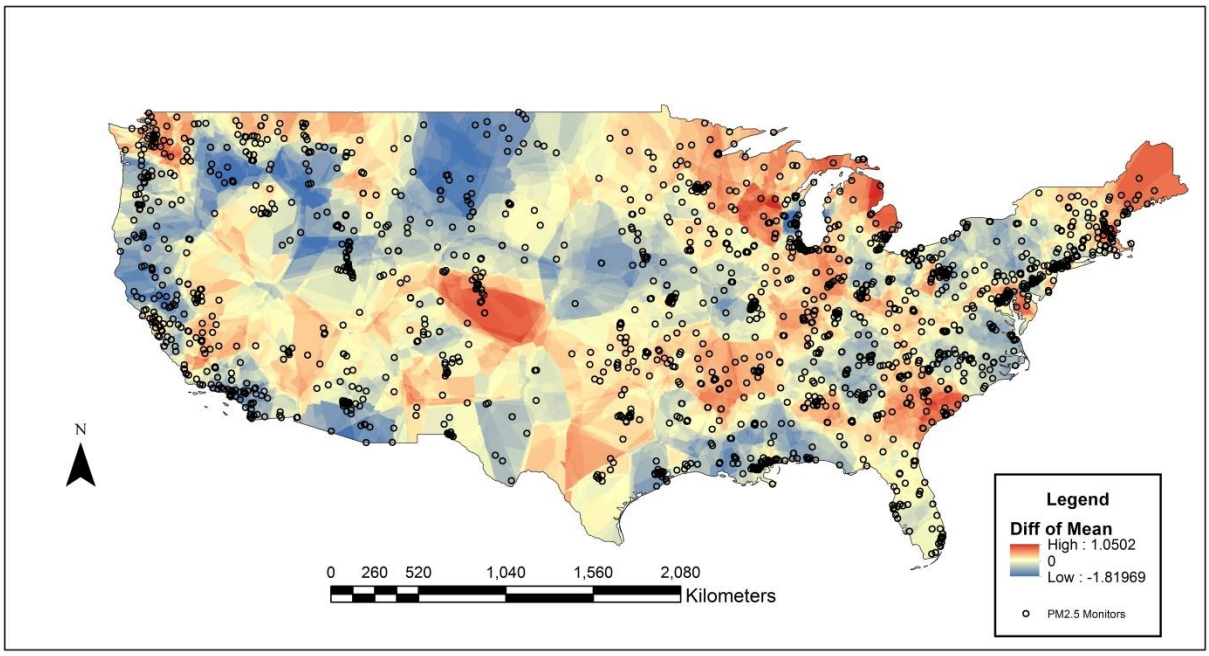


Figure S1-9. Difference between Predicted Long-Term Average and Monitored Long-Term Average

Note: We took the difference between long-term average of predicted $PM_{2.5}$ and long-term average of monitored $PM_{2.5}$ (from 2000 to 2012) at each monitoring site and interpolated it to places without monitors. After visualization, red color means predicted $PM_{2.5}$>monitored $PM_{2.5}$ in the long-term (overestimation) and blue color means the opposite. Light yellow indicates prediction and monitoring values are close. Unit in this map is $\mu g/m^3$.

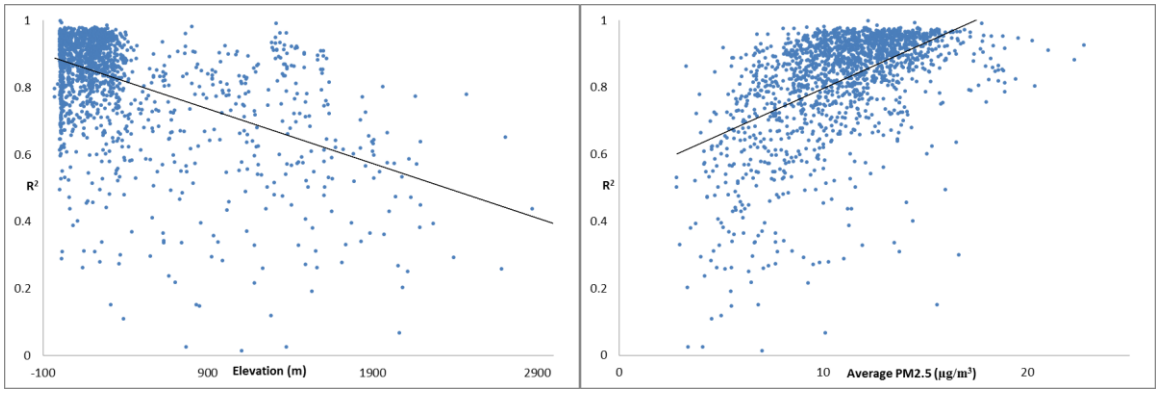


Figure S1-10. Model Performance Versus Elevation and Annual Average $PM_{2.5}$

References

1. Isakov V, Touma J, Touma J, et al. Estimating Fine Particulate Matter Component Concentrations and Size Distributions Using Satellite-Retrieved Fractional Aerosol Optical Depth: Part 2—A Case Study. Journal of the Air & Waste Management Association 2007;57:1360-9.

2. Liu Y. Mapping annual mean ground-level $PM_{2.5}$ concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. Journal of Geophysical Research 2004;109.

3. Di Q, Schwartz J. Using Chemical Transport Model to Fill in the Missingness of Satellite-Based AOD. Atmospheric Environment 2015.

4. Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. World map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 2006;15:259-63.

5. LeCun Y, Ranzato M. Deep learning tutorial. Tutorials in International Conference on Machine Learning (ICML'13); 2013: Citeseer.

6. Moré JJ. The Levenberg-Marquardt algorithm: implementation and theory. Numerical analysis: Springer; 1978:105-16.

7. Schalkoff RJ. Artificial neural networks. New York: McGraw-Hill; 1997.

8. Fletcher D, Goss E. Forecasting with neural networks. Information & Management 1993;24:159-67.

9. Kloog I, Koutrakis P, Coull BA, Lee HJ, Schwartz J. Assessing temporally and spatially resolved $PM_{2.5}$ exposures for epidemiological studies using satellite aerosol optical depth measurements. Atmospheric Environment 2011;45:6267-75.

10. van Donkelaar A, Martin RV, Park RJ. Estimating ground-level $PM_{2.5}$ using aerosol optical depth determined from satellite remote sensing. Journal of Geophysical Research 2006;111.

11. Kloog I, Chudnovsky AA, Just AC, et al. A new hybrid spatio-temporal model for estimating daily multi-year PM $_{2.5}$ concentrations across northeastern USA using high resolution aerosol optical depth data. Atmospheric Environment 2014;95:581-90.

12. Lee M, Kloog I, Chudnovsky A, et al. Spatiotemporal prediction of fine particulate matter using high-resolution satellite images in the Southeastern US 2003–2011. Journal of Exposure Science and Environmental Epidemiology 2015.

Chapter 2. Air Pollution and Mortality in the Medicare Population[1]

Qian Di[1], Yan Wang[1], Antonella Zanobetti[1], Yun Wang[2], Petros Koutrakis[1], Christine Choirat[2], Francesca Dominici[2], and Joel D. Schwartz[1]

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

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

[1] Reproduced with permission from Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F. and Schwartz, J.D., 2017. Air pollution and mortality in the Medicare population. *New England Journal of Medicine*, 376(26), pp.2513-2522. Copyright Massachusetts Medical Society.

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.)

Introduction

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.

Method

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, meteorological 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 2-1, 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 2-1, 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).

Result

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 2-1).

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 (Figure 2-1).

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 2-2). 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 2-2). Analysis of the MCBS 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 (Figure 2-2). Among black persons, the effect estimate for $PM_{2.5}$ was three times as high as that for the overall population (Table S2-3). 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 S2-8). 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 (Figure 2-3 and Figure S2-8).

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 factors, such as smoking and income, which could be important confounders. Still, our analysis of the MCBS subsample (Table S2-6) 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.

We also found that our results were robust when we excluded individual and ecologic covariates from the main analysis (Figure S2-2 and Table S2-2), when we stratified age at entry into 3-year and 4-year categories rather than the 5 years used in the main analysis (Figure S2-3), when we varied the estimation procedure (by means of a generalized estimating equation as opposed to mixed effects) (Table S2-3 and Table S2-4), 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 (Figure 2-3); 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 exceed70 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 $PM_{2.5}$ exposure 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 $PM_{2.5}$ (Figure S2-4). 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 S2-3).

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.

Table 2-1. Baseline Characteristics of the Entire Cohort

Variable	Entire Cohort	Ozone Concentration		$PM_{2.5}$ Concentration	
		≥50 ppb*	<50 ppb	≥12 μg/m^3	<12μg/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
$PM_{2.5}$ (μg/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

Table 2-1 (Continued)

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
Ecological Variables ‡					
BMI	28.2	27.9	28.4	28.0	28.4
Ever smoker (%)	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

Table 2-1 (Continued)

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
Meteorological 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 meteorological 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.

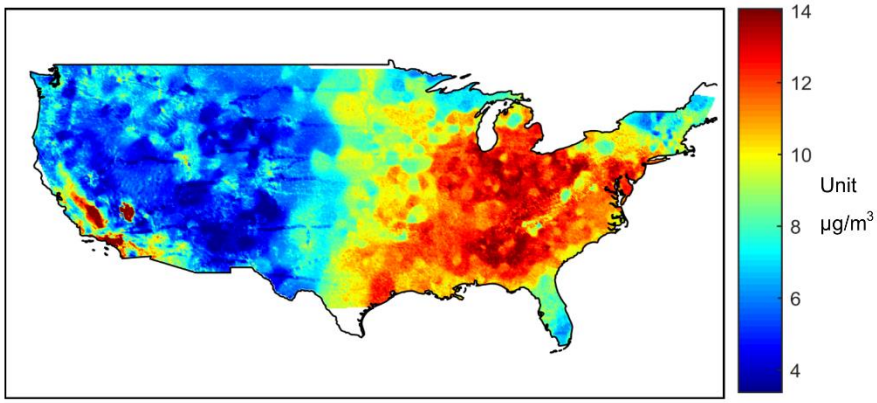
Table 2-2. Risk of Death Associated with an Increase of 10 µg per Cubic Meter in $PM_{2.5}$ or an Increase of 10 ppb in Ozone Concentration*

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)
Two-pollutant analysis at low exposure (Low-exposure analysis)	1.136 (1.131, 1.141)	1.010 (1.009, 1.011)
Single-pollutant analysis†	1.084 (1.081, 1.086)	1.023 (1.022, 1.024)
Two-pollutant analysis using the nearest monitoring site (Nearest-monitor analysis)‡	1.061 (1.059, 1.063)	1.001 (1.000, 1.002)

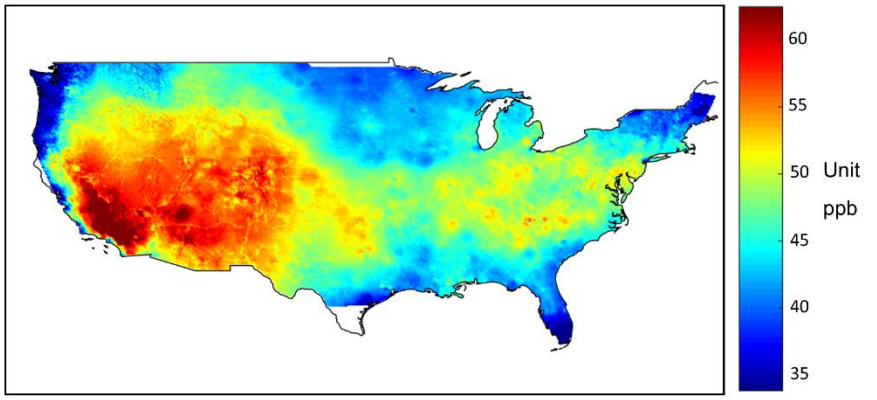
* Hazard ratios and 95% confidence intervals were calculated on the basis of an increase of 10 µg per cubic meter in exposure to $PM_{2.5}$ and an increase of 10 ppb in exposure to ozone.

† 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.



(a)



(b)

Figure 2-1. 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.

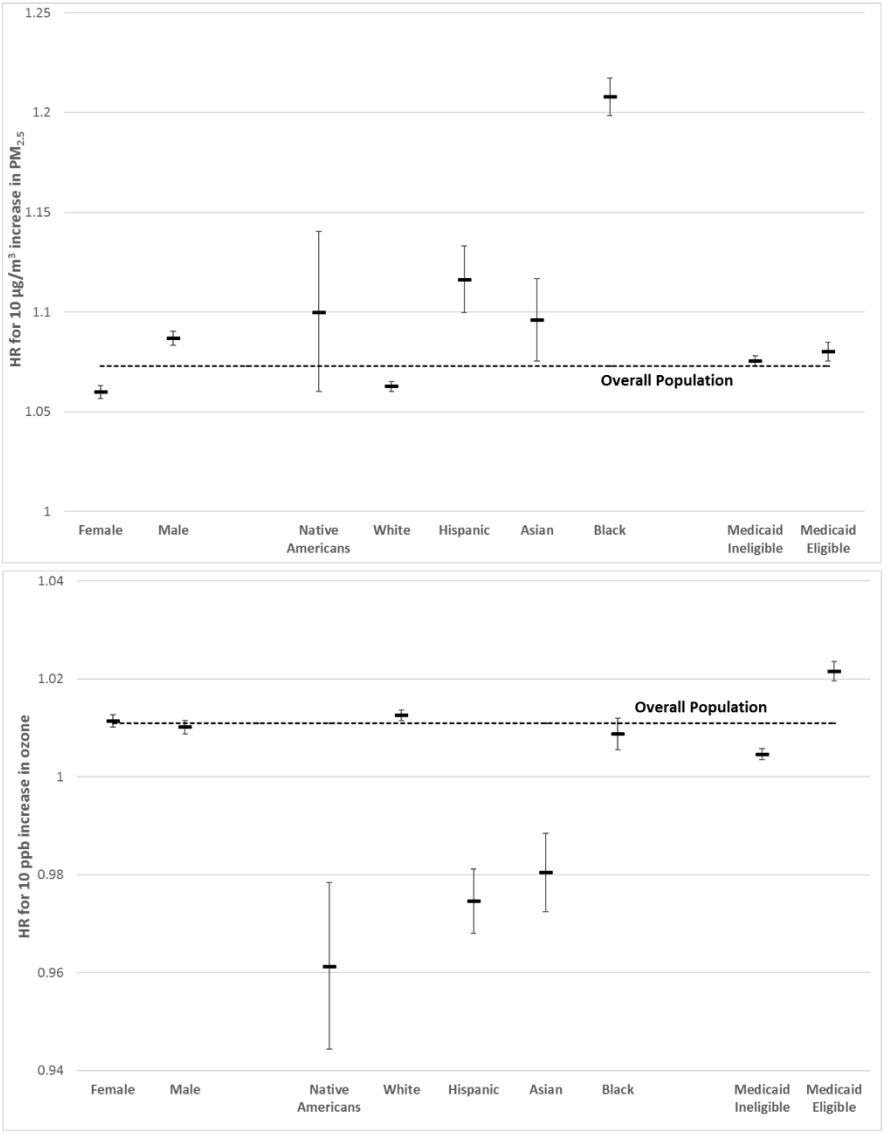


Figure 2-2. Risk of Death Associated with an Increase of 10 µg per Cubic Meter in $PM_{2.5}$ Concentrations and an Increase of 10 ppb in Ozone Exposure, According to Study Subgroups.

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 Table S2-3 and Table S2-4 in the Supplementary Appendix. Dashed lines indicate the estimated hazard ratio for the overall population.

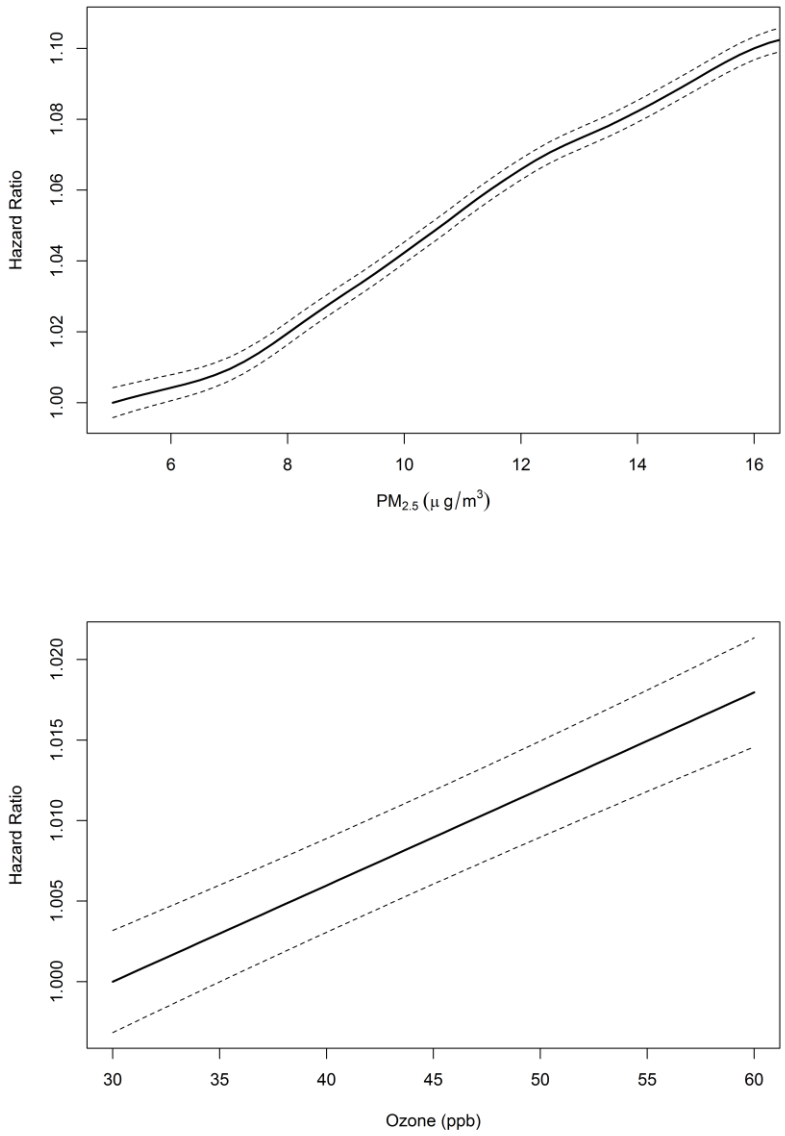


Figure 2-3. Concentration–Response Function of the Joint Effects of Exposure to $PM_{2.5}$ and Ozone on All-Cause Mortality.

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.

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 $PM_{2.5}$, O3, and $NO2_2$ 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 $PM_{2.5}$ 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 $PM_{2.5}$ 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 $PM_{2.5}$ 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 $PM_{2.5}$ 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. $PM_{2.5}$ 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 followup 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 Material

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 S2-1).

Table S2-1 summarizes the Pearson correlation coefficients between each pollutant and ecological covariate.

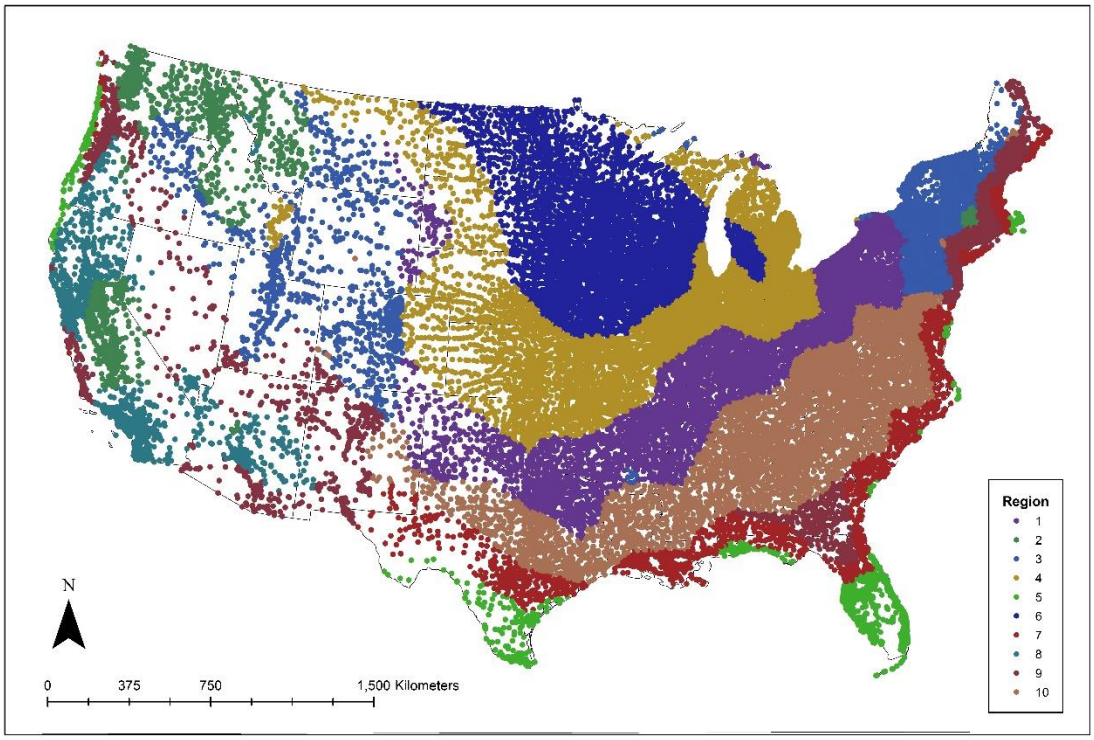


Figure S2-1. Regional Dummy Variable

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

Table S2-1. 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 S2-2 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 S2-2 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. -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.

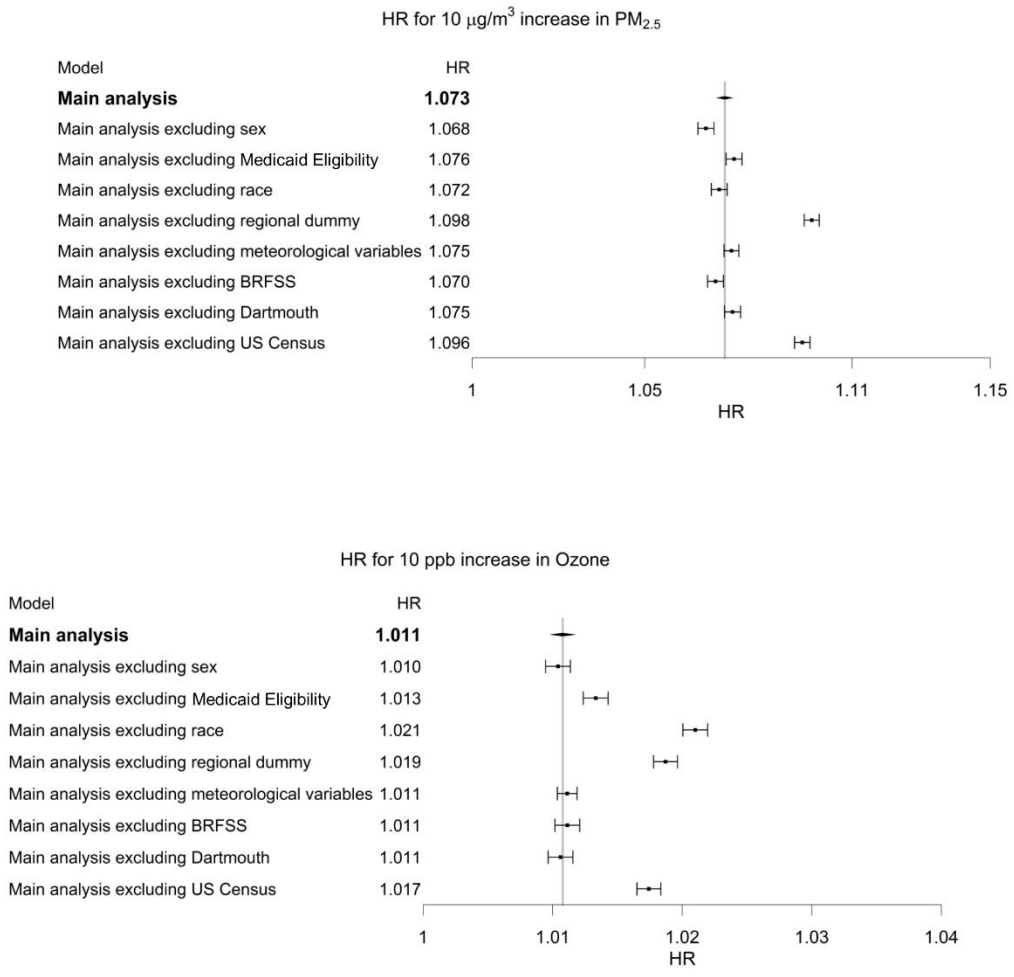


Figure S2-2. Estimated Risk of Death Associated with $PM_{2.5}$ and Ozone Exposure at Different Levels of Adjustment.

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 S2-3).

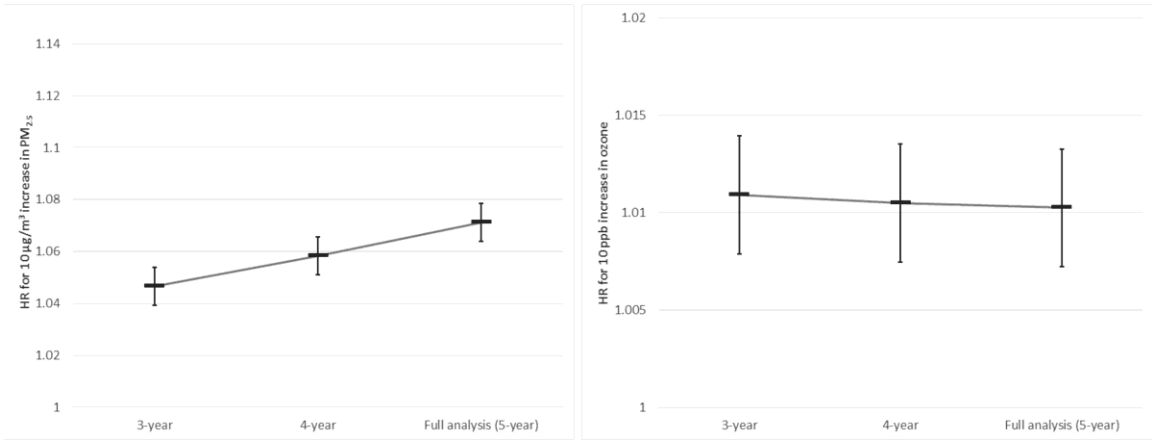


Figure S2-3. Risk of Death Associated with $PM_{2.5}$ and Ozone Exposure at Different Age Groups at Entry

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 S2-4). 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 S2-3 and Table S2-4 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.

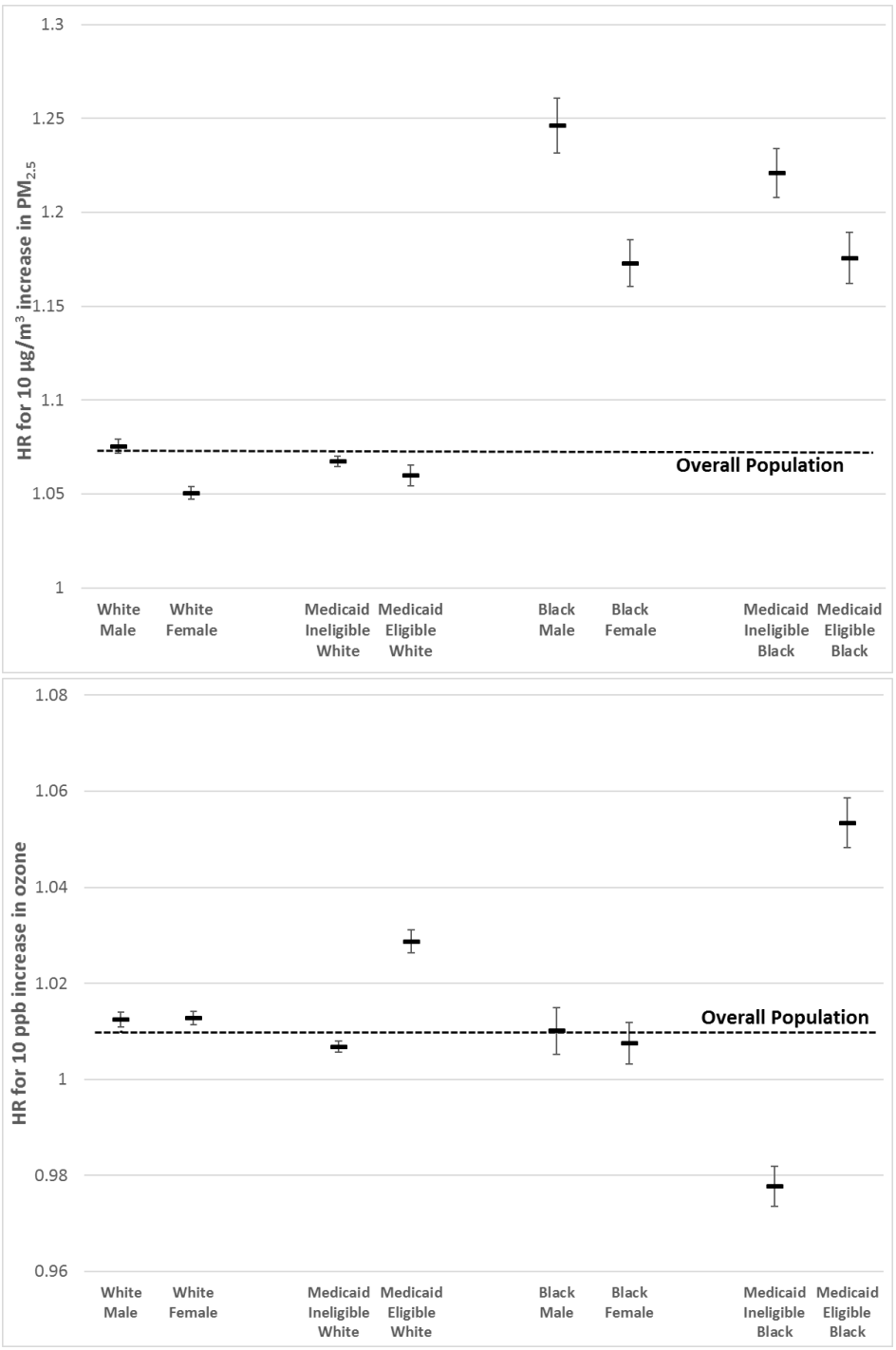


Figure S2-4. Risk of Death Associated with $PM_{2.5}$ and Ozone Exposure in Subgroups.

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 S2-3 and Table S2-4. Dashed line indicates the HR for the overall population.

Table S2-3. 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

Table S2-3 (Continued)

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			

Table S2-3 (Continued)

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

Table S2-3 (Continued)

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 S2-4. 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

Table S2-4 (Continued)

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

Table S2-4 (Continued)

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

Table S2-4 (Continued)

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 S2-3.

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 S2-5 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 S2-5 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 S2-6 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 S2-5. 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.

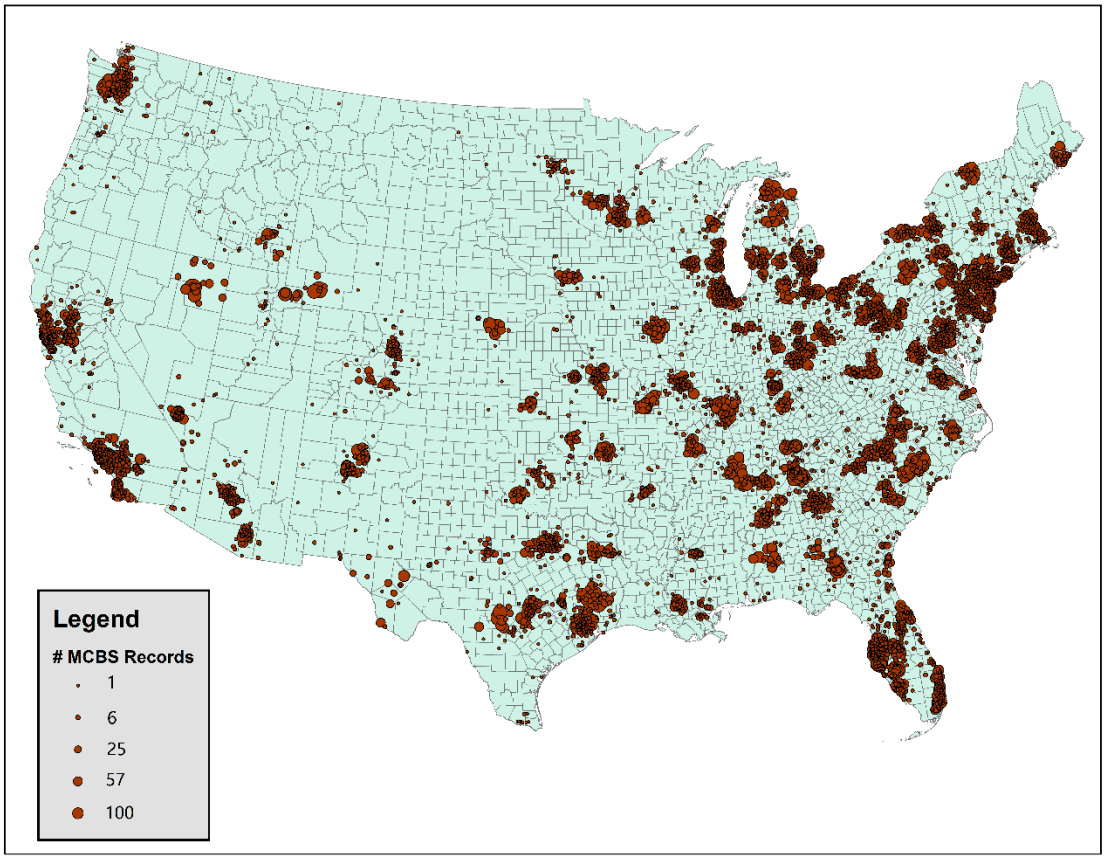


Figure S2-5. Geographic Distribution of MCBS Enrollees

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 S2-6. 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 S2-6).

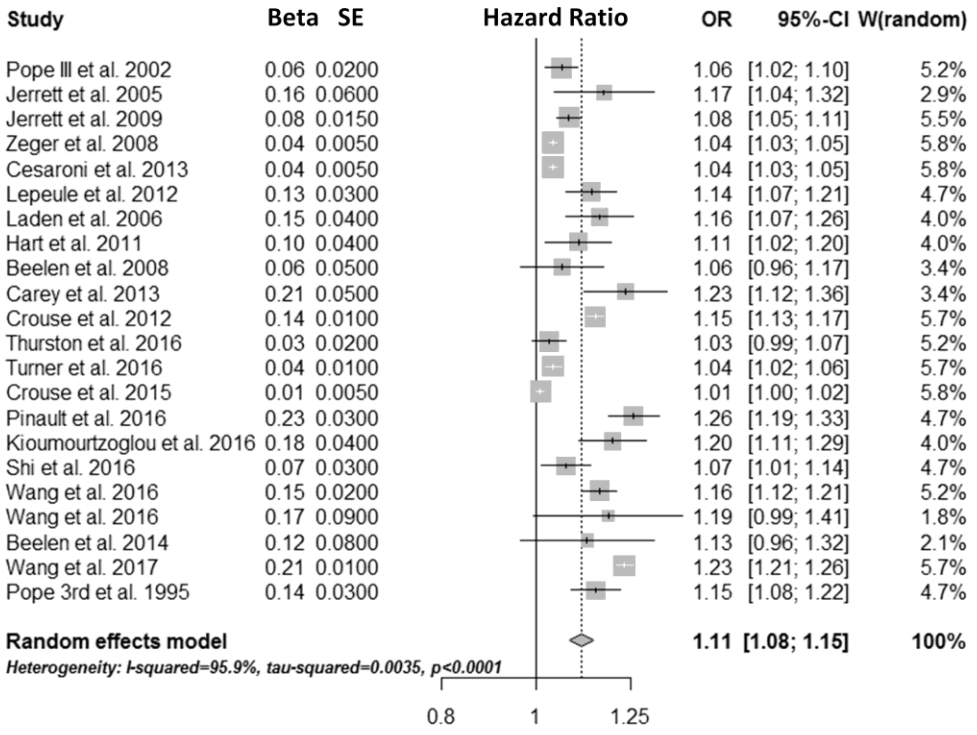


Figure S2-6. Forest Plot of Recent Studies on $PM_{2.5}$

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 S2-7 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 S2-7). This result is consistent with visual interpretation of the concentration-response curves (Figure S2-7). 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 S2-8). 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 S2-8, Figure 2-3).

Unlike the ozone concentration-response curve, the concentration-response surface displays no threshold effect (Figure S2-8, Figure 2-3). 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.

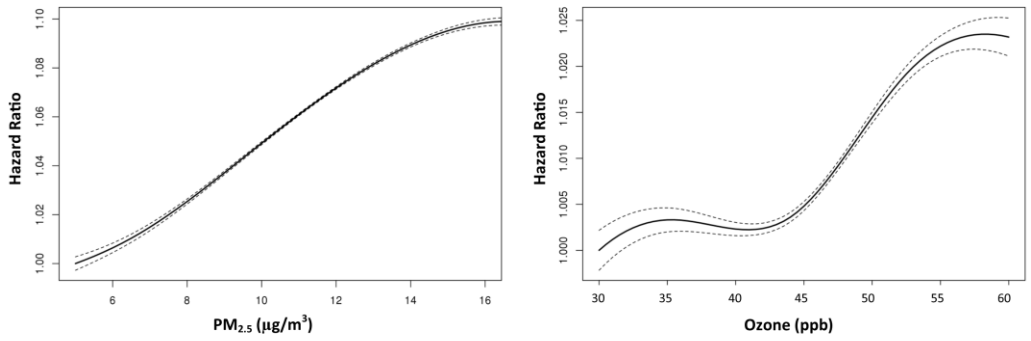


Figure S2-7. Concentration-Response Curves of $PM_{2.5}$ and Ozone on Mortality

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).

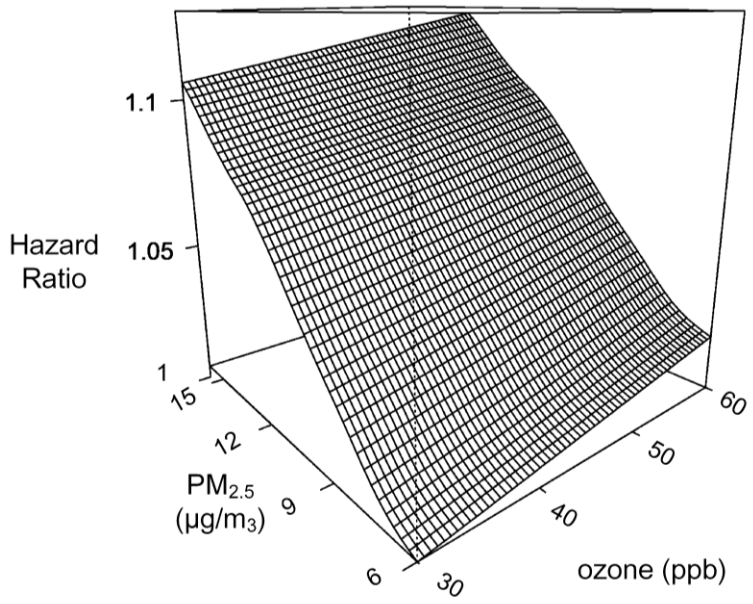


Figure S2-8. Concentration-Response Surface of $PM_{2.5}$ and Ozone on Mortality

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 S2-7. 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 S2-8. Interaction between $PM_{2.5}$ and Ozone

Variable	Beta	se	p
$PM_{2.5}$ ($\mu g/m^3$)	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 $\mu g/m^3$.

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 $PM_{2.5}$ 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.

Chapter 3. Association of Short-term Exposure to Air Pollution with Mortality in Older Adults[1]

Qian Di[1], Lingzhen Dai[1], Yun Wang[2], Antonella Zanobetti[1], Christine Choirat[2], Joel D. Schwartz[1], Francesca Dominici[2]

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

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

Author Contributions: Mr Di had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Mr Di and Dr Dai contributed equally to this study.

Concept and design: Di, Dai, Zanobetti, Schwartz, Dominici.

Acquisition, analysis, or interpretation of data: All authors. Drafting of the manuscript: Di, Dai, Choirat, Dominici.

Critical revision of the manuscript for important intellectual content: All authors. Statistical analysis: Di, Dai, Choirat, Schwartz, Dominici.

[1] Reproduced with permission from Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D. and Dominici, F., 2017. Association of short-term exposure to air pollution with mortality in older adults. *Jama*, *318*(24), pp.2446-2456. Copyright American Medical Association.

Obtained funding: Zanobetti, Schwartz, Dominici.

Administrative, technical, or material support: Wang, Choirat.

Supervision: Zanobetti, Schwartz, Dominici.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Mr Di reported receiving grants from the National Institutes of Health (NIH), Environmental Protection Agency (EPA), Health Effects Institute (HEI), and the National Cancer Institute. Dr Zanobetti reported receiving grants from the NIH, HEI, and EPA. Dr Choirat reported receiving grants from the NIH and EPA. Dr Schwartz reported receiving funding from the US Department of Justice, NIH, EPA, and HEI. Dr Schwartz is an expert consultant of the US Department of Justice regarding health impacts of Clean Air Act violations. No other disclosures were reported.

Funding/Support: This study was supported by grants R01 ES024332-01A1, ES-000002, ES024012, R01ES026217, and 4953-RFA14-3/16-4 from the NIH; grant 4953-RFA14-3/16-4 from the HEI; and grants 83587201-0 and RD-83479801 from the EPA.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The contents are solely the responsibility of the grantee and do not necessarily represent the official views of the funding agencies. Further, the funding agencies do not endorse the purchase of any commercial products or services related to this publication.

Additional Contributions: We thank Stacey C. Tobin, PhD, and Kathy L. Brenner, MAT, from Harvard T.H. Chan School of Public Health, for editorial assistance on the manuscript; Sarah L. Duncan, MDiv, and William J. Horka, BS, at the Institute for Quantitative Social Science, Harvard University, for their support with the Research Computing Environment; and Ista Zahn, MS, at the Institute for Quantitative Social Science, Harvard University, for programming support. Dr Tobin received compensation for editorial assistance.

Key Points

Question

What is the association between short-term exposure to air pollution below current air quality standards and all-cause mortality?

Finding

In a case-crossover study of more than 22 million deaths, each 10-$\mu g/m^3$ daily increase in fine particulate matter and 10-parts-per-billion daily increase in warm-season ozone exposures were associated with a statistically significant increase of 1.42 and 0.66 deaths per 1 million persons at risk per day, respectively.

Meaning

Day-to-day changes in fine particulate matter and ozone exposures were significantly associated with higher risk of all-cause mortality at levels below current air quality standards, suggesting that those standards may need to be reevaluated.

Abstract

Importance The US Environmental Protection Agency is required to reexamine its National Ambient Air Quality Standards (NAAQS) every 5 years, but evidence of mortality risk is lacking at air pollution levels below the current daily NAAQS in unmonitored areas and for sensitive subgroups.

Objective To estimate the association between short-term exposures to ambient fine particulate matter ($PM_{2.5}$) and ozone, and at levels below the current daily NAAQS, and mortality in the continental United States.

Design, Setting, and Participants Case-crossover design and conditional logistic regression to estimate the association between short-term exposures to $PM_{2.5}$ and ozone (mean of daily exposure on the same day of death and 1 day prior) and mortality in 2-pollutant models. The study included the entire Medicare population from January 1, 2000, to December 31, 2012, residing in 39 182 zip codes.

Exposures Daily $PM_{2.5}$ and ozone levels in a 1-km × 1-km grid were estimated using published and validated air pollution prediction models based on land use, chemical transport modeling, and satellite remote sensing data. From these gridded exposures, daily exposures were calculated for every zip code in the United States. Warm-season ozone was defined as ozone levels for the months April to September of each year.

Main Outcomes and Measures All-cause mortality in the entire Medicare population from 2000 to 2012.

Results During the study period, there were 22,433,862 case days and 76,143,209 control days. Of all case and control days, 93.6%had $PM_{2.5}$ levels below 25 $\mu g/m^3$, during which 95.2%of

deaths occurred (21,353,817 of 22,433,862), and 91.1% of days had ozone levels below 60 parts per billion, during which 93.4%of deaths occurred (20,955,387 of 22,433,862). The baseline daily mortality rates were 137.33 and 129.44 (per 1 million persons at risk per day) for the entire year and for the warm season, respectively. Each short-term increase of 10 $\mu g/m^3$ in $PM_{2.5}$ (adjusted by ozone) and 10 parts per billion (10^{-9}) in warm-season ozone (adjusted by $PM_{2.5}$) were statistically significantly associated with a relative increase of 1.05% (95%CI, 0.95%-1.15%) and 0.51%(95%CI, 0.41%-0.61%) in daily mortality rate, respectively. Absolute risk differences in daily mortality rate were 1.42 (95%CI, 1.29-1.56) and 0.66 (95%CI, 0.53-0.78) per 1 million persons at risk per day. There was no evidence of a threshold in the exposure-response relationship.

Conclusions and Relevance In the US Medicare population from 2000 to 2012, short-term exposures to $PM_{2.5}$ and warm-season ozone were significantly associated with increased risk of mortality. This risk occurred at levels below current national air quality standards, suggesting that these standards may need to be reevaluated.

Introduction

In the United States, the Clean Air Act[1] requires a review of National Ambient Air Quality Standards (NAAQS) for fine particulate matter ($PM_{2.5}$) and ozone every 5 years.[2] In 2012, the annual and 24-hour NAAQS for $PM_{2.5}$ were set to 12 $\mu g/m^3$ and 35 $\mu g/m^3$, respectively. With no annual standard for ozone, the 8-hour NAAQS for ozone was set to 70 parts per billion (ppb). Currently, the review of these standards is ongoing, with public comments expected in the fall of 2017.[3]

Several studies have provided evidence that short-term exposures to $PM_{2.5}$ and ozone were associated with mortality,[4-8] but these studies primarily included large and well monitored metropolitan areas. While the US Environmental Protection Agency (EPA) is considering more stringent NAAQS, evidence is needed to clarify the association between mortality risk and exposure levels below the daily NAAQS and in rural and unmonitored areas.

The Clean Air Act[1] also requires the US EPA to set standards to protect "sensitive subgroups." To estimate the health risk of short-term exposure to air pollution for specific subgroups (e.g., underrepresented minorities and those with low socioeconomic status, such as persons eligible for Medicaid), a large population is necessary to achieve maximum accuracy and adequate statistical power.

A case-crossover study was conducted to examine all deaths of Medicare participants in the continental United States from 2000 throughout 2012 and estimate the mortality risk associated with short-term exposures to $PM_{2.5}$ and ozone in the general population as well as in subgroups. The study was designed to estimate the association between daily mortality and air pollution at

levels below current daily NAAQS to evaluate the adequacy of the current air quality standards for $PM_{2.5}$ and ozone.

Methods

This study was approved by the institutional review board at the Harvard T.H. Chan School of Public Health. As a study of previously collected administrative data, it was exempt from informed consent requirements.

Study Population

Using claims data from the Centers for Medicare & Medicaid Services, all deaths among all Medicare beneficiaries were identified during the period 2000 to 2012, providing enough power to analyze the risk of mortality associated with $PM_{2.5}$ and ozone concentrations much lower than the current standards (Table 3-1). For each beneficiary, information was extracted on the date of death, age, sex, race, ethnicity, zip code of residence, and eligibility for Medicaid (a proxy for low income) to assess the associations of mortality with $PM_{2.5}$ and ozone concentrations in potentially vulnerable subgroups. Self-reported information on race and ethnicity was obtained from Medicare beneficiary files.

Outcome

The study outcome was all-cause mortality. Individuals with a verified date of death between January 1, 2000, and December 31, 2012, were included. Individuals with an unverified date of death, or still living after December 31, 2012, were excluded.

Study Design

We estimated the association between short-term exposure to $PM_{2.5}$ (adjusted by ozone) and short-term exposure to ozone (adjusted by $PM_{2.5}$) and all-cause mortality using a case-crossover design.[9] Specifically, "case day" was defined as the date of death. For the same person, we compared daily air pollution exposure on the case day vs daily air pollution exposure on "control days." Control days were chosen (1) on the same day of the week as the case day to control for potential confounding effect by day of week; (2) before and after the case day (bidirectional sampling) to control for time trend[10,11]; and (3) only in the same month as the case day to control for seasonal and subseasonal patterns.[10,12] Individual-level covariates and zip code–level covariates that did not vary day to day (e.g., age, sex, race/ethnicity, socioeconomic status, smoking, and other behavioral risk factors) were not considered to be confounders as they remain constant when comparing case days vs control days.

Environmental Data

Daily ambient levels of $PM_{2.5}$ and ozone were estimated from published and validated air pollution prediction models.[13,14] Combining monitoring data from the EPA, satellite-based measurements, and other data sets, neural networks were used to predict 24-hour $PM_{2.5}$ and 8-hour maximum ozone concentrations at each 1-km ×1-km grid in the continental United States, including locations with no monitoring sites. Cross-validation indicated good agreement between predicted values and monitoring values ($R^2 = 0.84$ for $PM_{2.5}$ and $R^2 = 0.76$ for ozone) and at low concentrations ($R^2 = 0.85$ when constraining to 24-hour $PM_{2.5}<25$ $\mu g/m^3$ and $R^2 = 0.75$ when constraining to daily 8-hour maximum ozone <60 ppb). Details have been published elsewhere.[13,14] Warm season was defined to be from April 1 to September 30, which is the specific time window to examine the association between ozone and mortality. Meteorological

variables, including air and dew point temperatures, were retrieved from North American Regional Reanalysis data and estimated daily mean values were determined for each 32-km × 32-km grid in the continental United States.[15]

For each case day (date of death) and its control days, the daily 24-hour $PM_{2.5}$, 8-hour maximum ozone, and daily air and dew point temperatures were assigned based on zip code of residence of the individual (Appendix 1 in the Supplement). Because we estimated air pollution levels everywhere in the continental United States, the number of zip codes included in this study was 39,182, resulting in a 33% increase compared with the number of zip codes with a centroid less than 50 km from a monitor (n = 26,115).

Statistical Analysis

The relative risk (RR) of all-cause mortality associated with short-term exposures to $PM_{2.5}$ (adjusted by ozone) and warm-season ozone (adjusted by $PM_{2.5}$) was estimated by fitting a conditional logistic regression to all pairs of case days and matched control days (Appendix 2 in the Supplement).9 The regression model included both pollutants as main effects and natural splines of air and dew point temperatures with 3 df to control for potential residual confounding by weather. For each case day, daily exposure to air pollution was defined as the mean of the same day of death (lag 0-day) and 1 day prior (lag 1-day), denoted as lag 01-day.[5,16,17] Relative risk increase (RRI) was defined as RR − 1. The absolute risk difference (ARD) of all-cause mortality associated with air pollution was defined as ARD = $\alpha \times (RR - 1)/RR$, where α denotes the baseline daily mortality rate (Appendix 3 in the Supplement).

The robustness of the analysis results was assessed with respect to (1) choosing the df used for the confounding adjustment for temperature, (2) using lag 01-day exposure as the exposure

metric, (3) the definition of warm season, and (4) using only air pollution measurements from the nearest EPA monitoring sites. Splines on meteorological variables with 6 and 9 df yielded results with a difference of less than 5% of the standard error (Figure S3-1). The main analysis, which used the lag 01-day exposure, yielded the lowest values of the Akaike Information Criteria values, indicating better fit to the data (Table S3-1). Different definitions of warm season yielded similar risk estimates (Appendix 4 in the Supplement), and using exposure measurements from the nearest monitors resulted in attenuated, but still significant, risk estimates (Table 3-2).

The subgroup analyses were conducted by sex (male and female), race/ethnicity (white, nonwhite, and others), age (≤69, 70-74, 75-84, and ≥85 years), eligibility for Medicaid, and population density (quartiles). We fitted separate conditional logistic regressions to the data for each subgroup and obtained subgroup-specific estimates of RR and ARD. We implemented a 2-sample test for assessing statistically significant differences in the estimated RR and ARD between categories within each subgroup (e.g., female vs male), based on the point estimate and standard error (se) (Appendix 5 in the Supplement):

$$Z = \frac{RR_{male} - RR_{famle}}{\sqrt{se(RR_{male})^2 + se\left(RR_{female}\right)^2}}$$

The goal was to estimate mortality rate increases (both RRI and ARD) at air pollution levels well below the current daily NAAQS. The analysis was restricted to days with daily air pollution concentrations below25 μg/m^3 for $PM_{2.5}$ and 60 ppb for ozone. We chose 25 μg/m^3 and 60 ppb instead of the current daily NAAQS (35 μg/m^3 for daily $PM_{2.5}$ and 70 ppb for 8-hour maximum ozone) because levels of $PM_{2.5}$ and ozone on most of the days included in the analysis were already below the current safety standards.

Exposure-response curves were estimated between$PM_{2.5}$ or ozone and mortality by replacing linear terms for the 2 pollutants with penalized splines for both $PM_{2.5}$ and ozone.

All analyses were performed in R software version 3.3.2 (R Foundation). Computations were run on (1) the Odyssey cluster supported by the Faculty of Arts and Sciences Division of Science, Research Computing Group at Harvard University and (2) the Research Computing Environment supported by the Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University.

Results

During the study period, there were more than 22 million case days (deaths) and more than 76 million control days (Table 3-1). Of all case and control days, 93.6% had $PM_{2.5}$ levels below 25 $\mu g/m^3$, during which 95.2% of deaths occurred (21,353,817 of 22,433,862), and 91.1% of days had ozone levels below 60 ppb, during which 93.4% of deaths occurred (20,955,387 of 22,433,862). The baseline daily mortality rates were 137.33 and 129.44 (per 1 million persons at risk per day [per 1M per day]) for the entire year and for the warm season, respectively. The mean time between case and control days was 12.55 days (range 7-28 days), with minimal differences in air and dew point temperatures between case and control days (0.003°C and 0.01°C, respectively). During the study period, the mean concentrations of $PM_{2.5}$ and ozone were 11.6 $\mu g/m^3$ and 37.8 ppb, respectively. Figure 3-1 and Figure 3-2 show the daily $PM_{2.5}$ and ozone time series by state, respectively.

Each 10-$\mu g/m^3$ and 10-ppb increase in the lag 01-day exposure for $PM_{2.5}$ and warm-season ozone was associated with an RRI of 1.05% (95% CI, 0.95%-1.15%) and 0.51% (95% CI, 0.41%-0.61%) in the daily mortality rate. The ARDs were 1.42 (95% CI, 1.29-1.56) and 0.66 (95% CI,

0.53-0.78) per 1M per day. These associations remained significant when examining days below 25 μg/m^3 for $PM_{2.5}$ and below 60 ppb for ozone, with larger effect size estimates for both $PM_{2.5}$ and ozone (RRI: 1.61% [95%CI, 1.48%-1.74%] and 0.58% [95%CI,0.46%-0.70%]; ARD: 2.17 [95% CI, 2.00-2.34] and 0.74 [95% CI, 0.59-0.90] per 1M per day, respectively) (Table 3-2). $PM_{2.5}$ was associated with higher mortality rate in some subgroups, including Medicaid eligible individuals (RRI: 1.49% [95%CI, 1.29%-1.70%]; ARD: 3.59 [95%CI, 3.11-4.08] per 1M per day; interaction: P < .001), individuals older than 70 years (e.g., for ≥85 years, RRI: 1.38% [95% CI, 1.23%-1.54%]; ARD: 5.35 [95% CI, 4.75-5.95] per 1M per day; interaction: P < .001), and females (RRI: 1.20% [95% CI, 1.07%-1.33%]; ARD: 1.56 [95%CI, 1.39-1.72] per1Mper day; interaction: P = .02) (Figure 3-3 and Figure 3-4). The effect estimates for $PM_{2.5}$ increased with age. The effect estimate for black individuals was higher than that for white individuals (P = .001; Figure S3-2 in the Supplement). For ozone, similar patterns were observed, but with less contrast between groups. No significant differences were found in the short-term associations between air pollution exposure ($PM_{2.5}$ and ozone) and mortality across areas with different population density levels (Figure 3-3 and Figure 3-4). Effect estimates using different lags of exposure are shown in Figure S3-3 in the Supplement.

Figure 3-5 shows the estimated exposure-response curves for $PM_{2.5}$ and ozone. The slope was steeper at $PM_{2.5}$ levels below 25 μg/m^3 (P < .001), consistent with the low-exposure analysis (Table 3-2). Both $PM_{2.5}$ and ozone exposure-responses were almost linear, with no indication of a mortality risk threshold at very low concentrations. Figure S3-4 in the Supplement shows the exposure-response curves for $PM_{2.5}$ when restricted to just the warm season and for ozone when not restricted to the warm season; results were similar.

Discussion

In this large case-crossover study of all Medicare deaths in the continental United States from 2000 to 2012, a 10-μg/m^3 daily increase in $PM_{2.5}$ and a 10-ppb daily increase in warm-season ozone exposures were associated with a statistically significant increase of 1.42 and 0.66 deaths per 1M per day, respectively. The risk of mortality remained statistically significant when restricting the analysis to days with $PM_{2.5}$ and ozone levels much lower than the current daily NAAQS.[18] This study included individuals living in smaller cities, towns, and rural areas that were unmonitored and thus excluded from previous time series studies. There were no significant differences in the mortality risk associated with air pollution among individuals living in urban vs rural areas. Taken together, these results provide evidence that short-term exposures to $PM_{2.5}$ and ozone, even at levels much lower than the current daily standards, are associated with increased mortality, particularly for susceptible populations.

The Clean Air Act[1] requires the administrator of the US EPA to set NAAQS at levels that provide "protection for at-risk populations, with an adequate margin of safety."[19] In this study, Medicaid-eligible individuals, females, and elderly individuals had higher mortality rate increases associated with $PM_{2.5}$ than other groups. Previous studies have found similar results in some subgroups.[20,21] Poverty, unhealthy lifestyle, poor access to health care, and other factors may make some subgroups more vulnerable to air pollution. The exact mechanism is worth exploring in future studies.

The current NAAQS for daily $PM_{2.5}$ is 35 μg/m^3. When restricting the analysis to daily $PM_{2.5}$ levels below 25 μg/m^3, the association between short-term $PM_{2.5}$ exposure and mortality remained but was elevated. The current daily NAAQS for ozone is 70 ppb; when restricting the

analysis to daily warm-season ozone concentrations below 60 ppb, the effect size also increased slightly. The exposure-response curves revealed a similar pattern. These results indicate that air pollution is associated with an increase in daily mortality rates, even at levels well below the current standards.

The exposure-response relationship between $PM_{2.5}$ exposure and mortality was consistent with findings of previous studies. One study combined exposure-response curves from 22 European cities and reported an almost linear relationship between $PM_{2.5}$ and mortality.[22] Another multicity study reported a linear relationship down to 2-$\mu g/m^3$ $PM_{2.5}$.[23] The present study found a similarly linear exposure-response relationship below 15-$\mu g/m^3$ $PM_{2.5}$ and a less steep slope above this level.

For ozone, the linear exposure-response curve with no threshold described in this study is consistent with earlier research. An almost linear exposure-response curve for ozone was previously reported with no threshold or a threshold at very low concentrations.[24] A study from the Netherlands also concluded that if an ozone threshold exists, it does so at very low levels.[25]

Findings from this study are also consistent with the literature regarding the observed effect sizes of both $PM_{2.5}$ [5,8,16,26-28] and ozone.[7,20,29,30] This study further demonstrates that in more recent years, during which air pollution concentrations have fallen, statistically significant associations between mortality and exposures to $PM_{2.5}$ and ozone persisted.

The association of mortality and $PM_{2.5}$ exposure is supported by a large number of published experimental studies in animals[31-33] and in humans exposed to traffic air pollution,[34,35] diesel particles,[36] and unfiltered urban air.[37] Similarly, a review of toxicological studies and a recent

panel study found that ozone exposure was associated with multiple adverse health outcomes.[38,39]

Strengths

This study has several strengths. First, to our knowledge, this is the largest analysis of daily air pollution exposure and mortality to date, with approximately 4 times the number of deaths included in a previous large study.[5] Second, this study assessed daily exposures using air pollution prediction models that provide accurate estimates of daily levels of $PM_{2.5}$ and ozone for most of the United States, including previously unmonitored areas. An analysis that relied only on exposure data from monitoring stations was found to result in a downward bias in estimates (Table 3-2). Third, the inclusion of more than 22 million deaths from 2000 to 2012 from the entire Medicare population provided large statistical power to detect differences in mortality rates in potentially vulnerable populations and to estimate mortality rates at very low $PM_{2.5}$ and ozone concentrations. Fourth, this study estimated the air pollution–mortality association well below the current daily NAAQS and in unmonitored areas, and it did not identify significant differences in the mortality rate increase between urban and rural areas. Fifth, this study used a case-crossover design that individually matched potential confounding factors by month, year, and other time-invariant variables and controlled for time-varying patterns, as demonstrated by the minimal differences in meteorological variables between case and control days.

Limitations

This study also has several limitations. First, the case-crossover design does not allow estimation of mortality rate increase associated with long-term exposure to air pollution. Long-term risks in the same study population have been estimated elsewhere.[40] Second, because this study used residential zip code to ascertain exposure level rather than exact home address or place of death,

some measurement error is expected. Third, the Medicare population primarily consists of individuals older than 65 years, which limits the generalizability of findings to younger populations. However, because more than two-thirds of deaths in the United States occur in people older than 65 years of age, and air pollution–related health risk rises with age, the Medicare population in this study includes most cases of air pollution–induced mortality. Fourth, Medicare files do not report cause-specific mortality. Fifth, the most recent data used in this study are nearly 5 years old, and it is uncertain whether exposures and outcomes would be the same with more current data.

Conclusions

In the US Medicare population from 2000 to 2012, short-term exposures to$PM_{2.5}$ and warm-season ozone were significantly associated with increased risk of mortality. This risk occurred at levels below current national air quality standards, suggesting that these standards may need to be reevaluated.

Funding/Support: This study was supported by grants R01 ES024332-01A1, ES-000002, ES024012, R01ES026217, and 4953-RFA14-3/16-4 from the NIH; grant 4953-RFA14-3/16-4 from the HEI; and grants 83587201-0 and RD-83479801 from the EPA.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The contents are solely the responsibility of the grantee and do not necessarily represent the official views of the funding agencies. Further, the funding agencies do not endorse the purchase of any commercial products or services related to this publication.

Table 3-1. Baseline Characteristics of Study Population (2000-2012)

Baseline Characteristics	
Case days (No.)	22,433,862
Control days (No.)	76,143,209
Among All Cases (n = 22 433 862), %	
Age at death	
≤69 years	10.38%
70 to 74 years	13.37%
75 to 84 years	38.48%
≥85 years	37.78%
Sex	
Male	44.73%
Female	55.27%
Race/ethnicity	
White	87.34%
Black	8.87%
Asian	1.03%
Hispanic	1.51%

Table 3-1 (Continued)

Native American	0.31%
Medicaid eligibility (n = 22 433 862), %	
Ineligible	77.36%
Eligible	22.64%

Table 3-2. Relative Risk and Absolute Risk Difference of Daily Mortality Associated with Each 10-μg/m^3 Increase in $PM_{2.5}$ and Each 10-ppb Increase in Ozone

	Relative Risk Increase, % (95% CI)		Absolute Risk Difference in Daily Mortality Rates, No. Per 1 Million Persons at Risk Per Day (95% CI)[a]	
Air Pollutant Analysis	**$PM_{2.5}$**	**Ozone[b]**	**$PM_{2.5}$**	**Ozone[b]**
Main Analysis[c]	1.05% (0.95%, 1.15%)	0.51% (0.41%, 0.61%)	1.42 (1.29, 1.56)	0.66 (0.53, 0.78)
Low-exposure Analysis[d]	1.61% (1.48%, 1.74%)	0.58% (0.46%, 0.70%)	2.17 (2.00, 2.34)	0.74 (0.59, 0.90)
Single-pollutant Analysis[e]	1.18% (1.09%, 1.28%)	0.55% (0.48%, 0.62%)	1.61 (1.48, 1.73)	0.71 (0.62, 0.79)
Nearest Monitors Analysis[f]	0.83% (0.73%, 0.93%)	0.35% (0.28%, 0.41%)	1.13 (0.99, 1.26)	0.45 (0.37, 0.53)

Abbreviations: $PM_{2.5}$, fine particulate matter; ppb, parts per billion.

[a] The daily baseline mortality rate was 137.33 per 1 million persons at risk per day; the warm-season daily baseline mortality rate was 129.44 per 1 million persons at risk per day.

[b] Ozone analyses included days from the warm season only (April 1 to September 30).

[c] The main analysis used the mean of daily exposure on the same day of death and 1 day prior (lag 01-day) as the exposure metric for both $PM_{2.5}$ and ozone, and controlled for natural splines

of air and dew point temperatures with 3 df. The main analysis considered the 2 pollutants jointly included into the regression model and estimated the percentage increase in the daily mortality rate associated with a 10-$\mu g/m^3$ increase in $PM_{2.5}$ exposure adjusted for ozone and the percentage increase in daily mortality rate associated with a 10-ppb increase in warm-season ozone exposure adjusted for $PM_{2.5}$.

[d] The low-exposure analysis had the same model specifications as the 2-pollutant analysis and was constrained for days when $PM_{2.5}$ was below 25 $\mu g/m^3$ or ozone below 60 ppb.

[e] The single-pollutant analysis estimated the percentage increase in the daily mortality rate associated with a 10-$\mu g/m^3$ increase in $PM_{2.5}$ exposure without adjusting for ozone and the percentage increase in the daily mortality rate associated with a 10-ppb increase in ozone exposure without adjusting for $PM_{2.5}$.

[f] $PM_{2.5}$ and ozone monitoring data were retrieved from the US Environmental Protection Agency Air Quality System, which provides the daily mean of $PM_{2.5}$ and daily 8-hour maximum ozone levels at each monitoring site. Daily ozone concentrations were averaged from April 1 to September 30. Individuals were assigned to the $PM_{2.5}$ and ozone levels from the nearest monitor site within 50 km. Those living 50 km from any monitoring site were excluded.

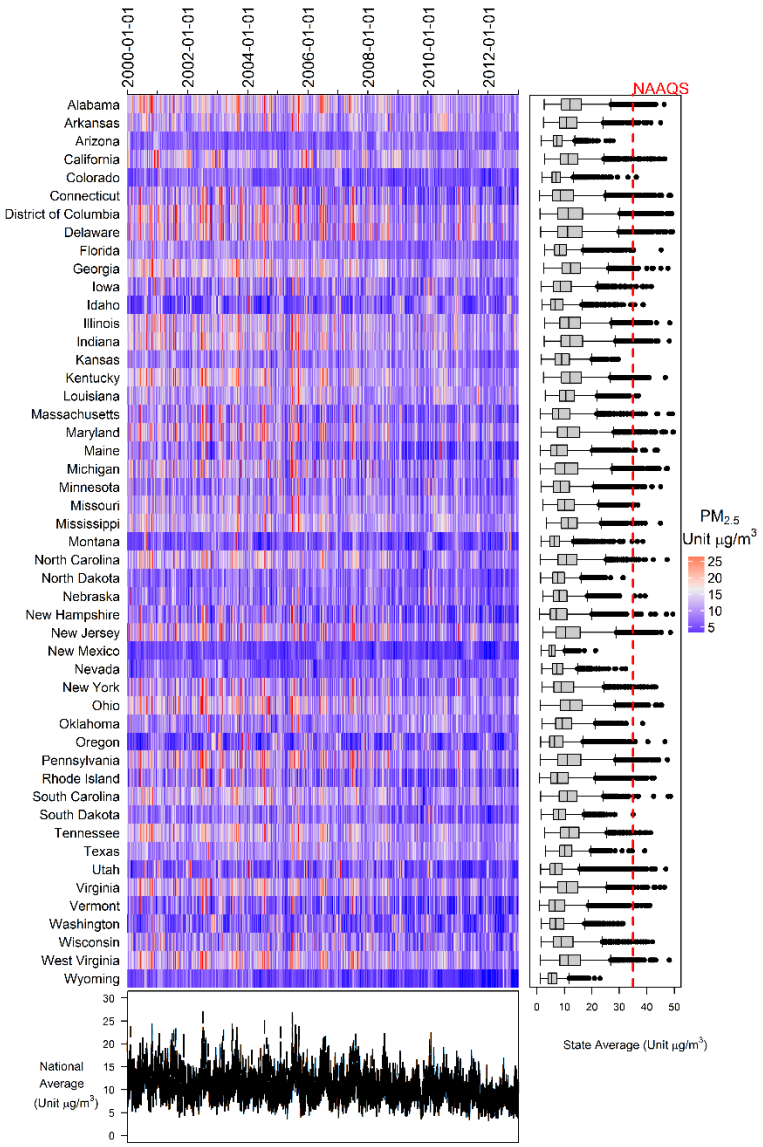


Figure 3-1. Daily Mean $PM_{2.5}$ Concentrations in the Continental United States, 2000-2012

Daily mean fine particulate matter ($PM_{2.5}$) concentrations were calculated and plotted by state. The time-series plot at the bottom indicates the national daily mean values across all locations. Boxplots show the distribution of daily $PM_{2.5}$ levels for each state. The blue dashed line indicates the daily National Ambient Air Quality Standards (NAAQS) for $PM_{2.5}$ (35 µg/m^3). The line across the box, upper hinge, and lower hinge represent the median value, 75th percentile (Q3),

and 25th percentile (Q1), respectively. The upper whisker is located at the smaller of the maximal value and Q3 + 1.5 × interquartile range; the lower whisker is located at the larger of the minimal value and Q1 – 1.5 × interquartile range. Any values that lie beyond the upper and lower whiskers are outliers.

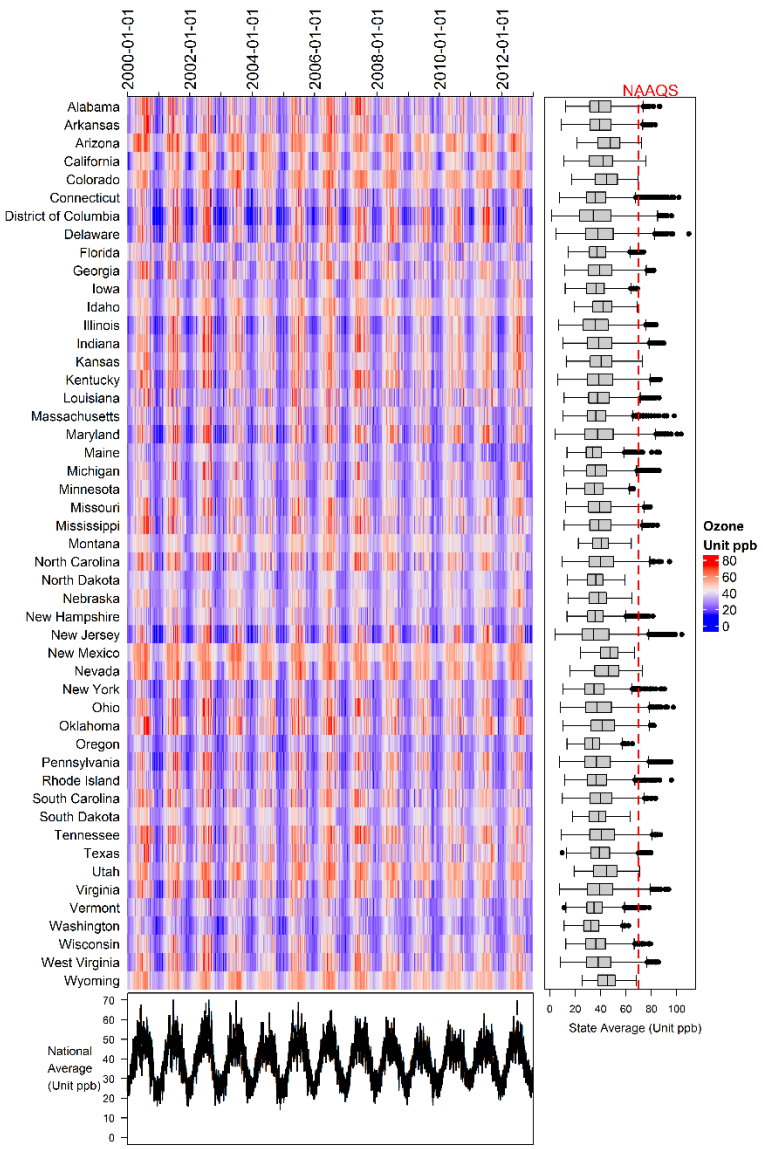


Figure 3-2. Daily 8-Hour Maximum Ozone Concentrations in the Continental United States, 2000-2012

Daily mean 8-hour maximum ozone concentrations were calculated and plotted by state. The time-series plot at the bottom indicates the national daily mean values across all locations. Boxplots show the distribution of daily ozone levels for each state. The blue dashed line indicates the daily National Ambient Air Quality Standards (NAAQS) for ozone (70 parts per

billion [ppb]). The line across the box, upper hinge, and lower hinge represent the median value, 75th percentile (Q3), and 25th percentile (Q1), respectively. The upper whisker is located at the smaller of the maximal value and Q3 + 1.5 × interquartile range; the lower whisker is located at the larger of the minimal value and Q1 – 1.5 × interquartile range. Any values that lie beyond the upper and lower whiskers are outliers.

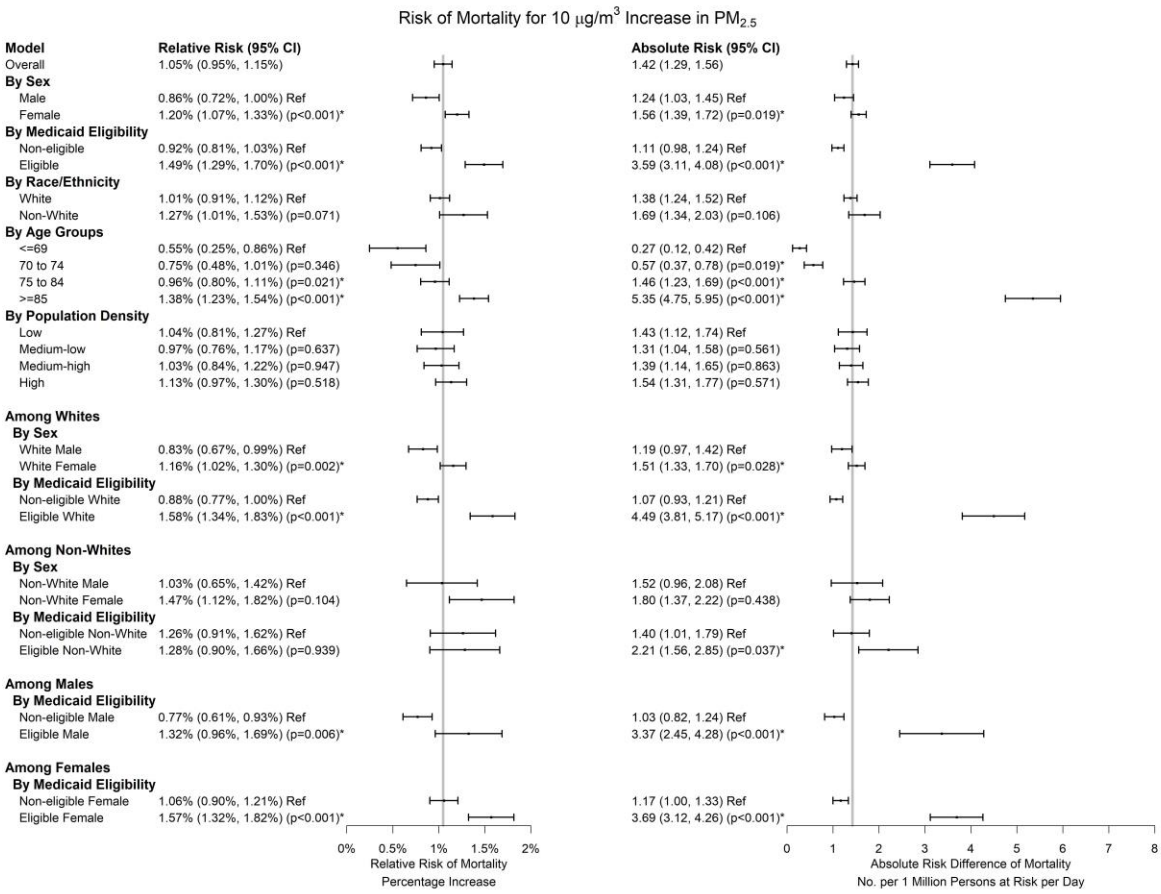


Figure 3-3. Relative Risk Increase and Absolute Risk Difference of Daily Mortality Associated With 10-μg/m^3 Increase in Fine Particulate Matter ($PM_{2.5}$)

For the main analysis, subgroup analyses used a 2-pollutant analysis (with both $PM_{2.5}$ and ozone), based on the mean of daily exposure on the same day of death and 1 day prior (lag 01-day) as the exposure metric for $PM_{2.5}$, and controlled for natural splines of air and dew point temperatures (each with 3 df). Vertical lines indicate effects for the entire study population. Subgroup analyses were conducted for each subgroup (e.g., male or female, white or nonwhite, Medicare eligible or Medicare ineligible, age groups, and quartiles of population density). For the main analysis and each subgroup, conditional logistic regressions were run to obtain relative

risk increases and calculated absolute risk difference based on baseline mortality rates (Appendix 2 in the Supplement). Numbers in the figure represent point estimates, 95% CIs, and P values for effect modifications. The reference groups were used when assessing effect modification.

[a] Statistically significant effect estimate (at 5% level) compared with the reference group.

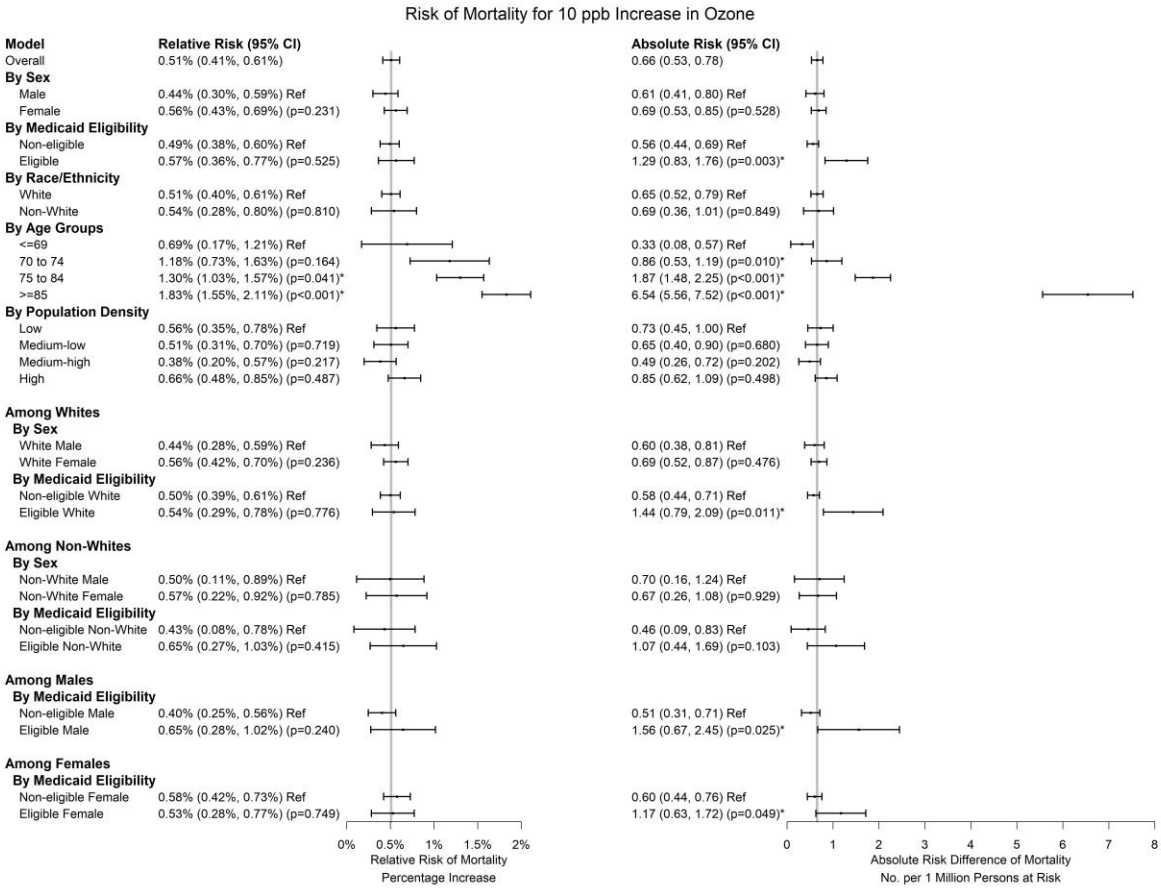


Figure 3-4. Relative Risk Increase and Absolute Risk Difference of Daily Mortality Associated With 10-Parts-per-Billion (ppb) Increase in Ozone

For the main analysis, subgroup analyses used a 2-pollutant analysis (with both $PM_{2.5}$ and ozone), based on the mean of daily exposure on the same day of death and 1 day prior (lag 01-day) as the exposure metric for ozone, and controlled for natural splines of air and dew point temperatures (each with 3 df). Vertical lines indicate effects for the entire study population. Subgroup analyses were conducted for each subgroup (e.g., male or female, white or nonwhite, Medicare eligible or Medicare ineligible, age groups, and quartiles of population density). For the main analysis and each subgroup, conditional logistic regressions were run to obtain relative

risk increases, and calculated absolute risk difference based on baseline mortality rates (Appendix 2 in the Supplement). For ozone, analyses were restricted to the warm season (April to September). Numbers in the figure represent point estimates, 95% CIs, and P values for effect modifications. The reference groups were used when assessing effect modification.

[a] Statistically significant effect estimate (at 5%level) compared with the reference group.

A Exposure-response curve for $PM_{2.5}$ **B** Exposure-response curve for ozone

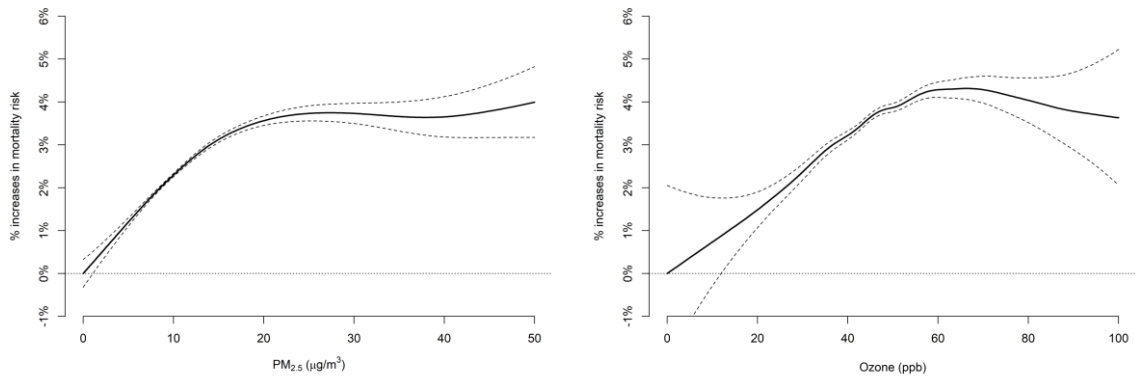


Figure 3-5. Estimated Exposure-Response Curves for Short-term Exposures to Fine Particulate Matter ($PM_{2.5}$) and Ozone

A 2-pollutant analysis with separate penalized splines on $PM_{2.5}$ (A) and ozone (B) was conducted to assess the percentage increase in daily mortality at various pollution levels. Dashed lines indicate 95% CIs. The mean of daily exposure on the same day of death and 1 day prior (lag 01-day) was used as metrics of exposure to $PM_{2.5}$ and ozone. Analysis for ozone was restricted to the warm season (April to September). Ppb indicates parts per billion.

References

1. Clean Air Act. 42 USC §7401 et seq (1970).

2. US Environmental Protection Agency. Process of reviewing the National Ambient Air Quality Standards. https://www.epa.gov/criteria-air-pollutants/process-reviewing-national-ambient-air-quality-standards. Accessed November 1, 2017.

3. US Environmental Protection Agency. Integrated review plan for the National Ambient Air Quality Standards for Particulate Matter. https://www3.epa.gov/ttn/naaqs/standards/pm/data/201612-final-integrated-review-plan.pdf Published December 2016. Accessed May 30, 2017.

4. Krall JR, Anderson GB, Dominici F, Bell ML, Peng RD. Short-term exposure to particulate matter constituents and mortality in a national study of US urban communities. Environ Health Perspect. 2013;121(10):1148-1153.

5. Zanobetti A, Schwartz J. The effect of fine and coarse particulate air pollution on mortality: a national analysis. Environ Health Perspect. 2009; 117(6):898-903.

6. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. JAMA. 2006;295(10):1127-1134.

7. Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F. Ozone and short-term mortality in 95 US urban communities, 1987-2000. JAMA. 2004; 292(19):2372-2378.

8. Schwartz J, Dockery DW, Neas LM. Is daily mortality associated specifically with fine particles? J Air Waste Manag Assoc. 1996;46(10):927-939.

9. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. Am J Epidemiol. 1991;133(2):144-153.

10. Bateson TF, Schwartz J. Control for seasonal variation and time trend in case-crossover studies of acute effects of environmental exposures. Epidemiology. 1999;10(5):539-544.

11. Maclure M, Mittleman MA. Should we use a case-crossover design? Annu Rev Public Health. 2000;21:193-221.

12. Levy D, Lumley T, Sheppard L, Kaufman J, Checkoway H. Referent selection in case-crossover analyses of acute health effects of air pollution. Epidemiology. 2001;12(2):186-192.

13. Di Q, Kloog I, Koutrakis P, Lyapustin A,Wang Y, Schwartz J. Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. Environ Sci Technol. 2016;50(9): 4712-4721.

14. 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(1):39-52.

15. Kalnay E, Kanamitsu M, Kistler R, et al. The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc. 1996;77(3):437-471.

16. Dai L, Zanobetti A, Koutrakis P, Schwartz JD. Associations of fine particulate matter species with mortality in the United States: a multicity time-series analysis. Environ Health Perspect. 2014; 122(8):837-842.

17. Ostro B, Broadwin R, Green S, Feng WY, Lipsett M. Fine particulate air pollution and mortality in nine California counties: results from CALFINE. Environ Health Perspect. 2006;114(1):29-33.

18. US Environmental Protection Agency. NAAQS table. https://www.epa.gov/criteria-air-pollutants/naaqs-table. Accessed May 30, 2017.

19. US Environmental Protection Agency. Criteria air pollutants. https://www.epa.gov/sites/production/files/2015-10/documents/ace3_criteria_air_pollutants.pdf. Updated October 2015. Accessed May 30, 2017.

20. Zanobetti A, Schwartz J. Is there adaptation in the ozone mortality relationship: a multi-city case-crossover analysis. Environ Health. 2008;7:22.

21. Baccini M, Mattei A, Mealli F, Bertazzi PA, Carugno M. Assessing the short-term impact of air pollution on mortality: a matching approach. Environ Health. 2017;16(1):7.

22. Samoli E, Analitis A, Touloumi G, et al. Estimating the exposure-response relationships between particulate matter and mortality within the APHEA multicity project. Environ Health Perspect. 2005;113(1):88-95.

23. Schwartz J, Laden F, Zanobetti A. The concentration-response relation between $PM_{(2.5)}$ and daily deaths. Environ Health Perspect. 2002;110 (10):1025-1029.

24. Bell ML, Peng RD, Dominici F. The exposure-response curve for ozone and risk of mortality and the adequacy of current ozone regulations. Environ Health Perspect. 2006;114(4): 532-536.

25. Hoek G, Schwartz JD, Groot B, Eilers P. Effects of ambient particulate matter and ozone on daily mortality in Rotterdam, the Netherlands. Arch Environ Health. 1997;52(6):455-463.

26. Alessandrini ER, Stafoggia M, Faustini A, et al; on behalf of the EpiAir2 Study Group. Association between short-term exposure to $PM_{2.5}$ and PM_{10} and mortality in susceptible subgroups: a multisite case-crossover analysis of individual effect modifiers. Am J Epidemiol. 2016;184(10):744-754.

27. Franklin M, Zeka A, Schwartz J. Association between $PM_{2.5}$ and all-cause and specific-cause mortality in 27 US communities. J Expo Sci Environ Epidemiol. 2007;17(3):279-287.

28. Franklin M, Koutrakis P, Schwartz P. The role of particle composition on the association between $PM_{2.5}$ and mortality. Epidemiology. 2008;19(5): 680-689.

29. Levy JI, Chemerynski SM, Sarnat JA. Ozone exposure and mortality: an empiric Bayes meta regression analysis. Epidemiology. 2005;16(4): 458-468.

30. Peng RD, Samoli E, Pham L, et al. Acute effects of ambient ozone on mortality in Europe and North America: results from the APHENA study. Air Qual Atmos Health. 2013;6(2):445-453.

31. Tamagawa E, Bai N, Morimoto K, et al. Particulate matter exposure induces persistent lung inflammation and endothelial dysfunction. Am J Physiol Lung Cell Mol Physiol. 2008;295(1):L79-L85.

32. Bartoli CR,Wellenius GA, Coull BA, et al. Concentrated ambient particles alter myocardial blood flow during acute ischemia in conscious canines. Environ Health Perspect. 2009;117(3):333-337.

33. Nemmar A, Hoet PH, Vermylen J, Nemery B, Hoylaerts MF. Pharmacological stabilization of mast cells abrogates late thrombotic events induced by diesel exhaust particles in hamsters. Circulation. 2004;110(12):1670-1677.

34. Hemmingsen JG, Rissler J, Lykkesfeldt J, et al. Controlled exposure to particulate matter from urban street air is associated with decreased vasodilation and heart rate variability in overweight and older adults. Part Fibre Toxicol. 2015;12:6.

35. Langrish JP, Mills NL, Chan JK, et al. Beneficial cardiovascular effects of reducing exposure to particulate air pollution with a simple facemask. Part Fibre Toxicol. 2009;6:8.

36. Mills NL, Törnqvist H, Gonzalez MC, et al. Ischemic and thrombotic effects of dilute diesel-exhaust inhalation in men with coronary heart disease. N Engl J Med. 2007;357(11):1075-1082.

37. Bräuner EV, Forchhammer L, Møller P, et al. Indoor particles affect vascular function in the aged: an air filtration-based intervention study. Am J Respir Crit Care Med. 2008;177(4):419-425.

38. Watkinson WP, Campen MJ, Nolan JP, Costa DL. Cardiovascular and systemic responses to inhaled pollutants in rodents: effects of ozone and particulate matter. Environ Health Perspect. 2001; 109 (suppl 4):539-546.

39. Chuang KJ, Chan CC, Su TC, Lee CT, Tang CS. The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults. Am J Respir Crit Care Med. 2007;176(4):370-376.

40. Di Q,Wang Y, Zanobetti A, et al. Air pollution and mortality in the Medicare population. N Engl J Med. 2017;376(26):2513-2522.

Supplementary Material

Appendix 1. Covariates and Spatial Join

In the subgroup analysis, we used 5 variables to define subgroups, including sex (male or female), race (White, non-White, and other groups), age at death (≤69, 70~74, 75~84, and ≥85), Medicaid eligibility (as a proxy for low socioeconomic status [SES]), and quartiles of population density. Sex, race, age at death, and Medicaid eligibility were either retrieved or calculated from Medicare data. Population density was obtained from the 2000 US Census and the 2010 US Census.

Daily air and dew point temperatures data were retrieved from the North American Regional Reanalysis data in approximate 32 km × 32 km grids. We acquired daily 1 km × 1 km gridded air pollution levels ($PM_{2.5}$ and ozone) from previously developed and validated air pollution prediction models.[1,2] The prediction models predict the daily mean of $PM_{2.5}$ and 8-hour maximum ozone. For each individual, we extracted the residential zip code at death and obtained air temperature, dew point temperature, $PM_{2.5}$, and ozone levels by taking the inverse-distance mean of the 4 nearest grid cells to the zip code's centroid.

We used air pollution monitoring data from the United States Environmental Protection Agency's Air Quality System for the Nearest Monitor Analysis.[3] We obtained the daily mean of $PM_{2.5}$ and daily 8-hour maximum ozone. 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 sites within 50 km, we treated the monitored exposure level as missing and excluded that zip code from the analysis.

Appendix 2. Pooled Analysis

We estimated the exposure-response curves between both air pollutants and mortality using penalized splines. Due to computational limitations, running a conditional logistic regression with penalized splines on the whole data set was not possible. Instead, we randomly divided the entire data set into 50 groups with equal probability and estimated exposure-response curves for each group separately. To combine exposure-response curves from group-level analyses, we applied the meta-smoothing approach that was used and modified in previous studies.[5-7] In each group, the predicted relative risk (RR) and its point-wise standard error were computed for each 1 μg/m^3 increment in $PM_{2.5}$ or 1 ppb increment in ozone. These group-level effect estimates ($\hat{\beta}_{ij}$=log RR) in each group *i* and for exposure level *j*, and corresponding standard error $se(\hat{\beta}_{ij})$ were combined by regressing $\hat{\beta}_{ij}$ 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 the indicator variable for exposure level *j*, and V_{ij} is the estimated variance in group *i* at exposure level *j*.

The meta-analysis was implemented with R package *mvmeta.*[8]

Appendix 3. Statistical Model

Statistical Model

Conditional logistic regression takes stratification into consideration.[4] The analysis included 1 case day and 3 or 4 control days in each stratum; we assumed 3 control days here for illustration purposes. The contribution of likelihood of a stratum was:

$$P(Y_{i1} = 1, Y_{i2} = 0, Y_{i3} = 0, Y_{i4} = 0 | \boldsymbol{X}_{i1}, \boldsymbol{X}_{i2}, \boldsymbol{X}_{i3}, \boldsymbol{X}_{i4})$$

$$= \frac{\exp(\alpha_i + \boldsymbol{\beta}^T \boldsymbol{X}_{i1})}{\exp(\alpha_i + \boldsymbol{\beta}^T \boldsymbol{X}_{i1}) + \exp(\alpha_i + \boldsymbol{\beta}^T \boldsymbol{X}_{i2}) + \exp(\alpha_i + \boldsymbol{\beta}^T \boldsymbol{X}_{i3}) + \exp(\alpha_i + \boldsymbol{\beta}^T \boldsymbol{X}_{i4})}$$

where $Y_{il} \in \{0,1\}$ is the death indicator of the lth observation in the ith stratum; $\boldsymbol{X}_{il}$ are the predictors, and included $PM_{2.5}$, ozone, splines of air temperature and dew point temperature; $\boldsymbol{\beta}$ are regression coefficients; α_i is stratum-specific intercepts, which cancels out.

We implemented conditional logistical regression in R software, version 3.3.2.

Estimation of Related Risk

The relative risk increase (RRI) of all-cause mortality for 10-μg/m^3 increase in $PM_{2.5}$ and 10-ppb increase in ozone was given by the conditional logistic regression: $RR_{PM2.5} = exp(10 * \beta_{PM2.5})$ and $RR_{ozone} = exp(10 * \beta_{ozone})$. We also calculated absolute risk difference (ARD), that is, the difference in the daily mortality rate associated with 10 μg/m^3 increase in short-term exposures to $PM_{2.5}$ as following. First, we calculated the baseline daily mortality rate as the daily death rate in the Medicare population during our study period, which we denoted as α. We then calculated the ARD associated with 10 μg/m^3 increase in $PM_{2.5}$ as $ARD_{PM2.5} = \alpha * \frac{RR_{PM2.5}-1}{\mathrm{RR_{PM2.5}}}$

and its standard error as $se(ARD_{PM2.5}) = \alpha * exp(-\beta_{PM2.5} * 10) * se(10 * \beta_{PM2.5})$ according to the delta method. For ozone, the ARD estimate was calculated in a similar way, but using baseline daily mortality rate only for the warm season (from April to September). We calculated subgroup-specific ARD by using subgroup-specific daily mortality rate.

Appendix 4. Sensitivity Analysis of Warm-Season Results

We restricted our analysis to the warm season (April 1 to September 30) when estimating the effect size of ozone in the main analysis. In a sensitivity analysis, we defined the warm season as May 1 to October 30 and restricted the analysis within this period. Results indicated that every 10-ppb increase in ozone was associated with a 0.59% (95% CI: 0.49%, 0.68%) increase in mortality, compared with a 0.51% (95% CI: 0.41%, 0.61%) increase in the main analysis (Table 3-2).

We reported the exposure-response relationship for $PM_{2.5}$ in Figure 3-5; here, we also reported the exposure-response relationship for $PM_{2.5}$ restricted to the warm season (April 1 to September 30) (Figure S3-4). We reported the exposure-response relationship for ozone during the warm season (April 1 to September 30) only in Figure 3-5; here, we also reported the exposure-response relationship for the entire year (Figure S3-4).

Appendix 5. Test for Interaction

To test for statistically significant difference in RR estimates across categories within subgroups, for example, in males vs females (H_0: $RR_{male} = RR_{female}$), we calculated: $Z = \frac{RR_{male} - RR_{female}}{\sqrt{se(RR_{male})^2 + se(RR_{female})^2}}$. We tested whether ARD is significantly different in a similar way, using point estimate and standard error of ARD.

We also tested whether RR estimates are significantly different below and above a certain air pollution threshold. Subgroups were defined in which one category of individuals died with exposure levels above the threshold and the second category died below the threshold. We repeated the above calculation to test whether RR estimates are significantly different.

Table S3-1. Sensitivity Analysis Using the Same Day Exposure (lag 0 day) and Previous Day Exposure (lag 1 day) and Mean of Daily Exposure on the Same Day of Death and One Day Prior (lag 01 day) of $PM_{2.5}$ and Ozone

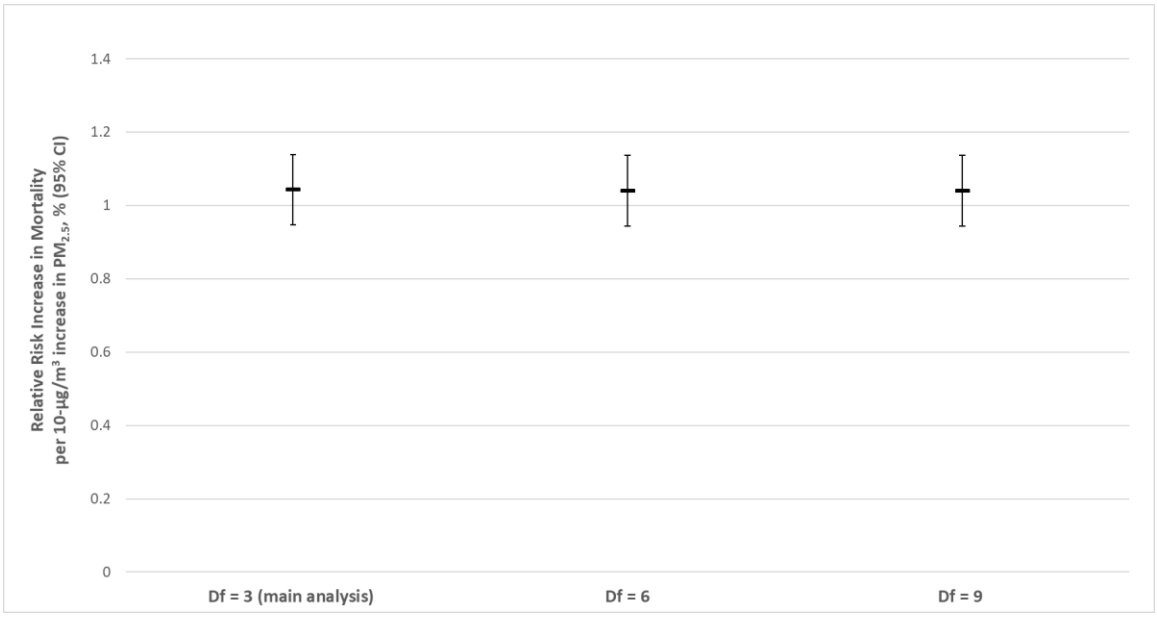
	$PM_{2.5}$[a]		Ozone[a]	
	Relative Risk Increase Estimate (95% CI)	**AIC[b]**	**Relative Risk Increase Estimate (95% CI)**	**AIC[b]**
Lag 01 Day[c]	1.05% (0.95%, 1.15%)	64,646,725	0.51% (0.41%, 0.61%)	30,635,577
Lag 1 Day[d]	0.83% (0.67%, 1.00%)	64,646,901	0.55% (0.38%, 0.72%)	30,635,801
Lag 0 Day[d]	0.79% (0.62%, 0.95%)	64,646,854	0.35% (0.19%, 0.51%)	30,635,663

[a] The analysis estimated $PM_{2.5}$ effect based on case days and control days from the entire year, while the ozone analysis used case days and control days from the warm season only (April 1 to September 30).

[b] Akaike information criterion

[c] The main analysis; results identical to Table 3-2.

[d] To conduct the sensitivity analysis, we repeated the main analysis, but used the same day exposure (lag 0 day) and previous day exposure (lag 1 day). This model also controlled for natural splines of air and dew point temperatures with 3 degrees of freedom. The 2-pollutant analysis estimated the percentage increase in daily mortality rate associated with each 10 $\mu g/m^3$ increase in $PM_{2.5}$ exposure adjusted for ozone and the percentage increase in the daily mortality rate associated with each 10 ppb increase in ozone exposure adjusted for $PM_{2.5}$.



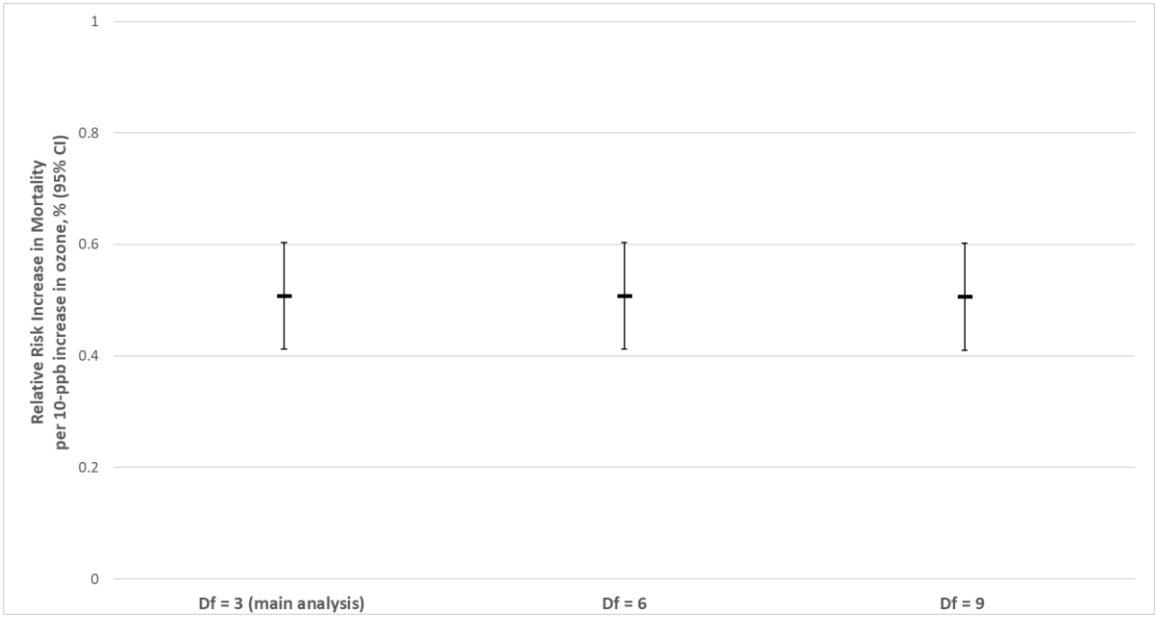
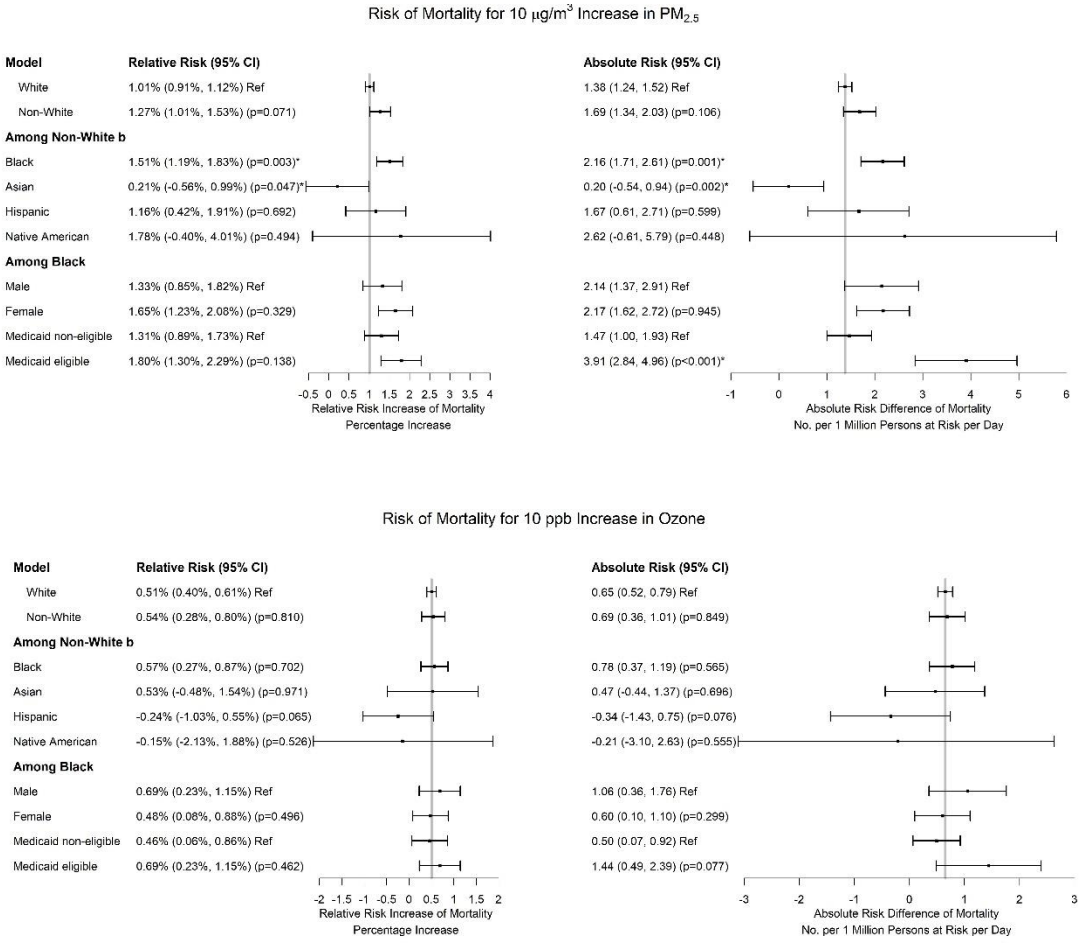


Figure S3-1. Sensitivity Analysis Using Splines on Meteorological Variables with More Degrees of Freedom.

We repeated the main analysis, but changed the natural splines on air temperature and dew point temperature to 6 degrees of freedom and 9 degrees of freedom. The sensitivity analysis estimated the percentage increase of mortality associated with each 10-µg/m^3 increase in $PM_{2.5}$ exposure

adjusted for ozone and the percentage increase of mortality associated with each 10 ppb increase in ozone exposure adjusted for $PM_{2.5}$. The error bar indicates 95% confidence interval.



6

Figure S3-2. Relative Risk Increase and Absolute Risk Difference of Daily Mortality Associated with Each 10 μg/m^3 Increase in $PM_{2.5}$ and 10 ppb Increase in Ozone among Non-Whites

Figure S3-2 was created using the same method as that described in Figure 3-2. The vertical lines were placed at the effect estimate for White individuals.

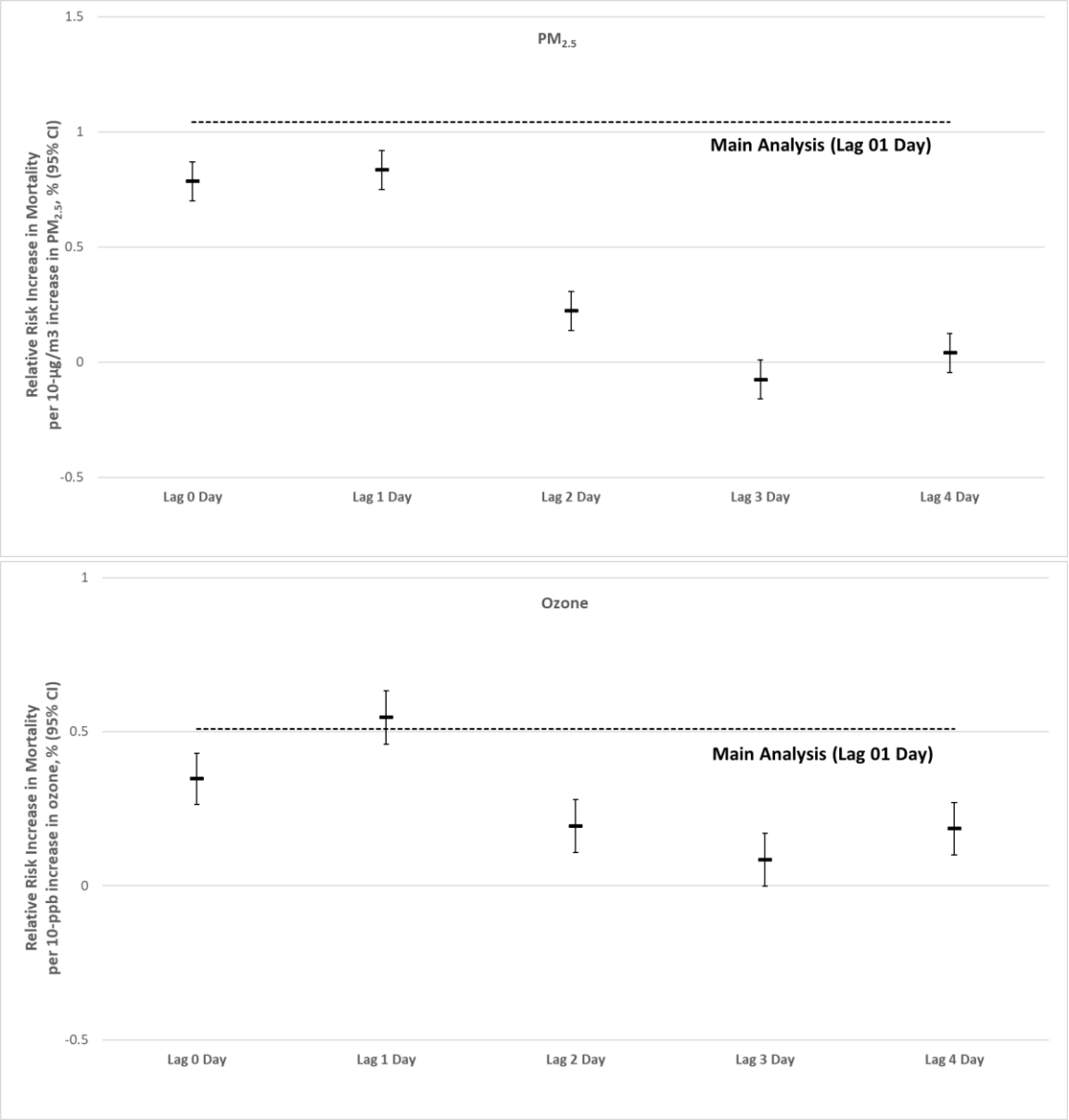


Figure S3-3. Relative Risk Increase Associated with Each 10 µg/m^3 Increase in $PM_{2.5}$ and 10 ppb Increase in Ozone for Single-lag Models.

As a sensitivity analysis of the exposure time window, we used single-lag 2-pollutant models and compared them to our main models that used the mean of daily exposure on the same day of death and one day prior (Lag 01 Day). For example, in the Lag 0 Day Model, we included ozone and $PM_{2.5}$ levels at day 0 (on the same day of death), temperature, and dew point temperature; in

the Lag 1 Day Model, we included ozone and $PM_{2.5}$ levels at day 1 (1 day before the date of death), temperature, and dew point temperature. We fit the single-lag models separately and obtained risk estimates. We considered air pollution from the same day (lag 0 day) to up to 4 days (lag 4 day). Figure S3-3 illustrates that lag 0 day and lag 1 day are most relevant to daily mortality. Air pollution concentrations 2 days prior to the date of death were less relevant to daily mortality. Based on the sensitivity analysis results (Table S3-1), we used the mean of daily exposure on the same day of death and one day prior (lag 01 day) as the exposure metric for both $PM_{2.5}$ and ozone.

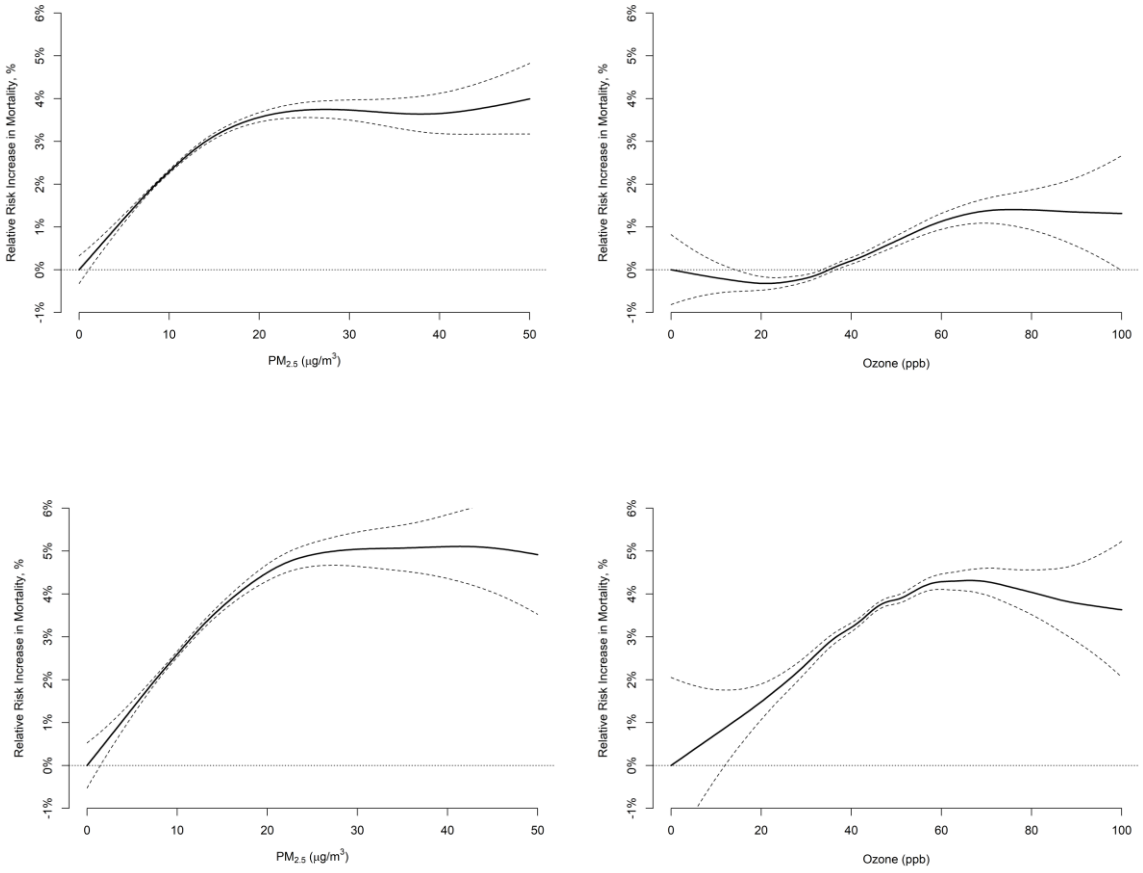


Figure S3-4. Estimated Exposure-response Curves for Short-term Exposures to $PM_{2.5}$ and Ozone for the Entire Year and Restricted to the Warm Season.

A 2-pollutant analysis with separate penalized splines on $PM_{2.5}$ (left panels) and ozone (right panels) was conducted to assess the percentage increase in daily mortality at various pollution levels. Dashed lines indicate 95% confidence intervals. The mean of daily exposure on the same day of death and one day prior (lag 01 day) were used as metrics of $PM_{2.5}$ and ozone. We plotted the exposure-response relationships for the entire year (upper panels). Analysis for both air pollutants were repeated and restricted to the warm season (April to September) (lower panels).

References

1. Di Q, Kloog I, Koutrakis P, Lyapustin A, Wang Y, Schwartz J. Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. *Environ Sci Technol.* 2016;50(9):4712-4721.

2. 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(1):39-52.

3. US Environmental Protection Agency. Air Quality System Data Mart (daily summary of $PM_{2.5}$ FRM/FEM Mass (88101), daily summary of ozone (44201)). 2016.

4. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. *Am J Epidemiol.* 1991;133(2):144-153.

5. 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(6):666-672.

6. Schwartz J, Ballester F, Saez M, et al. The concentration-response relation between air pollution and daily deaths. *Environmental health perspectives.* 2001;109:1001-1006.

7. Schwartz J, Laden F, Zanobetti A. The concentration-response relation between $PM_{(2.5)}$ and daily deaths. *Environmental health perspectives.* 2002;110(10):1025-1029.

8. Gasparrini A, Armstrong B, Kenward MG. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat Med.* 2012;31(29):3821-3839.

Conclusion Remarks

Air pollution is a serious public health concern, and a major cause of death and multiple adverse health outcomes. My colleagues and I investigated the association between all-cause mortality and exposures to air pollution in the long-term and short-term, particularly at places with air pollution levels below the current air quality standards. We identified subgroups that are more vulnerable to air pollution than the rest of the population. Our studies provide strong evidence to reevaluate the current air quality standards.

We found that together with neural network, satellite data, meteorological variables, land-use variables, and some other data sets are good predictors of ground-level $PM_{2.5}$ concentration. This air pollution model predicted daily $PM_{2.5}$ concentrations from 2000 to 2012 for every square kilometer in the continental United States with high accuracy. Model prediction and monitored $PM_{2.5}$ have a good agreement ($R^2 = 0.84$). Model performance for ozone was also good ($R^2 = 0.76$).

In the survival analysis on the long-term effect of air pollution, we found that increases of 10 $\mu g/m^3$ in annual $PM_{2.5}$ and of 10 ppb in warm-season ozone were associated with increases in all-cause mortality of 7.3% (95% confidence interval [CI], 7.1 - 7.5) and 1.1% (95% CI, 1.0 - 1.2), respectively. When restricting analyses to $PM_{2.5}$ below 12 $\mu g/m^3$ (the current annual air quality standard) and ozone below 50 ppb, the same increase in $PM_{2.5}$ and ozone were associated with all-cause mortality increase of 13.6% (95% CI, 13.1 - 14.1) and 1.0% (95% CI, 0.9 - 1.1), respectively. This low-exposure analysis provides strong evidence that the association between air pollution and mortality remains below the current air quality standards. The current air quality standards need to be reevaluated. Moreover, males, African Americans, Asian, Hispanic persons; and people with low socioeconomic status had higher risk of mortality associated with air pollution. These groups are more vulnerable to air pollution than the rest of the population.

In the case-crossover analysis on the short-term effect of air pollution, we found that each 10 $\mu g/m^3$ and 10 ppb increase in $PM_{2.5}$ and ozone (the mean exposure level on the same day of death and one day prior) was associated with increases in daily all-cause mortality of 1.05% (95% CI, 0.95%-1.15%) and 0.51% (95% CI, 0.41%-0.61%). When restricting analyses to days with $PM_{2.5}$ below 25 $\mu g/m^3$ and ozone below 60 ppb, much lower than the current daily air quality standards, the same increases in $PM_{2.5}$ and ozone were associated with daily mortality increase of 1.61% (95% CI, 1.48 – 1.74) and 0.58% (95% CI, 0.46 – 0.70), respectively. Again, this low-exposure analysis indicate that the current daily air quality standards need to be reevaluated. In subgroup analysis, we found that people of low socioeconomic status, older than 70 years, females, and African Americans had higher risk of mortality associated with air pollution.

Both the long-term and short-term analyses did not find any threshold level, at which level air pollution is safe and imposes no health effect. In other words, any levels of air pollution, no matter how low it is, is associated with increased risk of mortality.

In conclusion, air pollution is associated with increased risk of mortality. Moreover, long-term and short-term exposures to low levels of air pollution are associated with increased risk of mortality, providing strong evidence that the current air quality standards need to be reevaluated.