

RESEARCH

Open Access



Associations between long-term exposures to airborne PM_{2.5} components and mortality in Massachusetts: mixture analysis exploration

Tingfan Jin^{1*}, Heresh Amini², Anna Kosheleva¹, Mahdiah Danesh Yazdi³, Yaguang Wei¹, Edgar Castro¹, Qian Di⁴, Lihua Shi⁵ and Joel Schwartz¹

Abstract

Background: Numerous studies have documented PM_{2.5}'s links with adverse health outcomes. Comparatively fewer studies have evaluated specific PM_{2.5} components. The lack of exposure measurements and high correlation among different PM_{2.5} components are two limitations.

Methods: We applied a novel exposure prediction model to obtain annual Census tract-level concentrations of 15 PM_{2.5} components (Zn, V, Si, Pb, Ni, K, Fe, Cu, Ca, Br, SO₄²⁻, NO₃⁻, NH₄⁺, OC, EC) in Massachusetts from 2000 to 2015, to which we matched geocoded deaths. All non-accidental mortality, cardiovascular mortality, and respiratory mortality were examined for the population aged 18 or over. Weighted quantile sum (WQS) regression models were used to examine the cumulative associations between PM_{2.5} components mixture and outcomes and each component's contributions to the cumulative associations. We have fit WQS models on 15 PM_{2.5} components and a priori identified source groups (heavy fuel oil combustion, biomass burning, crustal matter, non-tailpipe traffic source, tailpipe traffic source, secondary particles from power plants, secondary particles from agriculture, unclear source) for the 15 PM_{2.5} components. Total PM_{2.5} mass analysis and single component associations were also conducted through quasi-Poisson regression models.

Results: Positive cumulative associations between the components mixture and all three outcomes were observed from the WQS models. Components with large contribution to the cumulative associations included K, OC, and Fe. Biomass burning, traffic emissions, and secondary particles from power plants were identified as important source contributing to the cumulative associations. Mortality rate ratios for cardiovascular mortality were of greater magnitude than all non-accidental mortality and respiratory mortality, which is also observed in cumulative associations estimated from WQS, total PM_{2.5} mass analysis, and single component associations.

Conclusion: We have found positive associations between the mixture of 15 PM_{2.5} components and all non-accidental mortality, cardiovascular mortality, and respiratory mortality. Among these components, Fe, K, and OC have been identified as having important contribution to the cumulative associations. The WQS results also suggests potential source effects from biomass burning, traffic emissions, and secondary particles from power plants.

*Correspondence:

Tingfan Jin
tingfanjin@hsph.harvard.edu

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Keywords: Air pollution, Particle components, Weighted quantile sum regression

Background

Numerous studies have documented the associations between both short-term [1–6] and long-term [4, 7–15] exposures to fine particulate matter (PM_{2.5}) and adverse health outcomes across different populations, especially on non-accidental mortality [4, 6–9, 11–15]. In the most updated Integrated Science Assessment (ISA) for Particulate Matter, the EPA has concluded a causal relationship between long-term PM_{2.5} exposure and mortality exists [16]. Determining the relative toxicity of PM_{2.5} components could be very helpful in facilitating PM_{2.5} pollution control efforts.

However, fewer studies have evaluated the associations between long-term exposure to PM_{2.5} components and mortality as well as the relative contributions of different components to the cumulative associations between PM_{2.5} and mortality. Inconsistent findings on the associations between PM_{2.5} components and mortality risks have been reported in previous studies [17–19]. A key limitation in examining PM_{2.5} components is the general lack of high-resolution exposure measurements. Some studies have obtained data from fixed monitoring sites [20–22]. However, compared to the total PM_{2.5} mass, there are fewer monitors for PM_{2.5} components, and sometimes the spatial heterogeneity of some components, such as elemental carbon, is much greater than for overall PM_{2.5}. Therefore, those studies were limited in capturing the spatial variation of PM_{2.5} components concentrations, which is critical to a study of long-term exposure on large population and at large geographic scale. More recently, exposure prediction models with higher spatial resolution have been used as an alternative, including chemical transport models [23–25] and land use regression models [17–19, 26, 27]. However, many of these models have relatively moderate prediction accuracies and spatial resolutions. Therefore, it is worthwhile to use a model with a higher spatial resolution and out-of-sample cross-validated prediction accuracy to better capture PM_{2.5} component exposure and further evaluate their links with adverse health outcomes.

Another key problem to study the potential associations between PM_{2.5} components and adverse health outcomes is the high correlation among different particle components which have been reported by some previous studies [21, 27, 28]. There are serious collinearity issues in comparing each component's contribution to the outcome using traditional regression methods. Weighted quantile sum (WQS) regression is one approach to estimate the cumulative associations between multiple components and adverse health outcomes and identify the “bad actors” (considered as most toxic components

among all) when there are collinearity issues [29, 30]. The WQS approach has already been applied in previous studies investigating the links between complex environmental mixtures and health outcomes [31, 32].

In this study, we have applied novel exposure prediction models which have generated high-resolution PM_{2.5} component concentration data across the contiguous U.S. to capture the exposure in Massachusetts. The exposure models combined land use, satellite observations, traffic, meteorological variables, chemical transport model output, and monitoring data for 15 PM_{2.5} components. We evaluated the associations between long-term exposures to PM_{2.5} components mixture and all non-accidental mortality, cardiovascular mortality, and respiratory mortality. The WQS regression models were fit to examine the cumulative associations and each component's contribution to the cumulative associations. We also conducted additional analysis based on the a priori identified source categories for the PM_{2.5} components, and the total PM_{2.5} mass analysis with single component associations examinations as supplemental secondary analysis.

Methods

This study was conducted at the Census tract level in the state of Massachusetts. The study period was from 2000 to 2015. The following exposure data, mortality data, and covariates data were collected and processed within this area and during this period.

Exposure data

15 PM_{2.5} components were selected in our study: zinc (Zn), vanadium (V), silicon (Si), lead (Pb), nickel (Ni), potassium (K), iron (Fe), copper (Cu), calcium (Ca), bromine (Br), sulfate (SO₄²⁻), nitrate (NO₃⁻), ammonium (NH₄⁺), organic carbon (OC), and elemental carbon (EC). Five of them were the major mass contributors to total PM_{2.5}: SO₄²⁻, NO₃⁻, NH₄⁺, OC, and EC, which will be referred to as “major components” in the following content. The other ten are trace elements. One reason to select these components was the number of monitors available to train the models. Another important reason to select these components is the relative toxicity of them. We chose to estimate the concentrations of 10 trace elements as well as the major components because there are evidence showing the trace elements are important for PM_{2.5} toxicity despite their low mass concentrations [33–40]. Moreover, the 10 trace elements represent various periodic table groups (groups 1 (K), 2 (Ca), 5 (V), 8 (Fe), 10 (Ni), 11 (Cu), 12 (Zn), 14 (Si and Pb), and 17 (Br)) as elements in the same group have similar physical or chemical characteristics.

We have generated annual mean predictions for the 15 $PM_{2.5}$ components' concentrations across the contiguous U.S. at a $50\text{ m} \times 50\text{ m}$ spatial resolution in urban areas and a $1\text{ km} \times 1\text{ km}$ spatial resolution in non-urban areas by year over the study period [41]. Specifically, we collected daily mean $PM_{2.5}$ component data from 987 monitoring sites across the United States (where 69 are located in Massachusetts). 166 predictors were used for modeling, including time and geography information (year, latitude, longitude, elevation), satellite observation data (vegetation, water index, nighttime lights, CO and CH₄, aerosol optical depth, etc.), meteorological data (temperature, humidity, wind velocity and direction, planetary boundary layer height, etc.), emitting/surrogate of emission sources (distance to power plants, distance to highways, traffic counts, burning index, etc.), and many others.

In non-urban areas, five machine-learning algorithms were fit to predict each component and then six super-learners and an ensemble weighted averaging model (ENWA) were applied to integrate predictions from the individual machine-learning algorithms. In urban areas, due to the computational limitations, we fit only three machine-learning algorithms and integrated predictions using four super-learners and an ENWA. The super-learner or ENWA with best performance was chosen as the final model, which varied by $PM_{2.5}$ component. With inclusion of a large number of monitoring sites across the US and many new predictors, this approach yielded excellent model performance at high spatial resolution, with out-of-sample validation R^2 for individual components ranging from 0.821 (Br in urban areas) to 0.957 (SO_4^{2-} in non-urban areas) [42, 43]. The root-mean-square error (RMSEs) calculated from regressing the monitored values against the predicted values specific in Massachusetts are presented in the supplemental materials (Table S1 in Additional File 1).

Each Census tract was matched to all the grid cells whose centroids were in it. We then averaged these grid cell-level predictions for each component at the Census tract level for each calendar year and assigned the Census tract-level annual means to each subject. Figure S1a and S1b in Additional File 1 shows the Census tract-level mean concentrations for each component in Massachusetts. Table S2 in Additional File 1 shows how the mean annual average concentrations across all census tracts in Massachusetts of the 15 $PM_{2.5}$ components vary from 2000 to 2015.

In addition to the 15 $PM_{2.5}$ components data, we also obtained the daily $PM_{2.5}$ (24-hour mean) predictions at $1\text{ km} \times 1\text{ km}$ spatial resolution from 2000 to 2015 in the state of Massachusetts from a spatial-temporal ensemble model integrating three machine learning algorithms (neural networks, random forest, gradient boosting) [44].

We then calculated the annual average $PM_{2.5}$ concentrations for each $1\text{ km} \times 1\text{ km}$ grid from 2000 to 2015 and aggregated them to the census tract level like the $PM_{2.5}$ components concentrations. These $PM_{2.5}$ concentrations were used for total $PM_{2.5}$ mass analyses supporting the interpretation of $PM_{2.5}$ components analyses results.

Mortality Data

We obtained the records of all deaths for years 2000–2015 from the Massachusetts Department of Public Health. These mortality records included the residential address (geocoded to longitude and latitude), age, sex, race, date of death, and cause of death (Underlying Cause of Death, ICD-10 codes). We restricted our analysis to non-accidental mortality (ICD-10 codes A-R) and dropped mortality records with missing age or residential address and restricted our study to decedents with age ≥ 18 years. This procedure left 787,352 deaths remaining in the final dataset. We aggregated the deaths by Census tract and year. We also calculated yearly counts of deaths from cardiovascular diseases (ICD-10 codes I00 to I99) and deaths from respiratory diseases (ICD-10 codes J00 to J99) based on the cause of death information.

Covariates

Meteorological Data – We obtained the daily maximum and minimum relative humidity and temperature data from gridMET at a $4\text{ km} \times 4\text{ km}$ spatial resolution (01/01/2000–12/31/2015) and calculated the daily mean values by averaging the maximum and minimum values [45]. We then calculated the annual average relative humidity and temperature. Each Census tract was linked to its nearest grid cell based on the distances between the centroids of the census tract and the grid cell, then the meteorological measures were assigned accordingly.

Census Data – Socioeconomic and demographic data were obtained from the 2000 and 2010 Decennial Censuses, and the 2009–2015 American Community Survey (ACS) 5-year estimate data retrieved from the IPUMS data service [46]. We pre-specified and collected data of potential confounders at the Census tract level including median household income, population density, proportion female, proportion African American, proportion Hispanic or Latino, proportion aged 65 or older, proportion whose income was below poverty level, and proportion with at least a high school degree or equivalent. We also obtained the total population with aged 18 or older for each Census tract. For years 2001–2008, the data were interpolated linearly from the observed data for each variable separately, using year as the sole predictor. This interpolation was applied to all census variables included in our analysis.

Lung Cancer Admission Rate Data – We computed the ZIP code-level lung cancer hospitalization rate data

from the MEDPAR files of the Center for Medicare and Medicaid Services. This lung cancer hospitalization rate variable was used as a surrogate for cumulative smoking. Each Census tract was first matched to the ZIP Code Tabulation Area (ZCTA) with maximum overlap, and then matched to the corresponding ZIP code.

Statistical methods

We had three outcomes of interest: all non-accidental mortality, cardiovascular mortality, and respiratory mortality. The main exposures of interest were the aforementioned 15 $PM_{2.5}$ components. The covariates selected for confounder control included the meteorological variables, socioeconomic and demographic variables, and lung cancer admission rates.

Our main analysis is assessing the cumulative associations between 15 $PM_{2.5}$ components and three different mortality outcomes by fitting weighted quantile sum (WQS) regression models. WQS is a supervised mixture analysis methods treating multiple pollutants as a mixture to which people were simultaneously exposed. In WQS, two parts of calculation are conducted. First, WQS estimates a composite index summarizing the pollutant mixture. This index is a weighted sum of each pollutant's quantiles (WQS term) based on each pollutant's relevance to the outcome being examined. Second, the WQS model estimates the cumulative association between the mixture and outcome by fitting a multivariate regression model with the WQS term as the exposure metric.

One advantage of WQS is it could reduce the impact of collinearity and extreme values from highly correlated pollutants in estimating the cumulative associations. In this study, we have 15 correlated $PM_{2.5}$ components (Figure S2a and S2b in Additional File 1). Moreover, we could also interpret each pollutant's contribution to the cumulative associations through the weights assigned to them in the WQS term [29]. The estimated weights could vary across different outcomes examined in WQS models.

For each outcome of interest (i.e., all non-accidental mortality, cardiovascular mortality, respiratory mortality), a separate WQS regression model was fit. More detailed explanation on how the WQS models were fit could be found in Paragraphs S1 in Additional File 1. Generally, the dataset would first be split into a training set (to estimate the WQS index and weights) and a validation set (to estimate the cumulative association), which we set as 50:50. We then scored the components concentrations into deciles and generated 250 bootstrap samples of the training set. In each bootstrap sample, a separate set of weights was estimated. Only the positive or null associations between $PM_{2.5}$ components and mortality were assessed in our study since we know of no evidence suggesting any of these components is beneficial for health through inhalation. Thus, the final weights of each

component were estimated by averaging the weight estimated from the bootstrap samples whose results showed a positive or null association between $PM_{2.5}$ components and mortality. The final WQS index could be presented as $\sum_1^{15} w_i q_i$ where w_i is the weight for the i th component concentration and q_i is the quantile score for that component concentration for that observation.

Then, the cumulative association between the mixture and outcome was estimated in the validation set through multivariate quasi-Poisson regression models. Several previous studies have used the Poisson regression models to examine the health effects of long-term air pollution exposures [47–51]. In our study, to address the overdispersion issue, we fit quasi-Poisson regression models all over the analyses. The formula for the regression models could be presented as:

$$\log(count) = \theta_0 + \theta_1 \cdot WQS + \theta_2' \cdot \mathbf{X} + offset(\log(population))$$

In this formula, *count* was the annual number of deaths; *WQS* was the WQS term, the exposure of interest; *X* were the covariates including the meteorological data, lung cancer admission data, socioeconomic data, and demographic data; and *offset(log(population))* was the offset term for mortality rate estimation. In WQS, each w_i was constrained between 0 and 1, and the sum of weights $\sum_1^{15} w_i$ was constrained to be 1. Therefore, one unit increase in final WQS index could be treated as approximately one decile increase in all components, and we interpreted the cumulative association estimates as the mortality rate ratio associated with this increase.

We also conducted several additional analyses of WQS models, and two additional sets of WQS models were fit. In each additional set, one separate WQS models was fit for each outcome of interest.

First, we noticed that there is evidence suggesting that interpreting the sum of the weights for a group of pollutants could be more reasonable than interpreting the weights for the single pollutants with highly correlated pollutants [29]. We have identified 8 source groups a priori: heavy fuel oil combustion, biomass burning, crustal matter, non-tailpipe traffic source, tailpipe traffic source, secondary particles from power plants, secondary particles from agriculture, and unknown source. We grouped the 15 $PM_{2.5}$ components into each category and calculated the sums of the weights of components for each source group to have a comparison among different sources.

Second, we noticed that there was a clear highly correlated cluster of 14 components except K in our study (Figure S2a and S2b in Additional File 1) and a cluster of highly correlated pollutants may result in underestimated weights in WQS models [29]. Therefore, it was

Table 1 Summary characteristics of mortality records in the final dataset

Characteristics	n (%)
Total	787,352 (100)
Gender	
Male	360,464 (45.8)
Female	426,873 (54.2)
Missing	15 (0.00019)
Race	
White	733,320 (93.1)
Black	31,387 (4.0)
Other	22,558 (2.9)
Missing	87 (0.01)
Cause of Death	
Cardiovascular Diseases	263,157 (33.4)
Respiratory Diseases	87,197 (11.1)
Other Non-accidental	436,998 (55.5)
Mean Age (Standard Deviation)	77.68 (14.15)

possible that K may “steal” the weights from other components. Hence, we also fit an additional set of WQS models without K to see how the results would change against this exclusion.

Finally, to further support the interpretation of we also fit a WQS models on the source groups aforementioned rather than directly on the 15 PM_{2.5} components. To do so, we first calculated the total mass of each source groups by summing the components mass of each group, and then fit WQS models taking the source groups as the individual pollutant in WQS models, with all other model settings and parameters the same as the components WQS models.

In addition to the WQS models, we conducted a total PM_{2.5} mass analysis through a quasi-Poisson regression model with PM_{2.5} as the exposure term. Moreover, we also fit single PM_{2.5} component models to see the individual-by-individual associations with the outcomes. In these quasi-Poisson models, penalized splines with penalty parameter chosen using restricted maximum likelihood estimation were used to adjust for potentially nonlinear confounding. In the single PM_{2.5} component models, confidence intervals were further adjusted for multiple comparisons using the Bonferroni correction.

All the data work and statistical analysis were conducted using R 4.1.1, and all WQS models were fit using gWQS 3.0.4 package [52].

Results

Of the 787,352 non-accidental deaths in the final dataset, 54.2% were female and 93.1% were white. The average age of death was 77.7 years. There were 263,157 deaths (33.4%) from cardiovascular diseases and 87,197 deaths (11.1%) from respiratory diseases. More detailed summary statistics are presented in Table 1.

Table 2 summarizes the characteristics of the Census tract-level annual mean PM_{2.5} components' concentrations, relative humidity, and temperature. SO₄²⁻ contributed the most to the total mixture mass, with a mean annual concentration of 2.21 µg/m³. Organic carbon also contributed a large portion with a mean annual concentration of 1.72 µg/m³. For the trace element components, Si, Fe, and K had the highest contribution to the total mixture mass with mean annual concentrations of 54.34 ng/m³, 50.80 ng/m³, and 47.81 ng/m³, respectively. The mean annual relative humidity was 64.5% and the mean annual temperature was 10.1 °C.

The correlations between PM_{2.5} components ranged from 0.04 for K-Si and K-Fe to 0.87 for NH₄⁺-SO₄²⁻. The correlations between K and other components were generally lower than the correlations among other 14 components. The detailed correlation matrix was presented in Figure S2a and S2b in Additional File 1.

WQS Models on 15 PM_{2.5} components

Table 3 summarizes the cumulative association estimates from WQS models on 15 PM_{2.5} components. We found positive associations between the PM_{2.5} components mixture and all three outcomes. The mortality rate ratio for cardiovascular mortality (RR=1.0533, 95% CI 1.0486–1.0581 per approximately 1 decile increase in all PM_{2.5} components) was of greater magnitude than respiratory (RR=1.0460, 95% CI 1.0392–1.0529) and all non-accidental mortality (RR=1.0303, 95% CI 1.0265–1.0341).

Figure 1 summarizes the weights of each component contributing to the associations with three outcomes. K and OC were found to have large weights to all three outcomes. Fe was found to have large weights to all non-accidental mortality and cardiovascular mortality. Besides, we also found large weights for Si, Ni, and SO₄²⁻ to cardiovascular mortality and for NO₃⁻ to respiratory mortality.

For the component mixture without K, we still found positive associations with all three outcomes and similar patterns across different causes. However, the mortality rate ratios were of less magnitude after dropping K from the mixture (Table S3 in Additional File 1). The orders of other components were relatively robust against to the dropping of K from the mixture. The components having large weights when including K in the components generally still had large weights when we dropped K (Figure S3 in Additional File 1).

Sums of weights for source groups

The PM_{2.5} components grouped by a priori identified source categories and the sum weights for each source group are presented in Table 4. We have identified 8 source groups. The source groups with high sum weights could be identified as important source groups

Table 2 Summary characteristics of particle component concentration, relative humidity, and temperature

Variable	Median	Mean	IQR	S.D.	75th	95th	Max
<i>PM_{2.5} Major Components</i>							
SO ₄ ²⁻ (µg/m ³)	2.36	2.21	1.41	0.75	2.87	3.25	3.76
NO ₃ ⁻ (µg/m ³)	0.80	0.82	0.26	0.18	0.94	1.15	1.57
NH ₄ ⁺ (µg/m ³)	0.75	0.76	0.45	0.25	0.96	1.17	1.49
OC (µg/m ³)	1.64	1.72	0.47	0.39	1.92	2.51	3.56
EC (µg/m ³)	0.45	0.48	0.16	0.14	0.55	0.76	1.23
<i>PM_{2.5} Trace Elements</i>							
Zn (ng/m ³)	9.29	9.49	3.82	2.72	11.36	14.12	19.18
V (ng/m ³)	1.43	1.70	2.00	1.29	2.54	4.15	6.84
K (ng/m ³)	47.63	47.81	6.20	4.77	50.87	55.83	67.21
Si (ng/m ³)	51.15	54.34	17.63	13.79	61.47	83.93	107.14
Pb (ng/m ³)	1.96	2.39	1.77	1.07	3.30	4.33	5.52
Ni (ng/m ³)	0.98	1.14	0.86	0.63	1.52	2.41	3.03
Fe (ng/m ³)	49.72	50.80	22.12	17.20	61.01	82.70	124.43
Cu (ng/m ³)	2.47	2.66	1.30	1.15	3.18	5.03	9.92
Ca (ng/m ³)	22.68	23.30	6.94	5.59	26.36	34.00	50.17
Br (ng/m ³)	2.63	2.63	0.41	0.32	2.83	3.15	3.68
<i>Total PM_{2.5} Mass</i>							
PM _{2.5} (µg/m ³)	8.75	8.73	2.80	1.94	10.08	11.96	16.39
<i>Meteorological variables</i>							
Relative humidity (%)	64.44	64.53	3.28	2.91	66.03	69.74	77.32
Temperature (°C)	10.09	10.09	1.36	1.05	10.83	11.72	12.86

Abbreviations: IQR for interquartile range; S.D. for standard deviation; 75th for 75th percentile; 95th for 95th percentile; Max for maximum value; SO₄²⁻ for sulfate; NO₃⁻ for nitrate; NH₄⁺ for ammonium; OC for organic carbon; EC for elemental carbon; Zn for zinc; V or vanadium; K for potassium; Si for silicon; Pb for lead; Ni for nickel; Fe for iron; Cu for copper; Ca for calcium; Br for bromine

Table 3 Cumulative associations between PM_{2.5} components and mortality estimated from WQS models^a

Cause-specific Mortality	Mortality Rate Ratio
<i>WQS Models on all 15 PM_{2.5} components</i>	
All Non-accidental Mortality	1.0303 (1.0265, 1.0341)
Cardiovascular mortality	1.0533 (1.0486, 1.0581)
Respiratory mortality	1.0460 (1.0392, 1.0529)
<i>WQS Models on the source groups identified for the 15 PM_{2.5} components</i>	
All Non-accidental Mortality	1.0277 (1.0241, 1.0313)
Cardiovascular mortality	1.0488 (1.0444, 1.0532)
Respiratory mortality	1.0435 (1.0368, 1.0504)

^aThe mortality rate ratios are per 1 unit increase in the WQS index, which could be interpreted as approximately 1 decile increase in all PM_{2.5} components or the source groups identified for the PM_{2.5} components

contributing to the cumulative association. Since OC has been assigned to both tailpipe traffic source and secondary particles (power plants), we have made an arbitrary split of OC’s weights giving 50% to each group.

We can see that for all non-accidental mortality, non-tailpipe traffic source and biomass burning had the highest weights. Non-tailpipe and tailpipe traffic sources

together consisted of over 50% of the total weights, indicating a strong contribution from traffic pollution to all non-accidental mortality. For respiratory mortality, we found highest weight for biomass burning, with secondary particles from power plants and the two traffic source groups also having large weights. The weights for cardiovascular mortality were more evenly distributed, with the weight for secondary particles from power plants as the highest. Biomass burning had the second highest weights, with traffic sources, heavy fuel oil combustion, and crustal matter also having weights over 0.1.

WQS Models on source groups

The cumulative association estimates from WQS models on source groups are also presented in Table 3. We have also made an arbitrary split of OC’s mass giving 50% to each group of tailpipe traffic source and secondary particles (power plants). Similar to WQS models on PM_{2.5} components, we also found positive associations between the mixture and all three outcomes, with mortality rate ratio for cardiovascular mortality (RR=1.0488, 95% CI 1.0444–1.0532 per approximately 1 decile increase in all PM_{2.5} components) of greater magnitude than respiratory (RR=1.0435, 95% CI 1.0368–1.0504) and all non-accidental mortality (RR=1.0277, 95% CI 1.0241–1.0313). The magnitudes of the mortality rate ratios estimated from

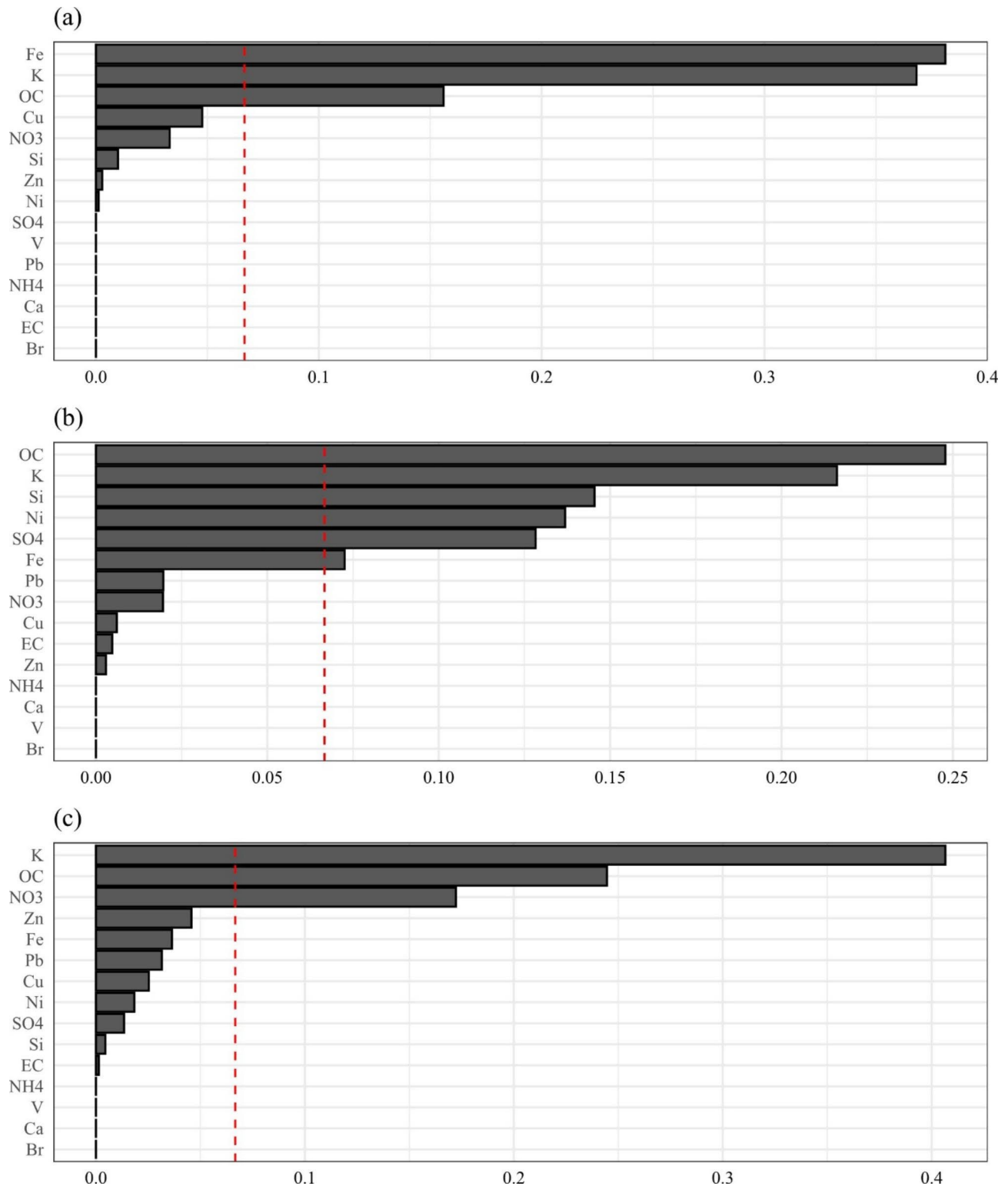


Fig. 1 The weights assigned to each PM_{2.5} component from all-component WQS models. (a): all non-accidental mortality; (b): cardiovascular diseases related mortality; (c): respiratory diseases related mortality; red dash line: reciprocal of the number of PM_{2.5} components in the mixture

WQS models on source groups were comparable to those estimated from WQS models on PM_{2.5} components.

The weights of each source group contributing to the association with three outcomes are presented in Table 4; Fig. 2. Compared to the sums of the weights estimated

Table 4 Source groups identified for PM_{2.5} components and WQS estimated weights assigned for the source groups

Sources	PM _{2.5} Components	All-cause	Cardio-vascular Mortality	Respiratory Mortality
<i>Weights estimated from WQS models on all 15 PM_{2.5} components, sum the weights for identified source groups</i>				
G1: Heavy fuel oil combustion	V, Ni	0.0012	0.1369	0.0184
G2: Biomass burning	K	0.3682	0.2162	0.4065
G3: Crustal matter	Si, Ca	0.0099	0.1454	0.0045
G4: Non-tailpipe traffic source	Zn, Pb, Fe, Cu	0.4316	0.1012	0.1389
G5: Tailpipe traffic source ^a	EC, OC	0.0780	0.1286	0.1237
G6: Secondary particles, power plants ^a	SO ₄ ²⁻ , NO ₃ ⁻ , OC	0.1111	0.2717	0.3080
G7: Secondary particles, agriculture	NH ₄ ⁺	1.3×10 ⁻⁸	3.2×10 ⁻⁸	6.1×10 ⁻⁹
G8: Unclear source	Br	8.8×10 ⁻⁹	1.8×10 ⁻⁹	1.8×10 ⁻⁹
<i>Weights estimated from WQS models on the identified source groups</i>				
G1: Heavy fuel oil combustion	V, Ni	1.4×10 ⁻⁶	0.0021	0.0002
G2: Biomass burning	K	0.4073	0.2342	0.4511
G3: Crustal matter	Si, Ca	0.0006	0.1069	0.0010
G4: Non-tailpipe traffic source	Zn, Pb, Fe, Cu	0.4472	0.0792	0.1586
G5: Tailpipe traffic source ^a	EC, OC	0.1445	0.2678	0.2674
G6: Secondary particles, power plants ^a	SO ₄ ²⁻ , NO ₃ ⁻ , OC	0.0004	0.3096	0.1209
G7: Secondary particles, agriculture	NH ₄ ⁺	6.7×10 ⁻⁶	0.0002	0.0006
G8: Unclear source	Br	6.9×10 ⁻⁷	6.9×10 ⁻¹⁰	0.0002

Abbreviations: All-cause for all non-accidental mortality, Cardiovascular Mortality for cardiovascular diseases related mortality; Respiratory Mortality for respiratory diseases related mortality

^aWe have made an arbitrary split of OC between tailpipe traffic source and secondary particles (power plants): for the sums of the weights, 50% of the OC weights were summed into tailpipe traffic source, and the other 50% were summed into secondary particles (power plants) source; for the WQS models on source groups, 50% of the OC mass were summed into tailpipe traffic source, and the other 50% were summed into secondary particles (power plants) source

from the WQS models on components, the source group with highest weights for each mortality outcome were still the same as non-tailpipe traffic source, secondary particles from power plants, and biomass burning. Additionally, the weights for tailpipe traffic source get higher and weights for heavy fuel oil combustion get lower. With higher sum weights for tailpipe and non-tailpipe traffic sources, the WQS models on source groups indicated more important contribution from traffic particles to the cumulative associations.

Total PM_{2.5} Mass Analysis and single component models

The total PM_{2.5} mass analysis results are presented in Table 5. We found similar pattern of associations compared to the cumulative associations estimated from WQS models: positive associations were observed between total mass with all three outcomes, with mortality rate ratio for cardiovascular mortality (RR=1.1089, 95% CI 1.0981–1.1199 per IQR increase) of greater magnitude than all non-accidental mortality (RR=1.0369, 95% CI 1.0292–1.0447) and respiratory mortality (RR=1.0754, 95%CI 1.0598–1.0911).

The results of single component's individual association estimates are presented in Table S4 in Additional File 1. Positive associations were observed between most of the components and all three outcomes except for Cu & all non-accidental mortality, Cu & respiratory mortality, and Br & all three outcomes. The general patterns across three outcomes were similar to WQS estimated cumulative associations and total PM_{2.5} mass analysis, that the mortality rate ratios were mostly of greater magnitude for cardiovascular mortality compared to the other two outcomes.

Discussion

We found positive cumulative associations between the PM_{2.5} component mixture with all three outcomes from both WQS models on PM_{2.5} components and source groups. The mortality rate ratios for cardiovascular mortality were of greater magnitude than the other two outcomes. This pattern was also observed in total PM_{2.5} mass analysis and single component models, with positive associations found for total PM_{2.5} mass and most components with all three outcomes. In WQS models on PM_{2.5} components, Fe, OC, and K have the highest weights contributing to the associations, and the sums of weights for the source groups indicating important sources as biomass burning, secondary particles from power plants, and non-tail pipe traffic source. In WQS models on source groups, the most important source group for each outcome were the same, while the weights for tailpipe traffic source get higher.

The 15 PM_{2.5} components included in this study could be considered as markers or surrogates for broader

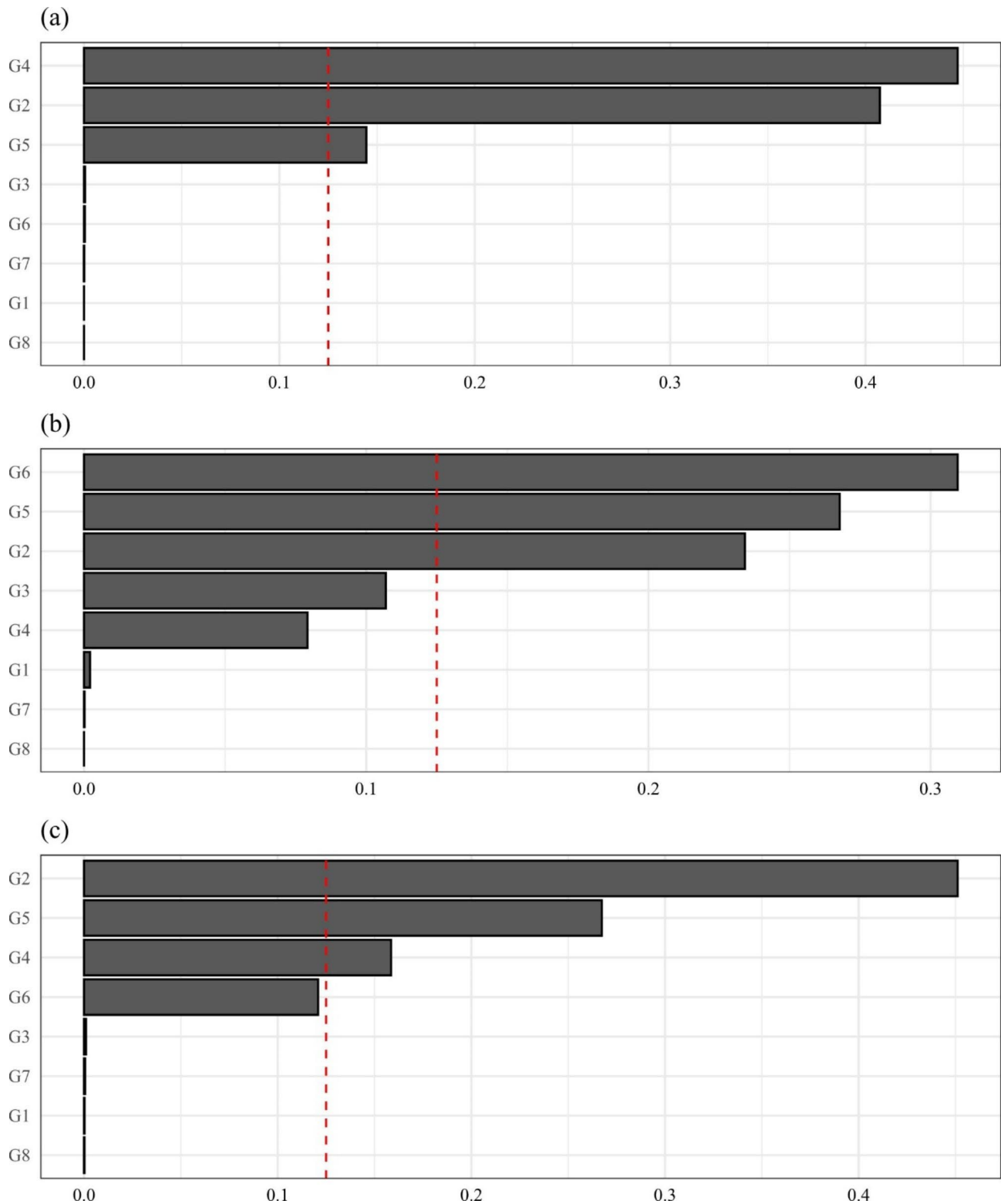


Fig. 2 The weights assigned to each PM_{2.5} source group from WQS models on identified source groups. (a): all non-accidental mortality; (b): cardiovascular diseases related mortality; (c): respiratory diseases related mortality; red dash line: reciprocal of the number of PM_{2.5} components in the mixture

source categories. For the most important components with highest weights identified from WQS models on

PM_{2.5} components, K is an indicator for biomass burning, Fe is an indicator for non-tailpipe traffic emissions,

Table 5 Associations between total PM_{2.5} mass and mortality outcomes^a

Cause-specific Mortality	Mortality Rate Ratio
All Non-accidental Mortality	1.0369 (1.0292, 1.0447)
Cardiovascular mortality	1.1089 (1.0981, 1.1199)
Respiratory mortality	1.0754 (1.0598, 1.0911)

^aPresented as the mortality rate ratio and 95% confidence intervals for interquartile range increase in concentrations

OC represents both secondary particles from power plants and tailpipe traffic emissions, Ni represents heavy fuel oil combustion, with other important components SO₄²⁻ and NO₃⁻ also represents secondary particles from power plants [18, 26, 53]. We have grouped all 15 PM_{2.5} components in this study into source categories as presented in Table 4. Based on this, we have calculated the sum of weights from WQS models on PM_{2.5} components for each source and also conducted WQS models directly on the mass of source groups. This provided us a clearer view of how the emission from each source contributing to the associations between components mixture and mortality. The high weights found for traffic emissions (include both tailpipe and non-tailpipe) is consistent with the intake fraction studies showing that traffic particles are more likely than average to be deposited in the lung because they are emitted closer to where people walk, live, and work [54, 55]. Therefore, the traffic particles may be more possible to induce adverse health outcomes and appear to have important contribution to the cumulative associations in WQS models, even if toxicity per unit concentration was the same. Besides the traffic emissions, the biomass burning and the secondary particles from power plants, which is mostly coal combustion, are also identified as important source groups. For the biomass burning, its marker component K has the highest weights among all components, although the mass of K is a small fraction of the mass of biomass combustion emissions. For the secondary particles from power plants, SO₄²⁻ is the major marker with the largest individual mortality rate ratio per IQR increase among the 5 major components.

The unexpectedly high weights for K estimated from WQS models is another issue of attention, which also indicate a need of further exploration of this component. This could be the result from possible measurement errors, residual confounding, and some other unclear reasons specifically related to K. In the future, it is worthwhile to conduct deeper investigation into it.

Several previous studies have also examined the health links of PM_{2.5} emitted from different sources. Most of these studies have identified the association between

long-term exposure to traffic-related particle components with adverse health outcomes, like mortality [56, 57], coronary heart disease[57], ischemic heart disease and stroke[58]. A nationwide cohort study in the U.S. reported a strong association between coal combustion particles with ischemic heart disease mortality [59]. Our study results suggest that both biomass burning and traffic-related emissions (both tailpipe and non-tailpipe) play an important role in PM_{2.5}'s association with mortality. These results may also suggest general linkages between sources indicative of combustion related activities and mortality since it is in common of all the sources identified with an important roles. Previous studies have also shown strong heterogeneity in PM_{2.5} emission source profiles across different regions [16, 53]. Therefore, our results may also only represent the exposure situation in Massachusetts.

There is heterogeneity in previous studies examining the associations between long-term exposures to PM_{2.5} components and mortality. Several multicity studies based on the ELAPSE (Effects of Low-level Air Pollution: A Study in Europe) and ESCAPE (European Study of Cohorts for Air Pollution Effects) projects have examined Cu, Fe, K, Ni, Si, V, and Zn. Some earlier studies have found null associations between these components and mortality with heterogeneities across different regions [18, 19], but a recently published study based on the ELAPSE project authors have reported positive associations between Cu, Fe, K, Ni, Si, V, Zn and mortality [17]. Some other studies have also supported associations between Zn, Si and mortality [21, 22, 26]. For Cu, Fe, and K, the previous studies have yielded more heterogeneous results, with some presenting positive associations [21, 26, 27] and some presenting null findings [22, 23, 27]. For the major components, the heterogeneity in previous studies still exist. A paper based on California Teacher Study has reported positive associations between EC, OC, SO₄²⁻, NO₃⁻ with all-cause mortality and cardiopulmonary mortality[21]. However, there are also studies reported null associations between those components with mortality [22, 23]. Some studies have examined NO₃⁻, SO₄²⁻, NH₄⁺ together as secondary inorganic aerosols (SIA) and shown associations with cardiovascular mortality but not with all-cause and respiratory mortality [24, 25]. Our study also reported a general pattern of greater mortality rate ratio magnitudes for cardiovascular mortality than all non-accidental and respiratory mortality. The heterogeneities in previous PM_{2.5} components studies may come from different levels of correlation amongst the individual components, different component profiles, meteorological conditions, population characteristics, prediction accuracies of exposure models, and the infiltration of air indoor[13]. The overall exposure differences between locations, different

exposure measurement method and data availabilities, different demographic characteristics, housing stock, and commuting patterns may also contribute to the heterogeneities. With these issues, the generalizability of the studies will be limited.

This is one of the few studies addressing health links of $PM_{2.5}$ components and thus inevitably bears several limitations. The first limitation concerns our mortality rate estimates. We have used the annual total population of each Census tract (>age 18) to approximate the annual person-year at risk at that Census tract. The total population data from 2001 to 2008 was linearly interpolated from U.S. Decennial Census and American Community Survey estimates, which is also an approximation of the actual population counts. These approximations could lead to inaccuracies in the mortality rate estimates. Second, although we have applied a high spatial resolution prediction model, there could still be exposure measurement errors come from lack of indoor air pollution measurements, commuting across different Census tracts, lack of information on residential history, and the residual prediction error of our model. We would expect this kind of exposure measurement error to be non-differential, and lead to a bias towards null though [60]. Moreover, the racial make-up of our study population is also a limitation to the results generalizability. Our study population is highly homogeneous with 93% non-Hispanic white, while non-Hispanic Whites are only about 70% in the whole population of Massachusetts and about 60% in the whole U.S. according to the most recently data [61, 62]. Non-Hispanic Whites are older than the population as a whole, hence their higher representation among all deaths. Different racial make-up could be linked with different exposure patterns [63]. Therefore, there may be vulnerable populations that could not be well represented by this study.

There are also some limitations in our analysis models. In WQS models, we have not incorporated the non-linearities of confounders in the WQS models, which could be a source of residual confounding. The interactions among different components in the mixture also have not been considered in WQS models [29]. Moreover, the source group we have used was based on a priori identified categories but not a formal source apportionment, which may have limitation in accuracy. Our secondary analysis of individual component models also has limitation in controlling the confounding by total $PM_{2.5}$ mass and other co-existing components [64]. It is worthwhile to explore this further in future studies.

This study also has advantages. In comparison with previous studies, one major advantage of our study is the application of a new exposure prediction model with high spatial resolution and prediction accuracy covering a long period from 2000 to 2015. These models

have yielded very high prediction accuracies. Therefore, although there were still residual exposure measurement errors, we could still capture the exposure with less errors than most previous simpler air pollution models. Another advantage of our study is that we have stronger study power by incorporating all deaths in Massachusetts from 2000 to 2015. Many previous studies on long-term $PM_{2.5}$ components exposures were cohort studies that identified several thousand deaths during the follow-up [18, 23, 24], or up to tens of thousands of cases [17, 25]. In our study, we have included nearly 800 thousand deaths, which provided a much stronger study power than the previous studies. Finally, we have applied a mixture analysis in our study treating the $PM_{2.5}$ components as a mixture people simultaneously exposed to. Using the WQS models, we could reduce the impact of collinearity of highly correlated components, and better estimate the cumulative associations and the relative importance of each component contributing to the associations in mixture analysis settings.

Conclusion

The WQS results showed strong positive association between the $PM_{2.5}$ components mixture with all three types of mortality, with Fe, OC, and K identified as potential important components, and biomass burning, traffic emission, and secondary particles from power plants as potential important sources having high contribution to the cumulative associations. The mortality rate ratios for cumulative associations were of greater magnitude for cardiovascular mortality than all non-accidental mortality and respiratory mortality. This pattern was also observed in total $PM_{2.5}$ analysis and single component models.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12940-022-00907-2>.

Supplementary Material 1

Acknowledgements

Heresh Amini is supported by Novo Nordisk Foundation Challenge Programme (NNF17OC0027812).

Author contribution

J.S. designed research and directed its implementation; T.J., H.A., A.K., M.Y., Y.W., E.C., Q.D. and L.S. prepared datasets; T.J. analyzed data, made the figures and tables, and wrote the paper; and all authors contributed to the revision of the manuscript.

Funding

This work was supported by the United States Environmental Protection Agency [grant number RD-83587201]; and the National Institutes of Health [grant numbers P30 ES000002, R01 ES032418-01]. The contents of this publication are solely the responsibility of the grantee and do not necessarily represent the official views of the US EPA. Further, the US EPA does not

endorse the purchase of any commercial products or services mentioned in the publication.

Data Availability

The mortality data has to been obtained from the Massachusetts Department of Public Health, and the applicants must complete the application process on their own to obtain it. The PM_{2.5} components concentration data are generated from the exposure models developed by Heresh Amini (heresh.amini@sund.ku.dk), and the personnel who would like to access this data may can directly contact Heresh Amini. The covariates data are available in the IPUM data service, the gridMET dataset, and the MEDPAR files of the Center for Medicare and Medicaid Services.

Declarations

Ethics approval and consent to participate

This study has been approved by the human subjects committee of Harvard T.H. School of Public Health.

Consent for publication

Not applicable.

Competing interests

Joel Schwartz has been an expert witness for the US Department of Justice in cases involving Clean Air Act violations.

Author details

¹Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

²Department of Public Health, University of Copenhagen, Copenhagen, Denmark

³Department of Family, Population, & Preventive Medicine, Program in Public Health, Stony Brook University, New York, NY, USA

⁴Vanke School of Public Health, Tsinghua University, Beijing, China

⁵Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

Received: 17 June 2022 / Accepted: 2 October 2022

Published online: 11 October 2022

References

- Dai L, Zanobetti A, Koutrakis P, Schwartz JD. Associations of Fine Particulate Matter Species with Mortality in the United States: A Multicity Time-Series Analysis. *Environ Health Perspect.* 2014;122(8):837–42.
- Zanobetti A, Schwartz J. Air pollution and emergency admissions in Boston, MA. *Journal of epidemiology and community health* (1979). 2006;60(10):890–5.
- Qiu X, Wei Y, Wang Y, Di Q, Sofer T, Awad YA, et al. Inverse probability weighted distributed lag effects of short-term exposure to PM_{2.5} and ozone on CVD hospitalizations in New England Medicare participants - Exploring the causal effects. *Environ Res.* 2020;182:109095–5.
- Kloog I, Ridgway B, Koutrakis P, Coull BA, Schwartz JD. Long- and Short-Term Exposure to PM_{2.5} and Mortality: Using Novel Exposure Models. *Epidemiol (Cambridge Mass).* 2013;24(4):555–61.
- Atkinson RW, Kang S, Anderson HR, Mills IC, Walton HA. Epidemiological time series studies of PM_{2.5} and daily mortality and hospital admissions: a systematic review and meta-analysis. *Thorax.* 2014 Jul 1;69(7):660–5.
- Shi L, Zanobetti A, Kloog I, Coull BA, Koutrakis P, Melly SJ, et al. Low-Concentration PM_{2.5} and Mortality: Estimating Acute and Chronic Effects in a Population-Based Study. *Environ Health Perspect.* 2016;124(1):46–52.
- Beelen R, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, Hoffmann B, et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. *The Lancet (British edition).* 2014;383(9919):785–95.
- Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME, et al. An Association between Air Pollution and Mortality in Six U.S. Cities. *N Engl J Med.* 1993;329(24):1753–9.
- Awad YA, Di Q, Wang Y, Choirat C, Coull BA, Zanobetti A, et al. Change in PM_{2.5} exposure and mortality among Medicare recipients: Combining a semi-randomized approach and inverse probability weights in a low exposure population. *Environ Epidemiol.* 2019;3(4):e054–4.
- Schwartz JD, Wang Y, Kloog I, Yitshak-Sade M, Dominici F, Zanobetti A. Estimating the Effects of PM_{2.5} on Life Expectancy Using Causal Modeling Methods. *Environ Health Perspect.* 2018;126(12):127002–2.
- Lepeule J, Laden F, Dockery D, Schwartz J. Chronic Exposure to Fine Particles and Mortality: An Extended Follow-up of the Harvard Six Cities Study from 1974 to 2009. *Environ Health Perspect.* 2012;120(7):965–70.
- Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. Estimating Causal Effects of Long-Term PM_{2.5} Exposure on Mortality in New Jersey. *Environ Health Perspect.* 2016;124(8):1182–8.
- Hoek G, Krishnan RM, Beelen R, Peters A, Ostro B, Brunekreef B, et al. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ Health.* 2013;12(1):43–3.
- Hart JE, Liao X, Hong B, Puett RC, Yanosky JD, Suh H, et al. The association of long-term exposure to PM_{2.5} on all-cause mortality in the Nurses' Health Study and the impact of measurement-error correction. *Environ Health.* 2015;14(1):38–8.
- Wang Y, Shi L, Lee M, Liu P, Di Q, Zanobetti A, et al. Long-term Exposure to PM_{2.5} and Mortality Among Older Adults in the Southeastern US. *Epidemiol (Cambridge Mass).* 2017;28(2):207–14.
- Integrated Science Assessment for Particulate Matter. Environmental Protection Agency Documents and Publications. Washington: Federal Information & News Dispatch, LLC; 2020.
- Chen J, Rodopoulou S, de Hoogh K, Strak M, Andersen ZJ, Atkinson R, et al. Long-Term Exposure to Fine Particle Elemental Components and Natural and Cause-Specific Mortality—a Pooled Analysis of Eight European Cohorts within the ELAPSE Project. *Environ Health Perspect.* 2021;129(4):47009–9.
- Wang M, Beelen R, Stafoggia M, Raaschou-Nielsen O, Andersen ZJ, Hoffmann B, et al. Long-term exposure to elemental constituents of particulate matter and cardiovascular mortality in 19 European cohorts: Results from the ESCAPE and TRANSPHORM projects. *Environ Int.* 2014;66:97–106.
- Beelen R, Hoek G, Raaschou-Nielsen O, Stafoggia M, Andersen ZJ, Weinmayr G, et al. Natural-cause mortality and long-term exposure to particle components: An Analysis of 19 European cohorts within the multi-center ESCAPE project. *Environ Health Perspect.* 2015;123(6):525–33.
- Chung Y, Dominici F, Wang Y, Coull BA, Bell ML. Associations between Long-Term Exposure to Chemical Constituents of Fine Particulate Matter (PM_{2.5}) and Mortality in Medicare Enrollees in the Eastern United States. *Environ Health Perspect.* 2015;123(5):467–74.
- Ostro B, Lipsett M, Reynolds P, Goldberg D, Hertz A, Garcia C, et al. Long-Term Exposure to Constituents of Fine Particulate Air Pollution and Mortality: Results from the California Teachers Study. *Environ Health Perspect.* 2010;118(3):363–9.
- Yang Y, Ruan Z, Wang X, Yang Y, Mason TG, Lin H, et al Short-term and long-term exposures to fine particulate matter constituents and health: A systematic review and meta-analysis. *Environmental pollution* (1987). 2019;247:874–82.
- Ostro B, Hu J, Goldberg D, Reynolds P, Hertz A, Bernstein L, et al. Associations of Mortality with Long-Term Exposures to Fine and Ultrafine Particles, Species and Sources: Results from the California Teachers Study Cohort. *Environ Health Perspect.* 2015;123(6):549–56.
- Sommar JN, Hvidtfeldt UA, Geels C, Frohn LM, Brandt J, Christensen JH, et al. Long-Term Residential Exposure to Particulate Matter and Its Components, Nitrogen Dioxide and Ozone—A Northern Sweden Cohort Study on Mortality. *Int J Environ Res Public Health.* 2021;18(16):8476–.
- Hvidtfeldt UA, Geels C, Sørensen M, Ketznel M, Khan J, Tjønneland A, et al. Long-term residential exposure to PM_{2.5} constituents and mortality in a Danish cohort. *Environ Int.* 2019;133:105268–8.
- Badaloni G, Cesaroni G, Cerza F, Davoli M, Brunekreef B, Forastiere F. Effects of long-term exposure to particulate matter and metal components on mortality in the Rome longitudinal study. *Environ Int.* 2017;109:146–54.
- Lavigne A, Sterrantino AF, Liverani S, Blangiardo M, de Hoogh K, Mollitor J, et al. Associations between metal constituents of ambient particulate matter and mortality in England: an ecological study. *BMJ open.* 2019;9(12):e030140–0.
- Lin H, Tao J, Du Y, Liu T, Qian Z, Tian L, et al Particle size and chemical constituents of ambient particulate pollution associated with cardiovascular mortality in Guangzhou, China. *Environmental pollution* (1987). 2016;208(Pt B):758–66.

29. Carrico C, Gennings C, Wheeler DC, Factor-Litvak P. Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting. *J agricultural Biol Environ Stat*. 2015;20(1):100–20.
30. Czarnota J, Gennings C, Wheeler DC. Assessment of Weighted Quantile Sum Regression for Modeling Chemical Mixtures and Cancer Risk. *Cancer Inform*. 2015;2015(S2):159–71.
31. Deyssenroth MA, Gennings C, Liu SH, Peng S, Hao K, Lambertini L, et al. Intrauterine multi-metal exposure is associated with reduced fetal growth through modulation of the placental gene network. *Environ Int*. 2018;120:373–81.
32. Romano ME, Eliot MN, Zoeller RT, Hoofnagle AN, Calafat AM, Karagas MR, et al. Maternal urinary phthalate metabolites during pregnancy and thyroid hormone concentrations in maternal and cord sera: The HOME Study. *Int J Hyg Environ Health*. 2018;221(4):623–31.
33. Frydas IS, Keremidou M, Tsave O, Salifoglou A, Sarigiannis DA. Unraveling the blood transcriptome after real-life exposure of Wistar-rats to PM2.5, PM1 and water-soluble metals in the ambient air. *Toxicol Rep*. 2020;7:1469–79.
34. Fang T, Guo H, Zeng L, Verma V, Nenes A, Weber RJ. Highly Acidic Ambient Particles, Soluble Metals, and Oxidative Potential: A Link between Sulfate and Aerosol Toxicity. *Environ Sci Technol*. 2017;51(5):2611–20.
35. Dai L, Mehta A, Mordukhovich I, Just AC, Shen J, Hou L, et al. Differential DNA methylation and PM2.5 species in a 450K epigenome-wide association study. *Epigenetics*. 2017;12(2):139–48.
36. GHIO AJ, DEVLIN RB. Inflammatory Lung Injury after Bronchial Instillation of Air Pollution Particles. *Am J Respir Crit Care Med*. 2001;164(4):704–8.
37. Carter JD, Ghio AJ, Samet JM, Devlin RB. Cytokine Production by Human Airway Epithelial Cells after Exposure to an Air Pollution Particle Is Metal-Dependent. *Toxicol Appl Pharmacol*. 1997;146(2):180–8.
38. Ghio AJ, Soukup JM, Case M, Dailey LA, Richards J, Bernsten J, et al. Exposure to wood smoke particles produces inflammation in healthy volunteers. Occupational and environmental medicine. (London England). 2012;69(3):170–5.
39. Kennedy T, Ghio AJ, Reed W, Samet J, Zagorski J, Quay J, et al. Copper-dependent Inflammation and Nuclear Factor-kappa B Activation by Particulate Air Pollution. *Am J Respir Cell Mol Biol*. 1998;19(3):366–78.
40. Samet JM, Graves LM, Quay J, Dailey LA, Devlin RB, Ghio AJ, et al. Activation of MAPKs in human bronchial epithelial cells exposed to metals. *Am J Physiol Lung Cell Mol Physiol*. 1998;275(3):551–8.
41. Amini H, Danesh-Yazdi M, Di Q, Requia W, Wei Y, AbuAwad Y, et al. Annual Mean PM2.5 Components (EC, NH4, NO3, OC, SO4) 50m Urban and 1km Non-Urban Area Grids for Contiguous U.S., 2000–2019, v1. Palisades: NASA Socioeconomic Data and Applications Center (SEDAC); 2022.
42. Amini H, Danesh Yazdi M, Di Q, Requia W, Abu Awad Y, Shi L, et al. Super-learning and ensemble weighted averaging models to predict hyperlocal long-term exposure to fine particulate matter components in the United States. *Environmental health perspectives Supplements*. 2021;2021(1).
43. Amini H, Danesh-Yazdi M, Di Q, et al. Hyperlocal super-learned PM2.5 components across the contiguous US. *Research Square*; 2022. DOI:<https://doi.org/10.21203/rs.3.rs-1745433/v1>.
44. Di Q, Amini H, Shi L, Kloog I, Silvern R, Kelly J, et al. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environ Int*. 2019;130:104909–9.
45. Abatzoglou JT. Development of gridded surface meteorological data for ecological applications and modelling. *Int J Climatol*. 2013;33(1):121–31.
46. Manson S, Schroeder J, Van Riper D, Kugler T, Ruggles S. IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. [Internet]. Minneapolis, MN: IPUMS; 2021. Available from: <https://doi.org/10.18128/D050.V16.0>.
47. Garcia CA, Yap PS, Park HY, Weller BL. Association of long-term PM2.5 exposure with mortality using different air pollution exposure models: impacts in rural and urban California. *Int J Environ Health Res*. 2016;26(2):145–57.
48. Wu X, Braun D, Schwartz J, Kioumourtzoglou MA, Dominici F. Evaluating the impact of long-term exposure to fine particulate matter on mortality among the elderly. *Sci Adv*. 2020;6(29):eaba5692–2.
49. Hales S, Atkinson J, Metcalfe J, Kuschel G, Woodward A. Long term exposure to air pollution, mortality and morbidity in New Zealand: Cohort study. *Sci Total Environ*. 2021;801:149660–0.
50. Eum KD, Kazemiparkouhi F, Wang B, Manjourides J, Pun V, Pavlu V, et al. Long-term NO2 exposures and cause-specific mortality in American older adults. *Environ Int*. 2019;124:10–5.
51. Kloog I, Zanobetti A, Nordio F, Coull BA, Baccarelli AA, Schwartz J. Effects of airborne fine particles (PM2.5) on deep vein thrombosis admissions in the northeastern United States. *J Thromb Haemost*. 2015;13(5):768–74.
52. Renzetti S, Curtin P, Just AC, Bello G, Gennings C. gWQS: Generalized Weighted Quantile Sum Regression. R package version 3.0.4. [Internet]. 2021. Available from: <https://CRAN.R-project.org/package=gWQS>.
53. Thurston GD, Ito K, Lall R. A source apportionment of U.S. fine particulate matter air pollution. *Atmospheric environment (1994)*. 2011;45(24):3924–36.
54. Taimisto P, Tainio M, Karvosenoja N, Kupiainen K, Porvari P, Karpinen A, et al. Evaluation of intake fractions for different subpopulations due to primary fine particulate matter (PM2.5) emitted from domestic wood combustion and traffic in Finland. *Air Qual Atmos Health*. 2011;4(3–4):199–209.
55. Tainio M, Sofiev M, Hujo M, Tuomisto JT, Loh M, Jantunen MJ, et al. Evaluation of the European population intake fractions for European and Finnish anthropogenic primary fine particulate matter emissions. *Atmospheric environment (1994)*. 2009;43(19):3052–9.
56. Beelen R, Hoek G, van den Brandt PA, Goldbohm RA, Fischer P, Schouten LJ, et al. Long-Term Effects of Traffic-Related Air Pollution on Mortality in a Dutch Cohort (NLCS-AIR Study). *Environ Health Perspect*. 2008;116(2):196–202.
57. Gan WQ, Koehoorn M, Davies HW, Demers PA, Tamburic L, Brauer M. Long-Term Exposure to Traffic-Related Air Pollution and the Risk of Coronary Heart Disease Hospitalization and Mortality. *Environ Health Perspect*. 2011;119(4):501–7.
58. Ljungman PLS, Andersson N, Stockfelt L, Andersson EM, Sommar JN, Eneroth K, et al. Long-Term Exposure to Particulate Air Pollution, Black Carbon, and Their Source Components in Relation to Ischemic Heart Disease and Stroke. *Environ Health Perspect*. 2019;127(10):107012–2.
59. Thurston GD, Burnett RT, Turner MC, Shi Y, Krewski D, Lall R, et al. Ischemic Heart Disease Mortality and Long-Term Exposure to Source-Related Components of US Fine Particle Air Pollution. *Environ Health Perspect*. 2016;124(6):785–94.
60. Wacholder S, Hartge P, Lubin JH, Dosemeci M. Non-differential misclassification and bias towards the null: a clarification. *Occupational and environmental medicine*. (London England). 1995;52(8):557–8.
61. U.S. Census Bureau QuickFacts. Massachusetts [Internet]. [cited 2022 May 30]. Available from: <https://www.census.gov/quickfacts/MA>.
62. U.S. Census Bureau. QuickFacts: United States [Internet]. [cited 2022 May 30]. Available from: <https://www.census.gov/quickfacts/fact/table/US/PST045221>.
63. Kravitz-Wirtz N, Crowder K, Hajat A, Sass V. THE LONG-TERM DYNAMICS, OF RACIAL/ETHNIC INEQUALITY IN NEIGHBORHOOD AIR POLLUTION EXPOSURE. *Du Bois review*. 2016;13(2):1990–2009. 237–59.
64. Mostofsky E, Schwartz J, Coull BA, Koutrakis P, Wellenius GA, Suh HH, et al. Modeling the Association Between Particle Constituents of Air Pollution and Health Outcomes. *Am J Epidemiol*. 2012;176(4):317–26.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.