Incident dementia and long-term exposure to constituents of fine particle air pollution: A national cohort study in the United States

Nelson Steenland, Emory University
William Caudle, Emory University
Thomas Wingo, Emory University
Howard Chang, Emory University
Jeremy Sarnat, Emory University
Liuhua Shi, Emory University
Qiao Zhu, Emory University
Y Wang, Rollins School of Public Health
H Hao, Rollins School of Public Health
H Zhang, Rollins School of Public Health

Only first 10 authors above; see publication for full author list.

**Journal Title:** Proceedings of the National Academy of Sciences of the United States of America

**Volume:** Volume 120, Number 1

**Publisher:** (publisher) | 2023-01-03, Pages e2211282119-e2211282119

**Type of Work:** Article

**Publisher DOI:** 10.1073/pnas.2211282119

**Permanent URL:** https://pid.emory.edu/ark:/25593/w565j

Final published version: http://dx.doi.org/10.1073/pnas.2211282119

Accessed July 6, 2023 7:46 AM EDT
Incident dementia and long-term exposure to constituents of fine particle air pollution: A national cohort study in the United States

Liuhua Shi1, Qiao Zhu1, Yifan Wang2, Hua Hao3, Haisu Zhang1, Joel Schwartz2,3,4, Heresh Amini3, Aaron van Donkelaar3,6, Randall V. Martin3, Kyle Steenland5, Jeremy A. Sarnat2, W. Michael Caudle3, Tszshan Ma7, Haomin Li8, Howard H. Chang4, Jeremiah Z. Liu9, Thomas Wingo8,10, Xiaobo Mao9, Armistead G. Russell8, Rodney J. Weber8, and Pengfei Liu1,11

Edited by George Thurston, New York University School of Medicine, NY; received June 30, 2022; accepted November 10, 2022 by Editorial Board Member Kenneth W. Wachter

Growing evidence suggests that fine particulate matter (PM$_{2.5}$) likely increases the risks of dementia, yet little is known about the relative contributions of different constituents. Here, we conducted a nationwide population-based cohort study (2000 to 2017) by integrating the Medicare Chronic Conditions Warehouse database and two independently sourced datasets of high-resolution PM$_{2.5}$ major chemical composition, including black carbon (BC), organic matter (OM), nitrate (NO$_3^-$), sulfate (SO$_4^{2-}$), ammonium (NH$_4^+$), and soil dust (DUST). To investigate the impact of long-term exposure to PM$_{2.5}$ constituents on incident all-cause dementia and Alzheimer’s disease (AD), hazard ratios for dementia and AD were estimated using Cox proportional hazards models, and penalized splines were used to evaluate potential nonlinear concentration–response (C-R) relationships. Results using two exposure datasets consistently indicated higher rates of incident dementia and AD for an increased exposure to PM$_{2.5}$ and its major constituents. An interquartile range increase in PM$_{2.5}$ mass was associated with a 6 to 7% increase in dementia incidence and a 9% increase in AD incidence. For different PM$_{2.5}$ constituents, associations remained significant for BC, OM, SO$_4^{2-}$, and NH$_4^+$ for both end points (even after adjustments of other constituents), among which BC and SO$_4^{2-}$ showed the strongest associations. All constituents had largely linear C-R relationships in the low exposure range, but most tailed off at higher exposure concentrations. Our findings suggest that long-term exposure to PM$_{2.5}$ is significantly associated with higher rates of incident dementia and AD and that SO$_4^{2-}$, BC, and OM related to traffic and fossil fuel combustion might drive the observed associations.

Globally, dementia is the seventh leading cause of death in the United States and a major cause of disability and dependency among older people, posing an urgent and significant public health challenge (1). More than 6 million Americans presently live with Alzheimer’s disease (AD) and dementia, leading to a massive economic burden; this number is projected to triple to approximately 14 million by 2060 (2). As medical treatments for the most common types of dementia remain challenging, identifying and mitigating modifiable risk factors are of paramount importance.

Growing evidence indicates that exposure to air pollution, specifically fine particulate matter (PM$_{2.5}$), plays a crucial role in the pathogenesis of AD and AD-related dementias (ADRD). A recent critical review by Delgado-Saborit et al. (2021) (3) summarized the epidemiological evidence of the effects of air pollution on ADRD, and a positive association between long-term exposure to PM$_{2.5}$ mass and increased ADRD risks was consistently reported across almost all studies. Relevant publications on air pollution and ADRD have doubled since this review, with the majority of studies finding positive associations between PM$_{2.5}$ mass and either all-cause dementia or AD (4–6).

To facilitate the targeting of pollution control efforts, the National Academy of Sciences and the World Health Organization have placed a high priority on determining which constituents of the PM$_{2.5}$ mass may be most hazardous (7, 8). Yang et al. (9) reviewed the literature regarding PM$_{2.5}$ constituents and both short-term and long-term health effects, focusing on all-cause mortality and cardiorespiratory morbidity, and found consistent associations with both black carbon (BC) and organic carbon (OC). However, the relative contributions of individual PM$_{2.5}$ constituents to ADRD risks remain largely unknown. Only one study to date has focused on PM$_{2.5}$ constituents and ADRD, which looked at PM$_{2.5}$ constituents and all-cause dementia in the northeastern United States (10).

Significance

Identifying the culprits of PM$_{2.5}$ constituents that are most responsible for elevated risks of neurodegeneration is of paramount importance. We perform a US nationwide cohort study of the associations between PM$_{2.5}$ constituents and dementia and AD. Long-term exposure to PM$_{2.5}$ mass and major constituents, particularly from traffic and fossil fuel combustion sources, is significantly associated with elevated dementia or AD incidence. All constituents had largely linear concentration–response relationships at low concentrations for both end points, implying no safe level of air pollution for brain health. Using two independent exposure datasets allows us to examine the robustness of findings and thus strengthen the credibility of the evidence for the associations. Our results will facilitate targeted source-specific pollution control strategies.

The authors declare no competing interest.

This article is a PNAS Direct Submission. G.T. is a guest editor invited by the Editorial Board.

Copyright © 2022 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1To whom correspondence may be addressed. Email: liuhua.shi@emory.edu or pengfei.liu@eas.gatech.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2211282119/-/DCSupplemental.

Published December 27, 2022.
Elucidating the potential relationship between PM$_{2.5}$ constituents and dementia has been challenging because of the sparsity of available, speciated chemical composition measurements, and the chronic nature of neurodegeneration. To cope with these challenges, long-term, high-resolution spatiotemporal PM$_{2.5}$ constituents’ estimates are needed, which require modeling with measurement constraints from ground observations. In addition, more complete health records, such as physician visits, inpatient visits, and outpatient visits, are needed to better capture disease incidence (5).

Here, we present a nationwide open cohort (i.e., dynamic cohort, meaning that members can leave or be added over time) study of the long-term PM$_{2.5}$ constituents’ exposure with incident dementia and AD among US older adults during 2000 to 2017. We used two high-resolution, multiple-species air pollution datasets and all Medicare claims across the contiguous United States to estimate the effect of PM$_{2.5}$ constituents on dementia risk.

**Results**

**Study Population Characteristics.** SI Appendix, Table S1 presents descriptive information on the dementia cohort and AD cohort between 2000 and 2017, with a 3-y clean period without events of interest. The dementia cohort included approximately 18.5 million individuals, and ~5.8 million individuals developed dementia. The AD cohort had approximately 19.2 million individuals, and ~2.8 million individuals developed AD. In both cohorts, most of the study population were White (~90%) and not eligible for Medicaid insurance (~89%), and about 60% were females. More detailed demographic characteristics on average and by PM$_{2.5}$ mass quintiles are presented in SI Appendix, Tables S1 and S2. County-level occurrences of first dementia and AD events per 100,000 Medicare beneficiaries across the contiguous United States (2000 to 2017) are presented in Fig. 1.

**Air Pollution Levels.** We accessed two high-resolution, speciated air pollution datasets of the contiguous United States from 2000 to 2017 from two independent sources, including BC, organic matter (OM), nitrate (NO$_3^-$), sulfate (SO$_4^{2-}$), ammonium (NH$_4^+$), and soil dust (DUST) (see Methods). Using exposure I, the dementia cohort had an average PM$_{2.5}$ mass concentration of 9.58 µg/m$^3$, with an interquartile range (IQR) of 3.68 µg/m$^3$ (11). The average concentrations of PM$_{2.5}$ major constituents and the corresponding IQRs are listed in SI Appendix, Table S1. The AD cohort shared a similar exposure with the dementia cohort (SI Appendix, Table S1). Two speciated air pollution datasets tend to have similar spatial distributions of major PM$_{2.5}$ constituents across the United States (Fig. 2). SI Appendix, Fig. S1 lists the correlation matrix among PM$_{2.5}$ mass and its six major constituents. PM$_{2.5}$ mass was highly correlated with NH$_4^+$, SO$_4^{2-}$, BC, and OM in both exposure sets (r values range from 0.60 to
0.83). BC and OM ($r = 0.74$ and 0.78 for exposures I and II, respectively) as well as $\text{SO}_4^{2-}$ and $\text{NH}_4^+$ ($r = 0.84$ and 0.86) also showed strong correlations in both exposure sets. The temporal trends of the annual population-weighted mean concentration of six PM$_{2.5}$ constituents are presented in SI Appendix, Fig. S2 of supplementary materials. Overall, from 2000 to 2017, national annual average concentrations showed a drastic downward trend for $\text{SO}_4^{2-}$, $\text{NH}_4^+$, and $\text{NO}_3^-$ in both exposure datasets, with more variability observed in $\text{SO}_4^{2-}$ concentrations from year to year, while other constituents showed a slighter temporal variation in different exposure datasets.

**Health Effect Estimates.** Fig. 3 provides the main results from the single-constituent Cox proportional hazards models stratified by individual characteristics, adjusting for neighborhood-level socioeconomic status (SES), behavioral risk factors, health-care capacity variables, and residual temporal and spatial trends (see Methods). Higher PM$_{2.5}$ mass and major constituents of interest were all consistently observed to be associated with increased dementia and AD incidence using exposure I. A null association was observed between DUST and dementia using exposure II. Per IQR increase in each pollutant in exposure I, hazard ratios (HRs) of dementia were 1.068 (95% CI: 1.063, 1.072) for PM$_{2.5}$ mass, 1.043 (1.040, 1.047) for BC, 1.036 (1.032, 1.039) for OM, 1.006 (1.003, 1.010) for DUST, 1.105 (1.099, 1.111) for $\text{SO}_4^{2-}$, 1.005 (1.000, 1.010) for $\text{NO}_3^-$, and 1.105 (1.099, 1.111) for $\text{NH}_4^+$ (SI Appendix, Table S4). Corresponding HRs of AD were 1.106 (1.099, 1.114) for PM$_{2.5}$ mass, 1.078 (1.071, 1.084) for BC, 1.062 (1.057, 1.068) for OM, 1.025 (1.020, 1.030) for DUST, 1.132 (1.123, 1.141) for $\text{SO}_4^{2-}$, 1.005 (0.997, 1.103) for $\text{NO}_3^-$, and 1.069 (1.060, 1.078) for $\text{NH}_4^+$ (SI Appendix, Table S5), respectively. Exposure II in general yielded slightly larger effect estimates than exposure I for both end points, and the association between NO$_3^-$ and AD became significant using exposure II. On a per $1 \mu g/m^3$ basis, BC had the highest associations with both end points across exposure datasets, followed by $\text{NH}_4^+$ and $\text{SO}_4^{2-}$.
Per 1 μg/m³ increase in BC, HRs of dementia were 1.123 (1.112, 1.135) using exposure I and 1.247 (1.223, 1.272) using exposure II (SI Appendix, Table S4), and HRs of AD were 1.227 (1.207, 1.247) using exposure I and 1.393 (1.358, 1.429) using exposure II (SI Appendix, Table S5).

Concentration–Response (C-R) Relationships. Fig. 4 presents the C-R relationships between each PM₂.₅ constituent of interest and outcomes of incident dementia and AD from single-constituent models. Linear relationships were observed with BC (exposure II), SO₄²⁻, and NH₄⁺, with no sign of threshold for both outcomes. The curves for BC with exposure I are essentially near linear for both end points until high and rarely occurring concentrations. Near-linear relationships were observed with OM (<5 μg/m³) and DUST (<1 μg/m³) in middle concentrations and then leveled out for higher concentrations for both end points. The C-R curves showed positive linear associations between both dementia and AD outcomes and NO₃⁻ at low concentrations; however, the relationships became unstable above about 1 μg/m³.

Effect Modifications. The subgroup-specific HRs in single-constituent models are presented in SI Appendix, Fig. S3. We found that BC and SO₄²⁻ were always positively associated with both dementia and AD across effect modifiers. The results obtained using estimates from exposure I and exposure II mostly showed similar patterns across the effect modifiers. For dementia, effect estimates for NO₃⁻, SO₄²⁻, and NH₄⁺ were higher among relatively older populations and males. Black individuals had a higher risk of dementia associated with NO₃⁻ and NH₄⁺, while races other than Blacks and Whites had a stronger association of dementia with DUST and SO₄²⁻. Additionally, for DUST, SO₄²⁻, and NH₄⁺, people eligible for Medicaid were at a significantly greater risk of dementia than those not eligible for Medicaid. Similar patterns were also observed in the association between AD and these constituents.

Sensitivity Analysis. Our results were robust in a series of sensitivity analyses. First, for both outcomes, multiconstituent models yielded similar results for most PM₂.₅ constituents, except that NO₃⁻ was observed to have a negative impact on dementia or AD after adjusting for other PM₂.₅ constituents using both exposure datasets (SI Appendix, Tables S4 and S5). The adjustment for the residual PM₂.₅ mass (i.e., subtracting the constituent of interest from total PM₂.₅ mass) yielded similar results, compared to the multiconstituent models (SI Appendix, Tables S4 and S5). Second, we considered the possible effect of outcome misclassification by 1) fitting a linear regression model based on incidence rates and 2) estimating the true number of cases misclassification by 1) fitting a linear regression model based on prior estimates of Medicare sensitivity and specificity, and the results of both methods remained largely stable toward the null (SI Appendix, Tables S6 and S7). Third, a stricter “clean” period of 5 y yielded roughly consistent results with the main analyses (SI Appendix, Table S8). In addition, the results of the nonmover cohort suggest little bias from residential mobility (SI Appendix, Table S9). Furthermore, the effect estimates remained unchanged regardless of the form of adjustment for “year” in the main models (SI Appendix, Table S10). Last, after accounting for differing exposure measurement error, we found that the relative effect estimates across the major PM₂.₅ constituents were consistent with our main analyses, and the conclusions were not affected, although the magnitudes were attenuated (SI Appendix, Table S11).
Discussion

Using two sets of high-resolution air pollution datasets, we consistently observed that long-term exposure to PM$_{2.5}$ mass and its major constituents were associated with higher risks of dementia and AD among older adults in a large US cohort. More specifically, higher exposure to four (i.e., SO$_4^{2-}$, NH$_4^+$, BC, and OM) of the six constituents explored were consistently associated with higher dementia and AD risks. SO$_4^{2-}$ and BC were associated with the highest dementia or AD risks, while NO$_3^-$ and DUST had relatively lower impacts. Overall, we observed stronger associations for AD than dementia. The relatively greater influence of particulate matter on AD may be because dementia encompasses a wide variety of disorders with different etiologies, some of which may be unrelated to fine particle pollution. The results from the multiconstituent models were only modestly changed from those in the single constituent models except for NO$_3^-$ suggesting that these associations were not confounded by other constituents.

The effect estimates per IQR increase on dementia and AD are the largest for SO$_4^{2-}$ and lowest for NO$_3^-$ Our single-constituent and multiconstituent models showed inconsistent results for NO$_3^-$ and we interpret this as evidence that there is no sufficient statistical evidence for the harmful effect of NO$_3^-$, and the protective effect seen in the multiconstituent models may be due to collinearity. Another plausible mechanism to explain the protective effect of NO$_3^-$ is that the presence of NO$_3^-$ is associated with lower aerosol acidity, which may reduce the solubility and bioavailability of transition metals (13). The per IQR effect estimates for NH$_4^+$ are in between, as expected, as NH$_4^+$ is chemically associated with SO$_4^{2-}$ and NO$_3^-$ The principal sources of SO$_2$ (i.e., SO$_2^{2-}$ precursor) are fossil fuel combustion (14). SO$_2^{2-}$ in the form of (NH$_4$)$_2$SO$_4$ or NH$_4$HSO$_4$ does not have strong acute neurotoxicity based on toxicological assessments (15). However,

![Fig. 4. The concentration–response curves for each PM$_{2.5}$ constituent and dementia (A) and Alzheimer's disease (B). The concentration–response curves, derived from the single-constituent models, are shown for the concentration ranges between 0.5th and 99.5th percentiles of the pollutants, i.e., with 1% poorly constrained extreme values excluded. For each constituent, the top panel used exposure I data (11) and the bottom panel used exposure II data (12).](https://doi.org/10.1073/pnas.2211282119)
SO$_4^{2-}$ may serve as a proxy for other combustion-emitted pollutants subsequently processed in the atmosphere, such as oxygenated organic species, which are known to cause oxidative stress (16). Excessive oxidative stress has been linked with the production of α-synuclein, amyloid-β, and hyperphosphorylated tau protein and may play a role in ADRD etiology (17).

Another plausible explanation is that the high aerosol acidity associated with SO$_4^{2-}$ may enhance the solubility and bioavailability of trace metals (18) and further generate reactive oxygen species in vivo by redox cycling and lead to oxidative stress (19). Studies on AD pathology found that the olfactory bulbs participate in the neuroinflammatory process related to exposure to metals in fine particles (20). The role of the olfactory nerve as a direct pathway of exposure to PM$_{2.5}$ is supported by rodent studies that have demonstrated exacerbation of AD-related pathology following exposure to well-characterized ambient air pollution (21). As rodents are obligate nose breathers, this exposure scenario isolates the olfactory system as a direct route through which PM$_{2.5}$ and its constituents could access the brain and contribute to AD pathology. In studies using the wild type and rats that express human AD risk genes, exposure to traffic-related air pollution resulted in an increase in PM$_{2.5}$ particles in the hippocampus of wild-type and transgenic rats, with a concomitant increase in AD-related pathology in both genotypes (22). Although focused on PM$_{2.5}$, these findings provide clear evidence for the olfactory nerve to serve as a conduit for PM$_{2.5}$ and its associated components to contribute to AD pathogenesis. Moreover, the synergistic effect of (NH$_4$)$_2$SO$_4$ and the presence of ultrafine particles may accelerate the aggregation of peptides that impact progression of neurodegenerative diseases (23). Collectively, these studies provide support for our findings about the relatively higher magnitude of SO$_4^{2-}$ compared to NO$_3^-$.

Among all examined components of PM$_{2.5}$, BC has the largest effect estimates on dementia and AD per 1 µg/m$^3$ increase. We note that exposure II has higher effect estimates per 1 µg/m$^3$ than exposure I, mainly because of the discrepancy in estimated BC concentrations between two datasets (SI Appendix, Table S1). This discrepancy in BC estimates can be explained by the fact that these two exposure sets relied on monitor data based on different techniques (i.e., thermal method vs. optical methods). (24, 25) BC are combustion-related particles, mainly derived from tailpipe traffic emissions and biomass burning (26). Our results are consistent with the literature that suggests an association between traffic-related air pollution (measured as BC levels) and poor cognition levels in older men (27). Possible explanations of the neurotoxicity of BC include increased levels of inflammatory mediators, markers of oxidative damage to DNA, and β-amyloid deposition as well as evidence of blood–brain barrier disruption (28). BC particles and other ultrafine particles can also be small enough to pass through the olfactory nerve, bypass the blood–brain barrier, and translocate to the brain (29), leading to oxidative stress and neuroinflammation (30). A recent study also highlights the translocation of fine BC particles from the lung to the brain through the inhalation/circulation/brain route (31). In addition, BC could act as a carrier for highly toxic OM species such as polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs), (32) and BC can be coemitted with nontailpipe traffic-related pollutants, such as road dust and metals and organics from brake and tire abrasion (33). These species can have adverse neurotoxic effects (34). On a per 1-µg/m$^3$ basis, NH$_4^+$ and SO$_4^{2-}$ also had high effect estimates on dementia and AD. However, as SO$_4^{2-}$ is usually highly correlated with NH$_4^+$, the HR per 1 µg/m$^3$ of (NH$_4$)$_2$SO$_4$ would be lower than the values reported for each individual constituent.

OM is another important PM$_{2.5}$ constituent that increases the risks of dementia and AD in our study. OM makes up a substantial fraction of PM$_{2.5}$ mass, consisting of primary OM directly emitted from combustion emissions and other sources and secondary OM formed by atmospheric oxidation of gas-phase species (35). Due to the complexity of OM in terms of sources and constitutions, the toxicity of some OM compounds has motivated greater scrutiny than other PM$_{2.5}$ compounds in health studies (36). OM includes highly toxic species such as PAHs (including products of atmospheric processing, such as oxy- and nitro-PAHs) and PCBs and contributes a large fraction of aerosol oxidative potential (37, 38). Ultrafine particles from traffic sources also largely consist of carcinogenic species, including OM and BC (39). However, research on the neurological effects of atmospheric OM pollutants is scarce, and more studies are warranted.

The linear relationships of BC (exposure II), SO$_4^{2-}$, and NH$_4^+$ with both outcomes were observed with no sign of threshold, and these results were consistent with the linear C-R relationships in the previously published “Northeastern” study (40). Interestingly, “bell-shaped” C-R curves were observed for the relationship between OM, DUST, and NO$_3^-$ exposure and dementia/AD, suggesting that the C-R curves are relatively steep at very low to moderate levels of exposure and leveled down at high levels of exposure. Numerous hypotheses have sought to explain the mechanisms leading to a nonlinear C-R relationship; these include errors in calculating exposure levels of pollution at higher concentrations, the existence of competing risks, and preferential avoidance based on symptoms (41). Moreover, the difference in population distributions across different constituents may also lead to different C-R curves.

Previous epidemiology studies have also demonstrated the associations between short-term and long-term exposure to PM$_{2.5}$ constituents for different morbidity and mortality. One recently published literature review has systematically reviewed 35 epidemiological studies (25 time series and 10 cohorts), and the authors observed the most robust and consistent associations between both BC and OC for all-cause mortality and cardiovascular mortality and morbidity. They also reported that NO$_3^-$ and SO$_4^{2-}$ were relevant for adverse cardiovascular and respiratory health outcomes. Since that review, another large city-level daily time series study of all-cause mortality has recently been published (42), covering 210 cities in 16 countries. They found that NH$_4^+$ had the largest effect estimates, while SO$_4^{2-}$ also increased daily mortality, but less markedly.

**Strengths and Limitations.** Our study has several strengths. This is a nationwide, population-based, open cohort study characterizing the health effects of ambient PM$_{2.5}$ constituent exposure on dementia and AD incidence. The large sample size gives us ample statistical power to identify the effects of long-term PM$_{2.5}$ constituent exposure on neurodegeneration. This study provides insights into composition-specific health effects of PM$_{2.5}$, while most studies focus on total PM$_{2.5}$ mass. Second, two independently sourced state-of-the-art exposure datasets allow us to examine the robustness of results against measurement error, thus strengthening the credibility of the findings. Third, using comprehensive Medicare claims (including physician visits) can better reflect incidence, as they can capture more cases, particularly earlier diagnosed cases that are often missed in hospitalization records.

Despite these strengths, we acknowledge that our study has several limitations. First, exposure measurement error has been inevitably introduced when using predicted ambient concentrations, albeit two datasets showed consistent effect estimates.
Additionally, coarse exposure concentrations at the ZIP code level may not allow some constituents with high spatial variability to be well characterized. However, due to the limitations in Medicare data resolution, health outcome information at the ZIP code level is the most granular information that can be obtained. Thus, analysis using higher resolution health information matched with fine PM$_{2.5}$ constituents’ exposure estimates is needed in future studies. Second, outcome misclassification is likely to occur when relying on administrative records. AD cases account for 48% of dementia diagnoses in our database, suggesting the undiagnosed nature of AD using administrative records (43). Moreover, dementia and AD have a long and insidious onset, and the exact timing of the disease onset is unknown. Another limitation is that we are able to adjust for only potential confounders that can be estimated based on neighborhood-level characteristics, and future work should incorporate individual-level characteristics associated with ADRD (44, 45). In addition, it is unclear whether the high effects of BC and SO$_4^{2−}$ are due to their intrinsic neurotoxicity or other culprit pollutants that are coemitted and correlated. Last, although PM$_{2.5}$ constituents provide useful information about sources, it is important to assess more source-specific effects of PM$_{2.5}$, as they can be readily translatable into effective abatement strategies.

In conclusion, long-term exposure to fine particle pollution is associated with higher risks of dementia and AD, and individual PM$_{2.5}$ constituents are associated with differences in risk. BC and SO$_4^{2−}$ have the strongest associations. Our findings imply that policies that target the reduction in ambient PM$_{2.5}$ concentrations, particularly primary and secondary particulate pollutants from sources such as traffic and sulfur-containing fossil fuel combustion, have a significant public health impact.

Materials and Methods

Study Population. We analyzed two nationwide, privacy-protected and publicly available databases from the Centers for Medicare and Medicaid Services (CMS), including the Medicare denominator file and the Medicare Chronic Conditions Warehouse (CCW), based on which we constructed separate cohorts for all-cause dementia and AD subtype in 2000 to 2017. The denominator file contains enrollment records for each Medicare beneficiary, including demographics, Medicaid insurance status (a proxy for SES), the date of death (if any), and ZIP code of residence, which were updated annually. The CCW claims data include predefined indicators for chronic conditions among the fee-for-service (FFS) Medicare beneficiaries and provides the date of the first occurrence with a diagnosis code for a specific condition.

Our study population comprised all Medicare beneficiaries enrolled at age 65 y or above living in the contiguous United States from 2000 to 2017, with continuous enrollment 1) in the Medicare FFS program and 2) in both Medicare Part A (hospital insurance) and Part B (medical insurance) over the follow-up period. These inclusion criteria were used because the CCW relies on FFS, Part A, and Part B to identify cases. In addition, we further required a clean period of 3 y of enrollment without dementia or AD to better approximate “incidence,” i.e., disease onset is unknown. Another limitation is that we are able to adjust for only potential confounders that can be estimated based on neighborhood-level characteristics, and future work should incorporate individual-level characteristics associated with ADRD (44, 45). In addition, it is unclear whether the high effects of BC and SO$_4^{2−}$ are due to their intrinsic neurotoxicity or other culprit pollutants that are coemitted and correlated. Last, although PM$_{2.5}$ constituents provide useful information about sources, it is important to assess more source-specific effects of PM$_{2.5}$, as they can be readily translatable into effective abatement strategies.

Outcome Classification. All-cause dementia and AD subtype were the two primary outcomes of this study. AD is a specific single disease that falls under the dementia umbrella, and AD is the major type of dementia that represents ~60 to 80% of dementia cases. The diagnoses of dementia and AD were identified and recorded as two distinct indicators using an algorithm created by CCW that incorporated information across the available Medicare claims, including inpatient and outpatient claims, Carrier files (primarily doctor visits), skilled nursing facility, and home health-care claims. This algorithm has been shown to be reasonably accurate in classifying diseases based on validation studies such as those by Taylor et al. (46, 47) who found high sensitivity for dementia (0.85) but lower for AD (0.65). CCW provides the date of the first occurrence with a dementia or AD diagnosis code. In the dementia cohort, the outcome was defined as the first occurrence of a dementia diagnosis code. In the AD cohort, the outcome was defined as either (1) the first occurrence of an AD diagnosis with no prior dementia diagnosis or (2) the first occurrence of a dementia diagnosis when there was a subsequent AD diagnosis (given that the original dementia diagnosis was probably AD). Given that previous studies (5) have found a greater effect of long-term PM$_{2.5}$ exposure on AD compared to dementia, it is likely that some causes of dementia may be less associated with air pollution, while AD with distinct disease assessment had a stronger association. The specific effects of different PM$_{2.5}$ constituents on dementia and AD were evaluated separately in this study to further assess whether AD progression is more related to PM$_{2.5}$ pollution compared to all-cause dementia.

Exposure Assessment. We accessed two high-resolution, specified air pollution datasets of the contiguous United States from 2000 to 2017 from two independent sources. The first set of annual mean concentrations including total PM$_{2.5}$ mass and its six major constituents (exposure I) was estimated at 1-km resolution using a previously validated PM$_{2.5}$ composition prediction model (11). Briefly, mean satellite-derived PM$_{2.5}$ total mass concentrations were first estimated at 1-km resolution by combining satellite retrievals of aerosol optical depth, chemical transport modeling (CTM), and ground-based observations. The six major constituents of PM$_{2.5}$ were calculated by decomposing the PM$_{2.5}$ total mass into individual chemical constituents using CTM output and further calibrated using ground-based observations. Our annual predictions of each constituent achieved good long-term spatial agreement compared with ground measurements, with cross-validated R$^2$ of 0.59 for BC, 0.86 for NO$_3^{−}$, 0.96 for SO$_4^{2−}$, 0.90 for NH$_4^{+}$, 0.57 for OM, and 0.60 for DUST.

The second set of annual mean predictions for PM$_{2.5}$ constituents (exposure II) was estimated using superlearning and ensemble weighted averaging models, with a 50-m spatial resolution in urban areas and a 1-km resolution in nonurban areas. Details about exposure II dataset can be found elsewhere (12). Briefly, PM$_{2.5}$ constituent training data were collected from 987 monitoring sites, and hundreds of predictors were used for six superlearning models and an ensemble weighted-averaging model. This approach achieved excellent model performance, with cross-validated R$^2$ for individual components used in this study ranging from 0.856 (OM) to 0.957 (SO$_4^{2−}$). In addition to PM$_{2.5}$ constituents, we have previously used a similar ensemble model that integrated multiple machine learners and hundreds of predictors to estimate PM$_{2.5}$ mass concentrations across the contiguous United States, with cross-validated R$^2$ of 0.89 for annual predictions (48). This PM$_{2.5}$ mass dataset has been widely used in previous epidemiological studies (5, 49, 50).

We finally averaged these gridded predictions for each constituent at the ZIP code level (i.e., the finest spatial resolution of Medicare data) for each year. Two sets of ZIP code-level annual means were assigned to each Medicare beneficiary, and any residential mobility changes by ZIP code were considered annually.

Covariates. Individual-level demographics (age, sex, and race) and Medicaid insurance status were obtained from the Medicare denominator file. We also included in the model neighborhood-level covariates, including ZIP code-level SES variables (population density, median household income, % Black population, % population living in rental house or apartment, % population aged 65 or above living below the poverty line, and % population with less than a high school education), county-level behavioral risk factors (smoking prevalence and mean body mass index), county-level health care capacity variables (number of hospitals and active medical doctors per 1,000 people), and a geographical region indicator. All covariates were included as linear terms in the models unless otherwise noted. Details and data sources of covariates are described in Shi et al. (2021)(49).

Statistical Analysis. We fit stratified Cox proportional hazards models with a generalized estimating equation (GEE) to estimate the associations between time-varying annual mean concentrations of PM$_{2.5}$ constituents on dementia or AD among older adults, with years of follow-up as the time scale. Given the potential multicollinearity among PM$_{2.5}$ constituents, we fit single-constituent models in the main analyses and estimated HRs per IQR increase and per 1 μg/m$^3$ increase in the annual mean concentrations of each PM$_{2.5}$ constituent. GEE allowed us to adjust...
for residual autocorrelation within the ZIP code and thus obtain more statistically robust CIs for the effect estimates. All models were stratified by age at entry (1–9 age categories), race (White, Black, and other), sex, and Medicaid insurance status and adjusted for the neighborhood-level covariates (see Covariates). To adjust for potential residual temporal and spatial trends, a linear term for calendar years and an indicator for geographical regions were included.

To account for possible nonlinearity in the CR relationships between each PM$_{2.5}$ constituent and dementia or AD, we fit penalized splines for each constituent in single-constituent models. The models adjusted for the same covariates as our main model. We further assessed potential effect modification by introducing an interaction term between the constituent and a modifier of individual-level characteristics, including age group (<75 y vs. >75 y), race, sex, and Medicaid eligibility.

We conducted several sensitivity analyses to assess the robustness of our main results. First, we fit multiconstituent models by including multiple PM$_{2.5}$ constituents in one model simultaneously. Considering the high correlations between BC and OM as well as between SO$_2$ and NO$_2$ (SI Appendix, Fig. S1), we fitted two separate multiconstituent models, 1) by including BC, DUST, SO$_2$, and NO$_2$ simultaneously and 2) by including OM, DUST, SO$_2$, and NO$_2$ simultaneously. Second, in single-constituent models, we additionally adjusted for the residual PM$_{2.5}$ mass (i.e., subtracting the constituent of interest from total PM$_{2.5}$ mass) other than the constituent of interest. Third, we assessed the impact of possible outcome misclassification in two ways: 1) fitting linear regression models for incidence rates of dementia or AD with GEE, which yielded additive effect estimates that were less sensitive to bias, because random misclassification of the outcome would be absorbed into the residual errors of the linear model for the true rates of events (51) and 2) considering the possible effect of outcome misclassification following methods similar to those described by Fox et al. (52) and adjusting the observed outcomes for each stratum based on the estimates of Medicare sensitivity and specificity from Taylor et al. (47) to estimate the expected true number of cases. In addition, we applied a clean period of 5 y by excluding anyone diagnosed with dementia or AD in their first 5 y of follow-up. Compared with the main analysis using a 3-y clean period, this would increase the possibility to capture the first diagnosis of dementia or AD, at a cost-reducing sample size. Moreover, we conducted a nonmover cohort analysis for subjects who did not move during the follow-up period to account for potential measurement error related to a change in the residential address. To further explore the impact of the temporal variation of exposure, we tested two other sets of adjustments for trend time, including 1) adding year as a spline term in the model with 4 degrees of freedom and 2) adding year as a categorical variable in the model. Finally, to investigate the impact of differing exposure measurement error across constituents on dementia and AD, we first extracted the regression mean square error for each constituent from the two exposure assessment models and then added an error term (following a log-normal distribution) to the original data and generated 20 perturbed datasets. Then, each dataset was used to fit the single-constituent models, and the results were summarized (53). To improve the computation efficiency, we bootstrapped 100 times for each perturbed dataset.

Software, version 4.0.2, was used for all analyses. Statistical significance was determined by two-sided $P < 0.05$.

Data, Materials, and Software Availability. The Medicare dataset was stored and analyzed in the Emory Rollins School secure cluster environment, with Health Insurance Portability and Accountability Act compliance. The rules governing the Medicare dataset prohibit any sharing of the health datasets being used for our epidemiologic research. Restricted by our Data Use Agreement with the US Centers for Medicare & Medicaid Services, the Medicare data that support the findings of this study are neither sharable nor publicly available from us. Academic and nonprofit researchers who are interested in using Medicare data should contact the US Centers for Medicare & Medicaid Services directly to obtain their own datasets upon completion of a Data Use Agreement.

Acknowledgments. We want to specially thank the Centers for Medicare & Medicaid Services for giving us access to the Medicare claims to conduct this study. We also gratefully acknowledge Caroline Owens for editorial support. This work was supported by NIH (R01 AG074357, R21 ES032666), the HERCULES Center (P30 ES019776), and the Emory Goizueta Alzheimer’s Disease Research Center (PS0 AG025688). J.S. and H.A. were supported by U.S. EPA (RD-83872101) and NIH (P30 ES000002, R01 ES032418-01). H.A. was supported by Novo Nordisk Foundation (17OC0027812). R.M.V. was supported by NASA HAOAST (Grant 80NSSC21K0308).

Author affiliations: “Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322; “Department of Environmental Health, Harvard Chan School of Public Health, Boston, MA 02115; “Department of Epidemiology, Harvard Chan School of Public Health, Boston, MA 02115; “Department of Public Health, Section of Environmental Health, University of Copenhagen, Copenhagen, Denmark 1014; “Department of Energy, Environmental & Chemical Engineering, Washington University at St. Louis, St. Louis, MO 63130; “Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322; “Department of Biostatistics, Harvard Chan School of Public Health, Boston, MA 02115; “Department of Neurology and Human Genetics, School of Medicine, Emory University, Atlanta, GA 30322; “Department of Neurology, Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205; “School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30318; and “School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30318