XAVIER BECERRA
Attorney General

State of California
DEPARTMENT OF JUSTICE

1300 I STREET, SUITE 125
P.O. BOX 944255
SACRAMENTO, CA 94244-2550

Public: (916) 445-9555
Telephone: (916) 210-7815
E-Mail: Scott.Lichtig@doj.ca.gov

February 16, 2021

Via Electronic Delivery[1] *and First Class Mail*

Jane Nishida, Acting Administrator
Office of the Administrator
U.S. Environmental Protection Agency
Room 3000, EPA WJC
1200 Pennsylvania Ave., NW
Washington, DC 20460
nishida.jane@epa.gov

Lars Perlmutt
Sector Policies and Programs Division
(Mail Code C539-04)
Office of Air Quality Planning & Standards
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711
perlmutt.lars@epa.gov

RE: Petition for Reconsideration of "Review of the National Ambient Air Quality Standards for Particulate Matter," 85 Fed. Reg. 82,684 (Dec. 18, 2020).

Dear Acting Administrator Nishida and Dr. Perlmutt:

Please find attached a Petition for Reconsideration submitted on behalf of the States of California (by and through Attorney General Xavier Becerra and the California Air Resources Board), Connecticut, Delaware, Illinois, Maryland, Minnesota, New Jersey, New York, Oregon, Rhode Island, Vermont, Washington, and Wisconsin; the Commonwealths of Massachusetts, Pennsylvania, and Virginia; and the City of New York, with respect to the above referenced action, Docket ID EPA–HQ–OAR–2015–0072.

[1] This Petition is submitted electronically in light of the COVID-19 pandemic and EPA's guidance with respect to hard copy submissions while Agency staff is teleworking. *Notice Regarding "Hard Copy" Submissions to EPA During the COVID-19 National Emergency* (May 12, 2020), https://www.epa.gov/aboutepa/notice-regarding-hard-copy-submissions-epa-during-covid-19-national-emergency.

Sincerely,

/s/ Scott J. Lichtig
SCOTT J. LICHTIG
Deputy Attorney General

For XAVIER BECERRA
Attorney General for the State of California

(Attachments)

cc:
Associate General Counsel for the Air and Radiation Law Office
Office of General Counsel (Mail Code 2344A)
U.S. Environmental Protection Agency
1200 Pennsylvania Ave. NW
Washington, DC 20460
A-and-R-Docket@epamail.epa.gov

BEFORE THE HONORABLE JANE NISHIDA, ACTING ADMINISTRATOR
UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

IN RE PETITION FOR RECONSIDERATION OF REVIEW OF THE NATIONAL AMBIENT AIR QUALITY STANDARDS FOR PARTICULATE MATTER, 85 FED. REG. 82,684 (DECEMBER 18, 2020).

Submitted by:

The States of California (by and through Attorney General Xavier Becerra and the California Air Resources Board), Connecticut, Delaware, Illinois, Maryland, Minnesota, New Jersey, New York, Oregon, Rhode Island, Vermont, Washington, and Wisconsin; the Commonwealths of Massachusetts, Pennsylvania, and Virginia; and the City of New York.

TABLE OF CONTENTS

INTRODUCTION

Pursuant to Clean Air Act Section 307(d), and for the reasons set forth below, the States of California (by and through Attorney General Xavier Becerra and the California Air Resources Board), Connecticut, Delaware, Illinois, Maryland, Minnesota, New Jersey, New York, Oregon, Rhode Island, Vermont, Washington, and Wisconsin; the Commonwealths of Massachusetts, Pennsylvania, and Virginia; and the City of New York (collectively, State Petitioners) hereby petition the U.S. Environmental Protection Agency (EPA) for reconsideration of its recent final action determining that strengthening the National Ambient Air Quality Standard (NAAQS) for particulate matter is not necessary to protect the public health. That action appears in the Federal Register, titled "Review of the National Ambient Air Quality Standards for Particulate Matter," published at 85 Fed. Reg. 82,684 (Dec. 18, 2020). Reconsideration is warranted here because State Petitioners' objections are based on information that arose after the end of the comment period concerning issues of central relevance to EPA's final determination. 42 U.S.C. § 7607(d)(7)(B). Since the close of the public comments, several critical new studies have both demonstrated that the current particulate matter standard is inadequate to protect the public health and have linked previously unidentified harms with increased exposure to particulate matter, health impacts that also disproportionately impact environmental justice communities throughout the nation. State Petitioners urge EPA to expeditiously reconsider the final rule and adopt new particulate matter standards that adequately account for the new information documented by these new studies.

EPA's decision not to strengthen the eight-year old particulate matter NAAQS fails to protect the public health and welfare from the effects of exposure to particulate matter as required under the Clean Air Act and is arbitrary and capricious, an abuse of discretion, and

otherwise in violation of the Act.[1] EPA's decision ignores substantial evidence demonstrating the serious harm to human health from particulate matter exposure at or below the existing NAAQS, repeatedly dismissing studies documenting such significant impacts due to purported "uncertainties." 85 Fed. Reg. at 82,685. In light of these new studies further documenting this harm, EPA should reconsider its decision to retain the standard and act swiftly to correct these fundamental—and ultimately deadly—errors. The public health and welfare cannot wait another five years.

The Clean Air Act requires that EPA reconsider its decision in light of new research released after the close of the public comment period—adding to the mountain of evidence demonstrating the serious harms related to particulate matter exposure and further underscoring the fallacy of any purported uncertainty asserted by EPA.[2] 42 U.S.C. § 7607(d)(7)(B). Specifically, a new study explicitly refutes EPA's assertion that insufficient evidence exists to prove health harm to human health at levels below the existing particulate matter NAAQS, demonstrating clear benefits from further strengthening the particulate matter standards.[3] Another study just published found that fine particulate matter less than 2.5 microns ($PM_{2.5}$) from fossil fuel combustion was responsible for *double* the amount of premature death worldwide than previously believed, and similarly showed significant health benefits from

[1] State Petitioners accordingly filed a petition for review of EPA's final rule on January 13, 2021.
[2] State Petitioners submitted several of these studies to EPA in a Supplemental Comment Letter on November 20, 2020, but EPA's final decision does not adequately address the new information provided.
[3] Joel Schwartz, et al., *A National Difference in Differences Analysis of the Effect of $PM_{2.5}$ on Annual Death Rates* (Journal Pre-Proof), at 20, ENV'T. RSCH., Vol. 194 (Mar. 2021), *available at*: https://www.sciencedirect.com/science/article/abs/pii/S0013935120315462?via%t3Dihub.

reductions in particulate matter even at low levels of exposure.[4] Further, multiple studies have found links between increased mortality from COVID-19 and exposure to $PM_{2.5}$, identifying additional significant potential injury to public health unaddressed in EPA's determination.[5] New research also demonstrates particulate matter exposure can lead to additional negative health impacts from serious neurodegenerative disorders such as Alzheimer's disease and other dementia in older Americans. This new information demonstrates the clear need for EPA to issue stronger standards as quickly as possible.[6]

These new studies further establish that EPA's decision not to lower the particulate matter NAAQS represents a massive failure to address one of the most critical environmental justice issues facing the nation. As noted in State Petitioners' original comments, a study recently estimated that the Black population's $PM_{2.5}$ burden was 54 percent higher than that of the general population, and the Latinx population's $PM_{2.5}$ burden was 20 percent higher.[7] The study also found that those living in poverty experience a $PM_{2.5}$ burden 35 percent greater than the general

---

[4] Karn Vohra, et al., *Global Mortality from Outdoor Fine Particle Pollution Generated by Fossil Fuel Combustion: Results from GEOS-Chem* (Journal Pre-Proof), ENV'T. RSCH. (Feb. 9, 2021), *available at*: http://acmg.seas.harvard.edu/publications/2021/vohra_2021_ff_mortality.pdf, and *at*: http://acmg.seas.harvard.edu/publications/2021/vohra_2021_ff_sup.pdf (for Supplemental Materials).

[5] Andrea. Pozzer, et al., *Regional and Global Contributions of Air Pollution to Risk of Death from COVID-19*, 116 CARDIOVASCULAR RSCH. 14, 2247 (Sept. 30, 2020), *available at*: https://academic.oup.com/cardiovascres/advance-article/doi/10.1093/cvr/cvaa288/5940460; Xiao Wu, et al., *Air Pollution and COVID-19 Mortality in the United States: Strengths and Limitations of an Ecological Regression Analysis*, SCI. ADVANCES, at 1 (Nov. 4, 2020), *available at*: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7673673/.

[6] Liuhua Shi et al., *Long-term Effects of $PM_{2.5}$ on Neurological Disorders in the American Medicare Population: A Longitudinal Cohort Study*, 4 THE LANCET E557, (Oct. 19, 2020), *available at*: https://www.thelancet.com/action/showPdf?pii=S2542-5196%2820%2930227-8.

[7] I. Mikati, Disparities in Distribution of Particulate Matter Emission Sources by Race and Poverty Status, 108(4) *Am. J. of Public Health* 480, 482 (Table 1) (Apr. 2018), *available at*: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5844406/.

population. Both COVID-19 and neurocognitive disease, the two illnesses identified as exacerbated by particulate matter exposure, already disproportionately impact Black, Indigenous, and persons of color (BIPOC). Continuing to expose these historically burdened communities to dangerous levels of particulate matter will only exacerbate the disparity in the rates of diseases already suffered by these communities.

EPA must address this cycle of environmental injustice. State Petitioners are encouraged that the White House has identified the particulate matter NAAQS determination as one of EPA's actions to be reviewed for consistency with President Biden's January 20, 2021, "Executive Order on Protecting Public Health and the Environment and Restoring Science to Tackle the Climate Crisis."[8] Promptly re-opening the proceedings would provide the opportunity for public comment on these significant harms and allow for EPA's consideration of this new substantial evidence showing the need for EPA to reduce exposure to particulate matter to protect the public health. Accordingly, the State Petitioners respectfully request that EPA convene proceedings for reconsideration of the final decision or take other appropriate action to efficiently and expeditiously correct the deficient standard. 42 U.S.C. § 7607(d)(7)(B).

LEGAL STANDARD

EPA must convene a reconsideration proceeding if a person raising an objection shows: (1) grounds for the objection arose after the public comment period; and (2) the objection "is of central relevance to the outcome of the rule." 42 U.S.C. § 7607(d)(7)(B). An objection is "of

[8] *See Executive Order on Protecting Public Health and the Environment and Restoring Science to Tackle the Climate Crisis*, Jan. 20, 2021, *available at*: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-protecting-public-health-and-environment-and-restoring-science-to-tackle-climate-crisis/; *see also Fact Sheet: List of Agency Actions for Review*, January 20, 2021, *available at*: https://www.whitehouse.gov/briefing-room/statements-releases/2021/01/20/fact-sheet-list-of-agency-actions-for-review/.

central relevance" if it provides "substantial support for the argument that the regulation should be revised." *Chesapeake Climate Action Network v. EPA*, 952 F.3d 310, 322 (D.C. Cir. 2020). Furthermore, EPA has discretion to reconsider its actions even where the standards for mandatory reconsideration are not met.

ARGUMENT

The Clean Air Act requires that EPA promulgate and revise the primary NAAQS to assure the standards are stringent enough "to protect public health," with "an adequate margin of safety." 42 U.S.C. § 7409(b)(1). The NAAQS must be based on air quality criteria incorporating the "latest scientific knowledge." *Id*. § 7408(a)(2). Courts have rejected EPA determinations that there is no need to lower the NAAQS level to protect public health or to provide an adequate margin of safety when the agency has failed to properly consider relevant new evidence. *Am. Farm Bureau Fed'n v. E.P.A.*, 559 F.3d 512, 520 (D.C. Cir. 2009); *see also Am. Lung Ass'n v. EPA*, 134 F.3d 388, 392-93 (D.C. Cir. 1998) (EPA must provide adequate explanation for failure to revise NAAQS in light of relevant evidence); *Lead Indus. Ass'n v. EPA*, 647 F.2d 1130, 1154 (D.C. Cir. 1980) (EPA must "err on the side of caution" in favor of more protective standards when setting NAAQS).

I. EPA SHOULD RECONSIDER THE FINAL RULE IN LIGHT OF ITS FAILURE TO PROVIDE THE REQUISITE PROTECTION OF PUBLIC HEALTH AND WELFARE.

The Final Rule is insufficiently protective of public health and welfare and is the product of arbitrary and capricious EPA decision making. Among other things, the agency arbitrarily gave little to no weight to epidemiological studies quantifying the mortality effects from particulate matter exposure, despite the centrality of that issue to the problem EPA is statutorily obligated to address.

Though there is no real doubt as to the value of the epidemiological studies, EPA chose to ignore them without good reason. As noted by Dr. Cristopher Frey, the head of the Independent Particulate Matter Review Panel and former member of the Clean Air Scientific Advisory Committee (CASAC) Review Panel for the 2012 revised PM NAAQS, "the weight of the evidence in this review is far stronger than what was available in the last review."[9] The Independent Particulate Matter Review Panel further noted that its conclusion that the particulate matter standards are insufficient to protect the public health is based on an *overwhelming* scientific consensus determined by conducting:

> [a] review of the scientific evidence from epidemiologic studies, toxicologic studies in animals, and controlled human exposure studies; this evidence is consistent within each discipline and coherent among the multiple disciplines in supporting a causal, biologically plausible relationship between ambient concentrations well below the current $PM_{2.5}$ standards and adverse health effects, including premature death. The epidemiologic evidence is consistent across studies with diverse designs, populations, pollutant mixtures, locations, and statistical approaches.[10]

Instead, EPA hid behind manufactured "uncertainties" alleged by some of members of the CASAC. 85 Fed. Reg. at 82,708, 82,717. But uncertainty is inherent in any scientific research, and can cut both ways—for instance, mortality effects from particulate matter exposure could well be *worse* than epidemiological studies suggest, not better. The statute explicitly requires EPA to account for such uncertainty when its sets the standards, telling EPA to select a standard that is sufficiently protective "with an adequate margin of safety." 42 U.S.C. § 7409(b)(1). EPA

---

[9] H. Christopher Frey, Ph.D., Public Comment on the Review of the National Ambient Air Quality Standards for Particulate Matter – The NAAQS PM Science Review Process and Outcome is Broken and Not Credible: EPA Should Follow the Science and the Law to Set Health Protective Annual and 24-Hour $PM_{2.5}$ Standards, 70 (June 29, 2020) (EPA-HQ-OAR-2015-0072-1006) [hereinafter Dr. Frey Comments on Proposed Rule].

[10] H. Christopher Frey, Ph.D., Comment Submitted by H. Christopher Frey, Chair, Independent Particulate Matter Review Panel et al. (Attachment 1 – The Need for a Tighter Particulate-Matter Air-Quality Standard) (June 22, 2020) (EPA-HQ-OAR-2015-0072-0669).

cannot use uncertainty alone to justify not lowering the standards while providing for the necessary margin of safety.

The agency also did not try to mitigate or lessen any uncertainty. Instead, EPA refused to meaningfully consider studies published after the Integrated Science Assessment (ISA), considering them only "provisionally." 85 Fed. Reg. 82,690 – 82,691. EPA considered studies published up to March 31, 2017, for analysis in the ISA, with only a "limited" update for studies published up to December 31, 2017. *Id*. 82,690. In limiting the scope of its review, EPA unlawfully failed to adequately consider substantial scientific evidence that clearly demonstrates the need to reduce ambient concentrations of particulate matter. *See Am. Farm Bureau Fed'n,* 559 F.3d at 520.

Further, as the CASAC commented "the Draft ISA does not provide a sufficiently comprehensive, systematic assessment of the available science relevant to understanding the health impacts of exposure to particulate matter." 85 Fed. Reg. 82,689. The CASAC recommended that EPA revise the ISA to consider a more comprehensive assessment of the available data, but EPA chose not to make any revisions, citing as justification the "back-to-basics" memo drafted by former EPA Administrator Scott Pruitt, in which EPA declared it would conduct the particulate matter NAAQS review in a manner to "ensure" any revisions were "finalized" by December 2020, the last month of the Trump Administration. *Id*. The record is clear that, at every turn, EPA had the opportunity and obligation to consider the full scope of the overwhelming evidence warranting the strengthening of the particulate matter NAAQS, but instead chose to ignore such evidence under the guise of "uncertainty."

Those errors and others *might* have been avoided had EPA not arbitrarily disbanded its particulate matter review panel or unlawfully purged CASAC of knowledgeable scientists whose work EPA had previously found deserving of grant funding, or even if EPA had simply given due consideration to the comments of the leading scientists that—having formerly made up the particulate matter review panel—formed the Independent Particulate Matter Review Panel. But EPA's final failure to promulgate appropriately protective particulate matter standards was nonetheless an unforced error, not an inevitability. The many flaws in EPA's process did not ultimately require that the agency arbitrarily ignore the strong scientific data demonstrating the connection between serious health risks and the current particulate matter standards. State Petitioners therefore request that EPA exercise its discretion to convene reconsideration proceedings to engage in a reasoned review of all of the evidence now before the agency, and that EPA promulgate tighter standards to adequately protect both human health, welfare, and the environment, as the Clean Air Act requires. As noted by Dr. Frey, "[i]f EPA followed the science using the same procedures and logic as in the last review, EPA would be proposing to revise the [particular matter NAAQS]."[11]

II. EPA MUST RECONSIDER THE FINAL RULE IN LIGHT OF NEW STUDIES DEMONSTRATING SIGNIFICANT LONG-TERM HEALTH RISKS FROM PARTICULATE MATTER EXPOSURE.

Further, EPA is statutorily required to convene reconsideration proceedings to evaluate important new evidence that arose after the close of public comment and is of central relevance to EPA's decision not to strengthen the NAAQS to adequately protect public health. These studies include those demonstrating significant public health risks at levels below the current

[11] Dr. Frey Comments on Proposed Rule, *supra* note 9, at 70.

NAAQS, in addition to studies linking particulate matter exposure to serious health impacts related to the COVID-19 pandemic and increases in dementia and other severe neurological disease.

A. Petitioners Were Unable to Raise These Objections During the Public Comment Period Because Key Studies Had Not Yet Been Published.

Since the close of public comments in June 2020, new research has raised serious doubt about the accuracy of EPA's prior justification for not strengthening the particulate matter NAAQS. State Petitioners present several critical new studies below, necessitating reconsideration of EPA's prior decision not to establish stronger standards that adequately protect the public health and safety.

1. A Key New Study Refutes EPA's Claim That Health Impacts from Particulate Matter Exposure Below the Current Standard May Be the Result of Unmeasured Confounders.

A new study to be published in March directly refutes EPA's assertion that strengthening the particulate matter NAAQS is unnecessary due to alleged uncertainties and biases in the available research. Public health researchers have determined that $PM_{2.5}$ exposure at levels below the current standard annual standard of 12 $\mu g/m^3$ is associated with increased mortality, concluding "that reducing $PM_{2.5}$ concentrations in the U.S. could save tens of thousands of premature deaths each year.[12] This new study was developed in direct response to the EPA Administrator's dismissal of the epidemiological studies showing a relationship between health impacts and $PM_{2.5}$ exposure.[13] Specifically, the new study responds to EPA's assertion that

---

[13] Joel Schwartz, et al., *supra* note 3, at 17 ("Some scientists . . . assert that studies using standard epidemiological methods should be given little weight in revising the NAAQS, and propose restricting to studies using causal methods .... Their main criticism is that traditional approaches

epidemiological studies "are not necessarily indicative of causal relationships," because such associations "'can reasonably be explained in light of uncontrolled confounding and other potential sources of error and bias.'" 85 Fed. Reg. 82,707. To address the Administrator's unsubstantiated concerns, the researchers developed a "causal method, controlling for temperature and socioeconomic status, and all individual and area level potential confounders, measured or unmeasured, that vary slowly over time."[14] Consistent with the prior epidemiological studies, the researchers found that exposure to $PM_{2.5}$ has significant negative impacts on public health, estimating that reducing the standard by 1 $\mu g/m^3$ could help avoid 239,900 early deaths each year.[15] In other words, the researchers concluded that 1 $\mu g/m^3$ of exposure, at levels below the current standard, for one year "results in an increased risk of dying of 4.26 per ten thousand."[16] Further, because this study's findings support the results of the epidemiological studies dismissed by the Administrator, those "other studies are unlikely to have been confounded by temperature, or slowly varying SES, racial, and behavioral factors which this study controlled for."[17]

---

only show associations that may be confounded... and do not inform causality .... [and] individual characteristics, socioeconomic status, and temperature may confound the published literature. EPA recently proposed not tightening the NAAQS for $PM_{2.5}$ relying on these arguments.").

[14] *Id.* at 17-18.

[15] *Id*. at 19.

[16] *Id*.

[17] *Id*.

2. Another New Study Highlights the Deadly Impacts of $PM_{2.5}$ Exposure and the Significant Health Benefits to Strengthening the NAAQS.

A new global study focusing on $PM_{2.5}$ attributable to the burning of fossil fuels demonstrates that exposure to such fine particulate matter causes significantly more disease and death than previously estimated.[18] The study estimates a global toll of 10.2 million annual premature deaths attributable to $PM_{2.5}$ exposure, including over 350,000 premature annual deaths in the United States, highlighting the astoundingly high health impacts from particulate matter exposure related to controllable sources like fossil fuel combustion. Critically, the study also indicates that decreases in $PM_{2.5}$ exposure at levels below the current NAAQS would substantially reduce mortality rates. At a mean exposure rate of 10 $\mu g/m^3$, each 1 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with a 1.29% increase in mortality, whereas that rate decreased significantly to 0.94% at a mean exposure of 20 $\mu g/m^3$, 0.81% at 30 $\mu g/m^3$, and 0.79% at 40 $\mu g/m^3$.[19] This new evidence supports the conclusion that a reduction in the particulate matter NAAQS below the existing standards will substantially reduce mortality rates from particulate matter exposure and is necessary to protect the public health and safety.

3. Newly Available Studies Link Particulate Matter Exposure to Significant Health Impacts from Respiratory Viruses Like COVID-19.

As of February 16, 2021, over 27 million people in the United States have been infected with COVID-19, a deadly virus that attacks the same respiratory and cardiovascular systems harmed by particulate matter exposure.[20] Despite this unprecedented public health crisis and its

[18] Karn Vohra, et al., *supra* note 4.
[19] *See id*. at supplemental materials 11-12.
[20] COVID Data Tracker, Centers for Disease Control and Prevention (as of Feb. 16, 2021), *available at*: https://covid.cdc.gov/covid-data-tracker/#datatracker-home.

potential long-term health impacts, EPA has determined that no reduction in the particulate matter NAAQS is necessary to protect public health. While preliminary data noted a relationship between COVID-19 and exposure to particulate matter, research published since the close of public comment has clearly established a disturbing link between increased COVID-19 mortality and particulate matter exposure. In light of this new information, EPA must reconsider its decision and strengthen the particulate matter standards.

One study published after the end of public comments found dramatic increases in COVID-19 mortality attributable to exposure to $PM_{2.5}$.[21] The analysis characterized global exposure levels of $PM_{2.5}$ to determine the extent to which such exposure was a cofactor increasing the risk of death from COVID-19. Like the NAAQS, the study focused only on anthropogenic fossil fuel related $PM_{2.5}$ to determine the impact of *avoidable* air pollution on COVID-19 mortality. The findings were stark. In countries with strict air quality standards and lower air pollution like Australia, where the annual $PM_{2.5}$ limit is 8 $\mu g/m^3$, the fraction of COVID-19 mortality attributable to human-made air pollution was 3%.[22] By comparison, in the United States, where the annual $PM_{2.5}$ limit will remain 12 $\mu g/m^3$ under EPA's existing action, the COVID-19 mortality attributable to human-made air pollution was 18%, *six times higher* than the percentage in Australia. A second new study found a direct link between the amount of a region's long-term $PM_{2.5}$ exposure and its COVID-19 mortality rate.[23] Analyzing county-level data in the United States, the study found that a mere 1 $\mu g/m^3$ increase in the long-term average

[21] A. Pozzer et al., *Regional and Global Contributions of Air Pollution to Risk of Death from COVID-19*, Cardiovascular Research, Sept. 30, 2020, *available at*: https://academic.oup.com/cardiovascres/advance-article/doi/10.1093/cvr/cvaa288/5940460.
[22] *See* Andrea. Pozzer, et al., *supra* note 5, at Supplementary Table, Table S1.
[23] Xiao Wu, et al., *supra* note 5, at 1.

$PM_{2.5}$ is associated with a statistically significant 11% increase in the county's COVID-19 mortality rate.

Such studies serve to highlight the enormous health benefits from reducing particulate matter exposure, which include protecting those who may continue to battle health impacts related to COVID-19's devastation of respiratory systems over the long-term, or "long-haulers."[24] While uncertainty remains about COVID-19's long-term health impacts, initial studies have found instances of severely impaired pulmonary capacities months after initial infections.[25] As the nation confronts the lasting effects of the COVID-19 pandemic, and addresses new mutations and future viruses, EPA also must consider whether the particulate matter NAAQS is sufficiently protective of the public health in light of the long-term respiratory problems faced by recovering patients.

Despite the ongoing respiratory pandemic impacting millions of Americans, EPA's only acknowledgment of COVID-19 comes in a few sentences in its response to comments. Notably, it does not address any of the numerous relevant studies that came out after the close of the comment period, several of which were submitted by State Petitioners in its supplemental letter. It does make the remarkable and wholly unsupported claim that early studies tying COVID-19 to particulate matter exposure were somehow "generally consistent with the evidence assessed in

---

[24] *See, e.g.*, Rita Rubin, *As Their Numbers Grow, COVID-19 "Long Haulers" Stump Experts*, JAMA (Sept. 23, 2020), *available at*: https://jamanetwork.com/journals/jama/fullarticle/2771111.

[25] *See* Chaolin Huang, et al., *6-Month Consequences of COVID-19 in Patients Discharged from Hospital: A Cohort Study*, 397 THE LANCET 220-232 (Jan. 8, 2021), *available at*: https://www.thelancet.com/action/showPdf?pii=S0140-6736%2820%2932656-8. *See also* European Lung Foundation, *COVID-19 Patients Suffer Long-Term Lung and Heart Damage but it Can Improve with Time*, SCIENCE DAILY (Sept. 6, 2020), *available at*: https://www.sciencedaily.com/releases/2020/09/200906202950.htm.

the [Integrated Science Assessment]," Response to Comments, p. 37, despite the fact that the analysis referenced was completed *before* the first reported COVID-19 outbreak.

4. New Studies Show Increases in Serious Neurodegenerative Diseases from Exposure to Particulate Matter.

An important new study demonstrating that increased particulate matter exposure exacerbates the risk of serious neurological disorders also warrants that EPA reconsider its decision not to lower the standards.[26] The study adds critical new evidence to emerging links between particulate matter air pollution and neurodegenerative conditions suffered by millions of American residents, including impaired cognitive function, accelerated cognitive decline, Parkinson's disease, Alzheimer's disease, and dementia.[27] This nationwide, long-term study found significant correlations between increased $PM_{2.5}$ exposure and first-time hospital admittances for diagnoses of Parkinson's disease, Alzheimer's disease, and related dementias in patients over 65.[28] Researchers found a strong linear relationship between incidences of neurodegenerative disease and increases in $PM_{2.5}$ exposure. This linear relationship was observed at $PM_{2.5}$ exposures down to 7$\mu g/m^3$—well below the current standard, indicating that lowering the $PM_{2.5}$ standard will have meaningful public health benefits across the country. This study released after the close of public comment further demonstrates the need for EPA to strengthen the particulate matter NAAQS to protect the State Petitioners' populations from the devastating impacts of serious cognitive disorders.

[26] Liuhua Shi et al., *supra* note 6.
[27] *Id*. at E564.
[28] *Id.* at E557.

B. The Identification of Serious New Threats to Health from Particulate Matter Exposure Are of Central Relevance to EPA's Unlawful Decision.

New evidence of decreased mortality at levels of $PM_{2.5}$ exposure below existing standards and the existence of significant new threats to public health is of "central relevance" to EPA's decision not to strengthen the particulate matter NAAQS, obligating EPA to reconsider the decision. The additional studies clearly "provide substantial support for the argument that the regulation should be revised." *Chesapeake Climate Action Network*, 952 F.3d at 322. To justify its decision that strengthening the particulate matter NAAQS is unnecessary to protect public health with an adequate margin of safety, EPA's decision *repeatedly* cites the "uncertainty" of the health benefits from the studies reviewed.[29] These new studies dispel that alleged uncertainty.

As noted above, EPA's decision rests significantly on its dismissal of several critical epidemiological studies due to assertions that they do not necessarily prove causal relationships between negative health impacts and levels of particulate matter below the current NAAQS standard, pointing to "uncontrolled confounding and other potential sources of error and bias.'" 85 Fed. Reg. 82,707. But the new study by Schwartz et al. cited above successfully demonstrates that uncontrolled factors were not the source of any error and bias, solidifying the link between particulate matter exposure and serious negative health impacts. The study is of central relevance because it addresses and refutes the primary rationale the Administrator relied on to not strengthen the current particulate matter NAAQS, and it also demonstrates significant public

[29] *See* 85 Fed. Reg. 82,685 ("For the primary $PM_{2.5}$ standards, the Administrator concludes that there are *important uncertainties in the evidence* for adverse health effects below the current standards and in the potential for additional public health improvements from reducing ambient $PM_{2.5}$ concentrations below those standards… [The Administrator] concludes that, based on the newly available evidence *with its inherent uncertainties*, the current primary PM_{10} standard is requisite to protect public health, with an adequate margin of safety, from effects of PM_{10} in ambient air, and should be retained, without revision" [emphasis added].)

health impacts from exposure to particulate matter below the current standard. The study conducted by Vohra et al. also cited above further demonstrates the likelihood of substantial public health improvements below the current $PM_{2.5}$ standards and the need for EPA to strengthen the particulate matter NAAQS. Even if purported uncertainties were originally a valid basis for EPA's decision, these new studies substantially lessen any uncertainty, undercutting the basis for EPA's decision and requiring reconsideration.

Further, whatever margin of safety that may be embodied in the current NAAQS is significantly reduced, if not eliminated entirely, once the relationship to COVID-19 and increases in neurocognitive diseases are considered. In light these facts, EPA must strengthen the NAAQS to a level that reintroduces the necessary buffer to provide an adequate margin of safety as required by the Clean Air Act. Even if the COVID-19 pandemic somehow subsides entirely, the long-term respiratory impacts noted above will likely leave significant portions of the public more vulnerable to particulate matter harms. Likewise, the links between particulate matter exposure and devastating neurological diseases like Alzheimer's and Parkinson's, as established in the Shi et al. study discussed above, are particularly relevant to the millions of Americans over 65, a population that is expected to double by 2060. Such a demographic shift in the American population will substantially increase the amount of harm from particulate matter exposure suffered by the overall population.[30] EPA must open reconsideration proceedings in order to fully evaluate these significant negative public health impacts unaddressed in the current particulate matter NAAQS.

---

[30] Kevin A. Matthews, et al., *Racial and Ethnic Estimates of Alzheimer's Disease and Related Dementias in the United States (2015–2060) in Adults Aged ≥ 65 Years*, 15 ALZHEIMER'S & DEMENTIA 17-24, (Jan. 2019), *available at*: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6333531/.

The new studies also present further evidence that EPA's decision will only exacerbate existing environmental injustices. As the pandemic enters its second year, COVID-19 has been shown to have a devastating impact on the very same disadvantaged communities already struggling from high particulate matter exposure.[31] Similarly, Alzheimer's disease and related dementias also disproportionately impact the Black and Latinx populations, compared to the non-Hispanic White population.[32] In accordance with President Biden's recent Executive Order committing to address environmental justice[33] and the longstanding disproportionate impacts federal policies have on historically marginalized and overburdened BIPOC communities, EPA should convene reconsideration proceedings in order to address the environmental injustice created by disproportionate exposure to particulate matter.

RELIEF REQUESTED

For the foregoing reasons, State Petitioners respectfully request that the Administrator immediately convene proceedings for reconsideration of the final action. 42 U.S.C. § 7607(d)(7)(B).

[31] Gregorio A. Millett MPH, et al., *Assessing Differential Impacts of COVID-19 on Black Communities*, 47 Annals of Epidemiology 37-44 (Jul. 2020), *available at*: https://doi.org/10.1016/j.annepidem.2020.05.003.
[32] Kevin A. Matthews et al., *supra* note 30, at 17-24.
[33] *See* Executive Order on Tackling the Climate Crisis at Home and Abroad (Jan. 27, 2021) ("Agencies shall make achieving environmental justice part of their missions by developing programs, policies, and activities to address the disproportionately high and adverse human health, environmental, climate-related and other cumulative impacts on disadvantaged communities, as well as the accompanying economic challenges of such impacts.") *Available at*: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/executive-order-on-tackling-the-climate-crisis-at-home-and-abroad/.

Dated: February 16, 2021

Respectfully Submitted,
FOR THE STATE OF CALIFORNIA

XAVIER BECERRA
Attorney General of California
ROBERT W. BYRNE
Senior Assistant Attorney General
DAVID A. ZONANA
Acting Senior Assistant Attorney General
MYUNG J. PARK
Supervising Deputy Attorney General
JONATHAN A. WIENER
SPARSH KHANDESHI
COREY MOFFAT
Deputy Attorneys General

/s/ Scott J. Lichtig
SCOTT J. LICHTIG
Deputy Attorney General
1300 I Street
Sacramento, CA 95814
Tel: (916) 210-7815
Scott.Lichtig@doj.ca.gov

Attorneys for Petitioner State of California, by and through its Attorney General Xavier Becerra, and California Air Resources Board

FOR THE STATE OF CONNECTICUT

WILLIAM TONG
Attorney General of Connecticut

/s/ Jill Lacedonia
JILL LACEDONIA
Assistant Attorney General
Office of the Attorney General
165 Capitol Avenue
Hartford, CT 06106
(860) 808-5250
Jill.Lacedonia@ct.gov
Counsel for the State of Connecticut

FOR THE STATE OF DELAWARE

KATHLEEN JENNINGS
Attorney General of Delaware

/s/ Christian Douglas Wright
CHRISTIAN DOUGLAS WRIGHT
Director of Impact Litigation
VALERIE EDGE
Deputy Attorney General
Delaware Department of Justice
102 W. Water Street, 3rd Floor
Dover, DE 19902
(302) 257-3219
Christian.Wright@delaware.gov
Counsel for the State of Delaware

FOR THE STATE OF ILLINOIS

KWAME RAOUL
Attorney General of Illinois

/s/ *Daniel I. Rottenberg*
MATTHEW J. DUNN
Chief, Environmental Enforcement/Asbestos Litigation Division
DANIEL I. ROTTENBERG
Assistant Attorney General
Office of the Attorney General
69 W. Washington St., 18th Floor
Chicago, IL 60602
(312) 814-3816
DRottenberg@atg.state.il.us
Counsel for the State of Illinois

FOR THE STATE OF MARYLAND

BRIAN E. FROSH
Attorney General of Maryland

/s/ Joshua M. Segal
JOSHUA M. SEGAL
Special Assistant Attorney General
Office of the Attorney General
200 St. Paul Place
Baltimore, MD 21202
(410) 576-6446
jsegal@oag.state.md.us
Counsel for the State of Maryland

FOR THE COMMONWEALTH OF MASSACHUSETTS

MAURA HEALEY
Attorney General of Massachusetts

/s/ Megan M. Herzog
CHRISTOPHE COURCHESNE
Assistant Attorney General and Deputy Chief
TURNER SMITH
Assistant Attorney General
MEGAN M. HERZOG
Special Assistant Attorney General
Energy and Environment Bureau
Office of the Attorney General
One Ashburton Place, 18th Fl.
Boston, MA 02108
(617) 727-2200
turner.smith@mass.gov
Counsel for the Commonwealth of Massachusetts

FOR THE STATE OF MINNESOTA

KEITH ELLISON
Attorney General of Minnesota

/s/ *Leigh K. Currie*
LEIGH K. CURRIE
Special Assistant Attorney General
Minnesota Office of the Attorney General
445 Minnesota Street Suite 900
Saint Paul, MN 55101
(651) 757-1061
peter.surdo@ag.state.mn.us
Counsel for the State of Minnesota

FOR THE STATE OF NEW JERSEY

GURBIR GREWAL
Attorney General of New Jersey

/s/ Daniel Resler
DANIEL RESLER
Deputy Attorney General
Division of Law
R.J. Hughes Justice Complex
25 Market Street, P.O. Box 093
Trenton, NJ 08625
Tel: (609) 376-2735
Email: Daniel.Resler@law.njoag.gov
Counsel for the State of New Jersey

FOR THE STATE OF NEW YORK

LETITIA JAMES
Attorney General of New York

/s/ Andrew G. Frank
MICHAEL J. MYERS
Senior Counsel
NICHOLAS C. BUTTINO
ANDREW G. FRANK
Assistant Attorneys General
LINDA M. WILSON
Staff Scientist
New York State Attorney General Office
28 Liberty Street
New York, New York 10005
(212) 416-8271
Nicholas.Buttino@ag.ny.gov
Counsel for the State of New York

FOR THE CITY OF NEW YORK

JAMES E. JOHNSON
Corporation Counsel for the City of New York

/s/ Christopher Gene King
Christopher Gene King
Senior Counsel
New York City Law Department
100 Church Street
New York, NY 10007
(917) 941-5603
cking@law.nyc.gov
Counsel for the City of New York

FOR THE STATE OF OREGON

ELLEN F. ROSENBLUM
Attorney General of Oregon

/s/ Paul Garrahan
PAUL GARRAHAN
Attorney-in-Charge
STEVE NOVICK
Special Assistant Attorney General
Natural Resources Section
Oregon Department of Justice
1162 Court Street NE
Salem, OR 97301-4096
(503) 947-4593
paul.garrahan@doj.state.or.us
steve.novick@doj.state.or.us
Counsel for the State of Oregon

FOR THE COMMONWEALTH OF PENNSYLVANIA

JOSH SHAPIRO
Attorney General of Pennsylvania

/s/ Ann Johnston
ANN JOHNSTON
Senior Deputy Attorney General
MICHAEL J. FISCHER
Chief Deputy Attorney General
Office of Attorney General
Strawberry Square, 14th Floor
Harrisburg, Pennsylvania 17120
(717) 705-6938
ajohnston@attorneygeneral.gov
Counsel for the Commonwealth of Pennsylvania

FOR THE STATE OF RHODE ISLAND

PETER F. NERONHA
Attorney General of Rhode Island

/s/ Gregory S. Schultz
GREGORY S. SCHULTZ
Special Assistant Attorney General
Rhode Island Office of Attorney General
150 South Main Street
Providence, RI 02903
(401) 274-4400
gschultz@riag.ri.gov
Counsel for the State of Rhode Island

FOR THE STATE OF VERMONT

THOMAS J. DONOVAN, JR.
Attorney General of Vermont

/s/ Nicholas F. Persampieri
NICHOLAS F. PERSAMPIERI
Assistant Attorney General
Office of the Attorney General
109 State Street
Montpelier, VT 05609
(802) 828-3171
nick.persampieri@vermont.gov
Counsel for the State of Vermont

FOR THE COMMONWEALTH OF VIRGINIA

MARK R. HERRING
Attorney General of Virginia

/s/ Caitlin C. G. O'Dwyer
PAUL KUGELMAN, JR.
Senior Assistant Attorney General,
Chief, Environmental Section
CAITLIN C. G. O'DWYER
Assistant Attorney General
Office of the Attorney General
202 North 9th Street
Richmond, Virginia 23219
(804) 786-1780
godwyer@oag.state.va.us
Counsel for the Commonwealth of Virginia

FOR THE STATE OF WASHINGTON

ROBERT W. FERGUSON
Attorney General of Washington

/s/ Christopher H. Reitz
CHRISTOPHER H. REITZ
Assistant Attorney General
Office of the Attorney General
P.O. Box 40117
Olympia, Washington 98504-0117
(360) 586-4614
chris.reitz@atg.wa.gov
Counsel for the State of Washington

FOR THE STATE OF WISCONSIN

JOSHUA L. KAUL
Attorney General of Wisconsin

/s/ Lorraine C. Stoltzfus
LORRAINE C. STOLTZFUS
EMILY M. ERTEL
Assistant Attorneys General
Wisconsin Department of Justice
Post Office Box 7857
Madison, WI 53707-7857
(608) 266-9226
stoltzfuslc@doj.state.wi.us
Counsel for the State of Wisconsin

A National Difference in Differences Analysis of the Effect of $PM_{2.5}$ on Annual Death Rates

Joel Schwartz, Yaguang Wei, Ma'ayan Yitshak-Sade, Qian Di, Francesca Dominici, Antonella Zanobetti

PII: S0013-9351(20)31546-2

DOI: https://doi.org/10.1016/j.envres.2020.110649

Reference: YENRS 110649

To appear in: *Environmental Research*

Received Date: 10 November 2020

Revised Date: 16 December 2020

Accepted Date: 17 December 2020

Please cite this article as: Schwartz, J., Wei, Y., Yitshak-Sade, M.'a., Di, Q., Dominici, F., Zanobetti, A., A National Difference in Differences Analysis of the Effect of $PM_{2.5}$ on Annual Death Rates, *Environmental Research*, https://doi.org/10.1016/j.envres.2020.110649.

A National Difference in Differences Analysis of the Effect of $PM_{2.5}$ on Annual Death Rates

Joel Schwartz[1,2],Yaguang Wei[1], Ma'ayan Yitshak-Sade[1,3], Qian Di[1,4], Francesca Dominici[5], Antonella Zanobetti[1]

1. Department of Environmental Health, Harvard TH Chan School of Public Health
2. Department of Epidemiology, Harvard TH Chan School of Public Health
3. Department of Environmental Medicine, Mount Sinai School of Medicine
4. Vanke School of Public Health, Tsinghua University, Beijing, China
5. Department of Biostatistics, Harvard TH Chan School of Public Health

Address communications to:

Joel Schwartz joel@hsph.harvard.edu. 401 Park Drive, Suite 404H, Boston MA 02215.

Abstract

Many studies have reported that $PM_{2.5}$ was associated with mortality, but these were criticized for unmeasured confounding, not using causal modeling, and not focusing on changes in exposure and mortality rates. Recent studies have used propensity scores, a causal modeling approach that requires the assumption of no unmeasured confounders.

We used differences in differences, a causal modeling approach that focuses on exposure changes, and controls for unmeasured confounders by design to analyze $PM_{2.5}$ and mortality in the U.S. Medicare population, with 623,036,820 person-years of follow-up, and 29,481,444 deaths. We expanded the approach by clustering ZIP codes into 32 groups based on racial, behavioral and socioeconomic characteristics, and analyzing each cluster separately. We controlled for multiple time varying confounders within each cluster. A separate analysis examined participants whose exposure was always below 12 $\mu g/m^3$.We found an increase of 1 $\mu g/m^3$ in $PM_{2.5}$ produced an increased risk of dying in that year of 3.85×10^{-4} (95% CI 1.95 $\times 10^{-4}$, 5.76 $\times 10^{-4}$). This corresponds to 14,000 early deaths per year per 1 $\mu g/m^3$. When restricted to exposures below 12 $\mu g/m^3$, the increased mortality risk was 4.26×10^{-4} (95% CI 1.43×10^{-4}, 7.09×10^{-4}). Using a causal modeling approach robust to omitted confounders, we found associations of $PM_{2.5}$ with increased death rates, including below U.S. and E.U. standards.

Key words: Air pollution, $PM_{2.5}$, causal, difference in differences, mortality

Authors contributions: JS conceived the approach, developed the methodology, and did the analysis and reviewed and edited the draft manuscript; YW wrote the first draft and prepared the data; MY-W helped with the programming, methodology and commented on the draft; FD obtained funding and the Medicare data; AZ obtained funding.

Funding: This study was funded by NIH grants ES000002 and ES024332, and U.S. EPA grant RD-83615601. The funders had no input into the design or conduct of the study.

Dr. Schwartz declares that he has testified as a health expert for the U.S. Department of Justice in a law suit over clean air act violations. No other authors have anything to declare.

This study was approved by the Human Subjects Committee of the Harvard School of Public Health.

1. Introduction

The Clean Air Act requires the US Environmental Protection Agency (EPA) to set National Ambient Air Quality Standards (NAAQS) to protect vulnerable populations with an adequate margin of safety. Many studies have reported associations of $PM_{2.5}$ and mortality and morbidity following long and short-term exposure(Abu Awad et al., 2019; Beelen et al., 2014; Crouse et al., 2015; Di et al., 2017; Hoek et al., 2013; Pinault et al., 2016; Pope et al., 2019; Vodonos et al., 2018). These were undertaken by many investigators with over 50 cohorts in the most recent $PM_{2.5}$ meta-analysis(Vodonos et al., 2018), and have resulted in EPA sequentially tightening the $PM_{2.5}$ standard. The global burden of disease ranks air pollution among the largest public health risks.

Recent studies have reported associations between $PM_{2.5}$ and mortality at concentrations below the 2012 U.S. EPA NAAQS or World Health Organization air quality guidelines(Di et al., 2017; Wang et al., 2016; Yang et al., 2012). However, some have criticized many of these studies for not using causal modeling approaches.

Causal modeling methods can aid in assessing causality. The general approach is to try to make an observational study closely mimic a randomized trial. In addition, causal methods provide marginal estimates of the effects of exposure, that do not depend on the distribution of the covariates in the study population(Imai and van Dyke, 2004). A common approach is to use propensity score matching or inverse probability weighting to make the exposure independent of all measured confounders(Baccini et al., 2017; Rubin, 1997). Recent studies have used that approach to examine the association of $PM_{2.5}$ with mortality, and provided robust findings(Abu Awad et al., 2019; Schwartz et al., 2018; Wei et al., 2020; Wu et al., 2020; Yitshak-Sade et al., 2019). However, propensity scores only control for measured confounders, and therefore do not address the

argument that there is unmeasured confounding. Hence it is important to complement that approach using methods that can address unmeasured confounders.

Approaches that control for unmeasured confounders by design include difference-in-differences (DID) analyses (Wang et al., 2016; Yitshak-Sade et al., 2019). In a classical DID model, the mean response is calculated for the exposed and non-exposed groups in pre-exposure and post-exposure periods. Since all slowly varying predictors of outcome such as socioeconomic status (SES), smoking, obesity, etc. are the same in each group in both periods, the difference between outcomes in the two periods in the exposed group cannot be confounded by those variables. The difference between pre-exposure and post-exposure periods in the unexposed group is a negative outcome control for the difference in the exposed group. It controls for changes in an outcome due to covariates that can change between periods similarly between the two locations. The difference in these pairs of differences is a causal estimate, assuming that no other exposure has affected the two groups differently over time(Donald and Lang, 2007). Because the DID approach examines the effect of changes in exposure (post vs pre periods) on change in outcome, it is precisely the type of study that EPA's CASAC says it prefers. The method has been generalized to look at more than two locations, more than two time periods, and continuous, time varying exposures(Wang et al., 2016). With multiple locations time-invariant omitted confounders are controlled using an indicator variable for each location. However, it still require the assumption that changes in mortality rates by year due to changing risk factors are common across locations. Here we simultaneously adopt two approaches to relax that assumption, and hence strengthen the evidence for causality. We apply them to assess whether changes in $PM_{2.5}$ are associated with changes in mortality rates in a national cohort of Medicare participants in the U.S. In addition, as

few previous cohort studies have controlled for temperature, we adjusted for mean warm season and mean cold season temperature.

2.0 Data and Methods

2.1 Medicare cohort

We obtained the Medicare beneficiary denominator file, which contains information on all Medicare participants in the U.S., from the Centers for Medicare and Medicaid Services(RESDAC, 2018). We followed all beneficiaries' ≥65 years in the contiguous U.S. from 2000 to 2016. Medicare insurance covers over 95% of the population ≥65 years of age in the United States. Medicare participants alive on January 1 of the year following their enrollment in Medicare entered the open cohort, and follow-up periods were calendar years. For the DID analysis, we computed an annual mortality rate in each ZIP code, in each group stratified by age (>84 or not), sex, race, and Medicaid coverage. This study was approved by the human subjects committee at the Harvard T. H. Chan School of Public Health.

2.2 Covariates

From the Medicare denominator file for each calendar year, we obtained the age, sex, race, ZIP code of residence for that year, eligibility for Medicaid for that year, and date of death (or censoring) of each participant. Age, ZIP code, and Medicaid eligibility were updated annually. 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 and 2010 Census Summary File 3(Bureau, 2010) at the ZIP code tabulation–area level (ZCTA) and the American Household Survey for each year after 2010. These included percent of people ≥ 65 living in poverty, median household income, median house value, percent

of owner occupied homes, percent black, percent Hispanic, population density, and education. We updated these variables for missing years by linearly extrapolating between the measured years. In addition, the county-level percentage of people who ever smoked and their mean body mass index (BMI) 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 (LDLC) test, a mammogram, an eye exam, and a visit to an annual checkup for each year in each hospital catchment area and assigned it to all ZCTAs in that area(Wennberg and Cooper, 1996). We also computed the distance from each ZIP code centroid to the nearest hospital. To capture long-term smoking history of Medicare participants in each ZIP code, we computed their hospitalization rate for lung cancer by ZIP code for each year. This risks over-control because air pollution has been associated with increased risk of lung cancer. To capture year-to-year changes in mortality rates due to temperature, we downloaded daily temperature data on a 12km grid from the NASA NLDAS-2 website (https://ldas.gsfc.nasa.gov/index.php/nldas/v2/models). We averaged all grid cells within the boundaries of a ZIP code, and constructed two measures for each year, the average temperature in the warm months (April-September) and in the cold months (October-March).

2.3 Exposure assessment

We estimated exposure using a validated prediction model calibrated to measurements at almost 2000 monitoring stations using an ensemble of machine learners that provided daily estimates for a 1km grid of the contiguous U.S.(Di et al., 2019; Di et al., 2020). In brief, the model used data from multiple sources including predictions of chemical transport models (GEOS-Chem, CMAQ, and MERRA-2), meteorological data, land-use terms, and satellite-based measures of aerosol

optical depth, surface reflectance, and absorbing aerosol index. We trained a neural network, a random forest, and a gradient boosting machine to monitoring data from the United States Environmental Protection Agency (EPA) Air Quality System to generate daily predictions on a 1×1 km grid. The models were fit using data from all years. The three predictions for $PM_{2.5}$ were combined in a nonlinear geographically weighted regression. The model showed good performance with ten-fold cross validation on held out monitoring sites yielding an out of sample R^2 of 0.89 for annual average predictions of $PM_{2.5}$. Penalized splines showed linear relationships between observed and predicted $PM_{2.5}$ from 0 to 60 μg/m^3. Predictions for all grid cells whose centroids were inside the ZIP code boundary were averaged for each year and assigned to participants in that ZIP code in that year.

2.4 Statistical analysis

The standard DID estimator for a continuous predictor posits that

$$E(Y_{ij}) = \beta_0 + \beta_1 PM_{2.5} + \beta_2 C_i + \beta_3 X_t \quad (1)$$

where Y_{ij} is the mortality rate in ZIP code i in demographic group (by age >84 or not, sex, race, and Medicaid coverage) j, C_i are the time-invariant or slowly changing confounders in ZIP code i, X_t are the time varying confounders that are common across ZIP codes. The C_i are controlled by fitting individual intercepts for each ZIP code. The time varying confounders are removed by fitting a nonlinear time trend; we used a natural spline function of year with 3 degrees of freedom, yielding:

$$E(Y_{ij}) = \beta_0 + \beta_1 PM_{2.5} + \delta_i + ns(year, 3) \quad (2)$$

where δ_i is a dummy variable for each ZIP Code. Since ZIP code is controlled, this model compares year-to-year variations around ZIP code average $PM_{2.5}$ and common time trend to year-

to-year variations of mortality rates about ZIP code average and common time trend. Differences in e.g. SES, smoking, or diabetes between ZIP code are removed by the dummy variable for ZIP code. For a causal interpretation of the DID estimate to hold, we must assume that all the ZIP codes have parallel long-term time trends in mortality rates, other than those caused by different time trends in $PM_{2.5}$.

If covariates producing different time trends in mortality rate by ZIP code are not correlated with ZIP code specific $PM_{2.5}$ trends, the interpretation still holds. It would be preferable to further weaken this assumption.

We added two methods to relax the parallel trends assumption, and combined them in our analysis. We added to equation (2) terms for confounders that we have measured that change over time, possibly differentially by ZIP code. This will control for any temporal trends due to changes in these covariates, which include the SES, race, demographic, behavioral, and health access variables described above. Second, we grouped ZIP codes based on the above covariates and fit separate time trends in each group. We think that ZIP codes that are similar in racial composition, percent living below the poverty level, population density, smoking rates etc. are more likely to have similar time trends in mortality rates than disparate ZIP codes. To accomplish this, we fit a principal component analysis to all the listed potential confounders and took the first 5 principal components. We classified each ZIP code into whether it was higher or lower than average on each of the 5 components, producing 32 categories of ZIP codes. In each of these 32 categories of ZIP codes, we fit separate splines for time trend and separate control for all of the covariates. This controls for time trends in measured covariates such as racial composition, median income, etc, and fits 32 separate time trends to the data to capture any trends unexplained by time trends in the measured covariates. It also allows the effects of the measured covariates to differ by the 32

different groups. In addition by performing analyses stratified by the 32 different groups we are also controlling by matching for the covariate clusters (e.g. SES and race) that characterize each group. Combining these, the final modeling approach is to fit 32 models (k=1:32) and meta-analyze the 32 values of β_1.

$$E(Y_{ij}, k) = \beta_0 + \beta_{1k}PM25 + \delta_i + \delta_{jk} + ns(year, 3, k) + \beta_3 X_{tk} \quad (3)$$

Here δ_i are indicator variables for each ZIP code, X_{tk} are the time varying covariates whose time trends may differ by ZIP code and by which group (k) the ZIP code is in, $ns(year, 3, k)$ is a natural spline for time trend with 3 degrees of freedom for each group k, and δ_{jk} is an indicator variable for each age-race-sex-Medicaid stratum in group k. Results were combined over strata using a random effects meta-analysis.

Finally, equation (3) embodies an additive, rather than multiplicative model for the rate of mortality in each ZIP code-demographic group. This allows us to estimate the additive effect of $PM_{2.5}$ on the probability of dying, provides more interpretative interaction terms, and provides a marginal effect estimate (i.e. not dependent on the distribution of the covariates, as a multiplicative model would be). Additive probability or rate models give unbiased estimates of effect just as the more usual logistic models, but biased estimates of standard errors(Caudill and Jackson, 1989). Therefore have used robust standard errors to estimate the confidence intervals. In a second analysis, we reran the analysis on data restricted to persons whose exposure was always below 12 $\mu g/m^3$, the U.S. standard for $PM_{2.5}$.

3.0 Results

Table 1 shows the characteristics of the Medicare cohort between 2000 and 2016. There were 623,036,820 person-years of follow-up during the study, and 29,481,444 deaths. 85.4% of the

participants were white, and 12.9% were covered by Medicaid, which provides additional benefits to the poor. The mean $PM_{2.5}$ during the study was 10.3 $\mu g/m^3$.

Table 1

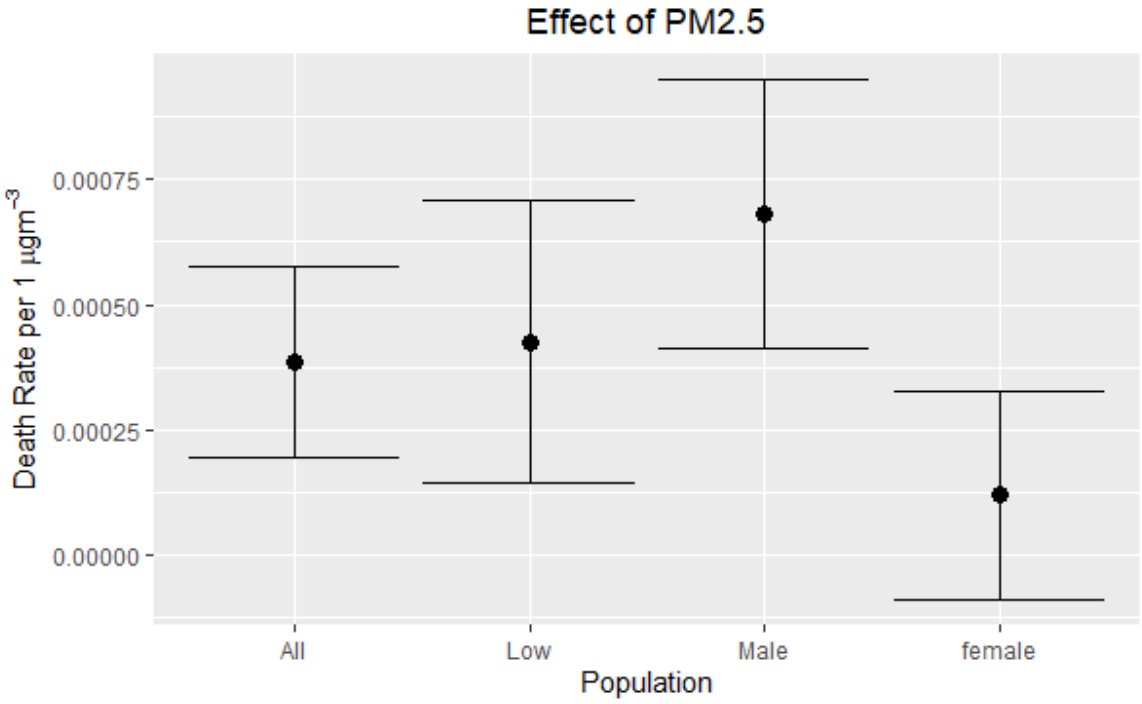
Variable	Values
Year	
Mean (SD)	2009.5 (4.90)
Median [25%, 75%]	2010 [2004, 2013]
Male	42.8 %
Race	
Black	8.4%
Other	6.2%
White	85.4%
Age > 84	13.2%
Medicaid Coverage	12.9%
ZIP Code Covariates	
Median Income	
Mean (SD)	$53,177 ($22,082)
Median [25%, 75%]	$47,998 [$38,030, $63,031]

Median House Value	
Mean (SD)	$200,139 ($159,728)
Median [25%, 75%]	$150,400 [$98,600, $240,300]
Percent ZIP code Black	
Mean (SD)	11% (17.9%)
Median [25%, 75%]	3.7% [1.1%, 12.0%]
Percent ZIP code Hispanic	
Mean (SD)	12.6% (16.4%)
Median [25%, 75%]	5.3% [2.1%, 14.6%]
Percent >65 below poverty	
Mean (SD)	9.5% (6.5%)
Median [25%, 75%]	7.9% [5.3%, 11.8%]
Percent Low Education	
Mean (SD)	25.3% (14.7%)
Median [25%, 75%]	22.6% [14.2%, 33.7%]
Percent with annual Mammogram	
Mean (SD)	63.7% (7.2%)
Median [25%, 75%]	63.9% [59.2%, 68.2%]

Percent with ambulatory Visit	
Mean (SD)	77.8% (6.2%)
Median [25%, 75%]	79.0% [74.4%, 82.1%]
Population Density (persons/mi^2)	
Mean (SD)	3397 (9032)
Median [25%, 75%]	967 [167, 3353]
Percent Owner Occupied	
Mean (SD)	68% (16%)
Median [25%, 75%]	70.8% [59.8%, 79.2%]
Mean BMI (kg/m^2)	
Mean (SD)	27.5 (1.58)
Median [25%, 75%]	27.3 [26.7, 28.0]
Distance to nearest hospital (km)	
Mean (SD)	6.5 (7.4)
Median [25%, 75%]	3.90 [1.98, 8.07]
Ever Smoker	
Mean (SD)	46.2% (6.8%)
Median [25%, 75%]	46.2% [41.8%, 50.4%]

Percent annual HbA1c test	
Mean (SD)	83.1% (4.9%)
Median [25%, 75%]	83.7% [80.5%, 86.3%]
Lung Cancer Rate (x 10^{-4})	
Mean (SD)	3.9 (2.8)
Median [25%, 75%]	3.3 [1.9, 4.9]
Percent annual LDL	
Mean (SD)	79.5% (6.2%)
Median [25%, 75%]	80.1% [76.1%, 83.5%]
Percent Annual Eye Exam	
Mean (SD)	67.4% (6.4%)
Median [25%, 75%]	67.1% [63.9%, 71.0%]
$PM_{2.5}$ ($\mu g/m^3$)	
Mean (SD)	10.3 (3.1)
Median [25%, 75%]	9.8 [7.9, 12.0]

In the meta-analysis of the results of the 32 strata-specific DID analyses, we found that the probability of dying in each year increased by $3.85x10^{-4}$ (95% CI $1.95 x10^{-4}$, $5.76 x10^{-4}$) for each 1 $\mu g/m^3$ increment in $PM_{2.5}$ in that year. The I^2 statistic for heterogeneity was 42%. When we restricted our analysis to persons whose exposure was always below 12 $\mu g/m^3$ during the follow-up period, we found a larger effect size, with the probability of dying in each year increased by 4.26 $x10^{-4}$ (95% CI $1.43x10^{-4}$, $7.09 x10^{-4}$) per 1 $\mu g/m^3$ increase in $PM_{2.5}$. Interaction terms for male gender, age > 84, and race were fit in the full data. Sex was a significant modifier (p for interaction <0.001), with larger effects in males (6.81x10-^{4}, 95% CI $4.14x10^{-4}$, $9.48x10^{-4}$) then females($1.20x10^{-4}$, 95% CI $-8.80x10^{-5}$, $3.29x10^{-4}$). These results are shown in Figure 1. There was no interaction by age. Interaction models for race did not converge because residential segregation in the U.S. resulted in groups with too few Blacks or Asians and other races. Instead, we reran the analysis without separate models for each of the 32 groups. There was no significant interaction by race.



4.0 Discussion

Using a difference in differences design applied to a linear rate model, we found that each 1 $\mu g/m^3$ increment in $PM_{2.5}$ was associated with a 3.85 $x10^{-4}$ increase in the probability of dying in a given year. If the difference in differences assumptions are met, this is a causal increase. We believe they are met for the following reasons. First, since this design controls for each ZIP code, all individual and neighborhood level confounders that change little over time are controlled, whether measured or not. This includes most of the variables (e.g. SES, smoking history, diet) that have be posited as potential confounders. Consequently, only time varying factors can be confounders. Second, we controlled for potential confounding in each ZIP code due to time trends in median household income, median home value, percent owner occupied housing, percent of ZIP code that is Black, percent of households that are Hispanic, percent of persons aged 65 or older living in poverty, smoking rate, BMI, Medicaid eligibility, educational attainment, population density, lung cancer rates, multiple measures of the adequacy of medical care, and summer and winter temperature. Third, we grouped the ZIP codes by these factors, and fit separate nonlinear time trends within each of 32 groups to capture any remaining time trends due to omitted confounders could differ between groups, but would be similar within group. This approach effectively looks at the within ZIP code fluctuations in exposure around the ZIP code mean, trend due to measured time-varying covariates, and common trends by group of ZIP codes. It compares that to the same deviation in mortality. Such an approach, looking at yearly deviations from trend and ZIP code mean in exposure and outcome, is inherently examining the relationship of changes in exposure to changes in outcome. This also addresses the issue of whether previous studies effects are due to primarily recent exposure, or reflect long term exposure, including when pollution concentrations were

higher. The dummy variable for each ZIP code controls for long term exposure at that ZIP code, and the removal of nonlinear time trends during the period under study focuses the exposure variable on the year of the death. EPA Regulatory Impact Analyses spread the estimated mortality effects out over a 20 year period. This study provides an estimate of immediate impact. Because our study incorporates 17 years of follow-up, each year has a new exposure, and a new effect.

In addition this paper adds to the sparse literature controlling for temperature in studies of long-term exposure to air pollution, which some have argued is an important confounder. Further, we estimated the probability of dying in a year, which is more easily interpretable than an instantaneous hazard rate, and by using an additive rather than multiplicative model we estimated the marginal effect if $PM_{2.5}$, not the conditional effect estimated by Cox's proportionate hazard model. This allows one to estimate attributable deaths in health impact assessments without making further, possibly implausible assumptions required when using a conditional estimate.

Some scientists, including the current chair of EPA's Clean Air Scientific Advisory Committee (CASAC), assert that studies using standard epidemiological methods should be given little weight in revising the NAAQS, and propose restricting to studies using causal methods, and particularly ones showing changes in air quality are associated with changes in mortality(CASAC, 2019). The recent meetings of EPA's CASAC highlighted the importance of these issues(CASAC., 2019). Their main criticism is that traditional approaches only show associations that may be confounded, vary depending on modeling approaches, and do not inform causality, which can only be addressed by causal methods. They also emphasize that unmeasured variables, particularly individual characteristics, socioeconomic status, and temperature may confound the published literature.(Cox and Popken, 2015) EPA recently proposed not tightening the NAAQS for $PM_{2.5}$, relying on these arguments. This paper provides an analysis using a causal method, controlling for

temperature and socioeconomic status, and all individual and area level potential confounders, measured or unmeasured, that vary slowly over time.

That EPA proposal also asserted there was insufficient evidence of a causal association at lower levels. When we restricted our analysis to include only persons who never experienced concentrations above 12 $\mu g/m^3$ during 17 years of follow-up we found a somewhat larger effect estimate of 4.26×10^{-4} (95% CI 1.43×10^{-4}, 7.09×10^{-4}) per 1 $\mu g/m^3$ increase in $PM_{2.5}$...5. This indicates that the current U.S. and E.U. standards are not sufficient to protect public health, and that the WHO standard of 10 $\mu g/m^3$ is unlikely to protect public health.

Other studies have applied causal modeling methods to air pollution, primarily propensity score methods(Abu Awad et al., 2019; Schwartz et al., 2018; Wang et al., 2017; Wei et al., 2020; Wu et al., 2020; Yitshak-Sade et al., 2019). These methods use the relationship between exposure and confounders to render the exposure independent of all of the measured confounders, and hence mimic a randomized trial. They have all reported that $PM_{2.5}$ increases mortality rates. The difference in differences approach complements those studies by its ability to deal with unmeasured confounders. All personal and small area time invariant or slowly varying confounders are removed by design, whether measured or not. All confounders whose time trends are due to measured time-varying confounders or similar within groups defined by race, SES, medical access, and behavioral characteristics are controlled whether measured or not. Hence, this paper adds assurance about many possible unmeasured confounders to the large literature of associational studies and smaller literature of propensity score-based models that provide causal estimates. Together, they provide strong evidence for a causal effect of $PM_{2.5}$ on mortality rates.

Since we estimate the probability of dying in each year and not a hazard rate, our effect sizes are not directly comparable to the other causal modeling studies. Compared to the larger literature, a

recent meta-analysis of then extant cohort studies estimated the effect size at 10 μg/m^3 (the mean concentration in this study) as a 1.29% increase in the rate per μg/m^3 increase in PM2.5, (95% CI 1.09%, 1.50%)(Vodonos et al., 2018).The annual mortality rate in the Medicare cohort was $4.7x10^{-2}$. A 1.29% increase in that rate is an additive increase of $6.1x10^{-4}$. The results from a previous Cox regression analysis of the Medicare cohort from 2000-2012 translate to an additive increase of $3.4x10^{-4}$. These are similar to our results. Hence these other studies are unlikely to have been confounded by temperature, or slowly varying SES, racial, and behavioral factors which this study controlled for.

Nor are these effects small. Multiplying our effect estimates by the total person-years in the Medicare cohort, we estimate that had everyone had 1 μg/m^3 lower exposure, 239,900 early deaths would have been avoided during the follow-up period. EPA's National Contingency Plan (40 C.F.R. § 300.430(d)(1)) states that the range of acceptable **lifetime risks** (of developing cancer) for carcinogens should be set between 1 in 10,000 and 1 in a million over a 70-year lifetime. Thus, when EPA considers regulations for carcinogens, it typically regulates if **lifetime risks** exceed 1 in a hundred thousand. In contrast, 1 μg/m^3 of exposure below the current EPA standard for only 1 year results in an increased risk of dying of 4.26 per ten thousand in our study.

Our finding has limitations. First, DID analyses depend on the change over time in other ZIP codes with different changes in $PM_{2.5}$ to serve as controls for changes over time in outcome that may have occurred independent of exposure. If the time trends in the ZIP codes are different, this control will fail. We have dealt with this by controlling for time trends in measured covariates and grouping ZIP codes into 32 groups that are similar on age, sex, race/ethnicity, SES, and access to medical care, and doing the analysis separately within each group, arguing that the time trends in

mortality rates will be similar within group. However, we cannot exclude the possibility that they are not. Second, our exposure estimates are not perfect. While an out of sample R^2 of 0.89 is high, there is still some exposure error, which may bias estimates. In addition, personal exposure within a neighborhood varies around the neighborhood ambient concentration. However, we believe most of that difference is likely to be Berksonian error, and hence not bias coefficients. Moreover, the principle reason for the differences between ambient and personal exposure are behavioral (more driving, more cooking, etc), and incorporating exposure related to those factors would require controlling for other related risk factors (e.g. stress from driving) that are not confounders of the neighborhood ambient concentrations. Hence, this exposure error is beneficial from the point of view of reducing confounding, as has been pointed out previously(Weisskopf and Webster, 2017).

In conclusion, we have found an effect of $PM_{2.5}$ on daily deaths using a causal modeling approach robust to unmeasured confounders. The effect size is similar to those reported in associational studies, suggesting that unmeasured confounders are not an issue with them, and is large enough to indicate that reducing $PM_{2.5}$ concentrations in the U.S. could save tens of thousands of premature deaths each year.

References

Abu Awad, Y., et al., 2019. Change in PM2.5 exposure and mortality among Medicare recipients. Environmental Epidemiology. 3, e054.

Baccini, M., et al., 2017. Assessing the short term impact of air pollution on mortality: a matching approach. Environ Health. 16, 7.

Beelen, R., 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, 785-95.

Bureau, U. S. C., US. Census 2000. Summary File 3 (SF 3). In: U. S. C. Bureau, (Ed.), 2010.

CASAC, Public Teleconference on Particulate Matter (PM). EPA Chartered Clean Air Scientific Advisory Committee (CASAC), Available at https://yosemite.epa.gov/sab/sabproduct.nsf/MeetingCalCASAC/4F40665AD1DDCEF6852583A000645464?OpenDocument., 2019.

CASAC., CASAC Review of the EPA's Integrated Science Assessment for Particulate Matter. . In: U. S. E. P. Agency, (Ed.), Washington, D.C., Apr. 2019, 2019.

Caudill, S., Jackson, J., 1989. Measuring Marginal Effects in Limited Dependent Variable Models. Journal of the Royal Statistical Society, Series B. 38, 203-6.

CDC, Behavioral Risk Factor Surveillance System. BRFSS 2013 Survey Data and Documentation. . In: C. f. D. Control, (Ed.), 2013.

Cox, L. A., Jr., Popken, D. A., 2015. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? Ann Epidemiol. 25, 162-73.

Crouse, D. L., et al., 2015. Ambient PM2.5, O3, and NO2 Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). Environ Health Perspect. 123, 1180-6.

Di, Q., et al., 2019. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. Environ Int. 130, 104909.

Di, Q., et al., 2020. Assessing NO2 Concentration and Model Uncertainty with High Spatiotemporal Resolution across the Contiguous United States Using Ensemble Model Averaging. Environmental Science & Technology. 54, 1372-1384.

Di, Q., et al., 2017. Air Pollution and Mortality in the Medicare Population. N Engl J Med. 377, 1498-9.

Donald, S., Lang, K., 2007. Inference with difference-in-differences and other panel data. . Rev Econ Stat. 89, 221–233.

Hoek, G., et al., 2013. Long-term air pollution exposure and cardio- respiratory mortality: a review. Environ Health. 12, 43.

Imai, K., van Dyke, D. A., 2004. Causal inference with general treatment regimes: Generalizing the Propensity Score. Journal of the American Statistical Association. 99, 854:866.

Pinault, L., et al., 2016. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort. Environmental Health. 15, 18.

Pope, C. A., 3rd, et al., 2019. Mortality Risk and Fine Particulate Air Pollution in a Large, Representative Cohort of U.S. Adults. Environ Health Perspect. 127, 77007.

RESDAC, Denominator File. 2018.

Rubin, D. B., 1997. Estimating causal effects from large data sets using propensity scores. Ann Intern Med. 127, 757-63.
Schwartz, J. D., et al., 2018. Estimating the Effects of PM2.5 on Life Expectancy Using Causal Modeling Methods. Environ Health Perspect. 126, 127002.
Vodonos, A., et al., 2018. The concentration-response between long-term PM2.5 exposure and mortality; A meta-regression approach. Environmental Research.
Wang, Y., et al., 2016. Estimating Causal Effects of Long-Term PM2.5 Exposure on Mortality in New Jersey. Environ Health Perspect. 124, 1182-8.
Wang, Y., et al., 2017. Doubly Robust Additive Hazards Models to Estimate Effects of a Continuous Exposure on Survival. Epidemiology (Cambridge, Mass.). 28, 771-779.
Wei, Y., et al., 2020. Causal effects of air pollution in Massachusetts. American journal of epidemiology. in press.
Weisskopf, M., Webster, T., 2017. Trade-offs of Personal Versus More Proxy Exposure Measures in Environmental Epidemiology. Epidemiology (Cambridge, Mass.). 635-43.
Wennberg, J., Cooper, M., 1996. The Dartmouth atlas of health care: . American Hospital Publishing Chicago, IL;.
Wu, X., et al., 2020. Evaluating the Impact of Long-term Exposure to Fine Particulate Matter on Mortality Among the Elderly. Science Advances.
Yang, C., et al., 2012. Alternative ozone metrics and daily mortality in Suzhou: the China Air Pollution and Health Effects Study (CAPES). Sci Total Environ. 426, 83-9.
Yitshak-Sade, M., et al., 2019. Estimating the causal effect of annual PM2.5 exposure on mortality rates in the Northeastern and mid-Atlantic states. Environmental Epidemiology. Latest Articles.

Figure Legend. Figure 1 shows the effect size estimate (absolute increase in the death rate for each 1 μg/m^3 increase in $PM_{2.5}$ exposure, and 95% Confidence Interval) for the entire Medicare Cohort in 2000-2016 (All), for only persons never exposed to $PM_{2.5}$ concentrations above 12 μg/m^3 (low) during the follow-up period, for males in the entire Medicare Cohort (males) and for females in the entire Medicare Cohort.

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Dr. Schwartz declares that he has testified as a health expert for the U.S. Department of Justice in a law suit over clean air act violations. No other authors have anything to declare.

Global mortality from outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem

Karn Vohra[1*], Alina Vodonos[2], Joel Schwartz[2], Eloise A. Marais[3,a], Melissa P. Sulprizio[4], Loretta J. Mickley[4]

[1] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

[2] Harvard T.H. Chan School of Public Health, Department of Environmental Health, Harvard University, Boston, MA, USA

[3] Department of Physics and Astronomy, University of Leicester, Leicester, UK

[4] John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

[a] Now at: Department of Geography, University College London, London, UK

* Corresponding author: Karn Vohra, Phone: +44 7716 496 867,

Email: kxv745@student.bham.ac.uk

Keywords; particulate matter, fossil fuel, mortality, health impact assessment

Abstract

The burning of fossil fuels – especially coal, petrol, and diesel – is a major source of airborne fine particulate matter ($PM_{2.5}$), and a key contributor to the global burden of mortality and disease. Previous risk assessments have examined the health response to total $PM_{2.5}$, not just $PM_{2.5}$ from fossil fuel combustion, and have used a concentration-response function with limited support from

the literature and data at both high and low concentrations. This assessment examines mortality associated with $PM_{2.5}$ from only fossil fuel combustion, making use of a recent meta-analysis of newer studies with a wider range of exposure. We also estimated mortality due to lower respiratory infections (LRI) among children under the age of five in the Americas and Europe, regions for which we have reliable data on the relative risk of this health outcome from $PM_{2.5}$ exposure. We used the chemical transport model GEOS-Chem to estimate global exposure levels to fossil-fuel related $PM_{2.5}$ in 2012. Relative risks of mortality were modeled using functions that link long-term exposure to $PM_{2.5}$ and mortality, incorporating nonlinearity in the concentration response. We estimate a global total of 10.2 (95% CI: -47.1 to 17.0) million premature deaths annually attributable to the fossil-fuel component of $PM_{2.5}$. The greatest mortality impact is estimated over regions with substantial fossil fuel related $PM_{2.5}$, notably China (3.9 million), India (2.5 million) and parts of eastern US, Europe and Southeast Asia. The estimate for China predates substantial decline in fossil fuel emissions and decreases to 2.4 million premature deaths due to 43.7% reduction in fossil fuel $PM_{2.5}$ from 2012 to 2018 bringing the global total to 8.7 (95% CI: -1.8 to 14.0) million premature deaths. We also estimated excess annual deaths due to LRI in children (0-4 years old) of 876 in North America, 747 in South America, and 605 in Europe. This study demonstrates that the fossil fuel component of $PM_{2.5}$ contributes a large mortality burden. The steeper concentration-response function slope at lower concentrations leads to larger estimates than previously found in Europe and North America, and the slower drop-off in slope at higher concentrations results in larger estimates in Asia. Fossil fuel combustion can be more readily controlled than other sources and precursors of $PM_{2.5}$ such as dust or wildfire smoke, so this is a clear message to policymakers and stakeholders to further incentivize a shift to clean sources of energy.

Introduction

The burning of fossil fuels – especially coal, petrol, and diesel – is a major source of airborne particulate matter (PM) and ground-level ozone, which have both been implicated as key contributors to the global burden of mortality and disease (Apte et al., 2015; Dedoussi and Barrett, 2014; Lim et al., 2013). A series of studies have reported an association between exposure to air pollution and adverse health outcomes (Brook et al., 2010), even at low exposure levels (< 10 μg m^{-3}, the current World Health Organization, WHO, guideline) (Di et al., 2017). The Global Burden of Diseases, Injuries, and Risk Factors Study 2015 (GBD 2015) identified ambient air pollution as a leading cause of the global disease burden, especially in low-income and middle-income countries (Forouzanfar et al., 2016). Recent estimates of the global burden of disease suggest that exposure to $PM_{2.5}$ (particulate matter with an aerodynamic diameter < 2.5 μm) causes 4.2 million deaths and 103.1 million disability-adjusted life-years (DALYs) in 2015, representing 7.6% of total global deaths and 4.2% of global DALYs, with 59% of these in east and south Asia (Cohen et al., 2017).

A series of newer studies conducted at lower concentrations and at higher concentrations have reported higher slopes than incorporated into the GBD using the integrated exposure–response (IER) curve (Burnett et al., 2014). These studies examined mortality due to exposure to $PM_{2.5}$ at concentrations below 10 μg m^{-3} in North America (Di et al., 2017; Pinault et al., 2016) and above 40 μg m^{-3} in Asia (Katanoda et al., 2011; Tseng et al., 2015; Ueda et al., 2012; Wong et al., 2015; 2016; Yin et al., 2017). Here we have used a concentration-response curve from a recently published meta-analysis of long-term $PM_{2.5}$ mortality association among adult populations which incorporates those new findings at high and low $PM_{2.5}$ concentrations (Vodonos et al.,

2018). We also focus our study on the health impacts of fossil-fuel derived $PM_{2.5}$. In contrast, GBD reports only the health impacts of total $PM_{2.5}$ and does not distinguish mortality from fossil-fuel derived $PM_{2.5}$ and that from other kinds of $PM_{2.5}$, including dust, wildfire smoke, and biogenically-sourced particles. We focus only on $PM_{2.5}$ since recent studies have provided mixed results on the link between ozone and mortality (Atkinson et al., 2016) and there does not exist a global coherent concentration-response function (CRF) for ozone.

The developing fetus and children younger than 5 years of age are more biologically and neurologically susceptible to the many adverse effects of air pollutants from fossil-fuel combustion than adults. This differential susceptibility to air pollution is due to their rapid growth, developing brain, and immature respiratory, detoxification, immune, and thermoregulatory systems (Bateson and Schwartz, 2008; Perera, 2018). Children also breathe more air per kilogram of body weight than adults, and are therefore more exposed to pollutants in air (WHO, 2006; Xu et al., 2012). The WHO estimated that in 2012, 169,000 global deaths among children under the age of 5 were attributable to ambient air pollution (WHO, 2016). Further estimation of the burden of mortality due to $PM_{2.5}$ (particularly from anthropogenic sources) among the young population would highlight the need for intervention aimed at reducing children's exposure.

Using the chemical transport model GEOS-Chem, we quantified the number of premature deaths attributable to ambient air pollution from fossil fuel combustion. Improved knowledge of this very immediate and direct consequence of fossil fuel use provides evidence of the benefits to current efforts to cut greenhouse gas emissions and invest in alternative sources of energy. It also helps quantify the magnitude of the health impacts of a category of $PM_{2.5}$ that can be more readily controlled than other kinds of $PM_{2.5}$ such as dust or wildfire smoke.

Materials and methods

Calculation of surface $PM_{2.5}$ concentrations

Previous studies examining the global burden of disease from outdoor air pollution have combined satellite and surface observations with models to obtain improved estimates of global annual mean concentrations of $PM_{2.5}$ (Shaddick et al., 2018). However, the goal of such studies was to quantify the health response to $PM_{2.5}$ from all sources, both natural and anthropogenic (Brauer et al., 2016; Cohen et al., 2017). Here the focus of our study is on surface ambient $PM_{2.5}$ generated by fossil fuel combustion, and for that we rely solely on the chemical transport model GEOS-Chem since current satellite and surface measurements cannot readily distinguish between the sources of $PM_{2.5}$. Results from GEOS-Chem have been extensively validated against surface, aircraft, and space-based observations around the world, including simulation of surface pollution over the United States (Drury et al., 2010; Ford and Heald, 2013; Heald et al., 2012; Leibensperger et al., 2012; Marais et al., 2016; Zhang et al., 2012), Asia (Koplitz et al., 2016; Lin et al., 2014), Europe (Protonotariou et al., 2013; Veefkind et al., 2011), and Africa (Lacey et al., 2017; Marais et al., 2014a; 2014b; 2016; 2019). The model has also been applied to previous studies quantifying the global burden of disease from particulate matter from all sources (Brauer et al., 2016; Cohen et al., 2017).

In this analysis we used GEOS-Chem with fossil fuel emissions from multiple sectors (power generation, industry, ships, aircraft, ground transportation, backup generators, kerosene, oil/gas extraction), detailed oxidant-aerosol chemistry, and reanalysis meteorology from the NASA Global Modeling and Assimilation Office. Fossil fuel emissions are from regional inventories where these are available for the US, Europe, Asia, and Africa, and from global

inventories everywhere else (such as Mexico, Australia, South America and Canada). More details of the specific fossil fuel inventories used in GEOS-Chem are in Table S1. Global-scale simulations in GEOS-Chem were carried out on a coarse spatial grid (2° × 2.5°, about 200 km × 250 km). Four regional simulations were also performed at fine spatial scale (0.5° × 0.67°, about 50 km × 60 km) for North America, Europe, Asia, and Africa using boundary conditions from the global model. The regional simulations allow for a better match with the spatial distribution of population, thus enhancing the accuracy of the estimates of health impacts. All simulations were set up to replicate 2012 pollution conditions. As described in the Supplemental Material, we find that globally, GEOS-Chem captures observed annual mean $PM_{2.5}$ concentrations with a spatial correlation of 0.70 and mean absolute error of 3.4 μg m^{-3}, values which compare well with those from other models (Shindell et al., 2018; Xing et al., 2015). We performed two sets of simulations: one set with fossil fuel emissions turned on and the other with such emissions turned off. We then assumed that the difference between the two sets of simulations represents the contribution of fossil fuel combustion to surface $PM_{2.5}$. More information on our choice of GEOS-Chem, the model setup, details of relevant anthropogenic emissions, and model validation is described in the Supplemental material.

Population and Health data

We used population data from the Center for International Earth Science Information Network (CIESIN) (CIESIN, 2018). The Gridded Population of the World, Version 4 Revision 11 (GPWv4.11) is gridded with an output resolution of 30 arc-seconds (approximately 1 km at the equator). Since the population data are provided only at five-year intervals, we applied 2015 population statistics to the results of our 2012 GEOS-Chem simulation. CIESIN population data

was then aggregated to the spatial scale of the model for the exposure estimates. Country/region level data on baseline mortality rates were from GBD data for 2015 (based on the 2017 iteration) (IHME, 2017). USA state-specific mortality rates were obtained from the CDC Wide-ranging Online Data for Epidemiologic Research (WONDER) compressed mortality files (WONDER). Canada death estimates by province were obtained from Statistics Canada, CANSIM (Canada, 2018).

$PM_{2.5}$ mortality concentration –response model

The risk of air pollution to health in a population is usually estimated by applying a concentration–response function (CRF), which is typically based on Relative Risk (RR) estimates derived from epidemiological studies. CRFs are necessary elements for the quantification of health impacts due to air pollution and require regular evaluation and update to incorporate new developments in the literature.

Global assessments of air pollution risk often use the Integrated Exposure-Response model (IER) (Burnett et al., 2014), which combined information on $PM_{2.5}$–mortality associations from non-outdoor $PM_{2.5}$ sources, including secondhand smoke, household air pollution from use of solid fuels, and active smoking. The IER used data from active smoking and passive smoking to address the limited number of outdoor $PM_{2.5}$ epidemiologic studies at $PM_{2.5}$ > 40 $\mu g\ m^{-3}$ available at the time. The IER formed the basis of the estimates of disease burden attributable to $PM_{2.5}$ (e.g., 4 million deaths in 2015 in GBD 2015). This function was then updated in 2018 using the Global Exposure Mortality Model (GEMM). In GEMM, data from 41 epidemiological cohort studies were applied (Burnett et al., 2018). Independently conducted analyses were conducted on 15 of these cohorts to characterize the shapes of $PM_{2.5}$–mortality associations in each cohort, using a specified functional form of the CRF. For the remaining 26

cohorts, the concentration-response was examined with a linear concentration hazard ratio model. A recent meta-analysis of the association between long-term $PM_{2.5}$ and mortality (Vodonos et al., 2018) applied techniques involving flexible penalized spline CRF in a multivariate random effects and meta-regression model. This approach allows the data to specify the shape of the CRF. The meta-regression pooled 135 estimates from 53 studies examining long-term $PM_{2.5}$ and mortality of cohorts aged 15 years and older. The estimate of the confidence intervals about the CRF includes a random variance component. This meta-analysis provided evidence of a nonlinear association between $PM_{2.5}$ exposure and mortality in which the exposure-mortality slopes decreases at higher concentrations (Figure S5 in Supplemental Material). We have chosen to use the dose-response function from the meta-analysis rather than the GEMM function as the meta-regression approach is more flexible and does not constrain the CRF to a specific functional form, it incorporates a random variance component in estimating the uncertainty around that curve, it is derived with more studies than previous approaches, and its estimates at high and low exposures are closer to the estimates in cohorts restricted to only very high and very low exposures. To ensure consistency with the concentration-response curve, premature mortality rates for the portion of the population >14 years of age were determined using the population and baseline mortality rates for different age groups from GBD data for 2015.

Health impact calculations

We estimated the number of premature deaths attributable to fossil fuel $PM_{2.5}$ using: (1) GEOS-Chem $PM_{2.5}$ estimated with all emission sources and GEOS-Chem $PM_{2.5}$ estimated without fossil fuel emissions, as a comparison against the first simulation, (2) total population above the age of 14 gridded to the GEOS-Chem grid resolution, (3) baseline all-cause mortality rates for

population above the age of 14 (per country or per state in the US and province in Canada), and (4) the meta-analysis CRF (Vodonos et al., 2018). All health impacts were calculated on a per-grid basis at the spatial resolution of the model. We applied the following health impact function to estimate premature mortality related to exposure to fossil fuel $PM_{2.5}$ in each GEOS-Chem grid cell:

$$\sum \Delta y = y_o * p * AF \quad (1)$$

$$AF = \frac{\exp^{(\bar{\beta}*\Delta x)}-1}{\exp^{(\bar{\beta}*\Delta x)}} \quad (2)$$

$$\bar{\beta}(PM_{2.5}) = \int_{PM_{2.5}\ \text{no fossil fuel}}^{PM_{2.5}\ \text{all emissions}} \beta(PM_{2.5}) \quad (3)$$

where Δy is the change in the number of premature deaths due to exposure to fossil fuel $PM_{2.5}$, y_o is the country/state/province specific baseline (all-cause) mortality rate, p is to the total population above the age of 14, AF is the attributable fraction of deaths (the fraction of total deaths attributable to $PM_{2.5}$ exposure), $\bar{\beta}$ is the mean estimate for long-term $PM_{2.5}$ mortality concentration-response over a range of concentrations from the penalized spline model in the recent meta-analysis, and Δx is the change in $PM_{2.5}$ concentration, calculated as the difference between GEOS-Chem $PM_{2.5}$ with all emissions and GEOS-Chem $PM_{2.5}$ without fossil fuel emissions.

For each country, we summed the change in premature deaths (Δy) in each grid cell over all grid cells in that country. To estimate the change in deaths between the two scenarios (with and without fossil fuel combustion), we computed the change in deaths in each grid cell, based on its

population, baseline rate, and exposure under the two scenarios (Equation (1)). The attributable fraction (AF), or proportion of deaths estimated as due to long-term exposure to $PM_{2.5}$ fossil fuel air pollution, was calculated using the concentration-response estimate, following the form shown in Equation (2) (Figure S5 in Supplemental material). Because these estimates of mortality concentration response (β) are a nonlinear function of concentration, we used the penalized spline model predictions from this meta-analysis to integrate the concentration-specific β in each grid cell from the low $PM_{2.5}$ scenario (without fossil fuel emissions) to the high $PM_{2.5}$ scenario (with all emissions, including fossil fuel). In this way, we could calculate a mean value of β for each grid cell. There exist insufficient epidemiological data to calculate a robust health response function specific to fossil-fuel $PM_{2.5}$. GEOS-Chem is a deterministic model. Therefore, our 95% confidence intervals (CI) for our estimates reflect only the 95% CI for the concentration response function.

Secondary analysis among children <5 years old

Lower respiratory infections (LRI), including pneumonia and bronchiolitis of bacterial and viral origin, are the largest single cause of mortality among young children worldwide and thus account for a significant global burden of disease worldwide (Nair et al., 2010). As mentioned previously, young children are more susceptible to the adverse effects of particulate air pollution than adults. Mehta et al. (2013) estimated the overall impact of $PM_{2.5}$ concentration with Relative Risk (RR) of 1.12 for LRI mortality per 10 μg m^{-3} increase in annual average $PM_{2.5}$ concentration, as compared to RR of 1.04 for respiratory mortality among adults (Vodonos et al., 2018). We estimated the number of premature deaths attributable to $PM_{2.5}$ among children under the age of 5 years due to a range of LRI classifications (ICD-10, International Classification of Diseases codes: A48.1, A70, J09-J15.8, J16-J16.9, J20-J21.9, P23.0-P23.4). Baseline numbers of deaths due to LRI were obtained from the GBD for 2015 (IHME, 2017). We used the Relative

Risk (RR) of 1.12 (1.03-1.30) for LRI occurrence per 10 μg m^{-3} increase in annual average $PM_{2.5}$ concentration (Mehta et al., 2013). Studies of longer-term exposure of $PM_{2.5}$ and LRI in that meta-analysis were conducted in only a few developed countries with relatively low levels of annual mean $PM_{2.5}$ (< 25 μg m^{-3}), specifically the Netherlands, Czech Republic, Germany, Canada and USA. We therefore calculated the number of premature LRI deaths attributable to $PM_{2.5}$ only in North America, South America, and Europe.

Results

Impact of fossil fuel use on $PM_{2.5}$

Figure 1 shows the difference between global GEOS-Chem $PM_{2.5}$ with and without fossil fuel emissions, plotted as the annual mean for 2012. Results show large contributions of 50-100 μg m^{-3} in $PM_{2.5}$ over China and India, with smaller increments of 10-50 μg m^{-3} over large swaths of the United States and Europe, industrialized countries in Africa (South Africa and Nigeria), and along the North African coastline due to European pollution.

Global assessment of mortality attributable to $PM_{2.5}$

Based on the annual $PM_{2.5}$ simulation with and without global fossil fuel emissions, we estimated the excess deaths and attributable fraction (AF %) for the population above 14 years old. Figure 2 shows the simulated annual global premature mortality due to exposure to ambient $PM_{2.5}$ from fossil fuel emissions. Greatest mortality is simulated over regions with substantial influence of fossil-fuel related $PM_{2.5}$, notably parts of Eastern North America, western Europe, and South-East Asia.

We estimated a total global annual burden premature mortality due to fossil fuel combustion in 2012 of 10.2 million (95% CI: -47.1 to 17.0 million). Table 1 reports the baseline number of deaths for people >14 years old, the annual $PM_{2.5}$ simulation with and without global fossil fuel emissions, the estimated excess deaths, and the attributable fraction for the populated continents. As shown in Table 1, we calculated 483,000 premature deaths in North America (95% CI: 284,000-670,000), 187,000 deaths in South America (95% CI: 107,000-263,000), 1,447,000 deaths in Europe (95% CI: 896,000-1,952,000), 7,916,000 deaths in Asia (95% CI: -48,106,000 to 13,622,000), and 194,000 deaths in Africa (95% CI: -237,000 to 457,000). The wide confidence intervals in Asia and Africa are due to the lack of data for areas where the exposure remains outside the range of the concentration response curve ($PM_{2.5} > 50$ μg m^{-3}; Figure S5). The population-weighted pollution concentrations presented in Table 1 are higher than the average $PM_{2.5}$ concentrations for each country, since fossil-fuel $PM_{2.5}$ is mainly emitted in populous areas. The two countries with the highest premature mortality are China with 3.91 million and India with 2.46 million. Supplemental Table S2 provides extended data of the health impact calculations for each country. For comparison, Table 1 also reports the number of premature deaths attributable to fossil fuel $PM_{2.5}$ when the GEMM function is applied to the GEOS-Chem output. For most regions, the number of premature deaths calculated with GEMM is significantly lower than that calculated with the new function from Vodonos et al. (2018). Globally, the GEMM function yields 6.7 million deaths in 2012 due to fossil fuel combustion.

Assessment of children (under the age of 5) LRI mortality attributable to $PM_{2.5}$

We estimated the number of premature deaths attributable to $PM_{2.5}$ among children under the age of 5 due to LRI only for those countries or regions with levels of annual $PM_{2.5}$

concentrations below 25 μg m^{-3}. These include North America, South America, and Europe. Based on the annual $PM_{2.5}$ simulation with and without fossil fuel emissions, we calculated 876 excess deaths due to LRI in North and Central America, 747 in South America, and 605 in Europe (Table 2). Using the GBD estimate of total deaths due to LRI (Institute for Health Metrics and Evaluation), we estimate that $PM_{2.5}$ from fossil fuel combustion accounted on average for 7.2% of LRI mortality among children under the age of 5 in these regions, with the largest proportion of 13.6% in Europe (95% CI -0.4 to 25.3%) .

Discussion

We used the chemical transport model GEOS-Chem to quantify the global mortality attributed to $PM_{2.5}$ air pollution from fossil fuel combustion. Using the updated concentration response relationship between relative mortality and airborne $PM_{2.5}$, we estimated global premature mortality in 2012 of 10.2 million per year from fossil fuel combustion alone. China has the highest burden of 3.91 million per year, followed by India with 2.46 million per year. These estimates carry large uncertainty (e.g., 95% CI of -47.1 to 17.0 million for the global estimate) from the concentration-response curve, as it is an improved function that provides a more realistic picture of the health consequences of $PM_{2.5}$ compared to previous studies.

Our estimate is for the year when fossil fuel emissions in China peaked and so predates large and dramatic reductions in fossil fuel emissions due to strict mitigation measures. These reductions led to a 30-50% decline in annual mean $PM_{2.5}$ across the country from 2013 to 2018 (Zhai et al., 2019). If we apply a 43.7% reduction in GEOS-Chem $PM_{2.5}$ concentrations from the simulation with all emission sources, premature mortality in China decreases from 3.91 million to

2.36 million. India has recently imposed controls on pollution sources, but there is not yet evidence of air quality improvements in densely populated cities like Delhi (Vohra et al., 2020). Consideration of the 2012-2018 decrease in $PM_{2.5}$ exposure in China reduces the total global premature mortality due to fossil fuel $PM_{2.5}$ from 10.2 million premature deaths each year to 8.7 (95% CI: -1.8 to 14.0) million.

In 2012, the population-weighted $PM_{2.5}$ is 72.8 μg m^{-3} for China and 52.0 μg m^{-3} for India from all sources and 9.9 μg m^{-3} for China and 9.0 μg m^{-3} for India without fossil fuel emissions. The low value of non-fossil fuel $PM_{2.5}$ is reasonable for southern India (Dey et al., 2012) but may be an underestimate in the Indo-Gangetic Plain where crop residue burning contributes to high levels of $PM_{2.5}$ (100-200 μg m^{-3}) during the post-monsoon season (Ojha et al., 2020). An increase in the concentration of non-fossil-fuel $PM_{2.5}$ would decrease our estimate of the number of premature deaths due to fossil fuel $PM_{2.5}$ in India and China, as this would decrease the risk of premature mortality with a unit change in $PM_{2.5}$ (Figure S5).

Comparison with previous estimates of global mortality attributable to outdoor $PM_{2.5}$

Previous estimates of the GBD for 2015 suggest that exposure to total $PM_{2.5}$ causes 4.2 million deaths (Cohen et al., 2017), whereas here we estimate more than double (10.2 million) the number of premature deaths from fossil fuel combustion alone in 2012. Differences between the current study and the 2015 GBD lower estimates are related mainly to the choice of the shape of the concentration-response function and the relative risk estimate. First, to provide information about exposure response at higher concentrations, the 2015 GBD study used the integrated exposure–response (IER) model in which active and second-hand smoking exposures were

converted to estimated annual $PM_{2.5}$ exposure equivalents using inhaled doses of particle mass (Burnett et al., 2014). Recent cohort studies from Asia indicate that this substantially underestimates the CRF at high concentrations. In contrast, in the current study we applied a CRF that was directly estimated from $PM_{2.5}$ studies alone, as described in a recent meta-analysis that included estimates from studies in countries like China with higher $PM_{2.5}$ concentrations than our included in previous derivations of CRFs (Vodonos et al., 2018). The CRF from this recent meta-analysis flattens out at higher concentrations, as does the IER curve. However, this flattening is not as great as in the IER, as Asian cohort studies at high $PM_{2.5}$ concentrations report larger effects than would be expected from the IER. Hence estimates of the global attributable fraction of deaths due to air pollution using the function from the recent meta-analysis are higher than the estimates using the IER function. In addition, at much lower concentrations ($< 10\ \mu g\ m^{-3}$), we applied higher slopes than assumed in the IER function. Recent studies at very low concentrations similarly show that the IER underestimated effects in this range (Pinault et al., 2016). Since GEOS-Chem estimated quite low concentrations in developed countries in Europe and North America, the number of premature deaths from $PM_{2.5}$ in these countries is greater than previous estimates.

Following an approach similar to the recent meta-analysis (Vodonos et al., 2018), Burnett et al. (2018) modeled the shape of the association between $PM_{2.5}$ and non-accidental mortality using data from 41 cohorts from 16 countries with GEMM. In that study, the uncertainty in a subset (15 cohorts) was characterized in the shape of the concentration-response parameter by calculating the Shape-Constrained Health Impact Function, a prespecified functional form. These estimated shapes varied across the cohorts included in the function. GEMM predicted 8.9 million (95% CI: 7.5–10.3) deaths in 2015 attributable to long-term exposure to $PM_{2.5}$ from all sources; 120% higher excess deaths than previous estimates, but still lower than our estimate of mortality from exposure

to fossil-fuel derived $PM_{2.5}$ for 2012. Lelieveld et al. (2019) estimated the global and regional mortality burden of fossil fuel attributable $PM_{2.5}$ by applying the GEMM CRF to a global chemistry-climate model that is overall coarser (~1.9° latitude and longitude) than the model used in this work. The authors reported 3.61 million deaths per year attributable to pollution from fossil fuel combustion and 5.55 million deaths per year due to pollution from all anthropogenic sources. The estimated deaths from fossil fuel combustion are much lower than those in the current study for several reasons. First, the meta-analysis function used in our work includes 135 coefficients of all-cause mortality for adults aged 14-64 years old, together with cause-specific mortality and all-cause mortality among adults aged 65 and older, thus incorporating many more studies in a meta-regression framework than the 41 cohorts and coefficients in the GEMM function. Second, the approach used to estimate the CRF in Vodonos et al. (2018) allows for additional flexibility in the shape of the function because of its use of penalized splines. In contrast, the GEMM pooled CRF integrates a set of 26 log-linear functions and 15 functions characterized by three parameters governing the shape of the function. Third, while Cohen et al. (2017), Lelieveld et al. (2019) and Burnett et al. (2018) accounted for mortality from five specific causes (ischemic heart disease, stroke, chronic obstructive pulmonary disease, lung cancer and acute respiratory infections), in the current analysis we estimated changes in deaths from all causes. Fourth, some of the difference in the mortality estimates may come from differences in the age range. Our approach considers a wider population age range of over 14 years old (Vodonos et al., 2018) compared to the other studies, which considered a population age range of over 25 years (Burnett et al., 2018; Cohen et al., 2017; Lelieveld et al., 2019). Our approach has wider age range since the age range for the studies in the meta-analysis (Vodonos et al., 2018) included people younger than 25 years old (Hart et al., 2011; Pinault et al., 2016) . Finally, the finer spatial resolution that GEOS-Chem

utilizes over much of the globe improves co-location of PM hotspots and population centers, yielding higher estimates of excess mortality compared to Lelieveld et al. (2019).

Limitations

There are a number of limitations that must be acknowledged. First, vulnerability to $PM_{2.5}$ exposure may vary by population characteristics such as ethnicity, socio-economic status (SES), risk behaviors such as smoking and underlying comorbidities (Krewski et al., 2000; Pope et al., 2004; Wang et al., 2017) and by different exposure characteristics. We were limited in our ability to undertake a comprehensive analysis of factors influencing the association between $PM_{2.5}$ and mortality since the global mortality data were not available by detailed age, ethnicity, SES, lifestyle, and underlying disease strata. In addition, the 95% CI of our estimates reflect the lower and upper bound of the CRF, which flattens out at higher concentrations. Regions with very high concentrations (>50 μg m^{-3}) are beyond the data range in the meta-analysis; thus, the lower limit of the CI for those regions (China, West and North Africa; Table 1) are much less than zero. Second, for LRI in children, we have restricted our analysis to developed countries with annual $PM_{2.5}$ < 25 μg m^{-3}, in accordance with the geographical locations of the studies included in the meta-analysis by Mehta et al. (2013). Developing countries have much higher LRI mortality rates, and this restriction doubtless results in an underestimate. Finally, GEOS-Chem estimates of $PM_{2.5}$ concentrations almost certainly contains errors in estimates of emissions of pollution precursors, meteorological effects on air quality, and representation of the complex physical and chemical formation pathways. In the absence of systematic bias, such model error may not produce large

aggregate errors in the mortality burden of $PM_{2.5}$, but bias may be present as well. In any event, it is challenging to estimate the true size of this error.

Conclusions

The effects of CO_2-driven climate change on human health and welfare are complex, ranging from greater incidence of extreme weather events, more frequent storm-surge flooding, and increased risk of crop failure (Duffy et al., 2019). One consequence of increasing reliance on fossil fuel as an energy source that has thus far received comparatively little attention is the potential health impact of the pollutants co-emitted with the greenhouse gas CO_2. Such pollutants include $PM_{2.5}$ and the gas-phase precursors of $PM_{2.5}$. This study demonstrates that the fossil fuel component of $PM_{2.5}$ contributes a large global mortality burden. By quantifying this sometimes overlooked health consequence of fossil fuel combustion, a clear message is sent to policymakers and stakeholders of the co-benefits of a transition to alternative energy sources.

Acknowledgments

This study was funded by the Wallace Global Fund, the Environment and Health Fund (EHF) Israel, and a University of Birmingham Global Challenges Fund PhD studentship awarded to KV.

Declaration of interests

We declare no competing interests.

Data availability. GEOS-Chem code and output are available at the GEOS-Chem website (http://acmg.seas.harvard.edu/geos_chem.html) and upon request.

Figures

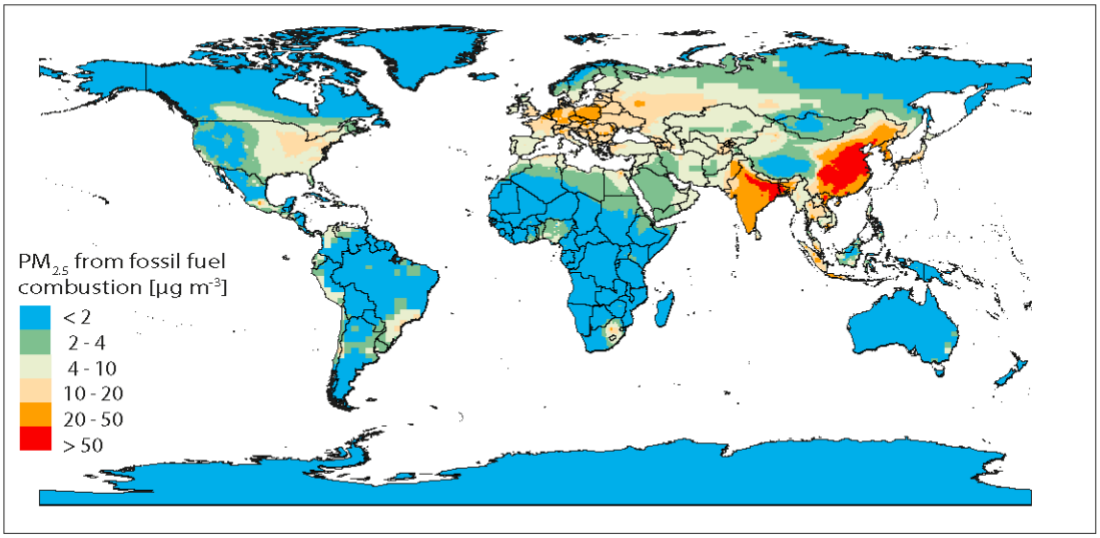


Figure 1: Contribution of fossil fuel combustion to surface $PM_{2.5}$, as calculated by the chemical transport model GEOS-Chem. The plot shows the difference in surface $PM_{2.5}$ concentrations from GEOS-Chem with and without fossil fuel emissions.

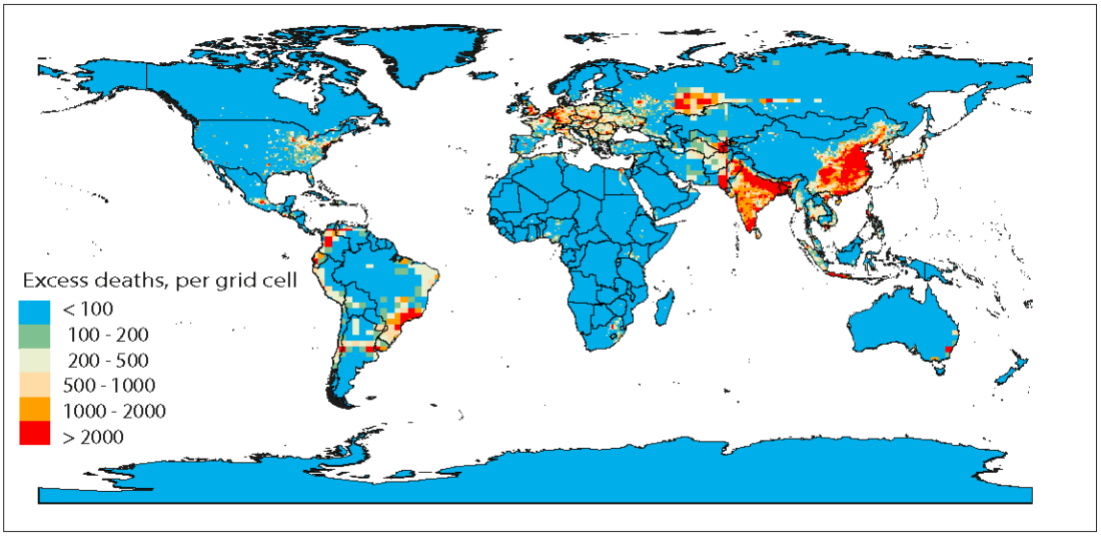


Figure 2. Estimated annual excess deaths due to exposure to ambient $PM_{2.5}$ generated by fossil fuel combustion.

Table 1. Number of deaths attributable to exposure to fine particulate matter ($PM_{2.5}$) generated by fossil fuel combustion for the population >14 years old

GEOS-Chem spatial grid resolution[a]	Region[b]		Total deaths >14 years old, in thousands	Population-weighted annual mean $PM_{2.5}$ concentration, μg m^{-3}			Mean attributable fraction of deaths, % (95% CI)[d]	Deaths attributable to fossil-fuel related $PM_{2.5}$, in thousands (95% CI)[c]	GEMM function deaths attributable to fossil-fuel related $PM_{2.5}$, in thousands (95% CI)[e]
				$PM_{2.5}$ from all emission sources	$PM_{2.5}$ without fossil fuel	Estimated $PM_{2.5}$ from fossil fuel, %			
Fine	North America	Central America & the Caribbean	1,148	10.06	3.03	7.03 (69.9)	8.2 (4.5-11.6)	94 (52-133)	80 (62-98)
		USA	2,705	11.81	2.15	9.66 (81.8)	13.1 (7.8-18.1)	355 (212-490)	305 (233-375)
		Canada	250	12.01	1.76	10.25 (85.4)	13.6 (8.0-18.7)	34 (20-47)	28 (22-35)
Coarse	South America		2,389	8.66	3.02	5.65 (65.2)	7.8 (4.5-11.0)	187 (107-263)	159 (121-195)
Fine	Europe		8,626	19.22	4.68	14.54 (75.7)	16.8 (10.4-22.6)	1,447 (896-1,952)	1,033 (798-1,254)
Fine	Asia	Eastern Asia	25,468	51.72	8.68	43.05 (83.2)	30.7 (-189.1-52.9)	7,821 (-48,150-13,478)	4,945 (3,943-5,826)
Coarse		Western Asia & the Middle East	1,456	26.95	20.73	6.22 (23.1)	6.5 (3.0-9.9)	95 (44-144)	54 (43-65)
Fine	Africa		5,274	32.98	28.98	4.00 (12.1)	3.7 (-4.5-8.7)	194 (-237-457)	102 (81-121)
Coarse	Australia & Oceania		189	4.17	2.19	1.98 (47.4)	3.2 (1.6-4.8)	6.0 (2.9-9.0)	6.4 (4.8-7.9)
	Global		47,506	38.01	11.14	26.87 (70.7)	21.5 (-99.0-35.7)	10,235 (-47,054-16,972)	6,713 (5,308-7,976)

[a] Fine spatial scale is 0.5° × 0.67°, or about 50 km × 60 km. Coarse spatial scale is 2° × 2.5°, or about 200 km × 250 km

[b] List of countries for each region and subregion is provided in supplemental Table S2

[c] Annual number of deaths attributable to long-term exposure to $PM_{2.5}$ derived from fossil fuel combustion. CI is the confidence interval.

[d] Mean proportion of all deaths which can be attributed to long-term exposure to $PM_{2.5}$ generated by fossil fuel combustion, averaged over the country or region. CI; confidence interval.

[e] Attributable deaths calculated with the Global Exposure Mortality Model (GEMM) concentration-response function. [44]

Table 2. Number of deaths due to lower respiratory infection (LRI) attributable to exposure to fine particulate matter ($PM_{2.5}$) from fossil fuel combustion for the population <5 years old

Region	Total deaths for children <5 years old due to LRI	LRI deaths attributable to fossil-fuel $PM_{2.5}$ (95% CI) [a]	Mean attributable fraction of deaths, % (95% CI) [b]
North America	**13,230**	**876 (-26-1,657)**	**6.6 (-0.2-12.5)**
Central America & the Caribbean	12,507	802 (-23-1,516)	6.4 (-0.2-12.1)
USA	672	69 (-2-131)	10.2 (-0.3-19.5)
Canada	50	5 (0-10)	10.8 (-0.3-20.5)
South America	**13,231**	**747 (-21-1,443)**	**5.7 (-0.2-10.9)**
Europe	**4,446**	**605 (-18-1,126)**	**13.6 (-0.4-25.3)**

[a] Annual number of deaths attributed to long-term exposure to $PM_{2.5}$ derived from fossil fuel combustion.

[b] Mean proportion of deaths due to long-term exposure to $PM_{2.5}$ generated by fossil fuel combustion. CI is the confidence interval.

References

1. Apte, J. S., Marshall, J. D., Cohen, A. J., et al., Addressing Global Mortality from Ambient $PM_{2.5}$, Environ Sci Technol, 49, 8057-8066, doi:10.1021/acs.est.5b01236, 2015.

2. Atkinson, R. W., Butland, B. K., Dimitroulopoulou, C., et al., Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies, Bmj Open, 6, doi:10.1136/bmjopen-2015-009493, 2016.

3. Bateson, T. F., Schwartz, J., Children's response to air pollutants, J Toxicol Env Heal A, 71, 238-243, doi:10.1080/15287390701598234, 2008.

4. Brauer, M., Freedman, G., Frostad, J., et al., Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013, Environ Sci Technol, 50, 79-88, doi:10.1021/acs.est.5b03709, 2016.

5. Brook, R. D., Rajagopalan, S., Pope, C. A., et al., Particulate Matter Air Pollution and Cardiovascular Disease An Update to the Scientific Statement From the American Heart Association, Circulation, 121, 2331-2378, doi:10.1161/CIR.0b013e3181dbece1, 2010.

6. Burnett, R., Chen, H., Szyszkowicz, M., et al., Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter, P Natl Acad Sci USA, 115, 9592-9597, doi:10.1073/pnas.1803222115, 2018.

7. Burnett, R., Pope, C. A., Ezzati, M., et al., An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure, Environ Health Persp, 122, 397-403, doi:10.1289/ehp.1307049, 2014.

8. Canada, S., Government of Canada. https://www150.statcan.gc.ca/n1/en/type/data, 2018.

9. CIESIN, Center for International Earth Science Information Network - Columbia University; Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY, https://doi.org/10.7927/H4PN93PB, 2018.

10. Cohen, A. J., Brauer, M., Burnett, R., et al., Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015, Lancet, 389, 1907-1918, doi:10.1016/S0140-6736(17)30505-6, 2017.

11. Dedoussi, I. C., Barrett, S. R. H., Air pollution and early deaths in the United States. Part II: Attribution of $PM_{2.5}$ exposure to emissions species, time, location and sector, Atmos Environ, 99, 610-617, doi:10.1016/j.atmosenv.2014.10.033, 2014.

12. Dey, S., Di Girolamo, L., van Donkelaar, A., et al., Variability of outdoor fine particulate ($PM_{2.5}$) concentration in the Indian Subcontinent: A remote sensing approach, Remote Sens Environ, 127, 153-161, doi:10.1016/j.rse.2012.08.021, 2012.

13. Di, Q., Wang, Y., Zanobetti, A., et al., Air Pollution and Mortality in the Medicare Population, N Engl J Med, 376, 2513-2522, doi:10.1056/NEJMoa1702747, 2017.

14. Drury, E., Jacob, D. J., Spurr, R. J. D., et al., Synthesis of satellite (MODIS), aircraft (ICARTT), and surface (IMPROVE, EPA-AQS, AERONET) aerosol observations over eastern North America to improve MODIS aerosol retrievals and constrain surface aerosol concentrations and sources, J Geophys Res-Atmos, 115, doi:10.1029/2009jd012629, 2010.

15. Duffy, P. B., Field, C. B., Diffenbaugh, N. S., et al., Strengthened scientific support for the Endangerment Finding for atmospheric greenhouse gases, Science, 363, 597-+, doi:10.1126/science.aat5982, 2019.

16. Ford, B., Heald, C. L., Aerosol loading in the Southeastern United States: reconciling surface and satellite observations, Atmos Chem Phys, 13, 9269-9283, doi:10.5194/acp-13-9269-2013, 2013.

17. Forouzanfar, M. H., Afshin, A., Alexander, L. T., et al., Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015, Lancet, 388, 1659-1724, doi:10.1016/S0140-6736(16)31679-8, 2016.

18. Hart, J. E., Garshick, E., Dockery, D. W., et al., Long-Term Ambient Multipollutant Exposures and Mortality, Am J Resp Crit Care, 183, 73-78, doi:10.1164/rccm.200912-1903OC, 2011.

19. Heald, C. L., Collett, J. L., Lee, T., et al., Atmospheric ammonia and particulate inorganic nitrogen over the United States, Atmos Chem Phys, 12, 10295-10312, doi:10.5194/acp-12-10295-2012, 2012.

20. IHME, Institute for Health Metrics and Evaluation. http://ghdx.healthdata.org/gbd-results-tool, 2017.

21. Katanoda, K., Sobue, T., Satoh, H., et al., An Association Between Long-Term Exposure to Ambient Air Pollution and Mortality From Lung Cancer and Respiratory Diseases in Japan, J Epidemiol, 21, 132-143, doi:10.2188/jea.JE20100098, 2011.

22. Koplitz, S. N., Mickley, L. J., Marlier, M. E., et al., Public health impacts of the severe haze in Equatorial Asia in September-October 2015: demonstration of a new framework for informing fire management strategies to reduce downwind smoke exposure, Environ Res Lett, 11, doi:10.1088/1748-9326/11/9/094023, 2016.

23. Krewski, D., Burnett, R. T., Goldberg, M. S., et al., Special report reanalysis of the Harvard six cities study and the American Cancer Society Study of particulate air pollution and mortality part II: Sensitivity Analyses Appendix C. Flexible Modeling of the Effects of Fine Particles and Sulphate on Mortality, Health Effects Institute, https://www.healtheffects.org/system/files/SR-PartIIAppC.pdf, 2000.

24. Lacey, F. G., Marais, E. A., Henze, D. K., et al., Improving present day and future estimates of anthropogenic sectoral emissions and the resulting air quality impacts in Africa, Faraday Discuss, 200, 397-412, doi:10.1039/c7fd00011a, 2017.

25. Leibensperger, E. M., Mickley, L. J., Jacob, D. J., et al., Climatic effects of 1950-2050 changes in US anthropogenic aerosols - Part 1: Aerosol trends and radiative forcing, Atmos Chem Phys, 12, 3333-3348, doi:10.5194/acp-12-3333-2012, 2012.

26. Lelieveld, J., Klingmuller, K., Pozzer, A., et al., Effects of fossil fuel and total anthropogenic emission removal on public health and climate, P Natl Acad Sci USA, 116, 7192-7197, doi:10.1073/pnas.1819989116, 2019.

27. Lim, S. S., Vos, T., Flaxman, A. D., et al., A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010 (vol 380, pg 2224, 2012), Lancet, 381, 628-628, <Go to ISI>://WOS:000315189300032, 2013.

28. Lin, J. T., van Donkelaar, A., Xin, J. Y., et al., Clear-sky aerosol optical depth over East China estimated from visibility measurements and chemical transport modeling, Atmos Environ, 95, 258-267, doi:10.1016/j.atmosenv.2014.06.044, 2014.

29. Marais, E. A., Jacob, D. J., Guenther, A., et al., Improved model of isoprene emissions in Africa using Ozone Monitoring Instrument (OMI) satellite observations of formaldehyde: implications for oxidants and particulate matter, Atmos Chem Phys, 14, 7693-7703, doi:10.5194/acp-14-7693-2014, 2014a.

30. Marais, E. A., Jacob, D. J., Jimenez, J. L., et al., Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: application to the southeast United States and co-benefit of SO_2 emission controls, Atmos Chem Phys, 16, 1603-1618, doi:10.5194/acp-16-1603-2016, 2016.

31. Marais, E. A., Jacob, D. J., Wecht, K., et al., Anthropogenic emissions in Nigeria and implications for atmospheric ozone pollution: A view from space, Atmos Environ, 99, 32-40, doi:10.1016/j.atmosenv.2014.09.055, 2014b.

32. Marais, E. A., Silvern, R. F., Vodonos, A., et al., Air Quality and Health Impact of Future Fossil Fuel Use for Electricity Generation and Transport in Africa, Environ Sci Technol, 53, 13524-13534, doi:10.1021/acs.est.9b04958, 2019.

33. Marais, E. A., Wiedinmyer, C., Air Quality Impact of Diffuse and Inefficient Combustion Emissions in Africa (DICE-Africa), Environ Sci Technol, 50, 10739-10745, doi:10.1021/acs.est.6b02602, 2016.

34. Mehta, S., Shin, H., Burnett, R., et al., Ambient particulate air pollution and acute lower respiratory infections: a systematic review and implications for estimating the global burden of disease, Air Qual Atmos Hlth, 6, 69-83, doi:10.1007/s11869-011-0146-3, 2013.

35. Nair, H., Nokes, D. J., Gessner, B. D., et al., Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis, Lancet, 375, 1545-1555, doi:10.1016/S0140-6736(10)60206-1, 2010.

36. Ojha, N., Sharma, A., Kumar, M., et al., On the widespread enhancement in fine particulate matter across the Indo-Gangetic Plain towards winter, Sci Rep-Uk, 10, doi:10.1038/s41598-020-62710-8, 2020.

37. Perera, F., Pollution from Fossil-Fuel Combustion is the Leading Environmental Threat to Global Pediatric Health and Equity: Solutions Exist, Int J Env Res Pub He, 15, doi:10.3390/ijerph15010016, 2018.

38. Pinault, L., Tjepkema, M., Crouse, D. L., et al., Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort, Environ Health-Glob, 15, doi:10.1186/s12940-016-0111-6, 2016.

39. Pope, C. A., Burnett, R. T., Thurston, G. D., et al., Cardiovascular mortality and long-term exposure to particulate air pollution - Epidemiological evidence of general pathophysiological pathways of disease, Circulation, 109, 71-77, doi:10.1161/01.Cir.0000108927.80044.7f, 2004.

40. Protonotariou, A. P., Bossioli, E., Tombrou, M., et al., Air Pollution in Eastern Mediterranean: Nested-Grid GEOS-CHEM Model Results and Airborne Observations. Advances in Meteorology, Climatology and Atmospheric Physics. Springer Atmospheric Sciences, Springer, Berlin, Heidelberg, 2013, pp. 1203-1209.

41. Shaddick, G., Thomas, M. L., Green, A., et al., Data integration model for air quality: a hierarchical approach to the global estimation of exposures to ambient air pollution, J R Stat Soc C-Appl, 67, 231-253, doi:10.1111/rssc.12227, 2018.

42. Shindell, D., Faluvegi, G., Seltzer, K., et al., Quantified, localized health benefits of accelerated carbon dioxide emissions reductions, Nat Clim Change, 8, doi:10.1038/s41558-018-0108-y, 2018.

43. Tseng, E., Ho, W. C., Lin, M. H., et al., Chronic exposure to particulate matter and risk of cardiovascular mortality: cohort study from Taiwan, Bmc Public Health, 15, doi:10.1186/s12889-015-2272-6, 2015.

44. Ueda, K., Nagasawa, S., Nitta, H., et al., Exposure to Particulate Matter and Long-term Risk of Cardiovascular Mortality in Japan: NIPPON DATA80, J Atheroscler Thromb, 19, 246-254, doi:10.5551/jat.9506, 2012.

45. Veefkind, J. P., Boersma, K. F., Wang, J., et al., Global satellite analysis of the relation between aerosols and short-lived trace gases, Atmos Chem Phys, 11, 1255-1267, doi:10.5194/acp-11-1255-2011, 2011.

46. Vodonos, A., Abu Awad, Y., Schwartz, J., The concentration-response between long-term $PM_{2.5}$ exposure and mortality; A meta-regression approach, Environ Res, 166, 677-689, doi:10.1016/j.envres.2018.06.021, 2018.

47. Vohra, K., Marais, E. A., Suckra, S., et al., Long-term trends in air quality in major cities in the UK and India: A view from space, Atmospheric Chemistry and Physics Discussions, doi:10.5194/acp-2020-342, 2020.

48. Wang, Y., Shi, L. H., Lee, M., et al., Long-term Exposure to $PM_{2.5}$ and Mortality Among Older Adults in the Southeastern US, Epidemiology, 28, 207-214, doi:10.1097/Ede.0000000000000614, 2017.

49. WHO, World Health Organization; Principles for evaluating health risks in children associated with exposure to chemicals. https://apps.who.int/iris/handle/10665/43604, 2006.

50. WHO, World Health Organization; Ambient air pollution: A global assessment of exposure and burden of disease. https://www.who.int/phe/publications/air-pollution-global-assessment/en/, 2016.

51. WONDER, C., Centers for Disease Control and Prevention Wide-ranging ONline Data for Epidemiologic Research https://wonder.cdc.gov/.

52. Wong, C. M., Lai, H. K., Tsang, H., et al., Satellite-Based Estimates of Long-Term Exposure to Fine Particles and Association with Mortality in Elderly Hong Kong Residents, Environ Health Persp, 123, 1167-1172, doi:10.1289/ehp.1408264, 2015.

53. Wong, C. M., Tsang, H., Lai, H. K., et al., Cancer Mortality Risks from Long-term Exposure to Ambient Fine Particle, Cancer Epidem Biomar, 25, 839-845, doi:10.1158/1055-9965.Epi-15-0626, 2016.

54. Xing, J., Mathur, R., Pleim, J., et al., Can a coupled meteorology-chemistry model reproduce the historical trend in aerosol direct radiative effects over the Northern Hemisphere?, Atmos Chem Phys, 15, 9997-10018, doi:10.5194/acp-15-9997-2015, 2015.

55. Xu, Z. W., Sheffield, P. E., Hu, W. B., et al., Climate Change and Children's Health-A Call for Research on What Works to Protect Children, Int J Env Res Pub He, 9, 3298-3316, doi:10.3390/ijerph9093298, 2012.

56. Yin, P., Brauer, M., Cohen, A., et al., Long-term Fine Particulate Matter Exposure and Nonaccidental and Cause-specific Mortality in a Large National Cohort of Chinese Men, Environ Health Persp, 125, doi:10.1289/Ehp1673, 2017.

57. Zhai, S. X., Jacob, D. J., Wang, X., et al., Fine particulate matter ($PM_{2.5}$) trends in China, 2013-2018: separating contributions from anthropogenic emissions and meteorology, Atmos Chem Phys, 19, 11031-11041, doi:10.5194/acp-19-11031-2019, 2019.

58. Zhang, L., Jacob, D. J., Knipping, E. M., et al., Nitrogen deposition to the United States: distribution, sources, and processes, Atmos Chem Phys, 12, 4539-4554, doi:10.5194/acp-12-4539-2012, 2012.

Supplemental Material

Global mortality from outdoor fine particle pollution generated by fossil fuel combustion: Results from GEOS-Chem

Karn Vohra[1*], Alina Vodonos[2], Joel Schwartz[2], Eloise A. Marais[3,a], Melissa P. Sulprizio[4], Loretta J. Mickley[4]

[1] School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

[2] Harvard T.H. Chan School of Public Health, Department of Environmental Health, Harvard University, Boston, MA, USA

[3] Department of Physics and Astronomy, University of Leicester, Leicester, UK

[4] John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

[a] Now at: Department of Geography, University College London, London, UK

* Corresponding author: Karn Vohra, Phone: +44 7716496867,

Email: kxv745@student.bham.ac.uk

$PM_{2.5}$ mortality concentration –response model

Figure S5. Estimates for long-term $PM_{2.5}$ mortality dose-response, drawn from the meta-analysis of long-term association between $PM_{2.5}$ and mortality.

Figure S6. Hazard Ratio based on GEMM function (Burnett et al., 2018) compared to the Hazard Ration based on the meta-analysis.

Table S2. Extended data. Global regions, number of deaths, attributable fraction (%) for the population above 14 years old attributable to fine particulate matter ($PM_{2.5}$) exposure.

References

Description of GEOS-Chem.

GEOS-Chem is a three-dimensional chemical transport model that includes detailed oxidant-aerosol chemistry in the troposphere and is used by more than 80 groups worldwide (www.geos-chem.org). The model is widely cited in the peer-reviewed literature – e.g., more than 4000 times in the year 2017 alone (http://acmg.seas.harvard.edu/geos/geos_pub.html). The model has been frequently applied to interpret observed $PM_{2.5}$ in regions dominated by anthropogenic sources – e.g., China (Aunan et al., 2018), Korea (Lee et al., 2017), India (Venkataraman et al., 2018), and the US (Di et al., 2016; Silvern et al., 2017); and validation has been performed for specific source sectors – e.g., transportation (Travis et al., 2016), biogenic sources (Marais et al., 2017), and power plants (S. W. Wang et al., 2012). Here we use GEOS-Chem v10-01, driven by 2012 GEOS-5 meteorology (gmao.gsfc.nasa.gov/GEOS_systems/). The GEOS-5 data are produced at 0.5°×0.667° horizontal resolution and are re-gridded here to 2°×2.5° for the global simulation. We also perform four regional simulations – for Europe, North America, Africa, and Asia – and for these simulations we keep the native grid resolution. Boundary conditions at 2°×2.5° from the global simulation are applied to these regional simulations. Most fine-scale, regional models, such as the Community Multiscale Air Quality Model, rely on chemical boundary conditions from global models with different chemical schemes, but our approach permits application of a consistent scheme across the globe. The 0.5°×0.667° horizontal resolution in GEOS-Chem over key regions is, however, relatively coarse compared to that in some other regional models. Y. Li et al. (2016) show that application of coarse resolution leads

to an underestimate of health impacts of 8%, implying that our mortality estimates are conservative. Our choice of 2012 as the simulation year is discussed below.

GEOS-Chem simulates the mass concentrations of key particle types including sulfate, nitrate, and ammonium (Park et al., 2004; L. Zhang et al., 2012), organic carbon (Heald et al., 2006; 2011) black carbon (Q. Q. Wang et al., 2014), dust (Fairlie et al., 2007), and sea salt (Jaegle et al., 2011). Particle chemistry is coupled to gas-phase chemistry as described by (Mao et al., 2013). Gas/particle partitioning of sulfate, nitrate and ammonium (SNA) particles is computed with the ISORROPIA II thermodynamic module (Fountoukis and Nenes, 2007; Pye et al., 2009). Wet and dry deposition of particles follow Liu et al. (2001) and L. M. Zhang et al. (2001), respectively.

Emissions in GEOS-Chem are computed by the Harvard-NASA Emission Component (HEMCO) (Keller et al., 2014), which combines and regrids ensembles of user-selected emission inventories. We apply global anthropogenic emissions but supersede these with regional emissions where the latter are more reliable (Table 1). Fossil fuel emissions in Africa include (1) industry and power plants from the global inventories and (2) diffuse and inefficient combustion sources (diesel and petrol generators, ad-hoc oil refining, gas flares, kerosene use, cars, and motorcycles) from the DICE-Africa inventory (Marais and Wiedinmyer, 2016). We scale all anthropogenic inventories to 2012, as described by van Donkelaar et al. (2008). Biogenic emissions are from MEGAN v2.1 for volatile organic compounds (Guenther et al., 2012) and from Hudman et al. (2012) for soil nitrogen oxides. Lightning emissions of nitrogen oxides are computed as a function of cloud top height as described by Murray et al. (2012). Dust entrainment and deposition follow the DEAD scheme of Zender et al. (2003) as implemented in GEOS-Chem by Fairlie et al. (2007). Biomass burning emissions are from the Global Fire Emissions Database version 4 (GFED4) (Giglio et al., 2013).

For this study, we first calculate the surface fine particle mass concentrations ($PM_{2.5}$), with all emissions sources turned on. For consistency with the $PM_{2.5}$ measurement protocol set by the U.S. Environmental

Protection Agency, we assume 35% relative humidity everywhere (except for Europe) and standard ambient conditions, with temperature of 298.15 K and surface pressure of 1013.25 hPa. In Europe, we assume 50% relative humidity, as is the protocol there. We then perform the identical simulation with emissions arising from fossil fuel combustion turned off. The same meteorological fields are applied for both simulations – i.e., the simulation does not allow feedbacks from particles onto meteorology. In the no-fossil-fuel case, all fossil fuel sources are turned off in both the nested simulations and in the global simulation providing boundary conditions. The difference between the two simulations (with and without fossil fuel) represents the contribution of fossil-fuel combustion to surface $PM_{2.5}$. This approach assumes a linear response of surface $PM_{2.5}$ to changes in emissions.

Our choice of 2012 as the simulation year requires explanation. Air quality is influenced not just by emissions but also by meteorological variables such as surface temperature and wind speed, which can vary greatly on inter-annual timescales. Ideally, our analysis would involve multi-year simulations on both the coarse- and fine-scale grids, but such effort would be computationally expensive. We choose instead to do a one-year simulation for a year not influenced by El Niño conditions, which can worsen or ameliorate air pollution, depending on the region (e.g., Chang et al. (2016), Shen and Mickley (2017)). To gauge the error implied by our choice to simulate just one year rather than a span of years, we examine the inter-annual variation in total $PM_{2.5}$ concentrations at the surface estimated from the Dalhousie University archive (van Donkelaar et al., 2016). The $PM_{2.5}$ values in the Dalhousie archive are calculated by first combining satellite observations with GEOS-Chem estimates, and then calibrating the resulting concentrations with available ground-based observations (mostly Europe, the US, India and China). We find that the global mean average of the relative standard deviation of total $PM_{2.5}$ in the Dalhousie archive over 2008 to 2016 is just 7%. Averaged over large regions on the continental scale, the relative standard deviation ranges from 4% over Australia to 11% over the Asia nested grid domain (Figure S1). Inter-annual variability in this metric is greatest (> 60%) for smaller regions influenced by wildfires or biomass burning – e.g., Indonesia and remote areas at high northern latitudes where few people live. To test our choice of 2012 as a representative year,

we calculate the 2012 anomaly in the Dalhousie $PM_{2.5}$ time series (Figure S2). Again on a continental scale, we find that 2012 concentrations range from 0.7 μg m^{-3} less to 0.4 μg m^{-3} greater than the 2008-2016 average (Figure S2). Given the relatively small inter-annual variability in surface $PM_{2.5}$ in the Dalhousie archive over most populated regions, as well as the small anomalies in $PM_{2.5}$ in 2012 relative to the long-term mean, we conclude that the 2012 GEOS-Chem simulation provides a representative snapshot of global air quality.

To validate the 2012 $PM_{2.5}$ results from GEOS-Chem, we rely on archived $PM_{2.5}$ concentrations from the World Health Organization database (WHO). We find that GEOS-Chem captures the observed annual mean $PM_{2.5}$ concentrations with a correlation of 0.70, mean absolute error of 3.4 μg m^{-3}, and normalized mean bias of 27% (Figure S3). Our high bias in the US (where most North American WHO data are located) is opposite to the low bias estimated by Ford and Heald (2016) in urban (-25%) and rural (-6%) areas; such biases may be due to differences in US emission inventories for both gas-phase aerosol precursors and primary particles (Xing et al., 2015). A caveat in our comparison is that most observations (95%) in the WHO database with at least 75% temporal coverage in 2012 are in North America and Europe. We add to Figure S3 the 2012 observations from the US embassy in Shanghai (those for Beijing are already in the WHO dataset), and national monitoring sites embassies in Delhi (Cusworth et al., 2018), and the Highveld region in South Africa (South African Air Quality Information System; data obtained by request from the South African Weather Service in July 2018). Over the European domain in Figure S1, we find that GEOS-Chem yields a correlation of 0.60, mean absolute error of 5.2 μg m^{-3} and a normalized mean bias of 33% in surface $PM_{2.5}$; over the North American domain in Figure S1, these values are 0.52, 1.8 μg m^{-3} and 20% (Figure S4). Taken together, these validation statistics are similar to those reported by other studies examining surface $PM_{2.5}$ in global models (e.g., Shindell et al. (2018)) and regional models (e.g., Xing et al. (2015)) .

Table S1. GEOS-Chem anthropogenic emissions. All emissions are scaled to 2012 conditions.

Region	Inventory name	Species	Reference
Global	EDGAR v4.2[a,c]	NO, CO, SO_2, sulfate, ammonia	Olivier and Berdowski (2001)
Global	RETRO[a,c]	Non-methane VOCs	Schultz et al. (2007)
Global	---	Ethane	Xiao et al. (2008)
Global	GEIA	Biofuel ammonia	www.geiacenter.org
Global	BOND[a,c]	Carbonaceous particles	Bond et al. (2007)
Global	AEIC v2.0 aircraft	NO, CO, etc.	Stettler et al. (2011)
Global	ARTCAS ship	SO_2	Eyring et al. (2005)
Global	ICOADS ship	CO	C. Wang et al. (2008)
Global	PARANOX ship	NO	Vinken et al. (2011)
United States	NEI 2011[a,b,c]	Many species	US EPA, www3.epa.gov/airtrends
Europe	EMEP[b,c]	Many species	www.emep.int
Asia	MIX[c]	Many species	M. Li et al. (2017), Venkataraman et al. (2018), X. Li et al. (2018)
Africa	DICE[c,d]	Many species	Marais and Wiedinmyer (2016)
Africa	---	Open waste burning species	Wiedinmyer et al. (2014)

[a] Includes biofuel sources

[b] Includes ship emissions

[c] Includes land-based transport emissions

[d] Includes only diffuse and inefficient sources of anthropogenic emissions – residential fuelwood, diesel and petrol generators, ad-hoc oil refining, gas flares, kerosene use, charcoal production and use, road transport (including motorcycles). For emissions from formal industry and powerplants, we use the global inventories.

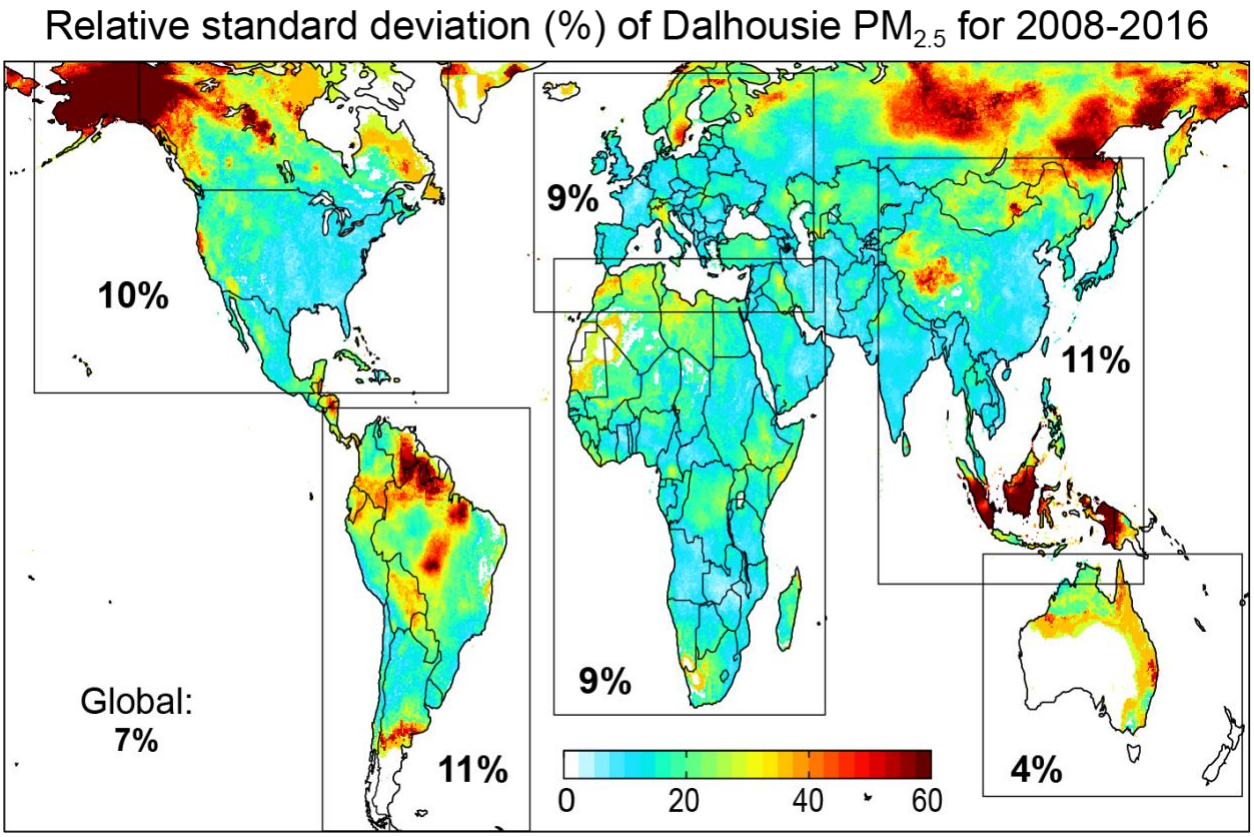


Figure S1. Uncertainty in 2012 $PM_{2.5}$ due to interannual variability. Interannual variability is estimated as the relative standard deviation of the Dalhousie satellite-derived $PM_{2.5}$ product (van Donkelaar et al., 2016) for 2008-2016 at 0.1°×0.1°. Values inset are the domain mean relative standard deviations for North America, South America, Western Europe (including portions of North Africa and the Middle East), Africa (including a portion of the Middle East), Southeast Asia, and Australia.

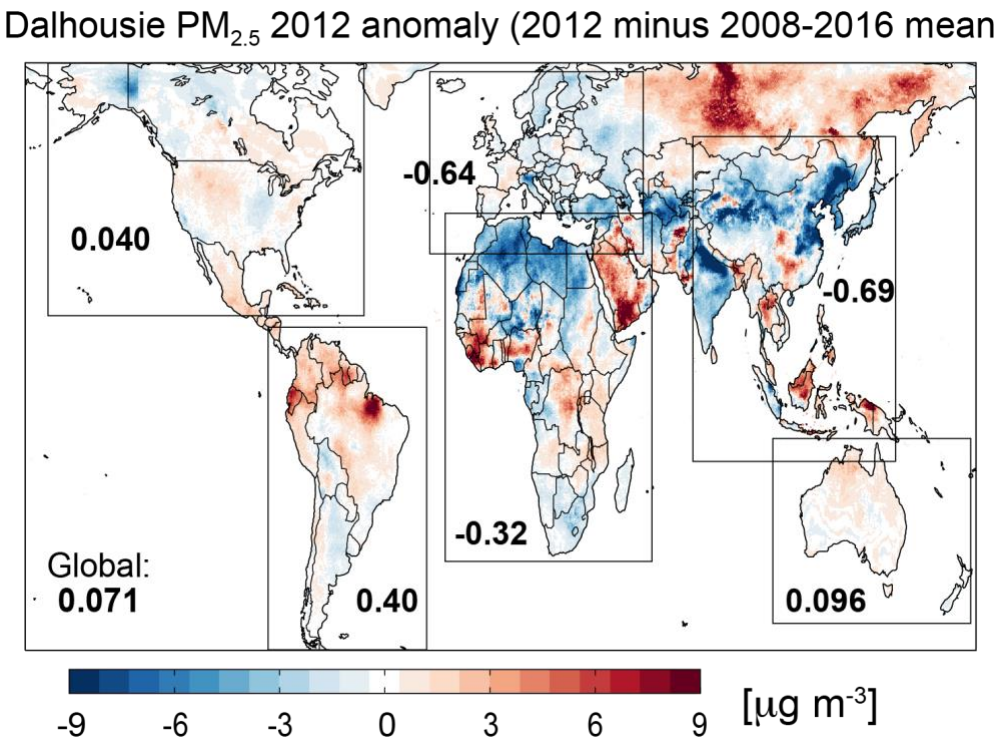


Figure S2. Representativeness of $PM_{2.5}$ in 2012, calculated as the absolute difference in 2012 and 2008-2016 mean $PM_{2.5}$ from Dalhousie (van Donkelaar et al., 2016) at 0.1°×0.1°. Values inset are domain mean anomalies for North America, South America, Western Europe (including portions of North Africa and the Middle East), Africa (including a portion of the Middle East), Southeast Asia, and Australia.

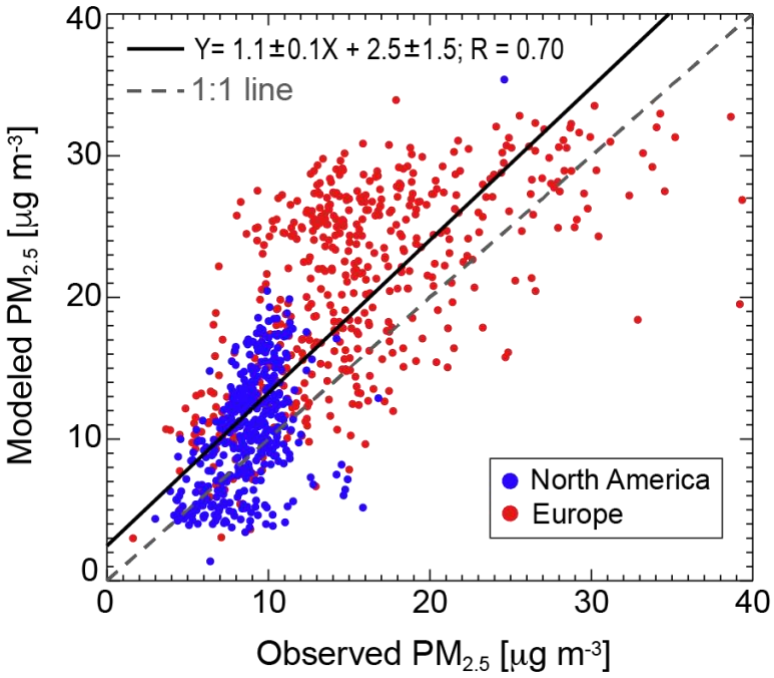


Figure S3. Evaluation of GEOS-Chem $PM_{2.5}$. Points are annual mean $PM_{2.5}$ for coincident 0.5°×0.667° grid squares with at least 75% temporal coverage in the observations. GEOS-Chem $PM_{2.5}$ is estimated at 50% relative humidity (RH) in Europe and 35% RH everywhere else, following standard protocols in measurements of $PM_{2.5}$. Reduced major axis (RMA) regression line (solid black line) and statistics, and the Pearson's correlation coefficient for all coincident grid squares are given inset. Points in red are in Europe and in blue are in North America. Only 7 out of 957 points exceed the range shown.

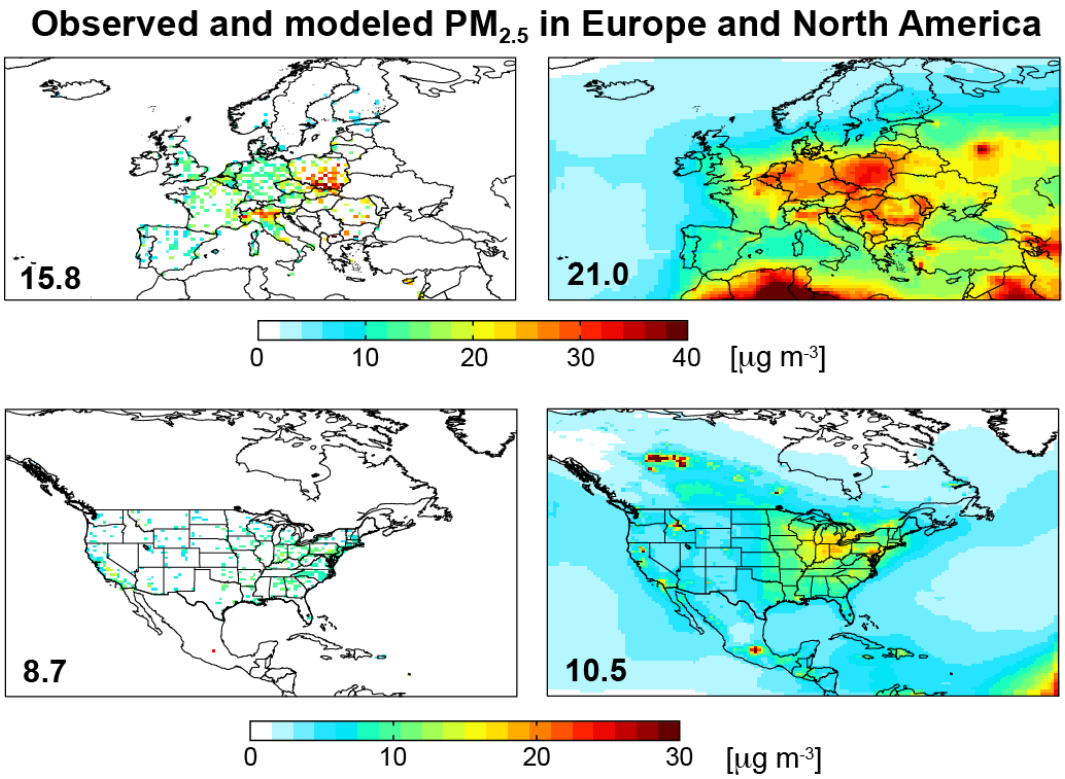


Figure S4. Comparison of the spatial distribution of observed and modeled $PM_{2.5}$ in Europe and North America. Data are on a uniform 0.5°×0.667° grid. Only observations with at least 75% temporal coverage are used. $PM_{2.5}$ are obtained at 50% RH in Europe and 35% RH in North America. Data for the two domains are plotted on different scales. Mean $PM_{2.5}$ for coincident grid squares is given inset

$PM_{2.5}$ mortality concentration –response model

We estimated the number of premature deaths attributable to fossil-fuel related $PM_{2.5}$ using a health impact function. To estimate the excess number of deaths associated with $PM_{2.5}$ exposure one requires estimates of exposure, the size of the population exposed, the mortality rate for that population, and the fraction of total deaths attributable to that exposure (AF%).

Recent meta-analysis of the association between long-term $PM_{2.5}$ and mortality (Vodonos et al., 2018) applied a multivariate linear random effects meta-analysis and meta-regression models that polled 135 hazard ratio estimates derived from 53 studies examined long-term $PM_{2.5}$ and mortality. This meta-analysis provided an evidence of a nonlinear association where the exposure-mortality slopes decreased at higher concentrations (**Figure S5)**. For example, each 1 μg m^{-3} increase in $PM_{2.5}$ was associated with a 1.29% increase in all-age all-cause mortality (95%CI 1.09-1.50) at a mean exposure of 10 μg m^{-3}, which decreased to 0.94 % (95%CI 0.76-1.12) at a mean exposure of 20 μg m^{-3}, to 0.81% (95%CI 0.52-1.12) at 30 μg m^{-3} and to 0.79% (95%CI 0.40-1.13) at 40 μg/$m^{3.}$

Hence, for examining a reduction of $PM_{2.5}$ levels from 15 to 10 μg/m^3, we calculated the mean slope as area under the curve between 0.014 and 0.011= 0.0125. A reduction of $PM_{2.5}$ levels from 30 to 20 μg/m^3, the mean slope was calculated as area under the curve between 0.009 and 0.008 = 0.00814

Mean value of estimates of mortality ($\bar{\beta}$) for each grid cell was calculated as area under the curve for the concentration-specific β in each grid cell from the low $PM_{2.5}$ scenario (without fossil fuel emissions) to the high $PM_{2.5}$ scenario (with all emissions, including fossil fuel) following the form shown in Equation

$$\bar{\beta}(PM_{2.5})=\int_{PM_{2.5}\ \text{no fossil fuel}}^{PM_{2.5}\ \text{all emissions}} \beta(PM_{2.5})$$

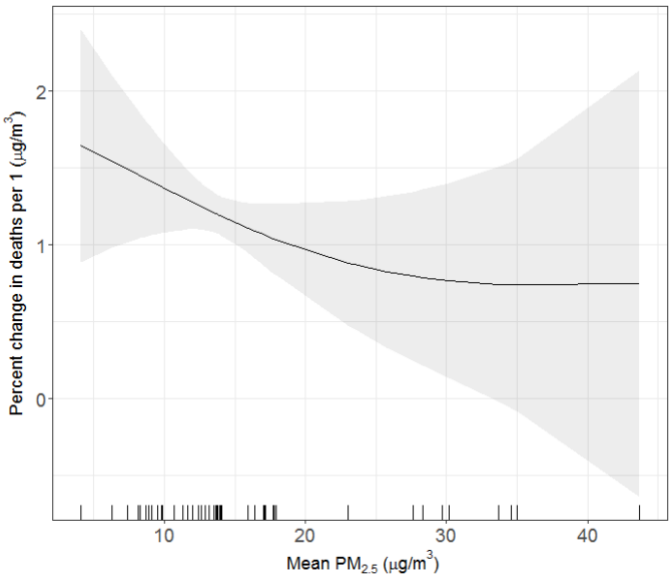


Figure S5. Estimates for long-term $PM_{2.5}$ mortality dose-response, drawn from the meta-analysis of long-term association between $PM_{2.5}$ and mortality (Vodonos et al., 2018).

Table S2. Extended data. Global regions, number of deaths, attributable fraction (%) for the population above 14 years old attributable to fine particulate matter ($PM_{2.5}$) exposure in 2012

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
North America						
Bermuda	488	3	1.9	1.1	9	1.8
Greenland	472	1.2	0.9	0.3	3	0.6
Central America & the Caribbean						
Antigua and Barbuda	538	4.4	4.1	0.3	2	0.4
Bahamas	2,347	4.1	2.8	1.4	53	2.3
Barbados	2,523	4.9	4.7	0.2	7	0.3
Belize	1,530	5	4	1.1	26	1.7
Costa Rica	38,094	5.4	2.9	2.6	1,557	4.1
Cuba	95,635	5.3	3.8	1.5	2,334	2.4
Dominica	668	4.9	4.7	0.2	2	0.3
Dominican Republic	60,949	11.2	5.3	6	4,925	8.1
El Salvador	44,036	9.7	3.4	6.3	4,029	9.1
Grenada	983	4.6	4.3	0.4	6	0.6
Guatemala	67,426	9.7	3.2	6.5	6,205	9.2
Haiti	70,013	8.2	4.9	3.3	3,409	4.9
Honduras	40,564	7.9	3.5	4.4	2,620	6.5
Jamaica	18,511	9.1	4.7	4.4	1,183	6.4

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Mexico	615,874	11.8	2.4	9.5	65,871	10.7
Nicaragua	20,467	5.4	3.5	1.9	614	3.0
Panama	16,364	4.7	2.5	2.2	594	3.6
Puerto Rico	28,717	5.5	4.6	0.9	409	1.4
Saint Lucia	1,191	5	4.8	0.2	4	0.3
Saint Vincent and the Grenadines	913	4.7	4.5	0.2	3	0.3
Trinidad and Tobago	19,561	5.4	4.5	0.9	277	1.4
United States Virgin Islands	1,202	4.6	4.2	0.4	7	0.6
South America						
Argentina	306,979	7.9	3.4	4.5	20,385	6.6
Bolivia	50,854	5.7	4.4	1.3	1,095	2.2
Brazil	1,161,922	8.9	2.9	6.1	94,216	8.1
Chile	108,995	10	2.4	7.6	11,202	10.3
Colombia	247,981	8.2	2.7	5.5	20,045	8.1
Ecuador	74,588	6.7	2.1	4.6	5,357	7.2
Guyana	4,830	8	6.6	1.4	96	2.0
Paraguay	29,665	9.2	6	3.2	1,374	4.6
Peru	120,778	7.3	1.8	5.5	10,209	8.5
Suriname	3,667	6.9	6.2	0.7	36	1.0
Uruguay	30,980	6.5	2.4	4.1	1,967	6.3
Venezuela	247,407	10.6	4.3	6.2	21,185	8.6

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Europe						
Albania	20,072	19.8	8.6	11.2	2,458	12.2
Andorra	654	13.4	5.8	7.6	65	9.9
Austria	79,627	21.4	4.3	17.1	15,018	18.9
Belarus	115,131	20.6	2.9	17.8	23,397	20.3
Belgium	108,113	25.5	2.8	22.7	25,633	23.7
Bosnia and Herzegovina	36,427	21	6.8	14.2	5,628	15.5
Bulgaria	106,938	20.2	7.2	13	15,346	14.4
Croatia	52,156	20.2	5.6	14.6	8,454	16.2
Cyprus	7,171	15.4	9.2	6.3	543	7.6
Czech Republic	109,205	26.2	3.4	22.8	25,467	23.3
Denmark	51,600	16.3	2.1	14.2	9,202	17.8
Estonia	14,761	12.6	1.6	11	2,227	15.1
Finland	50,553	8.6	1.3	7.3	5,506	10.9
France	562,481	18.1	3.4	14.7	97,242	17.3
Georgia	51,550	23.3	10.2	13.1	6,670	12.9
Germany	896,319	23.9	3.2	20.7	198,569	22.2
Greece	116,757	15.6	8.1	7.5	10,616	9.1
Hungary	128,981	24.7	4.7	20	26,863	20.8
Iceland	1,891	2.6	1.6	1	31	1.6
Ireland	30,421	8.3	2	6.4	2,902	9.5
Italy	622,080	18.8	6	12.8	89,412	14.4
Kazakhstan	126,168	17.1	9.2	7.9	11,343	9.0

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Latvia	31,672	16.2	2	14.3	5,719	18.1
Lithuania	40,380	21.4	2.3	19.1	8,729	21.6
Malta	3,593	16	11.4	4.6	193	5.4
Moldova	43,245	25.4	5.2	20.2	8,922	20.6
Montenegro	6,223	18	7.9	10.1	724	11.6
Netherlands	143,387	24.2	2.7	21.5	32,972	23.0
Norway	29,299	5.9	1.4	4.5	2,065	7.0
Poland	393,724	26.5	3.1	23.4	93,842	23.8
Portugal	104,738	8.9	3.7	5.2	8,032	7.7
Romania	269,933	23.9	6.2	17.7	49,583	18.4
Russia	1,833,839	19	4.9	14.1	289,922	15.8
Serbia	100,172	24.8	6.9	17.9	18,076	18.0
Slovakia	53,258	24.9	4.1	20.8	11,522	21.6
Slovenia	19,680	21.7	5.3	16.3	3,528	17.9
Spain	418,063	12.9	4.8	8.1	44,603	10.7
Sweden	88,058	10	1.6	8.5	10,548	12.0
Switzerland	62,993	20.3	4.6	15.8	11,196	17.8
Turkey	361,723	18.2	8.1	10.1	41,811	11.6
Ukraine	731,672	19.4	5.1	14.3	120,217	16.4
United Kingdom	579,747	15.4	2	13.5	99,069	17.1

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Africa						
Algeria	142,304	31.4	20.5	10.9	13,295	9.3
Angola	100,845	15.4	14.1	1.3	1,537	1.5
Benin	42,616	40.4	36.2	4.2	1,450	3.4
Botswana	12,721	8.2	6	2.1	397	3.1
Burkina Faso	84,040	55.9	54.6	1.3	855	1.0
Burundi	44,973	16.2	15.4	0.8	419	0.9
Cameroon	118,759	39.7	38.2	1.5	1,520	1.3
Cape Verde	2,545	66.9	66	0.9	18	0.7
Central African Republic	41,111	30.7	30.1	0.6	178	0.4
Chad	56,523	59.8	58.7	1	460	0.8
Comoros	3,878	1.6	1.4	0.1	9	0.2
Congo	21,705	20.6	19.3	1.3	287	1.3
Cote d'Ivoire	111,211	29.3	28.2	1.1	1,065	1.0
Democratic Republic of the Congo	419,021	21.3	20.7	0.6	2,261	0.5
Djibouti	4,509	21.2	17.5	3.8	164	3.6
Egypt	392,226	56.7	40.2	16.5	46,783	11.9
Equatorial Guinea	4,679	10	9.5	0.5	32	0.7
Eritrea	20,386	31.3	28.5	2.8	444	2.2
Ethiopia	287,855	17	15.2	1.8	5,657	2.0
Gabon	13,783	11	10.5	0.5	90	0.7
Gambia	9,610	58	56	2	151	1.6

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Ghana	149,177	31.5	28.9	2.6	3,361	2.3
Guinea	63,691	49.7	48.8	1	467	0.7
Guinea-Bissau	9,223	51.9	50.6	1.3	89	1.0
Kenya	219,806	8.3	6.4	2	6,035	2.7
Lesotho	25,223	12.6	7.5	5.1	1,689	6.7
Liberia	19,482	25.3	24.7	0.7	113	0.6
Libya	26,745	42.3	34.3	8	1,565	5.9
Madagascar	97,088	3.7	3.3	0.4	641	0.7
Malawi	83,919	9.9	9.4	0.6	681	0.8
Mali	69,737	60.3	59.3	1	555	0.8
Mauritania	13,520	98.7	97.4	1.3	159	1.2
Mauritius	9,564	1.6	1.3	0.3	43	0.4
Morocco	186,609	23.8	16.9	6.9	12,436	6.7
Mozambique	163,474	6.8	6.3	0.5	1,309	0.8
Namibia	12,923	11.1	10.2	0.9	159	1.2
Niger	63,052	73.3	71.6	1.7	844	1.3
Nigeria	689,902	59.7	54.9	4.8	25,282	3.7
Rwanda	43,547	16.4	15.2	1.2	557	1.3
Sao Tome and Principe	821	5.5	5.4	0.1	2	0.2
Senegal	61,877	71.2	69.3	1.8	916	1.5
Seychelles	702	1.5	1.2	0.3	4	0.6
Sierra Leone	33,549	42	41	0.9	230	0.7
Somalia	47,288	9.5	8.3	1.3	789	1.7

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
South Africa	487,500	21.9	11.7	10.2	45,134	9.3
Sudan[c]	165,624	35.3	33.6	1.7	2,197	1.3
Swaziland	9,954	10.6	6.7	3.9	534	5.4
Tanzania	202,713	6.9	6.4	0.5	1,660	0.8
Togo	34,797	36.6	34.4	2.1	617	1.8
Tunisia	59,521	25.5	17.1	8.3	4,711	7.9
Uganda	127,825	13.1	11.8	1.3	2,018	1.6
Zambia	71,697	12.7	12.2	0.6	511	0.7
Zimbabwe	88,229	10.5	9	1.6	1,797	2.0
Western Asia & the Middle East						
Afghanistan	148,817	20.9	13.9	7	11,153	7.5
Armenia	25,420	22.6	11.9	10.7	2,721	10.7
Azerbaijan	85,764	29.8	17.6	12.2	8,733	10.2
Bahrain	3,315	33.1	30.2	2.9	73	2.2
Iran	330,324	28.5	23.8	4.7	13,168	4.0
Iraq	95,874	30.1	26.4	3.7	2,942	3.1
Israel	40,291	21.2	14.4	6.9	2,776	6.9
Jordan	13,031	22.9	16.6	6.2	766	5.9
Kuwait	5,120	37.4	34.4	3	110	2.1
Kyrgyzstan	29,441	17.3	8.4	8.9	3,041	10.3
Lebanon	27,756	18	11.7	6.3	1,931	7.0
Oman	7,482	46.5	40.6	5.8	321	4.3

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Palestine	12,562	22.7	15.6	7.1	853	6.8
Qatar	4,252	35.2	31.7	3.5	109	2.6
Saudi Arabia	82,403	32.6	29.6	3	1,893	2.3
Syria	140,751	19.4	12.7	6.7	10,159	7.2
Tajikistan	38,948	21.7	9.6	12.1	4,914	12.6
Turkmenistan	51,096	31.7	26.4	5.3	2,124	4.2
United Arab Emirates	16,636	54	45.8	8.1	1,000	6.0
Uzbekistan	205,829	24.8	12.8	12	23,912	11.6
Yemen	90,616	23	19.9	3.1	2,520	2.8
Eastern Asia						
Bangladesh	692,081	58.9	6.7	52.3	252,927	36.5
Bhutan	2,909	23.6	5.7	17.9	516	17.7
Brunei	1,684	6.1	3.3	2.8	72	4.3
Cambodia	85,803	20.9	11.6	9.2	8,445	9.8
China	9,720,397	72.8	9.9	62.9	3,910,916	40.2
China (2018)[d]	9,720,397	41	9.7	31.2	2,355,579	24.2
India	8,009,357	52	9	42.9	2,458,384	30.7
Indonesia	1,495,066	20.9	5.7	15.3	230,097	15.4
Japan	1,284,769	22.6	4.6	18	242,561	18.9
Laos	33,822	19.6	8	11.6	4,404	13.0
Malaysia	154,090	18.9	5.3	13.6	22,228	14.4
Maldives	865	5.9	2.3	3.7	50	5.8

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Mongolia	12,013	8.4	4.8	3.5	628	5.2
Myanmar	340,623	16.4	7.4	9	36,978	10.9
Nepal	168,690	38.8	9.5	29.3	39,066	23.2
North Korea	201,841	35.8	5.3	30.5	52,942	26.2
Pakistan	1,115,784	36.7	15.1	21.7	188,406	16.9
Papua New Guinea	63,224	3.1	2.9	0.2	168	0.3
Philippines	559,792	8.7	2.1	6.7	51,203	9.1
Singapore	14,100	21.9	4.9	16.9	2,616	18.6
South Korea	265,641	44	5.3	38.8	80,962	30.5
Sri Lanka	116,032	13.4	3.5	9.9	14,998	12.9
Taiwan	164,488	14.5	3.2	11.3	23,711	14.4
Thailand	418,824	20.6	4.7	15.9	71,184	17.0
Timor-Leste	5,381	6.4	5	1.4	115	2.1
Vietnam	541,064	31.7	4.4	27.4	127,614	23.6
Australia & Oceania						
American Samoa	301	0.7	0.7	0	0	0.0
Australia	142,935	4.9	2.4	2.5	5,686	4.0
Federated States of Micronesia	679	0.7	0.7	0	1	0.1
Fiji	5,538	1.3	1.2	0.1	9	0.2
Guam	1,112	1.2	1	0.2	4	0.4
Kiribati	852	0.8	0.8	0	0	0.0
Marshall Islands	336	1.1	1.1	0	0	0.0

Country name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
New Zealand	29,923	2.2	1.5	0.6	320	1.1
Northern Mariana Islands	249	1.3	1.1	0.3	1	0.4
Samoa	960	0.7	0.7	0	0	0.0
Solomon Islands	3,286	1.2	1.2	0	2	0.1
Tonga	657	1.2	1.1	0.1	1	0.2
Vanuatu	1,791	2.2	2.2	0.1	2	0.1

USA						
State name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Alabama	50,411	9.4	2.6	6.9	5,067	10.1
Alaska	3,384	2.2	1.4	0.9	51	1.5
Arizona	56,565	7.9	4	3.9	3,263	5.8
Arkansas	26,345	10.3	2.6	7.6	2,887	11.0
California	259,363	12.2	2.4	9.8	34,081	13.1
Colorado	36,885	6.8	3	3.8	2,140	5.8
Connecticut	32,639	12.1	1.7	10.5	4,749	14.6
Delaware	4,436	13.2	1.7	11.5	694	15.6
Florida	191,646	6.6	2.4	4.2	12,483	6.5
Georgia	75,518	11.3	2.5	8.8	9,290	12.3

State name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Hawaii	11,032	2.6	2.1	0.4	83	0.8
Idaho	13,006	6.2	3.3	2.8	581	4.5
Illinois	102,593	16.6	1.9	14.7	18,952	18.5
Indiana	66,979	17	1.9	15.1	12,637	18.9
Iowa	33,378	11.9	2.1	9.8	4,562	13.7
Kansas	33,671	10.4	1.9	8.5	4,094	12.2
Kentucky	52,325	14.3	2	12.4	8,500	16.2
Louisiana	42,176	10.4	2.8	7.5	4,505	10.7
Maine	14,555	7.7	1.6	6.1	1,350	9.3
Maryland	40,784	15.8	1.8	14.1	7,336	18.0
Massachusetts	53,851	11.8	1.6	10.2	7,654	14.2
Michigan	93,585	16.7	1.8	14.9	17,438	18.6
Minnesota	39,674	13.3	2.2	11.1	5,877	14.8
Mississippi	40,360	10	2.6	7.3	4,263	10.6
Missouri	48,205	11.2	2.1	9.1	6,161	12.8
Montana	9,520	5.1	3.4	1.7	266	2.8
Nebraska	13,881	9	2.1	7	1,432	10.3
Nevada	23,541	6.7	3.4	3.3	1,192	5.1
New Hampshire	12,314	10	1.6	8.3	1,495	12.1
New Jersey	97,747	15.7	1.6	14.1	17,646	18.1
New Mexico	21,308	4.9	2.2	2.7	938	4.4
New York	129,489	14.6	1.6	13	21,931	16.9
North Carolina	95,239	12.5	2.2	10.3	13,357	14.0

State name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
North Dakota	4,070	6.9	2	4.9	309	7.6
Ohio	115,955	16.8	1.7	15	21,818	18.8
Oklahoma	40,908	8.7	1.9	6.8	4,190	10.2
Oregon	38,128	8.1	2.4	5.6	3,152	8.3
Pennsylvania	133,771	17.1	1.7	15.4	25,382	19.0
Rhode Island	4,910	10	1.6	8.3	597	12.2
South Carolina	51,014	10.9	2.5	8.4	6,048	11.9
South Dakota	7,036	7.4	2.1	5.4	574	8.2
Tennessee	67,804	11.4	2.1	9.3	8,844	13.0
Texas	183,885	8.4	1.9	6.4	17,663	9.6
Utah	16,534	6.5	2.7	3.8	981	5.9
Vermont	6,415	9.8	1.6	8.2	770	12.0
Virginia	71,555	13.9	2	11.9	11,206	15.7
Washington	50,955	7.7	2.3	5.4	4,138	8.1
West Virginia	22,500	11	1.9	9.1	2,900	12.9
Wisconsin	59,470	14.7	2	12.7	9,842	16.5
Wyoming	3,642	4.7	2.8	1.9	114	3.1

Canada						
Province name	Total Deaths >14 years old	Mean population weighted annual $PM_{2.5}$ ($\mu g\ m^{-3}$)			Attributable deaths[a]	Mean attributable fraction (%)[b]
		With all emission sources	Without fossil fuel	Estimated fossil fuel $PM_{2.5}$		
Alberta	21,535	8	2	6	1,958	9.1
British Columbia	33,403	8.7	1.9	6.8	3,237	9.7
Manitoba	9,868	7.9	2.7	5.2	778	7.9
New Brunswick	7,095	4.8	1.5	3.4	391	5.5
Newfoundland & Labrador	1,588	2.4	1.4	1	27	1.7
Northwest Territories	172	3.2	2.8	0.4	1	0.6
Nova Scotia	9,158	4.9	1.6	3.3	497	5.4
Nunavut	129	1.2	0.8	0.4	1	0.8
Ontario	90,996	15	1.6	13.4	15,728	17.3
Prince Edward Island	1,269	4.3	1.4	2.9	61	4.8
Quebec	66,494	13.9	1.6	12.3	10,645	16.0
Saskatchewan	8,515	7.5	2.4	5.2	678	8.0
Yukon Territory	193	1.1	0.9	0.3	1	0.5

[a] Annual number of deaths attributed to long term exposure to $PM_{2.5}$ generated by fossil fuel combustion.

[b] Mean proportion of deaths attributed to long term exposure to fossil-fuel related $PM_{2.5}$.

[c] Includes South Sudan

[d] Estimates derived after applying a 43.7% reduction to $PM_{2.5}$ from all sources for China

References

1. Aunan, K., Ma, Q., Lund, M. T., et al., Population-weighted exposure to $PM_{2.5}$ pollution in China: An integrated approach, Environ Int, 120, 111-120, doi:10.1016/j.envint.2018.07.042, 2018.

2. Bond, T. C., Bhardwaj, E., Dong, R., et al., Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850-2000, Global Biogeochem Cy, 21, doi:10.1029/2006gb002840, 2007.

3. Burnett, R., Chen, H., Szyszkowicz, M., et al., Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter, P Natl Acad Sci USA, 115, 9592-9597, doi:10.1073/pnas.1803222115, 2018.

4. Chang, L. Y., Xu, J. M., Tie, X. X., et al., Impact of the 2015 El Nino event on winter air quality in China, Sci Rep-Uk, 6, doi:10.1038/srep34275, 2016.

5. Cusworth, D. H., Mickley, L. J., Sulprizio, M. P., et al., Quantifying the influence of agricultural fires in northwest India on urban air pollution in Delhi, India, Environ Res Lett, 13, doi:10.1088/1748-9326/aab303, 2018.

6. Di, Q., Koutrakis, P., Schwartz, J., A hybrid prediction model for $PM_{2.5}$ mass and components using a chemical transport model and land use regression, Atmos Environ, 131, 390-399, doi:10.1016/j.atmosenv.2016.02.002, 2016.

7. Eyring, V., Kohler, H. W., van Aardenne, J., et al., Emissions from international shipping: 1. The last 50 years, J Geophys Res-Atmos, 110, doi:10.1029/2004jd005619, 2005.

8. Fairlie, T. D., Jacob, D. J., Park, R. J., The impact of transpacific transport of mineral dust in the United States, Atmos Environ, 41, 1251-1266, doi:10.1016/j.atmosenv.2006.09.048, 2007.

9. Ford, B., Heald, C. L., Exploring the uncertainty associated with satellite-based estimates of premature mortality due to exposure to fine particulate matter, Atmos Chem Phys, 16, 3499-3523, doi:10.5194/acp-16-3499-2016, 2016.

10. Fountoukis, C., Nenes, A., ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K^+-Ca^{2+}-Mg^{2+}-Nh_4^+-Na^+-SO_4^{2-}-NO_3^--Cl^--H_2O aerosols, Atmos Chem Phys, 7, 4639-4659, doi:10.5194/acp-7-4639-2007, 2007.

11. Giglio, L., Randerson, J. T., van der Werf, G. R., Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), J Geophys Res-Biogeo, 118, 317-328, doi:10.1002/jgrg.20042, 2013.

12. Guenther, A. B., Jiang, X., Heald, C. L., et al., The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geosci Model Dev, 5, 1471-1492, doi:10.5194/gmd-5-1471-2012, 2012.

13. Heald, C. L., Coe, H., Jimenez, J. L., et al., Exploring the vertical profile of atmospheric organic aerosol: comparing 17 aircraft field campaigns with a global model, Atmos Chem Phys, 11, 12673-12696, doi:10.5194/acp-11-12673-2011, 2011.

14. Heald, C. L., Jacob, D. J., Turquety, S., et al., Concentrations and sources of organic carbon aerosols in the free troposphere over North America, J Geophys Res-Atmos, 111, doi:10.1029/2006jd007705, 2006.

15. Hudman, R. C., Moore, N. E., Mebust, A. K., et al., Steps towards a mechanistic model of global soil nitric oxide emissions: implementation and space based-constraints, Atmos Chem Phys, 12, 7779-7795, doi:10.5194/acp-12-7779-2012, 2012.

16. Jaegle, L., Quinn, P. K., Bates, T. S., et al., Global distribution of sea salt aerosols: new constraints from in situ and remote sensing observations, Atmos Chem Phys, 11, 3137-3157, doi:10.5194/acp-11-3137-2011, 2011.

17. Keller, C. A., Long, M. S., Yantosca, R. M., et al., HEMCO v1.0: a versatile, ESMF-compliant component for calculating emissions in atmospheric models, Geosci Model Dev, 7, 1409-1417, doi:10.5194/gmd-7-1409-2014, 2014.

18. Lee, H. M., Park, R. J., Henze, D. K., et al., $PM_{2.5}$ source attribution for Seoul in May from 2009 to 2013 using GEOS-Chem and its adjoint model, Environ Pollut, 221, 377-384, doi:10.1016/j.envpol.2016.11.088, 2017.

19. Li, M., Zhang, Q., Kurokawa, J., et al., MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP, Atmos Chem Phys, 17, 935-963, doi:10.5194/acp-17-935-2017, 2017.

20. Li, X., Wu, J. R., Elser, M., et al., Contributions of residential coal combustion to the air quality in Beijing-Tianjin-Hebei (BTH), China: a case study, Atmos Chem Phys, 18, 10675-10691, doi:10.5194/acp-18-10675-2018, 2018.

21. Li, Y., Henze, D. K., Jack, D., et al., The influence of air quality model resolution on health impact assessment for fine particulate matter and its components, Air Qual Atmos Health, 9, 51-68, doi:10.1007/s11869-015-0321-z, 2016.

22. Liu, H. Y., Jacob, D. J., Bey, I., et al., Constraints from ^{210}Pb and ^{7}Be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields, J Geophys Res-Atmos, 106, 12109-12128, doi:10.1029/2000jd900839, 2001.

23. Mao, J. Q., Paulot, F., Jacob, D. J., et al., Ozone and organic nitrates over the eastern United States: Sensitivity to isoprene chemistry, J Geophys Res-Atmos, 118, 11256-11268, doi:10.1002/jgrd.50817, 2013.

24. Marais, E. A., Jacob, D. J., Turner, J. R., et al., Evidence of 1991-2013 decrease of biogenic secondary organic aerosol in response to SO_2 emission controls, Environ Res Lett, 12, doi:10.1088/1748-9326/aa69c8, 2017.

25. Marais, E. A., Wiedinmyer, C., Air Quality Impact of Diffuse and Inefficient Combustion Emissions in Africa (DICE-Africa), Environ Sci Technol, 50, 10739-10745, doi:10.1021/acs.est.6b02602, 2016.

26. Murray, L. T., Jacob, D. J., Logan, J. A., et al., Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data, J Geophys Res-Atmos, 117, doi:10.1029/2012jd017934, 2012.

27. Olivier, J. G. J., Berdowski, J. J. M., Global emission sources and sinks. The Climate System. A.A. Balkema Publishers/Swets & Zeitlinger Publishers, Lisse, The Netherlands, 2001, pp. 33-78.

28. Park, R. J., Jacob, D. J., Field, B. D., et al., Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, J Geophys Res-Atmos, 109, doi:10.1029/2003jd004473, 2004.

29. Pye, H. O. T., Liao, H., Wu, S., et al., Effect of changes in climate and emissions on future sulfate-nitrate-ammonium aerosol levels in the United States, J Geophys Res-Atmos, 114, doi:10.1029/2008jd010701, 2009.

30. Schultz, M. G., Backman, L., Balkanski, Y., et al., REanalysis of the TROpospheric chemical composition over the past 40 years: Final report. http://hdl.handle.net/11858/00-001M-0000-0011-FBDD-B, 2007.

31. Shen, L., Mickley, L. J., Effects of El Nino on Summertime Ozone Air Quality in the Eastern United States, Geophys Res Lett, 44, 12543-12550, doi:10.1002/2017gl076150, 2017.

32. Shindell, D., Faluvegi, G., Seltzer, K., et al., Quantified, localized health benefits of accelerated carbon dioxide emissions reductions, Nat Clim Change, 8, doi:10.1038/s41558-018-0108-y, 2018.

33. Silvern, R. F., Jacob, D. J., Kim, P. S., et al., Inconsistency of ammonium-sulfate aerosol ratios with thermodynamic models in the eastern US: a possible role of organic aerosol, Atmos Chem Phys, 17, 5107-5118, doi:10.5194/acp-17-5107-2017, 2017.

34. Stettler, M. E. J., Eastham, S., Barrett, S. R. H., Air quality and public health impacts of UK airports. Part I: Emissions, Atmos Environ, 45, 5415-5424, doi:10.1016/j.atmosenv.2011.07.012, 2011.

35. Travis, K. R., Jacob, D. J., Fisher, J. A., et al., Why do models overestimate surface ozone in the Southeast United States?, Atmos Chem Phys, 16, 13561-13577, doi:10.5194/acp-16-13561-2016, 2016.

36. van Donkelaar, A., Martin, R. V., Brauer, M., et al., Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors, Environ Sci Technol, 50, 3762-3772, doi:10.1021/acs.est.5b05833, 2016.

37. van Donkelaar, A., Martin, R. V., Leaitch, W. R., et al., Analysis of aircraft and satellite measurements from the Intercontinental Chemical Transport Experiment (INTEX-B) to quantify long-range transport of East Asian sulfur to Canada, Atmos Chem Phys, 8, 2999-3014, doi:10.5194/acp-8-2999-2008, 2008.

38. Venkataraman, C., Brauer, M., Tibrewal, K., et al., Source influence on emission pathways and ambient $PM_{2.5}$ pollution over India (2015-2050), Atmos Chem Phys, 18, 8017-8039, doi:10.5194/acp-18-8017-2018, 2018.

39. Vinken, G. C. M., Boersma, K. F., Jacob, D. J., et al., Accounting for non-linear chemistry of ship plumes in the GEOS-Chem global chemistry transport model, Atmos Chem Phys, 11, 11707-11722, doi:10.5194/acp-11-11707-2011, 2011.

40. Vodonos, A., Abu Awad, Y., Schwartz, J., The concentration-response between long-term $PM_{2.5}$ exposure and mortality; A meta-regression approach, Environ Res, 166, 677-689, doi:10.1016/j.envres.2018.06.021, 2018.

41. Wang, C., Corbett, J. J., Firestone, J., Improving spatial representation of global ship emissions inventories, Environ Sci Technol, 42, 193-199, doi:10.1021/es0700799, 2008.

42. Wang, Q. Q., Jacob, D. J., Spackman, J. R., et al., Global budget and radiative forcing of black carbon aerosol: Constraints from pole-to-pole (HIPPO) observations across the Pacific, J Geophys Res-Atmos, 119, 195-206, doi:10.1002/2013jd020824, 2014.

43. Wang, S. W., Zhang, Q., Streets, D. G., et al., Growth in NO_x emissions from power plants in China: bottom-up estimates and satellite observations, Atmos Chem Phys, 12, 4429-4447, doi:10.5194/acp-12-4429-2012, 2012.

44. WHO, Ambient and household air pollution and health. World Health Organization, http://www.who.int/airpollution/data/en/.

45. Wiedinmyer, C., Yokelson, R. J., Gullett, B. K., Global Emissions of Trace Gases, Particulate Matter, and Hazardous Air Pollutants from Open Burning of Domestic Waste, Environ Sci Technol, 48, 9523-9530, doi:10.1021/es502250z, 2014.

46. Xiao, Y. P., Logan, J. A., Jacob, D. J., et al., Global budget of ethane and regional constraints on US sources, J Geophys Res-Atmos, 113, doi:10.1029/2007jd009415, 2008.

47. Xing, J., Mathur, R., Pleim, J., et al., Can a coupled meteorology-chemistry model reproduce the historical trend in aerosol direct radiative effects over the Northern Hemisphere?, Atmos Chem Phys, 15, 9997-10018, doi:10.5194/acp-15-9997-2015, 2015.

48. Zender, C. S., Bian, H. S., Newman, D., Mineral Dust Entrainment and Deposition (DEAD) model: Description and 1990s dust climatology, J Geophys Res-Atmos, 108, doi:10.1029/2002jd002775, 2003.

49. Zhang, L., Jacob, D. J., Knipping, E. M., et al., Nitrogen deposition to the United States: distribution, sources, and processes, Atmos Chem Phys, 12, 4539-4554, doi:10.5194/acp-12-4539-2012, 2012.

50. Zhang, L. M., Gong, S. L., Padro, J., et al., A size-segregated particle dry deposition scheme for an atmospheric aerosol module, Atmos Environ, 35, 549-560, doi:10.1016/S1352-2310(00)00326-5, 2001.

ESC European Society of Cardiology

Cardiovascular Research (2020) **116**, 2247–2253
doi:10.1093/cvr/cvaa288

Regional and global contributions of air pollution to risk of death from COVID-19

Andrea Pozzer [1,2], Francesca Dominici[3], Andy Haines[4], Christian Witt [5], Thomas Münzel [6,7]*, and Jos Lelieveld [2,8]*

[1]International Center for Theoretical Physics, Trieste, Italy; [2]Max Planck Institute for Chemistry, Atmospheric Chemistry Department, Mainz, Germany; [3]Harvard T.H. Chan School of Public Health, Department of Biostatistics, Boston, MA, USA; [4]Centre for Climate Change and Planetary Health, London School of Hygiene and Tropical Medicine, London, UK; [5]Charité University Medicine, Pneumological Oncology and Transplantology, Berlin, Germany; [6]University Medical Center of the Johannes Gutenberg University, Mainz, Germany; [7]German Center for Cardiovascular Research, Mainz, Germany; and [8]The Cyprus Institute, Climate and Atmosphere Research Center, Nicosia, Cyprus

Received 24 June 2020; revised 3 October 2020; editorial decision 23 September 2020; accepted 30 September 2020; online publish-ahead-of-print 26 October 2020

Time for primary review: 6 days

Aims The risk of mortality from the coronavirus disease that emerged in 2019 (COVID-19) is increased by comorbidity from cardiovascular and pulmonary diseases. Air pollution also causes excess mortality from these conditions. Analysis of the first severe acute respiratory syndrome coronavirus (SARS-CoV-1) outcomes in 2003, and preliminary investigations of those for SARS-CoV-2 since 2019, provide evidence that the incidence and severity are related to ambient air pollution. We estimated the fraction of COVID-19 mortality that is attributable to the long-term exposure to ambient fine particulate air pollution.

Methods and results We characterized global exposure to fine particulates based on satellite data, and calculated the anthropogenic fraction with an atmospheric chemistry model. The degree to which air pollution influences COVID-19 mortality was derived from epidemiological data in the USA and China. We estimate that particulate air pollution contributed ~15% (95% confidence interval 7–33%) to COVID-19 mortality worldwide, 27% (13 – 46%) in East Asia, 19% (8–41%) in Europe, and 17% (6–39%) in North America. Globally, ~50–60% of the attributable, anthropogenic fraction is related to fossil fuel use, up to 70–80% in Europe, West Asia, and North America.

Conclusion Our results suggest that air pollution is an important cofactor increasing the risk of mortality from COVID-19. This provides extra motivation for combining ambitious policies to reduce air pollution with measures to control the transmission of COVID-19.

Keywords COVID-19 • Air pollution • Fine particulate matter • comorbidity • mortality

1. Introduction

Poor air quality, especially from fine particulate matter with a diameter <2.5 μm ($PM_{2.5}$), is one of the leading risk factors, and responsible for many excess deaths.[1,2] The global loss of life expectancy from long-term exposure to ambient air pollution exceeds that of infectious diseases, and is comparable with that of tobacco smoking.[1–3] The mortality from COVID-19 depends on comorbidities, including conditions that increase cardiovascular risks such as arterial hypertension, diabetes mellitus, obesity, and established coronary artery disease, as well as respiratory conditions such as asthma and chronic obstructive pulmonary disease (COPD), being similar to those that are influenced by air pollution.[3–6] The risk of death is strongly related to age, being particularly high in those aged >70. It is also higher amongst males, economically disadvantaged populations, and in some ethnic groups. In assessing the relationships between exposures to risk factors and outcomes, potential confounders therefore need to be accounted for in the design of studies and in data analysis. These include the age distribution of the population, availability of hospital beds (and intensive care capacity), and the proportion of the population living in poverty.

* Corresponding authors. Jos Lelieveld: Tel: +49 6131 305 4000, Fax: +49 6131 305 4019, Email: jos.lelieveld@mpic.de or Thomas Münzel: Tel: +49 6131 17 7250, Fax: +49 6131 17 6615, Email: tmuenzel@uni-mainz.de

A recent study, using an ecological design, assessed how environmental influences modify the severity of COVID-19 outcomes in the USA.[7] Potential confounders were identified, and statistical models were used to relate long-term exposure to ambient $PM_{2.5}$ to COVID-19 deaths. The computed mortality rate ratios (MRRs) express the relative increase in COVID-19 deaths for each microgram per cubic meter increment of $PM_{2.5}$ in ambient air. The $PM_{2.5}$ data were derived from satellite and ground-based measurements combined with atmospheric modelling,[8] and the confounders were determined from county-level censuses, homeland infrastructure, and meteorological data. Here we test the assumption that the derived MRRs are representative for the populations of other countries (China) and consider the global impact. In the present study, we apply the MRRs to estimate the excess mortality, i.e. the fraction of COVID-19 deaths that could be avoided if the population were exposed to lower counterfactual air pollution levels without fossil fuel-related and other anthropogenic emissions. We emphasize that our results are provisional, based on epidemiological data collected up to the third week of June 2020, and a comprehensive evaluation will need to follow after the COVID-19 pandemic.

1.1 SARS and air pollution

In the early 2000s, the first severe acute respiratory syndrome coronavirus (SARS-CoV-1) appeared in China (Guangdong Province). The virus was zoonotic, as it originally developed in bats.[9] The World Health Organization (WHO) reported that it resulted in a SARS epidemic with >8000 cases in 26 countries, mostly in south-east Asia and in Canada.[8] The disease emerged in November 2002 and was contained in July 2003. SARS-CoV-1 and SARS-CoV-2 have many similarities, as their RNA genomes are closely related and the viruses enter the host cells by binding to the same entry receptor angiotensin-converting enzyme 2 (ACE2).[10–12] About 2–14 days after infection, the systemic symptoms of both diseases are alike, and a similar fraction of patients develops severe symptoms with a mortality rate that increases strongly with advanced age.[13–16] In China alone, >5000 cases of SARS-CoV-1 were reported, leading to nearly 350 fatalities. Since the exposure to ambient air pollution is associated with respiratory and cardiovascular diseases, it was hypothesized that health outcomes of SARS were aggravated by poor air quality. A study in 2003 corroborated that in parts of China with moderate levels of air pollution, the risk of dying from the disease was >80% higher compared with areas with relatively clean air, while in heavily polluted regions the risk was twice as high.[17]

1.2 COVID-19 and air pollution

In 2019, the related second virus strain appeared (SARS-CoV-2) in China (Hubei Province), which also developed in bats,[4] causing COVID-19, which grew from an epidemic into a pandemic in the early part of 2020. A Chinese analysis indicated that the risk of symptomatic infection typically increases by ∼4% for each year of age between 30 and 60, and that the lethality is highest for individuals >60 years.[15]COVID-19 is associated with a combination of respiratory and cardiovascular complications, which may comprise myocardial infarction, heart failure, venous thrombo-embolisms, and increases in biomarkers,[18] which are also found in connection with high levels of air pollutants.[5] In a recent analysis of 5700 patients hospitalized with COVID-19 in the New York City area, the most common comorbidities were hypertension (57%), obesity (42%), and diabetes (34%),[19] representing cardiovascular risk factors that are also observed in relation to elevated $PM_{2.5}$ concentrations,[5,20] suggesting additive or synergistic effects on the cardiovascular system. In addition, advanced age is a strong risk factor for cardiovascular disease, and the effects on immune function may be equally important for COVID-19 susceptibility. The age dependency coincides with that of excess mortality from $PM_{2.5}$.[3,15] The COVID-19 mortality rate has been estimated to be ∼4% in symptomatic cases, in part because pre-existing conditions such as cardiovascular and respiratory disorders increase the risk.[21]

Considering the cardiovascular and respiratory health impacts of air pollution, the relationship to COVID-19 mortality is not unexpected. Preliminary studies addressed the influence of air pollution on COVID-19 in different regions. In China, the incidence of COVID-19 was found to be significantly enhanced by $PM_{2.5}$,[22] while a correlation between ambient $PM_{2.5}$ and the mortality rate was also established.[23] In Italy, it was found that the high pollution concentrations that are typical for the Po valley, especially in the Lombardy region of which Milan is the capital, were associated with a high mortality rate.[24] As mentioned above, in the USA the severity of COVID-19 outcomes was linked to $PM_{2.5}$ exposure, making use of Medicare data for >60 million people and nationwide air quality measurements.[7] Data were collected for 98% of the population in 3087 of the total number of 3142 counties, of which ∼42% had reported COVID-19 deaths up to the third week of April 2020. The death counts relied on data from the Coronavirus Resource Center of the Johns Hopkins University.[25] The study accounted for 20 potential confounding factors including population size, age distribution, population density, time period since the beginning of the outbreak, time elapsed since the home confinements, hospital beds, number of individuals tested, meteorological conditions, and socioeconomic and risk factors such as obesity and smoking.[7] The results showed significant overlap between the causes of death in COVID-19 patients and those that lead to mortality from $PM_{2.5}$. The MRR, i.e. the percentage increase of COVID-19 mortality risk per $\mu g/m^3$ increase of exposure to $PM_{2.5}$, was found to be 8%, with a 95% confidence interval of 2–15%.[7] The calculations are continually updated based on the most recent data (up to 18 June at the time of writing), showing no significant changes in the MRR in the preceding 4 months.

2. Methods

2.1 Global model and data

We applied a global atmospheric chemistry general circulation model (EMAC) which comprehensively simulates atmospheric chemical and meteorological processes and interactions with the oceans and the biosphere, in the same set-up as in recent studies on climate change, air pollution, and public health.[3,26] In addition to the standard simulation, we performed two sensitivity calculations: (i) with fossil fuel-related emissions removed and (ii) with all anthropogenic emissions removed. The model results were used to estimate the ratio of fine particulates in simulation (i) and (ii) and the standard simulation. The annual atmospheric near-surface $PM_{2.5}$ concentrations were taken from model-integrated satellite data, for the year 2019.[8,27] The horizontal resolution is 0.01 by 0.01 degrees, corresponding to a grid size of ∼1 km × 1 km. The near-surface concentrations of $PM_{2.5}$ for fossil fuel-related and all anthropogenic emissions are estimated by scaling this data set to the ratios (i) and (ii) obtained with the EMAC model simulations.

2.2 Relative risk

To estimate the relative risk (RR or hazard ratio) of excess COVID-19 mortality from the long-term exposure to air pollution, we used the exposure–response function of the WHO,[28]

$$RR = \left(\frac{X+1}{X_0+1}\right)^{\beta},$$

RR is a function of the concentration of air pollutants, which specifies annual average exposure dependent on location (grid cell) derived from the data mentioned above. X is the pollutant ($PM_{2.5}$) and X_0 is the pollutant threshold concentration below which exposure does not have implications for public health. Both β and X_0 are estimated by fitting to data from the literature with a least square method (*Figure 1*). We adopted the threshold $PM_{2.5}$ concentration (X_0) from Burnett *et al.*[2] (i.e. < 2.4 μg/m^3 $PM_{2.5}$), forcing the curve fitting into this range. We tested different exposure–response functions, e.g. of Burnett *et al.*,[2] and values for X_0, and find that the results are not sensitive to these assumptions.

Because the COVID-19 mortality rate ratio due to air pollution, based on data in the USA alone,[7] may not represent countries with very high fine particle concentrations (associated with a lack of observations in such regions), we investigated the effect of including data from the enhanced mortality rate derived for the Chinese SARS epidemic in 2003.[17] We make the assumption that SARS and COVID-19 mortality are similarly affected by long-term exposure to air pollution. Since the analysis for SARS was based on the Chinese Air Pollution Index (API), we converted the API to $PM_{2.5}$ concentrations following empirical relationships from the literature.[29,30] The large uncertainty range in the fitting function to a large degree derives from those in these relationships (black squares and ranges in *Figure 1*). In spite of uncertainties, the curves for the USA only and those that include the Chinese results are almost identical, providing confidence in the function derived for conditions in the USA only.

2.3 Attributable fraction

We calculated RR globally using $PM_{2.5}$ distributions calculated under the standard scenario. The attributable fraction (AF) of COVID-19 mortality to air pollution is calculated from the RR by AF = 1 – 1/RR. From the globally distributed, gridded AFs, we aggregated into regional and country-level AFs, weighted according to the population density, in order to account for the varying population distributions within regions and countries. The population data for the year 2020 were obtained from the NASA Socioeconomic Data and Applications Center (SEDAC), hosted by the Columbia University Center for International Earth Science Information Network (CIESIN).[31] Our definition of AF does not imply a direct cause–effect relationship between air pollution and COVID-19 mortality (although it is possible). Instead it refers to relationships between the two, direct and indirect, i.e. by aggravating comorbidities that could lead to fatal health outcomes of the virus infection.

3. Results

3.1 Attribution of COVID-19 mortality

To estimate the AF from exposure to ambient $PM_{2.5}$ to COVID-19 mortality, we used the epidemiological data from the USA (red curve in *Figure 1*). The chronic exposure to $PM_{2.5}$ in the years prior to the COVID-19 outbreak was estimated on the basis of satellite observations over the year 2019. The anthropogenic and fossil fuel-related fractions were calculated with the global EMAC model. Here we focus on anthropogenic and fossil fuel-related $PM_{2.5}$ to determine the impact of potentially avoidable air pollution on COVID-19 mortality. *Figure 2* and *Table 1* present the average fractions of COVID-19 mortality attributed to the exposure to $PM_{2.5}$ pollution, both globally and regionally. *Table S1* (available as Supplementary material online) lists the results for all countries. To account for the different population distributions within countries, e.g. between rural and urban areas, the averages have been weighted accordingly.

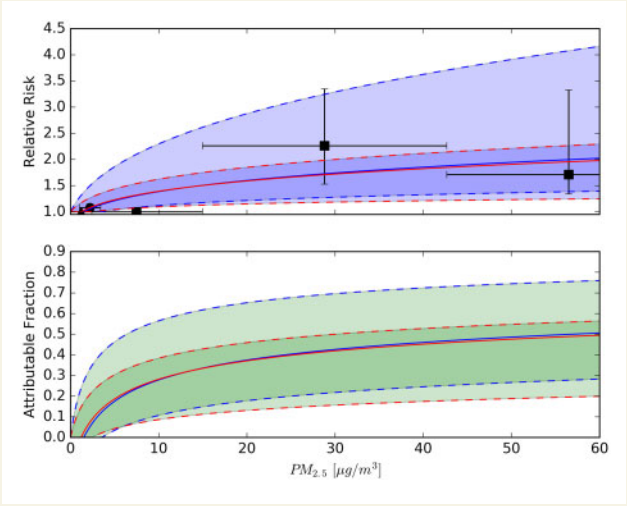


Figure 1 Exposure-response dependencies, based on a log-normal relationship[28]. The relative risk (or hazard ratio), from which the attributable fraction has been derived, is based on mortality rate ratios attributed to air pollution in the COVID-19 pandemic[7] and the SARS epidemic[17], indicated by the black bullet and squares, respectively. The triangle represents the threshold concentration below which $PM_{2.5}$does not have health implications[2]. The red curves depict the function fitted to the data from COVID-19 in the USA only[7], plus the threshold[2] (triangle and bullet). The blue curves depict the function fitted to all data[2,7,17]. The colored ranges show the 95% confidence intervals, which are wider after including the SARS-related results (blue), mostly due to uncertainty from converting Chinese API's into $PM_{2.5}$concentrations (black squares).

In regions with strict air quality standards and relatively low levels of air pollution, such as Australia, the attributable fraction by human-made air pollution to COVID-19 mortality is found to be a few percent only. Relatively high fractions occur in parts of east Asia (~35%), central Europe (~25%), and eastern USA (~25%). The country-level contribution to COVID-19 that we find for China, i.e. 27% (95% confidence interval 13 – 47%), agrees well with that found for the SARS epidemic in 2003.[17] The largest country-average fractions are found in the Czech Republic, Poland, China, North Korea, Slovakia, Austria, Belarus, and Germany, all above 25% (Supplementary material, *Table S1*). Globally, anthropogenic air pollution contributes ~15% (7 – 33%) to COVID-19 mortality, which could have been largely prevented, for example by adopting the air quality regulations applied in Australia (annual $PM_{2.5}$ limit of 8 μg/m^3). The global mean contribution of fossil fuel use to the anthropogenic fraction is ~56%, being highest in North America (83%), West Asia (75%), and Europe (68%) (*Table 1*).

4. Discussion

4.1 Pathophysiological aspects

Both the air pollutant $PM_{2.5}$ and the SARS-CoV-2 virus enter the lungs via the bronchial system (portal organ), with potential systemic health impacts through the blood circulation. Both $PM_{2.5}$ and SARS-CoV-2 cause vascular endothelial dysfunction, oxidative stress, inflammatory responses, thrombosis, and an increase in immune cells.[32–36] The SARS-CoV-2 infection facilitates the induction of endothelial inflammation in

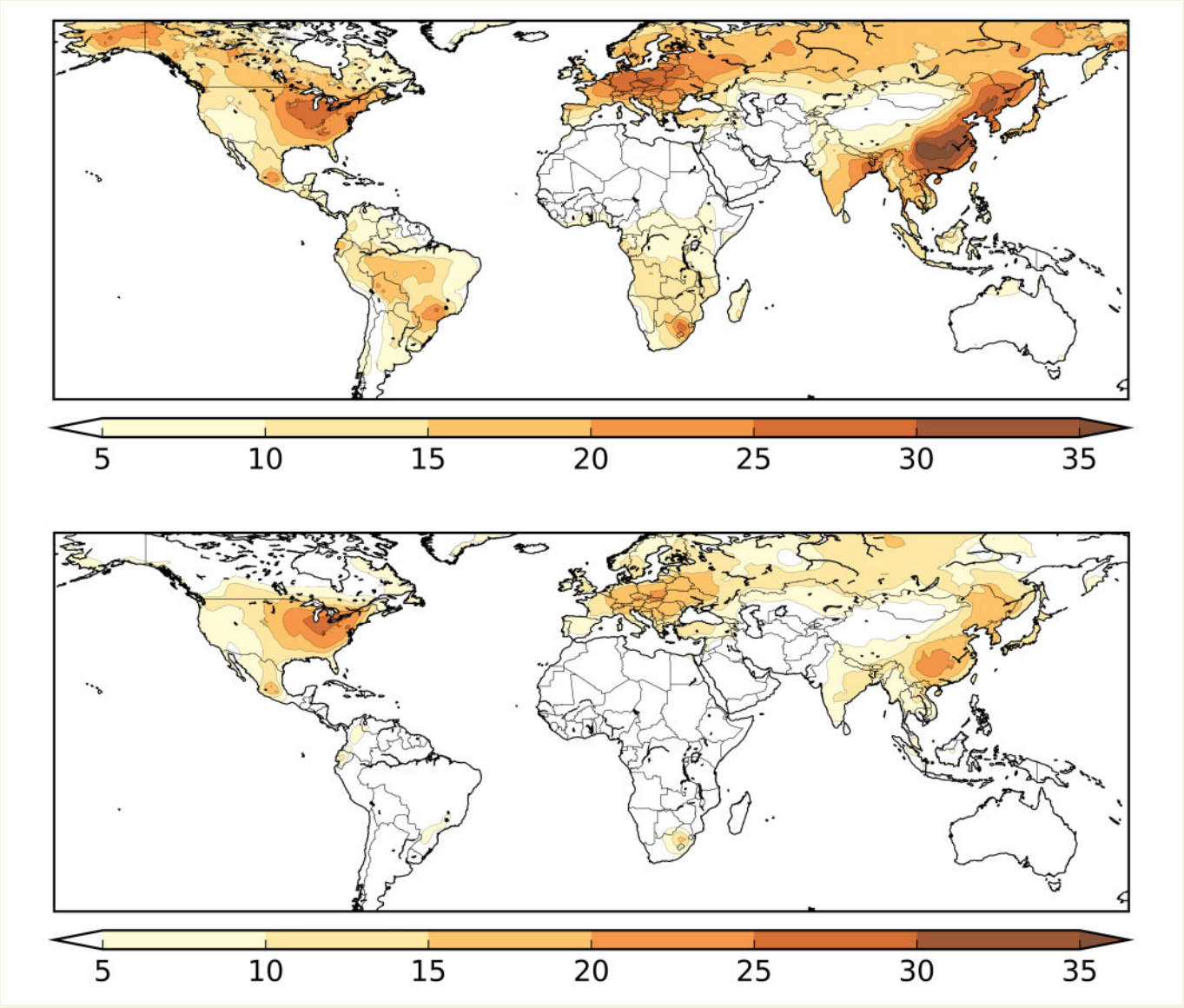


Figure 2 Estimated percentages of COVID-19 mortality attributed to air pollution from all anthropogenic sources (top), and from fossil fuel use only (bottom). The regions with high attributable fractions coincide with high levels of air pollution. The mapped results account for population density, thus reflecting population weighted exposure to $PM_{2.5}$.

several organs as a direct consequence of viral cytotoxic effects and the host inflammatory response, which can aggravate pre-existing chronic respiratory and vascular (coronary) dysfunction, and cause lung injury by alveolar damage, as well as stroke and myocardial infarction by inducing plaque rupture.[37] Potential common pathophysiological mechanisms of increased risk thus relate to endothelial injury[33,38] and pathways that regulate immune function.[39,40] Further, there are strong indications of increased susceptibility to viral infections from exposure to air pollution.[41–46]

Lung injuries, including the life-threatening acute respiratory distress syndrome and respiratory failure, as well as acute coronary syndrome, arrhythmia, myocarditis, and heart failure, were shown to be clinically dominant, leading to critical complications of COVID-19.[47,48] Recent studies in China, the USA, as well as Europe indicate that patients with cardiovascular risk factors or established cardiovascular disease and other comorbid conditions are predisposed to myocardial injury during the course of COVID-19.[19,46,49–52] From the available information, it thus follows that air pollution-induced inflammation leads to greater vulnerability and less resiliency, and the pre-conditions increase the host vulnerability. Air pollution causes adverse events through myocardial infarction and stroke, and it is an additional factor capable of increasing blood pressure, while there is emerging evidence for a link with type 2 diabetes and a possible contribution to obesity and enhanced insulin resistance.[36] Bronchopulmonary and cardiovascular pre-conditions, including hypertension, diabetes, coronary artery disease, cardiomyopathy, asthma, COPD, and acute lower respiratory illness, all negatively influenced by air pollution, lead to a substantially higher mortality risk in COVID-19. Furthermore, it seems likely that fine particulates prolong the atmospheric lifetime of infectious viruses, thus favouring transmission.[53] It is possible that future research will reveal additional pathways that mediate the relationship between air pollution and the risk of death from COVID-19.

Table 1 Regional percentages of COVID-19 mortality attributed to fossil fuel-related and all anthropogenic sources of air pollution

Region	Population (million)	COVID-19 mortality fraction attributed to air pollution (%)	
		Fossil fuel-related emissions	All anthropogenic emissions
Europe	628	13 (6–33)	19 (8–41)
Africa	1345	2 (1–19)	7 (3–25)
West Asia	627	6 (3–25)	8 (4–27)
South Asia	2565	7 (3–22)	15 (8–31)
East Asia	1685	15 (8–32)	27 (13–46)
North America	525	14 (6–36)	17 (6–39)
South America	547	3 (1–23)	9 (4–30)
Oceania	28	1 (0–20)	3 (1–23)
World	7950	8 (4–25)	15 (7–33)

The 95% confidence levels are given in parentheses.

4.2 Limitations

Our results indicate that the long-term exposure to high levels of fine particulate matter is a significant cofactor that influences the severity of COVID-19 outcomes. Since $PM_{2.5}$ in China and the USA, from which epidemiological data have been used, is dominated by anthropogenic sources that are potentially preventable, we focus our analysis on this fraction of $PM_{2.5}$. The good agreement of our results for the USA and China is in line with recent studies, showing that the association between air pollution and excess mortality is valid for many different countries.[2,55] Nevertheless, the calculations of RRs (hazard ratios) and the AF to mortality rely on the use of data from an ecological study design that has limitations, even though 19 county-level variables and one state-level variable, some of which are more important than air pollution, were considered as potential confounders in the analysis—and the $PM_{2.5}$ exposure data have been extensively cross-validated.[7] However, we acknowledge that residual confounding cannot be excluded. While cross-sectional ecological studies do not allow conclusions about cause–effect relationships, the biological mechanisms of air pollution-related disorders, acting as comorbidities in COVID-19, are well documented.[56,57] Recent studies in England and The Netherlands corroborate the positive relationships between air pollution and the number of COVID-19 cases, hospital admissions, and mortality.[58–60] The reported MRRs for $PM_{2.5}$ range from 1–7% to 13–21% (we applied 2–15%), which confirms the significant role of air pollution but emphasizes the large uncertainty ranges. Furthermore, our approach is likely to realistically approximate the contribution of fossil fuels and other anthropogenic sources to the total excess deaths through long-term ambient $PM_{2.5}$ air pollution exposure.

We reiterate that the data used for China are associated with substantial uncertainty, and underly the assumption that comorbidity and mortality from air pollution in COVID-19 are the same as in SARS. Nonetheless, using these data does not change the results, providing confidence in the robustness of our findings. We emphasize that the data relevant to the present study are from upper-middle and high-income countries, and the representativeness of our results for low-income countries may be limited, and uncertainties are likely to exceed the 95% confidence intervals. It is expected that in countries with high levels of aeolian dust, e.g. in Africa and West Asia, $PM_{2.5}$ pollution is also a cofactor but with less contribution from human activities. Household air pollution is also likely to be important, being of particular relevance in low-income countries.[61] It will be critical to collect epidemiological evidence from many regions with different socio-economic and environmental conditions, to support analyses of the COVID-19 pandemic and investigate the role of environmental factors. The uncertainty ranges that accompany our results are considerable but, taking into account the biological plausibility of the relationship and the strong evidence of the impact of air pollution on conditions that are known to increase COVID-19 mortality, they can nevertheless inform policy decisions.

4.3 Short- and long-term health impacts

A new, though preliminary, finding of the present study is that a significant fraction of worldwide COVID-19 mortality is attributable to anthropogenic air pollution, of which ~50 – 60% is related to fossil fuel use (~70 – 80% in Europe, West Asia, and North America). This represents potentially avoidable, excess mortality. The links between economic activity, traffic, energy use, and public health have been illustrated by the strong reduction of air pollution in many locations during the lockdown measures.[62,63] There is ample evidence for a relationship between short-term exposure to $PM_{2.5}$ and adverse health effects, including excess mortality from cardiovascular and respiratory diseases.[55] While it is in principle possible to disentangle the acute from the chronic outcomes from short- and long-term exposure to air pollution,[64] at this stage it is difficult to make that distinction for $PM_{2.5}$-induced comorbidity and mortality from COVID-19. Generally, short-term associations between air pollution and mortality are substantially less than those from long-term exposure, due to the more persistent, cumulative effects from the latter.[65] By relating air pollution anomalies to short-term health outcomes during the COVID-19-induced societal lockdown, it was found that in China alone >4600 excess deaths may have been avoided.[62] This can be viewed as a health co-benefit from the containment measures, which may reduce air pollution-induced COVID-19 mortality. Such benefits could also be achieved after the COVID-19 lockdown. Both perspectives of air pollution during the pandemic underscore the important role of fossil fuel-related and other anthropogenic emissions.

4.3 Future directions

Our results suggest the potential for substantial benefits from reducing air pollution exposure even at relatively low $PM_{2.5}$ levels. Refinement of the exposure–response relationship and reducing uncertainties will require additional data analyses, including from large cohort studies as the COVID-19 pandemic evolves, but may appear too late to guide

decision-making. A lesson from our environmental perspective of the COVID-19 pandemic is that the quest for effective policies to reduce anthropogenic emissions, which cause both air pollution and climate change, needs to be accelerated. The pandemic ends with the vaccination of the population or with herd immunity through extensive infection of the population. However, there are no vaccines against poor air quality and climate change. The remedy is to mitigate emissions. The transition to a green economy with clean, renewable energy sources will further both environmental and public health locally through improved air quality and globally by limiting climate change.

Supplementary material

Supplementary material is available at *Cardiovascular Reseach* online.

Funding

We thank the Mainz Heart Foundation for continuous support. T.M. is the principal investigator of the DZHK (German Center for Cardiovascular Research), Partner Site Rhine-Main, Mainz, Germany.

Conflict of interest: none declared.

Data availability

The data underlying this article will be shared upon reasonable request to the corresponding author.

References

1. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, Balakrishnan K, Brunekreef B, Dandona L, Dandona R, Feigin V, Freedman G, Hubbell B, Jobling A, Kan H, Knibbs L, Liu Y, Martin R, Morawska L, Pope CA 3rd, Shin H, Straif K, Shaddick G, Thomas M, van Dingenen R, van Donkelaar A, Vos T, Murray CJL, Forouzanfar MH. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet* 2017;**389**:1907–1918.
2. Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA, Apte JS, Brauer M, Cohen A, Weichenthal S, Coggins J, Di Q, Brunekreef B, Frostad J, Lim SS, Kan H, Walker KD, Thurston GD, Hayes RB, Lim CC, Turner MC, Jerrett M, Krewski D, Gapstur SM, Diver WR, Ostro B, Goldberg D, Crouse DL, Martin RV, Peters P, Pinault L, Tjepkema M, van Donkelaar A, Villeneuve PJ, Miller AB, Yin P, Zhou M, Wang L, Janssen NAH, Marra M, Atkinson RW, Tsang H, Quoc Thach T, Cannon JB, Allen RT, Hart JE, Laden F, Cesaroni G, Forastiere F, Weinmayr G, Jaensch A, Nagel G, Concin H, Spadaro JV. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc Natl Acad Sci USA* 2018;**115**:9592–9597.
3. Lelieveld J, Pozzer A, Pöschl U, Fnais M, Haines A, Münzel, T. Comparison of mortality from ambient air pollution with other risk factors: a worldwide perspective. *Cardiov Res* 2020;doi: 10.1093/cvr/cvaa025.
4. World Health Organization. *Report of the WHO–China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. Geneva: WHO; 2020.
5. Miller MR. Oxidative stress and the cardiovascular effects of air pollution. *Free Radic Biol Med* 2020;**151**:69–87.
6. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, Curtis HJ, Mehrkar A, Evans D, Inglesby, P, Cockburn J, McDonald HI, MacKenna B, Tomlinson L, Douglas IJ, Rentsch RT, Mathur R, Wong AYS, Grieve R, Harrison D, Forbes H, Schultze, A, Croker R, Parry J, Hester F, Harper S, Perera R, Evans SJW, Smeeth L, Goldacre B. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;**584**:430–436.
7. Wu X, Nethery RC, Sabath MB, Braun D, Dominici F. Exposure to air pollution and COVID-19 mortality in the United States. *MedRxiv* 2020; doi: 10.1101/2020.04.05.20054502.
8. van Donkelaar A, Martin RV, Li C., Burnett RT. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environ Sci Technol* 2019;**53**:2595-2611.
9. World Health Organization. SARS (Severe Acute Respiratory Syndrome) Disease Information. Geneva: WHO; 2020. https://www.who.int/ith/diseases/sars/en/
10. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, Li F. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci USA* 2020;**117**:11727–11734.
11. Li Y, Liu B, Cui J, Wang Z, Shen Y, Xu Y, Yao K, Guan Y. Similarities and evolutionary relationships of COVID-19 and related viruses. *arXiv* 2020;2020030316 (doi: 10.20944/preprints202003.0316.v1).
12. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020;**579**:270–273.
13. Chan-Yeun M, Xu R-H. SARS: epidemiology. *Respiration* 2003;**8**:S9–S14.
14. Yang M, Li CK, Li K, Hon KLE, Ng MHL, Chan PKS, Fok TF. Hematological findings in SARS patients and possible mechanisms (Review). *Int J Mol Med* 2004;**14**:311–315
15. Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, Cowling BJ, Lipsitch M, Leung GM, Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med* 2020;**26**:506–510.
16. Tian S, Hu N, Lou J, Chen K, Kang X, Xiang Z, Chen H, Wang D, Liu N, Liu D, Chen G, Zhang Y, Li D, Li J, Lian H, Niu S, Zhang L, Zhang J. Characteristics of COVID-19 infection in Beijing. *J Infect* 2020;**80**:401–406.
17. Cui Y, Zhang Z-F, Froines J, Zhao J, Wang H, Yu S-Z, Detels R. Air pollution and case fatality of SARS in the People's Republic of China: an ecologic study. *Environ Health* 2003;**2**:15.
18. Driggin E, Madhavan MV, Bikdeli B, Chuich T, Laracy J, Biondi-Zoccai G, Brown TS, Nigoghossian CD, Zidar DA, Haythe J, Brodie D, Beckman JA, Kirtane AJ, Stone GW, Krumholz HM, Parikh SA. Cardiovascular considerations for patients, health care workers, and health systems during the COVID-19 pandemic. *J Am Coll Cardiol* 2020;**75**: 2352–2371.
19. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, and the Northwell C-RC, Barnaby DP, Becker LB, Chelico JD, Cohen SL, Cookingham J, Coppa K, Diefenbach MA, Dominello AJ, Duer-Hefele J, Falzon L, Gitlin J, Hajizadeh N, Harvin TG, Hirschwerk DA, Kim EJ, Kozel ZM, Marrast LM, Mogavero JN, Osorio GA, Qiu M, Zanos TP. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA* 2020;**323**:2052–2059.
20. Münzel T, Sorensen M, Gori T, Schmidt FP, Rao X, Brook J, Chen LC, Brook RD, Rajagopalan S. Environmental stressors and cardio-metabolic disease: part I—epidemiologic evidence supporting a role for noise and air pollution and effects of mitigation strategies. *Eur Heart J* 2017;**38**:550–556.
21. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet* 2020;**395**:470–473.
22. Wang B, Liu J, Fu S, Xu X, Li L, Ma Y, Zhou J, Yao J, Liu X, Zhang X, He X, Yan H, Shi Y, Ren X, Niu J, Luo B, Zhang K. An effect assessment of airborne particulate matter pollution on COVID-19: a multi-city study in China. *MedRxiv* 2020;https://doi.org/10.1101/2020.04.09.20060137.
23. Yao Y, Pan J., Wang W, Liu X, Kan H., Meng X, Wang W. Spatial correlation of particulate matter pollution and death rate of COVID-19. *MedRxiv* 2020;https://doi.org/10.1101/2020.04.07.20052142.
24. Conticini E, Frediani B, Caro D. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? *Environ Poll* 2020;**261**:114465.
25. Coronavirus Resource Center. Johns Hopkins University and Medicine, Baltimore, MD, USA, 2020; https://coronavirus.jhu.edu.
26. Lelieveld J, Klingmüller K, Pozzer A, Burnett RT, Haines A, Ramanathan V. Effects of fossil fuel and total anthropogenic emission removal on public health and climate. *Proc Natl Acad Sci USA* 2019;**116**:7192–7197.
27. van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, Lyapustin A, Sayer AM, Winker D.M. Global estimates of fine particulate matter using a combined geophysical–statistical method with information from satellites, models, and monitors. *Environ Sci Technol* 2016;**50**:3762–3772.
28. Ostro B. *Outdoor Air Pollution: Assessing the Environmental Burden of Disease at National and Local Levels. World Health Organization Protection of the Human Environment*. Geneva: WHO; 2004.
29. Guo JP, Zhang X-Y, Che H-Z, Gong S-L, An X, Cao C-X, Guang J, Zhang H, Wang Y-Q, Zhang XC, Xue M, Li X-W. Correlation between PM concentrations and aerosol optical depth in eastern China. *Atmos Environ* 2009;**43**:5876–5886.
30. Zheng S, Cao CX, Singh RP. Comparison of ground based indices (API and AQI) with satellite based aerosol products. *Sci Total Environ* 2014;**488**:398–412.
31. Center for International Earth Science Information Network, CIESIN, Columbia University. *Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11. Palisades, NY*. NASA Socioeconomic Data and Applications Center (SEDAC) 2020; https://doi.org/10.7927/H4JW8BX5.
32. Pope CA III, Bhatnagar A, McCracken JP, Abplanalp W, Conklin DJ, O'Tool T. Exposure to fine particulate air pollution is associated with endothelial injury and systemic inflammation. *Circ Res* 2016;**119**:1204–1214.
33. Münzel T, Gori, T, Al-Kindi S, Deanfield J, Lelieveld J, Daiber A, Rajagopalan S. Effects of gaseous and solid constituents of air pollution on endothelial function. *Eur Heart J* 2018;**39**:3543–3550.

34. Varga Z, Flammer AJ, Steiger P, Haberecker M, Andermatt R, Zinkernagel AS, Mehra MR, Schuepbach RA, Ruschitzka F, Moch H. Endothelial cell infection and endotheliitis in COVID-19. *Lancet* 2020;**395**:1417–1418.
35. Tsai D-H, Riediker M, Berchet A, Paccaud F, Waeber G, Vollenweider P, Bochud M. Effects of short- and long-term exposures to particulate matter on inflammatory marker levels in the general population. *Environ Sci Poll Res* 2019;**26**:19697–19704.
36. Thurston GD, Kipen H, Annesi-Maesano I, Balmes J, Brook RD, Cromar K, De Matteis S, Forastiere F, Forsberg B, Frampton MW, Grigg J, Heederik D, Kelly FJ, Kuenzli N, Laumbach R, Peters A, Rajagopalan ST, Rich D, Ritz B, Samet JM, Sandstrom T, Sigsgaard T, Sunyer J, Brunekreef B. A joint ERS/ATS policy statement: what constitutes an adverse health effect of air pollution? An analytical framework. *Eur Respir J* 2017;**49**:1600419.
37. Wenzel P, Kopp S, Göbel S, Jansen T, Geyer M, Hahn F, Kreitner K-F, Escher F, Schultheiss H-P, Münzel T. Evidence of SARS-CoV-2 mRNA in endomyocardial biopsies of patients with clinically suspected myocarditis tested negative for COVID-19 in nasopharyngeal swab. *Cardiovasc Res* 2020;**116**:1661–1663.
38. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, Vanstapel A, Werlein C, Stark H, Tzankov A, Li WW, Li VW, Mentzer SJ, Jonigk D. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. *N Engl J Med* 2020; **383**:120–128.
39. Kelly FJ, Fussell JC. Linking ambient particulate matter pollution effects with oxidative biology and immune responses. *Ann N Y Acad Sci* 2015;**1340**:84–94.
40. O'Driscoll CA, Owens LA, Gallo ME, Hoffmann EJ, Afrazi A, Han M, Fechner JH, Schauer JJ, Bradfield CA, Mezrich JD. Differential effects of diesel exhaust particles on T cell differentiation and autoimmune disease. *Part Fibre Toxicol* 2018;**15**: 35.
41. Becker S, Soukup JM Exposure to urban air particulates alters the macrophage-mediated inflammatory response to respiratory viral infection. *J Toxicol Environ Health A* 1999;**57**:445–457.
42. Harrod KS, Jaramillo RJ, Rosenberger CL, Wang S-Z, Berger JA, McDonald JD, Reed MD. Increased susceptibility to RSV infection by exposure to inhaled diesel engine emissions. *Am J Respir Cell Mol Biol* 2003;**28**:451–463.
43. Kaan PM, Hegele RG. Interaction between respiratory syncytial virus and particulate matter in guinea pig alveolar macrophages. *Am J Respir Cell Mol Biol* 2003;**28**:697–704.
44. Lambert AL, Mangum JB, DeLorme MP, Everitt JI. Ultrafine carbon black particles enhance respiratory syncytial virus-induced airway reactivity, pulmonary inflammation, and chemokine expression. *Toxicol Sci* 2003;**72**:339–346.
45. Ye Q, Fu J, Mao J, Shang S. Haze is a risk factor contributing to the rapid spread of respiratory syncytial virus in children. *Environ Sci Poll Res* 2016;**23**:20178–20185.
46. Liang Y, Fang L, Pan H, Zhang K, Kan H, Brook JR, Sun Q. PM2.5 in Beijing—temporal pattern and its association with influenza. *Environ Health* 2014;**13**:102.
47. Chen T, Wu D, Chen H, Yan W, Yang D, Chen G, Ma K, Xu D, Yu H, Wang H, Wang T, Guo W, Chen J, Ding C, Zhang X, Huang J, Han M, Li S, Luo X, Zhao J, Ning Q. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ* 2020;**368**:m1091.
48. Guzik TJ, Mohiddin SA, Dimarco A, Patel V, Savvatis K, Marelli-Berg FM, Madhur MS, Tomaszewski M, Maffia P, D'Acquisto F, Nicklin SA, Marian AJ, Nosalski R, Murray EC, Guzik B, Berry C, Touyz RM, Kreutz R, Wang DW, Bhella D, Sagliocco O, Crea F, Thomson EC, McInnes IB. COVID-19 and the cardiovascular system: implications for risk assessment, diagnosis, and treatment options. *Cardiovasc Res* 2020;**116**: 1666–1687.
49. Guo T, Fan J, Chen M, Wu X, Zhang L, He T, Wang H, Wan J, Wang X, Lu Z. Cardiovascular implications of fatal outcomes of patients with Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol* 2020;**5**:811–818.
50. Shi S, Qin M, Shen B, Cai Y, Liu T, Yang F, Gong W, Liu X, Liang J, Zhao Q, Huang H, Yang B, Huang C. Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. *JAMA Cardiol* 2020;**5**:802–810.
51. Inciardi RM, Adamo M, Lupi L, Cani DS, Di Pasquale M, Tomasoni D, Italia L, Zaccone G, Tedino C, Fabbricatore D, Curnis A, Faggiano P, Gorga E, Lombardi CM, Milesi G, Vizzardi E, Volpini M, Nodari S, Specchia C, Maroldi R, Bezzi M, Metra M. Characteristics and outcomes of patients hospitalized for COVID-19 and cardiac disease in Northern Italy. *Eur Heart J* 2020;**41**:1821–1829.
52. European Society of Cardiology (ESC). ESC guidance for the diagnosis and management of CV disease during the COVID-19 pandemic; https://www.escardio.org/Education/COVID-19-and-Cardiology/ESC-COVID-19-Guidance.
53. Frontera A, Martin C, Vlachos K, Sgubin G. Regional air pollution persistence links to COVID-19 infection zoning. *J Infect* 2020;**81**:318–356.
54. Yan J, Grantham M, Pantelic M, Bueno de Mesquita PJ, Albert B, Liu F, Ehrman S, Milton DK, EMIT Consortium. Infectious virus in exhaled breath of symptomatic seasonal influenza cases from a college community. *Proc Natl Acad Sci USA* 2018;**115**: 1081–1086.
55. Liu C, Chen R, Sera F, Vicedo-Cabrera AM, Guo Y, Tong S, Coelho MSZS, Saldiva PHN, Lavigne E, Matus P, Valdes Ortega N, Osorio Garcia S, Pascal M, Stafoggia M, Scortichini M, Hashizume M, Honda Y, Hurtado-Díaz M, Cruz J, Nunes B, Teixeira JP, Kim H, Tobias A, Íñiguez C, Forsberg B, Åström C, Ragettli MS, Guo Y-L, Chen B-Y, Bell ML, Wright CY, Scovronick N, Garland RM, Milojevic A, Kyselý J, Urban A, Orru H, Indermitte E, Jaakkola JJK, Ryti NRI, Katsouyanni K, Analitis A, Zanobetti A, Schwartz J, Chen J, Wu T, Cohen A, Gasparrini A, Kan H. Ambient particulate air pollution and daily mortality in 652 cities. *N Engl J Med* 2019;**381**:705–715.
56. Brook RD, Rajagopalan S, Pope CA III, Brook JR, Bhatnagar A, Diez-Roux AV, Holguin F, Hong Y, Luepker RV, Mittleman MA, Peters A, Siscovick D, Smith SCJrWhitsel, L Kaufman, JD; on behalf of the American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism. Particulate matter air pollution and cardiovascular disease. An update to the scientific statement from the American Heart Association. *Circulation* 2010;**121**:2331–2378.
57. Landrigan PJ, Fuller R, Acosta NJR, Adeyi O, Arnold R, Basu N, Baldé AB, Bertollini R, Bose-O'Reilly S, Boufford JI, Breysse PN, Chiles T, Mahidol C, Coll-Seck AM, Cropper ML, Fobil J, Fuster V, Greenstone M, Haines A, Hanrahan D, Hunter D, Khare M, Krupnick A, Lanphear B, Lohani B, Martin K, Mathiasen KV, McTeer MA, Murray CJL, Ndahimananjara JD, Perera F, Potočnik J, Preker AS, Ramesh J, Rockström J, Salinas C, Samson LD, Sandilya K, Sly PD, Smith KR, Steiner A, Stewart RB, Suk WA, van Schayck OCP, Yadama GN, Yumkella K, Zhong M. *The Lancet Commission on pollution and health. Lancet* 2018;**391**:464–512.
58. Office for National Statistics. *Coronavirus (COVID-19) Related Mortality Rates and the Effects of Air Pollution in England*. 2020; https://www.ons.gov.uk/economy/environmentalaccounts/methodologies/coronaviruscovid19relatedmortalityratesandtheeffectsofairpollutioninengland.
59. Cole MA, Ozgen C, Strobl E. *Air Pollution Exposure and COVID-19*. IZA Institute of Labor Economics 2020; IZA DP No. 13367: https://www.iza.org/publications/dp/13367/air-pollution-exposure-and-covid-19.
60. Travaglio M, Yu Y, Popovic R, Selley L, Santos Leal N, Martins LM. Links between air pollution and COVID-19 in England. https://www.medrxiv.org/content/10.1101/2020.04.16.20067405v5.
61. Smith KR, Bruce N, Balakrishnan K, Adair-Rohani H, Balmes J, Chafe Z, Dherani M, Hosgood HD, Mehta S, Pope D, Rehfuess E, and others in the HAP CRA Risk Expert Group. Millions dead: how do we know and what does it mean? Methods used in the Comparative Risk Assess of Household Air Pollution. *Ann Rev Pub Health* 2014;**35**: 185–206.
62. Chen K, Wang M, Huang C, Kinney PL, Anastas PT. Air pollution reduction and mortality benefit during the COVID-19 outbreak in China. *Lancet Planet Health* 2020; 10.1016/S2542-5196(20)30107-8.
63. Venter ZS, Aunan K, Chowdhury S, Lelieveld J. COVID-19 lockdowns cause global air pollution declines. *Proc Natl Acad Sci USA* 2020;**117**:18984–18990.
64. Eftim S, Dominici F. Multisite time-series studies versus cohort studies: methods, findings, and policy implications. *J Toxicol Environ Health A* 2005;**68**;1191–1205.
65. Beverland IJ, Cohen GR, Heal MR, Carder M, Yap C, Robertson C, Hart CL, Agius RM. A comparison of short-term and long-term air pollution exposure associations with mortality in two cohorts in Scotland. *Environ Health Perspect* 2012; **120**,1280–1285.

Translational perspective

COVID-19 infections and air pollution cause excess mortality from cardiovascular and pulmonary diseases. We estimated the fraction of COVID-19 mortality attributable to the long-term exposure to ambient fine particulate air pollution ($PM_{2.5}$). Global exposure to $PM_{2.5}$ was characterized based on satellite data, and the anthropogenic fraction was calculated with an atmospheric chemistry model. $PM_{2.5}$ contributed ~15% to COVID-19 mortality worldwide, 27% in East Asia, 19% in Europe, and 17% in North America. Globally ~50–60% of the attributable, anthropogenic fraction is related to fossil fuel use, and 70–80% in Europe/West Asia/North America, indicating the potential for substantial health benefits from reducing air pollution exposure.

CORONAVIRUS

Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis

X. Wu[1]*, R. C. Nethery[1]*, M. B. Sabath[1], D. Braun[1,2], F. Dominici[1†]

Assessing whether long-term exposure to air pollution increases the severity of COVID-19 health outcomes, including death, is an important public health objective. Limitations in COVID-19 data availability and quality remain obstacles to conducting conclusive studies on this topic. At present, publicly available COVID-19 outcome data for representative populations are available only as area-level counts. Therefore, studies of long-term exposure to air pollution and COVID-19 outcomes using these data must use an ecological regression analysis, which precludes controlling for individual-level COVID-19 risk factors. We describe these challenges in the context of one of the first preliminary investigations of this question in the United States, where we found that higher historical $PM_{2.5}$ exposures are positively associated with higher county-level COVID-19 mortality rates after accounting for many area-level confounders. Motivated by this study, we lay the groundwork for future research on this important topic, describe the challenges, and outline promising directions and opportunities.

INTRODUCTION

The suddenness and global scope of the coronavirus disease 2019 (COVID-19) pandemic have raised urgent questions that require coordinated investigation to slow the disease's devastation. A critically important public health objective is to identify key modifiable environmental factors that may contribute to the severity of health outcomes [e.g., intensive care unit (ICU) hospitalization and death] among individuals with COVID-19. Numerous scientific studies reviewed by the U.S. Environmental Protection Agency (EPA) have linked fine particles ($PM_{2.5}$; particles with diameter, ≤ 2.5 μm) to a variety of adverse health events (*1*) including death (*2*). It has been hypothesized that because long-term exposure to $PM_{2.5}$ adversely affects the respiratory and cardiovascular systems and increases mortality risk (*3*–*5*), it may also exacerbate the severity of COVID-19 symptoms and worsen the prognosis of this disease (*6*).

Epidemiological studies to estimate the association between long-term exposure to air pollution and COVID-19 hospitalization and death is a rapidly expanding area of research that is attracting attention around the world. Two studies have been published using data from European countries (*7*, *8*), and many more are available as preprints. However, because of the unprecedented nature of the pandemic, researchers face serious challenges when conducting these studies. One key challenge is that, to our knowledge, individual-level data on COVID-19 health outcomes for large, representative populations are not publicly available or accessible to the scientific community. Therefore, the only way to generate preliminary evidence on the link between $PM_{2.5}$ and COVID-19 severity and outcomes using these aggregate data is to use an ecological regression analysis. With this study design, publicly available area-level COVID-19 mortality rates are regressed against area-level air pollution concentrations while accounting for area-level potential confounding factors. Here, we discuss the strengths and limitations of conducting ecological regression analyses of air pollution and COVID-19 health outcomes and describe additional challenges related to evolving data quality, statistical modeling, and control of measured and unmeasured confounding, paving the way for future research on this topic. We discuss these challenges and illustrate them in the context of a specific study, in which we investigated the impact of long-term $PM_{2.5}$ exposure on COVID-19 mortality rates in 3089 counties in the United States, covering 98% of the population.

Illustration of an ecological regression analysis of historical exposure to $PM_{2.5}$ and COVID-19 mortality rate

We begin by describing how to conduct an ecological regression analysis in this setting. COVID-19 death counts (a total of 116,747 deaths) were obtained from the Johns Hopkins University Coronavirus Resource Center and were cumulative up to 18 June 2020. We used data from 3089 counties, of which 1244 (40.3%) had reported zero COVID-19 deaths at the time of our analysis. Daily $PM_{2.5}$ concentrations were estimated across the United States on a $0.01° \times 0.01°$ grid for the period 2000–2016 using well-validated atmospheric chemistry and machine learning models (*9*). We used zonal statistics to aggregate $PM_{2.5}$ concentration estimates to the county level and then averaged across the period 2000–2016 to perform health outcome analyses. Figure 1 illustrates the spatial variation in 2000–2016 average (hereafter referred to as "long-term average") $PM_{2.5}$ concentrations and COVID-19 mortality rates (per 1 million population) by county.

We fit a negative binomial mixed model using COVID-19 mortality rates as the outcome and long-term average $PM_{2.5}$ as the exposure of interest, adjusting for 20 county-level covariates. We conducted more than 80 sensitivity analyses to assess the robustness of the findings to various modeling assumptions. We found that an increase of 1 $\mu g/m^3$ in the long-term average $PM_{2.5}$ is associated with a statistically significant 11% (95% CI, 6 to 17%) increase in the county's COVID-19 mortality rate (see Table 1); this association continues to be stable as more data accumulate (fig. S3). We also found that population density, days since the first COVID-19 case was reported, median household income, percent of owner-occupied housing, percent of the adult population with less than high school

[1]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [2]Department of Data Sciences, Dana-Farber Cancer Institute, Boston, MA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: fdominic@hsph.harvard.edu

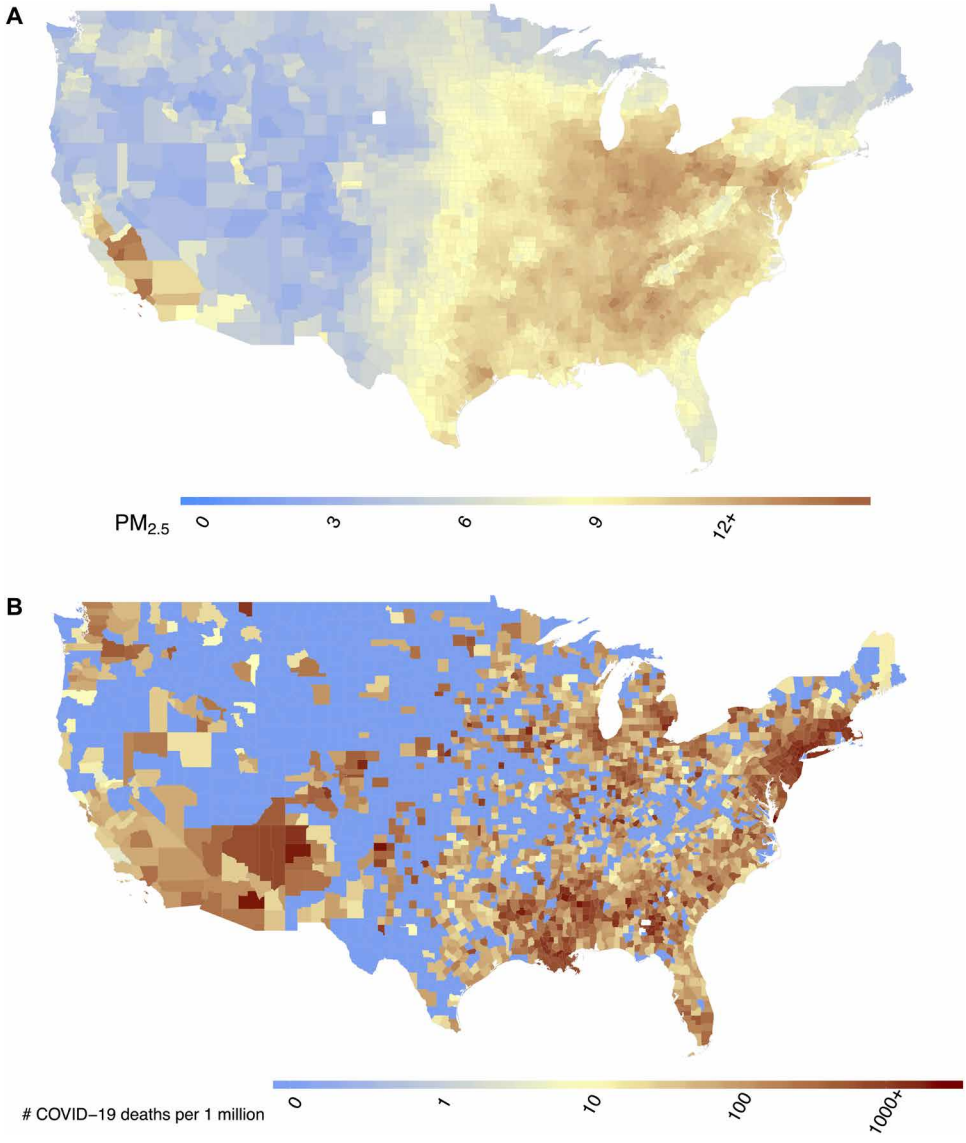


Fig. 1. National maps of historical $PM_{2.5}$ concentrations and COVID-19 deaths. Maps show (**A**) county-level 17-year long-term average of $PM_{2.5}$ concentrations (2000–2016) in the United States in $\mu g/m^3$ and (**B**) county-level number of COVID-19 deaths per 1 million population in the United States up to and including 18 June 2020.

education, age distribution, and percent of Black residents are important predictors of the COVID-19 mortality rate in the model. We found a 49% (95% CI, 38 and 61%) increase in COVID-19 mortality rate associated with a 1-SD (per 14.1%) increase in percent Black residents of the county. Details on the data sources, statistical methods, and analyses are summarized in the Supplementary Materials. All data sources used in the analyses, along with fully reproducible code, are publicly available at https://github.com/wxwx1993/PM_COVID.

Strengths and limitations of an ecological regression analysis

Ecological regression analysis provides a simple and cost-effective approach for studying potential associations between historical exposure to air pollution and increased vulnerability to COVID-19 in large representative populations, as illustrated in our study in the previous section. This approach is regularly applied in many areas of research (*10*). Using our study as an example, we summarize in Table 2 the strengths, limitations, and opportunities considering (i) study design, (ii) COVID-19 health outcome data, (iii) historical exposure to air pollution, and (iv) measured and unmeasured confounders, with the goal of paving the way for future research.

Among the key limitations, by design, ecological regression analyses are unable to adjust for individual-level risk factors (e.g., age, race, and smoking status); when individual-level data are unavailable, this approach leaves us unable to make conclusions regarding individual-level associations. In the context of COVID-19 health outcomes, this is a severe limitation, as individual-level risk factors are known to affect COVID-19 health outcomes. It is important to note that confusion between ecological associations and individual associations may present an ecological fallacy. In extreme

cases, this fallacy can lead to associations detected in ecological regression that do not exist or are in the opposite direction of true associations at the individual level. However, ecological regression analyses still allow us to make conclusions at the area level, which can be useful for policy-making (*11*). For the association between COVID-19 health outcomes and $PM_{2.5}$ exposure, we argue that area-level conclusions are valuable, as they can inform important immediate policy actions that will benefit public health, such as (i) prioritization of precautionary measures [e.g., personal protective equipment (PPE) allocations and hospital beds] to areas with historical higher air pollution and (ii) further strengthening the scientific argument for lowering the U.S. National Ambient Air Quality Standards for $PM_{2.5}$ and other pollutants. To completely avoid potential ecological bias, a representative sample of individual-level data is necessary. While this may not be feasible in the near future, as some COVID-19 outcome data become available at the individual level, existing approaches that augment county-level data with individual-level data (*12*) could be used to correct for ecological bias.

Furthermore, air pollution exposure misclassification, due to between-area mobility and within-area variation, is another potential source of bias that could affect the ecological regression results described in our example study. Methods to account for the propagation of exposure error into the ecological regression model (*13*) could be applied to help mitigate the impact of measurement error. Outcome misclassification is another limitation that can be partially overcome by accessing nationwide registry data with the validated cause of death (*14*). As in all observational studies, adjustment for measured and unmeasured confounding presents another key challenge in ecological regression analyses, which may be exacerbated when dealing with dynamic pandemic data, as in our study. Conducting studies using both traditional regressions and methods for causal inference as in Wu *et al.* (*2*) is necessary to assess the robustness of the findings.

Increasing the scientific rigor of research in this area requires access to representative, individual-level data on COVID-19 health outcomes, including information about patients' residential address, demographics, and individual-level confounders. This is an enormous challenge that will require consideration of many privacy, legal, and ethical trade-offs (*14*). Future areas of research also include the application of statistical methods to quantify and correct for ecological bias and measurement error, reproducible methods for causal inference, and sensitivity analysis of measured and unmeasured confounding bias as suggested above. These strengths and limitations are illustrated further in the context of our own study (see the Supplementary Materials).

Table 1. Mortality rate ratios (MRR), 95% confidence intervals (CI), and *P* values for all variables in the main analysis. Details of the statistical models are available in section S2. Q, quintile.

	MRR	95% CI	*P* value
$PM_{2.5}$	1.11	(1.06–1.17)	0.00
Population density (Q2)	0.91	(0.71–1.15)	0.42
Population density (Q3)	0.91	(0.71–1.16)	0.45
Population density (Q4)	0.74	(0.57–0.95)	0.02
Population density (Q5)	0.92	(0.69–1.23)	0.56
% In poverty	1.04	(0.96–1.12)	0.31
Log(median house value)	1.13	(0.99–1.29)	0.07
Log(median household income)	1.19	(1.04–1.35)	0.01
% Owner-occupied housing	1.12	(1.04–1.20)	0.00
% Less than high school education	1.20	(1.10–1.32)	0.00
% Black	1.49	(1.38–1.61)	0.00
% Hispanic	1.06	(0.97–1.16)	0.23
% ≥ 65 years of age	1.04	(0.93–1.17)	0.46
% 45–64 years of age	0.77	(0.67–0.90)	0.00
% 15–44 years of age	0.76	(0.68–0.85)	0.00
Days since stay-at-home order	1.18	(0.92–1.52)	0.20
Days since first case	2.40	(2.05–2.80)	0.00
Rate of hospital beds	1.00	(0.93–1.08)	0.95
% Obese	0.96	(0.90–1.03)	0.32
% Smokers	1.13	(1.00–1.28)	0.05
Average summer temperature (°F)	1.11	(0.95–1.30)	0.20
Average winter temperature (°F)	0.86	(0.69–1.07)	0.19
Average summer relative humidity (%)	0.93	(0.80–1.09)	0.38
Average winter relative humidity (%)	0.97	(0.87–1.07)	0.52

DISCUSSION

Ecological regression analyses are crucial to stimulate innovations in a rapidly evolving area of research. Ongoing research has already focused on overcoming some aspects of these limitations (*8*, *15*). For example, ecological regression analysis of air pollution and COVID-19, using data with finer geographic resolution, is being conducted for different countries and regions around the world. Cole *et al.* (*8*) published an ecological regression analysis using data in Dutch municipalities and found results consistent with our own investigation; the California Air Resources Board (CARB) is planning to conduct a similar study at the census tract level (*15*). Although an ecological regression analysis cannot provide insight into the mechanisms underlying the relationship between $PM_{2.5}$ exposure and COVID-19 mortality, studies are starting to shed light on the potential biological mechanisms that may explain the relationship between air pollution and viral infection outcomes (*16*). For example, it has been hypothesized that chronic exposure to $PM_{2.5}$ causes alveolar angiotensin-converting enzyme 2 (ACE-2) receptor overexpression and impairs host defenses (*17*). This could cause a more severe form

Table 2. Strengths and limitations of ecological regression analyses applied to research on air pollution and COVID-19 and opportunities for future research.

	Strengths	Limitations	Future research
Study design: ecological regression	Feasible, timely, and cost-effective	Cannot be used to make inference about individual-level associations, doing so leads to ecological fallacy	Augment county-level data with individual-level data to adjust for ecological bias (*12*)
	Data are representative of the entire U.S. population	Cannot adjust for individual-level risk factors such as age, gender, and race (*19*–*21*)	Conduct studies of individual-level health records using traditional regression and causal inference methods as in Wu *et al.* (*2*)
	Allows inference at the area level, which can be useful for policy-making (*11*)	Results are sensitive to the assumptions of the statistical model (*11*)	
	Computationally efficient and can be conducted daily to allow for the dynamic nature of the data and observe temporal trends; see fig. S3		
	Facilitates comparison of results across countries		
Outcome: COVID-19 deaths aggregated at the county level	Publicly available data updated almost daily	Potential for outcome misclassification (*22*), particularly differential misclassification over time and space, which could bias results	Access to nationwide registry data with the validated cause of death (*14*)
			Analyses using county excess deaths as the outcome (*23*)
Exposure: 2000–2016 average exposure to $PM_{2.5}$ at the county level	Use of well-validated atmospheric chemistry models and machine learning models (*9*, *24*)	Aggregation assumes that everyone in a county experiences the same exposures, leading to exposure misclassification, especially for the largest counties	Individual-level data on COVID-19 deaths with geocoded addresses to link to air pollution data at the place of residence
	$PM_{2.5}$ exposure estimated at fine grids, which can be aggregated to the county level to assess exposure even in unmonitored areas (*24*)	Can be used to assess historical exposures to air pollution but not real-time exposures	Additional statistical methods to account for the propagation of exposure error into the ecological regression model (*13*)
	As opposed to using monitor data, aggregation of modeled estimates ensures that county $PM_{2.5}$ exposure estimates represent the distribution across the entire area		
Measured confounders	More than 20 area-level variables capture age distribution, race distribution, socioeconomic status, population density, behavioral risk factors, epidemic stage, and stay-at-home orders (see tables S1 and S2)	County average features may not represent the features of COVID-19 patients, leading to inadequate adjustment	Causal inference approaches to adjust for measured confounding bias, producing results that are less sensitive to statistical modeling assumptions
	These overlap with the confounder sets used in much of the previous literature on air pollution and health (*25*, *26*)	Difficult to formalize the notion of "epidemic stage," which may be an important confounder	
		The threat of unmeasured confounding bias still present	Causal inference approaches to assess covariate balance (*2*)
		Sensitive to the form of the statistical model specified (i.e., assumptions of linearity and no effect modification)	Individual-level data on key measured confounders such as smoking and body mass index
Unmeasured confounders	Leverage existing approaches, such as the calculation of the E-value (*27*), to assess how strong the effect of an unmeasured confounder would need to be to explain away the associations detected (see section S3)	The most important threat to the validity of any observational study Even measures like the E-value cannot inform us about the likelihood that a strong unmeasured confounder exists; this must be evaluated on the basis of subject matter knowledge	Natural experiment designs and instrumental variables can be used to reduce the threat of unmeasured confounding but are less common

of COVID-19 in ACE-2–depleted lungs, increasing the likelihood of poor outcomes, including death (*18*).

The associations detected in ecological regression analyses provide strong justification for follow-up investigations as more and higher-quality COVID-19 data become available. Such studies would include validation of our findings with other data sources and study types, as well as investigations into mediating factors and effect modifiers, biological mechanisms, impacts of $PM_{2.5}$ exposure timing, and relationships between $PM_{2.5}$ and other COVID-19 outcomes such as hospitalization. Research on how modifiable factors may exacerbate COVID-19 symptoms and increase mortality risk is essential to guide policies and behaviors to minimize fatality related to the pandemic. Such research could also provide a strong scientific argument for revision of the U.S. Ambient Air Quality Standards for $PM_{2.5}$ and other environmental policies in the midst of a pandemic.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/45/eabd4049/DC1

REFERENCES AND NOTES

1. U.S. Environmental Protection Agency (EPA), Integrated Science Assessment (ISA) for Particulate Matter (Final Report, 2019). U.S. Environmental Protection Agency, EPA/600/R-19/188 (2019); https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=347534 [accessed 15 September 2020].
2. X. Wu, D. Braun, J. Schwartz, M. A. Kioumourtzoglou, F. Dominici, Evaluating the impact of long-term exposure to fine particulate matter on mortality among the elderly. *Sci. Adv.* **6**, eaba5692 (2020).
3. R. D. Brook, S. Rajagopalan, C. Arden Pope III, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, Y. Hong, R. V. Luepker, M. A. Mittleman, A. Peters, D. Siscovick, S. C. Smith Jr., L. Whitsel, J. D. Kaufman, Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **121**, 2331–2378 (2010).
4. C. A. Pope III, R. T. Burnett, G. D. Thurston, M. J. Thun, E. E. Calle, D. Krewski, J. J. Godleski, Cardiovascular mortality and long-term exposure to particulate air pollution: Epidemiological evidence of general pathophysiological pathways of disease. *Circulation* **109**, 71–77 (2004).
5. C. A. Pope III, N. Coleman, Z. A. Pond, R. T. Burnett, Fine particulate air pollution and human mortality: 25+ years of cohort studies. *Environ. Res.* **183**, 108924 (2020).
6. T. Benmarhnia, Linkages between air pollution and the health burden from COVID-19: Methodological challenges and opportunities. *Am. J. Epidemiol.* , kwaa148 (2020).
7. E. Conticini, B. Frediani, D. Caro, Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? *Environ. Pollut.* **261**, 114465 (2020).
8. M. A. Cole, C. Ozgen, E. Strobl, Air pollution exposure and COVID-19 in Dutch municipalities. *Environ. Resour. Econ. (Dordr)* , 1–30 (2020).
9. A. van Donkelaar, R. V. Martin, C. Li, R. T. Burnett, Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environ. Sci. Technol.* **53**, 2595–2611 (2019).
10. G. King, M. A. Tanner, O. Rosen, *Ecological Inference: New Methdological Strategies* (Cambridge Univ. Press, 2004).
11. A. Gelman, D. K. Park, S. Ansolabehere, P. N. Price, L. C. Minnite, Models, assumptions and model checking in ecological regressions. *J. Royal Stat. Soc. Ser. A* **164**, 101–118 (2001).
12. C. Jackson, N. Best, S. Richardson, Improving ecological inference using individual-level data. *Stat. Med.* **25**, 2136–2159 (2006).
13. J. Richmond-Bryant, T. C. Long, Influence of exposure measurement errors on results from epidemiologic studies of different designs. *J. Expo. Sci. Environ. Epidemiol.* **30**, 420–429 (2020).
14. D. F. Sittig, H. Singh, COVID-19 and the need for a national health information technology infrastructure. *JAMA* **323**, 2373–2374 (2020).
15. InisdeEPA.com, CARB study on pollution link to virus deaths may spur push for PM cuts (May 6, 2020); https://insideepa.com/daily-news/carb-study-pollution-link-virus-deaths-may-spur-push-pm-cuts.
16. J. Ciencewicki, I. Jaspers, Air pollution and respiratory viral infection. *Inhal. Toxicol.* **19**, 1135–1146 (2007).
17. L. Miyashita, G. Foley, S. Semple, J. Grigg, Traffic-derived particulate matterand angiotensin-converting enzyme 2 expression in human airway epithelial cells. *bioRxiv* 2020.2005.2015.097501 , (2020).
18. A. Frontera, L. Cianfanelli, K. Vlachos, G. Landoni, G. Cremona, Severe air pollution links to higher mortality in COVID-19 patients: The "double-hit" hypothesis. *J. Infect.* **81**, 255–259 (2020).
19. S. Greenland, H. Morgenstern, Ecological bias, confounding, and effect modification. *Int. J. Epidemiol.* **18**, 269–274 (1989).
20. S. Greenland, J. Robins, Invited commentary: Ecologic studies–Biases, misconceptions, and counterexamples. *Am. J. Epidemiol.* **139**, 747–760 (1994).
21. S. Richardson, I. Stücker, D. Hémon, Comparison of relative risks obtained in ecological and individual studies: Some methodological considerations. *Int. J. Epidemiol.* **16**, 111–120 (1987).
22. New York City Department of Health; Mental Hygiene Covid-Response Team, Preliminary estimate of excess mortality during the COVID-19 outbreak–New York City, March 11–May 2, 2020. *MMWR Morb. Mortal Wkly. Rep.* **69**, 603–605 (2020).
23. R. J. Acosta, R. A. Irizarry, Monitoring health systems by estimating excess mortality. *medRxiv* 2020.06.06.20120857 , (2020).
24. Q. Di, I. Kloog, P. Koutrakis, A. Lyapustin, Y. Wang, J. Schwartz, Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. *Environ. Sci. Technol.* **50**, 4712–4721 (2016).
25. Q. Di, Y. Wang, A. Zanobetti, Y. Wang, P. Koutrakis, C. Choirat, F. Dominici, J. D. Schwartz, Air pollution and mortality in the Medicare population. *N. Engl. J. Med.* **376**, 2513–2522 (2017).
26. D. W. Dockery, C. A. Pope III, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris Jr., F. E. Speizer, An association between air pollution and mortality in six US cities. *N. Engl. J. Med.* **329**, 1753–1759 (1993).
27. S. Haneuse, T. J. VanderWeele, D. Arterburn, Using the E-value to assess the potential effect of unmeasured confounding in observational studies. *JAMA* **321**, 602–603 (2019).
28. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
29. Q. Di, H. Amini, L. Shi, I. Kloog, R. Silvern, J. Kelly, M. B. Sabath, C. Choirat, P. Koutrakis, A. Lyapustin, Y. Wang, L. J. Mickley, J. Schwartz, An ensemble-based model of $PM_{2.5}$ concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **130**, 104909 (2019).
30. J. G. Booth, G. Casella, H. Friedl, J. P. Hobert, Negative binomial loglinear mixed models. *Stat. Model.* **3**, 179–191 (2003).
31. R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2018); www.R-project.org/.
32. X. Zhang, H. Mallick, Z. Tang, L. Zhang, X. Cui, A. K. Benson, N. Yi, Negative binomial mixed models for analyzing microbiome count data. *BMC Bioinformatics* **18**, 4 (2017).
33. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
34. Q. H. Vuong, Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* **57**, 307–333 (1989).
35. T. J. VanderWeele, P. Ding, Sensitivity analysis in observational research: Introducing the E-value. *Ann. Intern. Med.* **167**, 268–274 (2017).
36. M. B. Mathur, P. Ding, C. A. Riddell, T. J. VanderWeele, Web site and R package for computing E-values. *Epidemiology* **29**, e45–e47 (2018).
37. D. D. Ingram, S. J. Franco, 2013 NCHS urban-rural classification scheme for counties. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics (2014).
38. Health Effects Institute, State of global air 2019: A special report on global exposure to air pollution and its disease burden (2019); www.stateofglobalair.org/sites/default/files/soga_2019_report.pdf [accessed 15 September 2020].
39. R. D. Brook, B. Franklin, W. Cascio, Y. Hong, G. Howard, M. Lipsett, R. Luepker, M. Mittleman, J. Samet, S. C. Smith Jr., I. Tager, Air pollution and cardiovascular disease: A statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association. *Circulation* **109**, 2655–2671 (2004).
40. F. Dominici, R. D. Peng, M. L. Bell, L. Pham, A. McDermott, S. L. Zeger, J. M. Samet, Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA* **295**, 1127–1134 (2006).
41. R. C. Puett, J. E. Hart, J. D. Yanosky, C. Paciorek, J. Schwartz, H. Suh, F. E. Speizer, F. Laden, Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. *Environ. Health Perspect.* **117**, 1697–1701 (2009).
42. R. J. Šrám, B. Binková, J. Dejmek, M. Bobak, Ambient air pollution and pregnancy outcomes: A review of the literature. *Environ. Health Perspect.* **113**, 375–382 (2005).
43. G. A. Wellenius, M. R. Burger, B. A. Coull, J. Schwartz, H. H. Suh, P. Koutrakis, G. Schlaug, D. R. Gold, M. A. Mittleman, Ambient air pollution and the risk of acute ischemic stroke. *Arch. Intern. Med.* **172**, 229–234 (2012).

44. J. Rhee, F. Dominici, A. Zanobetti, J. Schwartz, Y. Wang, Q. Di, J. Balmes, D. C. Christiani, Impact of long-term exposures to ambient $PM_{2.5}$ and ozone on ARDS risk for older adults in the United States. *Chest* **156**, 71–79 (2019).
45. Y. Cui, Z.-F. Zhang, J. Froines, J. Zhao, H. Wang, S.-Z. Yu, R. Detels, Air pollution and case fatality of SARS in the People's Republic of China: An ecologic study. *Environ. Health* **2**, 15 (2003).
46. C. A. Pope III, Respiratory disease associated with community air pollution and a steel mill, Utah Valley. *Am. J. Public Health* **79**, 623–628 (1989).
47. D. P. Croft, W. Zhang, S. Lin, S. W. Thurston, P. K. Hopke, E. van Wijngaarden, S. Squizzato, M. Masiol, M. J. Utell, D. Q. Rich, Associations between source-specific particulate matter and respiratory infections in New York State adults. *Environ. Sci. Technol.* **54**, 975–984 (2020).
48. B. D. Horne, E. A. Joy, M. G. Hofmann, P. H. Gesteland, J. B. Cannon, J. S. Lefler, D. P. Blagev, E. K. Korgenski, N. Torosyan, G. I. Hansen, D. Kartchner, C. A. Pope III, Short-term elevation of fine particulate matter air pollution and acute lower respiratory infection. *Am. J. Respir. Crit. Care Med.* **198**, 759–766 (2018).
49. Q. Di, L. Dai, Y. Wang, A. Zanobetti, C. Choirat, J. D. Schwartz, F. Dominici, Association of short-term exposure to air pollution with mortality in older adults. *JAMA* **318**, 2446–2456 (2017).
50. D.-H. Tsai, M. Riediker, A. Berchet, F. Paccaud, G. Waeber, P. Vollenweider, M. Bochud, Effects of short- and long-term exposures to particulate matter on inflammatory marker levels in the general population. *Environ. Sci. Pollut. Res. Int.* **26**, 19697–19704 (2019).
51. K. F. Morales, J. Paget, P. Spreeuwenberg, Possible explanations for why some countries were harder hit by the pandemic influenza virus in 2009 – A global mortality impact modeling study. *BMC Infect. Dis.* **17**, 642 (2017).
52. Z. Xu, W. Hu, G. Williams, A. C. A. Clements, H. Kan, S. Tong, Air pollution, temperature and pediatric influenza in Brisbane, Australia. *Environ. Int.* **59**, 384–388 (2013).
53. K. Clay, J. Lewis, E. Severnini, Pollution, infectious disease, and mortality: Evidence from the 1918 Spanish influenza pandemic. *J. Econ. Hist.* **78**, 1179–1209 (2018).
54. C. A. Pope III, A. Bhatnagar, J. P. McCracken, W. Abplanalp, D. J. Conklin, T. O'Toole, Exposure to fine particulate air pollution is associated with endothelial injury and systemic inflammation. *Circ. Res.* **119**, 1204–1214 (2016).
55. S. Becker, J. M. Soukup, Exposure to urban air particulates alters the macrophage-mediated inflammatory response to respiratory viral infection. *J. Toxicol. Environ. Health A* **57**, 445–457 (1999).
56. P. M. Kaan, R. G. Hegele, Interaction between respiratory syncytial virus and particulate matter in guinea pig alveolar macrophages. *Am. J. Respir. Cell Mol. Biol.* **28**, 697–704 (2003).
57. A. L. Lambert, J. B. Mangum, M. P. DeLorme, J. I. Everitt, Ultrafine carbon black particles enhance respiratory syncytial virus-induced airway reactivity, pulmonary inflammation, and chemokine expression. *Toxicol. Sci.* **72**, 339–346 (2003).
58. L. Peng, X. Zhao, Y. Tao, S. Mi, J. Huang, Q. Zhang, The effects of air pollution and meteorological factors on measles cases in Lanzhou, China. *Environ. Sci. Pollut. Res. Int.* **27**, 13524–13533 (2020).
59. Q. Ye, J.-f. Fu, J.-h. Mao, S.-q. Shang, Haze is a risk factor contributing to the rapid spread of respiratory syncytial virus in children. *Environ. Sci. Pollut. Res. Int.* **23**, 20178–20185 (2016).

Acknowledgments: The computations in this paper were run on (i) the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University and (ii) the Research Computing Environment supported by Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University. We gratefully acknowledge support from the 2020 Star-Friedman Challenge for Promising Scientific Research, the Climate Change Solutions Fund at Harvard University, and the Fernholz Foundation. We would like to thank L. Goodwin and S. Tobin for editorial assistance in the preparation of this manuscript. **Funding:** This work was made possible by support from NIH grants R01 ES024332-01A1, P50MD010428, RO1 ES026217, RO1 ES028033, R01 ES030616, R01 AG066793-01, and R01 MD012769; Health Effects Institute grant (HEI) 4953-RFA14-3/16-4: and US EPA grant 83587201-0. The funding sources did not participate in the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript. The research described in this article was conducted under contract to the HEI, an organization jointly funded by the EPA (Assistance Award No. R-83467701), and certain motor vehicle and engine manufacturers. The contents of this article do not necessarily reflect the views of HEI or its sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and engine manufacturers. **Author contributions:** X.W. and R.C.N. contributed equally to the paper. X.W. and R.C.N. contributed to formulation of the idea, data preparation, data analysis, data interpretation, and writing of the manuscript. M.B.S. and D.B. contributed to data preparation, data interpretation, and review of the manuscript. F.D. contributed to formulation of the idea, study design, data interpretation, funding, and writing of the manuscript. All authors contributed to the interpretation of the results and critical revision of the manuscript for important intellectual content and approved the final version of the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. F.D. is the guarantor. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Data and code are publicly available at https://github.com/wxwx1993/PM_COVID. Additional data related to this paper may be requested from the authors.

Submitted 25 June 2020
Accepted 18 September 2020
Published 4 November 2020
10.1126/sciadv.abd4049

Citation: X. Wu, R. C. Nethery, M. B. Sabath, D. Braun, F. Dominici, Air pollution and COVID-19 mortality in the United States: Strengths and limitations of an ecological regression analysis. *Sci. Adv.* **6**, eabd4049 (2020).

Long-term effects of $PM_{2.5}$ on neurological disorders in the American Medicare population: a longitudinal cohort study

Liuhua Shi, Xiao Wu*, Mahdieh Danesh Yazdi, Danielle Braun, Yara Abu Awad, Yaguang Wei, Pengfei Liu, Qian Di, Yun Wang, Joel Schwartz, Francesca Dominici, Marianthi-Anna Kioumourtzoglou†, Antonella Zanobetti†*

Summary

Background Accumulating evidence links fine particulate matter ($PM_{2.5}$) to premature mortality, cardiovascular disease, and respiratory disease. However, less is known about the influence of $PM_{2.5}$ on neurological disorders. We aimed to investigate the effect of long-term $PM_{2.5}$ exposure on development of Parkinson's disease or Alzheimer's disease and related dementias.

Methods We did a longitudinal cohort study in which we constructed a population-based nationwide open cohort including all fee-for-service Medicare beneficiaries (aged ≥65 years) in the contiguous United States (2000–16) with no exclusions. We assigned $PM_{2.5}$ postal code (ie, ZIP code) concentrations based on mean annual predictions from a high-resolution model. To accommodate our very large dataset, we applied Cox-equivalent Poisson models with parallel computing to estimate hazard ratios (HRs) for first hospital admission for Parkinson's disease or Alzheimer's disease and related dementias, adjusting for potential confounders in the health models.

Findings Between Jan 1, 2000, and Dec 31, 2016, of 63 038 019 individuals who were aged 65 years or older during the study period, we identified 1·0 million cases of Parkinson's disease and 3·4 million cases of Alzheimer's disease and related dementias based on primary and secondary diagnosis billing codes. For each 5 μg/m³ increase in annual $PM_{2.5}$ concentrations, the HR was 1·13 (95% CI 1·12–1·14) for first hospital admission for Parkinson's disease and 1·13 (1·12–1·14) for first hospital admission for Alzheimer's disease and related dementias. For both outcomes, there was strong evidence of linearity at $PM_{2.5}$ concentrations less than 16 μg/m³ (95th percentile of the $PM_{2.5}$ distribution), followed by a plateaued association with increasingly larger confidence bands.

Interpretation We provide evidence that exposure to annual mean $PM_{2.5}$ in the USA is significantly associated with an increased hazard of first hospital admission with Parkinson's disease and Alzheimer's disease and related dementias. For the ageing American population, improving air quality to reduce $PM_{2.5}$ concentrations to less than current national standards could yield substantial health benefits by reducing the burden of neurological disorders.

Funding The Health Effects Institute, The National Institute of Environmental Health Sciences, The National Institute on Aging, and the HERCULES Center.

Lancet Planet Health 2020; 4: e557–65

Published Online October 19, 2020 https://doi.org/10.1016/S2542-5196(20)30227-8

*Contributed equally

†Co-senior authors

Department of Environmental Health (L Shi ScD, M Danesh Yazdi PhD, Y Wei MS, Prof J Schwartz PhD, A Zanobetti PhD) **and Department of Biostatistics** (X Wu MS, D Braun PhD, Y Wang PhD, Prof F Dominici PhD), **Harvard T H Chan School of Public Health, Boston, MA, USA; Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA** (L Shi)**; Department of Data Sciences, Dana-Farber Cancer Institute, Boston, MA, USA** (D Braun)**; Department of Psychology, Concordia University, Montreal, QC, Canada** (Y Abu Awad ScD)**; School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA** (P Liu PhD)**; Vanke School of Public Health, Tsinghua University, Beijing, China** (Q Di ScD)**; and Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA** (M-A Kioumourtzoglou ScD)

Correspondence to: Dr Antonella Zanobetti, Department of Environmental Health, Harvard T H Chan School of Public Health, Boston, MA 02215, USA **azanobet@hsph.harvard.edu**

Introduction

Globally, neurological disorders are the leading group-cause of disability and the second leading group-cause of death, posing an urgent and substantial worldwide public health challenge.[1] Parkinson's disease and Alzheimer's disease and related dementias are the most prevalent neurodegenerative diseases.[2] Worldwide, an estimated 6 million people have Parkinson's disease and 44 million people have Alzheimer's disease and related dementias.[1] The Global Burden of Diseases, Injuries, and Risk Factors Study 2016 analysis estimated that, since 1990, the prevalence of Parkinson's disease has increased by 145% and Alzheimer's disease and related dementias have increased by 117%. The prevalence of these conditions is expected to continue to increase due to lengthening life expectancy.[1] As no cure exists yet for these conditions, the identification of modifiable risk factors, such as environmental exposures, should be a top research priority.

Concern is mounting that air pollution increases the risk for neurological disorders. Emerging evidence has shown that particulate air pollution is associated with impaired cognitive function,[3,4] accelerated cognitive decline,[5,6] Parkinson's disease, Alzheimer's disease, and dementia.[7–9] Research suggests that air pollution might contribute to the potential onset of neurodegeneration via mechanisms such as oxidative stress, systemic inflammation, and neuroinflammation, among others.[10–12] There is also evidence that air pollution might exacerbate disease progression by accelerating these biological pathways or worsening intermediate processes.[13,14] Therefore, the first hospital admission with a relevant diagnosis code is occurring sooner than expected. Previous studies that used hospital admission data to assess the effect of air pollution exposure on progression of Parkinson's disease and Alzheimer's disease and related dementias included populations residing in the southeastern US region,[7,15] the

Research in context

Evidence before this study

Air pollution is a known risk factor for poorer human health. Concern is mounting that air pollution increases the risk for neurological disorders, the leading group-cause of disability and the second leading group-cause of death according to the Global Burden of Diseases, Injuries, and Risk Factors Study 2016. We searched PubMed and Google Scholar for studies examining associations of air pollution exposure with neurological disorders published from database inception until Aug 25, 2020. We used the keywords: ("PM2.5" OR "fine particulate matter" OR "fine particles" OR "air pollution" OR "air pollutants") AND ("neurological" OR "neurodegeneration" OR "neurodegenerative" OR "cognitive" OR "Parkinson's disease" OR "Alzheimer's disease" OR "dementia"). Although toxicological evidence links long-term $PM_{2·5}$ exposure with adverse effects on the nervous system, the epidemiological evidence remains scarce. Emerging evidence has shown that particulate matter air pollution is associated with impaired cognitive function, accelerated cognitive decline, Parkinson's disease, Alzheimer's disease, and dementia. The studies using hospital admission data to look at the effect of particulate matter air pollution on these conditions generally included populations residing in well monitored urban areas, or in a single region of the USA, or in a province of Canada. To date, no study has been done nationwide in the USA. Previous studies also focused on older data; air pollution concentrations in the USA have been steadily decreasing, so it is essential to establish whether these associations persist even at lower concentrations.

Added value of this study

To our knowledge, this is the first nationwide cohort study of the association between $PM_{2·5}$ exposures and neurodegenerative disease in the USA. Our findings provide strong epidemiological evidence for the association between air pollution and neurological disorders. We showed that long-term $PM_{2·5}$ exposures were significantly associated with an increased risk of first hospital admission with primary or secondary diagnosis codes for Parkinson's disease and Alzheimer's disease and related dementias. In addition, we observed that risk of first hospital admission with a diagnosis code for these conditions, as a proxy for neurodegeneration, linearly increased with increasing $PM_{2·5}$ concentrations less than the current national standards (annual mean 12 μg/m³), suggesting that no safe threshold exists for health-harming pollution concentrations. One highlight of this paper is that we are leveraging an unparalleled amount of data compared with any previous air pollution study to our knowledge, to provide robust epidemiological evidence with the highest possible scientific rigour. Another key feature is the use of innovative computational approaches to accommodate our very large datasets, which can be applicable to other epidemiological studies that face similar challenges in the era of big data.

Implications of all the available evidence

Our study adds to the small but emerging evidence base indicating that long-term air pollution exposures are linked to an increased risk of neurological health deterioration, even at $PM_{2·5}$ concentrations less than the current national standards. Our findings suggest that policies that result in further reductions in ambient $PM_{2·5}$ concentrations can yield substantial health benefits in the ageing American population, even for those already exposed to low $PM_{2·5}$ concentrations.

Ontario province of Canada,[8] and well monitored urban areas in the northeastern USA.[9] To the best of our knowledge, no study to date has been done in the whole US population. Previous studies also focused on older data; as air pollution concentrations have been steadily decreasing in the past few decades in the USA although increases have been seen in some regions, it is essential to establish whether these associations persist even at low concentrations. Hence, evidence remains scarce for the health effects of long-term exposure to low amounts of air pollution across the USA, including locations with sparse or no monitoring.

We aimed to investigate the effect of long-term exposure to fine particulate matter ($PM_{2·5}$) on hospital admissions with a Parkinson's disease or an Alzheimer's disease and related dementias diagnosis code. We leveraged a nationwide comprehensive dataset integrating highly accurate and well validated high-resolution $PM_{2·5}$ prediction models and health data for all fee-for-service Medicare beneficiaries across the contiguous United States (2000–16). To address the computational challenges, we applied a novel computationally scalable re-parameterised Cox-equivalent Poisson model.

Methods

Study design and participants

We did a longitudinal cohort study in which we constructed a cohort including all Medicare-fee-for-service beneficiaries who were aged 65 years or older in the USA from Jan 1, 2000, to Dec 31, 2016, using the Medicare part A data. We obtained the Medicare inpatient hospital claims from the Medicare Provider and Analysis Review files, which include one record per hospital admission. People are eligible to enter Medicare after they turn 65 years of age, and for each beneficiary, follow-up started on Jan 1, 2000, or Jan 1 of the year following entry into the cohort, until first admission with diagnosis codes for each outcome separately (ie, Parkinson's disease or Alzheimer's disease and related dementias), death, or the end of the study period, whichever came first. We extracted age, sex, race, postal code (ie, ZIP code) of residence, and Medicaid eligibility for each

beneficiary in each follow-up year. Medicaid (which is distinct from Medicare) is a joint federal–state insurance programme that provides health and nursing home coverage to Americans of all ages on low incomes or with disabilities. Medicaid eligibility is a proxy for individual-level socioeconomic status—ie, a Medicare beneficiary eligible for Medicaid is likely to have lower socioeconomic status. This study was done under a protocol approved by the Human Subjects Committee of the Harvard T H Chan School of Public Health. Written informed consent of individuals was not required due to the nature of the study.

Procedures

We used International Classification of Diseases (ICD) codes to identify Parkinson's disease (ICD-9: 332; ICD-10: G20, G21·11, G21·19, and G21·8) or Alzheimer's disease and related dementias (ICD-9: 331·0, 290; ICD-10: G30·9, and F05) admissions as principal or secondary diagnoses during the study period (appendix p 3). We observed some overlap in diagnoses for Alzheimer's disease and dementia. We found that 298 461 (24·2%) of 1 233 132 Medicare-fee-for-service beneficiaries with a dementia diagnosis also received an Alzheimer's disease diagnosis, while of 2 490 431 Medicare-fee-for-service beneficiaries with Alzheimer's disease diagnoses, 298 461 (12·0%) also received a dementia diagnosis. This overlap in diagnoses probably reflects different classifications in different medical centres, but not diagnostic misclassification, as routinely collected health data have been shown to achieve high positive predictive values.[15] Therefore, following the literature,[8] we combined Alzheimer's disease and dementia into one outcome for the main analysis and treated them separately as a sensitivity analysis. Thus, separate analyses were done for the two outcomes: Parkinson's disease and Alzheimer's disease and related dementias.

We obtained daily $PM_{2·5}$ predictions at a 1 km² spatial resolution across the contiguous United States from a well validated ensemble model.[16] The model included more than 100 predictor variables from satellite data, land-use and meteorological variables, and chemical transport model simulations. The model was calibrated with daily $PM_{2·5}$ concentrations measured at 2156 monitors—data obtained from the US Environmental Protection Agency's Air Quality System database and IMPROVE monitoring network—and had excellent performance (10-fold cross-validated r^2 of 0·86 for $PM_{2·5}$ predictions across the USA, ranging from 0·77 in the mountainous USA to 0·92 in the eastern Midwestern USA). The technical details, including information on the model validation, have been previously published.[16] Using daily $PM_{2·5}$ predictions at 1 km² grid cells, we calculated the daily mean $PM_{2·5}$ concentration for each postal code, by averaging the predictions at the grid cells whose centroids fell within the boundary of that postal code. Based on these results, we estimated annual postal code means and assigned the postal code-wide annual $PM_{2·5}$ concentration means to Medicare enrollees according to the postal code of residence and calendar year. In the USA, the mean population per postal code is around 7500. Each postal code can cover a small area in cities but can be larger in rural areas. The median land area of a postal code is around 92 km².

Statistical analysis

We collected neighbourhood-level socioeconomic status variables, available at county level or postal code tabulation areas level, which have both been associated with ambient air pollution and implicated in neurological health.[17,18] These variables were derived from the 2000–16 Behavioral Risk Factor Surveillance System, the 2000 and 2010 US Census, and the American Community Survey for each year from 2005 to 2016 (appendix pp 3–4). Region was classified as northeast, southeast, midwest, southwest, and west.

Given the very large dataset, we applied a Cox-equivalent re-parameterised Poisson approach for each of the two outcomes, coupled with parallel computing, to address the computational challenges (eg, inadequate memory size and lengthy computational time) faced by the conventional Cox proportional hazards model. Specifically, we proposed and fit a stratified quasi-Poisson model to estimate associations between the rate of first hospital admissions with neurological-related diagnosis codes (Parkinson's disease or Alzheimer's disease and related dementias) and time-varying annual mean $PM_{2·5}$ concentrations. The dependent variable was the count of outcome-related hospital admissions in each follow-up year, calendar year, and postal code location within strata specified by individual characteristics, using the corresponding total person-time of Medicare-fee-for-service beneficiaries as the offset. By stratifying on individual characteristics—ie, sex, race, Medicaid eligibility, and age at study entry in 2-year categories—we allowed for flexible strata-specific baseline rates. Mathematically, this stratified Poisson model is equivalent to a time-varying Cox proportional hazard model under an Anderson-Gill representation (appendix pp 4–5).[19] Importantly, the Cox-equivalent Poisson models also allowed use of parallel computing techniques that are not available for Cox models, further reducing the computation time. To account for within postal code correlated observations across years, we applied an m-out-n bootstrap method using postal code units to calculate statistically robust CIs.[20]

To adjust for potential confounding, we also included neighbourhood-level socioeconomic status factors in our analyses. To account for potential residual confounding by spatial and temporal trends, we included indicator variables for region and calendar years. We also estimated effects at low concentrations of $PM_{2·5}$, by restricting analyses to the subset of the cohort with annual exposures always lower than the current national standards

See Online for appendix

For the **Environmental Protection Agency's air quality data** see https://www.epa.gov/outdoor-air-quality-data

(ie, 12 μg/m³) over the study period (low-exposure analysis). Finally, to evaluate any potential deviations from linearity in the concentration–response curves, we included penalised splines for the $PM_{2·5}$ term in the models.

	Full cohort (n=63 038 019)	Low-exposure cohort (n=21 928 573)
Age at entry, years		
65–74	48 784 857 (77·4%)	17 010 757 (77·6%)
75–84	10 550 039 (16·7%)	3 673 343 (16·8%)
85–94	3 327 268 (5·3%)	1 134 507 (5·2%)
95–104	375 708 (0·6%)	109 934 (0·5%)
105–114	147 (<0·1%)	32 (<0·1%)
Mean (SD)	69·9 (7·2)	69·8 (7·1)
Sex		
Men	28 295 987 (44·9%)	10 084 588 (46·0%)
Women	34 742 032 (55·1%)	11 843 985 (54·0%)
Race		
White	53 229 370 (84·4%)	19 776 603 (90·2%)
Black	5 513 530 (8·7%)	663 313 (3·0%)
Other*	4 295 119 (6·8%)	1 488 657 (6·8%)
Medicaid eligibility		
Eligible	7 853 739 (12·5%)	2 405 354 (11·0%)
Ineligible	55 184 280 (87·5%)	19 523 219 (89·0%)
$PM_{2·5}$ concentration, μg/m³	9·7 (3·2)	7·2 (2·3)
Body-mass index, kg/m²	27·5 (1·1)	27·3 (1·0)
Ever smoked, %	47·1 (7·7)	48·1 (7·8)
Hispanic, %	9·2 (16·7)	9·2 (16·3)
Black, %	9·1 (17·3)	2·7 (7·5)
Median household income, US$1000	48·0 (21·7)	47·5 (18·9)
Median home value, $1000	159·0 (141·9)	153·9 (131·8)
Below poverty level, %	11·0 (10·9)	9·7 (10·2)
Not graduated from high school, %	28·7 (18·8)	24·2 (17·1)
Owner-occupied housing, %	71·1 (18·8)	75·2 (14·8)
Population density, people per mile²	1601·2 (5233·1)	595·1 (1595·8)

Data are n (%) or mean (SD). *Other included Asian, Hispanic, American Indian or Alaskan Native, and unknown.

Table 1: **Cohort characteristics**

	Parkinson's disease	Alzheimer's disease and related dementias
Main analyses		
Number of admissions	1 033 669	3 425 102
Total person-years	478 335 593	473 696 618
Median follow-up year	7	7
HR per 5 μg/m³ $PM_{2·5}$	1·13 (1·12–1·14)	1·13 (1·12–1·14)
Low-exposure analyses (<12 μg/m³)		
Number of admissions	301 227	939 035
Total person-years	156 287 478	155 139 930
Median follow-up year	6	6
HR per 5 μg/m³ $PM_{2·5}$	1·14 (1·12–1·16)	1·18 (1·15–1·21)

Data are n or HR (95% CI). HR=hazard ratio.

Table 2: **Cause-specific admissions for Parkinson's disease and Alzheimer's disease and related dementias, 2000–16**

To identify subpopulations who might be particularly susceptible, we assessed potential effect modification by sex (men *vs* women), race (white people *vs* Black people *vs* other [Asian, Hispanic, American Indian or Alaskan Native, and unknown]), age (≥80 years *vs* <80 years), Medicaid eligibility (dual *vs* non-dual eligibility) as a surrogate for individual-level socioeconomic status, and urbanicity (quartiles of population density), by including interaction terms between these potential modifiers and $PM_{2·5}$. Specifically, we calculated the effect of $PM_{2·5}$ in each category of the effect modifier and assessed significance of the interaction term. We chose the age of 80 years as a cutoff to distinguish the young and middle-old from the old-old.[21]

We did a series of sensitivity analyses to assess the robustness of our results to confounding, inclusion of prevalent cases, potential outcome misclassification, and exposure time window (appendix pp 5–8). Given that these neurodegenerative diseases are age-dependent, as additional sensitivity analysis we also considered stratification by age at entry using 1-year intervals. To remove potentially prevalent cases, we ran additional analyses excluding anyone who had a first admission for these outcomes in their first 2 years of follow-up and repeated our analyses. As information in Medicare is only available after beneficiaries turn 65 years old, it is possible that some study participants had a Parkinson's disease or Alzheimer's disease and related dementias hospital admission before enrolling to Medicare. This sensitivity analysis—excluding cases with an admission during their first 2 years of enrolment—increases the probability that we are capturing the first admission with a related code. To evaluate whether the associations we observed can be attributed to a different outcome also linked to air pollution, we excluded the subset of Parkinson's disease and Alzheimer's disease and related dementias cases with the most frequent category of primary discharge codes (ie, circulatory system disease [ICD-9: 390–459; ICD-10: I00–I99]) from analyses. The primary discharge code appeared in 392 588 (41·1%) cases of Parkinson's disease and 1 323 044 (45·3%) cases of Alzheimer's disease and related dementias. Additionally, we added a sensitivity analysis restricting cases only to those with primary diagnoses codes for Parkinson's disease or Alzheimer's disease and related dementias. Finally, we considered an alternative exposure window with 1-year lag period (ie, using the annual mean exposure during the year preceding the outcome). Considering that chemical composition of $PM_{2·5}$ mass (and thus relative toxicity) can vary markedly among different regions in the USA, we also did a subgroup analysis by region.

The computations of the analyses of this study were done on the Research Computing Environment, which is supported by the Institute for Quantitative Social Science at Harvard University. We used R software, version 3.3.2 for all analyses.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Data were analysed for Medicare beneficiaries who were 65 years and older between Jan 1, 2000, and Dec 31, 2016. The full cohort included 63 038 019 individuals living in 39 065 postal codes (table 1). The mean age at entry was 69·9 years (SD 7·2). There were 478·3 million person-years of follow-up for Parkinson's disease and 473·7 million for Alzheimer's disease and related dementias (table 2). The total number of first admissions was 1·0 million for Parkinson's disease and 3·4 million for Alzheimer's disease and related dementias. The median follow-up was 7 years (IQR 8). Of the Parkinson's disease cases, 77 016 (7·5%) of 1 033 669 had Parkinson's disease as the primary discharge diagnosis code and, of the Alzheimer's disease and related dementias cases, 502 565 (14·7%) of 3 425 102 did. For the cases identified with secondary diagnoses of these conditions, we examined the distribution of primary diagnostic codes and found that the primary conditions were predominantly circulatory system diseases (appendix p 8).

The low-exposure cohort subset included 21 928 573 individuals living in 15 775 postal codes, with a mean entry age of 69·8 years (SD 7·1). For Parkinson's disease, the total person-years of follow-up was 156·3 million and for Alzheimer's disease and related dementias, it was 155·1 million. The number of first admissions was 0·3 million for Parkinson's disease and 0·9 million for Alzheimer's disease and related dementias (table 2).

The mean annual $PM_{2·5}$ concentration over the study period was 9·7 μg/m³ (SD 3·2, IQR 4·3, 5th to 95th percentile 5·2–15·9; figure 1A). $PM_{2·5}$ concentrations were generally higher in eastern USA than in western USA (except California). Figures 1B and 1C present the occurrence of first hospital admissions with a Parkinson's disease or an Alzheimer's disease and related dementias diagnosis code, per 100 000 Medicare beneficiaries across the contiguous United States (2000–16).

Overall, long-term exposure to $PM_{2·5}$ was significantly positively associated with both neurodegenerative outcomes in both the entire cohort and the low-exposure subset. Specifically, in the entire cohort we observed a hazard ratio (HR) of 1·13 (95% CI 1·12–1·14) for Parkinson's disease admissions and an HR of 1·13 (1·12–1·14) for Alzheimer's disease and related dementias admissions per 5 μg/m³ increase in annual $PM_{2·5}$ concentrations. In the low-exposure subset, we found a slightly elevated association (HR 1·14, 95% CI 1·12–1·16) for Parkinson's disease and an elevated association (HR 1·18, 1·15–1·21) for Alzheimer's disease and related dementias admissions per 5 μg/m³ $PM_{2·5}$ increase (table 2).

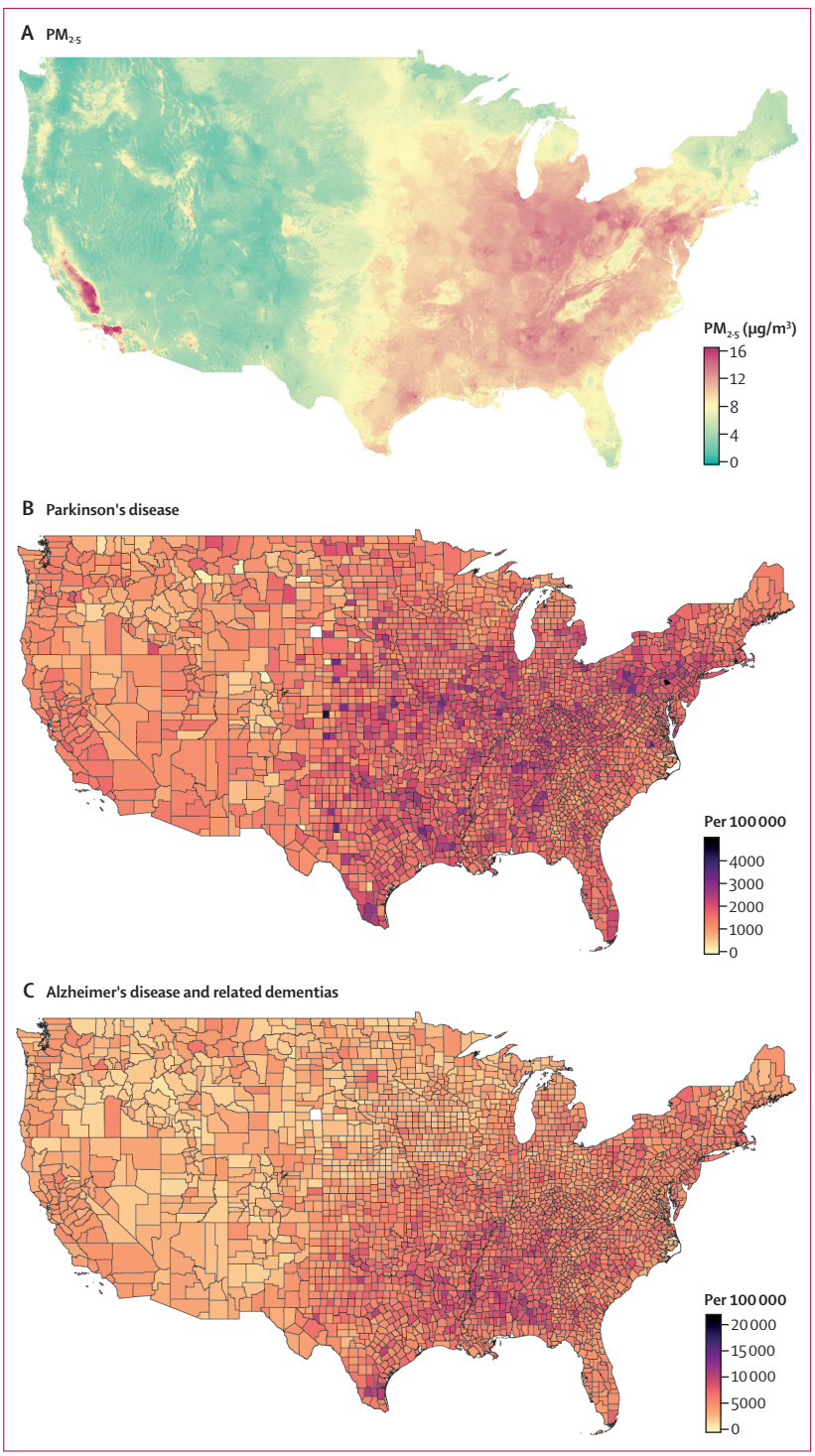


Figure 1: **Nationwide concentrations of $PM_{2·5}$, and occurrences of Parkinson's disease and Alzheimer's disease and related dementias across the contiguous United States**

(A) 17-year mean of annual $PM_{2·5}$ concentrations (μg/m³). (B) Occurrence of first Parkinson's disease hospital admissions per 100 000 Medicare beneficiaries. (C) Occurrence of first Alzheimer's disease and related dementias hospital admissions per 100 000 Medicare beneficiaries (2000–16).

Figure 2 shows the concentration–response relationships for Parkinson's disease and Alzheimer's disease and related dementias. We observed a strong linear relationship for annual mean $PM_{2·5}$ concentrations less than 16 μg/m³, followed by a plateaued association with increasingly larger confidence bands for both outcomes. However, less than 5% of the distribution of the $PM_{2·5}$ concentrations were greater than 16 μg/m³.

Among the effect modifiers, we found $PM_{2·5}$ effect estimates that were significantly larger in magnitude among individuals in more urban areas versus those in less urban areas (as expressed in quartiles of population density). We also observed higher HRs among those who identified as white than those who identified as Black or Asian, Hispanic, American Indian or Alaskan Native, and unknown, and for women compared with men (figure 3).

For both Parkinson's disease and Alzheimer's disease and related dementias, all sensitivity analyses yielded similar results to the main analyses (appendix pp 5–8). When excluding potentially prevalent cases (ie, excluding those who had a first admission in the first 2 years of follow-up), both effect estimates were slightly elevated. The sensitivity analysis in which Alzheimer's disease and dementia were treated as separate outcomes also yielded significant and positive associations between $PM_{2·5}$ and the two separate outcomes of interest. However, the effect estimates for Alzheimer's disease (HR 1·17, 95% CI 1·16–1·18) were higher than those for dementia (HR 1·06, 1·05–1·07). Our results were robust to confounding adjustment—ie, the results were almost unchanged when we excluded different sets of covariates in alternative models compared with the main one. Additionally, both exclusion of all cases identified through secondary diagnostic codes and exclusion of those secondary diagnostic cases with circulatory system disease as the primary diagnosis code did not change the main results. Finally, our results were robust to the use of a different exposure window. The 1-year lagged exposure analysis (eg, using annual mean $PM_{2·5}$ in 2005 to link the outcome in 2006) yielded results nearly identical to the findings from our main analysis.

All region-specific results consistently suggested a link between $PM_{2·5}$ and first Parkinson's disease and Alzheimer's disease and related dementias hospital admissions, although effect estimates varied by geographical region. In summary, we observed the highest HR for first Parkinson's disease hospital admission among Medicare enrollees in the northeastern USA and for first Alzheimer's disease and related dementias hospital admissions in the midwestern USA.

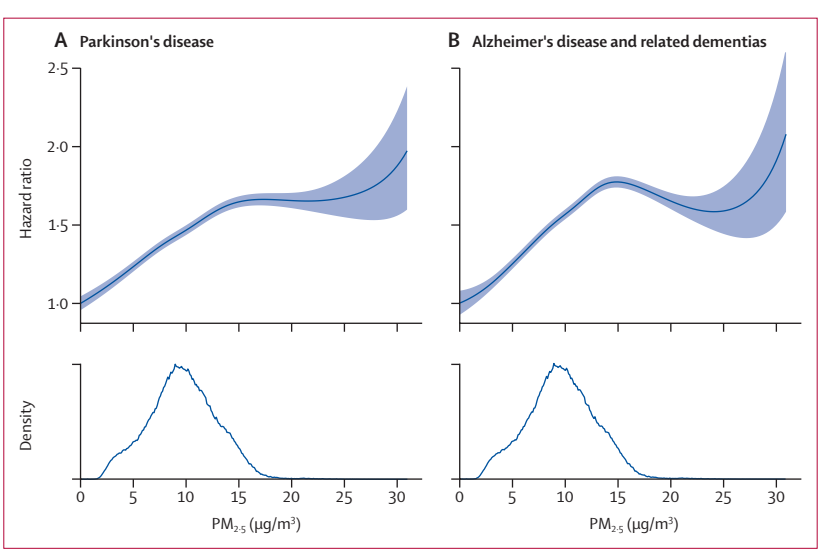


Figure 2: **Concentration–response curves of the association between long-term $PM_{2·5}$ exposure and neurological disorders**
Parkinson's disease (A) and Alzheimer's disease and related dementias (B).

Discussion

In this large, nationwide prospective cohort of all Medicare-fee-for-service beneficiaries, long-term exposure to $PM_{2·5}$, an indicator for the air pollution mixture at each postal code, was associated with an increased risk of first hospital admission with a Parkinson's disease or an Alzheimer's disease and related dementias diagnosis code, even at concentrations less than the current annual national standards (12 μg/m³). We also identified women, white people, and more urbanised populations as particularly susceptible subgroups. These findings suggest that improving air quality, with $PM_{2·5}$ concentrations even lower than current national standards, could yield public health benefits.

The shape of the concentration–response relationship between air pollution and neurodegeneration has rarely been assessed in the literature. Only one previous study simply assessed non-linearity using quartiles and found no evidence of deviation.[9] This result was in agreement with our results, had we used quartiles. Use of splines allowed for a more detailed characterisation of the shape across the $PM_{2·5}$ concentration range. Risk of first hospital admission with a Parkinson's disease or an Alzheimer's disease and related dementias diagnosis code, as a proxy for neurodegeneration progression, linearly increased with increasing $PM_{2·5}$ concentrations less than the current annual standards (12 μg/m³), suggesting no safe threshold for harmful pollution. Although we detected some deviations from linearity at concentrations greater than 16 μg/m³, less than 5% of the observations were higher than that. It is possible that any deviation at such high concentrations could indicate that the flexible penalised splines are sensitive to potential outlying observations with high leverage.

Our findings regarding associations between $PM_{2·5}$ and Alzheimer's disease and related dementias are consistent with previous research, both in terms of direction and magnitude; of these, one was done in Ontario's Canadian population,[8] and the other two were done in regional subpopulations of US Medicare enrollees.[7,9] Mixed results,

including both positive and null findings, however, were reported for the association between $PM_{2·5}$ and Parkinson's disease in the literature.[9,22,23] It is worth noting that a comprehensive city-level study in 50 northeastern US cities among Medicare enrollees found higher estimates in magnitude for Parkinson's disease and Alzheimer's disease and related dementias than the ones estimated in this study,[9] which matches our finding of significantly higher $PM_{2·5}$ effects among urban dwellers. Other studies also found similar results in the urban populations they investigated.[7,8] The observed associations for the other grouping within race are not clear and more work is needed to understand these results. We note, however, that the percentage of the population aged older than 65 years in the USA who are not white or Black is 6·8%.

Both examined diseases have long insidious onsets and the exact timing of disease onset is not known.[24,25] Furthermore, disease diagnosis probably occurs at a neurologist's office and not during a hospital admission. Therefore, use of an administrative dataset does not allow investigation of the association between $PM_{2·5}$ and onset of these outcomes. That is, with our analysis we cannot examine true onset incidence or incidence of diagnosis. Our analysis estimates incidence of first hospital admission, which can be interpreted as increased susceptibility to hospital admissions among this patient population and accelerated disease progression. In support of our hypothesis and main findings, the sensitivity analysis excluding people that had a first admission in the first 2 years of the cohort (ie, potentially prevalent cases) resulted in larger in magnitude effect estimates.

In our main analysis, 956 653 (92·5%) of the Parkinson's disease cases and 2 922 537 (85·3%) of the Alzheimer's disease and related dementias cases that we identified were based on secondary causes and had a different primary cause of admission. Although this observation was expected, as these outcomes are monitored by neurologists and do not necessarily show up as primary hospital admissions, we were concerned that the observed effect could reflect the signal with the most common primary diagnosis code for these outcomes. Exclusion, however, of Parkinson's disease and Alzheimer's disease and related dementias cases with a primary diagnosis for circulatory system disease did not change our results.

Toxicological studies suggest various potential mechanisms via which air pollution might contribute to neurodegenerative progression. Systemic and brain inflammation, for example, enhance the pathogenic alteration of α-synuclein, accelerating the progression of Parkinson's disease[14] and Alzheimer's disease.[13] Oxidative stress, in addition, is also involved both in initiation and progression, and plays an important role in accelerating Parkinson's disease progression.[26] Air pollution might play a key role in neuroinflammation and further exacerbate or initiate dysfunctional protein handling, in the context of amyloid plaques, tau hyperphosphorylation,

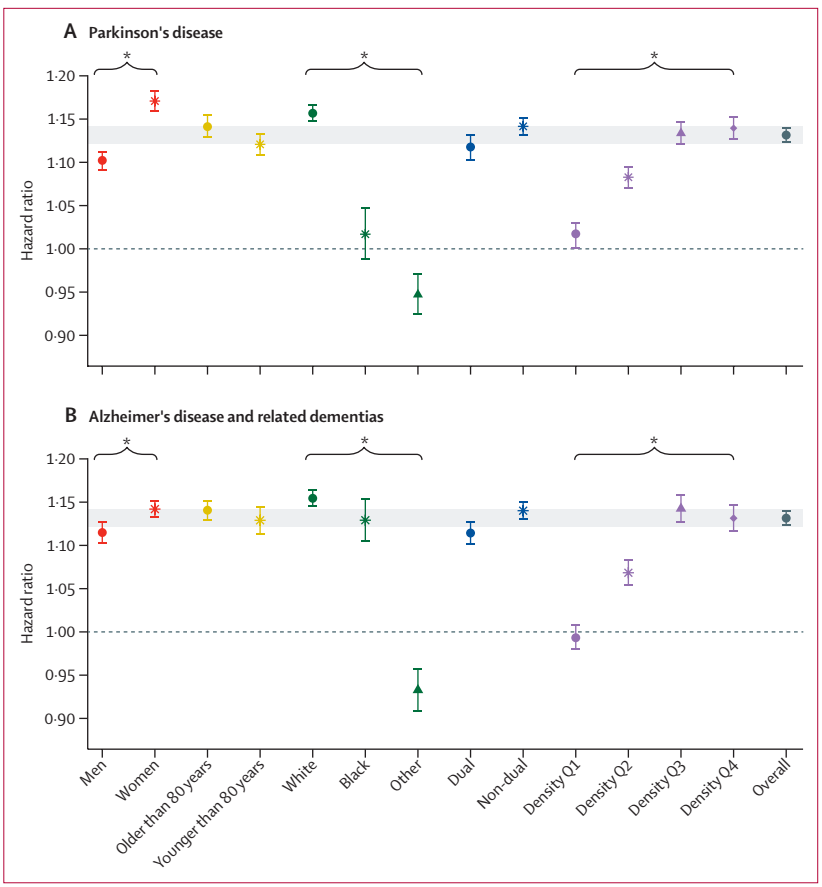


Figure 3: **Identification of vulnerable subpopulations**
Hazard ratios for Parkinson's disease (A) and Alzheimer's disease and related dementias (B) associated with a 5 μg/m³ increase in $PM_{2·5}$ concentrations by study subgroups. The shading represents the estimated main effects for the overall population. Dual or non-dual refers to eligibility for Medicaid. Density Q1–Q4 denote quartiles of population density—ie, low population density, low to medium population density, medium to high population density, and high population density. Other included Asian, Hispanic, American Indian or Alaskan Native, and unknown race. *Significant modification (at α=0·05 level).

and neurofibrillary tangles.[27] Several air pollutants, including $PM_{2·5}$ and ultrafine (<0·1 μm) particulate matter, have been shown to easily cross the blood–brain barrier, providing an important route for air pollutants to interact with the CNS. Indeed, increases in air pollution can elicit increases in the inflammatory response in the prefrontal lobes, with concomitant increases in oxidative damage and amyloid β deposition. The origin of these pathological markers could arise from the direct interaction of air pollutants with microglia in the brain, resulting in a release of pro-inflammatory signals that further facilitate neuronal damage and protein aggregation. Elevation in pro-inflammatory signals can mediate dysfunctional protein handling, in the form of elevations in amyloid β and hyperphosphorylation of tau.[27] Given the pathogenesis of Alzheimer's disease and other neurogenerative diseases that are defined by neuroinflammation, oxidative damage, and protein misfolding, exposure to air pollution might serve as an important risk factor in the development and progression of

Alzheimer's disease and Parkinson's disease pathology and concomitant neurobehavioural deficits. For all neurological outcomes, we observed significantly higher effects of $PM_{2·5}$ among individuals in urban areas versus rural areas. One possible reason might be the abundance of metal-bearing nanoparticles in the urban atmosphere, which have very small diameters and can access the brain directly.[28] The higher estimates among white people and women could be attributed to a longer life expectancy in these groups—ie, the chance of competing risks among non-white individuals or men is greater, including the probability of death before developing Parkinson's disease or Alzheimer's disease and related dementias.[29]

Our study data and methods have several advantages. First, our study population of all Medicare-fee-for-service beneficiaries in the USA gives us ample power to detect effects. This statistical power is particularly useful in environmental studies in which exposures are highly prevalent but effect estimate sizes are often small. Second, our study assessed the whole of the USA, which has greater generalisability than previous smaller cohort studies that were geographically restricted, although our study might not be generalisable to other countries. Furthermore, the aggregation of data into strata of shared individual characteristics not only allowed us to create a more efficient model but also allowed us to analyse a very large dataset with a far smaller computational burden. Given the increase in the use of very large datasets, this novel analytical approach might be useful in other research as well.

Our findings, however, should be interpreted in light of some potential limitations. First, reliance on an administrative cohort did not allow us to examine the relationship between $PM_{2·5}$ and disease onset. Parkinson's disease and Alzheimer's disease and related dementias are diseases that do not require hospital admission for diagnosis and treatment; usually, hospital admission occurs at more advanced stages of the disease for treating complications or for adjusting the therapeutic plan. Thus, the hospital admission records cannot represent disease incidence and we probably underestimate the case number when using first hospital admission as a proxy for neurodegeneration. In addition, a positive predictive value of 0·65 for Parkinson's disease[30] and about 0·75 for Alzheimer's disease and related dementias[31] has been reported when Medicare claims were used, indicating the under-diagnosed nature of neurological conditions using claims records. Furthermore, our results only represent the Medicare-fee-for-service population, which does not include all Medicare beneficiaries. Specifically, earlier in our study period (eg, in 2003), the Medicare-fee-for-service population covered up to 29 230 838 (84·9%) of 34 423 305 Medicare beneficiaries, while in 2016 it was 30 974 063 (65·8%) of 47 099 370 Medicare beneficiaries. It is possible that Medicare-fee-for-service beneficiaries switched to Medicare-HMO (Medicare managed care plan) and back, potentially resulting in some missed cases in our data, as we have no information on Medicare-HMO claims records. Our findings, thus, might not be generalisable to the entire Medicare population. Second, the use of predicted concentrations for exposure assessment might have resulted in some exposure measurement error. However, the prediction model we used is considered to have excellent predictive accuracy,[16] substantially reducing potential exposure measurement error. In our study, exposure measurement error is likely to be non-differential because the error in the predicted ambient $PM_{2·5}$ concentrations is probably independent of outcome status. Thus, any resulting bias would be towards the null.[32] Third, we cannot exclude the possibility of potential residual confounding bias. We did, however, adjust all our models for multiple neighbourhood-level socioeconomic status variables, and thus any potential residual bias is expected to be very small. Individual-level risk factors for neurological disorders, such as smoking, are not available in Medicare. However, we used postal code mean predicted $PM_{2·5}$ to assign exposures, which could only covary with individual-level factors through postal code-level socioeconomic status,[33] for which we carefully adjusted, thus effectively minimising this potential source of bias. Fourth, our ensemble model predicts total $PM_{2·5}$ mass concentrations, but not all particles have the same toxicity; some studies have shown that traffic-related pollution might be particularly toxic.[34] Future studies should aim to disentangle specific effects of regional versus local particles.

In conclusion, our study provides strong epidemiological evidence that long-term exposure to air pollution is significantly associated with a higher risk of neurological health deterioration, even at concentrations less than the current national standards. Our findings suggest that policies that result in further reductions in ambient $PM_{2·5}$ concentrations can yield substantial health benefits in the ageing US population, even for those already exposed to low $PM_{2·5}$ concentrations.

Contributors

AZ and M-AK designed the research and directed its implementation. MDY, DB, YAA, YWe, YWa, PL, QD, JS, and FD prepared datasets. LS and XW analysed data. LS, XW, and PL made the figures. LS, XW, M-AK, and AZ wrote the paper, and all authors contributed to the revision of the manuscript.

Declaration of interests

We declare no competing interests.

Acknowledgments

We thank Benjamin Sabath for the support with the Research Computing Environment and William Michael Caudle for fruitful discussion. This study was supported by the Health Effects Institute (4953-RFA14-3/16-4), the National Institute of Environmental Health Sciences (NIEHS R01 ES024332, R01 ES028805, R21 ES028472, P30 ES009089, P30 ES000002), the National Institute on Aging (NIA/NIH R01 AG066793-01, P50 AG025688), and the HERCULES Center (P30ES019776). Research described in this Article was done under contract to the Health Effects Institute, an organisation jointly funded by the US Environmental Protection Agency (assistance award number R-83467701) and some motor vehicle and engine manufacturers. The contents of this Article do not necessarily reflect the views of the Health Effects Institute, or its sponsors, nor do they necessarily reflect

the views and policies of the US Environmental Protection Agency or motor vehicle and engine manufacturers. The computations in this paper were run on the Research Computing Environment supported by the Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University.

References

1 GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019; **18:** 459–80.

2 Maragakis NJ, Rothstein JD. Mechanisms of disease: astrocytes in neurodegenerative disease. *Nat Clin Pract Neurol* 2006; **2:** 679–89.

3 Ailshire J, Karraker A, Clarke P. Neighborhood social stressors, fine particulate matter air pollution, and cognitive function among older U.S. adults. *Soc Sci Med* 2017; **172:** 56–63.

4 Tzivian L, Dlugaj M, Winkler A, et al. Long-term air pollution and traffic noise exposures and mild cognitive impairment in older adults: a cross-sectional analysis of the Heinz Nixdorf recall study. *Environ Health Perspect* 2016; **124:** 1361–68.

5 Cacciottolo M, Wang X, Driscoll I, et al. Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models. *Transl Psychiatry* 2017; **7:** e1022.

6 Weuve J, Puett RC, Schwartz J, Yanosky JD, Laden F, Grodstein F. Exposure to particulate air pollution and cognitive decline in older women. *Arch Intern Med* 2012; **172:** 219–27.

7 Lee M, Schwartz J, Wang Y, Dominici F, Zanobetti A. Long-term effect of fine particulate matter on hospitalization with dementia. *Environ Pollut* 2019; **254:** 112926.

8 Chen H, Kwong JC, Copes R, et al. Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: a population-based cohort study. *Lancet* 2017; **389:** 718–26.

9 Kioumourtzoglou M-A, Schwartz JD, Weisskopf MG, et al. Long-term $PM_{2.5}$ exposure and neurological hospital admissions in the northeastern United States. *Environ Health Perspect* 2016; **124:** 23–29.

10 Bondy SC. Anthropogenic pollutants may increase the incidence of neurodegenerative disease in an aging population. *Toxicology* 2016; **341–343:** 41–46.

11 Calderón-Garcidueñas L, Solt AC, Henríquez-Roldán C, et al. Long-term air pollution exposure is associated with neuroinflammation, an altered innate immune response, disruption of the blood-brain barrier, ultrafine particulate deposition, and accumulation of amyloid β-42 and α-synuclein in children and young adults. *Toxicol Pathol* 2008; **36:** 289–310.

12 Block ML, Elder A, Auten RL, et al. The outdoor air pollution and brain health workshop. *Neurotoxicology* 2012; **33:** 972–84.

13 Bhatt DP, Puig KL, Gorr MW, Wold LE, Combs CK. A pilot study to assess effects of long-term inhalation of airborne particulate matter on early Alzheimer-like changes in the mouse brain. *PLoS One* 2015; **10:** e0127102.

14 Gao H-M, Zhang F, Zhou H, Kam W, Wilson B, Hong J-S. Neuroinflammation and α-synuclein dysfunction potentiate each other, driving chronic progression of neurodegeneration in a mouse model of Parkinson's disease. *Environ Health Perspect* 2011; **119:** 807–14.

15 Mosconi L, Tsui WH, Herholz K, et al. Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer's disease, and other dementias. *J Nucl Med* 2008; **49:** 390–98.

16 Di Q, Amini H, Shi L, et al. An ensemble-based model of $PM_{2.5}$ concentration across the contiguous United States with high spatiotemporal resolution. *Environ Int* 2019; **130:** 104909.

17 Baumgart M, Snyder HM, Carrillo MC, Fazio S, Kim H, Johns H. Summary of the evidence on modifiable risk factors for cognitive decline and dementia: a population-based perspective. *Alzheimers Dement* 2015; **11:** 718–26.

18 Wirdefeldt K, Adami H-O, Cole P, Trichopoulos D, Mandel J. Epidemiology and etiology of Parkinson's disease: a review of the evidence. *Eur J Epidemiol* 2011; **26** (suppl 1)**:** S1–58.

19 Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Stat* 1982; **10:** 1100–20.

20 Bickel PJ, Götze F, van Zwet WR. Resampling fewer than n observations: gains, losses, and remedies for losses. *Statistica Sinica* 1997; **7:** 1.

21 Forman DE, Berman AD, McCabe CH, Baim DS, Wei JY. PTCA in the elderly: the "young-old" versus the "old-old". *J Am Geriatr Soc* 1992; **40:** 19–22.

22 Liu R, Young MT, Chen J-C, Kaufman JD, Chen H. Ambient air pollution exposures and risk of Parkinson disease. *Environ Health Perspect* 2016; **124:** 1759–65.

23 Kirrane EF, Bowman C, Davis JA, et al. Associations of ozone and $PM_{2.5}$ concentrations with Parkinson's disease among participants in the Agricultural Health Study. *J Occup Environ Med* 2015; **57:** 509–17.

24 Liu S-Y, Chan P, Stoessl AJ. The underlying mechanism of prodromal PD: insights from the parasympathetic nervous system and the olfactory system. *Transl Neurodegener* 2017; **6:** 4.

25 Tschanz JT, Corcoran CD, Schwartz S, et al. Progression of cognitive, functional, and neuropsychiatric symptom domains in a population cohort with Alzheimer dementia: the Cache County Dementia Progression study. *Am J Geriatr Psychiatry* 2011; **19:** 532–42.

26 Zhou C, Huang Y, Przedborski S. Oxidative stress in Parkinson's disease: a mechanism of pathogenic and therapeutic significance. *Ann N Y Acad Sci* 2008; **1147:** 93–104.

27 Leyns CEG, Holtzman DM. Glial contributions to neurodegeneration in tauopathies. *Mol Neurodegener* 2017; **12:** 50.

28 Gonet T, Maher BA. Airborne, vehicle-derived fe-bearing nanoparticles in the urban environment: a review. *Environ Sci Technol* 2019; **53:** 9970–91.

29 Shrestha LB. Life expectancy in the United States. Congressional Information Service, Library of Congress, 2005. https://www.everycrsreport.com/files/20050303_RL32792_ca5374f4af13937c82f8749ffee41be74a3f5909.pdf (accessed Oct 2, 2020).

30 Noyes K, Liu H, Holloway R, Dick AW. Accuracy of Medicare claims data in identifying Parkinsonism cases: comparison with the Medicare current beneficiary survey. *Mov Disord* 2007; **22:** 509–14.

31 Taylor DH Jr, Østbye T, Langa KM, Weir D, Plassman BL. The accuracy of Medicare claims as an epidemiological tool: the case of dementia revisited. *J Alzheimers Dis* 2009; **17:** 807–15.

32 Kioumourtzoglou M-A, Spiegelman D, Szpiro AA, et al. Exposure measurement error in $PM_{2.5}$ health effects studies: a pooled analysis of eight personal exposure validation studies. *Environ Health* 2014; **13:** 2.

33 Weisskopf MG, Webster TF. Trade-offs of personal vs. more proxy exposure measures in environmental epidemiology. *Epidemiology* 2017; **28:** 635–43.

34 Turner MC, Jerrett M, Pope CA 3rd, et al. Long-term ozone exposure and mortality in a large prospective study. *Am J Respir Crit Care Med* 2016; **193:** 1134–42.