Original Research

ENVIRONMENTAL EPIDEMIOLOGY

OPEN

Estimating the causal effect of annual $PM_{2.5}$ exposure on mortality rates in the Northeastern and mid-Atlantic states

Maayan Yitshak-Sade[a], Itai Kloog[b], Antonella Zanobetti[a], Joel D. Schwartz[a,c]

Background: Dozens of cohort studies have associated particulate matter smaller than 2.5 μm in diameter ($PM_{2.5}$) exposure with early deaths, and the Global Burden of Disease identified $PM_{2.5}$ as the fifth-ranking mortality risk factor in 2015. However, few studies have used causal modeling techniques. We assessed the effect of annual $PM_{2.5}$ exposure on all-cause mortality rates among the Medicare population in the Northeastern and mid-Atlantic states, using the difference-in-differences approach for causal modeling.

Methods: We obtained records of Medicare beneficiaries 65 years of age or more who reside in the Northeastern or mid-Atlantic states from 2000 to 2013 and followed each participant from the year of enrollment to the last year of follow-up. We estimated the causal effect of annual $PM_{2.5}$ exposure on mortality rates using the difference-in-differences approach in the Poisson survival analysis. We controlled for individual confounders, for spatial differences using dummy variables for each ZIP code and for time trends using a penalized spline of year.

Results: We included 112,376,805 person-years from 15,401,064 people, of whom 37.4% died during the study period. The interquartile range (IQR) of the annual $PM_{2.5}$ concentration was 3 μg/m^3, and the mean annual $PM_{2.5}$ concentration ranged between 6.5 and 14.5 μg/m^3 during the study period. An IQR incremental increase in $PM_{2.5}$ was associated with a 4.04% increase (95% CI = 3.49%, 4.59%) in mortality rates.

Conclusions: Assuming no omitted predictors changing differently across ZIP codes over time in correlation with $PM_{2.5}$, we found a causal effect of $PM_{2.5}$ on mortality incidence rate.

Keywords: Causal modeling; PM2.5; Mortality; Air pollution

Introduction

Numerous epidemiological studies have concluded that long-term exposure to particulate matter smaller than 2.5 μm in diameter ($PM_{2.5}$) increases the risk of mortality, shortens life expectancy, and increases the risk of cardiovascular and respiratory morbidity.[1–13] The Global Burden of Diseases, Injuries, and Risk Factors study ranked air pollution as the fifth highest mortality risk factor in 2015, mostly from ischemic heart disease and cerebrovascular disease.[1]

Randomized controlled trials (RCTs) are not feasible when studying population-based air pollution effects. Therefore, the majority of the evidence relies on classical observational approaches. Unlike RCTs, where the randomization assures that the exposure of interest is independent of all other parameters at the time of randomization, classical observational approaches are prone to confounding bias.[14] The causal modeling approach overcomes this limitation by using methods designed to mimic the randomization process as much as possible.

Causal modeling methods follow the potential outcome approach and aim to estimate the difference between the outcome of an exposed population, and the expected outcome of the same population had they received different exposure. Because that counterfactual outcome is missing by definition, causal modeling methods seek unbiased ways to estimate these unobserved potential outcomes.[15–17]

What this study adds

Numerous studies found higher mortality risks associated with long-term exposure to particulate matter smaller than 2.5 μm in diameter ($PM_{2.5}$). However, because randomized controlled trials are not feasible when studying population-based air pollution effects, the majority of the evidence originates from observational approaches. This cohort study of 112,376,805 person-years shows a significant effect of $PM_{2.5}$ on the risk of mortality among the elderly population. Therefore, the associations between $PM_{2.5}$ and mortality remain when using a casual modeling approach designed to overcome the limitations of a nonrandomized observational study.

[a]Exposure, Epidemiology, and Risk Program, Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; [b]Department of Geography and Environmental Development, Faculty of Humanities and Social Sciences, Ben-Gurion University, Beer-Sheva, Israel; and [c]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Restricted by our Data Use Agreement with the U.S. Centers for Medicare and Medicaid Services, the Medicare data that support the findings of this study are neither sharable nor publicly available. Academic and nonprofit researchers who are interested in using Medicare data shall contact with the US Centers for Medicare and Medicaid Services directly. The computer code can be made available upon request in writing to myitshak@hsph.harvard.edu.

National Institutes of Health NIH/NIEHS grants P30 ES000002, RO1 ES024332; EPA grants RD-83587201, RD-83615601, and RD-83587201. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the United States Environmental Protection Agency (USEPA). Furthermore, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication.

*Corresponding author: Department of Environmental Health Exposure, Epidemiology, and Risk Program Harvard T.H. Chan School of Public Health 401 Park Drive Boston, MA 02215. E-mail: myitshak@hsph.harvard.edu (M. Yitshak-Sade).

Copyright © 2019 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of Environmental Epidemiology. All rights reserved. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Environmental Epidemiology (2019) 3:e052

Received: 14 March 2019; Accepted 19 May 2019

Published online 17 June 2019

DOI: 10.1097/EE9.0000000000000052

Studies that used these causal modeling methods for air pollution have become more common in recent years,[15,18,19] but are still scarce. A common approach is the use of propensity score matching or inverse probability weighting, to make the exposure of interest independent of all measured confounders.[15,16] Makar et al[18] used a propensity score approach in a representative sample of 32,119 Medicare beneficiaries. Wang et al[19] applied a doubly robust causal modeling estimate with inverse probability weights to look at $PM_{2.5}$ and mortality in the Southeastern United States. The disadvantage of these methods is that it can only account for measured confounders and for unmeasured confounders that are highly correlated with the measured confounders.[20]

The difference-in-differences (DID) approach overcomes this limitation and allows controlling for unmeasured confounders by design. In recent decades, the increasing availability of routinely collected electronic health records has provided the opportunity to conduct large population-based studies. However, because the data are not collected for specific study aims, many potential confounders remain unmeasured and may bias the studied associations. Hence, the importance of controlling for these unmeasured confounders.[21]

In a classical DID model, mean outcomes are calculated for the exposed and nonexposed groups in the preexposure and postexposure periods. The difference between preexposure and postexposure periods in the unexposed group is a negative outcome control for the difference in the exposed group, and the difference in these pairs of differences is a causal estimate, assuming that no other exposure has affected the two groups differently over time.[22] This is because the difference between death rates in the same location between two periods controls for all slowly varying predictors of mortality such as socioeconomic status (SES), smoking, obesity, etc. The difference between the two periods in the control location controls for time trends in an outcome that is similar between the two locations.

In a previous analysis done by our group, Wang et al[23] applied a variant of the DID approach among all mortality cases in New Jersey. We applied the DID approach of Wang et al[23] to the Medicare population of the Northeastern and mid-Atlantic states, extended the analysis to incorporate individual covariates, and assessed the effect of annual $PM_{2.5}$ exposure on all-cause mortality rates.

Methods

Study population

We included all the Medicare beneficiaries 65 years of age or more who reside in the following states from 2000 to 2013: Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Delaware, Pennsylvania, Maryland, Washington, D.C., Virginia, and West Virginia. Participants entered the cohort on January 1 of the year after they became Medicare participants, and were followed for each calendar year until death or end of the study. This study was approved by the human subjects committee at the Harvard T. H. Chan School of Public Health.

Exposure data

Our group generated highly spatially resolved $PM_{2.5}$ data (1 × 1 km spatial resolution) from a hybrid satellite-based model incorporating daily satellite remote sensing Aerosol Optic Depth data and classic land-use regression methodologies. We used out-of-sample 10-fold cross-validation to quantify the accuracy of the model predictions and found excellent model performance for both days with available satellite data (mean out-of-sample R^2 = 0.88) and days without satellite observations (mean out-of-sample R^2 = 0.87). For more in-depth description, please refer to Kloog et al.[24]

We averaged the exposure values of all grid cells located within each ZIP code and a 500 meters buffer annually and assigned exposure data to ZIP codes using ArcGIS based on spatial location.

Covariates

The Medicare denominator file includes data on the participant's age, sex, race, date of death, ZIP code of residence, and eligibility for Medicaid, which provides additional coverage for low-income participants.

We obtained the mean annual temperature in the summer and winter seasons from the North American Regional Reanalysis data and estimated daily mean values for each ZIP code in our study area.[25]

Statistical analysis

We used a causal modeling approach to assess the effect of long-term $PM_{2.5}$ exposure on the rate of death. Specifically, we used the DID approach in a Poisson survival analysis using the Anderson–Gill formulation with time-varying covariates.[26] We have extended the classical DID approach to continuous exposures and multiple time periods.[23] We used a similar approach in this study, but added some controls for individual covariates. We followed Medicare participants in 7600 ZIP codes in the Northeastern and mid-Atlantic states. We begin with the potential outcomes framework of the Rubin Causal Model.[27] Let $Y^{A=a}_{i,c,t}$ be the potential outcome in person I, in ZIP code c if exposed to $A = a$ in year t, and let $Y^{A=a'}_{i,c,t}$ be the potential outcome under the alternative exposure $A = a'$. We assumed that the incidence of death depends on predictors in the following manner:

$$Log\left(E\left(Y^{a}_{i,c,t}\right)\right) = \beta_0 + \beta_1\, a_t + \beta_2\, Z_c + \beta_3 U_t + \beta_4 W_{i,} + \beta_5 W_{i,t} \qquad (1)$$

where Z_c is small area level confounders that vary between ZIP codes but not over the study period; U_t is temporal confounders that vary over time but not between ZIP codes; W_i is the person's individual confounders that do not vary over time and $W_{i,t}$ is the person's individual confounders that do vary over time.

A dummy variable for each ZIP code will effectively remove confounding by variables that only vary across exposure groups (Z_c and W_i). Similarly, a penalized spline for the year of study will remove all the temporal confounders that vary between periods (U_t and $W_{i,t}$). The adjustment to the individual characteristics will account for individual confounders that may not be captured by adjustment to the ZIP code of residence. Then, the remaining variability of the annual $PM_{2.5}$ levels can be treated as the difference in these pairs of differences (i.e., the causal estimate). The benefit of this approach is that we are able to control for unmeasured confounders, because they are removed by the time trend and ZIP code intercepts. This is a generalization of the original DID approach where there were only two treatment units and two time periods, and the causal effect of treatment was the interaction term. It reduces to the same thing in that case.[23,28]

The DID assumptions

For a causal interpretation of the DID estimate, the following assumptions must hold[29]:

1. The intervention (or the exposure, in this case) is unrelated to the outcome at baseline. Meaning that the level of annual $PM_{2.5}$ exposure is not determined by the mortality rates.
2. All the ZIP codes have parallel trends in mortality rates. Meaning that, had the $PM_{2.5}$ levels not been changing the same over time, the difference in mortality rates between the ZIP codes will be stable over time as well. This can be weakened to allow differences in the trend in mortality rates over time by ZIP code due to other factors, as long as those differences are not correlated with differences in $PM_{2.5}$ over time by ZIP code.

For the first assumption to hold, mortality in a current year should not affect the current exposure and should not be affected by past exposures. In a classical DID approach, there is a defined driver for the change in exposure between the preintervention and postintervention periods (usually an intervention), which can potentially be affected by the outcome. This seems unlikely in this case, because EPA policy changes are not impacted by Northeast mortality rates.

Regarding the second assumption, we must assume that differences from year to year in $W_{i,t}$ which are not common across ZIP codes are not correlated with $PM_{2.5}$. That is, as long as the annual *within* ZIP code changes in individual smoking rates, diet, etc., which are different from the same changes elsewhere are not correlated with annual *within* ZIP code changes in $PM_{2.5}$ that are different from the same changes elsewhere, then we will obtain a causal estimate of the effect of $PM_{2.5}$. Because this assumption cannot be tested statistically, it must be judged based on knowledge and plausibility from outside the study. The only factors that may potentially confound the association are factors like smoking rates or SES that may vary differently with time across ZIP codes. However, for these variables to confound the association, it requires a highly unlikely scenario in which these variations are correlated with variations in $PM_{2.5}$ levels from the average ZIP code level and the average annual level.

The long-term trend in exposure is removed by the penalized spline, and the exposure contrast is the local deviation from that trend. That can be parsed into two parts: difference between the long-term trend in the ZIP code from the trend in the Northeastern United States and fluctuations about that trend from year to year. The major driver of trends in $PM_{2.5}$ in the Northeast in this time period was EPA required reductions in NOx and SO_2 emissions at upwind power plants. ZIP code trends in $PM_{2.5}$ concentrations will vary depending on the extent to which the ZIP codes are downwind of the plants that put on controls during the period, and when the controls were installed. All these factors are unlikely related to trends in SES, smoking, or any other potential confounder.[23] That said, as an additional measure to control for potential confounding, we adjusted the models for individual characteristics that may be related with SES. The remaining variability in $PM_{2.5}$ depends mostly on meteorological factors, such as year-to-year differences in mixing height and wind speed, which seem random with respect to year-to-year differences from trend in ZIP code mortality rates.

However, it is possible that mortality in the current year will be related to exposures in earlier years. We, therefore, added two sensitivity analyses aimed to make sure that the effects of current year exposure are not confounded by earlier exposures. First, we repeated our model, assessing the effect of $PM_{2.5}$ in the year of death (time t), with adjustment for $PM_{2.5}$ exposure in the previous year (time $t-1$). Second, we obtained the residuals from our main analysis, which assessed the effect of $PM_{2.5}$ in time t on mortality, adjusting for time trends and ZIP code differences. We then assessed the mean within ZIP code correlation between the residuals in time t and time $t-1$.

The one variable that has the potential to confound the tested association despite the adjustment for the spatial, temporal, and individual confounders is temperature. We, therefore, adjusted the models to seasonal temperatures.

To support our findings, we added a sensitivity analysis analyzing the association between $PM_{2.5}$ exposure and alcoholic fatty liver disease, as a negative outcome control. This condition is a result of alcohol consumption, and therefore, theoretically should not be related to $PM_{2.5}$ exposure. Therefore, an association would suggest residual confounding by unmeasured confounders as shown in the Directed acyclic graph (Figure 1).

To further address the possibility of confounding by differential changes in SES across ZIP codes, we conducted an additional sensitivity analysis controlling for annual estimates of population density, percent over 65 years of age and living in poverty, and median household income.

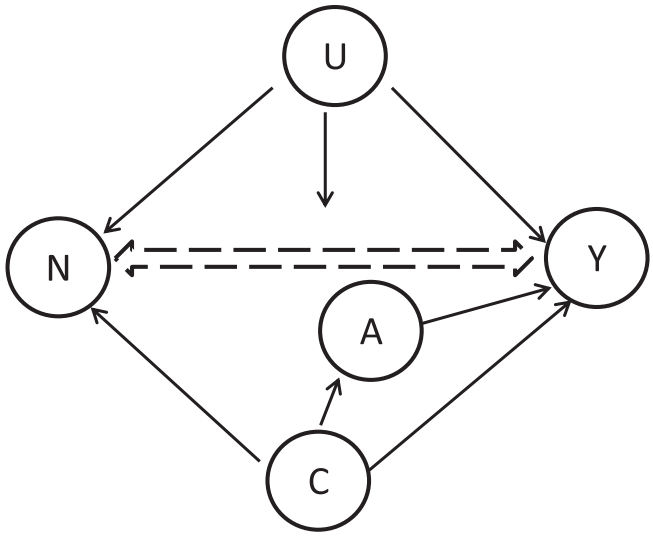

Figure 1. Directed acyclic graph for the causal association between air pollution exposure (A), mortality as the primary outcome (Y), fatty liver disease as the negative outcome control (N), measured confounders (C), and unmeasured confounders (U).

Due to computational limitations arising from the large sample size, we randomly split the data into subsets and applied our models on each subset. The effect estimates were then pooled using a fixed-effect meta-analysis. Results are presented as percent change for IQR increase in annual exposure.

We assessed modification of the $PM_{2.5}$ and mortality association by adding multiplicative interaction terms with race and eligibility to Medicaid services. Each interaction term was assessed separately.

Results

We included 112,376,805 person-years from 15,401,064 people, of whom 37.4% have died during the study period. The mean age of the study population was 78 years, and the mean age of death was 82 years, 42.8% were males, 85.2% were white, and 16.2% had dual eligibility for the Medicare and Medicaid services (Table 1). The IQR of the annual $PM_{2.5}$ concentration was 3 $\mu g/m^3$. Figure 2 shows the states and ZIP codes included in the study and the centiles of the mean annual $PM_{2.5}$ concentration, ranging between 6.5 and 14.5 $\mu g/m^3$ throughout the study period (Figure 2).

After adjustment for personal characteristics and the spatial and temporal annual variation in $PM_{2.5}$, an IQR incremental increase in $PM_{2.5}$ was associated with a 4.04% increase (95% CI = 3.49%, 4.59%; $P < 0.001$) in mortality rates. The effect was modified by eligibility to Medicaid insurance and race (interaction P value <0.001 for both), with larger effects among people who are eligible to Medicaid Services (5.99%; 95% CI = 4.38%, 7.62%) and among blacks (10.10%; 95% CI = 8.56%, 11.67%) (Table 2).

We conducted several sensitivity analyses to assess whether the assumptions, required for a causal interpretation of the estimate, are met. First, when we repeated our model with adjustment to the $PM_{2.5}$ exposure in the previous year. The effect of $PM_{2.5}$ in the current year remained unchanged (4.10% increase; 95% CI = 3.55%, 4.65%; $P < 0.001$), and we observed no association with the earlier exposure (0.33% increase; 95% CI = −0.19%, 0.86%; $P = 0.189$). In addition, after removing the variability in the outcome that is related to $PM_{2.5}$ in the current year, ZIP code differences, time trend, and individual

Table 1

Population characteristics (n = 15,401,064 people)

Personal characteristics	
Death, n (%)	5,765,395 (37.4)
Male, n (%)	6,594,913 (42.8)
Race, n (%)	
White	13,119,522 (85.2)
Black	1,495,027 (9.7)
Asian	231,128 (1.5)
Other	487,913 (3.2)
Medicaid eligibility, n (%)	2,491,551 (16.2)
Mean age, mean ± SD	78.0 ± 8.5

confounders – the mean Pearson's correlation – of the within ZIP code residuals, was very low (r = –0.02). Second, when we tested the association with alcoholic fatty liver as a negative outcome, we found no association with $PM_{2.5}$ (3.0% increase; 95% CI = –4.7%, 11.4%; P = 0.456). Third, when we added adjustment for annual estimates of percent of persons 65 years of age and more living in poverty, population density, and median household income, the effect was similar (3.88%; 95% CI = 3.33%, 4.43%) to the base analysis. In addition, the correlation between median income in the first and last year of the study by ZIP code was 0.86, suggesting that there has been little trend in status by ZIP code during the period.

Discussion

This study of 14 years of follow-up involved all Medicare beneficiaries of 14 US states, includes a large number of persons and deaths, and also captures participants who reside in rural areas. We found a significant effect of $PM_{2.5}$ on the risk of mortality using a method that controls for many, and possibly all, omitted confounders. These findings suggest that the associations between $PM_{2.5}$ and mortality remain when using a casual modeling approach designed to overcome the limitations of a nonrandomized observational study. The specific approach used in this study only examines the association of year-to-year changes in $PM_{2.5}$ within ZIP code with year-to-year changes in mortality risk within ZIP code, after removing time trends from both, and control for some key individual covariates.

These results are in line with results from another study conducted among this study population, which estimated the

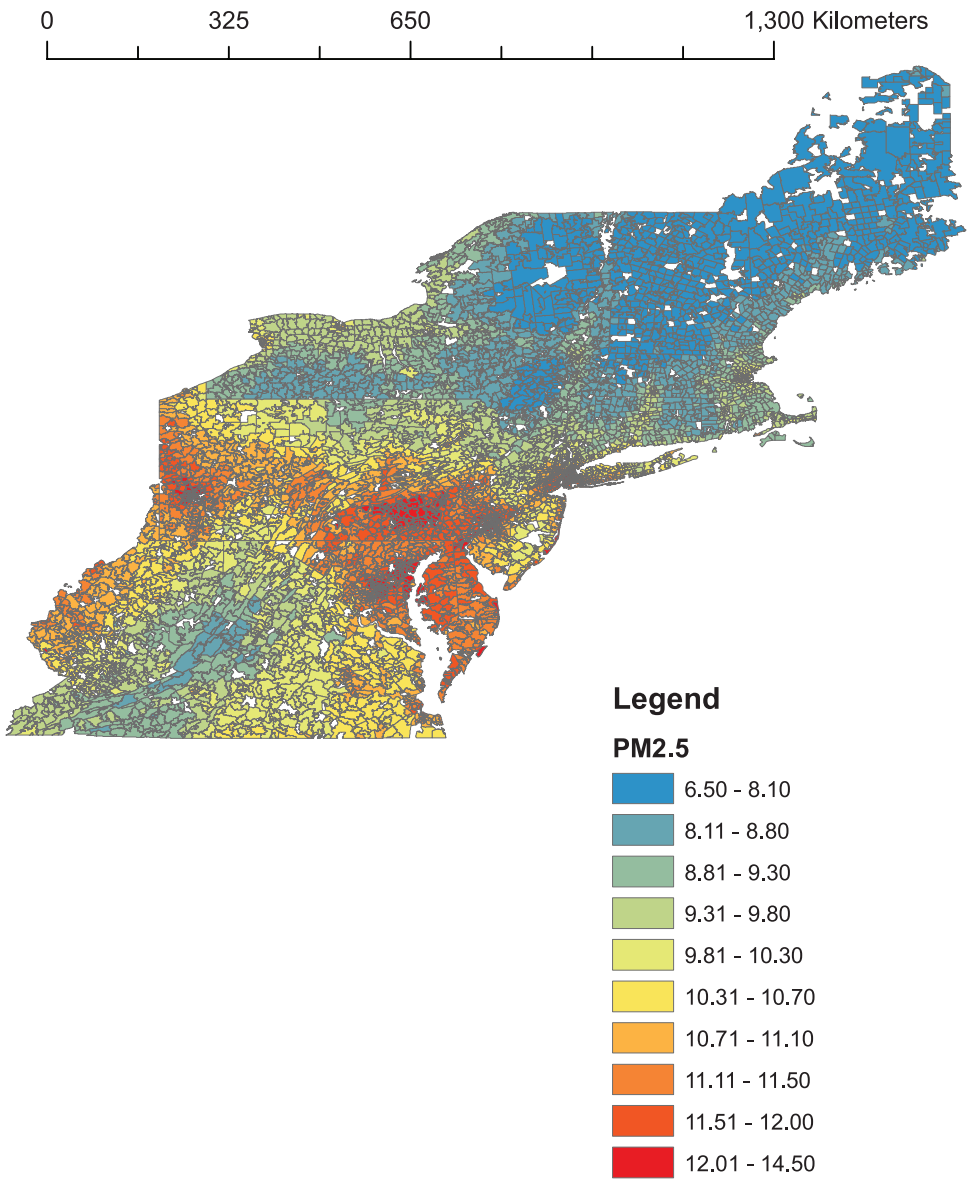

Figure 2. The ZIP codes included in the study and the centiles of the mean annual $PM_{2.5}$.

Table 2

Percent increase in total mortality, associated with interquartile range increase (3 μg/m³) in annual $PM_{2.5}$ exposure

	% Change (95% confidence intervals)
Main effect	4.04% (3.49%; 4.59%)
Interaction with Medicaid eligibility (*P* value)	<0.001
No Medicaid eligibility	3.54% (2.93%; 4.15%)
Medicaid eligibility	5.99% (4.38%; 7.62%)
Interaction with race (*P* value)	<0.001
White	3.35% (2.87%; 3.84%)
Black	10.10% (8.56%; 11.67%)

distribution of life expectancy as a function of annual $PM_{2.5}$ exposure. Comparing the percent of the population that would die before the age of 76 years, the authors found higher mortality rates among those exposed to $PM_{2.5}$ at 12 μg/m³ (23.5%) than those exposed to $PM_{2.5}$ at 7 μg/m³ (20.1%). While this study uses a different methodology and a different outcome, it support our findings of the causal effect of $PM_{2.5}$ on mortality.[30]

We believe the trends within ZIP code exposures found in our study are likely to be independent of potential confounders, and therefore, the association identified is likely causal. It is possible that events occurred during the study period that affected both the exposure and the outcome differentially across ZIP code, and therefore may still confound the association. For example, if a major industrial site was closed, this has the potential to affect both $PM_{2.5}$ levels and mortality rate only among ZIP code which are located closer to the factory. However, because our analysis only included the elderly population, this scenario is less likely. Furthermore, such a change would likely be felt through a change in SES, and control for annual SES in each ZIP code did not modify our results. In addition, since most of the change in $PM_{2.5}$ in our study period is due to nonlocal sources, we think that differential temporal changes by zip codes and other predictors of health are not likely to be correlated. The decline in $PM_{2.5}$ in the United States over this time period derives from two basic sources. First, SO_2, and NOx controls were retrofit on many coal-burning power plants during this time period. This reduced concentrations of sulfate, nitrate, and secondary organic aerosols downwind of those plants and the impact was felt primarily in ZIP codes hundreds of miles downwind. However, because that assignment to more or less secondary particles is random with respect to socioeconomic changes, or other predictors of health, and is rather the result of wind patterns, we do not expect it to confound the association. The second source of decline was the imposition of particle controls on new Diesel engines, and the gradual retirement of older trucks. Because this reduction in emissions is due to lower emissions rates per truck, and not less truck traffic, it is less likely to impact socioeconomic patterns.

Sofer et al[31] suggested the negative outcome controls can be used to detect confounding bias. A negative outcome can be any measured outcome that is not causally related to the exposure of interest and is influenced by the same unmeasured confounders of the exposure–outcome association of interest.[31] We chose alcoholic fatty liver as a negative outcome and found no association with $PM_{2.5}$. This suggests that unmeasured confounders do not confound the association of interest in our study.

The association between particulate pollution, morbidity, and mortality has been well documented in many observational studies.[8,11,32–35] Even with the decreasing trend of air pollution, $PM_{2.5}$ exposure is still linked to excess mortality.[32] For example, a recent cohort study of all Medicare beneficiaries in the United States found a 7.3% increase in all-cause mortality associated with 10 μg/m³ increase in annual $PM_{2.5}$ exposure.[12] Another example is a study that included seven US southeastern states and found a 23% increase in all-cause mortality associated with 10 μg/m³ increase in annual $PM_{2.5}$ exposure.[36] These associations, reported repeatedly in many settings, different populations, and geographical areas are compelling. However, by definition, these studies describe associations and do not prove causation.

Controlled exposure studies can assess underlying mechanisms and prove causation.[37] A few controlled exposure studies that assessed the effect of air pollution on ECG changes[38] or heart rate variability[39] were conducted over the years. Another trial has randomized people to either functional or sham in home particle filters for a year and documented a 7–mm Hg difference in blood pressure.[40] These trials can provide indirect support for the effect of air pollution exposure on mortality by identifying mechanisms. Despite the important advantage of randomization and controlled exposure, the major limitation of these studies is the ethical issues that arise when exposing individuals to these toxic components.

Animal toxicology studies can also provide mechanistic support for a conclusion of causality. For example, mice on a high-fat diet exposed for 6 months to an average particle concentrations of 15.2 μg/m³ had almost twice the atherosclerotic plaque as animals exposed to filtered air.[41] Another study reported that lung function in mice exposed to outdoor air, with average particle concentrations of 16.8 μg/m³, was lower than in mice exposed to filtered air, with particle concentrations of 2.9 μg/m³.[42]

The use of causal modeling techniques allowed us to conduct a large-scale population-based study while obtaining an estimate for the $PM_{2.5}$ effect that is causal if our assumptions are met. Similar to our study, a previous analysis done by our group Wang and colleagues[23] has estimated the causal effects of long-term exposure to $PM_{2.5}$ on mortality. The authors found 3% increase in all-cause mortality, associated with an IQR increase (2 μg/m³) in annual $PM_{2.5}$ exposure, among mortality cases in New Jersey.[23]

Another study that estimated the causal effect of 2-year exposure to $PM_{2.5}$ on mortality among a representative sample of Medicare beneficiaries found that increases in $PM_{2.5}$ were associated with increases of all-cause and cause-specific hospital admission rates but did not find evidence of an increase in mortality.[18] However, that study had only a one-year follow-up on 32,119 individuals and the number of deaths in that follow-up period was small. The different inferences may be due to the use of individual survival data, the size of the cohort and the number of deaths analyzed, the selected population, and the adjustment for personal confounders.

In addition to the methods and study design, another term required for the determination of a causal effect is biological plausibility. Two possible mechanisms that link $PM_{2.5}$ exposure with mortality are commonly described in the literature. First, the development of pulmonary oxidative stress may lead to respiratory mortality or an indirect effect on the circulatory system. The later may cause cardiovascular events, arrhythmia, and even death as a result of the inflammatory responses, vascular dysfunction, and the accumulation of cytokines and clotting factors. Second, ultrafine particles can penetrate directly into the circulatory system through the pulmonary system and cause local oxidative stress, which may lead to the same cardiovascular and mortality outcomes.[32,43]

We found larger effects among people with eligibility for Medicaid services and blacks. People dual eligibility to Medicaid are often of lower SES.[23,44,45] The larger effects found in these populations may be related to the higher likelihood of non-white and low-income populations to be exposed to higher pollution levels.[44,46] It can also be related to limited financial and physical access to healthcare services.[44,47–49]

The large sample size and geographical area covered in this study are the major strengths of the study. Using a satellite-based model to estimate $PM_{2.5}$ exposure, we were able to include both urban and rural population and to obtain exposure

in high spatial resolution. In addition, if our assumption that all the time-varying confounders were accounted for in the models, holds, we were able to identify the causal association between $PM_{2.5}$ and mortality.

Our study also has limitations. First, we were not able to directly adjust the models to predictors such as comorbidities, smoking or BMI. To the extent that those predictors varied slowly during the follow-up period, or similarly within ZIP codes, they were removed by the DID approach. We believe any remaining variation in such predictors will be uncorrelated with within ZIP code year-to-year variations in $PM_{2.5}$, but that is an untestable hypothesis. An examination of a random sample of the Medicare cohort that did have individual-level information on such covariates found no association between them and $PM_{2.5}$ even before using the DID approach.[12] Trying to minimize confounding by individual factors, we did control for the measured personal and socioeconomic predictors. Second, despite the use of highly spatially resolved exposure models, because we assigned the same annual average to each ZIP code, exposure misclassification may still be present. However, because the potential error is Berksonian, it may increase the confidence intervals but should not bias the estimates.

In conclusion, assuming no predictors changing differentially across ZIP codes over time other than those accounted for in our model, we found a causal effect of $PM_{2.5}$ on the mortality incidence rate. The negative outcome control approach provides farther strength to our findings.

Acknowledgments

The authors declare that they have no conflicts of interest with regard to the content of this report.

References

1. Cohen AJ, Brauer M, Burnett R, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet*. 2017;389:1907–1918.
2. Samoli E, Peng R, Ramsay T, et al. Acute effects of ambient particulate matter on mortality in Europe and North America: results from the APHENA study. *Environ Health Perspect*. 2008;116:1480–1486.
3. Barnett AG, Williams GM, Schwartz J, et al. The effects of air pollution on hospitalizations for cardiovascular disease in elderly people in Australian and New Zealand cities. *Environ Health Perspect*. 2006;114:1018–1023.
4. Brook RD, Rajagopalan S, Pope CA III, et al; American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation*. 2010;121:2331–2378.
5. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*. 2006;295:1127–1134.
6. Kloog I, Coull BA, Zanobetti A, Koutrakis P, Schwartz JD. Acute and chronic effects of particles on hospital admissions in New-England. *PLoS One*. 2012;7:e34664.
7. Kloog I, Ridgway B, Koutrakis P, Coull BA, Schwartz JD. Long- and short-term exposure to PM2.5 and mortality: using novel exposure models. *Epidemiology*. 2013;24:555–561.
8. Pope CA III, Dockery DW. Health effects of fine particulate air pollution: lines that connect. *J Air Waste Manag Assoc*. 2006;56:709–742.
9. Yitshak-Sade M, Kloog I, Novack V. Do air pollution and neighborhood greenness exposures improve the predicted cardiovascular risk? *Environ Int*. 2017;107:147–153.
10. Yitshak Sade M, Novack V, Ifergane G, Horev A, Kloog I. Air pollution and ischemic stroke among young adults. *Stroke*. 2015;46: 3348–3353.
11. Hoek G, Krishnan RM, Beelen R, et al. Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environ Health*. 2013;12:43.
12. Di Q, Wang Y, Zanobetti A, et al. Air pollution and mortality in the medicare population. *N Engl J Med*. 2017;376:2513–2522.
13. Zanobetti A, Schwartz J. Particulate air pollution, progression, and survival after myocardial infarction. *Environ Health Perspect*. 2007;115:769–775.
14. Dominici F, Zigler C. Best practices for gauging evidence of causality in air pollution epidemiology. *Am J Epidemiol*. 2017;186:1303–1309.
15. 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;16:7.
16. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127(8 pt 2):757–763.
17. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology*. 2008;19:766–779.
18. 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*. 2017;28:627–634.
19. Wang Y, Lee M, Liu P, et al. Doubly robust additive hazards models to estimate effects of a continuous exposure on survival. *Epidemiology*. 2017;28:771–779.
20. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci*. 2010;25:1–21.
21. Streeter AJ, Lin NX, Crathorne L, et al. Adjusting for unmeasured confounding in nonrandomized longitudinal studies: a methodological review. *J Clin Epidemiol*. 2017;87:23–34.
22. Donald SG, Lang K. Inference with difference-in-differences and other panel data. *Rev Econ Stat*. 2007;89:221–233.
23. Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. Estimating causal effects of long-term PM2.5 exposure on mortality in New Jersey. *Environ Health Perspect*. 2016;124:1182–1188.
24. Kloog I, Chudnovsky AA, Just AC, et al. A new hybrid spatio-temporal model for estimating daily multi-year PM2.5 concentrations across northeastern USA using high resolution aerosol optical depth data. *Atmos Environ (1994)*. 2014;95:581–590.
25. Kalnay E, Kanamitsu M, Kistler R, et al. The NCEP/NCAR 40-year reanalysis project. *B Am Math Soc*. 1996;77:437–471.
26. Andersen PK, Gill RD. Cox regression model for counting processes a large sample study. *Ann Stat*. 1982;10:1100–1120.
27. Rubin DB. Practical implications of modes of statistical inference for causal effects and the critical role of the assignment mechanism. *Biometrics*. 1991;47:1213–1234.
28. Donohue JJ, Ho eD. The impact of damage caps on malpractice claims: randomization inference with difference-in-differences. *J Empir Leg Stud*. 2007;4:69–102.
29. Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ. Impact Evaluation in Practice. Washington, DC: The world bank; 2011:1–248.
30. Schwartz JD, Wang Y, Kloog I, Yitshak-Sade M, Dominici F, Zanobetti A. Estimating the effects of PM2.5 on life expectancy using causal modeling methods. *Environ Health Perspect*. 2018;126:127002.
31. Sofer T, Richardson DB, Colicino E, Schwartz J, Tchetgen Tchetgen EJ. On negative outcome control of unobserved confounding as a generalization of difference-in-differences. *Stat Sci*. 2016;31:348–361.
32. Brook RD. Cardiovascular effects of air pollution. *Clin Sci (Lond)*. 2008;115:175–187.
33. 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;329:1753–1759.
34. Laden F, Neas LM, Dockery DW, Schwartz J. Association of fine particulate matter from different sources with daily mortality in six U.S. cities. *Environ Health Perspect*. 2000;108:941–947.
35. Pope CA III, Burnett RT, Thun MJ, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*. 2002;287:1132–1141.
36. Wang Y, Shi L, Lee M, et al. Long-term exposure to PM2.5 and mortality among older adults in the southeastern US. *Epidemiology*. 2017;28:207–214.
37. Langrish JP, Bosson J, Unosson J, et al. Cardiovascular effects of particulate air pollution exposure: time course and underlying mechanisms. *J Intern Med*. 2012;272:224–239.
38. 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;357:1075–1082.

39. Devlin RB, Ghio AJ, Kehrl H, Sanders G, Cascio W. Elderly humans exposed to concentrated air pollution particles have decreased heart rate variability. *Eur Respir J Suppl*. 2003;40:76s–80s.
40. Chuang HC, Ho KF, Lin LY, et al. Long-term indoor air conditioner filtration and cardiovascular health: A randomized crossover intervention study. *Environ Int*. 2017;106:91–96.
41. Sun Q, Wang A, Jin X, et al. Long-term air pollution exposure and acceleration of atherosclerosis and vascular inflammation in an animal model. *JAMA*. 2005;294:3003–3010.
42. Mauad T, Rivero DH, de Oliveira RC, et al. Chronic exposure to ambient levels of urban particles affects mouse lung development. *Am J Respir Crit Care Med*. 2008;178:721–728.
43. Shrey K, Suchit A, Deepika D, Shruti K, Vibha R. Air pollutants: the key stages in the pathway towards the development of cardiovascular disorders. *Environ Toxicol Pharmacol*. 2011;31:1–9.
44. Bell ML, Ebisu K. Environmental inequality in exposures to airborne particulate matter components in the United States. *Environ Health Perspect*. 2012;120:1699–1704.
45. Bravo MA, Anthopolos R, Bell ML, Miranda ML. Racial isolation and exposure to airborne particulate matter and ozone in understudied US populations: Environmental justice applications of downscaled numerical model output. *Environ Int*. 2016;92-93:247–255.
46. Martenies SE, Milando CW, Williams GO, Batterman SA. Disease and health inequalities attributable to air pollutant exposure in Detroit, Michigan. *Int J Environ Res Public Health*. 2017;14:24.
47. Makri A, Stilianakis NI. Vulnerability to air pollution health effects. *Int J Hyg Environ Health*. 2008;211:326–336.
48. Forastiere F, Stafoggia M, Tasco C, et al. Socioeconomic status, particulate air pollution, and daily mortality: differential exposure or differential susceptibility. *Am J Ind Med*. 2007;50:208–216.
49. Bell ML, Ebisu K, Peng RD, Dominici F. Adverse health effects of particulate air pollution: modification by air conditioning. *Epidemiology*. 2009;20:682–686.