Journal of Exposure Science and Environmental Epidemiology (2007) 17, 488–497
© 2007 Nature Publishing Group All rights reserved 1559-0631/07/$30.00
www.nature.com/jes

Source location of air pollution and cardiac autonomic function: Trajectory cluster analysis for exposure assessment

SUNG KYUN PARK[a], MARIE S. O'NEILL[b], BARBARA J.B. STUNDER[c], PANTEL S. VOKONAS[d], DAVID SPARROW[d], PETROS KOUTRAKIS[a] AND JOEL SCHWARTZ[a]

[a]*Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA*
[b]*Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan, USA*
[c]*NOAA Air Resources Laboratory, Silver Spring, Mary Land, USA*
[d]*VA Normative Aging Study, Veterans Affairs Boston Healthcare System and the Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA*

Although many studies report that exposure to air pollution harms health, few have examined associations between pollution sources and health outcomes. We hypothesized that pollution originating in different locations has different associations with heart rate variability (HRV) among 497 men from the Normative Aging Study in Boston, Massachusetts. We identified the paths that air masses traveled ('back-trajectories') before arriving in Boston on the days the men were examined. Next, we classified these trajectories into six clusters. We examined whether the association of measured air pollutants with HRV (standard deviation of normal-to-normal intervals, high-frequency power (HF) and low-frequency power (LF), and LF/HF ratio) differed by cluster. We also examined whether the clusters alone (not considering air pollution measurements) showed different associations with HRV. The effects of black carbon (BC) on all HRV measures were strongest on days with southwest trajectories. Subjects who were examined on days where air parcels came from west had the strongest associations with ozone. All particle pollutants (particulate matter $<2.5\,\mu$m in aerodynamic diameter ($PM_{2.5}$), BC, and sulfates) were associated with increased LF/HF ratio on days with relatively short trajectories, which are related to local, slow-moving air masses. We also observed significant increases in LF/HF in days where air came from the northwest and west, compared to north trajectory days. Health effects associated with exposure to air pollution can be evaluated using pollutant concentrations as well as aspects of the pollution mixture captured by identifying locations where air masses originate. Independent effects of both these indicators of pollution exposure were seen on cardiac autonomic function.
Journal of Exposure Science and Environmental Epidemiology (2007) **17,** 488–497; doi:10.1038/sj.jes.7500552; published online 14 February 2007

Keywords: *air pollution, autonomic nervous system, back-trajectory, cluster analysis, heart rate variability.*

Introduction

Air pollution has been consistently associated with cardiovascular mortality and morbidity in a number of epidemiologic studies (Schwartz, 1999; Samet et al., 2000; Pope et al., 2002). Inhalation of ambient pollutants can induce an inflammatory response in the lung that triggers endothelial dysfunction, alterations of the autonomic nervous modulation, atherosclerosis, and thrombosis (Utell et al., 2002). These physiologic effects may vary according to the physical and chemical characteristics of the pollutants. Previous studies have evaluated whether certain components of airborne particles, such as transition metal elements, black carbon (BC), and/or sulfate (SO_4^{2-}) are responsible for health effects observed in epidemiologic research (Godleski et al., 2000; Laden et al., 2000; Magari et al., 2002; O'Neill et al., 2005; Schwartz et al., 2005). A panel study performed in Boston observed a significant association between BC, a marker of traffic particles, and decreased heart rate (HR) variability (HRV), a marker of cardiac autonomic dysfunction (Schwartz et al., 2005).

There is an increasing interest by policy-makers to link health effects with specific sources of air pollution (Committee on Research Priorities for Airborne Particulate Matter, 2004). Toward this end, scientists have used some particle components as surrogates for specific sources, for example BC for traffic particles and SO_4^{2-} for regional transported secondary particles from coal-burning power plants. In addition, examining the location from which air masses originate can provide insights into potentially toxic sources located upwind of the receptor site. This can be accomplished

1. Abbreviations: BC, black carbon; BMI, body mass index; CI, confidence interval; ECG, electrocardiogram; HF, high frequency power; HR, heart rate; HRV, heart rate variability; IQR, interquartile range; LF, low frequency power; NOAA, National Oceanic and Atmospheric Administration; O_3, ozone; $PM_{2.5}$, particulate matter $<2.5\,\mu$m in aerodynamic diameter; SD, standard deviation; SDNN, standard deviation of NN intervals; SO_4^{2-}, sulfate

2. Address all correspondence to: Dr. Sung Kyun Park, Department of Environmental Health Sciences, University of Michigan School of Public Health, 109 S. Observatory Street SPH II-M6240, Ann Arbor, MI 48109, USA. Tel.: +734 936 1719. Fax: +734 763 8095.
E-mail: sungkyun@umich.edu
Received 19 July 2006; accepted 4 December 2006; published online 14 February 2007

using air mass back-trajectories, which have been used in the past to link atmospheric concentrations to pollution sources (Stohl, 1998). However, little has been reported to date on associations between the origin of air masses and health outcomes. In an experimental study, Godleski et al. (2000) evaluated whether the location of air masses arriving in Boston modified cardiac responses in normal dogs exposed to ambient particles collected near a busy road in Boston (Godleski et al., 2000). They found significant increases in the standard deviation (SD) of HR and parasympathetic function and a significant decrease in average HR on days on which the air came from the northwest *versus* other directions. The association of air pollution and HRV among elderly individuals in Baltimore varied depending on the direction of air-mass trajectories for certain days (Creason et al., 2001). These findings and other evidence support the notion that air-mass origin can modify the effects of air pollution, and may also serve as a surrogate for complex mixtures from the sources in that location.

In a previous community-based cross-sectional study, we reported that particulate matter $<2.5\,\mu$m in aerodynamic diameter ($PM_{2.5}$), ozone (O_3), and BC were associated with reduced cardiac autonomic function, measured by HRV (Park et al., 2005). In this paper, we examine whether pollution originating in different locations has different associations with HRV using data from our previous study (Park et al., 2005) and from a trajectory cluster analysis.

Data and methods

Study Subjects

Participants were drawn from the Normative Aging Study, a longitudinal study of aging established by the US Veterans Administration in 1963 (Bell et al., 1972). The methods are described in detail elsewhere (Park et al., 2005). Briefly, healthy male volunteers from the Greater Boston area were screened at entry and accepted into the study if they had no history of known chronic diseases. Between 1963 and 1968, 2280 men aged 21–81 years who met the entry criteria were enrolled. At each visit, extensive physical examination, laboratory, anthropometric, and questionnaire data were collected. Beginning in November 2000, HRV was measured for 603 subjects during the course of each continuing participant's regularly scheduled evaluation at the Department of Veterans Affairs Outpatient Clinic in Boston. Participants visited the study center in the morning, after an overnight fast and abstinence from smoking. We collected data related to cardiovascular risk including body mass index (BMI), HR, systolic and diastolic blood pressures, fasting blood glucose, diabetes, hypertension, and prevalent ischemic heart disease status. Cigarette smoking, alcohol consumption, and subjects' use of medications were obtained by questionnaire. Temperature of the room where the HRV measurement was taken was recorded.

HRV Measurement

HRV measurement was performed between 0600 and 1300 at room temperature between 20°C and 30°C using a two-channel (five-lead) electrocardiogram (ECG) monitor (Trillium 3000; Forest Medical, East Syracuse, NY, USA) according to published protocol (Pope et al., 2001). The ECG was recorded for approximately 7 min with the subject seated after a 5-min rest (sampling rate of 256 Hz per channel). The best four consecutive minute intervals were used for the HRV calculations. The ECG digital recordings were processed, and HR and HRV measures were calculated using PC-based software (Trillium 3000 PC Companion Software for MS Windows; Forest Medical), which conforms to established guidelines (Task Force, 1996). Beats were automatically detected and assigned tentative annotations, which were then reviewed by an experienced scanner to correct for any mislabeled beats or artifacts. We used only normal-to-normal (NN) beat intervals, and calculated SD of NN intervals (SDNN), high-frequency power (HF) (0.15–0.4 Hz) and low-frequency power (LF) (0.04–0.15 Hz), and LF/HF ratio.

Trajectory Cluster Analysis

Air mass back-trajectories indicate transport paths along which air and pollutants travel to reach the location of interest, in this case, Boston, Massachusetts. Analytically, it was necessary to group individual trajectories by origination region with common air pollution sources, so that we could evaluate cardiovascular responses among study participants on days with air coming from those regions. The trajectories were clustered, or classified into groups of similar trajectories, using the method described in Stunder (1996) to identify several broad pollutant transport paths to Boston.

Back-trajectories were calculated using meteorological model output from the National Oceanic and Atmospheric Administration (NOAA), National Centers for Environmental Prediction, Eta Data Assimilation System (EDAS), as archived by the NOAA Air Resources Laboratory (NOAA, 2004). The meteorological data from the model included three-dimensional wind components, temperature, and other parameters, reported on a horizontal grid with 80 km resolution for each cell in the grid, over the continental US and surrounding areas. In the vertical dimension, the model output data were reported at various pressure levels, every 3 h. The trajectory computation was performed by the time integration of the position of an air parcel as it is transported by the three-dimensional winds. Back-trajectories were calculated using the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT-4) model (Draxler, 1996; Draxler and Hess, 1998).

The clusters were created using hourly trajectory end points (latitude, longitude). Given a set of M number of trajectories, each was defined initially as a separate cluster with zero spatial variance. Cluster spatial variance was the

sum of the squared distances between the end points of the individual trajectories in the cluster and the cluster-mean trajectory. Total spatial variance was the sum of the cluster spatial variances. Successive computational passes through the clusters paired the most similar clusters, or those that resulted in the least increase of the total spatial variance. Through this iterative process, similar clusters continued to be paired, and the change in total spatial variance was noted. Finally, only one cluster was left, having paired all trajectories. The record of the change in total spatial variance was reviewed to determine the resulting number of clusters. For much of the pairing process, the change in total spatial variance was small, but when two "dissimilar" clusters were paired, the change was large. This objective process gave the possible cluster outcome(s). If more than one outcome was possible, the final outcome had to be chosen subjectively from those outcomes. If statistical variance did not provide guidance, a visual inspection of plots showing the trajectories in each cluster provided information for assigning a given trajectory to the cluster representing a plausible source region. As noted by Stohl (1998), trajectory cluster analysis is not fully objective.

In this analysis, 48-h back-trajectories were calculated for air masses arriving at heights of 500 and 1000 m above ground level at the Harvard School of Public Health monitoring site, Boston, Massachusetts (latitude: 42.34 N, longitude: 71.10 W). We chose 500 and 1000 m as the starting heights because these would be representative of long-range transport and likely free of local wind patterns caused by buildings, trees and other objects. The air masses' arrival time was 0900 (14 or 13 coordinated universal time at standard or daylight savings time, respectively) on the dates on which ECG measurements were performed on study participants. The back-trajectories were generated for the 258 days on which patients were seen. An example of a back-trajectory calculation is shown in Figure 1. Different colors represent back-trajectories at different starting heights.

As back-trajectories from Boston are an idealization of three-dimensional transport and dispersion from pollutant sources to Boston, we chose to investigate the transport using two starting heights within the boundary layer. However, many of these 48-h back-trajectories reached the edge of the meteorological model domain before 48 h and so their originating point could not be calculated. The clustering methodology requires all trajectories to have the same duration, so we chose to use 36-h back-trajectories, which resulted in fewer unclassifiable days.

In the set of trajectories for this analysis, two additional criteria to determine the final number of clusters were set. One was that the cluster outcome occurred just before the first occurrence of pairing "dissimilar" clusters. The second was that the final number of clusters was <14. Dissimilar clusters were identified by a change in percentage increase of the total spatial variance of more than 30%. Fourteen

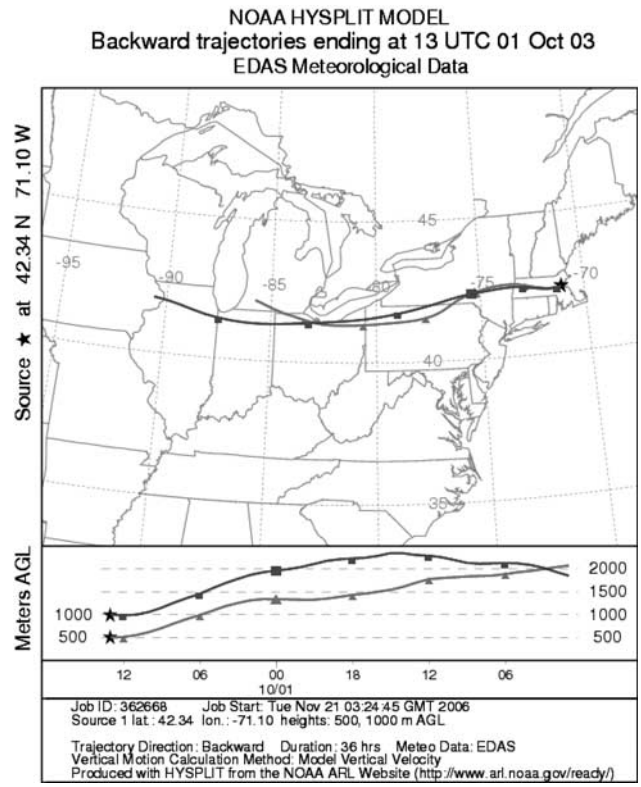

Figure 1. An example of a back-trajectory calculation.

clusters were arbitrarily chosen but could allow a combination of seven transport directions (e.g. northwest, east, etc.) by two wind speeds (e.g. slow or fast), which should allow for an adequate meteorological classification. Hence, trajectory cluster analysis produced seven clusters for each starting height that best represented the direction and speed (i.e. trajectory distance) differences in the trajectories. Figure 2 shows a sample cluster including all individual back-trajectories from a certain direction (west).

The average back-trajectories for the seven clusters at both 500 and 1000 m above ground level heights are presented in Figure 3. For further reference, each cluster is abbreviated according to its general direction: north (N), northwest (NW), west (W), etc. The abbreviation L stands for "local" because back-trajectories in this cluster are more associated with local, slow-moving air masses. The S cluster, which had even shorter cluster-mean trajectory than the L at 1000 m, is also more likely to have local sources. Overall, just over half (140/258 days) of the air masses arriving in Boston at the 500-m height had the same trajectory cluster classification as the air masses arriving at the 1000-m height.

Boundary layer height can vary by season (Stull, 1988). In the warm season, the boundary layer tends to be higher, whereas in the cool season it tends to have a lower height. As wind speed and direction tend to vary with height above ground, trajectories starting at different heights can be

different, particularly within a frontal system. Therefore, we chose the 1000 m as the starting height for a "warm season" (April–October) and the 500 m for a "cold season" (November–March).

The number of subjects who visited on the days corresponding to air masses from the NE cluster was too small (3.1%) to be compared with other clusters. We combined this cluster with the N cluster because characteristics of air pollution and meteorological data between these two clusters were similar. Hence, six clusters were used in the statistical analyses.

We also calculated back-trajectories for the air masses that arrived in Boston 1 day before the HRV examination (1-day lag) to assess if there was a delayed effect in relation to the 1-day-lagged back-trajectories.

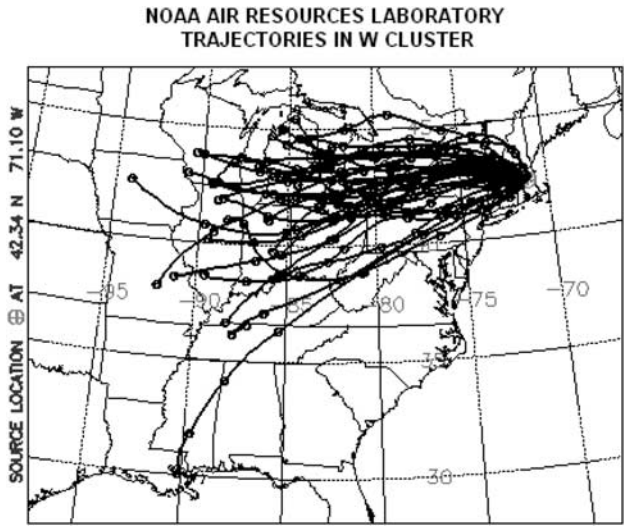

Figure 2. A sample cluster of back-trajectories (W cluster).

Air Pollution and Meteorological Data

Particulate air pollutants ($PM_{2.5}$, BC, and SO_4^{2-}) were measured at the Harvard School of Public Health monitoring site, 1 km from the exam site. Gaseous pollutants (O_3, sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO)) and meteorological data (temperature and dew-point temperature) were obtained from the Massachusetts Department of Environmental Protection local monitoring sites. To control for weather in analyses of HRV, we used apparent temperature, defined as a person's perceived air temperature (Steadman, 1979; Kalkstein and Valimont, 1986).

Statistical Methods

As the distribution of HRV was skewed to the left, $\log_{10}$-transformed HRV measures were used. Linear regression analyses were conducted controlling for age, BMI, fasting blood glucose, mean arterial pressure, use of antihypertensive medications, cigarette smoking, room temperature, season, and a cubic spline of apparent temperature with three degrees of freedom. We first examined whether each pollutant had a different effect by the six clusters. We constructed single pollutant linear regression models including interaction terms between categories of cluster and each pollutant along with the main effects. We estimated the percentage changes in each HRV parameter for interquartile range (IQR) increases in each air pollutant. Additionally, we fit linear regression models using dummy variables for the clusters as main effects and compared HRV on days where air masses came from different directions to HRV on days in the N cluster, because

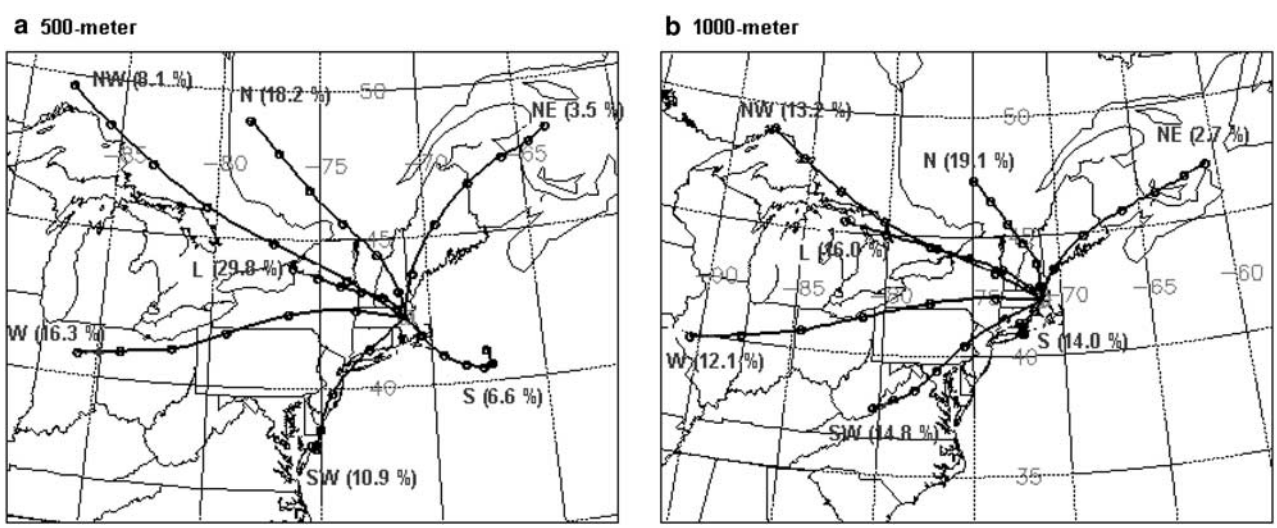

Figure 3. Thirty-six-hour average back-trajectories for the seven clusters at Harvard School of Public Health, Boston (latitude: 42.34 N, longitude: 71.10 W) with starting height at 500 and 1000 m above ground level. Percentages in parentheses reflect percentage of total days contributing to the averaged trajectory. A total of 6.6 and 8.2% of the days were unclassifiable in 500 and 1000 m, respectively.

this area includes air from the Atlantic Ocean (NE cluster) and there are fewer pollutant sources (industrial/urban) in that region; air from this cluster was expected to be relatively clean. We also ran the same regression models with 1-day-lagged clusters to assess the lagged effect of trajectory clusters. All statistical analyses were performed using SAS statistical software using PROC GAM (SAS Institute, 2001).

Results

Among 603 participants with HRV measurements, we excluded 92 with problematic HR measurements (atrial fibrillation, atrial bigeminy and trigeminy, pacemakers, irregular rhythm, irregular sinus rhythm, frequent ventricular ectopic activity, ventricular bigeminy, multifocal atrial tachycardia, or measurement time <3.5 min) and 14 with missing values of the potential confounding factors. As mentioned previously, 47 of the back-trajectories starting at either 500 or 1000 m terminated before 36 h and were excluded because they could not be assigned to a cluster. Accounting for these unclassified days and the other exclusion criteria, 450 participants were available for the analyses.

The demographic and clinical characteristics of the population ($n=497$) have been previously reported (Park et al., 2005). The characteristics of 450 participants included in the present study were substantially the same as previously described. The study subjects were all male, with average age 72.7 years (SD = 6.6 years).

The days assigned to the N cluster had the highest number of participants examined ($n=120$, 26.7%). The days when air masses came from the NW cluster had the fewest patients visiting ($n=49$, 10.9%). Subjects who visited on the days corresponding to air masses from the S cluster had somewhat lower HRV levels than those who visited on days corresponding to the other clusters (Table 1). No important differences in other characteristics were observed, comparing participant characteristics across the clusters.

Table 2 shows 24-h moving averages of air pollution and temperature levels at 0900 by the clusters. Northerly (N and NW) transports had lower concentrations of $PM_{2.5}$, BC, SO_4^{2-}, and NO_2 compared to the other transports, whereas airflows from the S, SW, and W resulted in higher concentrations. However, ratios of BC/$PM_{2.5}$ and SO_4^{2-}/$PM_{2.5}$ were highest in N and NW clusters, respectively. A total of 17.6% of the polluted air was associated with air masses from L cluster. The L cluster had high concentrations

Table 1. Characteristics and HRV measures [mean (SD) for continuous variables and *n* (%) for categorical variables] of the study participants by the clusters.

Variable	S ($n=56$)[a]	SW ($n=68$)	W ($n=75$)	L ($n=82$)	NW ($n=49$)	N ($n=120$)
Continuous variables						
Age (years)	73.2 (7.6)	72.1 (6.1)	72.3 (6.0)	72.3 (6.4)	73.4 (7.0)	73.1 (6.5)
BMI (kg/m^2)	28.1 (4.5)	27.9 (3.8)	29.3 (5.0)	27.9 (3.9)	28.5 (3.7)	28.1 (4.0)
SBP (mmHg)	131.3 (14.9)	131.0 (15.9)	132.5 (17.8)	126.8 (16.0)	134.3 (15.3)	131.8 (15.9)
DBP (mmHg)	76.3 (9.2)	74.7 (9.3)	75.7 (9.4)	74.3 (9.0)	76.8 (10.0)	76.4 (9.8)
MAP (mmHg)	94.6 (10.2)	93.5 (10.3)	94.6 (10.6)	91.8 (9.5)	96.0 (9.9)	94.9 (10.2)
Fasting glucose (mg/dl)	107.9 (25.6)	104.8 (23.9)	108.1 (34.2)	105.5 (21.0)	111.1 (37.9)	110.3 (32.8)
HRV						
Log_{10} SDNN (ms)	1.48 (0.22)	1.53 (0.27)	1.55 (0.25)	1.55 (0.26)	1.48 (0.25)	1.51 (0.24)
Log_{10} HF (ms^2)	1.75 (0.58)	1.90 (0.75)	1.92 (0.63)	1.99 (0.67)	1.78 (0.64)	1.92 (0.64)
Log_{10} LF (ms^2)	1.91 (0.55)	2.07 (0.58)	2.05 (0.54)	1.97 (0.51)	2.01 (0.49)	1.96 (0.51)
Log_{10} LF/HF	0.16 (0.46)	0.17 (0.48)	0.13 (0.45)	−0.02 (0.56)	0.23 (0.47)	0.04 (0.49)
Frequency, n (%)						
Smoking						
Never smoker	19 (33.9)	20 (29.4)	16 (21.3)	27 (32.9)	17 (34.7)	45 (37.5)
Former smoker	35 (62.5)	47 (69.1)	51 (68.0)	50 (61.0)	30 (61.2)	71 (59.2)
Current smoker	2 (3.6)	1 (1.5)	8 (10.7)	5 (6.1)	2 (4.1)	4 (3.3)
Diabetes mellitus	11 (19.6)	12 (17.6)	10 (13.3)	12 (14.6)	6 (12.2)	15 (12.5)
Hypertension	35 (62.5)	45 (66.2)	59 (78.7)	53 (64.6)	32 (65.3)	81 (67.5)
Use of β-blocker	14 (25.0)	23 (33.8)	30 (40.0)	25 (30.5)	11 (22.4)	46 (38.3)
Use of calcium channel blocker	5 (8.9)	10 (14.7)	13 (17.3)	13 (15.9)	4 (8.2)	19 (15.8)
Use of ACE inhibitor	11 (19.6)	12 (17.6)	22 (29.3)	19 (23.2)	10 (20.4)	20 (16.7)

ACE, angiotensin-converting enzyme; BMI, body mass index; DBP, diastolic blood pressure; HF, high-frequency power; HRV, heart rate variability; LF, low-frequency power; MAP, mean arterial pressure; SBP, systolic blood pressure; SDNN, standard deviation of normal-to-normal intervals. Each source location is named according to its general direction except for L, which stands for "local".
[a]The numbers in parentheses are the number of subjects in each cluster.

Table 2. 24-h moving averages of outdoor air pollutants and apparent temperature levels at 0900 on the dates of the HRV measurement by the clusters.

	S ($n=30$)[a]	SW ($n=34$)	W ($n=39$)	L ($n=42$)	NW ($n=28$)	N ($n=65$)
$PM_{2.5}$ ($\mu g/m^3$)	14.4 (12.3)[b]	14.1 (7.5)[b]	13.7 (6.5)[b]	15.1 (8.8)[b]	8.5 (6.0)	8.5 (4.9)
BC ($\mu g/m^3$)	1.1 (0.55)[b]	0.99 (0.42)	0.99 (0.39)	1.1 (0.45)[b]	0.63 (0.32)	0.79 (0.44)
SO_4^{2-} ($\mu g/m^3$)	4.2 (5.6)	4.0 (3.0)	3.8 (2.6)	4.0 (2.9)[b]	2.7 (1.8)	2.3 (1.3)
$BC/PM_{2.5}$ (%)	9.7 (3.7)	8.0 (3.3)	8.0 (3.0)	8.7 (3.8)	8.5 (4.3)	10.9 (7.4)
$SO_4^{2-}/PM_{2.5}$ (%)	25.1 (11.5)	27.1 (9.8)	26.9 (7.9)	27.2 (9.7)	33.2 (11.5)	28.2 (10.7)
O_3 (ppb)	29.1 (17.0)	25.4 (17.1)	17.0 (12.9)	22.8 (13.2)	22.9 (9.6)	21.3 (8.8)
NO_2 (ppb)	24.0 (5.9)	23.7 (5.7)	24.9 (7.1)[b]	24.5 (6.6)[b]	19.9 (4.0)	20.9 (4.8)
SO_2 (ppb)	4.3 (2.1)	4.6 (4.2)	6.7 (4.3)	6.0 (4.7)	4.0 (1.8)	4.7 (2.3)
CO (ppm)	0.45 (0.15)	0.54 (0.31)	0.62 (0.26)[b]	0.57 (0.27)[b]	0.49 (0.14)	0.43 (0.16)
ATemp (°C)	15.4 (9.3)	13.1 (8.7)	8.0 (9.4)	13.5 (12.0)	7.0 (8.6)	10.0 (9.5)

ATemp, apparent temperature (a person's perceived air temperature) = −2.653 + (0.994 × air temperature) + (0.0153 × dew-point temperature). Each source location is named according to its general direction except for L, which stands for "local".
[a]The numbers in parentheses are the number of days in each cluster. These numbers are different from those in Table 1 because more than one subject had HRV measurements on a given day.
[b]$P<0.05$ compared to N cluster (ANOVA-Bonferroni multiple comparisons).

of pollution, especially $PM_{2.5}$ and BC, because it was a cluster composed of relatively slow-moving air masses. The S cluster showed the highest concentration of SO_4^{2-} and similar BC concentrations as the L cluster. The maximum O_3 concentration was seen in the S cluster. Airflows from W, NW, and N had lower apparent temperature and lower O_3 concentrations, because the majority of them occurred in the cold season.

Figure 4 presents the estimated percentage changes of HRV in relation to various air pollutants by the clusters. The effects of BC on all HRV parameters were strongest in the SW cluster. One IQR increase in the 48-h moving average of BC (0.6 $\mu g/m^3$) was associated with 29.7% (95% confidence interval (CI), 10.6–44.7%), 62.4% (95% CI, 29.2–80.0%), and 44.8% (95% CI, 9.0–66.5%) decreases in SDNN, HF, and LF, respectively. Subjects who were examined on days where air parcels came from the W had the strongest associations with O_3. LF and SDNN decreased by 33.4% (95% CI, 0.9–55.3%) and 17.1% (95% CI, −0.3 to 31.5%), respectively, per one IQR increase in the 4-h average of O_3 (17 ppb). In addition, we found significant increases in LF/HF ratio in the L cluster in relation to $PM_{2.5}$ (47.0% (95% CI, 8.0–100.1%), IQR = 7 $\mu g/m^3$), BC [74.5% (95% CI, 14.4–166.2%)), and SO_4^{2-} (42.7% (95% CI, 6.6–91.1%), IQR = 3.1 $\mu g/m^3$). SO_4^{2-} was associated with a 67.7% (95% CI, 2.4–174.6%) increase in LF/HF in the NW cluster.

We also assessed the independent effect of trajectory clusters on HRV. Figure 5 shows the relative estimated percentage changes of HRV parameters in different clusters compared to those in the N cluster. After adjustment for potential confounders, 58.3% (95% CI, 10.2–127.4%) and 41.1% (95% CI, 1.4–96.3%) increases in LF/HF were found in the NW and W clusters, respectively. LF/HF ratio for days in the SW cluster was also borderline significantly higher than that for the N cluster. No significant difference in associations by cluster was found for other HRV measures.

In the analysis with 1-day-lagged clusters, $PM_{2.5}$, BC, and SO_4^{2-} were associated with increased LF/HF in the S cluster. $PM_{2.5}$ was associated with decreased HF and increased LF/HF in the NW clusters. There was no difference in the independent effect of the 1-day-lagged clusters on HRV parameters compared to the N cluster. Overall, both the effect modification and the independent effect of the clusters were less prominent compared to the results of the clusters with no lag (data not shown).

Discussion

To our knowledge, this is the first epidemiologic study to evaluate an association between source location and health outcomes in a community-based population. This approach to exposure assessment is consistent with recent statements by the US Environmental Protection Agency and the National Research Council calling for research that focuses on the role of biologically important components and characteristics of particulate matter and the consideration of pollution source indices (Committee on Research Priorities for Airborne Particulate Matter, 2004).

This study provides evidence that pollution arriving in Boston from different locations has different associations with cardiac autonomic function, as measured by HRV. Two consistent patterns were observed in our study. The strongest association of BC with reduced HRV occurred on days when the air mass was coming from the SW. This air traveled over the metropolitan areas of Washington, DC, Philadelphia, PA, New York, NY and other eastern seaboard cities before reaching Boston. We also found significant associations between LF/HF ratio and all particulate pollutants in the L

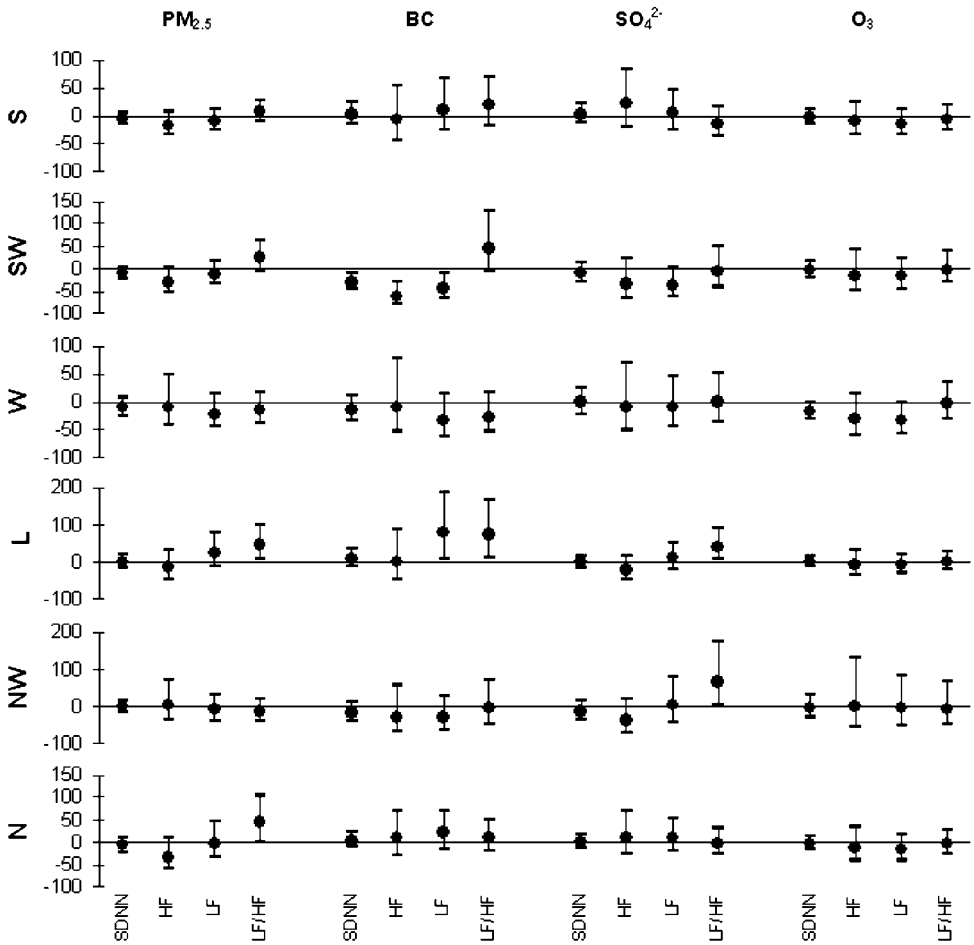

Figure 4. Estimated percentage changes and 95% CIs in HRV associated with one-interquartile increases in $PM_{2.5}$ (7 $\mu g/m^3$), BC (0.6 $\mu g/m^3$), SO_4^{2-} (3.1 $\mu g/m^3$), or O_3 (17 ppb) by the clusters, adjusted for age, BMI, mean arterial pressure, fasting blood glucose, cigarette smoking, use of β-blocker, calcium channel blocker, and/or angiotensin-converting enzyme (ACE) inhibitor, room temperature, season, and cubic smoothing splines (three degrees of freedom) of apparent temperature. The models were fit for each pollutant separately (single pollutant models).

cluster. The air of this cluster was mainly influenced by local urban sources, and the fact that we saw relatively high concentrations of $PM_{2.5}$ and BC in this cluster confirms the relation. BC is normally considered a surrogate for local, primary traffic particles, and indeed the diurnal pattern of hourly BC in Boston peaks in the early morning when trucks come into the area.

The effects of O_3 on LF and SDNN were greatest when the air mass over Boston originated from the W, traveling over Illinois, Indiana, and Ohio. However, the O_3 concentration in this cluster was low. This suggests either that O_3 on those days is capturing the effects of other, secondary and/or transported pollutants from the coal belt, which may modify the effect of O_3, or that the relation between ambient O_3 and exposure to O_3 is stronger during those periods. O_3 and SO_4^{2-} are produced by similar photochemical reactions and both are also regional transported pollutants. In the present study, no association between SO_4^{2-} and HRV was seen in the W cluster, but a significant relation between SO_4^{2-} and LF/HF was found in the NW, another westerly cluster next to W. The relatively stronger effects on LF/HF observed in these two directions compared to the N cluster (Figure 5) suggest that there is more than just SO_4^{2-} involved in this effect modification. On the other hand, the mean apparent temperature on those days is low (Table 2); so the participants in the study may have been keeping their windows open and not using air conditioning during the 36 h before their exams. Thus the personal exposure to ambient pollution is better represented by the ambient monitor measurements on those days *versus* the warmer days represented in other clusters. A recent study supports the fact that home ventilation is an important modifier of the association between personal and ambient pollution exposure (Sarnat et al., 2006).

The strongest main effect of the NW cluster and the significant association between SO_4^{2-} and increased LF/HF in this cluster suggest that the composition of particles from this location may be more toxic. Increased LF/HF reflects a condition of sympathetic excitation and parasympathetic (vagal) withdrawal, which may lead to sudden cardiac events,

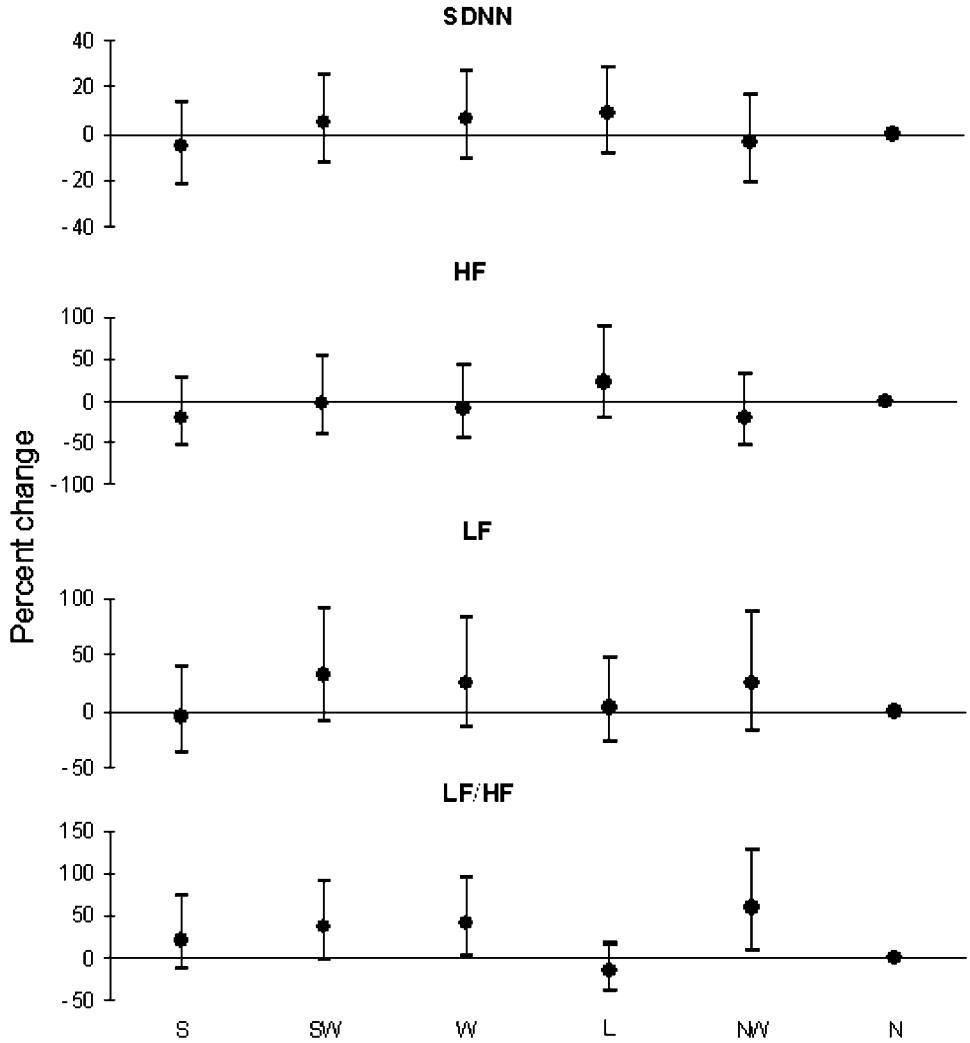

Figure 5. Estimated percentage changes and 95% CIs in HRV by the clusters, compared to north (N) cluster, adjusted for age, BMI, mean arterial pressure, fasting blood glucose, cigarette smoking, use of β-blocker, calcium channel blocker, and/or ACE inhibitor, room temperature, season, and cubic smoothing splines (three degrees of freedom) of apparent temperature.

such as ventricular arrhythmias and myocardial infarction (Task Force, 1996). The concentrations of SO_4^{2-} and $PM_{2.5}$ in this cluster were relatively lower than those in southerly and westerly clusters, but the ratio of SO_4^{2-} to $PM_{2.5}$ was the highest (Table 2). Transition metals bound to particles exist as metal oxides, such as vanadium sulfate (VSO_4), and therefore higher concentrations of SO_4^{2-} may be paralleled by higher concentrations of toxic metals on days classified in the NW trajectory. A study using elemental particle concentration data from Boston in 1995 and 1996 found that the nickel concentrations were higher among particles originating in the northwest than in particles coming from any other directions, and vanadium and nickel were associated with oil combustion sources (Oh, 2000). We could not identify whether the concentrations of these metals were really higher in this cluster because trace element concentration data were not available for our study period. Further studies to identify source elements and emission inventories that may have differential effects on health outcomes are needed.

It is important to place our work in the context of previous efforts suggesting that source-specific particle types have differential toxicity. Several studies have used source apportionment to estimate source-specific effects of particles (Laden et al., 2000; Mar et al., 2000, 2006; Janssen et al., 2002; Lippmann et al., 2005; Ito et al., 2006; Lanki et al., 2006). These studies have consistently showed that particles from motor vehicles and/or long-range transported secondary sulfate were significantly associated with adverse human health including increased total and cardiovascular mortality and hospital admissions, and depressed cardiac function. It is not possible to compare these studies with our results directly because we evaluated the directions from which air pollution came, not the types of particle sources. However, the present study supports the notion that particle toxicity is attributable

to traffic and coal combustion sources. Other studies have also reported that specific components of particles are associated with cardiovascular health outcomes. Magari et al. (2002) observed significant increases in SDNN in relation to lead and vanadium concentrations in a healthy occupational cohort. In a recent population-based study, O'Neill et al. (2005) also found independent effects of sulfate and BC on depressed vascular reactivity among patients with diabetes. Moreover, Schwartz et al. (2005) reported that BC was more strongly associated with decreased HRV than were other particle components.

The trajectory clustering approach does not focus on individual composition of particles, but on the flow of air masses. The studies described above used mass concentrations as an indicator of source-specific particle exposure, and thus enabled calculation of dose–response or thresholds in particle concentration–health relationships. By contrast, the trajectory clustering approach provides estimates of the associations with health end points on days with different air-mass sources, compared to days when the air mass came from the reference location. This approach estimates the effect of air pollution in terms of the regional variation in the particle composition and pollution mixture. Therefore, the trajectory clustering approach for exposure assessment provides complementary information to the source-specific particle approach in air pollution epidemiology.

Several limitations to the present study should be considered. We wished to use source locations of air masses 48 h before the examination of HRV, because in the previous study we examined lagged effects of particle pollutants from 4 up to 72 h averages and found stronger associations between particle pollutants and HRV in the 48-h moving average (Park et al., 2005). However, we had to compromise and use 36-h duration trajectories because many back-trajectories reached the edge of the meteorological grid before 48 h and therefore could not be classified. Most other studies chose a 24 h or even a shorter exposure window. Therefore, a 36-h exposure window is relevant to evaluate the effect of air pollution on HRV.

Long-range transport (regional scale, 100–1000 or more km) may be more relevant to evaluate health effects due to particles than shorter transports (urban scale, 10–100 or more km). In eastern North America, typical average regional contributions are approximately 75% of the mean urban $PM_{2.5}$ mass concentrations, whereas local urban aerosol contributes the remaining portion. On average, SO_4^{2-} is strongly regional in eastern North America, representing $\sim$75–95% of the urban SO_4^{2-} concentrations (Brook et al., 1999). The 36-h back-trajectories we used were able to capture some of this long-range transport because the distance of air mass transport depends on wind speed, and the trajectory clusters shown in Figure 3 extended considerable distances.

Trajectories were calculated to characterize the air-mass/pollutant source region associated with each subjects' physical examination. The start time for each trajectory was set to correspond to the time of the subjects' physical exam to clearly identify a source region affecting exposure to air pollution just before this visit. However, measurable effects on subjects' health may require many hours' exposure to toxic substances, in which case trajectories calculated from a start time before the exam may be more representative. Particularly on days where there was a cold or warm frontal passage, the trajectories used may not have captured the change in wind direction. Typically frontal (storm) systems pass about every 4 days (Anthes et al., 1978); so some of the trajectories may not have appropriately characterized the source region for that day.

Another limitation is potential error of choosing starting heights depending on season. When we analyzed trajectory clusters separately by different starting heights (500 and 1000 m), there were consistent associations of HRV with BC in the SW cluster and of O_3 in the W cluster in both heights. However, the relations with other pollutants in other clusters were not consistent according to the starting heights, and were difficult to interpret. Hence, to combine the trajectory clusters generated separately in each starting height, we chose 1000 m for the warm season and 500 m for the cold season, because boundary-layer height depends on meteorological conditions (Stull, 1988). We cannot rule out misclassification of trajectory clusters because some days classified in one cluster could be classified in other clusters if we change the definition of warm and cold seasons. Nevertheless, the consistent findings observed in this combined analysis and the analyses carried out separately by each starting height suggest that these results may not be simply due to chance.

Lastly, trajectories represent approximations to the air/pollutant path (Stohl, 1998). Turbulence affects atmospheric transport, but trajectories are calculated using mean winds, not the turbulent component. For instance, pollutants traveling to Boston would not travel along a curved line (i.e. a trajectory), but would disperse horizontally and vertically in the atmosphere. Nevertheless, trajectories are a useful tool and have been used frequently to identify source regions (Wotawa and Kroger, 1999; Butler et al., 2003).

Using back-trajectories for exposure assessment may not always be feasible in geographic regions where air pollution sources are relatively similar in all directions around the 'exposure' area of interest. Boston has the advantage of distinct source regions with contrasting pollutant mixtures.

Polluted air from the south and west of Boston, where sources include vehicular traffic and coal-burning power plants, is associated with cardiac autonomic dysfunction. Although the concentrations of air pollutants on days in the NW cluster were relatively low, the strong associations with HRV suggest that the composition of air pollution from that area may be more toxic. The use of the trajectory clustering approach for exposure assignment provides complementary information to the individual pollutant analysis in air pollution epidemiology.

Acknowledgements

This work was supported by the National Institute of Environmental Health Sciences (NIEHS) (ES00002, PO1 ES009825) and the US Environmental Protection Agency (EPAR827353). The VA Normative Aging Study is supported by the Cooperative Studies Program/Epidemiology Research and Information Center of the US Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research and Information Center, Boston. S.K.P. and M.S.O. were supported by training grant T32 ES07069 from the NIEHS, National Institutes of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIEHS. M.S.O. was supported by the Robert Wood Johnson Foundation Health & Society Scholars program.

We thank E.R. Dibbs and J.D. Awerbach for their invaluable assistance in conducting the heart rate variability measurements and other contributions to the VA Normative Aging Study.

References

Anthes R.A., Panofsky H.A., Cahir J.J., and Rango A. The changing weather outside the tropics. *The Atmosphere*, 2nd edn., Chapter 5 Charles E. Merrill Publishing Company: Columbus, Ohio, 1978, 137–170.

Bell B., Rose C., and Damon A. The Normative Aging Study: an interdisciplinary and longitudinal study of health and aging. *Aging Human Dev* 1972: 3: 4–17.

Brook J.R., Dann T.F., and Bonvalot Y. Observations and interpretations from the Canadian fine particle monitoring program. *J Air Waste Manag Assoc* 1999: 49: PM-35–PM-44.

Butler T.J., Likens G.E., Vermeylen F.M., and Stunder B.J.B. The relation between NO_x emissions and precipitation NO_3^- in the eastern USA. *Atmos Environ* 2003: 37: 2093–2104.

Committee on Research Priorities for Airborne Particulate Matter, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies. Synthesis of research progress on particulate matter. *Research Priorities for Airborne Particulate Matter: IV. Continuing Research Progress*. The National Academy Press: Washington, DC, 2004, 64–117.

Creason J., Neas L., Walsh D., Williams R., Sheldon L., Liao D., et al. Particulate matter and heart rate variability among elderly retirees: the Baltimore 1998 PM study. *J Expo Anal Environ Epidemiol* 2001: 11(2): 116–122.

Draxler R.R. Trajectory optimization for balloon flight planning. *Weather Forcast* 1996: 11: 111–114.

Draxler R.R., and Hess G.D. An overview of the Hysplit_4 modeling system for trajectories, dispersion, and deposition. *Aust Met Mag* 1998: 47: 295–308.

Godleski J.J., Verrier R.L., Koutrakis P., Catalano P., Coull B., Reinisch U., et al. Mechanisms of morbidity and mortality from exposure to ambient air particles. *Res Rep Health Eff Inst* 2000: (91): 5–88, discussion 89–103.

Ito K., Christensen W.F., Eatough D.J., Henry R.C., Kim E., Laden F., et al. PM source apportionment and health effects: 2. An investigation of intermethod variability in associations between source-apportioned fine particle mass and daily mortality in Washington, DC. *J Expo Sci Environ Epidemiol* 2006: 16(4): 300–310.

Janssen N.A., Schwartz J., Zanobetti A., and Suh H.H. Air conditioning and source-specific particles as modifiers of the effect of PM(10) on hospital admissions for heart and lung disease. *Environ Health Perspect* 2002: 110(1): 43–49.

Kalkstein L.S., and Valimont K.M. An evaluation of summer discomfort in the United States using a relative climatological index. *Bull Am Meteorol Soc* 1986: 67: 842–848.

Laden F., Neas L.M., Dockery D.W., and Schwartz J. Association of fine particulate matter from different sources with daily mortality in six US cities. *Environ Health Perspect* 2000: 108(10): 941–947.

Lanki T., de Hartog J.J., Heinrich J., Hoek G., Janssen N.A., Peters A., et al. Can we identify sources of fine particles responsible for exercise-induced ischemia on days with elevated air pollution? The ULTRA study. *Environ Health Perspect* 2006: 114(5): 655–660.

Lippmann M., Hwang J.S., Maciejczyk P., and Chen L.C. PM source apportionment for short-term cardiac function changes in ApoE−/− mice. *Environ Health Perspect* 2005: 113(11): 1575–1579.

Magari S.R., Schwartz J., Williams P.L., Hauser R., Smith T.J., and Christiani D.C. The association of particulate air metal concentrations with heart rate variability. *Environ Health Perspect* 2002: 110(9): 875–880.

Mar T.F., Ito K., Koenig J.Q., Larson T.V., Eatough D.J., Henry R.C., et al. PM source apportionment and health effects 3. Investigation of inter-method variations in associations between estimated source contributions of $PM_{2.5}$ and daily mortality in Phoenix, AZ. *J Expo Sci Environ Epidemiol* 2006: 16(4): 311–320.

Mar T.F., Norris G.A., Koenig J.Q., and Larson T.V. Associations between air pollution and mortality in Phoenix, 1995–1997. *Environ Health Perspect* 2000: 108(4): 347–353.

NOAA A.R.L. Gridded Meteorological Data Archives. *EDAS Archive*, Available at: http://www.arl.noaa.gov/ss/transport/archives.htmlAccessed August 9, 2004.

Oh J.A. *Characterization and source apportionment of air pollution in Nashville, TN and Boston, MA* ScD thesis, Department of Environmental Health, Harvard School of Public Health, Boston, 2000.

O'Neill M.S., Veves A., Zanobetti A., Sarnat J.A., Gold D.R., Economides P.A., et al. Diabetes enhances vulnerability to particulate air pollution-associated impairment in vascular reactivity and endothelial function. *Circulation* 2005: 111(22): 2913–2920.

Park S.K., O'Neill M.S., Vokonas P.S., Sparrow D., and Schwartz J. Effects of air pollution on heart rate variability: The VA Normative Aging Study. *Environ Health Perspect* 2005: 113(3): 304–309.

Pope III C.A., Burnett R.T., Thun M.J., Calle E.E., Krewski D., Ito K., et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 2002: 287(9): 1132–1141.

Pope III C.A., Eatough D.J., Gold D.R., Pang Y., Nielsen K.R., Nath P., et al. Acute exposure to environmental tobacco smoke and heart rate variability. *Environ Health Perspect* 2001: 109(7): 711–716.

Samet J.M., Dominici F., Curriero F.C., Coursac I., and Zeger S.L. Fine particulate air pollution and mortality in 20 US cities, 1987–1994. *N Engl J Med* 2000: 343(24): 1742–1749.

Sarnat S.E., Coull B.A., Schwartz J., Gold D.R., and Suh H.H. Factors affecting the association between ambient concentrations and personal exposures to particles and gases. *Environ Health Perspect* 2006: 114(5): 649–654.

SAS Institute. *The SAS System for Windows, Release 8*. SAS Institute Inc., Cary, NC, 2001.

Schwartz J. Air pollution and hospital admissions for heart disease in eight US counties. *Epidemiology* 1999: 10(1): 17–22.

Schwartz J., Litonjua A., Suh H., Verrier M., Zanobetti A., Syring M., et al. Traffic related pollution and heart rate variability in a panel of elderly subjects. *Thorax* 2005: 60(6): 455–461.

Steadman R.G. The assessment of sultriness. Part II: effects of wind, extra radiation, and barometric pressure on apparent temperature. *J Appl Meteor* 1979: 18: 874–885.

Stohl A. Computation, accuracy and applications of trajectories — a review and bibliography. *Atmos Environ* 1998: 32(6): 947–966.

Stull R.B. Mean boundary layer characteristics. *An Introduction to Boundary Layer Meteorology*. Kluwer Academic Publishers: Dordrecht, The Netherlands, Chapter 1 1988, 1–25.

Stunder B.J.B. An assessment of the quality of forecast trajectories. *J Appl Meteor* 1996: 35: 1319–1331.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* 1996: 93(5): 1043–1065.

Utell M.J., Frampton M.W., Zareba W., Devlin R.B., and Cascio W.E. Cardiovascular effects associated with air pollution: potential mechanisms and methods of testing. *Inhal Toxicol* 2002: 14(12): 1231–1247.

Wotawa G., and Kroger H. Testing the ability of trajectory statistics to reproduce emission inventories of air pollutants in cases of negligible measurement and transport errors. *Atmos Environ* 1999: 33: 3037–3043.