

Exposure to air pollution and COVID-19 mortality in the United States

Xiao Wu MS, Rachel C. Nethery PhD, M. Benjamin Sabath MA, Danielle Braun PhD, Francesca

Dominici PhD

All authors are part of the Department of Biostatistics, Harvard T.H. Chan School of Public

Health, Boston, MA, 02115, USA

Lead authors: Xiao Wu and Rachel C. Nethery

Corresponding and senior author: Francesca Dominici, PhD
Clarence James Gamble Professor of Biostatistics, Population and Data Science
Harvard T.H. Chan School of Public Health
Co-Director Harvard Data Science Initiative
677 Huntington Avenue Boston, MA 02115
410.258.5886
Email: fdominic@hsph.harvard.edu

Esposizione all'inquinamento atmosferico e mortalità da Covid-19 negli Stati Uniti

Xiao Wu MS, Rachel C. Nethery PhD, M. Benjamin Sabath MA, Danielle Braun PhD, Francesca Dominici PhD

Tutti gli autori fanno parte del Dipartimento di Biostatistica, Harvard T.H. Chan School of Public Health, Boston MA 02115 Stati Uniti

Aggiornato al 5 aprile 2020

Sintesi

Premessa: gli esperti del governo statunitense stimano che il Covid-19 potrebbe provocare la morte di un numero di americani compreso tra 100.000 e 240.000. La maggior parte delle condizioni pregresse che aumentano il rischio di morte da Covid-19 sono le stesse patologie legate a un'esposizione prolungata all'inquinamento atmosferico. Abbiamo condotto ricerche per

verificare se l'esposizione prolungata media al particolato fine (PM_{2,5}) incrementa il rischio di morte da Covid-19 negli Stati Uniti.

Metodi: Sono stati raccolti dati relativamente a circa 3.000 contee statunitensi, pari al 98% della popolazione, fino al 4 aprile 2020 compreso. Abbiamo adattato modelli misti binomiali negativi "a inflazione di zeri" (zero-inflated) considerando i decessi da Covid-19 a livello di contea come effetto e il valore medio di esposizione prolungata a PM_{2,5} a livello di contea come esposizione prescelta. Abbiamo effettuato aggiustamenti rispetto a numero di abitanti, numero di posti letto negli ospedali, numero di test individuali effettuati, variabili comportamentali, socioeconomiche e meteorologiche tra cui figurano, ma non solo, obesità e abitudine al fumo. Abbiamo incluso una intercetta casuale per singolo stato al fine di tener conto di potenziali correlazioni in contee all'interno del medesimo stato.

Risultati: Abbiamo riscontrato che un aumento anche solo di 1 µg/m³ dell'esposizione media prolungata a PM_{2,5} è associato a un incremento del 15% del tasso di mortalità da Covid-19, intervallo di confidenza (CI) 95% (5%, 25%). I risultati sono statisticamente significativi e si sono rivelati fondati anche a seguito di analisi secondarie e di sensibilità.

Conclusioni: un piccolo aumento dell'esposizione prolungata a PM_{2,5} induce un forte incremento del tasso di mortalità da Covid-19, una percentuale di 20 volte superiore a quella stimata per la mortalità generale legata a PM_{2,5}. I risultati del presente studio sottolineano anche l'importanza della continuità nell'applicazione delle norme in vigore sull'inquinamento atmosferico sia durante che dopo l'emergenza Covid-19. I dati e il codice sono di dominio pubblico.

Introduzione

L'emergenza sanitaria Covid-19 è un fenomeno che non ha eguali nella nostra epoca e che avrà delle serie conseguenze economiche e sociali. La repentinità e la portata globale della pandemia hanno sollevato problemi urgenti che richiedono la disponibilità di informazioni attendibili e coordinate per rallentarne la furia devastatrice. Un obiettivo di importanza critica per la sanità pubblica consiste nell'individuazione di fattori ambientali chiave modificabili come, per esempio, l'inquinamento dell'aria ambiente, che potrebbero aumentare la gravità degli effetti sulla salute (es.: ricovero in terapia intensiva e decesso) per le persone affette da Covid-19.

La determinazione delle cause della morte di pazienti che hanno contratto il Covid-19 è ancora in fase di evoluzione. I primi dati provenienti dalla Cina indicano che la maggior parte dei decessi provocati dal virus si sono registrati in adulti di età superiore a 60 anni e in soggetti affetti da gravi patologie pregresse (1,2). Le analisi preliminari del decorso di pazienti colpiti dal Covid-19 negli Stati Uniti concordano nel rilevare gli indici di letalità più elevati tra persone di età superiore a 65 anni (3).

In un report diffuso il mese scorso dalla missione congiunta della Organizzazione mondiale della sanità (4) si legge che un paziente su sette soffre di difficoltà respiratorie e di altre gravi complicazioni. Generalmente questi pazienti presentano insufficienza respiratoria e altri problemi ad apparati vitali. Il Covid-19 può provocare polmonite virale con ulteriori manifestazioni e complicazioni extrapolmonari tra cui la sindrome da distress respiratorio acuto (ARDS) che ha un tasso di mortalità compreso tra il 27% e il 45% (5). Tra i fattori associati alla mortalità figurano il sesso (maschile), l'età avanzata (oltre i 65 anni) e la presenza di patologie concomitanti quali ipertensione, diabete mellito, malattie cardiovascolari e patologie cerebrovascolari. Il Covid-19 è associato anche a un carico infiammatorio elevato che può indurre infiammazioni vascolari, miocarditi e aritmie cardiache (6).

Sebbene l'epidemiologia del Covid-19 sia ancora in fase di evoluzione, abbiamo riscontrato che esiste un'ampia sovrapposizione tra le cause del decesso di pazienti che hanno contratto il Covid-19 e patologie legate a un'esposizione prolungata al particolato fine (PM_{2,5}). Nel *Global Burden of Disease Study* l'inquinamento atmosferico viene indicato quale fattore di rischio per la mortalità totale e indotta da patologie cardiovascolari e ritenuto all'origine di 5,5 milioni di morti premature all'anno su scala mondiale (7). Il PM_{2,5} contiene particelle solide microscopiche o goccioline liquide talmente piccole da poter essere inalate provocando gravi problemi di salute. Il 26 marzo 2020 l'Agenzia per la protezione ambientale (EPA) statunitense ha annunciato un allentamento generalizzato delle norme ambientali in risposta alla pandemia da coronavirus che consente a centrali elettriche, fabbriche e altre strutture di determinare autonomamente la propria conformità ai livelli di legge nel comunicare i dati relativi a inquinamento atmosferico e delle acque.

Noi ipotizziamo che, poiché l'esposizione prolungata a PM_{2,5} ha effetti negativi sugli apparati respiratorio e circolatorio, essa possa anche aggravare i sintomi dell'infezione da Covid-19 e aumentare il rischio di morte per i pazienti che hanno contratto il virus. E' ben noto negli Stati Uniti e a livello mondiale il legame tra PM_{2,5} e stato di salute, compresi effetti in gravidanza, patologie respiratorie, cardiovascolari e neurocognitive (8,9,10,11,12,13).

Da uno studio recente effettuato dal nostro gruppo è anche emersa un'associazione statisticamente significativa tra l'esposizione prolungata a PM_{2,5} e ozono e il rischio di ARDS tra la popolazione più anziana negli Stati Uniti (14). Numerosi studi scientifici presi in esame dalla EPA statunitense hanno collegato il PM_{2,5} a tutta una serie di condizioni patologiche tra cui figurano: morte prematura di persone affette da disturbi cardiaci o polmonari, infarti non letali, battito irregolare, asma aggravata, riduzione della funzione polmonare e incremento di sintomi respiratori quali infiammazioni, irritazioni delle vie aeree, tosse e difficoltà respiratorie (15).

Il nostro studio prende in considerazione 3.080 contee negli Stati Uniti e copre il 98% della popolazione statunitense. Abbiamo fatto uso di nostri studi precedenti che stimavano gli effetti a lungo termine dell'esposizione a PM_{2,5} sul tasso di mortalità tra 60 milioni di cittadini statunitensi che rientrano nel programma di assistenza sanitaria Medicare (16,17,18). Abbiamo utilizzato una piattaforma di dati di ricerca ampiamente sperimentata che raccoglie, armonizza e mette in relazione a livello nazionale dati relativi a inquinamento atmosferico, dati demografici e altre potenziali variabili di disturbo (confounding variables) con dati relativi agli effetti sulla salute. Abbiamo ampliato questa piattaforma con l'apporto di dati relativi al Covid-19 di recente acquisizione provenienti da fonti autorevoli (19). Tutte le fonti di dati utilizzate in queste analisi, unitamente al codice interamente riproducibile, sono di pubblico dominio per facilitare ulteriori ricerche sui suddetti rapporti a fronte dell'evoluzione della pandemia e della disponibilità di una maggiore quantità di dati.

Metodi

Nella **Tabella 1** sono sintetizzate le fonti che abbiamo utilizzato e la loro provenienza, compresi link da cui estrarre direttamente i dati grezzi.

Decessi da Covid-19: il numero di decessi da Covid-19 per ciascuna contea statunitense è stato reso noto dalla John Hopkins University, Center for System Science and Engineering Coronavirus Resource Center (20), che fornisce i dati aggiornati più completi sul Covid-19 a livello di contea comunicati dagli US Centers for Disease Control and Prevention e dai dipartimenti della Sanità dei singoli stati. Questi dati forniscono in tempo reale il numero di decessi e il numero di casi confermati in ciascuna contea statunitense. Abbiamo raccolto il numero cumulativo di decessi in ciascuna contea fino al 4 aprile 2020 compreso. Nelle nostre analisi i tassi di mortalità da Covid-19 a livello di contea corrispondono al rapporto tra il numero di decessi da Covid-19 e il numero di abitanti della contea considerata.

Esposizione all'inquinamento atmosferico: abbiamo calcolato l'esposizione prolungata a PM_{2,5} a livello di contea (valore medio relativo agli anni 2000-2016) ricorrendo a un modello predittivo di esposizione consolidato (21). I livelli di esposizione a PM_{2,5} sono stimati su base mensile a una risoluzione di griglia pari a 0,01° x 0,01° in tutto il territorio continentale degli Stati Uniti combinando dati modellati, monitorati e rilevati da satelliti in una regressione ponderata geograficamente. Queste stime sono state oggetto di numerose validazioni incrociate (22). Abbiamo calcolato valori medi di PM_{2,5} su base temporale (2000-2016) a livello di contea facendo la media dei valori di PM_{2,5} stimati all'interno di

una determinata contea. Abbiamo calcolato in modo analogo l'esposizione a PM_{2,5} nel 2016 in ciascuna contea da utilizzare nelle analisi di sensitività.

Variabili di disturbo potenziali: abbiamo considerato variabili di disturbo sedici variabili a livello di contea e una variabile a livello statale, come segue: densità di popolazione, percentuale di abitanti di età superiore a 65 anni, percentuale di persone che vivono in stato di povertà, reddito familiare medio, percentuale di popolazione di colore, percentuale di popolazione ispanica, percentuale di popolazione adulta con livello di istruzione inferiore alla scuola superiore, valore medio dell'abitazione, percentuale di proprietari di prime case, indice di massa corporea (BMI) medio (quale indicatore di obesità), percentuale di fumatori abituali, numero di posti letto negli ospedali e i valori medi giornalieri di temperatura e di umidità relativa in estate (giugno-settembre) e in inverno (dicembre-febbraio) per ciascuna contea e il numero di test Covid-19 eseguiti a livello di stato. Per ulteriori dettagli sulla creazione di tutte le variabili utilizzate nell'analisi si rimanda ai materiali supplementari.

Metodi statistici

Abbiamo adattato modelli misti binomiali negativi "a inflazione di zeri" (zero-inflated) (ZNB)(23) considerando i decessi da Covid-19 come effetto e PM_{2,5} come esposizione prescelta. Lo ZNB si compone di due sotto-modelli, un *sotto-modello di conteggio* che stima l'associazione tra i decessi da Covid-19 e PM_{2,5} (aggiustata per le covariate) nelle contee dove potrebbe registrarsi un decesso da Covid-19 (es.: per presenza di casi di contagio confermati) e un *sotto-modello zero* che tiene conto degli zeri in eccesso che potrebbero essere generati da contee dove potrebbero anche non registrarsi decessi da Covid-19 (es.: per assenza di casi di contagio confermati) e dove probabilmente non ci sono stati decessi da Covid-19 fino al 4 aprile 2020. Per ulteriori dettagli di modellazione si rimanda ai materiali supplementari. Abbiamo incluso un offset relativo al numero di abitanti e abbiamo effettuato aggiustamenti in relazione a tutte le variabili elencate in precedenza. Abbiamo incluso anche una intercetta casuale per singolo stato per tener conto di una potenziale correlazione in contee all'interno di uno stesso stato per via di caratteristiche simili in termini socio-culturali e comportamentali e di sistemi di assistenza sanitaria nonché di similitudini relativamente alla risposta all'emergenza Covid-19 e alle pratiche di test. Riportiamo esclusivamente il risultato del *sotto-modello di conteggio*. Nello specifico riportiamo stime di parametri esponenziati, i rapporti tra tassi di mortalità (MRR) e un intervallo di confidenza 95% provenienti dal *sotto-modello di conteggio*. Il MRR può essere interpretato quale incremento relativo del tasso di mortalità da Covid-19 associato a un aumento di 1 µg/m³ dell'esposizione media prolungata a PM_{2,5} tra le contee dove potrebbe registrarsi

un decesso da Covid-19. Non abbiamo riportato risultati provenienti dal *sottomodello zero*. Tutte le analisi sono state effettuate con il software statistico R e per l'adattamento (fitting) dei modelli è stato utilizzato il pacchetto NBZIMM (24).

Analisi secondaria

Abbiamo effettuato sei analisi secondarie per verificare la fondatezza dei risultati ottenuti rispetto alla serie di variabili di disturbo utilizzate, a potenziali variabili di disturbo non misurate e a valori anomali (outlier).

- Primo, poiché fino a oggi lo stato di New York risulta il più colpito dal Covid-19 all'interno degli Stati Uniti e ha registrato un numero di decessi collegati al virus di cinque volte superiore a quello dello stato immediatamente successivo, riteniamo che possa influenzare in larga misura la nostra analisi e quindi l'abbiamo ripetuta escludendo tutte le contee dello stato di New York.
- Secondo, i dati attualmente disponibili sul numero dei contagi e dei test effettuati non bastano a tracciare un quadro accurato dell'estensione della pandemia in una determinata contea e l'impossibilità di introdurre aggiustamenti completi in relazione a questo fattore potrebbe falsare in certo qual modo l'analisi (es.: se le contee che presentano un'esposizione elevata a PM2,5 tendono anche a registrare grandi estensioni della pandemia rispetto al numero di abitanti, i loro tassi di mortalità per unità di popolazione potrebbero apparire elevati in modo differenziale, dando origine a una correlazione spuria con PM2,5). Per verificare le possibili ricadute sui risultati ottenuti, abbiamo effettuato analisi escludendo le contee che hanno registrato un numero di casi confermati inferiore a 10.
- Terzo e ultimo, abbiamo adattato i modelli omettendo le seguenti variabili di disturbo potenziali (separatamente):
 - ° numero di posti letto negli ospedali
 - ° numero di test eseguiti in relazione al Covid-19
 - ° indice di massa corporea (BMI) medio della popolazione e percentuale di fumatori (BRFSS – Behavioral Risk Factor Surveillance System)
 - ° temperatura e umidità relativa in estate e inverno (condizioni meteorologiche)

Analisi di sensitività

Abbiamo effettuato numerose analisi di sensitività per verificare la fondatezza dei risultati ottenuti rispetto alle scelte in termini di dati e modelli.

* Primo, abbiamo ripetuto tutte le analisi ricorrendo a metodi diversi per stimare l'esposizione a PM_{2,5} (25)

* Secondo, poiché il nostro studio si basa su dati osservazionali, i risultati ottenuti potrebbero essere sensibili alle scelte di modellazione (es.: ipotesi di distribuzione o ipotesi di linearità). Abbiamo valutato la sensibilità alle suddette scelte effettuando delle analisi dove:

° il PM_{2,5} è stato considerato una variabile categorica (categorizzata in quintili empirici)

° è stato effettuato un aggiustamento in relazione alla densità di popolazione quale variabile categorica (categorizzata in quintili empirici)

° è stato utilizzato un modello binomiale negativo senza tener conto dell'inflazione di zeri.

I risultati delle analisi di sensibilità sono riportati nei materiali supplementari.

Risultati

Il nostro studio ha utilizzato dati provenienti da 3.080 contee, di cui 2.395 (77,8%) non avevano registrato decessi da Covid-19 al momento dell'analisi. Nella **Tabella 2** sono descritti i dati utilizzati nelle nostre analisi. Tutti i conteggi dei decessi da Covid-19 sono conteggi cumulativi fino al 4 aprile 2020.

Nella **Figura 1** sono visualizzati la variazione spaziale dell'esposizione prolungata a PM_{2,5} e i tassi di mortalità da Covid-19 (per milione di abitanti) per contea. Da un'analisi visiva emerge che i tassi di mortalità da Covid-19 sono più elevati nella regione medio-atlantica, nel Midwest superiore, nella costa del golfo e nelle regioni della costa occidentale. In generale questo andamento spaziale dei tassi di mortalità da Covid-19 va di pari passo con valori elevati in termini di densità di popolazione e di esposizione a PM_{2,5}. Nei materiali supplementari vengono fornite ulteriori visualizzazioni e diagnostica dei dati che giustificano l'uso del modello ZNB nelle nostre analisi. Dopo l'eliminazione delle contee con variabili di disturbo mancanti, l'analisi principale è stata effettuata prendendo in considerazione 1.783 contee che rappresentano il 90% dei decessi da Covid-19 fino al 4 aprile 2020 compreso. Nella **Figura 2** sono riportati i rapporti tra tassi di mortalità (MRR) e l'intervallo di confidenza (CI) 95% per le nostre analisi principale e secondarie. Nei materiali supplementari sono riportati i risultati di tutte le analisi di sensibilità. Nell'analisi principale il MRR stimato è pari a 1,15 (1,05 , 1,25), il che significa che abbiamo riscontrato che l'aumento anche solo di 1 µg/m³ dell'esposizione media prolungata a PM_{2,5} è associato a un incremento statisticamente significativo del 15% del numero di decessi da Covid-19.

Questo sta ad indicare che, se nella contea di New York (Manhattan) l'esposizione media prolungata a PM_{2,5} fosse stata ridotta anche solo di 1 µg/m³, il numero di decessi da Covid-19, che ha raggiunto un totale di 1905 al 4 aprile 2020, sarebbe diminuito di 248 (= 1905 x (1 - 1/1,15)) casi.

Nelle analisi secondarie illustrate nella **Figura 2** abbiamo: a) eliminato il numero di posti letto negli ospedali dalla serie di aggiustamenti; b) eliminato il numero di casi testati in ogni stato dalla serie di aggiustamenti; c) eliminato la percentuale di fumatori e l'indice di massa corporea (BMI) provenienti dal Behavioral Risk Factor Surveillance System (BRFSS) dalla serie di aggiustamenti; d) eliminato la temperatura e l'umidità relativa in estate e inverno dalla serie di aggiustamenti; e) escluso i dati provenienti dallo stato di New York; f) escluso le contee che hanno registrato un numero di casi confermati inferiore a 10. I risultati di queste analisi sono in linea con quelli dell'analisi principale tranne in caso di eliminazione del numero di posti letto negli ospedali dalla serie di aggiustamenti che fa diminuire la stima puntuale e riduce la significatività statistica. Questo fatto starebbe a indicare che il numero di posti letto negli ospedali è una variabile di disturbo forte.

Discussione

Questo è il primo studio condotto a livello nazionale negli Stati Uniti che si propone di stimare il rapporto esistente tra l'esposizione prolungata a PM_{2,5} e i tassi di mortalità da Covid-19. Abbiamo riscontrato un'evidenza statisticamente significativa che l'aumento di 1 µg/m³ dell'esposizione media prolungata a PM_{2,5} è associato a un incremento del 15% del tasso di mortalità da Covid-19. I nostri risultati sono stati aggiustati rispetto a un'ampia serie di variabili di disturbo di natura socioeconomica, demografica, meteorologica, comportamentale e relative al sistema sanitario e si sono rivelati fondati in numerose analisi di sensitività.

Nel nostro precedente studio relativo a 60 milioni di cittadini americani di età superiore a 65 anni avevamo riscontrato che l'aumento di 1 µg/m³ dell'esposizione media prolungata a PM_{2,5} è associato a un incremento dello 0,73% del tasso di mortalità generale. Ne consegue che un piccolo aumento di valore dell'esposizione prolungata a PM_{2,5} induce un forte incremento del tasso di mortalità da Covid-19, una percentuale di 20 volte superiore a quella stimata per la mortalità generale.

La presente analisi fornisce una caratterizzazione tempestiva del rapporto tra esposizione storica all'inquinamento atmosferico e numero di decessi da

Covid-20 negli Stati Uniti. E' di fondamentale importanza condurre ricerche sul modo in cui fattori modificabili possono aggravare i sintomi dell'infezione da Covid-19 e aumentare il rischio di morte al fine di orientare politiche e comportamenti in modo tale da ridurre al minimo la letalità legata alla pandemia. La nostra analisi si basa su dati aggiornati relativi all'espansione del Covid-19 a livello di popolazione e su parametri di misurazione dell'esposizione all'inquinamento atmosferico ampiamente validati (26).

Tra i punti di forza della presente analisi figurano l'aggiustamento rispetto a un'ampia gamma di variabili di disturbo e la dimostrata fondatezza dei risultati anche a fronte dell'applicazione di modelli diversi. Inoltre, l'analisi si basa esclusivamente su dati e codice di pubblico dominio, per cui rappresenta una piattaforma che consente alla comunità scientifica di aggiornare e ampliare queste analisi in linea con l'evoluzione della pandemia e la disponibilità di ulteriori dati.

L'impossibilità di quantificare con precisione il numero di casi da Covid-19 a causa della limitata capacità di test rappresenta una limitazione potenziale in quanto non siamo in grado di effettuare un aggiustamento completo rispetto all'estensione della pandemia in una determinata contea, il che potrebbe comportare la mancata misurazione di un certo numero di variabili di disturbo. Abbiamo in alternativa effettuato un aggiustamento per numero totale di abitanti e riteniamo che l'uso di un modello a inflazione di zeri possa sopperire almeno in parte a questo inconveniente in quanto il MRR è stato stimato tenendo conto del fatto che alcune contee non avrebbero probabilmente fatto registrare alcun decesso da Covid-19 poiché il numero di casi confermati al loro interno era ancora pari a zero. Inoltre, le misure in risposta alla pandemia, in costante evoluzione, danno origine a uno scenario fluido per quanto riguarda la disponibilità di risorse mediche in una determinata area. E' probabile che la disponibilità di queste risorse influisca sui tassi di letalità. Poiché non sono disponibili dati in tempo reale sulle suddette misure e risorse, non siamo in grado di effettuare aggiustamenti rispetto a questo elemento, il che potrebbe rivelarsi di conseguenza un'altra fonte di variabili di disturbo.

I risultati del presente studio indicano che l'esposizione prolungata all'inquinamento atmosferico aumenta la vulnerabilità agli esiti più gravi di un'infezione da Covid-19. Queste conclusioni sono in linea con il noto legame tra esposizione a PM_{2,5} e numerose patologie respiratorie e cardiovascolari concomitanti che incrementano notevolmente il rischio di morte per i pazienti affetti da Covid-19. Sono anche in linea con risultati che dimostrarono come

l'esposizione all'inquinamento atmosferico avesse aumentato in larga misura il rischio di morte durante lo scoppio nel 2003 dell'epidemia SARS (Severe Acute Respiratory Syndrome) che è provocata da un altro tipo di coronavirus (27). Il presente studio rappresenta uno stimolo per ricerche di follow-up di più ampio respiro a fronte dell'aumentata disponibilità di dati qualitativamente e quantitativamente migliori sulla pandemia da Covid-19, il che dovrebbe comportare la validazione di questi risultati in altre fonti di dati e tipi di studi e la realizzazione di studi sui meccanismi biologici, sull'impatto dei tempi di esposizione a PM_{2,5} e sui rapporti tra PM_{2,5} e altri effetti del Covid-19 come, per esempio, i ricoveri ospedalieri.

I risultati del presente studio sottolineano anche l'importanza della continuità nell'applicazione delle norme in vigore sull'inquinamento atmosferico anche durante l'emergenza Covid-19. In caso contrario, sulla base dei risultati ottenuti, prevediamo che si potrebbe verificare un potenziale incremento dei decessi da Covid-19 e dei ricoveri ospedalieri con ulteriore carico per il nostro sistema sanitario e sottrazione di risorse che potrebbero essere impiegate per i pazienti affetti da Covid-19.

Ringraziamenti:

This work was made possible by the support from the NIH grant R01 ES024332-01A1, P50MD010428, ES024012, ES026217, ES028033; MD012769, HEI grant 4953-RFA14-3/16-4, and USEPA grants 83587201-0, RD-83479801. The contents are solely the responsibility of the grantee and do not necessarily represent the official views of the funding agencies. Further, funding agencies do not endorse the purchase of any commercial products or services related to this publication. The computations in this paper were run on (1) the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University; (2) the Research Computing Environment supported by the Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University. The authors would like to thank Lena Goodwin, for editorial assistance in the preparation of this manuscript.

Table 1: Data Sources used in the analysis (all publicly available)

	Source	Data
Outcome: COVID-19 Deaths	Johns Hopkins University the Center for Systems Science and Engineering (CSSE) Coronavirus Resource Center (https://coronavirus.jhu.edu/)	County level COVID-19 death count up to and including April 4, 2020
Exposure: PM _{2.5} concentrations	Atmospheric Composition Analysis Group (http://fizz.phys.dal.ca/~atmos/martin)	0.01° × 0.01° grid resolution PM _{2.5} prediction, averaged across the period 2000-2016 and averaged across grid cells in each county
Confounders	US Census/American Community Survey (https://www.census.gov/programs-surveys/acs/data.html)	County level socioeconomic and demographic variables for 2012-2016 American Community Survey
	US CDC BRFSS (https://www.cdc.gov/brfss/)	County level behavioral risk factor variables for 2011 (most recent year with county-level data available)
	Homeland Infrastructure Foundation-Level Data (HIFLD) (https://hifld-geoplatform.opendata.arcgis.com/datasets/hospitals)	County level number of hospital beds in 2019
	The COVID tracking project (https://covidtracking.com/)	State level number of COVID-19 tests performed up to and including April 4, 2020
	Gridmet via google earth engine https://developers.google.com/earth-engine/datasets/catalog/IDAHO_EPSC_OR_GRIDMET	4km × 4km temperature and relative humidity predictions, summer and winter averaged across the period 2000-2016 and averaged across grid cells in each county.

Table 2: Characteristics for the Study Cohort up to and including April 04, 2020

Variable	Entire Cohort (3080 counties)	Exclude NY State (3018 counties)	Exclude County < 10 Cases (1047 counties)	With HIFLD available (2272 counties)	With BRFSS available (2214 counties)

Number of deaths	7,054	4,662	6,968	6,342	7,024
Number of hospital beds	827,799	770,716	706,051	827,799	811,070
Number of testing (State level)	1,620,460	1,336,839			
PM _{2.5} (µg/m ³) (SD, standard deviation)	8.4 (2.5)	8.4 (2.5)	9.5 (2.3)	8.5 (2.5)	9.0 (2.3)
County level risk factors characteristics (SD)					
% Ever smoked	47.0 (9.2)	47.0 (9.2)	45.9 (8.0)	46.8 (8.9)	47.0 (9.2)
Mean BMI (kg/m ²)	28.1 (1.3)	28.1 (1.3)	27.9 (1.1)	28.1 (1.3)	28.1 (1.3)
% Below poverty level	10.5 (5.7)	10.6 (5.7)	9.7 (4.7)	10.3 (5.3)	10.3 (5.1)
% Below high school	21.2 (10.5)	21.3 (10.5)	19.5 (8.6)	20.6 (9.8)	21.4 (9.7)

education					
% Owner-occupied housing	74.2 (8.9)	74.2 (8.8)	70.1 (10.3)	74.1 (9.0)	73.4 (9.1)
% Hispanic	7.7 (12.3)	7.7 (12.4)	8.6 (11.3)	7.3 (11.4)	7.5 (11.5)
% Black	8.2 (14.2)	8.3 (14.3)	12.1 (15.4)	8.0 (13.6)	9.2 (13.9)
% older than 65 year-old	15.9 (4.1)	16.0 (4.1)	13.7 (3.6)	15.8 (4.0)	14.9 (3.5)
Population density (persons/mile ²)	446.0 (2121.62)	377.4 (1162.27)	1044.7 (2854.5)	514.1 (2384.8)	608.3 (2479.9)
Median household income (1000 \$)	49.0 (13.1)	48.8 (13.0)	54.7 (15.6)	49.6 (12.7)	50.1 (13.6)
Median home value (1000 \$)	136.6 (90.5)	135.2 (87.9)	179.8 (114.7)	134.7 (80.4)	148.8 (94.5)

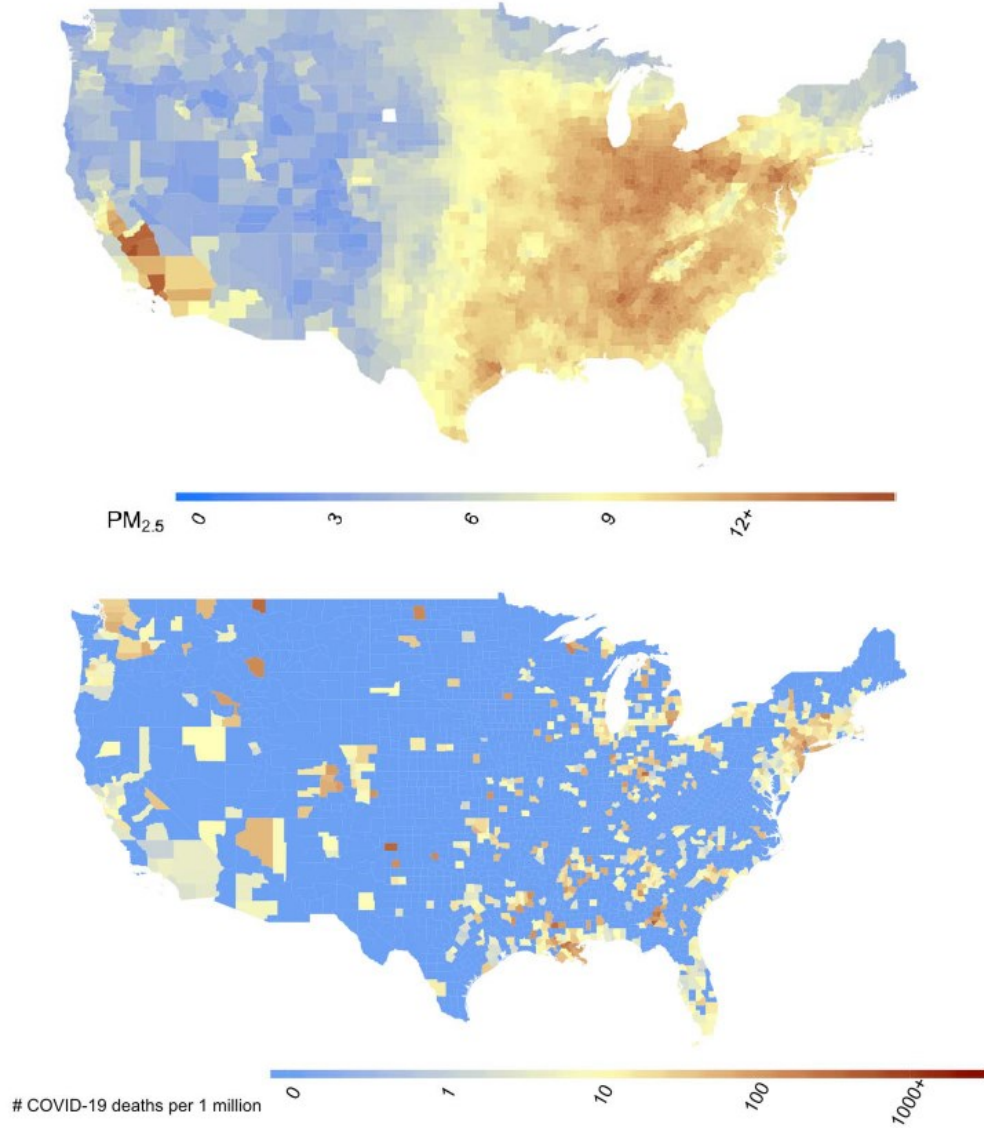


Figure 1: Maps show (a) county level 17-year long-term average of PM_{2.5} concentrations (2000-2016) in the US in g/m³ and (b) county level number of COVID-19 deaths per one million population in the US up to and including April 4, 2020.

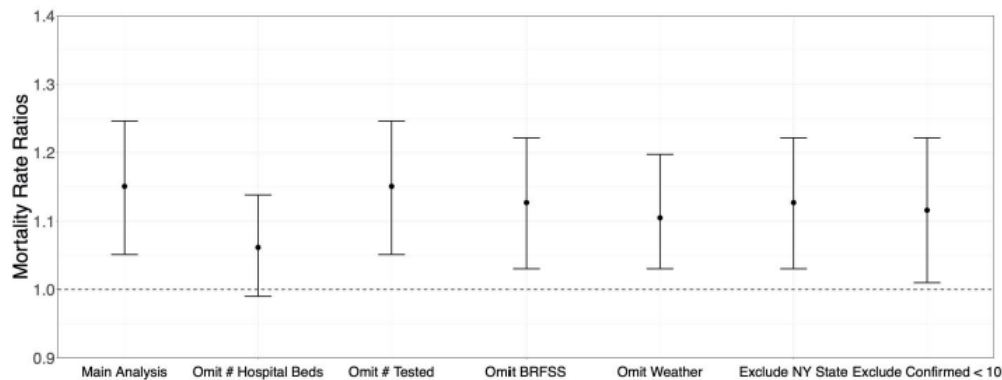


Figure 2: Mortality Risk Ratios (MRR) and 95% confidence intervals. The MRR can be interpreted as percentage increase in the COVID-19 death rate associated with a $1 \mu\text{g}/\text{m}^3$ increase in long term average $\text{PM}_{2.5}$. The MRR from the main analysis is adjusted by seventeen potential confounders (population density, percent of the population >65, percent living in poverty, median household income, percent black, percent Hispanic, percent of the adult population with less than a high school education, median house value, percent of owner-occupied housing, population mean BMI, percent ever-smokers, summer/winter temperature/relative humidity, and number of hospital beds for each county, and number of COVID-19 tests performed for each state). In secondary analyses, we (a) omit number of hospital beds; (b) omit number of tested cases in each state; (c) omit smoking rate and BMI from BRFSS; (d) omit summer/winter temperature/relative humidity; (e) exclude data from New York state; (f) exclude data from counties with less than 10 confirmed cases. All COVID-19 death counts are cumulative counts up to and including April 04, 2020.

References

1. Wei-jie Guan, Ph.D., Zheng-yi Ni, M.D., Yu Hu, M.D., Wen-hua Liang, Ph.D., Chun-quan Ou, Ph.D., Jian-xing He, M.D., Lei Liu, M.D., Hong Shan, M.D., Chun-liang Lei, M.D., David S.C. Hui, M.D., Bin Du, M.D., Lan-juan Li, M.D., et al. for the China Medical Treatment Expert Group for Covid-19. *NEJM*. 2020 DOI: 10.1056/NEJMoa2002032.
(<https://www.nejm.org/doi/full/10.1056/NEJMoa2002032>)
2. Wu, J.T., Leung, K., Bushman, M. *et al.* Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med* (2020).
(<https://www.nature.com/articles/s41591-020-0822-7>)

3. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) — United States, February 12–March 16, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:343–346. DOI: (<http://dx.doi.org/10.15585/mmwr.mm6912e>)
4. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19) (<https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>)
5. Diamond M, Peniston Feliciano HL, Sanghavi D, et al. Acute Respiratory Distress Syndrome (ARDS) [Updated 2020 Jan 5]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK436002/>.
6. Preliminary Estimates of the Prevalence of Selected Underlying Health Conditions Among Patients with Coronavirus Disease 2019 — United States, February 12–March 28, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:382–386. DOI: (<http://dx.doi.org/10.15585/mmwr.mm6913e>)
7. <https://www.stateofglobalair.org/report>
8. Brook RD, Franklin B, Cascio W et al. 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*. 2004;109(21):2655-71.
9. Di Q, Wang Y, Zanobetti A et al. Air pollution and mortality in the Medicare population. *NEJM*. 2017;376(26):2513-22.
10. Puett RC, Hart JE, Yanosky JD et al. Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. *EHP*. 2009;117(11):1697-701.
11. Dominici F, Peng R, Bell M et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*. 2006;295(10):1127–34.
12. Wellenius G. Ambient air pollution and the risk of acute ischemic stroke. *Arch Intern Med*. 2012;172(3):229–34.

13. Radim J. Šrám, Blanka Binková, Jan Dejmek, and Martin Bobak. Ambient Air Pollution and Pregnancy Outcomes: A Review of the Literature. *Vol. 113, No. 1* April 2005
(<https://doi.org/10.1289/ehp.6362>)
14. Rhee, Jongeun et al. Impact of Long-Term Exposures to Ambient PM_{2.5} and Ozone on ARDS Risk for Older Adults in the United States CHEST, Volume 156, Issue 1, 71 - 79 ([https://journal.chestnet.org/article/S0012-3692\(19\)30711-1/abstract](https://journal.chestnet.org/article/S0012-3692(19)30711-1/abstract))
15. U.S. EPA. Integrated Science Assessment (ISA) for Particulate Matter (Final Report, 2019). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-19/188, 2019.
16. Zanobetti A, Dominici F, Wang Y, Schwartz JD. (2014). A national case-crossover analysis of the short-term effect of PM_{2.5} on hospitalizations and mortality in subjects with diabetes and neurological disorders. *Environmental Health*, 13(1):38. PMID: PMC4064518.
17. Kioumourtzoglou MA, Schwartz J, Weisskopf M, Melly S, Wang Y, Dominici F, Zanobetti (2016) Long-term PM_{2.5} exposure and neurological hospital admissions in the Northeastern United States. *Environmental Health Perspectives*, 124(1):23-9. PMID: PMC4710596.
18. Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Dominici F, Schwartz J. (2017). Air Pollution and Mortality in the Medicare Population. *New England Journal of Medicine*, 376:2513-2522, June 29, 2017. PMID: 29020593 DOI: 10.1056/NEJMoa1702747 PMID: 5766848.
19. 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. *Environmental science & technology*. 2019;53(5):2595–2611.
20. <https://coronavirus.jhu.edu>
21. 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. *Environmental science & technology*.2019;53(5):2595–2611.
22. Venables, W. N. and Ripley, B. D. (2002) "Modern Applied Statistics with S". Fourth edition. Springer.
23. Xinyan Zhang, Himel Mallick, Xiangqin Cui, Andrew K. Benson, and Nengjun Yi (2017) Negative Binomial Mixed Models for Analyzing Microbiome Count Data. *BMC Bioinformatics* 18(1):4.
24. <https://rdr.io/github/nyuab/NBZIMM/man/glmm.zinb.html>
25. Zhang, X., Mallick, H., Tang, Z., Zhang, L., Cui, X., Benson, A.K. and Yi, N., 2017. Negative binomial mixed models for analyzing microbiome count data. *BMC bioinformatics*, 18(1), p.4.
26. Di Q, Amini H, Shi L, et al. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environ Int* 2019; 130: 104909.
27. http://fizz.phys.dal.ca/~atmos/martin/?page_id=140