A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP3130.

Estimating the Effects of $PM_{2.5}$ on Life Expectancy Using Causal Modeling Methods

Joel D. Schwartz,[1] Yan Wang,[1] Itai Kloog,[2] Ma'ayan Yitshak-Sade,[1] Francesca Dominici,[3] and Antonella Zanobetti[1]

[1]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
[2]Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer-Sheva, Israel
[3]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

Background: Many cohort studies have reported associations between $PM_{2.5}$ and the hazard of dying, but few have used formal causal modeling methods, estimated marginal effects, or directly modeled the loss of life expectancy.

Objective: Our goal was to directly estimate the effect of $PM_{2.5}$ on the distribution of life span using causal modeling techniques.

Methods: We derived nonparametric estimates of the distribution of life expectancy as a function of $PM_{2.5}$ using data from 16,965,154 Medicare beneficiaries in the Northeastern and mid-Atlantic region states (129,341,959 person-years of follow-up and 6,334,905 deaths). We fit separate inverse probability-weighted logistic regressions for each year of age to estimate the risk of dying at that age given the average $PM_{2.5}$ concentration at each subject's residence ZIP code in the same year, and we used Monte Carlo simulations to estimate confidence intervals.

Results: The estimated mean age at death for a population with an annual average $PM_{2.5}$ exposure of $12\,\mu g/m^3$ (the 2012 National Ambient Air Quality Standard) was 0.89 y less (95% CI: 0.88, 0.91) than estimated for a counterfactual $PM_{2.5}$ exposure of $7.5\,\mu g/m^3$. In comparison, life expectancy at 65 y of age increased by 0.9 y between 2004 and 2013 in the United States. We estimated that 23.5% of the Medicare population would die before 76 y of age if exposed to $PM_{2.5}$ at $12\,\mu g/m^3$ compared with 20.1% if exposed to an annual average of $7.5\,\mu g/m^3$.

Conclusions: We believe that this is the first study to directly estimate the effect of $PM_{2.5}$ on the distribution of age at death using causal modeling techniques to control for confounding. We find that reducing $PM_{2.5}$ concentrations below the 2012 U.S. annual standard would substantially increase life expectancy in the Medicare population. https://doi.org/10.1289/EHP3130

Introduction

Over 50 cohort studies have reported that long-term exposure to airborne particulate matter (PM) with aerodynamic diameter of $\leq 2.5\,\mu m$ ($PM_{2.5}$) is associated with higher mortality rates (Beelen et al. 2014; Crouse et al. 2015; Kioumourtzoglou et al. 2016; Krewski et al. 2009; Lepeule et al. 2012; GBD 2013 Mortality and Causes of Death Collaborators 2015; Puett et al. 2009; Shi et al. 2016; Wang et al. 2016, 2017; Vodonos et al. 2018). This is supported by a substantial toxicological literature showing that particle exposure produces endothelial dysfunction, atherosclerosis, systemic inflammation, decreased plaque stability, and electrocardiogram abnormalities (Adar et al. 2010; Bräuner et al. 2008; Brook 2008; Brook et al. 2009; Gareus et al. 2008; Hansen et al. 2007; Soares et al. 2009; Sun et al. 2008). Consequently, the 2015 Global Burden of Disease (GBD) study included ambient PM air pollution exposure among the largest worldwide contributors to avoidable early deaths (GBD 2013 Risk Factors Collaborators et al. 2015). Recent studies have reported associations between $PM_{2.5}$ and mortality at concentrations below the 2012 U.S. EPA National Ambient Air Quality Standard (NAAQS) (Di et al. 2017; Wang et al. 2017). However, to date few of these studies have used the approaches of causal modeling and none has directly estimated effects on life expectancies.

Causal modeling methods represent a valuable approach to advance the argument for causality. The general approach is to try to make an observational study closely mimic a randomized trial. In addition, unlike traditional methods, causal modeling methods provide marginal estimates of the effects of $PM_{2.5}$, that is, estimates that do not depend on the distribution of the covariates in the study population. As such, their use in quantitative risk assessments, such as the GBD estimates, is more straightforward. Specifically, the coefficients of a standard Cox regression analysis, when applied to an individual, produce the marginal effect of an increment in exposure, holding all covariates constant, but only for that individual. However, because of the nonlinearity and lack of collapsibility of the proportionate hazard model, the mean of the individual marginal effects is not the population marginal effect (Greenland and Pearl 2011). In contrast, inverse probability–weighted (IPW) approaches do produce population marginal effect estimates (Robins et al. 2000).

The results of most survival analyses, including previous air pollution cohort studies, are presented as hazard ratios (HRs) associated with a given increment of exposure. The concept of an HR is not clear to many policy makers or to the public. Nor does it directly translate into quality-adjusted life years (QUALYs) or disability-adjusted life years (DALYs). No one has ever asked their physician, "If I quit smoking, what will that do to my hazard of dying?" They want to know how it will affect their life span. Providing a more direct estimate of the effect of exposure on life expectancy, therefore, would be valuable. Here we present an IPW survival model to estimate the marginal effect of $PM_{2.5}$ exposure on the distribution of life expectancy in the United States, which, under appropriate conditions, is a causal estimate.

Data and Methods

We obtained the Medicare beneficiary denominator file, which contains information on all Medicare participants in the United States, from the Center for Medicare and Medicaid Services (ResDAC 2018). We constructed an open cohort using all beneficiaries ≥ 65 y of age in the Northeastern and mid-Atlantic region states (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Delaware, Pennsylvania, Maryland, Washington, DC, Virginia, and West Virginia) from 2000 to 2013, and examined survival of those beneficiaries as our

Address correspondence to J.D. Schwartz, 401 Park Dr., Landmark Center 4th Floor West, Boston, MA 02215 USA. Telephone: (617) 384-8752. Email: joel@hsph.harvard.edu

J.D.S. is a consultant to the U.S. Department of Justice in cases involving Clean Air Act violations. There are no other competing financial interests. The other authors declare they have no actual or potential competing financial interests.

Received 20 November 2017; Revised 13 November 2018; Accepted 15 November 2018; Published 5 December 2018.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

outcome. Medicare insurance covers over 95% of the population ≥65 y of age in the United States. Medicare participants alive on 1 January of the year following their enrollment in Medicare were entered into the open cohort for survival, and follow-up periods were calendar years. Persons alive and enrolled prior to 2000 were entered into the cohort in 2000. The data set included 16,965,154 Medicare participants. Among these, we had 129,341,959 person-years of follow-up and 6,334,905 deaths.

Exposure Data

We used a previously published exposure model to estimate annual average concentrations of $PM_{2.5}$ at each ZIP code in the Northeast (Kloog et al. 2014). This model has previously been used in multiple epidemiological studies (Chiu et al. 2014; Fleisch et al. 2014; Kloog et al. 2015). It is a hybrid model that integrates land use, meteorological, and satellite remote sensing data. Aerosol optical depth (AOD) is an optical measurement of the extinction of light by particles in the air. We used the 1×1 km gridded AOD data that are available daily from the NASA Aqua and Terra satellites, which are processed using the Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm (Lyapustin et al. 2011). In brief, we calibrated the AOD–$PM_{2.5}$ relationship on each day during the study period using data from grid cells with both ground $PM_{2.5}$ concentrations from the U.S. EPA or Interagency Monitoring of Protected Visual Environments (IMPROVE) networks and AOD measurements (FED 2018), as well as land use and meteorological variables, and predicted $PM_{2.5}$ on the remaining grid cells. We used inverse probability weighting to address selection bias due to nonrandomly missing AOD measurements due to snow and clouds or other factors. Finally, we filled in grid cell–days without AOD by regressing the nonmissing measurements for that grid cell against nearby monitors and local land use and meteorological variables. $PM_{2.5}$ predictions were validated using 10-fold cross-validation. We found a high out-of-sample R^2 of 0.89 (Kloog et al. 2014). We used these 1×1 km predictions to compute the annual average $PM_{2.5}$ in each ZIP code by averaging the predictions within ZIP code.

Covariates. From the Medicare denominator file for each calendar year, we obtained the age, sex, race, ZIP code of residence for that year, and date of death (or censoring) of each participant. Age and ZIP code were updated annually. Race and sex were self reported at enrollment. We also obtained annual information about their eligibility for Medicaid, which provides additional coverage for low-income persons. This file is publicly available from the Centers for Medicare and Medicaid Services (ResDAC 2018).

We obtained small area–level social, economic, and housing characteristic variables from the U.S. Census Bureau 2000 Census Summary File 3 (U.S. Census Bureau 2010) at the ZIP code tabulation–area level (ZCTA). A ZCTA converts a ZIP code, which is a set of line segments for postal delivery into an area measure for computing census variables. The variables used were ZCTA percentage of the population that was black, Hispanic, ≥65 y of age living in poverty, living in owner-occupied housing, and with less than a high school education as well as median household income, median value of owner-occupied housing, and population density; all of these variables were updated each year. Annual updates were obtained by linearly extrapolating between the census years. To capture long-term smoking history of Medicare participants in each ZIP code, we used the Medicare data to compute their hospitalization rate for lung cancer by ZIP code for each year. In addition, the county-level percentage of people who ever smoked and their mean body mass index (BMI) scores were obtained from the CDC Behavioral Risk Factor Surveillance survey (CDC 2013), which were then assigned to each ZCTA within the county and updated each year. From the Dartmouth Health Atlas, we obtained percentage of Medicare participants who had a hemoglobin A1c test, a low-density lipoprotein cholesterol (LDL-C) test, a mammogram, and a visit to a primary care physician for each year in each hospital catchment area and assigned it to all ZCTAs in that area (Wennberg and Cooper 1996). Dartmouth catchment areas are nonoverlapping and, in denser populated areas, can include multiple hospitals. As covariates, we used individual age, sex, race (white, black, Asian, other), and Medicaid eligibility as well as ZCTA-level percentage of the population that was black, Hispanic, ≥65 y of age living in poverty, living in owner-occupied housing, with less than a high school education, with an annual physician visit, and with tests for HbA1c, LDL-C, mammography as well as median household income, median value of owner-occupied housing, population density, lung cancer rate, smoking rate. Race and sex were time-invariant covariates; all other covariates were updated each year.

One advantage of the Medicare cohort is that the dropout rate is lower than in most cohorts. People only lose eligibility by dying, and other than administrative censoring in 2013, the only loss to follow-up occurred when people moved to a different region of the country. Of the 16,965,154 participants 469,996 (2.77%) moved out of the region during follow-up, and we assumed they were missing at random, conditional on their exposure and covariates. An additional 1,881,578 participants moved within the region and were assigned exposure and covariates for their new addresses after the move.

Statistical Methods

Survival data with time-varying covariates are typically analyzed using the Andersen-Gill formulation of Cox's proportional hazard model with one observation per person per follow-up period and estimated hazard rates, not survival times. Approaches that directly model the failure time, such as accelerated failure-time models, generally require an assumption about the distribution of the failure times. A nonparametric approach to estimating the distribution is particularly advantageous in cohort studies where there is left censoring at age of entry and a skewed distribution of life expectancy.

If a separate logistic regression is fit for death at each year of age, we can obtain an estimate of the probability of failing at t y of age, conditional on the covariates, and on the person having survived until t y of age. These estimates make no parametric assumption about the distribution of the survival times. Further, they allow for the effect of both the exposure of interest and of all of the covariates to differ by year of age.

Causal modeling seeks to make the analysis of observational data mimic a randomized trial as closely as possible. In a randomized trial, the randomization assures that (at the time of randomization) the exposure of interest is independent of the covariates. Propensity score methods seek to recover that property by making the distribution of the exposure independent of the covariates. For a continuous exposure, the generalized propensity score fits a linear regression of that exposure against the measured covariates (Imai and van Dyk 2004). The probability density of the residual for each observation is the probability density of the subject receiving their observed exposure level in that year given their covariates in that year. This can be used for propensity-score matching or for IPW analyses. In IPW, this density is the denominator of the IPW, and the numerator is the marginal probability density of exposure (Lunceford and Davidian 2004). If the logistic regression for surviving each year of age is weighted by the IPW and the linear regression used to derive the weights was correctly specified (i.e., it included the necessary interactions and accounted for nonlinearity), the exposure should be independent of the covariates in the weighted sample. This can be confirmed by computing the standardized mean difference in each covariate for observations above

versus below the mean level of exposure, and values below a level of 0.1–0.2 are generally taken as indications of balance. If the standardized mean difference for one or more covariates is larger than that threshold, then the linear regression can be modified by including additional interactions or splines until the weights are appropriately balanced vs. If all relevant confounders are included and positivity is met (i.e., each person had a nonzero probability of receiving any exposure), then analysis of the weighted sample will produce a causal marginal estimate of the effect of exposure.

For example, if obese people tend to have higher exposure, we can render exposure independent of obesity by giving more weight to observations from obese people with lower exposure. Because the covariates no longer have to be controlled for in the regression for the outcome (being independent of exposure in the weighted sample), the effect estimate is not conditional on the covariates but is, instead, a marginal estimate.

We fit a separate logistic regression to predict death at each year of age, using the corresponding annual average $PM_{2.5}$ concentration at each subject's residential ZIP code as the exposure and age-specific IPW weights to allow the influence of confounders to change with age. We used robust variance estimates from the *sandwich* package in R (Version 2.3; R Development Core Team) to estimate the confidence intervals (CIs) to account for correlations in the errors that could be induced by either spatial correlation in the residuals or the IP weights.

A key advantage of propensity score models is the ability to check whether all of the covariates are independent of exposure in the weighted sample. We first fit a model with linear terms for all covariates and checked the balance. Based on those results, we added interaction terms and splines for covariates that were not balanced and iterated until good balance was achieved. The final model included interaction terms of black race with Medicaid eligibility, ZIP code percentage black with population density, and male with population density and Medicaid eligibility. Natural cubic splines with 3 df were used with the percentage of ZIP code who was black, median household income, percentage with annual physician visit, percentage below poverty level, population density, percentage living in owner-occupied housing, median value of housing, smoking rate, percentage with mammogram screening, and percentage with hemoglobin A1c screening.

This approach estimated the condition probability of dying at t y of age given exposure x among people that survived to t y of age. We can compute that survival probability (to t y of age) as the cumulative product of $(1-p_i)$, for i up to t y of age, where p_i is the probability of dying at i y of age. Multiplying the conditional probability p_t by the survival probability to t y of age gives the unconditional probability of dying at t y of age. Given the unconditional probabilities of dying at each age at a given exposure, we can compute the mean age at death given that exposure, and the probability of living to only ≤ 75 y of age, or of living past 85 y of age. We estimated these quantities for two counterfactual levels of $PM_{2.5}$ exposure: the 2012 U.S. National Ambient Air Quality Standard (https://www.epa.gov/pm-pollution/2012-national-ambient-air-quality-standards-naaqs-particulate-matter-pm) of 12 $\mu g/m^3$ and an alternative annual average $PM_{2.5}$ concentration of 7.5 $\mu g/m^3$. We estimated CIs for these quantities using 10,000 Monte Carlo simulations from the asymptotic distribution of the predicted probabilities estimated for each age.

Although the goal of this paper was to directly estimate effects on the distribution of life span, which is inherently different from previous studies that estimated hazard rates, it is useful to compare results across studies. To produce an estimate akin to a hazard rate, we computed a weighted average of our age-specific coefficients for $PM_{2.5}$, weighting each coefficient proportional to the number of deaths in that year of age and inversely by the variance of the coefficient, and report this as well.

This study was reviewed and granted an exemption as use of previously collected administrative data by the institutional review board of the Harvard School of Public Health.

Results

Table 1 shows the characteristics of the study population. Because this is an open cohort, the variables change each year and the averages are across all years. The mean age was 75.5 y, 11% were receiving Medicaid, and the population was predominantly white. The mean $PM_{2.5}$ exposure was 10.3 $\mu g/m^3$, the 25th and 97.5th percentiles of exposure were 7.33 $\mu g/m^3$ and 12.94 $\mu g/m^3$, respectively, and 70% of the person-years had exposures of <12 $\mu g/m^3$. The population inhabited 7,600 ZIP codes, with a mean number of participants of 2,232. There were 430 hospital service areas with an average of 39,454 participants each, and 398 counties with an average of 42,626 participants.

Figure 1A shows the weighted standardized mean difference in covariates above versus below the mean $PM_{2.5}$ concentration to evaluate the balance achieved by the propensity score. In general, the standardized mean differences are near zero, except for males. For example, the standardized difference between high and low $PM_{2.5}$ for lung cancer rate was 0.0005, and for percentage of ZIP code below poverty level was -0.0356. In contrast, for males, it was 0.3255, indicating that subjects with $PM_{2.5}$ exposure <10.3 $\mu g/m^3$ were more likely to be male. Because of this, and because of the significant difference in life expectancy between men and women, we repeated our analysis separately for

Table 1. Distribution of variables in all Medicare beneficiaries ≥ 65 y of age residing in the states of Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Delaware, Pennsylvania, Maryland, Washington, DC, Virginia, and West Virginia who were enrolled during 2000–2013 ($n = 16{,}965{,}154$; person-years $= 129{,}341{,}959$).

	Percentile			
Variable	25th	50th	75th	Mean
Age (y)	69.00	74.00	81.00	75.5
Male (%)	41.0	44.2	47.6	44.8
Medicaid coverage (%)	4.8	8.2	13.7	11.0
Black	0	0.6	4.4	6.9
Asian	0	0.3	0.9	0.9
Other race	0	0.3	1.0	1.5
Area-based covariates				
Median income (USD)[a]	38,000	48,600	66,900	55,700
Percentage below poverty level[a]	4.8	7.7	12.0	10.6
Population density (people–km^{-2})[a]	26	114	752	1453
Median housing value (USD)[a]	95,900	160,600	263,800	208,000
Percentage owner occupied[a]	68	78	84	72
Percentage <high school[a]	20	28	38	30
BMI (kgm^{-2})[b]	26.9	27.5	28.2	27.6
HbA1c (% screened)[c]	81.3	84.0	86.3	83.6
LDL-C (% screened)[c]	77.7	80.7	82.9	80.2
Mammography (%)[c]	62.9	65.6	70.0	65.2
Lung cancer rate ($\times 10^{-5}$)[d]	27.6	39.2	53.2	48.0
Ever Smoker (%)[b]	45.4	48.6	51.3	48.2
Percentage ZIP code black[e]	0.3	1.6	7.6	8.0
Percentage ZIP code Hispanic[e]	0.9	1.8	4.9	5.2
Percentage annual checkup[c]	74.9	78.3	81.6	77.5
$PM_{2.5}$ ($\mu g/m^3$)	9.2	10.4	11.4	10.3

Note: Percentiles of the distribution of the variables are over all observations.
[a]ZIP code data (U.S. Census Bureau 2010).
[b]County level (CDC 2013).
[c]Dartmouth Health Atlas Hospital Catchment Area data.
[d]Rate of hospital admissions of all Medicare enrollees in the ZIP code, computed from CMS data (ResDAC 2018).
[e]Percentage of all inhabitants in ZIP code, not just Medicare cohort participants (U.S. Census Bureau 2010).

each sex. Figure 1B,C shows the balance achieved for women and men separately.

In the full population, the distribution of the probability of death according to age is shifted to the left for the 12-$\mu g/m^3$ $PM_{2.5}$ exposure scenario relative to the 7.5-$\mu g/m^3$ scenario, resulting in a higher probability of dying at a younger age (Figure 2A), and an estimated difference in average life expectancy between the two scenarios of 0.89 y (95% CI: 088, 0.91) (Table 2). The higher exposure scenario resulted in an additional 3.4% of the population dying before 76 y of age (95% CI: 3.5%, 3.3%) and 3.7% (95% CI: 3.6%, 3.8%) fewer people living past 85 y of age.

The shift was more pronounced for men than for women (Figure 2B,C), resulting in a larger difference in average life expectancy between the higher- and lower-exposure scenarios for men [1.17 y (95% CI: 1.14, 1.19)] than for women [0.74 y (95% CI: 0.72, 0.77)] and a larger difference in the probability of death at $\leq$75 y of age under the higher-versus lower-exposure scenario for men [4.7% higher (95% CI: 4.8, 4.6)] than for women [2.6% higher (95% CI: 2.7, 2.5)] (Table 2). The 0.40-y difference (95% CI: 0.44, 0.36) in the effect of $PM_{2.5}$ on life expectancy between men and women was significant.

The weighted average of the age-specific coefficients for death with a 1-$\mu g/m^3$ increase in $PM_{2.5}$ was 0.0225, which is

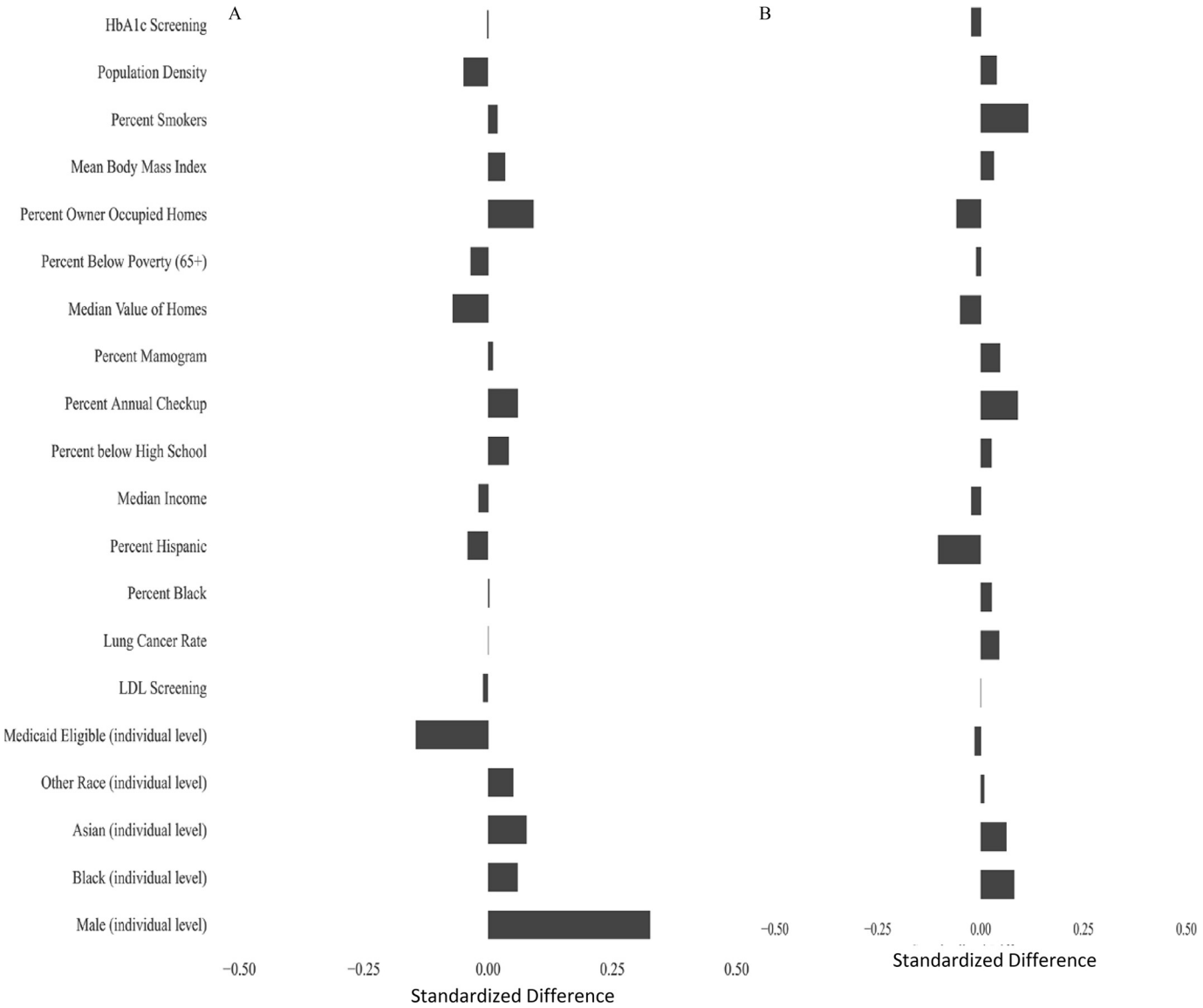

Figure 1. Standardized mean differences in covariates between observations above and below the mean annual $PM_{2.5}$ concentrations of 10.4 $\mu g/m^3$ after weighting using the propensity score. (*A*) Standardized differences in the entire cohort; (*B*) standardized differences in women; and (*C*) standardized differences in men. The propensity score was fit using the following individual covariates male, black, Asian, other race, Medicaid eligible and the following area-based variables percentage of people >65 y of age who had screening for low-density lipoprotein cholesterol that year, percentage of women >65 y of age who had a mammogram that year, percentage of people >65 y of age who had hemoglobin A1c measured that year, percentage of people >65 y of age who had an annual checkup that year, all by hospital catchment area; lung cancer hospitalization rate in the Medicare population, percentage of population that is black, percentage of population that is Hispanic, median household income, median value of owner-occupied housing, percentage of housing occupied by owner, percentage of persons >65 y of age with less than a high school education, and population density, all by ZIP code; and mean body mass index in the county and smoking rate in the county. In addition, interaction terms were included for percentage black × population density, Medicaid eligibility × population density, and male sex ×population density. Nonlinear terms were used for percentage of the population that was black, median household income, percentage with less than a high school education, percentage with an annual checkup, median value of housing, percentage below poverty level, population density, percentage owner-occupied housing, percentage with HbA1c screening, percentage of women with mammograms, and percentage of smokers.

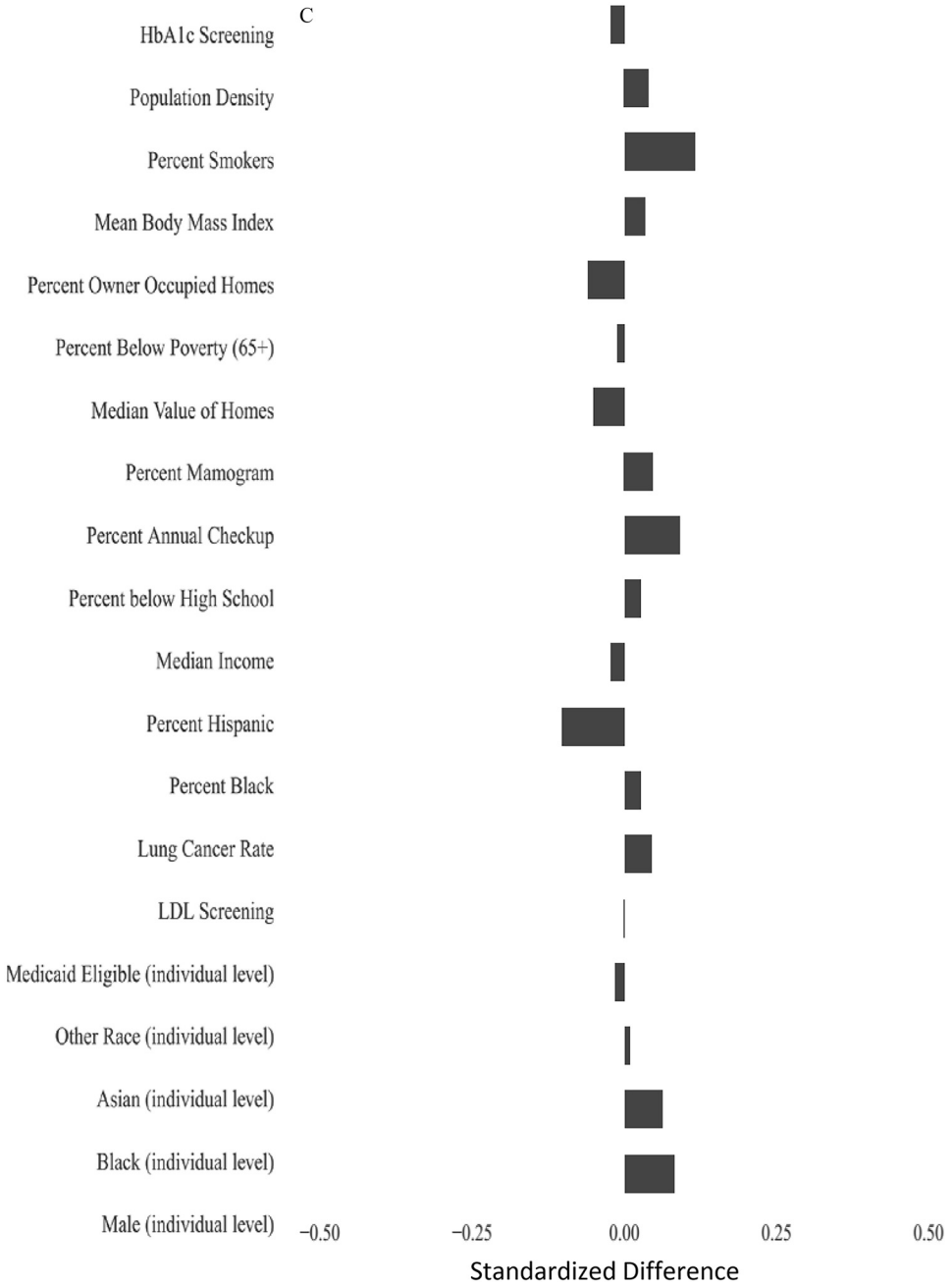

Figure 1. (*Continued.*)

consistent with an HR of 1.25 for a 10-$\mu g/m^3$ increment. This estimate is higher than those reported for earlier studies done at higher $PM_{2.5}$ concentrations [e.g., the HR of 1.14 reported by Laden et al. (2006)] but is consistent with the higher HRs reported by recent studies at lower exposures such as ours (e.g., the HR of 1.26 reported by Pinault et al. (2016)].

Discussion

We believe that this is the first study to directly estimate the effect of $PM_{2.5}$ on the distribution of age at death modeling that distribution nonparametrically and using causal modeling techniques to control for confounding. We found a highly significant difference in the mean age of death [0.89 y (95% CI: 088, 0.91)] for a 4.5-$\mu g/m^3$ difference in exposure, going from an exposure that met the 2012 ambient standard of 12 $\mu g/m^3$ to 7.5 $\mu g/m^3$. If our propensity score model, which allows the effect of confounders to vary by year of age, is valid, this is a causal estimate. Moreover, our methods allowed us to estimate how $PM_{2.5}$ exposure influences the full distribution of life expectancy in addition to changing the mean and allows for a different effect of $PM_{2.5}$ exposure at each age. We estimated that 23.5% of the Medicare population would die before 76 y of age if they were exposed to the 2012 ambient standard $PM_{2.5}$ concentration of 12 $\mu g/m^3$ compared with 20.1% if the Medicare population was exposed to an annual average of 7.5 $\mu g/m^3$. In a population of 16,965,154, this translates to over half a million extra Medicare participants dying before 76 y of age. Similarly, we estimated that 40.8% of Medicare recipients would live past 85 y of age if exposed to 12-$\mu g/m^3$ $PM_{2.5}$ compared with 44.5% at 7.5-$\mu g/m^3$ $PM_{2.5}$.

To put these results in perspective, the National Center for Health Statistics reported that, between 2004 and 2013, the U.S.

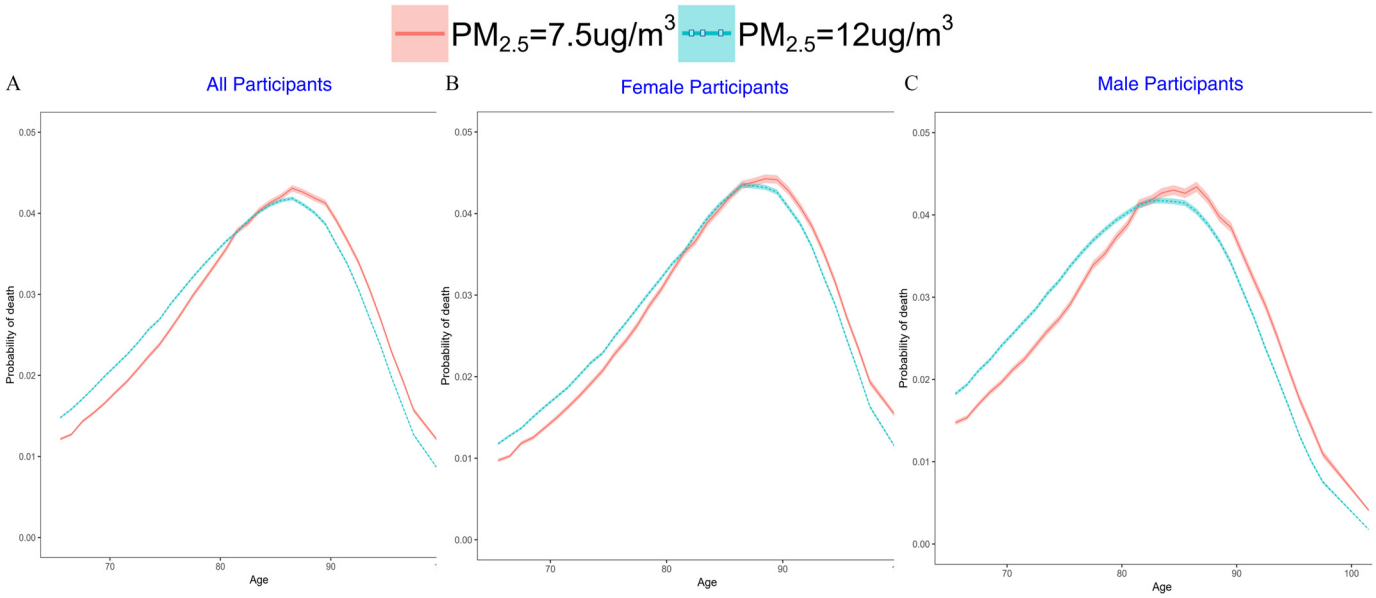

Figure 2. Nonparametric estimate of the probability of death according to age (in years) under two different counterfactual annual average $PM_{2.5}$ concentrations just meeting the 2012 U.S. EPA standard of 12. $\mu g/m^3$ vs. 7.5. $\mu g/m^3$ for (*A*) the entire cohort, (*B*) women, and (*C*) men. Estimates were generated using separate logistic regressions for each year of age. The shaded area displays the 95% confidence intervals about the curves.

Table 2. Effect size estimates (mean ± confidence interval) for the difference in life expectancy and of percentage of the population dying before 76 y of age in Medicare beneficiaries ≥65 y of age residing in the northeast and mid-Atlantic states, at two counterfactual levels of $PM_{2.5}$ concentration: 12 $\mu g/m^3$ and 7.5 $\mu g/m^3$.

Exposure category	Mean age at death (y)	Percentage dying at <76 y of age
12 $\mu g/m^3$	82.51 (82.50, 82.53)	23.5 (23.4, 23.6)
7.5 $\mu g/m^3$	83.41 (83.39, 83.42)	20.1 (20.0, 20.2)
Difference	0.89 (0.88 0.91)	3.4 (3.5, 3.3)
Females		
12 $\mu g/m^3$	83.64 (83.63, 83.65)	19.5 (19.4, 19.6)
7.5 $\mu g/m^3$	84.38 (84.37, 84.40)	16.9 (16.8, 17.0)
Difference	0.74 (0.72, 0.77)	2.59 (2.70, 2.49)
Males		
12 $\mu g/m^3$	81.14 (81.13, 81.16)	28.3 (28.2, 28.3)
7.5 $\mu g/m^3$	82.31 (82.29, 82.33)	23.5 (23.4, 23.6)
Difference	1.17 (1.14, 1.19)	4.7 (4.8, 4.6)

life expectancy at 65 y of age increased by 0.9 y (National Center for Health Statistics 2017). During that decade, the average $PM_{2.5}$ concentration in our cohort decreased from 11.74 to 8.79 $\mu g/m^3$, or by 2.95 $\mu g/m^3$. Assuming proportionality, this decrease may have accounted for 0.58 of the 0.90-y increase in life expectancy, a substantial fraction. Is this plausible? We believe it is. For decades, cardiovascular mortality rates have been falling, principally because of the decrease in smoking rates, the increased use of statins and hypertensive medication, and the introduction of emergency catheterization labs to interrupt myocardial infarctions, among other factors. However, these changes had mostly occurred by 2004, whereas the increase in life in the next decade was unabated or reduced in magnitude. This suggests other causes played an increasing role, of which air pollution could be one. Further study is clearly warranted to see whether air pollution improvements accounted for two-thirds of the change in that decade and what the effect is elsewhere.

This paper adds to an existing literature of many cohort studies that have reported associations between $PM_{2.5}$ exposure and the hazard rate for death (Vodonos et al. 2018). Because it estimates changes in mean life expectancy and in the proportion of the population who die early (≤75 y of age) it cannot be compared numerically to those results. However, as noted above, the results are broadly consistent with newer studies at lower $PM_{2.5}$ concentrations (as this one). The Vodonos meta-analysis confirmed that across 53 cohorts, the hazard rate per 1 $\mu g/m^3$ was higher at lower $PM_{2.5}$ concentrations. Together, these studies provide strong evidence for the causality of the association because they were conducted by multiple investigators in multiple countries, with a wide range of covariate control strategies. This study adds to the evidence by using a causal modeling approach to control confounding. Using inverse probability weights that change for each person and year of follow-up generates causal estimates if all relevant confounders are included and appropriately specified in the propensity score analysis.

To be a confounder, a variable must be a predictor of the exposure as well as the outcome. A key issue with this study is the lack of information on many individual covariates, such as smoking and BMI. A recent analysis of a random sample of the Medicare population showed that neither of these predictors of mortality were associated with $PM_{2.5}$ exposure at the ZIP code level and, hence, are not confounders (Di et al. 2017). In addition, 86% of the beneficiaries were nonsmokers.

However, most individual predictors of health are unlikely to be predictors of $PM_{2.5}$. For example, individual smoking causes a trivial impact on ambient $PM_{2.5}$. Why then is smoking a possible confounder? It is because neighborhoods with more smokers might also be neighborhoods with higher ambient $PM_{2.5}$. In that case, imagine I moved a nonsmoker from a neighborhood with few smokers to a neighborhood with many. By hypothesis, her $PM_{2.5}$ exposure is likely to be higher in the new neighborhood. However, this has nothing to do with her individual smoking status. The hypothesized confounding is by neighborhood smoking status, and controlling for neighborhood smoking is the appropriate control. The same holds for many other individual predictors of health, such as cholesterol levels and BMI, which are not causes of $PM_{2.5}$. They can only be associated with neighborhood-level $PM_{2.5}$ because the same (e.g., socioeconomic) factors that cause people with high or low levels of the individual predictors to cluster in neighborhoods may also be predictors of exposure. Hence, control

of neighborhood-level covariates, and particularly neighborhood-level socioeconomic variables, is the key to confounding control in this scenario, not control for individual risk factors. Our study controlled for a large number of neighborhood-level socioeconomic variables and other confounders and updated them annually, which is a strength of the study. The issue of neighborhood versus personal confounding and the advantage of neighborhood exposure in this regard has also been discussed elsewhere (Weisskopf and Webster 2017).

Our estimates are larger than estimates from two ecological studies of associations between changes in life expectancy and air pollution over time (Correia et al. 2013; Pope et al. 2009). This may be explained in part by better exposure classification (ZIP code vs. county level, and accounting for residential mobility) and better control of confounding, including individual-level and ZIP code–level factors, in the present study.

Supporting this view, the American Cancer Society (ACS) study and Women's Health Initiative cohorts both found stronger associations between $PM_{2.5}$ and mortality for within-versus between-county contrasts (Jerrett et al. 2005; Miller et al. 2007), and an updated analysis of the ACS study found that controlling for small-area socioeconomic variables increased effect size estimates for $PM_{2.5}$ and mortality relative to estimates that did not control for them (Krewski et al. 2009; Miller et al. 2007). Consequently, we believe the larger effect size we report reflects less exposure error and better control for confounding.

Controlled human exposure studies and randomized trials support causal effects of air pollution on health. For example, a randomized trial of air filtration versus sham filtration for 48 h in the dormitory rooms of 35 college students reported that air filtration was associated with increased Long Interspersed Nuclear Element-1 (LINE-1) methylation and with methylation of genes involved in inflammation, coagulation, and vasoconstriction (Chen et al. 2016). A similar randomized trial of 55 college students reported that levels of cortisol, cortisone, epinephrine, norepinephrine, glucose, membrane eicosanoids, 8-hydroxy-2-deoxyguansine, malondialdehyde, iso-prostaglandin F2α, and superoxide dismutase as well as systolic blood pressure and insulin resistance were lower after 9 d of air filtration compared with sham filtration (Li et al. 2017). A study of 50 healthy adults reported a reduction in nitroglycerin-induced vasodilation, increased sympathetic tone, and decreased parasympathetic tone following 5 h of exposure to air from a busy street ($PM_{2.5}$ at $24\,\mu g/m^3$) versus exposure to filtered air ($PM_{2.5}$ at $3\,\mu g/m^3$) (Hemmingsen et al. 2015). A study of 15 healthy adults who wore a continuous blood pressure monitor while walking in central Beijing reported that systolic blood pressure was lower when participants wore a particle-filtering mask during their walk (Langrish et al. 2009). In longer-term exposure studies, Chuang et al. (2017) randomized 200 participants to a particle filter versus a sham filter for a year; the sham filter resulted in a 7.8-mmHg increase in systolic blood pressure among the participants.

Animal studies also support an effect of $PM_{2.5}$ on mortality and are discussed in detail in the U.S. EPA Integrated Science Assessment for Particulate Matter (U.S. EPA 2018). To highlight a few relevant studies, mice exposed to ambient air ($16.8\,\mu g/m^3$) had lower lung function than those exposed to filtered air ($2.9\,\mu g/m^3$) (Mauad et al. 2008). Other studies of long-term exposure in mice reported increased atherosclerotic plaque, increased macrophage counts and tissue factor in plaques, increased vasoconstriction, and increased oxidation of LDL-C (Sun et al. 2008; Soares et al. 2009).

This study has limitations. First, we assumed that the propensity score model was well specified. We checked the balance for each covariate, which provides reasonable assurance on that point. We also assumed there were no important omitted cofounders. This assumption is common to all epidemiology studies whether they use causal modeling techniques or not. However, we clearly have only a limited number of individual-level covariates. As noted above, there can be few, if any, individual variables that are predictors of both death and $PM_{2.5}$. Obesity does not produce particles. Rather, to confound the association, those predictors must be correlated with exposure because both have a common antecedent, and that occurs on an area level. That confounding is because both risk factors and exposure cluster by area. Further, the antecedents that produced this area-level confounding are likely to be race and socioeconomic status. In that case, controlling for the antecedents blocks the confounding. In our study, we controlled for individual-level race and poverty and small area–level race/ethnicity as well as many measures of socioeconomic status such as percentage of elderly living in poverty, median income, median house value, education, BMI, and smoking. Given the extent of area-level antecedents we controlled for, we believe the potential for confounding to be limited.

Another limitation of this paper is that we assumed a linear association between $PM_{2.5}$ and the risk of dying at each age. However, the 2.5–97.5% range of exposure is from 7.3 to 12.9 $\mu g/m^3$, and we believe any reasonable concentration–response function is essentially linear within that range.

Another limitation is our use of a single year of $PM_{2.5}$ exposure. In the Nurses' Health Study, a moving average of different numbers of months was used, and no further information was gained by using monthly averages of longer than 48 months (Puett et al. 2009). Other studies have used annual averages and gotten results similar to those of studies that used longer exposure periods (Di et al. 2017). Using longer exposure periods would have resulted in dropping years of follow-up from our study.

Finally, causality is a conclusion of humans, not the output of a statistical model. Our model produced causal estimates if certain conditions are met, and it is not possible to verify that those conditions were truly met. That is why judgments of causality must consider consistency with other studies and experimental evidence.

Conclusions

We found that exposure of the Medicare population to $PM_{2.5}$ at the 2012 National Ambient Standard is associated with a substantial reduction in life expectancy, a substantial increase in the proportion who die at $\leq$75 y of age, and a substantial decrease in the proportion who live past 85 y of age. These results are consistent with HRs estimated by previous observational studies and with a wide range of experimental studies that support causality. The means for reducing ambient $PM_{2.5}$ concentrations [including scrubbers and oxides of nitrogen (NO_x) controls on electric generating facilities, NO_x controls on gasoline engines, NO_x and particle controls on Diesel engines] are known and available; hence, significant public health gains are within reach.

Acknowledgments

This publication was made possible by U.S. Environmental Protection Agency (EPA) grant RD-8358720. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the U.S. EPA. Further, the U.S. EPA does not endorse the purchase of any commercial products or services mentioned in the publication. This publication was also made possible by National Institutes of Health (NIH)/National Institute of Environmental Health Sciences (NIEHS) grants ES-000002, R01 ES024332-01, and R21 ES024012, and by Health Effects Institute grant 4953.

References

Adar SD, Klein R, Klein BE, Szpiro AA, Cotch MF, Wong TY, et al. 2010. Air pollution and the microvasculature: a cross-sectional assessment of in vivo retinal images in the population-based Multi-Ethnic Study of Atherosclerosis (MESA). PLoS Med 7(11):e1000372, PMID: 21152417, https://doi.org/10.1371/journal.pmed.1000372.

Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. 2014. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. Lancet 383(9919):785–795, PMID: 24332274, https://doi.org/10.1016/S0140-6736(13)62158-3.

Bräuner EV, Forchhammer L, Møller P, Barregard L, Gunnarsen L, Afshari A, et al. 2008. Indoor particles affect vascular function in the aged: an air filtration-based intervention study. Am J Respir Crit Care Med 177(4):419–425, PMID: 17932377, https://doi.org/10.1164/rccm.200704-632OC.

Brook RD. 2008. Cardiovascular effects of air pollution. Clin Sci (Lond) 115(6):175–187, PMID: 18691154, https://doi.org/10.1042/CS20070444.

Brook RD, Urch B, Dvonch JT, Bard RL, Speck M, Keeler G, et al. 2009. Insights into the mechanisms and mediators of the effects of air pollution exposure on blood pressure and vascular function in healthy humans. Hypertension 54(3):659–667, PMID: 19620518, https://doi.org/10.1161/HYPERTENSIONAHA.109.130237.

CDC (Centers for Disease Control and Prevention). 2013. Behavioral Risk Factor Surveillance System. BRFSS 2013 Survey Data and Documentation. https://www.cdc.gov/brfss/annual_data/annual_2013.html [accessed 28 November 2018].

Chen R, Meng X, Zhao A, Wang C, Yang C, Li H, et al. 2016. DNA hypomethylation and its mediation in the effects of fine particulate air pollution on cardiovascular biomarkers: a randomized crossover trial. Environ Int 94: 614–619, PMID: 27397927, https://doi.org/10.1016/j.envint.2016.06.026.

Chiu YH, Coull BA, Sternthal MJ, Kloog I, Schwartz J, Cohen S, et al. 2014. Effects of prenatal community violence and ambient air pollution on childhood wheeze in an urban population. J Allergy Clin Immunol 133(3):713–722.e4, PMID: 24200349, https://doi.org/10.1016/j.jaci.2013.09.023.

Chuang H-C, Ho K-F, Lin L-Y, Chang T-Y, Hong G-B, Ma C-M, et al. 2017. Long-term indoor air conditioner filtration and cardiovascular health: a randomized cross-over intervention study. Environ Int 106:91–96, PMID: 28624750, https://doi.org/10.1016/j.envint.2017.06.008.

Correia AW, Pope CA III, Dockery DW, Wang Y, Ezzati M, Dominici F. 2013. Effect of air pollution control on life expectancy in the United States: an analysis of 545 U.S. counties for the period from 2000 to 2007. Epidemiology 24(1):23–31, PMID: 23211349, https://doi.org/10.1097/EDE.0b013e3182770237.

Crouse DL, Peters PA, Hystad P, Brook JR, van Donkelaar A, Martin RV, et al. 2015. Ambient $PM_{2.5}$, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect 123(11):1180–1186, PMID: 26528712, https://doi.org/10.1289/ehp.1409276.

Di Q, Dominici F, Schwartz JD. 2017. Air pollution and mortality in the Medicare population. N Engl J Med 377:1498–1499, PMID: 29020593, https://doi.org/10.1056/NEJMc1709849.

FED. 2018. Database Query Wizard. http://views.cira.colostate.edu/fed/QueryWizard/Default.aspx [accessed 28 November 2018].

Fleisch AF, Gold DR, Rifas-Shiman SL, Koutrakis P, Schwartz JD, Kloog I, et al. 2014. Air pollution exposure and abnormal glucose tolerance during pregnancy: the project Viva cohort. Environ Health Perspect 122(4):378–383, PMID: 24508979, https://doi.org/10.1289/ehp.1307065.

Gareus R, Kotsaki E, Xanthoulea S, van der Made I, Gijbels MJ, Kardakaris R, et al. 2008. Endothelial cell-specific NF-κB inhibition protects mice from atherosclerosis. Cell Metab 8(5):372–383, PMID: 19046569, https://doi.org/10.1016/j.cmet.2008.08.016.

GBD 2013 Mortality and Causes of Death Collaborators. 2015. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 385(9963):117–171, PMID: 25530442, https://doi.org/10.1016/S0140-6736(14)61682-2.

GBD 2013 Risk Factors Collaborators, Forouzanfar MH, Alexander L, Anderson HR, Bachman VF, Biryukov S, et al. 2015. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 386(10010):2287–2323, PMID: 26364544, https://doi.org/10.1016/S0140-6736(15)00128-2.

Greenland S, Pearl J. 2011. Adjustments and their consequences—collapsibility analysis using graphical models. Int Stat Rev 79(3):401–426, https://doi.org/10.1111/j.1751-5823.2011.00158.x.

Hansen CS, Sheykhzade M, Møller P, Folkmann JK, Amtorp O, Jonassen T, et al. 2007. Diesel exhaust particles induce endothelial dysfunction in apoE$^{-/-}$ mice. Toxicol Appl Pharmacol 219(1):24–32, PMID: 17234226, https://doi.org/10.1016/j.taap.2006.10.032.

Hemmingsen JG, Rissler J, Lykkesfeldt J, Sallsten G, Kristiansen J, Møller PP, et al. 2015. 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 12:6, PMID: 25890359, https://doi.org/10.1186/s12989-015-0081-9.

Imai K, van Dyk DA. 2004. Causal inference with general treatment regimes: generalizing the propensity score. J Am Stat Assoc 99(467):854–866, https://doi.org/10.1198/016214504000001187.

Jerrett M, Burnett RT, Ma R, Pope CA III, Krewski D, Newbold KB, et al. 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16(6):727–736, PMID: 16222161, https://doi.org/10.1097/01.ede.0000181630.15826.7d.

Kioumourtzoglou MA, Schwartz J, James P, Dominici F, Zanobetti A. 2016. PM2.5 and mortality in 207 US cities: modification by temperature and city characteristics. Epidemiology 27(2):221–227, PMID: 26600257, https://doi.org/10.1097/EDE.0000000000000422.

Kloog I, Chudnovsky AA, Just AC, Nordio F, Koutrakis P, Coull BA, et al. 2014. 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. Atmos Environ (1994) 95:581–590, PMID: 28966552, https://doi.org/10.1016/j.atmosenv.2014.07.014.

Kloog I, Zanobetti A, Nordio F, Coull BA, Baccarelli AA, Schwartz J. 2015. Effects of airborne fine particles ($PM_{2.5}$) on deep vein thrombosis admissions in the northeastern United States. J Thromb Haemost 13(5):768–774, PMID: 25678264, https://doi.org/10.1111/jth.12873.

Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, et al. 2009. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. Res Rep Health Eff Inst 140:5–114, PMID: 19627030.

Laden F, Schwartz J, Speizer FE, Dockery DW. 2006. Reduction in fine particulate air pollution and mortality: extended follow-up of the Harvard Six Cities study. Am J Respir Crit Care Med 173(6):667–672, PMID: 16424447, https://doi.org/10.1164/rccm.200503-443OC.

Langrish JP, Mills NL, Chan JK, Leseman DL, Aitken RJ, Fokkens PH, et al. 2009. Beneficial cardiovascular effects of reducing exposure to particulate air pollution with a simple facemask. Part Fibre Toxicol 6:8, PMID: 19284642, https://doi.org/10.1186/1743-8977-6-8.

Lepeule J, Laden F, Dockery D, Schwartz J. 2012. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. Environ Health Perspect 120(7):965–970, PMID: 22456598, https://doi.org/10.1289/ehp.1104660.

Li H, Cai J, Chen R, Zhao Z, Ying Z, Wang L, et al. 2017. Particulate matter exposure and stress hormone levels: a randomized, double-blind, crossover trial of air purifcation. Circulation 136(7):618–627, https://doi.org/10.1161/CIRCULATIONAHA.116.026796.

Lunceford JK, Davidian M. 2004. Stratiflcation and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat Med 23(19):2937–2960, PMID: 15351954, https://doi.org/10.1002/sim.1903.

Lyapustin A, Wang Y, Laszlo I, Kahn R, Korkin S, Remer L, et al. 2011. Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm. J Geophys Res Atmos 116:D03211, https://doi.org/10.1029/2010JD014986.

Mauad T, Rivero DH, de Oliveira RC, Lichtenfels AJ, Guimarães ET, de Andre PA, et al. 2008. Chronic exposure to ambient levels of urban particles affects mouse lung development. Am J Respir Crit Care Med 178(7):721–728, PMID: 18596224, https://doi.org/10.1164/rccm.200803-436OC.

Miller KA, Siscovick DS, Sheppard L, Shepherd K, Sullivan JH, Anderson GL, et al. 2007. Long-term exposure to air pollution and incidence of cardiovascular events in women. N Engl J Med 356(5):447–458, PMID: 17267905, https://doi.org/10.1056/NEJMoa054409.

National Center for Health Statistics. 2017. *Health, United States, 2016: With Chartbook on Long-Term Trends in Health*. Hyattsville, MD:National Center for Health Statistics.

Pinault L, Tjepkema M, Crouse DL, Weichenthal S, van Donkelaar A, Martin RV, et al. 2016. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort. Environ Health 15:18, PMID: 26864652, https://doi.org/10.1186/s12940-016-0111-6.

Pope CA III, Ezzati M, Dockery DW. 2009. Fine-particulate air pollution and life expectancy in the United States. N Engl J Med 360(4):376–386, PMID: 19164188, https://doi.org/10.1056/NEJMsa0805646.

Puett RC, Hart JE, Yanosky JD, Paciorek C, Schwartz J, Suh H, et al. 2009. Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. Environ Health Perspect 117(11):1697–1701, PMID: 20049120, https://doi.org/10.1289/ehp.0900572.

ResDAC (Research Data Assistance Center). 2018. Master Beneficiary Summary File (MBSF) Base. https://www.resdac.org/cms-data/files/mbsf-base [accessed 28 November 2018].

Robins JM, Hernán MA, Brumback B. 2000. Marginal structural models and causal inference in epidemiology. Epidemiology 11:550–560, PMID: 10955408, https://doi.org/10.1097/00001648-200009000-00011.

Shi L, Zanobetti A, Kloog I, Coull BA, Koutrakis P, Melly SJ, et al. 2016. Low-concentration $PM_{2.5}$ and mortality: estimating acute and chronic effects in a population-based study. Environ Health Perspect 124(1):46–52, PMID: 26038801, https://doi.org/10.1289/ehp.1409111.

Soares SR, Carvalho-Oliveira R, Ramos-Sanchez E, Catanozi S, da Silva LF, Mauad T, et al. 2009. Air pollution and antibodies against modified lipoproteins are associated with atherosclerosis and vascular remodeling in hyperlipemic mice. Atherosclerosis 207(2):368–373, PMID: 19486979, https://doi.org/10.1016/j.atherosclerosis.2009.04.041.

Sun Q, Yue P, Kirk RI, Wang A, Moatti D, Jin X, et al. 2008. Ambient air particulate matter exposure and tissue factor expression in atherosclerosis. Inhal Toxicol 20(2):127–137, PMID: 18236227, https://doi.org/10.1080/08958370701821482.

U.S. Census Bureau. 2010. U.S. Census 2000. Summary file 3 (sf 3). https://www.census.gov/census2000/sumfile3.html [accessed 27 November 2018].

U.S. EPA (U.S. Environmental Protection Agency). 2018. "Integrated Science Assessment (ISA) for Particulate Matter (External Review Draft)." EPA/600/R-18/179. Washington, DC:U.S. EPA.

Vodonos A, Awad YA, Schwartz J. 2018. The concentration-response between long-term $PM_{2.5}$ exposure and mortality; a meta-regression approach. Environ Res 166:677–689, PMID: 30077140, https://doi.org/10.1016/j.envres.2018.06.021.

Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. 2016. Estimating causal effects of long-term $PM_{2.5}$ exposure on mortality in New Jersey. Environ Health Perspect 124(8):1182–1188, PMID: 27082965, https://doi.org/10.1289/ehp.1409671.

Wang Y, Lee M, Liu P, Shi L, Yu Z, Awad YA, et al. 2017. Doubly robust additive hazards models to estimate effects of a continuous exposure on survival. Epidemiology 28(6):771–779, PMID: 28832358, https://doi.org/10.1097/EDE.0000000000000742.

Weisskopf M, Webster T. 2017. Trade-offs of personal versus more proxy exposure measures in environmental epidemiology. Epidemiology 28(5):635–643, PMID: 28520644, https://doi.org/10.1097/EDE.0000000000000686.

Wennberg J, Cooper M. 1996. *The Dartmouth Atlas of Health Care*. Chicago, IL: American Hospital Publishing.