HHS Public Access

Author manuscript

Environ Int. Author manuscript; available in PMC 2015 April 02.

Published in final edited form as:

Environ Int. 2014 October ; 71: 13–19. doi:10.1016/j.envint.2014.05.023.

Health effects of multi-pollutant profiles

Antonella Zanobetti[1], **Elena Austin**[1], **Brent A. Coull**[2], **Joel Schwartz**[1], and **Petros Koutrakis**[1]

[1]Department of Environmental Health, Harvard School of Public Health, Boston

[2]Department of Biostatistics, Harvard School of Public Health, Boston

Abstract

Background—The association between exposure to particle mass and mortality is well established; however, there are still uncertainties as to whether certain chemical components are more harmful than others. Moreover, understanding the health effects associated with exposure to pollutants mixtures may lead to new regulatory strategies.

Objectives—Recently we have introduced a new approach that uses cluster analysis to identify distinct air pollutant mixtures by classifying days into groups based on their pollutant concentration profiles. In Boston during the years 1999–2009, we examined whether the effect of $PM_{2.5}$ on total mortality differed by distinct pollution mixtures.

Methods—We applied a time series analysis to examine the association of $PM_{2.5}$ with daily deaths. Subsequently, we included an interaction term between $PM_{2.5}$ and the pollution mixture clusters.

Results—We found a 1.1 % increase (95% CI: 0.0, 2.2) and 2.3% increase (95% CI: 0.9–3.7) in total mortality for a 10 μg/m^3 increase in the same day and the two-day average of $PM_{2.5}$ respectively. The association is larger in a cluster characterized by high concentrations of the elements related to primary traffic pollution and oil combustion emissions with a 3.7% increase (95% CI: 0.4, 7.1) in total mortality, per 10 μg/m^3 increase in the same day average of $PM_{2.5}$.

Conclusions—Our study shows a higher association of $PM_{2.5}$ on total mortality during days with a strong contribution of traffic emissions, and fuel oil combustion. Our proposed method to create multi-pollutant profiles is robust, and provides a promising tool to identify multi-pollutant mixtures which can be linked to the health effects.

Keywords

Total mortality; fine particulate air pollution; pollutant mixtures

1. Introduction

There is mounting evidence that long- and short-term exposures to ambient air pollution are associated with both acute and chronic mortality risk (Katsouyanni et al. 1997; Samet et al.

Address for correspondence: Antonella Zanobetti, Department of Environmental Health, Exposure Epidemiology and Risk Program, Harvard School of Public Health, 401 Park Drive, Landmark Center, Suite 415, P.O. Box 15698, Boston, MA 02215, Phone: 617 384-8751, Fax: 617 384-8745, azanobet@hsph.harvard.edu.

2000; Samoli et al. 2008; Zanobetti and Schwartz 2009), but there is only limited data on the effects of pollutant mixtures.

The importance of describing, understanding, and regulating multi-pollutant mixtures has been highlighted by the US National Academy of Science (NAS) (NRC, 2004) and the Environmental Protection Agency (EPA). The importance of developing a multi-pollutant air quality management plan is discussed in their Multi-Pollutant Report of 2008 (U.S. EPA, 2008).

One limitation in examining $PM_{2.5}$ components is the irregularity of the $PM_{2.5}$ speciation data available from the EPA Speciation Trends Network (STN) monitoring sites. The STN was established in 2000 and reports data for every third or sixth day, but almost all work using these data are based single pollutant models evaluating associations between health and each element one-at-a time. Consequently, there is still uncertainty as to which components are most harmful. Across different urban areas, air quality in the US may differ in levels and in the composition of pollutants. An understanding of common profiles of urban air quality can potentially lead to better air quality management; moreover, furthering our understanding of the health effects associated with exposure to pollutant mixtures may lead to new regulatory strategies.

Developing a multi-pollutant approach is extremely challenging due to the highly complex interactions between source emissions, atmospheric processes and effects on human health and ecosystems. Investigating the multivariate relationship between pollutants at a given site will enhance our understanding of the interaction between pollutants as well as the human health effects related to these complex mixtures.

The EPA has considered a variety of ways in which air pollutants might jointly affect health, including additive, multiplicative and antagonistic effects. (Mauderly et al. 2010). Populations are exposed daily to complex mixtures of pollutants, some of which are known or suspected to cause health effects at ambient concentrations. Understanding the effect of the mixtures on health, rather than the effect of the individual components is a crucial step that must be undertaken in order to further our knowledge of this field.

A limited number of studies have examined the effect of multi-pollutant mixtures, as independent effects of individual pollutants within mixtures are often effectively unobservable due to highly collinear components. Some studies have examined the health effects of individual species. In a multi-site time series study of 119 US communities during the period 2000 to 2005, Peng et al (Peng et al. 2009) found that ambient levels of Elemental Carbon (EC) and Organic Carbon (OC) were associated with the largest risks of emergency hospitalization across the major chemical constituents of fine particles. Ostro et al (Ostro et al. 2007) examined the associations between 19 $PM_{2.5}$ components and daily mortality in six California counties, and found that $PM_{2.5}$ mass and EC, OC, NO_3^-, Fe, K, and Ti were associated with cardiovascular deaths. Franklin and co-authors (Franklin et al. 2008), and Zanobetti and co-authors (Zanobetti et al. 2009) used a hierarchical approach to determine whether the association between daily $PM_{2.5}$ mass and mortality and hospitalizations, respectively, were modified by $PM_{2.5}$ composition in 25 US communities. They found that

effect estimates for $PM_{2.5}$ mass and cardiovascular hospital admissions were higher when the $PM_{2.5}$ content of Br, Cr, Ni, or Na^+ was higher. In a similar study, Bell et al (Bell et al. 2009) found a different set of effect modifiers (EC, V, or Ni).

Furthermore, studies have examined the effects of sources using source apportionment methods (Laden et al. 2000; Mar et al. 2000; Thurston et al. 2005; Sarnat et al. 2008; Ostro et al. 2011) The source apportionment methods group pollutants according to how their daily concentrations co-vary, presumably due to emission from common sources.

Recently, we have introduced a new approach that uses cluster analysis to identify distinct air pollutant mixtures (Austin et al. 2012). While cluster analysis has been mostly used to group pollutants based on their source origins, in this paper cluster analysis was used to classify sampling days into groups based on their pollutant concentration profiles.

While the source apportionment method groups pollutants according to how their daily concentrations co-vary, the cluster method groups days with similar multi-pollutants profile. In the present study, the identified pollutant profiles or clusters are applied in a health effect study examining the effect of $PM_{2.5}$ mass on total mortality in the greater Boston area.

2. Data and Methods

2.1 Health Data

We obtained individual mortality data for Boston, MA from the National Center for Health Statistics (NCHS) for the years 1999 to 2006, and from the Massachusetts Department of Public Health for the years 2006–2009. Boston was defined as the “greater Boston area” which included the counties of Middlesex, Norfolk and Suffolk.

The mortality files provided information on the exact date of death, and the underlying cause of death. Our outcome was all-cause non accidental daily mortality (ICD-9: 0–799). We focused on this cause of mortality in order to yield a sufficiently high number of deaths per day for adequate statistical power.

2.2 Air quality data

$PM_{2.5}$ (particles with an aerodynamic diameter ≤2.5μm), PN (particle number), and BC (black carbon) were collected at the Harvard Supersite in Boston, MA. The site is located on the roof of the Countway Library of the Harvard Medical School near downtown Boston, within one block of a four-lane street with truck traffic and with two major highways nearby: Interstate 90 (I-90) is approximately 1.5 km to the north and Interstate 93 (I-93) is approximately 3 km to the south. Data used in this analysis were collected between October 15, 1999 and December 31, 2009.

We computed the two-day moving average of $PM_{2.5}$ concentrations as the average of the same and previous days of $PM_{2.5}$ in the complete time series; if one day is missing then the two days average is missing. Daily integrated $PM_{2.5}$ samples were collected on Teflon filters using Harvard Impactors and were analyzed for elements by X-Ray Fluorescence at the Harvard School of Public Health.

Particle number was measured hourly by a condensation particle counter (CPC, TSI Inc., Model 3022A, Shoreview, MN). Finally, an Aethalometer (Magee Scientific, Berkeley, CA) was used to measure black carbon.

The concentrations of the gaseous pollutants including nitrogen oxides (NO and NO_2), sulfur dioxide (SO_2) and ozone (O_3) were obtained from the Greater Boston area monitoring sites, and exposures were estimated by averaging data from the available sites.

The pollutant measures used in the clustering analysis were PN, CO, NO, NO_2, O_3, SO_2, BC, and particulate S, Cu, Fe, Zn, Ni, V, Ti, Mg, K, Si, Na, Cl, Ca, Br, Sr, Pb, and Mn. Other elements obtained as part of the speciation analysis of the filters were considered as possible clustering variables but were excluded either because the analytical measurement was judged to be unreliable or the element had a large proportion of the measurements below the method detection limit. From the total 4124 possible study days, there were 1186 days excluded from the analysis because they had missing data in the pollutants of interest (259 days with missing elemental data, 306 days with missing PN data, 55 days missing both elemental and PN data, 81 days missing BC data, 9 days missing elemental and BC data, 137 days missing PN and BC data and, 158 missing PN, BC and elemental data and 75 days missing gas data). In addition, 93 days were excluded from the analysis because they contained one or more species whose measured values were six standard deviations away from their respective means. The $PM_{2.5}$ mass was not a variable used in the clustering, although it was later used to subjectively interpret the resulting clusters of days.

We obtained local meteorological data measured at the Logan airport station, including temperature and dew point temperature, from the National Oceanic and Atmospheric Administration (NOAA).

2.3 Clusters

Using the method described by Austin et al. (Austin et al. 2012), days were grouped into 5 clusters corresponding to distinct pollution profiles. Austin et al. (Austin et al. 2012)clustered Boston data collected between 2004 and 2009. Because additional years of XRF elemental data have since become available, the current analysis spans the time period 1999–2009. In addition, the XRF analysis was redone on all available data so there is consistency across the entire data set. Due to the improved precision of the analytical method we were able to include more elements in the cluster analysis than in Austin et al. The goal of the clustering was to produce groups of days with similar multi-pollutant profiles. For each cluster of days we estimated pollutant concentrations and elemental ratios to better characterize differences among the 5 groups. In addition, because pollutant concentration relationships should be similar on days with similar local meteorological conditions, for each cluster we estimated the mean values of different meteorological parameters.

Prior to running the clustering algorithm, each $PM_{2.5}$ constituent was normalized using a modified Z-score (median) transformation, in order to prevent bias due to widely varying scales of the individual variables. The advantage of using a modified Z-score is that outlier

values have less influence on the clustering results. The clustering algorithm was applied to the modified z-score values.

The method used to produce the clustering is briefly described below. The Hartigan and Wong (Hartigan and Wong 1979) k-means algorithm was used for the analysis. The Hartigan-Wong algorithm minimizes the within cluster sum of squares and uses this measure of dissimilarity to categorize the observations. In order to determine the prevailing origin of air-masses for each identified cluster, the NOAA HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) model was used to plot the frequency distribution of the hourly backwards trajectories (up to 84 hours) of the cluster members from 2004–2009 (Rolph and Draxler 1990; Rolph 2003). The model used to predict the back trajectories was computed on a 40-km resolution grid for which data is available from 2004 onwards. Therefore, days prior to 2004 were not included in this analysis.

Normalized concentrations were calculated in order to better compare the clusters. These normalized concentrations represent the enrichment of a given constituent (element) of $PM_{2.5}$ within a cluster as compared to the entire sample

Normalized Concentration

$$NC_{ij} = \frac{\overline{S}}{\overline{PM_{2.5}}} \div \frac{\overline{S}}{\overline{PM_{2.5}}}$$

Where: NC_{ij} represents the Normalized Concentration of species i in cluster j

S_{ij} represents the mean Species Concentration j (Fe, OC, Na^+, etc.) in cluster i

S_j represents the mean Species Concentration over all clusters

$PM_{2.5}$ represents the concentration of $PM_{2.5}$

For each cluster we plotted back-trajectory points that were between 0 and 1,000 meters above ground on wind rose graphs with 8 quadrants corresponding to the direction of the air mass trajectories with respect to Boston. We restricted the plots to trajectories that were 24–48 hours prior to the sample day. The results were plotted using the heR.misc R package developed by (Klepeis 2004). The distance of the trajectory with respect to Boston was displayed using colored bars on each radial segment. These findings were used to validate the results of the cluster analysis as well as to investigate the predominant direction of the air-masses within each cluster of days.

Overall, this analysis yielded a solution that was both robust to outlier points and interpretable based on chemical, physical and meteorological characteristics.

2.4 Analytical strategy

We investigated the association between the same day $PM_{2.5}$ and total mortality with a time series analysis. We first applied a Poisson generalized additive model, controlling for long-term trend and seasonality with a natural cubic regression spline with 5 degrees of freedom for each year; day of the week using indicator variables; and for weather using a natural

cubic spline with three degrees of freedom for the same and previous day temperature and for dew point temperature.

In order to determine whether multi-pollutant profiles characterized by the clusters modifies the effect of $PM_{2.5}$ in this model, we then included an interaction term between $PM_{2.5}$ (same day) and the categorical variable for each cluster (same day) in our model. Upon centering $PM_{2.5}$ at its mean, the model is:

$$Cases_t \sim Poisson(\exp\{\beta_0 + f(temp_t) + g(temp_{t-1}) + h(dewp_t) + k(season_t) + \ldots + \gamma_1 clusters_2 + \ldots + \gamma_4 clusters_5 + \beta_1 PM_{25} + \delta_1 PM_{25} * clusters_2 + \ldots + \delta_4 PM_{25} * clusters_5\})$$

Where *f, g, h* and *k* are the smoothing functions of same day temperature, previous day temperature, same day dew point, and seasonality, respectively, $\gamma_1 \ldots \gamma_4$ are the main effects of each cluster (cluster 1 is the reference) at the average $PM_{2.5}$ level (due to centering $PM_{2.5}$), the $\delta_1 .. \delta_4$ are the differences between the $PM_{2.5}$ effect in cluster 1 and cluster 2–5, respectively; and β_1 is the main effect of $PM_{2.5}$, which represents the effect of $PM_{2.5}$ in cluster 1. We then computed the $PM_{2.5}$ effect in each of the five clusters by summing β_1 and each δ; for example the $PM_{2.5}$ effect in cluster 2 is: $\beta_1 + \delta_1$; with standard error:

$$\sqrt{\mathrm{var}(\beta_1) + \mathrm{var}(\delta_1) + 2\mathrm{cov}(\delta_1, \beta_1)}$$

Because it has been previously reported that the two days average of $PM_{2.5}$ is more strongly associated with mortality than same day $PM_{2.5}$, we also investigated the association between total mortality and the two days average $PM_{2.5}$, and examined whether the cluster variable, derived by applying the clustering algorithm to the two-day averages, modified the effect of the two days $PM_{2.5}$ with the same model described above.

As sensitivity analysis we tested whether differences in effects across clusters could be driven by differences in the effects across seasons by adding a main effect of season and a season* $PM_{2.5}$ interaction.

The data were analyzed using R 2.15.1 (http://www.R-project.org). The effect estimates are expressed as a percent increase in mortality for a 10 $\mu g/m^3$ increase in $PM_{2.5}$ mass concentration.

3. Results

Table 1 presents the means, standard deviations and number of observations for total mortality, $PM_{2.5}$ exposure, and weather variables for years 1999–2009, in total and by cluster. $PM_{2.5}$ concentrations were low, with an average of 10 $\mu g/m^3$, and varied by cluster with concentrations in cluster 4 (“Regional Summer”) being the highest (Figure 1). Clusters were missing in 1186 days over the 11 years period. Table S1 in the supplemental material presents the frequency distribution of clusters by season.

We selected a solution with 5 clusters to describe the Boston data from 1999–2009. This was the most parsimonious solution that minimized the ratio of the within cluster to between cluster variability in the multivariate pollutant vector (SSW/SSB) (Figure 2). After

examining the 4 cluster and 6 cluster solutions, the 5 cluster solution was the most interpretable based on weather and chemical characteristics. Summary statistics for each of the clusters are presented in Table 2. Some elements cluster means are negative due to small negative values being reported when the concentration on the filter is below the limit of detection and lower than those measured on a blank filter. We retain these negative values in the dataset so as to not alter the distribution profile of the elements. A small negative mean value suggests that a cluster has no measurable concentration for that element. Important pollutant ratios by cluster are presented in Table 3. Table 4 presents the mean pollutant normalized concentrations by cluster as compared to the entire dataset.

The clusters obtained can be described based on their chemical composition, the ratio of species/$PM_{2.5}$ and important species ratios, weather patterns, and seasonal distribution. Cluster 1, which we termed “Low Particles – High Ozone” occurred mostly during the spring and mid-fall and was characterized by low $PM_{2.5}$ concentrations, high normalized concentrations for O_3 and above average normalized concentration of NO_2 and PN. The BC/$PM_{2.5}$ ratio is quite low in this cluster, as is the NO/NO_2 ratio, suggesting low impact of primary traffic and combustion sources. Cluster 2, which we termed “Crustal” occurred less frequently in colder months. This cluster is particularly enriched in Si, Ca, Br and Ti. Fe was also elevated in this cluster. These elements, when present together indicate that a larger proportion of the mass observed on these days is attributable to crustal particles. Cluster 3, which we termed “Winter – Primary” was enriched in the elements Ni, V, Zn and Br and the gases CO, NO and SO_2. It also has high NO/NO_2 and BC/$PM_{2.5}$ ratios. These all indicate the strong contribution of traffic and oil combustion emissions on these days. The composition of days in cluster 4, which we termed “Regional Summer”, was dominated by transported regional pollution based on the high S, $PM_{2.5}$ and Se concentrations and lower PN/$PM_{2.5}$ ratio (used as a particle size indicator with a lower value indicating a relatively larger particle size). It occurred almost entirely during the warm season. Cluster 5, which we termed “Winter – Low Primary, Higher O_3” is similar to cluster 3 (Winter – Primary) in general composition, although it has significantly lower $PM_{2.5}$ mass and higher O_3. It occurred mostly in the winter months and had the lowest average daily temperature. The particle mass is characterized by higher composition in Ni, V and K, and a higher PN/$PM_{2.5}$ ratio, suggesting above average contributions of combustion sources such as oil and wood burning. This cluster was also enriched in CO, NO, NO_2, SO_2, which is consistent with primary combustion emissions. The Fe/Si ratio in cluster 5 was 1.0 and the concentrations of NO and NO_2 were lower than cluster 3 (Winter – Primary), which indicates a lower traffic contribution in this cluster. Although both clusters 3 and 5 have high normalized concentrations of Ni and V, which are associated with the heating oil combustion, cluster 5 has a much higher size index (PN/$PM_{2.5}$).This suggests that clusters 3 and 5 are composed of different mixtures of primary pollutants, and have important differences in the size distribution of the particles.

Back-trajectory analysis of the different clusters (from 2004–2009), similar to what was performed in the original Austin et al (2012) paper, suggest that cluster 4 is associated with air masses originating from the SW of Boston, while cluster 3 (Winter – Primary) is associated with air masses originating from the W of Boston (Figure 3). Each of these two clusters have more than 50% of the hourly back-trajectories within that cluster originating in

the same direction. Other clusters show less association between the cluster type and air mass direction.

We performed different sensitivity analyses to ensure that the clustering produced was meaningful and robust to changes in the initial dataset. We compared the solution obtained here for the years 1999–2009, to that reported in Austin et al. (Austin et al. 2012). Although the earlier clustering analysis included fewer elements and only started in 2004, the results are fairly similar (72% of days appearing in both analyses were classified into the same cluster). We also performed a sensitivity test to determine whether removing 10% of the observation days resulted in significant differences in the clustering result. We created 100 new datasets, with 10% of the days missing at random. The clustering results were highly comparable, with an average of adjusted Rand Index of 0.9 (sd 0.1). The Rand Index is a measure of similarity between two different partitions of the same data set; it ranges between 0 and 1 where 0 indicates that two data clusters do not agree on any pair of points and 1 indicating that the data clusters are identical; values greater than 0.9 reflect excellent agreement.

Figure 4 presents the estimated associations between $PM_{2.5}$ and mortality in total and by cluster. Across all days, we found a 1.1% increase (95% CI: 0.0, 2.2) in total mortality for 10 $\mu g/m^3$ increase in the same day $PM_{2.5}$. When looking at the effect modification by cluster, the association of the same day $PM_{2.5}$ on total mortality is significant in cluster 3 (Winter – Primary) with a 3.7% increase (95% CI: 0.4, 7.1) in total mortality (Figure 4). An overall test comparing the models with and without the interactions of $PM_{2.5}$ with the clusters was marginally significant (P-value=0.088).There was strong evidence that the effect of $PM_{2.5}$ in cluster 3 was larger (P-value=0.045) than that in cluster 1 as tested by the interaction term, with no evidence of differences in effect estimates across the other clusters.

When we examined the two day average $PM_{2.5}$ across all days we found a 2.3% increase (95% CI: 0.9, 3.7) in total mortality for 10 $\mu g/m^3$ increase in the two-days average of $PM_{2.5}$. The association of the two days average of $PM_{2.5}$ on total mortality is significant in cluster 3 (Winter – Primary) with a 4.8% increase (95% CI: 0.8, 9.0) in total mortality, in cluster 4 (Regional summer) with a 3.4% increase (95% CI: 0.4, 6.6), and in cluster 5 (Aged winter) with a 3.7% increase (95% CI: 0.1, 7.4), in total mortality for 10 $\mu g/m^3$ increase in the two-days average of $PM_{2.5}$. The overall 4 degree of freedom test of an interaction between $PM_{2.5}$ and the clusters was not significant.

As mentioned above, cluster 3 (Winter – Primary) represents days with higher levels in Ni, V, Zn, Br, CO, NO and SO_2. The days in this cluster are impacted by fresh traffic and oil combustion (possibly domestic heating) emissions.

In sensitivity analysis we found that adjusting for effect modification by season didn't change our results (Figure S1 supplemental material) and, therefore, the interaction between $PM_{2.5}$ and the clusters is not confounded by effect modification of $PM_{2.5}$ by season.

We also examined the effect of $PM_{2.5}$ and the effect modification by cluster in cardiovascular mortality (Figure S2 in the supplementary material). Across all days, we found a positive non-significant association between CVD and the two days average $PM_{2.5}$.

Upon including the interaction between $PM_{2.5}$ and the clusters, we found that the effect of $PM_{2.5}$ was elevated in clusters 4 and 5, but were not significantly different from the same effect in cluster 1.

4. Discussion

This is the first study to investigate the health effects of multi-pollutant mixtures using a novel clustering framework that groups days based on their pollutant concentration profile. In Boston we found that the effect of $PM_{2.5}$ on total mortality was significant and that days characterized by high concentrations of the species representing primary traffic pollution and fuel oil combustion had a higher mortality-associated risk.

One important finding is that the method used to create multi-pollutant profiles proposed by Austin et al. (Austin et al. 2012) is robust, and provides a promising tool to identify multi-pollutant mixtures which can be linked to the health effects. In fact, the new cluster analysis, which included additional elements and years, yielded similar results with 72% of days that appeared in both analyses classified into the same cluster.

The effects estimates for $PM_{2.5}$ found in this study are similar to those previously reported for the Normative Aging Study. We have shown that pollutant mixtures, as indexed by air masses with specific back trajectories, influence the health effects of particles (Park et al. 2007). Specifically, the authors identified the paths that air masses traveled ('back-trajectories') before arriving in Boston, and they classified these trajectories into six clusters. The authors then examined whether the association of measured air pollutants with HRV differed by cluster. They found the strongest association of BC with reduced HRV during days when the air masses were coming from the SW, and days when concentrations of $PM_{2.5}$ and BC are relatively high due to the influence of local urban sources. These results are comparable to ours in that we found higher effects in cluster 3 characterized by days with strong contribution of traffic emissions.

We found that cluster 3 (Winter – Primary) was enriched in the elements Ni, V, Zn and Br and the gases CO, NO and SO_2, while days in cluster 5 (Winter – Low Primary, Higher O_3) were enriched by Ni, V, and K, and days in cluster 4 had high S, $PM_{2.5}$, and Se. This is consistent with previous investigations that examined the effects of individual particle species. For example, Franklin and co-authors (Franklin et al. 2008) and Zanobetti and co-authors (Zanobetti et al. 2009) found that effect estimates for $PM_{2.5}$ mass were higher when the $PM_{2.5}$ content of Br, Cr, Ni, or Na^+ was higher. In Bell et al (Bell et al. 2009) communities with higher $PM_{2.5}$ content of EC, V, or Ni had higher risk of hospitalizations associated with short-term exposure to $PM_{2.5}$.

Previously, source apportionment studies have shown increases in total and cardiovascular mortality associated with traffic emissions (characterized by elevated BC) as well as regional pollution (containing secondary pollutants), but not with soil particles (Laden et al. 2000; Mar et al. 2000; Sarnat et al. 2008). This is consistent with the results presented here. Furthermore, in Boston, short-term exposure to traffic and to a lesser extent SO_4^{2-} particles was associated with short-term mortality (Maynard et al. 2007). Ventricular tachyarrhythmia

were associated with $PM_{2.5}$, CO, NO_2, and BC and marginally with SO_4^{2-} (Dockery et al. 2005). Finally, in toxicological studies of concentrated ambient particles (CAPS), SO_4^{2-} and BC concentrations were associated with subtle alterations in pulmonary and systemic cell profiles (Clarke et al. 2000).

4.1 Limitations

One limitation of this study is that the framework proposed to define the multi-pollutants clusters can only be extended to communities where daily speciation data is available. The EPA Speciation Trends Network (STN) collects samples only one day in three or six; therefore, due to the irregularity of these measurements this analysis cannot be extended to other communities where only the public EPA data is available.

A possible issue is the use of two days average for the exposure, as the clusters represent the multi-pollutant profile at lag 0, while the two days average of $PM_{2.5}$ represent the average concentration of $PM_{2.5}$ over two days, lag 0 and lag 1; nevertheless we found very similar results when using the same day or the two days average of $PM_{2.5}$ mass.

Another limitation is the use of only one monitor for air pollution. Even though regional pollution is the major contributor to the pollutant concentrations observed in Boston, species concentrations measured at central site will have differing amounts of measurement error, depending on the spatial heterogeneity of a given species across the study region. Moreover, it would be important to have enough years of data in order to have adequate statistical power to examine other causes of mortality or of hospital admissions.

5. Conclusions

Our study shows a higher effect of $PM_{2.5}$ on total mortality during days with a strong contribution of primary traffic emissions and oil combustion. This important finding illustrates that mass alone is not a sufficient metric to use when evaluating health effects of PM exposure.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was made possible in part by USEPA grant RD-83479801, awarded by the U.S. Environmental Protection Agency. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA. Further, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication. This study was also funded by NIEHS grant ES000002

References

Austin E, Coull B, Thomas D, Koutrakis P. A framework for identifying distinct multipollutant profiles in air pollution data. Environment international. 2012; 45:112–121. [PubMed: 22584082]

Bell ML, Ebisu K, Peng RD, Samet JM, Dominici F. Hospital Admissions and Chemical Composition of Fine Particle Air Pollution. Am J Respir Crit Care Med. 2009; 179(12):1115–1120. [PubMed: 19299499]

Clarke RW, Coull B, Reinisch U, Catalano P, Killingsworth CR, Koutrakis P, et al. Inhaled concentrated ambient particles are associated with hematologic and bronchoalveolar lavage changes in canines. Environ Health Perspect. 2000; 108(12):1179–1187. [PubMed: 11133399]

Dockery DW, Luttmann-Gibson H, Rich DQ, Link MS, Mittleman MA, Gold DR, et al. Association of air pollution with increased incidence of ventricular tachyarrhythmias recorded by implanted cardioverter defibrillators. Environ Health Perspect. 2005; 113(6):670–674. [PubMed: 15929887]

Franklin M, Koutrakis P, Schwartz P. The role of particle composition on the association between PM2.5 and mortality. Epidemiology. 2008; 19(5):680–689. [PubMed: 18714438]

Hartigan J, Wong M. A k-means clustering algorithm. Journal of the Royal Statistical Society Series C. 1979; 28(1):100–108.

Katsouyanni K, Zmirou D, Spix C, Sunyer J, Schouten JP, Ponka A, et al. Short-term effects of air pollution on health: a European approach using epidemiologic time series data. The APHEA Project. Air Pollution Health Effects--A European Approach. Public Health Rev. 1997; 25(1):7–28. [PubMed: 9170962]

Klepeis NE. The heR Software Project | Exposure Science. Exposure Science Website. 2004

Laden F, Neas LM, Dockery DW, Schwartz J. Association of fine particulate matter from different sources with daily mortality in six U.S. cities. Environ Health Perspect. 2000; 108(10):941–947. [PubMed: 11049813]

Mar TF, Norris GA, Koenig JQ, Larson TV. Associations between air pollution and mortality in Phoenix, 1995–1997. Environ Health Perspect. 2000; 108(4):347–353. [PubMed: 10753094]

Mauderly JL, Burnett RT, Castillejos M, Ozkaynak H, Samet JM, Stieb DM, et al. Is the air pollution health research community prepared to support a multipollutant air quality management framework? Inhalation toxicology. 2010 Jun; 22(Suppl 1):1–19. [PubMed: 20462389]

Maynard D, Coull BA, Gryparis A, Schwartz J. Mortality risk associated with short-term exposure to traffic particles and sulfates. Environ Health Perspect. 2007; 115(5):751–755. [PubMed: 17520063]

Ostro B, Feng WY, Broadwin R, Green S, Lipsett M. The effects of components of fine particulate air pollution on mortality in california: results from CALFINE. Environ Health Perspect. 2007; 115(1):13–19. [PubMed: 17366813]

Ostro B, Tobias A, Querol X, Alastuey A, Amato F, Pey J, et al. The effects of particulate matter sources on daily mortality: a case-crossover study of Barcelona, Spain. Environ Health Perspect. 2011; 119(12):1781–1787. ed. [PubMed: 21846610]

Park SK, O'Neill MS, Stunder BJ, Vokonas PS, Sparrow D, Koutrakis P, et al. Source location of air pollution and cardiac autonomic function: trajectory cluster analysis for exposure assessment. J Expo Sci Environ Epidemiol. 2007; 17(5):488–497. [PubMed: 17299527]

Peng RD, Bell ML, Geyh AS, McDermott A, Zeger SL, Samet JM, et al. Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. Environ Health Perspect. 2009; 117(6):957–963. [PubMed: 19590690]

Rolph, G. NOAA Air Resources Laboratory. Silver Spring, MD: Silver Spring; 2003. READY - Real-time Environmental Applications and Display sYstem.

Rolph GD, Draxler RR. Sensitivity of Three-Dimensional Trajectories to the Spatial and Temporal Densities of the Wind Field. Journal of Applied Meteorology. 1990; 29(10):1043–1054.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, et al. The National Morbidity, Mortality, and Air Pollution Study. Part II: Morbidity and mortality from air pollution in the United States. Res Rep Health Eff Inst. 2000; 94(Pt 2):5–9.

Samoli E, Peng R, Ramsay T, Pipikou M, Touloumi G, Dominici F, et al. Acute effects of ambient particulate matter on mortality in Europe and North America: results from the APHENA study. Environ Health Perspect. 2008; 116(11):1480–1486. [PubMed: 19057700]

Sarnat JA, Marmur A, Klein M, Kim E, Russell AG, Sarnat SE, et al. Fine particle sources and cardiorespiratory morbidity: an application of chemical mass balance and factor analytical source-apportionment methods. Environmental health perspectives. 2008; 116(4):459–466. [PubMed: 18414627]

Thurston GD, Ito K, Mar T, Christensen WF, Eatough DJ, Henry RC, et al. Workgroup report: workshop on source apportionment of particulate matter health effects--intercomparison of results and implications. Environ Health Perspect. 2005; 113(12):1768–1774. [PubMed: 16330361]

Zanobetti A, Franklin M, Koutrakis P, Schwartz J. Fine particulate air pollution and its components in association with cause-specific emergency admissions. Environ Health. 2009; 8:58. [PubMed: 20025755]

Zanobetti A, Schwartz J. The effect of fine and coarse particulate air pollution on mortality: a national analysis. Environ Health Perspect. 2009; 117(6):898–903. [PubMed: 19590680]

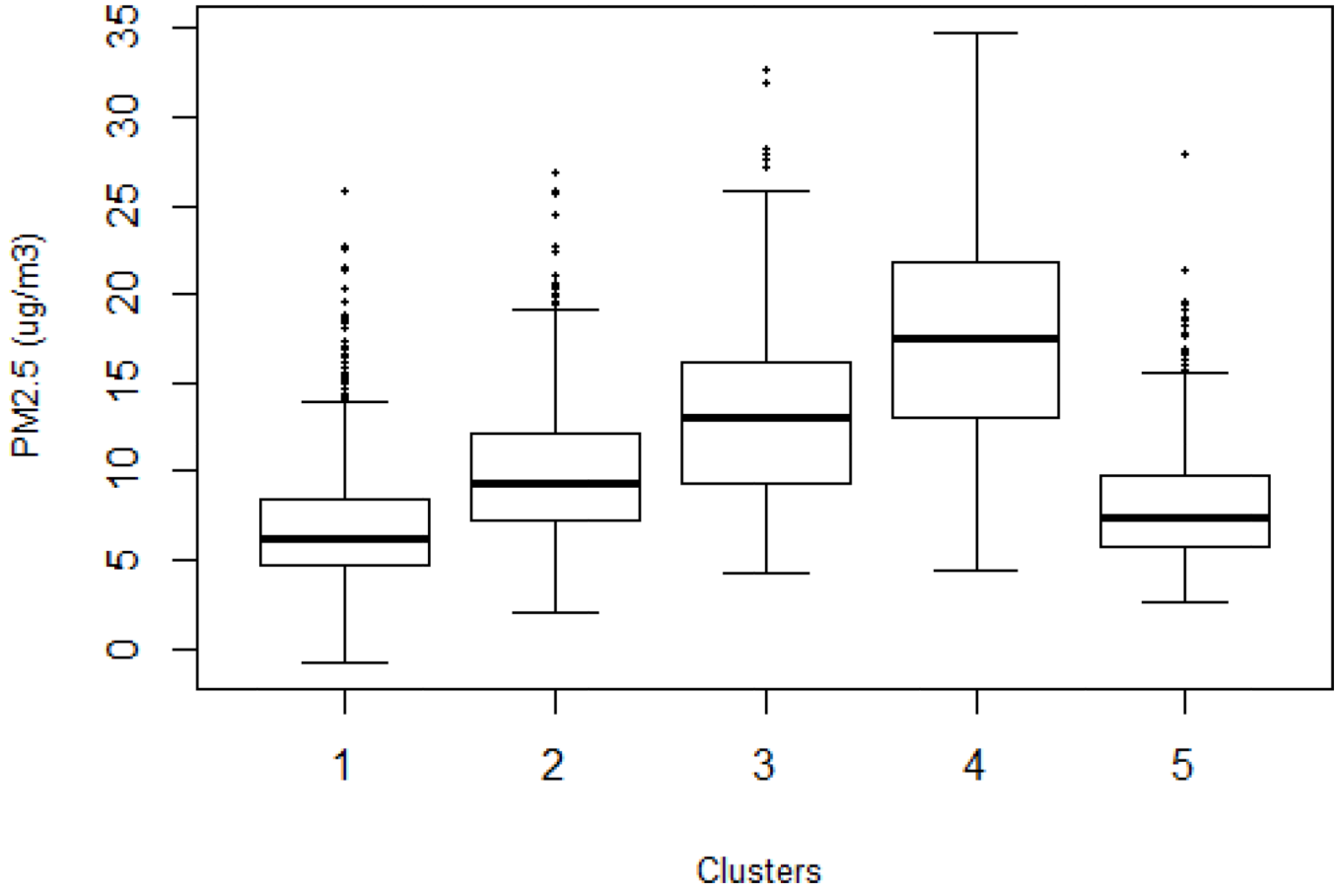

Figure 1.
Distribution of the $PM_{2.5}$ concentrations by clusters, years 1999–2009.

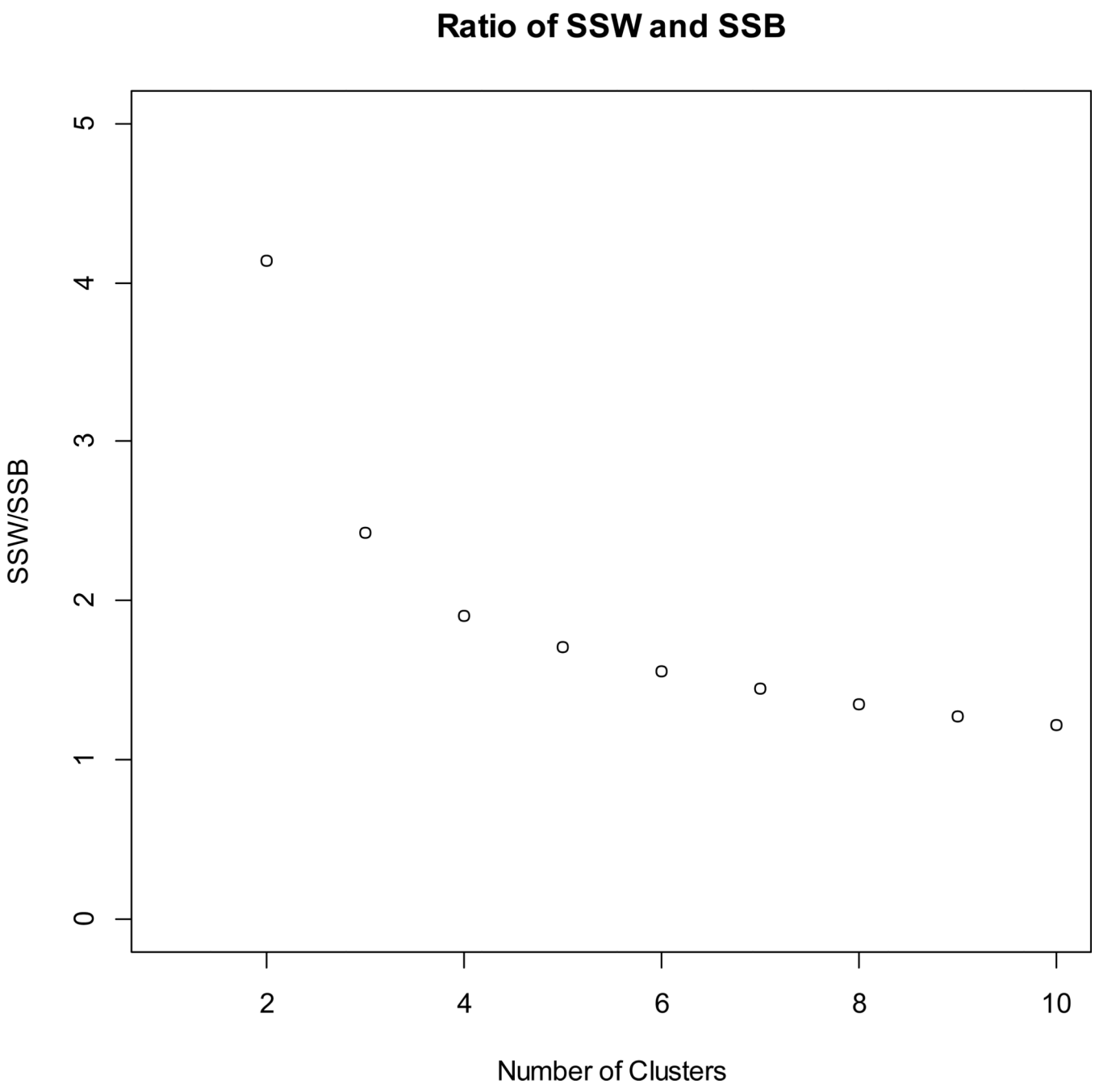

Figure 2.
Selecting the appropriate value of clusters: ratio of the within cluster to between cluster variability in the multivariate pollutant vector (SSW/SSB) for different values of k (number of clusters)

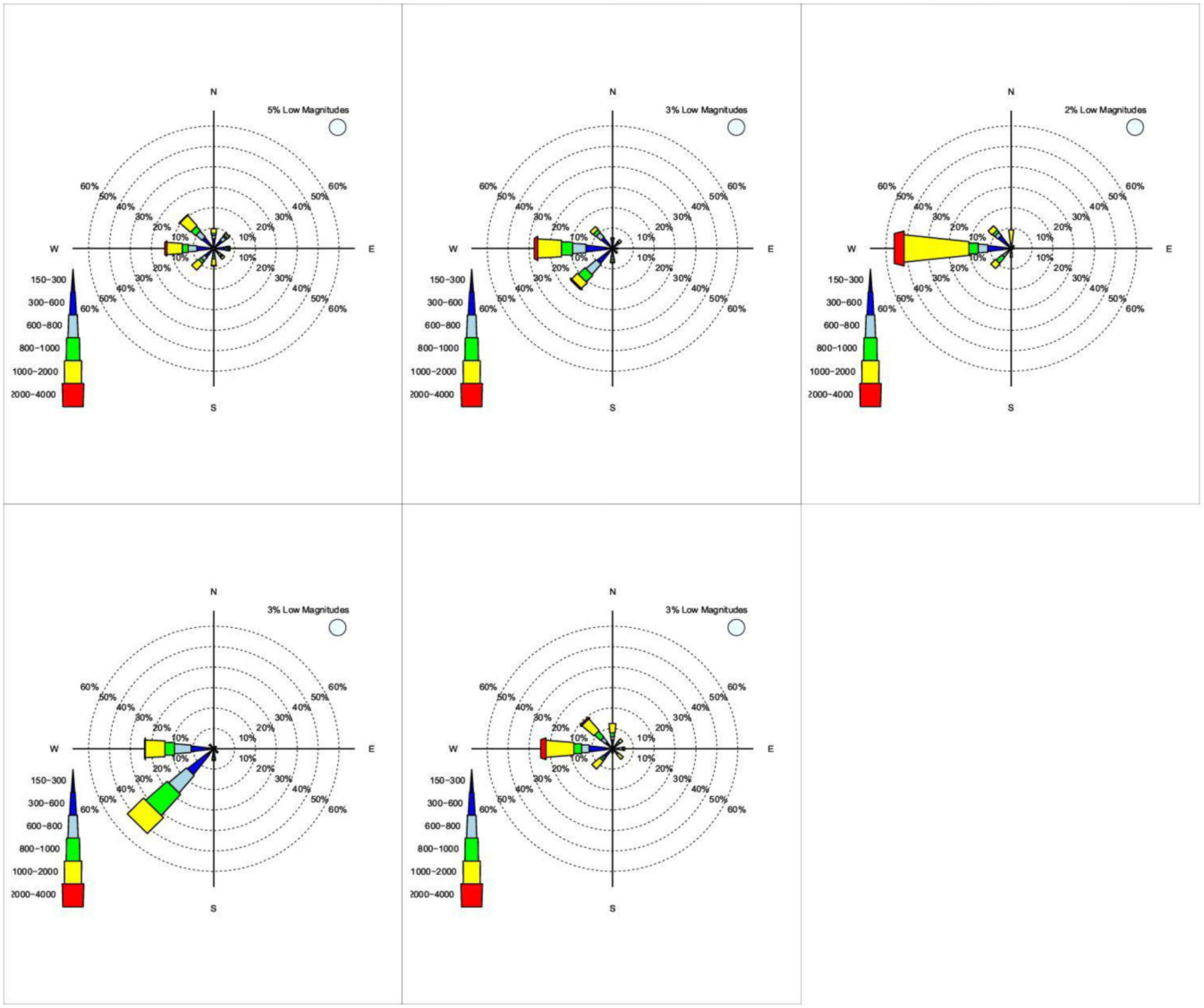

Figure 3.
Back-trajectory analysis of the different clusters; years 2004–2009

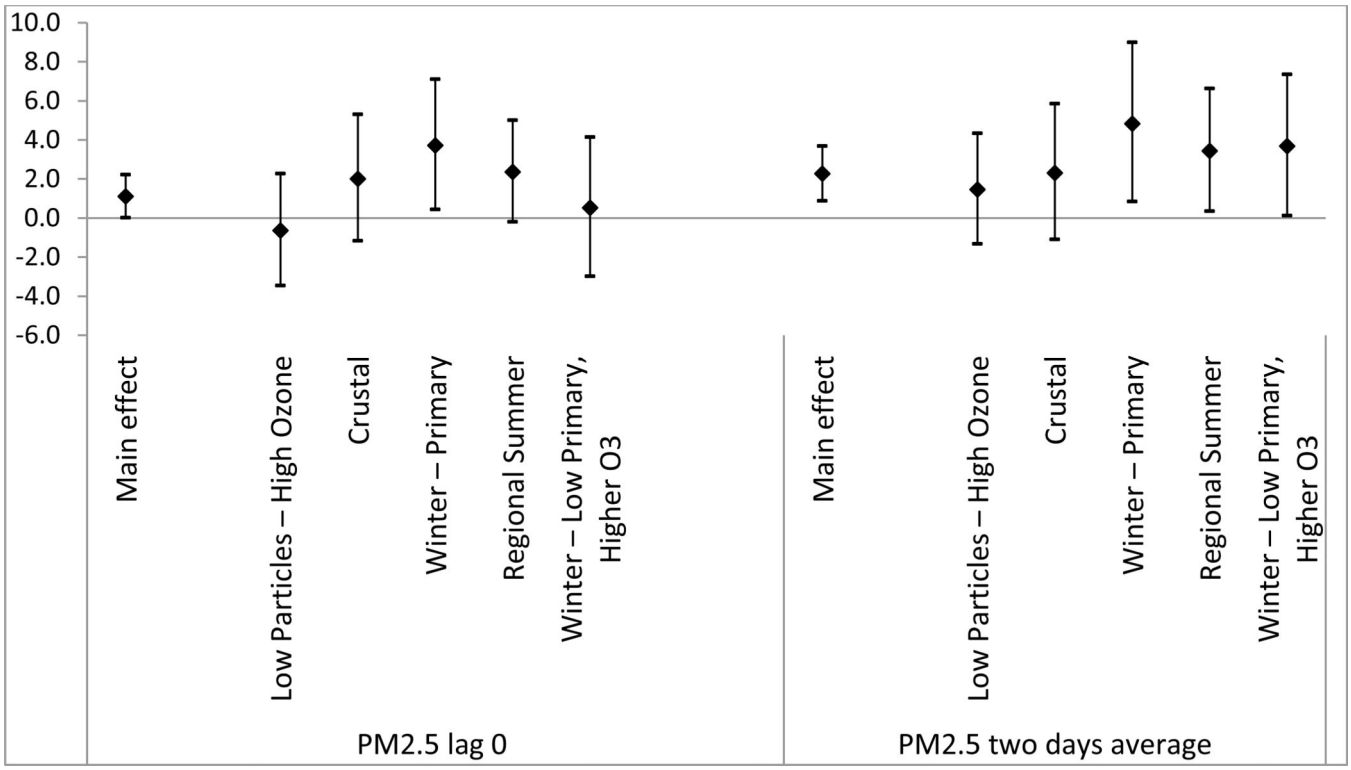

Figure 4.

Percent increase, 95% confidence intervals, in total mortality for 10 μg/m^3 increase in the same day and two days average of $PM_{2.5}$: main effect and effect of $PM_{2.5}$ in each cluster

Table 1

Boston 1999–2009; descriptive statistics of mortality, and exposure variables in total and by cluster

	Cluster 1 1068 days		Cluster 2 526 days		Cluster 3 272 days		Cluster 4 376 days		Cluster 5 603 days		Total 4018 days		
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**N**
Total mortality	53.7	9.0	54.2	8.4	61.2	9.3	53.5	8.2	60.0	9.4	56.0	9.5	4018
Temperature (°C)	12.0	8.4	13.8	7.8	3.3	6.5	20.6	6.6	3.0	7.1	10.8	9.4	4017
Dew point temp (°C)	5.4	9.5	6.9	9.0	−2.3	8.0	14.4	6.8	−3.5	9.5	4.4	10.4	4017
$PM_{2.5}$ ($\mu g/m^3$)	6.9	3.4	10.1	4.1	13.4	5.2	17.7	6.4	8.2	3.3	9.8	5.5	3683

Table 2

Cluster Description

	Low Particles - High O_3	Crustal	Winter Primary	Regional Summer	Winter - Low Primary, Higher O_3
	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
Frequency	1068	526	272	376	603
PM2.5 (ug/m3)	6.0 (2.7)	10.3 (3.7)	15.4 (5.3)	18.9 (7.2)	8.3 (3.1)
PN (#)	15322 (6750)	19222 (7575)	38322 (12350)	12368 (5429)	32090 (10757)
CO (ppm)	244.8 (119.1)	382.4 (139.7)	682.7 (222.4)	310.1 (135.2)	390.3 (151.9)
NO (ppb)	14.1 (5.7)	25.5 (12.2)	57.2 (23.7)	11.9 (6.2)	25.5 (10.6)
NO2 (ppb)	17.3 (4.0)	26.1 (5.1)	33.6 (5.8)	22.0 (5.7)	25.3 (4.6)
SO2 (ppb)	2.4 (1.5)	4.2 (2.0)	10.3 (3.9)	3.0 (1.7)	5.9 (2.4)
O3 (ppb)	20.8 (8.1)	18.1 (9.1)	8.0 (6.0)	34.8 (10.8)	14.7 (7.2)
BC (ug/m3)	0.5 (0.2)	0.9 (0.3)	1.2 (0.4)	0.9 (0.4)	0.6 (0.2)
S (ug/m3)	0.6 (0.4)	1.0 (0.5)	1.3 (0.5)	2.5 (1.1)	0.8 (0.4)
Cu (ng/m3)	2.4 (2.2)	4.1 (2.4)	4.7 (2.5)	3.6 (2.3)	2.8 (2.2)
Fe (ng/m3)	42.0 (15.9)	88.5 (24.7)	98.2 (37.7)	74.5 (24.4)	51.2 (16.0)
Zn (ng/m3)	6.5 (4.1)	14.3 (7.9)	22.2 (8.3)	11.2 (5.8)	11.0 (5.2)
Ni (ng/m3)	1.2 (1.1)	2.3 (1.6)	8.2 (4.1)	2.2 (1.5)	3.9 (2.5)
V (ng/m3)	1.6 (1.4)	3.0 (2.1)	8.2 (3.5)	3.6 (2.6)	4.0 (2.6)
Ti (ng/m3)	2.3 (1.2)	4.7 (1.7)	4.3 (1.8)	4.7 (2.1)	2.4 (1.1)
Mg (ng/m3)	43.4 (19.3)	50.4 (18.8)	56.0 (21.2)	87.0 (25.0)	48.5 (19.1)
K (ng/m3)	27.1 (16.0)	41.3 (16.5)	58.2 (21.0)	59.3 (49.4)	34.1 (12.6)
Si (ng/m3)	52.5 (30.3)	102.0 (54.4)	78.4 (42.8)	104.6 (67.8)	51.1 (24.9)
Na (ng/m3)	148.3 (98.9)	167.2 (85.9)	237.3 (96.6)	379.7 (140.9)	183.5 (91.3)
Cl (ng/m3)	12.0 (42.1)	4.9 (11.2)	10.1 (32.1)	13.1 (52.6)	7.3 (28.5)
Ca (ng/m3)	22.1 (8.7)	39.8 (14.8)	38.6 (14.5)	37.2 (14.7)	25.5 (8.2)
Br (ng/m3)	−0.1 (1.2)	0.7 (1.6)	1.7 (1.9)	1.1 (1.7)	0.1 (1.3)
Sr (ng/m3)	1.6 (1.2)	1.8 (1.2)	1.7 (1.2)	2.1 (1.5)	1.5 (1.2)
Pb (ng/m3)	4.7 (2.7)	6.3 (3.1)	8.1 (3.2)	6.4 (3.1)	5.2 (2.7)

	Low Particles - High O_3	Crustal	Winter Primary	Regional Summer	Winter - Low Primary, Higher O_3
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Mn (ng/m3)	−1.4 (1.8)	0.1 (2.1)	0.7 (2.7)	−0.5 (2.0)	−0.8 (1.9)
Temp (oC)	11.7 (8.7)	14.0 (7.6)	4.0 (6.3)	21.0 (6.4)	2.9 (7.1)
RH (%)	65.3 (16.4)	67.1 (15.5)	71.6 (14.8)	70.7 (12.3)	65.4 (18.1)
Wind Speed (m/s)	5.5 (1.7)	4.0 (1.2)	3.9 (1.4)	4.9 (1.3)	5.3 (1.6)

Table 3

Pollutant Ratios

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Fe/Si	0.80	0.87	1.25	0.71	1.00
K/Si	0.52	0.40	0.74	0.57	0.67
NO/NO2	0.81	0.98	1.70	0.54	1.01
S/PM2.5	0.10	0.09	0.09	0.13	0.10
BC/PM2.5	0.07	0.09	0.08	0.05	0.07
Ca/Si	0.42	0.39	0.49	0.36	0.50
Ni/V	0.75	0.76	1.01	0.62	0.97
PN/PM2.5	2534	1861	2486	654	3851

Table 4

Mean of the Normalized Concentrations of each element by Cluste

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
PN	1.2	0.9	1.2	0.3	1.8
CO	1.1	1.0	1.2	0.5	1.3
NO	1.0	1.1	1.6	0.3	1.4
NO2	1.2	1.1	0.9	0.5	1.3
SO2	0.9	0.9	1.5	0.4	1.6
O3	1.7	0.9	0.3	0.9	0.9
BC	1.1	1.3	1.1	0.7	1.0
S	1.0	0.9	0.8	1.3	0.9
Cu	1.2	1.2	1.0	0.6	1.1
Fe	1.1	1.4	1.0	0.6	1.0
Zn	1.0	1.2	1.3	0.5	1.2
Ni	0.7	0.8	1.9	0.4	1.7
V	0.8	0.9	1.6	0.6	1.5
Ti	1.2	1.4	0.8	0.8	0.9
Mg	1.3	0.9	0.7	0.9	1.1
K	1.2	1.0	1.0	0.8	1.1
Si	1.2	1.4	0.7	0.8	0.9
Na	1.2	0.8	0.8	1.0	1.1
Cl	2.0	0.5	0.7	0.7	0.9
Ca	1.2	1.3	0.8	0.7	1.0
Br	−0.2	1.5	2.4	1.3	0.4
Sr	1.5	1.0	0.7	0.6	1.1
Pb	1.4	1.1	0.9	0.6	1.1

*Mn was removed from this table due to the large number of reported negative values.