Bayesian Hierarchical Modeling of Cardiac Response to Particulate Matter Exposure

Sandra J. McBride
Nicholas School of the Environment and Earth Sciences, Duke University
Durham, NC 27708-0251
Gary Norris
National Exposure Research Laboratory, USEPA, Research Triangle Park, NC 27711
Ron Williams
National Exposure Research Laboratory, USEPA, Research Triangle Park, NC 27711
Lucas Neas
National Health and Environmental Effects Research Laboratory, USEPA, Research Triangle Park, NC 27711

Acknowledgements

The authors wish to thank John Creason (US EPA) and Debra Walsh (US EPA) for their assistance with health measures data as well as Charles Rodes and the staff of RTI International for their collection of field exposure data.

Disclaimer

This manuscript is now being subjected to external peer review and has not been cleared for publication by the U.S. Environmental Protection Agency. The U.S. Environmental Protection Agency through its Office of Research and Development funded and conducted the research described here through contract 3D-5925-WATX and 4D-5895-WATX to Dr. Sandra McBride. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Abstract

Studies have linked increased levels of particulate air pollution with decreased autonomic control, as measured by heart rate variability (HRV) particularly in susceptible populations such as the elderly. In the present study we utilize data from the 1998 USEPA epidemiology-exposure longitudinal panel study of elderly adults in a Baltimore retirement home to examine the relationship between heart rate variability and $PM_{2.5}$ personal exposure. We consider $PM_{2.5}$ personal exposure in the aggregate as well as personal exposure to the components of $PM_{2.5}$ as estimated in two ways using receptor models by Hopke et al. (2003). We develop a Bayesian hierarchical model for heart rate variability as a function of personal exposure to $PM_{2.5}$ which integrates heart rate variability (HRV) measurements as well as data from personal, indoor and

outdoor $PM_{2.5}$ monitoring and meteorological data. We found a strong relationship between decreased HRV (HF, LF, r-MSSD, and SDNN) and total personal exposure to $PM_{2.5}$ at lag one day. Using the personal exposure monitoring (PEM) apportionment results of Hopke et al. (2003), we examined the relative importance of ambient and non-ambient personal $PM_{2.5}$ exposure to HRV and found the effect of internal non-ambient sources of $PM_{2.5}$ on HRV to be minimal. Using the PEM apportionment data, a consistent effect of soil at short time scales (lag 0) was found across all five HRV measures, and an effect of sulfate on HRV was seen for HF and r-MSSD at the moving average of lags 0 and 1 day. Hopke et al. (2003)'s ambient site apportionment data indicated effects of nitrate on HRV at lags 1 day, and moving averages of days 0 and 1 and days 0-2 for all but the ratio LF/HF. Sulfate had an effect on HRV at lag 1 day for four HRV measures (HF, LF, r-MSSD, SD of NN) and for LF/HF at a moving average of days 0-2.

1 Introduction

Studies have linked increased levels of particulate air pollution with increased cardiovascular mortality and morbidity in susceptible populations such as the elderly (Dockery et al., 1993; Dockery, 2001). However, the potential physiological mechanisms of this association are still unknown. One hypothesis is that particulate exposure may alter cardiac autonomic control as measured through heart rate variability (HRV), a measure of naturally occurring, beat-to-beat variations in heart rate. Declines in HRV have been associated with increased risk of myocardial infarction and sudden cardiac death in the elderly and those with compromised cardiovascular health (La Rovere et al., 1998; Dekker et al., 1997).

Animal studies (Godleski et al., 2000) and a number of panel studies have shown an association between increased exposure to total airborne particulate matter of diameter less than 2.5 microns ($PM_{2.5}$) mass and lowered HRV over time scales of up to 48 hours. Magari et al. (2001) monitored forty male boilermakers during a work shift using an ambulatory electrocardiogram monitor and a personal exposure monitor (PEM) for $PM_{2.5}$ and found that workers experienced decreased HRV (as measured by the 5-minute standard deviation of the normal-to-normal intervals (SDNN)) as a function of moving $PM_{2.5}$ averages taken from two hours to seven hours after exposure after adjustment for heart rate. In a study by Devlin et al. (2003), healthy elderly volunteers exposed to concentrated air pollution particles for a two hour period were found to have decreased HRV in the time and frequency domains immediately following exposure, with some changes persisting up to 24 hours later. Pope et al. (2004) examined the relationship between daily HRV and daily average ambient $PM_{2.5}$ levels in 88 elderly residents of 3 communities in Utah, using repeated 24-hour ambulatory ECG monitoring during periods of high and low air pollution. After controlling for temperature and humidity, consistent declines in HRV were seen as $PM_{2.5}$ levels increased. Cavallari et al. (2008) monitored 36 male boilermaker welders using ambulatory electrocardiograms

and PEMs, and found an inverse association between SDNN and work-related $PM_{2.5}$ exposures in each of the 14 hours after work ended, suggesting an early phase response, at 2-3 hours, and a later phase response, at 9-13 hours.

More recent studies have examined associations between HRV measures and the components of $PM_{2.5}$. For 497 men in the Normative Aging Study in greater Boston, Park et al. (2005) examined the relationship between HRV and 4-hour, 24-hour, and 48-hour moving averages of air pollution at an ambient location and found decreases in HRV measures over all three time scales. HRV measurements included SDNN, high and low frequency power (HF, LF, respectively) and the ratio of LF to HF. Park found several significant associations between $PM_{2.5}$ mass and ozone, but found no significant association of HRV with particle number concentration, NO_2, SO_2, and CO for any of the averaging time periods. Luttman-Gibson et al. (2006) conducted a panel study of 32 non-smoking senior adults over two seasons, examining the relationship between 24-hour integrated $PM_{2.5}$ concentrations at an ambient site and HRV measures including SDNN, the mean square of differences between adjacent RR intervals (r-MSSD), and the frequency domain HRV measures (HF and LF). Luttman-Gibson et al. (2006) also examined concentrations of sulfate (SO_4^{-2}), elemental carbon (EC) and gaseous pollutants (O_3, NO_2, SO_2). Findings included (1) an association between the four HRV measures and mean $PM_{2.5}$ during the day previous to HRV measurement, (2) a significant association between SO_4^{-2} and HRV at lag one day, (3) an association between nonsulfate $PM_{2.5}$ and SDNN and r-MSSD, and (4) no association between the elemental carbon (EC) fraction or gaseous components with HRV measures. Sarnat et al. (2008) examined the relationship between source-apportionment estimates and cardiorespiratory morbidity in Atlanta using Poisson generalized linear models, and found positive associations between same-day $PM_{2.5}$ and mobile source and biomass combustion sources, as well as between sulfate-rich secondary $PM_{2.5}$ and respiratory emergency department visits.

In the present study we utilize data from the 1998 USEPA epidemiology-exposure lon-

gitudinal panel study of elderly adults in a Baltimore retirement home to examine the relationship between HRV and $PM_{2.5}$ personal exposure. Two analyses of USEPA panel study data for elderly adults in Baltimore have already been published. For 26 elderly residents of a retirement home, Liao et al. (1999) examined the relationship between HRV and daily $PM_{2.5}$ concentrations measured at a central indoor site and an outdoor location over a three week period in early 1997. Using a series of mixed effects models, Liao et al. (1999) found an inverse association between daily $PM_{2.5}$ concentrations and HRV. Liao's study formed the pilot study for a second panel study conducted on the same population with more extensive personal $PM_{2.5}$ monitoring. Full details of the 1998 Baltimore Epidemiology-Exposure Study are given in Williams et al. (2000b) and Williams et al. (2000a). Briefly, HRV measures were taken over a period of one month for 56 respondents, 21 of whom wore personal $PM_{2.5}$ exposure monitors. Creason et al. (2001) reported a small negative association between HRV and outdoor $PM_{2.5}$ on the previous day after adjusting for age, sex and cardiovascular status in mixed effects models. Findings based on $PM_{2.5}$ concentrations at a central indoor site were similar.

In the present study, we expand on the work of Liao et al. (1999) and Creason et al. (2001) to include personal $PM_{2.5}$ measurements, and we develop our models in a Bayesian hierarchical framework, which integrates data from personal, indoor and outdoor monitoring and meteorological data. We develop a sequence of nested probability models that integrate different types of data at multiple levels and bring together multiple sources of variation in one probabilistic framework. The joint distribution of the parameters links estimation in a unified way, such that parameter estimates "borrow strength" from available information on related parameters elsewhere in the model. Markov Chain Monte Carlo (MCMC) methods make possible a range of inferences about quantities at different levels of the hierarchy. Controlling for apparent temperature, age and cardiovascular health, we then relate HRV measures in study subjects to the posterior distribution of personal $PM_{2.5}$ exposure of am-

bient and non-ambient origins. We then expand upon the Bayesian hierarchical framework to relate HRV to particular sources by incorporating receptor modeling results from the Baltimore study by Hopke et al. (2003).

The paper is organized as follows. In Section 2 we provide more details on the panel study data and we lay out the hierarchical model, prior distributions and implementation details. In Section 3, we discuss posterior inference with the hierarchical model, and we perform model checking and sensitivity analysis. In Section 4, we discuss the implications of the findings.

2 Methods

2.1 The 1998 Baltimore Epidemiology-Exposure Study

In this analysis, we utilize data on $PM_{2.5}$ measurements and health endpoints indicative of cardiac autonomic control for 56 subjects enrolled in the 1998 Baltimore Epidemiology-Exposure Study. Full details of the design, materials and methods are given in Williams et al. (2000b), Williams et al. (2000a) and Creason et al. (2001). The study was conducted at an 18-story retirement facility in central Baltimore county (Towson, Maryland), about 15 km from downtown Baltimore, over a four week period from July 26 to August 21, 1998. The self-contained retirement facility included its own bank, cafeteria and dining hall, recreational rooms, on-site medical unit and sundries shop. Apartments in the facility contained 1-2 bedrooms, a kitchen/dining room, living room and bathroom.

All 56 subjects were self-sufficient and ambulatory white non-smokers ranging in age from 72-97 with a mean age of 82. Subjects excluded from the study included those with physician-diagnosed uncontrolled hypertension, coronary bypass surgery and/or heart attack within the past year, episodes of syncope in the past year, dementia, dialysis treatment, need for supplemental oxygen, having a pacemaker or being a current cigarette smoker. Of the 11

men and 45 women in the study, 36 were classified as having some degree of cardiovascular compromise, which included thyroid disease (13%), coronary disease (16%) and hypertension (43%). Eight subjects had physician-diagnosed chronic obstructive pulmonary disease. Eighteen subjects had none of the above conditions. Cooking and use of tobacco products, the two major indoor sources of $PM_{2.5}$ (Özkaynak et al., 1996), were rarely performed by study participants (Williams et al., 2000b). Subjects spent 94% of their time either inside their apartments, the retirement facility or other indoor locations, and exposure to indoor sources such as cooking, vacuuming, dusting or tobacco products totaled less than 0.5 h/day (Williams et al., 2000a).

Subjects were scheduled to visit a health-monitoring clinic in a vacant apartment in the facility in two groups of approximately 30 subjects for examination on alternate days, 3 days per week at the same time each day over the study period. This paper focuses on 5 measures of HRV in the time and frequency domain recorded at each daily visit. Six minutes of resting supine beat-to-beat heart rate data were collected after the subjects had rested in the supine position for ten minutes. HRV measures included: (1) the high frequency (HF) and (2) low frequency (LF) spectral power component of the power spectral density curve and (3) the ratio of LF/HF, (4) the standard deviation of all normal to normal (NN) beat-to-beat R wave to R wave (RR) time intervals (referred to as SDNN), and (5) the square root of the mean of squared differences between adjacent normal RR (r-MSSD). Summary statistics for HRV measurements are given in Table 1. Each of the 56 subjects participated in 9 to 12 HRV measurement sessions, with 82% participating in 12 measurement sessions.

[Table 1 about here.]

We hypothesize that lowered HRV occurs within 24 hours following exposure to higher $PM_{2.5}$ concentrations. Thus, the main model we describe examines the relationship between each day's HRV measurements and unknown mean personal $PM_{2.5}$ concentrations at time

lag of 1 day. We also consider time lags of 0 days, the moving average of days 0 and 1, and the moving average of days 0, 1 and 2. Further analysis in this paper utilizes the modeled $PM_{2.5}$ source apportionment results of Hopke et al. (2003) to examine relationships between HRV and the components of $PM_{2.5}$. Below we describe in more detail the available $PM_{2.5}$ datasets and modeled apportionment results used in the analysis.

A subset of 10 subjects wore personal exposure monitors, or PEMs, which collected daily personal $PM_{2.5}$ measurements using inertial impactor samplers. Personal monitoring was conducted on 23 days of the 27 day study period. PEMs also provided daily measurements of sulfur at the personal (10 subjects, 7-10 days/subject), apartment (10 apartments, 7-10 days/apartment) and central indoor locations (28 days). We utilize the ratio of personal to outdoor sulfur to calculate infiltration of ambient $PM_{2.5}$ indoors in Subsection 2.2.3. Daily ambient $PM_{2.5}$ and sulfur monitoring were conducted at a site 4 km north of downtown Baltimore. Hourly relative humidity, temperature and vector averaged wind speed were also collected at the ambient monitoring site. Each weather variable was averaged into daily values for this analysis.

Hopke et al. (2003) provide source apportionment results for the personal PEM data described above, as well as for a dichotomous Versatile Air Pollutant Sampler, or VAPS (URG Corporation, Chapel Hill, NC), located at the ambient site. To analyze the PEM data, Hopke et al. (2003) used the multilinear engine (ME) model of Paatero (1999) to apportion the personal PEM data into three external and three internal sources. The three external factors were identified as (1) secondary sulfate, (2) soil, and (3) unknown, which estimated unmeasured nitrate and carbon mass concentration. The three internal factors were comprised of (1) dust from gypsum board or drywall, with a high concentration of calcium and sulfur, (2) $PM_{2.5}$ associated with personal activities including time outside the facility, which was primarily unknown mass (79%) and sulfur (3%), and (3) personal care products, with a high concentration of Zn (possibly linked to talc use), Si and Ti. Tables

2 and 3 show source contribution estimates in percentage terms and in units of $\mu g/m^3$, respectively, for the PEM samples. Among the external sources, sulfate predominated, and among internal sources, $PM_{2.5}$ due to personal activities predominated. We will refer to this model output as PEM apportionment data.

To analyze the VAPS data at the ambient site, Hopke et al. (2003) utilized a PMF3 model (Paatero, 1997), a least squares approach to factor analysis. The four factors identified were (1) a combination of ammonium sulfate and ammonium nitrate, typically observed as the product of atmospheric processing of SO_2 and NO_x (2) secondary sulfate with a small (<1%) contribution of NO_3^-, (3) organic carbon, (4) motor vehicle exhaust, which includes organic carbon, elemental carbon and NO_3^-. Summary statistics for modeled components of the VAPS data are given in Table 4 in percentage terms, and in Table 5 in units of $\mu g/m^3$. We refer to this model output as ambient site apportionment data.

[Table 2 about here.]

[Table 3 about here.]

[Table 4 about here.]

[Table 5 about here.]

2.2 Bayesian hierarchical model

Three datasets describing $PM_{2.5}$ exposure are used to explore the relationship between HRV and $PM_{2.5}$ and its components: (1) total $PM_{2.5}$ measured using PEMs worn by respondents and PEMs located at an ambient site location, referred to as "total personal $PM_{2.5}$," (2) modeled output from the multilinear engine model which apportions personal PEM measurements from 10 subjects into 3 internal and 3 external components, referred to as "PEM apportionment," and (3) modeled output from the PMF3 model which partitions VAPS data

at the ambient site location into four components, referred to as "ambient site apportionment." Below we describe in detail the model used in the analyses of the total personal $PM_{2.5}$ data; modifications to this model to accommodate the other two datasets are described in Section 3.

The Bayesian hierarchical model is comprised of nested probability models organized in stages, and includes measurement error models for recorded measurements, models relating the unknown means to fixed and random covariates, and prior distributions for parameters. We lay out the hierarchical model in three parts. First, in Subsection 2.2.1, we model the unknown mean health effect as a function of personal exposure to $PM_{2.5}$, subject-specific fixed covariates, apparent temperature, random subject effects and a correlated error term. Second, in Subsection 2.2.2, personal exposure to $PM_{2.5}$ is modeled as a function of its personal ambient and non-ambient components, where personal ambient $PM_{2.5}$ exposure is expressed as a function of ambient $PM_{2.5}$ and an indoor infiltration factor. Third, in Subsection 2.2.3, sulfur concentrations at the personal and outdoor locations are used to model the indoor infiltration factor.

Each of the three parts of the hierarchical model can be described in terms of a directed graphical model (Ntzoufras, 2009; Richardson and Best, 2003), as seen in Figures 1, 2, and 3. In a directed graphical model, all modeled quantities are represented as nodes in a directed graph. Given the parent node, each node is independent of all other nodes in the graph except the descendants of that node. Ellipses denote either stochastic nodes which have a distribution or deterministic nodes which are logical functions of other nodes. Rectangles denote constants which are fixed by the design of the study. Arrows between nodes indicate which variables directly influence those nodes. A solid arrow indicates stochastic dependence while two-lined arrow denotes a logical function. Repeated parts of the graph are indicated with large boxes around relevant quantities, indicating loops through subjects (i) or time points (t).

2.2.1 Model for the unknown mean health effect

First, we specify a measurement error model for measured HRV. For subjects $i = 1, \ldots, I$, measured on days $t = 1, \ldots, T$, let $Z_{i,t}$ be the normally distributed HRV measurement for subject i on day t, with unknown mean $H_{i,t}$ and variance τ^Z.

$$Z_{i,t}|H_{i,t}, \tau^Z \sim N(H_{i,t}, \tau^Z) \tag{1}$$

Measures of HRV that we consider are log 10 transformed values of SDNN, r-MSSD, LF and HF as well as the ratio of LF to HF; log 10 transformations of these variables are common in the literature, reflecting the right skewness of the sampling distribution of these measurements. These five measures of HRV are treated in separate models; for simplicity, we refer to each of them in general terms as $Z_{i,t}$. In the graphical model shown in Figure 1, HRV measurements $Z_{i,t}$ are represented as a rectangle, with solid arrows indicating stochastic dependence between $Z_{i,t}$ and the parameters of its normal distribution, the mean, $H_{i,t}$, and variance, τ^Z.

[Figure 1 about here.]

The mean HRV for subject i on day t, $H_{i,t}$, is related to fixed and random covariates via a linear model in equation (2). Fixed effects in the model include the overall mean, the age of subject i, AGE_i, an indicator of cardiovascular compromise for subject i, CV_i, and the gender of subject i, SEX_i. We include a subject level random intercept, b_i, to represent subject specific permanent effects for subject i.

Unknown mean HRV on day t is also taken to be a function of apparent or "perceived" temperature on day t, A.TEMP_t, which we calculate as linear in temperature (TEMP) and quadratic in dew point temperature (TEMP.DEW) as follows: $-2.653 + 0.994(\text{TEMP}) +$

0.0153(TEMP.DEW)^2 (O'Neill et al., 2003). Here, dew-point temperature was calculated using the well-known Magnus-Tetens approximation. We represent nonlinearity in temperature using a natural cubic spline (Hastie and Tibshirani, 1990) as in Park et al. (2005). The basis, $\boldsymbol{h}()$, consists of four basis functions with knots at the median and quartiles. $\boldsymbol{\Omega}$ is a vector of coefficients multiplying the associated vector of the natural spline basis function $\boldsymbol{h}()$. Other random terms in equation (2) include unknown total personal $PM_{2.5}$ exposure on the day previous to the HRV measurement, $M_{i,t-1}^{P.TOT}$, and an error term, $\varepsilon_{i,t}^{H}$.

$$\begin{aligned} H_{i,t} &= \theta_0 + \theta_1 \text{AGE}_i + \theta_2 \text{CV}_i + \theta_3 \text{SEX}_i + \boldsymbol{h}(\text{A.TEMP}_t)^T \boldsymbol{\Omega} + \theta_4 M_{i,t-1}^{P.TOT} \\ &\quad + b_i + \varepsilon_{i,t}^{H} \end{aligned} \tag{2}$$

As seen in Figure 1, the unknown total personal exposure to $PM_{2.5}$ for subject i at time $t-1$, $M_{i,t-1}^{P.TOT}$, impacts measured HRV, $Z_{i,t}$, through the unknown mean HRV, $H_{i,t}$. In Subsection 2.2.2 we describe how $M_{i,t-1}^{P.TOT}$ depends on its ambient and non-ambient components (Figure 2).

Errors, $\varepsilon_{i,t}^{H}$, follow a continuous time AR(1) autocorrelation structure with autocorrelation function $\rho^H(\Delta)$.

$$\varepsilon_{i,t}^{H} = \rho^H(\Delta)\varepsilon_{i,t-1}^{H} + u_0^H \tag{3}$$

We define the function $\rho^H(\Delta) = \exp(-\alpha\Delta)$, where Δ is the distance in days between the HRV measurement at time t and the previous HRV measurement for subject i (Diggle, 1988). The parameter α is taken to be common across subjects and time. u_0^H is defined as a white noise process.

2.2.2 Model for unknown total personal exposure to $PM_{2.5}$

In Figure 1, we model the links between unknown total personal exposure to $PM_{2.5}$ ($M_{i,t-1}^{P.TOT}$), unknown HRV($H_{i,t}$), and measured HRV ($Z_{i,t}$). In this subsection, we describe the links between the indoor infiltration factor (γ), unknown total personal exposure to $PM_{2.5}$ ($M_{i,t-1}^{P.TOT}$), measured total personal $PM_{2.5}$ exposure ($Y_{i,t}^{P.TOT}$), as shown in Figure 2. The measured total personal $PM_{2.5}$ concentration received by individual i at time t, $Y_{i,t}^{P.TOT}$, is represented by a rectangle in Figure 2. $Y_{i,t}^{P.TOT}$ is taken to be normally distributed with mean $M_{i,t}^{P.TOT}$ and variance $\tau^{M.P.TOT}$.

$$Y_{i,t}^{P.TOT} | M_{i,t}^{P.TOT}, \tau^{M.P.TOT} \sim N(M_{i,t}^{P.TOT}, \tau^{M.P.TOT}) \tag{4}$$

[Figure 2 about here.]

Daily personal $PM_{2.5}$ concentrations for 10 subjects are available, totaling 119 observations. These observations are used to characterize personal $PM_{2.5}$ concentrations for the remaining 46 subjects who had HRV measurements but no personal $PM_{2.5}$ concentrations.

Measured ambient $PM_{2.5}$ concentrations at time t, Y_t^A, are taken to be normally distributed with mean M_t^A and variance $\tau^{M.A}$.

$$Y_t^A | M_t^A, \tau^{M.A} \sim N(M_t^A, \tau^{M.A}) \tag{5}$$

Other distributional choices, such as the lognormal, for values of recorded $PM_{2.5}$ concentrations and for values of recorded sulfur observations described in Subsection 2.2.3 are possible and would more appropriately reflect the non-negativity in pollutant measurements and deviations from normality across more general situations. However, for the data available in

this 27-day single season study, the normality assumption does not raise serious problems in terms of capturing the behavior of mean $PM_{2.5}$ levels and their association with HRV responses.

As in Wallace and Williams (2005), the unknown total personal exposure to $PM_{2.5}$ for subject i at time t, $M_{i,t}^{P.TOT}$, can be broken down into its ambient ($M_t^{P.A}$) and non-ambient components ($M^{P.NA}$).

$$M_{i,t}^{P.TOT} = M_t^{P.A} + M^{P.NA} \tag{6}$$

This logical relationship is depicted in Figure 2 with two-lined arrows. The ambient component is due to outdoor sources, and the non-ambient component is due to nonoutdoor sources, such as indoor sources in the home and other locations, and sources associated with resuspension of particles on clothes and indoor surfaces.

Unknown personal $PM_{2.5}$ of non-ambient origin, $M^{P.NA}$, which is difficult to measure directly, is taken to have a common distribution across subjects and days.

$$M^{P.NA}|\mu^{MPNA}, \tau^{MPNA} \sim N(\mu^{MPNA}, \tau^{MPNA}) \tag{7}$$

The unknown personal $PM_{2.5}$ exposure of ambient origin, $M_t^{P.A}$, is taken to have a common distribution across subjects for each day t, and is the product of a $PM_{2.5}$ infiltration factor, γ, and the concurrent unknown $PM_{2.5}$ concentration at the ambient monitoring site, M_t^A. Estimation of the infiltration factor is described in Subsection 2.2.3.

$$M_t^{P.A} = \gamma M_t^A \tag{8}$$

The unknown ambient $PM_{2.5}$ concentration time series, M_t^A, is taken to be normally distributed with mean, $\mu_t^{M.A}$, and variance,$\tau^{mu.M.A}$.

$$M_t^A|\mu_t^{M.A}, \tau^{mu.M.A} \sim N(\mu_t^{M.A}, \tau^{mu.M.A}) \tag{9}$$

Meteorological covariates determining the unknown mean $PM_{2.5}$ time series $\mu_t^{M.A}$ include: vector averaged wind speed (W_t) and its one-day lag, which accounts for the magnitude and direction of particle sources and day-to-day carry-over of $PM_{2.5}$ concentrations; relative humidity (U_t) and its one-day lag, which may increase available water vapor to condense on aerosol particles, allowing uptake of sulfates and nitrates (Finlayson-Pitts and Pitts, 1999); and a weekday/weekend effect (D_t), which accounts for traffic patterns. In equation (10), $\mu_t^{M.A}$ is written as a linear combination of these covariates. Temperature data were not incorporated into the model for unknown mean $PM_{2.5}$ because early model runs incorporating temperature and its one-day lag showed high posterior cross-correlation between its coefficients and the coefficients of humidity and its one-day lag. Autocorrelation in errors is modeled using an AR(1) structure.

$$\mu_t^{M.A} = \beta_0^M + \beta_1^M U_t + \beta_2^M U_{t-1} + \beta_3^M W_t + \beta_4^M W_{t-1} + \beta_5^M D_t + \rho^{M.A} M_{t-1}^A + u_0^{M.A} \tag{10}$$

where $\rho^{M.A}$ models the autocorrelation between successive observations, and $u_0^{M.A}$ is a normally distributed white noise sequence. Other models relating outdoor $PM_{2.5}$ concentrations to meteorological variables are possible (Huang et al., 2005; Holloman et al., 2004) and may more accurately reflect spatial and temporal variation of $PM_{2.5}$.

Equation (2) as written accounts for an effect due to personal total $PM_{2.5}$ at lag one day. Modification to account instead for an effect at lag 0 is straightforward. We consider a

moving average of personal total $PM_{2.5}$ concentrations over lags 0 and 1 by replacing $M_{i,t-1}^{P.TOT}$ in equation (2) with $M_{i,01MA}^{P.TOT}$, where

$$M_{i,01MA}^{P.TOT} = (M_{i,t}^{P.TOT} + M_{i,t-1}^{P.TOT})/2 \tag{11}$$

A moving average of personal $PM_{2.5}$ concentrations over lags 0, 1 and 2 was constructed in a similar way.

2.2.3 Estimation of the unknown indoor infiltration factor

In the previous section we modeled the links between the indoor infiltration factor (γ), unknown total personal exposure to $PM_{2.5}$ ($M_{i,t-1}^{P.TOT}$), and measured total personal $PM_{2.5}$ exposure ($Y_{i,t}^{P.TOT}$). As illustrated in Figure 3, we now show how the indoor infiltration factor is linked to measured personal sulfur concentrations (X_t^P) and at the ambient site (X_t^A) through the unknown means of sulfur concentrations at the personal level (S_t^P) and at the ambient site (S_t^A), respectively.

As seen in Figure 2, the indoor infiltration factor, γ in equation (8), influences the unknown total personal $PM_{2.5}$ exposure through the unknown ambient personal $PM_{2.5}$ exposure. In this subsection, we describe estimation of the indoor infiltration factor. The ratio of unknown indoor to unknown outdoor sulfur concentrations is used to approximate the $PM_{2.5}$ infiltration factor (Wilson and Brauer, 2006; Strand et al., 2006; Wallace and Williams, 2005). This is a valid approximation provided there are no indoor sources of sulfur and the particle size distributions are similar.

$$S_t^P = \gamma S_t^A + \varepsilon^S \tag{12}$$

We include a normally distributed white noise error term, ε^S. The logical relationship between the indoor infiltration factor and its personal and ambient sulfur counterparts is shown in Figure 3, where unknown personal exposure to sulfur (S_t^P) is influenced by both the indoor infiltration factor (γ) and the unknown ambient sulfur concentration (S_t^A). As seen in Figure 3, changes in meteorology (daily humidity, vector wind speed and temperature and their respective one-day lags) are linked to the indoor infiltration factor through the unknown ambient (S_t^A) and unknown personal (S_t^P) sulfur concentrations.

Available indoor sulfur data include measurements at personal, apartment and central indoor locations; Wallace and Williams (2005) state that these three are very similar for the purposes of determining indoor infiltration of ambient $PM_{2.5}$. In the analysis, we refer to "personal sulfur concentrations" as measurements taken at the personal, apartment and central indoor locations.

Multiple human and environmental exposure factors have the potential for influencing daily $PM_{2.5}$ residential infiltration. As discussed in Wallace and Williams (2005), the indoor infiltration factor is known to vary across days within seasons for a single detached residence. For the single-season retirement home study considered here, we take γ to be common over subjects, apartments and days. An assumption of a common indoor infiltration factor is reasonable for these data for a number of reasons: (1) the study was conducted in a communal living situation with interior entry doors indicating good probability of consistent study population behavior with respect to residence heating and air conditioning operations as well as other personal exposure factors (cooking, cleaning, grooming type of behaviors); (2) $PM_{2.5}$ concentrations in individual apartments were highly correlated with those at the ambient site monitor; (3) there was little overall variability in outdoor temperatures over the study period (24.7 +/- 3.6 degrees Celsius) (Wallace et al., 2006). In addition, in another analysis of these data, Landis et al. (2001) report that the apparent variability of indoor/outdoor sulfate ratios over the nearly month-long monitoring period rarely differed

by more than 10% among all participants on a given day.

Measurement error models for recorded concentrations of personal and ambient sulfur are as follows. Let X_t^P denote the normally distributed measured personal sulfur concentration at time t, with mean S_t^P and variance $\tau^{S.P}$. Sulfur concentrations at the ambient monitoring site at time t, X_t^A, are taken to be normally distributed with mean S_t^A and variance $\tau^{S.A}$.

$$X_t^P | S_t^P, \tau^{S.P} \sim N(S_t^P, \tau^{S.P}) \tag{13}$$

$$X_t^A | S_t^A, \tau^{S.A} \sim N(S_t^A, \tau^{S.A}) \tag{14}$$

[Figure 3 about here.]

Similar to the modeling of ambient $PM_{2.5}$ concentrations, the unknown ambient sulfur concentration, S_t^A, is taken to be normally distributed with mean, $\mu_t^{S.A}$, and variance, $\tau^{mu.S.A}$.

$$S_t^A | \mu_t^{S.A}, \tau^{mu.S.A} \sim N(\mu_t^{S.A}, \tau^{mu.S.A}) \tag{15}$$

The unknown mean of the ambient sulfur time series, $\mu^{S.A}$, is modeled as a linear combination of meteorological covariates and their one-day lags including temperature (TEMP_t), vector averaged wind speed (W_t), and relative humidity (U_t), as well as a weekday/weekend effect (D_t). An AR(1) structure is used to model autocorrelation in errors.

$$\begin{aligned} \mu_t^{S.A} &= \beta_0^S + \beta_1^S \mathrm{U}_t + \beta_2^S \mathrm{U}_{t-1} + \beta_3^S \mathrm{W}_t + \beta_4^S \mathrm{W}_{t-1} + \beta_5^S \mathrm{TEMP}_t + \beta_6^S \mathrm{TEMP}_{t-1} \\ &\quad + \beta_7^S \mathrm{D}_t + \rho^{S.A} S_{t-1}^A + u_0^{S.A} \end{aligned} \tag{16}$$

where $\rho^{S.A}$ models the autocorrelation between successive observations, and $u_0^{S.A}$ is a normally distributed white noise sequence.

2.3 Prior Distributions

Non-informative prior distributions were used in model runs where prior information about parameter values was not available. To define priors on many of the variance parameters in our model, we follow Gelman (2006), who suggests using a uniform prior on the hierarchical standard deviation. The standard deviation of measured HRV responses, $\sqrt{\tau^Z}$ in equation (1), had a Uniform(0.0,1.0E4) prior. Standard deviation terms for recorded HRV measures (eq. (1)) and measured pollutant concentrations (eqs. (4), (5), (13), (14)),were taken to be uniformly distributed on (0.0,1.0E4). Regression coefficients in equations (2), (16) and (10) had N(0,100) prior distributions. Subject level random intercepts, b_i, in equation (2) are taken to have normal priors with mean 0 and variance τ_b, where $\sqrt{\tau_b} \sim$ Uniform(0.0,1.0E4). The standard deviation of the total measured personal $PM_{2.5}$, $\sqrt{\tau^{M.P.TOT}}$, and the standard deviation of the total measured ambient $PM_{2.5}$, $\sqrt{\tau^{M.A}}$, had Uniform(0.0,1.0E4) priors, as did the standard deviation of the mean personal $PM_{2.5}$ of non-ambient origin, $M^{P.NA}$ in equation (7). Mean personal $PM_{2.5}$ of non-ambient origin, μ^{MPNA} in equation (7), was taken to have a N(5.0,100) prior, where 5.0 μg/m^{3}was used as a rough estimate of a daily personal non-ambient $PM_{2.5}$ exposure. In the error equation (3), the parameter α was given a Uniform(0,20) prior. $PM_{2.5}$ indoor infiltration (eqs. (8), (12)) had a Uniform(0.0, 1.0) prior, with error term $\varepsilon^S \sim N(0, 100)$.

The standard deviations of measured personal sulfur concentrations, $\sqrt{\tau^{S.P}}$ in equation (13), measured ambient sulfur concentrations, $\sqrt{\tau^{S.A}}$ in equation (14), and mean sulfur concentrations, $\sqrt{\tau^{mu.S.A}}$ in equation (15), were each given Uniform(0.0,1.0E4) priors. In the equation for mean ambient site sulfur concentrations as a function of meteorological variables (eq. 16), priors were $\rho^{S.A} \sim$ Uniform(-1.0,1.0) and $u_0^{S.A} \sim N(0, 100)$. To achieve

an AR structure in the error for mean sulfur(eq. 16), a value of mean sulfur at time 0, $\mu_0^{S.A}$, was needed; this value had prior $\mu_0^{S.A} \sim N(2.3, 100)$, where 2.3 μg/m^{3}was the mean of the sulfur values across the available days.

The standard deviations of measured personal $PM_{2.5}$ concentrations, $\sqrt{\tau^{M.P}}$ in equation (4), measured ambient site $PM_{2.5}$ concentrations, $\sqrt{\tau^{M.A}}$ in equation (5), and mean $PM_{2.5}$ concentrations, $\sqrt{\tau^{mu.M.A}}$ in equation (9), were taken to have Uniform(0.0,1.0E4) priors. In equation (10), priors were $\rho^{M.A} \sim$ Uniform(-1.0,1.0) and $u_0^{M.A} \sim N(0, 100)$. $\mu_0^{M.A}$ had a N(18,100) prior, where 18 μg/m^{3}was the mean of the available $PM_{2.5}$ values.

2.4 Implementation Details

Posterior distributions of parameters were obtained using Markov Chain Monte Carlo (MCMC) methods as implemented in WinBUGS software (Spiegelhalter et al., 2003) using an interface with R, an Open Source system for statistical computing and graphics (Gelman et al. (2003), `http://www.r-project.org/`). WinBUGS code is given in Appendix A.

The MCMC algorithm was run using 3 chains for at least 5000 iterations each and up to 10,000 iterations each. For 5000 iteration runs, the first 2500 draws were used to assess burn-in; sample traces suggested convergence to the stationary distribution for all parameters. To create approximately independent samples, inferences about model parameters are based on every 8th sample. For each parameter, 1000 samples from the posterior distribution were retained for inference. For these 1000 samples, the estimate of Monte Carlo error as calculated by consistent batch means (Jones et al., 2006) was less than 5% of its respective standard deviation, indicating that estimation error was significantly less than uncertainty in the true parameter values. Also calculated for each model variable was the potential scale reduction factor (Gelman and Hill, 2007), $\hat{R}$, which approximates the variance of the mixture of the three chains divided by the within chain variance. Values of $\hat{R}$ less than 1.1 indicate approximate convergence of the algorithm and adequate mixing of the parallel chains. For

each model variable, a crude measure of effective sample size was calculated; values of at least 100 indicated convergence of the algorithm and usually corresponded to $\hat{R}$ values less than 1.1.

Convergence in MCMC parameter estimates was achieved by fixing four parameters which were poorly identified by the model. In equation (3), $\varepsilon_{i,0}^{H}$, the error in the mean HRV response at time 0 for subject i, was fixed at 0.0 μg/m^3. The mean of the normally distributed first observation of the unknown ambient site $PM_{2.5}$ time series, $\mu_0^{M.A}$ in equation (10), was fixed at 18 μg/m^3. The autocorrelation between successive outdoor $PM_{2.5}$ observations, $\rho^{M.A}$ in equation (10), was fixed at 0.0. The variance of the measured ambient site $PM_{2.5}$ concentrations, $\tau^{M.A}$ in equation (5), was fixed at 9.0. Sensitivity analysis of model output from combinations of these parameters set at different values showed little impact on resultant posterior HRV estimates.

3 Results

Inference in Bayesian hierarchical models is based on posterior distributions which allow for direct probability statements about parameters of interest. As a result, we are able to construct posterior intervals, or Bayesian confidence intervals, which give the probability that the parameter lies in an interval given the data. We note that the usual frequentist confidence interval does not allow this type of interpretation. In this section, we quantify and characterize the strength of the hypothesized inverse relationship between HRV and personal $PM_{2.5}$ exposure and its components. In terms of Bayesian inference, we present probabilistic statements providing evidence that coefficients describing $PM_{2.5}$ and its components are negative.

3.1 Total personal $PM_{2.5}$

3.1.1 Analysis and model checking for the lag 1 model

We begin with a detailed analysis of the model described in Section 2.2 for total personal $PM_{2.5}$ at lag one day. Figure 4 and Table 6 show the posterior percentage change in HRV associated with a 6.5 $\mu g/m^3$(1 SD) increase in personal total $PM_{2.5}$ exposure at lag one day. For all HRV measures except for the ratio of LF to HF, an effect of lowered HRV with increased personal exposure to $PM_{2.5}$ is seen at lag one day; the posterior probability that the coefficient of personal total $PM_{2.5}$ (θ_4 in eq. (2)) is negative is at least 0.72 for four HRV measures.

[Table 6 about here.]

[Figure 4 about here.]

Runs of the model for the five health effects at lag 1 gave the posterior distribution of the unitless indoor infiltration rate, γ in equations (8) and (12), mean 0.38, standard deviation 0.02, and 95% posterior interval (0.35, 0.42). Consideration of other lags (lag 0, and moving averages of lags 0 and 1 as well as 0-2 days) gave posterior means for γ between 0.37-0.39, with posterior standard deviation 0.02. These agree well with previous calculations of indoor infiltration that were based only on indoor/outdoor $PM_{2.5}$ concentrations (not sulfur data) reported in McBride et al. (2007); posterior mean infiltration values for individual residences had overall mean 0.37, with standard deviations between 0.04-0.06, and infiltration at the central indoor site had posterior mean 0.38 with standard deviation 0.03.

Across the five HRV measures, the non-ambient component of personal $PM_{2.5}$ exposure, μ^{MPNA} in equation (7), had posterior mean between 4.90 - 5.25 $\mu g/m^3$, with posterior standard deviation between 8.02 - 8.52 $\mu g/m^3$. Partitioning the effects of personal $PM_{2.5}$ of ambient origin and of non-ambient origin on HRV is difficult using this dataset for total

personal $PM_{2.5}$ exposure because personal $PM_{2.5}$ of non-ambient origin was not measured directly. Runs of the model with separate regression coefficients in equation (2) for ambient and non-ambient personal $PM_{2.5}$ were unable to resolve and separately estimate the two components; high posterior cross-correlations among regression coefficients were seen. We address the role of ambient and non-ambient personal $PM_{2.5}$ components in our analysis of the personal PEM apportionment modeling output in Section 3.2.

Quality of model fit at different stages of the hierarchical model was assessed by calculation of posterior distributions of residuals for HRV, total personal $PM_{2.5}$ exposure, and $PM_{2.5}$ exposure at the ambient site. Draws from the posterior distribution of each of these posterior quantities were subtracted from observed values, and 95% posterior intervals were calculated. Of note, there were 658 posterior intervals calculated for HRV across 56 subjects, while there were 89 posterior intervals for total personal $PM_{2.5}$ exposure across 10 subjects. For HRV as well as total personal $PM_{2.5}$ exposure, across all five HRV measures, these intervals covered zero roughly half the time, with the remaining half split roughly equally between underfit (intervals lying above zero) and overfit (intervals lying below zero). Results were similar when stratifying by day or subject, and patterns in lack of fit by day or subject were not evident. These results suggest that the model may not reflect the extremes in individual $PM_{2.5}$ exposure, likely because total $PM_{2.5}$ exposure is modeled in equation (6) as the sum of a non-ambient component, common across days and subjects, and an ambient component, common across subjects and varying across days. More detailed models of total $PM_{2.5}$ exposure might attempt separate estimation of total $PM_{2.5}$ exposure by subject and day, possibly including interior sources, time varying air exchange rates, as well as individual activity patterns and individual apartment $PM_{2.5}$ measurements (McBride et al., 2007). The inclusion of more subject-specific parameters was not well accommodated in the current modeling framework due to poorly identified parameters, but could likely be achieved with more data availability. However, posterior predictive checks using the posterior predictive

distribution of measured HRV values $Z_{i,t}$ in equation (1) showed that across the five HRV measures, 95% to 96% of posterior predictive intervals for $Z_{i,t}$ contained the measured HRV values, indicating that the HRV values could plausibly have come from the model.

For $PM_{2.5}$ exposure at the ambient site location, all posterior residual intervals covered zero. Figure 5 shows the time series of measured $PM_{2.5}$ at the ambient site (dots) coplotted with 95% posterior intervals for $PM_{2.5}$ at ambient site (grey bands) and modeled mean $PM_{2.5}$ at ambient site (dotted line). Varying degrees of uncertainty in posterior estimates of $PM_{2.5}$ at the ambient site are reflected in the width of the grey bands.

[Figure 5 about here.]

Across the five HRV measures, positive biases were seen in the posterior distributions of the white noise sequences in the models for the mean of ambient $PM_{2.5}$ (eq. (10)) and for the mean of ambient sulfur (eq. (16)). The posterior mean of $u^{M.A}$ ranged from 7.8 - 8.5 $\mu g/m^3$, with standard deviations between 8.9-9.2 $\mu g/m^3$. The posterior mean of $u^{S.A}$ ranged from -7.5 to -7.1 $\mu g/m^3$with standard deviations between 7.2-7.7 $\mu g/m^3$. The cause of the bias is likely that the regression models for mean pollutant concentrations as a function of meteorology are underspecified given the complexity of pollutant formation.

While relationships were apparent between the four time HRV responses and personal total $PM_{2.5}$ at lag one day, consideration of other lags only showed a relationship between the ratio of low to high frequency heart rate variability (LF/HF) and personal total $PM_{2.5}$. Effects were found for total personal $PM_{2.5}$ at lag zero, at the moving average of lag zero and lag one day, and at the moving average of lags zero, lag one and lag two days. Results for LF/HF are shown in Table 7.

[Table 7 about here.]

3.2 Personal PEM apportionment

Hopke et al. (2003) calculated source apportionment results for personal PEM data, identifying three external and three internal factors. We begin our analysis by comparing the relative importance of the total internal versus total external sources in order to assess the impact on HRV of $PM_{2.5}$ of non-ambient origin versus $PM_{2.5}$ of ambient origin, an important question from a regulatory policy perspective. For a subset of 10 subjects, 20 days on average of apportionment data are available (a range of 12 - 23 days per subject). We adapt the model for the unknown total personal exposure to $PM_{2.5}$ in Subsection 2.2.2 as follows. We replace the equation for measured total personal $PM_{2.5}$ (eq. (4)) and measured total ambient $PM_{2.5}$ (eq. (5)) with equations for measured personal external and personal internal source $PM_{2.5}$ concentrations.

$$Y_{i,t}^{P.EXT} | M_t^{P.A}, \tau^{M.P.EXT} \sim N(M_t^{P.A}, \tau^{M.P.EXT}) \tag{17}$$

$$Y_{i,t}^{P.INT} | M_t^{P.NA}, \tau^{M.P.INT} \sim N(M_t^{P.NA}, \tau^{M.P.INT}) \tag{18}$$

where the variance terms $\tau^{M.P.EXT}$ and $\tau^{M.P.INT}$ are given non-informative uniform priors. In the regression equation (2), we include separate coefficients for internal $M_t^{P.NA}$ and external $M_t^{P.A}$ factors. $M_t^{P.NA}$ is then modeled as in equation (7). $M_t^{P.A}$ is modeled as in equation (9), with mean driven by meteorology as in equation (10). Since the measured personal internal and external $PM_{2.5}$ components account for personal $PM_{2.5}$ exposure after infiltration of $PM_{2.5}$ indoors, we omit modeling of sulfur infiltration as described in Subsection 2.2.3 (eqs. (13), (14), (8), (12), (15), (16)).

In terms of the graphical model, we alter Figure 2 by eliminating links to $Y_{i,t}^{P.TOT}$ as well as the link to the indoor infiltration factor γ and the link to Figure 3. Two new data sources, $Y^{P.EXT}$ and $Y^{P.INT}$ are incorporated, as seen in Figure 6. Figure 1 is unchanged.

[Figure 6 about here.]

Results from the adapted model for four time lags and five HRV response variables (20 models) are given in Table 8. Posterior estimates are given for the percentage change in HRV associated with a 4 $\mu g/m^3$(1 SD) increase in personal $PM_{2.5}$ exposure due to external sources. Also shown are posterior probabilities that coefficients of external personal and internal personal $PM_{2.5}$ concentrations are negative. Across the four lags for the responses HF, LF , LF/HF and SDNN, the posterior probabilities that the internal source coefficients are negative are at most 0.56, providing little evidence to suggest that personal $PM_{2.5}$ of non-ambient origin has an effect on these HRV responses. For the HRV response variables SDNN and r-MSSD, there may be some association between increased personal $PM_{2.5}$ of non-ambient origin and lowered r-MSSD values, with posterior probabilities ranging between 0.64-0.70 across the four lags. Given the overall weak association between personal $PM_{2.5}$ of non-ambient origin and HRV for the majority of the HRV measures, we drop the internal source component from further modeling.

[Table 8 about here.]

We next consider the relationships between HRV and the three external factors in the PEM apportionment data, identified as (1) secondary sulfate, (2) soil, and (3) unknown, which estimated unmeasured nitrate and carbon mass concentration. We modify the adapted model described above by dropping equation (18) for personal $PM_{2.5}$ of non-ambient origin, and substituting measurements for each of the three external components in equation (17). Table 9 shows posterior estimates of percentage change in HRV associated with a 1 $\mu g/m^3$increase in personal exposure to $PM_{2.5}$ components based on personal PEM apportionment. A 1 $\mu g/m^3$increase was chosen since the standard deviations of the three components ranged from 0.2 - 4.6 $\mu g/m^3$(Table 3). Also shown are posterior probabilities that coefficients of personal $PM_{2.5}$ component concentrations are negative. Of 60 possible models

(5 HRV measures, 4 time lags, 3 $PM_{2.5}$ components), the 11 models shown have posterior probabilities greater than 70% that coefficients of personal $PM_{2.5}$ component concentrations are negative. Based on these results, there appears to be a relationship between the five HRV measures and the soil component of personal $PM_{2.5}$ of ambient origin at lag 0 days, with posterior probabilities ranging from 0.73 - 0.91. There appears also to be an association between LF/HF and soil for a moving average of lags 0 and 1 day. Sulfate shows an effect for the HRV measures HF and r-MSSD at a moving average of lags 0 and 1 day. The component labeled unknown, comprised of unmeasured nitrate and carbon mass concentration, appears to have an effect on LF/HF at 3 different lags.

[Table 9 about here.]

3.3 Ambient site apportionment

We consider the relationships between HRV and the four $PM_{2.5}$ components identified in Hopke et al. (2003)'s analysis of the ambient site VAPS data. The four factors identified were (1) a combination of ammonium sulfate and ammonium nitrate, typically observed as the product of atmospheric processing of SO_2 and NO_x (2) secondary sulfate with a small (<1%) contribution of NO_3^-, (3) organic carbon, (4) motor vehicle exhaust, which includes organic carbon, elemental carbon and NO_3^-.

To accommodate the ambient site apportionment data, we make minor modifications to the model for unknown total personal exposure to $PM_{2.5}$ in Subsection 2.2.2. First, equation (4) is eliminated since component-wise personal $PM_{2.5}$ concentrations are not available from VAPS monitors. Based on the findings of Subsection 3.2 on the weak relationship between internal sources and HRV, we do not consider $PM_{2.5}$ of non-ambient origin in this analysis, thus eliminating equations (6) and (7). In each model run we let the term $M_{i,t-1}^{P.TOT}$ in regression equation (2) refer to unknown mean personal concentration for a single $PM_{2.5}$

component, and we let Y_t^A refer to the measured ambient concentration for that component. We thus modify the graphical model in Figure 2 by removing the link to $Y_{i,t}^{P.TOT}$ as well as the link to $M^{P.NA}$.

Of 80 possible models (5 HRV measures, 4 time lags, 4 $PM_{2.5}$ components), 17 models had posterior probabilities greater than 75% that the coefficients of personal $PM_{2.5}$ component concentrations are negative. For these 17 models, Table 10 shows posterior estimates of percentage change in HRV associated with a 3 $\mu g/m^3$increase in personal exposure to $PM_{2.5}$ components based on ambient site apportionment. A 3 $\mu g/m^3$increase was chosen because standard deviations for the four components ranged from 0.14 $\mu g/m^3$for organic carbon to 9.57 $\mu g/m^3$for sulfate (Table 5). Nitrate is seen to have an effect at all lags except lag 0 for the HRV measures of HF, LF, r-MSSD and SDNN. Sulfate appears to have an effect on HRV at lag one day for all effects except LF/HF. A relationship between LF/HF and sulfate is seen for the moving average of days 0, 1 and 2. One other result of note was an effect of nitrate at lag 0 for r-MSSD, which showed a posterior probability of 0.72 that the coefficient of personal nitrate concentrations was negative. The effect of organic carbon (OC) on HRV was not clear from model results, since 8 of 20 models incorporating OC showed some issues with convergence even after 20,000 iterations. All other OC models gave posterior probabilities less than 0.68 that the coefficient of personal OC concentrations were negative. None of the models involving motor vehicles showed a relationship with HRV.

[Table 10 about here.]

4 Discussion

The Bayesian hierarchical model presented here integrates data from personal HRV measurements, $PM_{2.5}$ concentrations on personal monitors and at an ambient site, and sulfur data from indoor and ambient site locations. In this way, the model allows for propagation

of all sources of uncertainty in each of the three parts of the model seen in Figures 1, 2, 3 onto estimation of key exposure and health effect parameters.

We found a strong relationship between decreased HRV, as measured by HF, LF, r-MSSD, and SDNN, and total personal exposure to $PM_{2.5}$ at lag one day. This agrees with and expands upon the mixed effects modeling results for the same data by Creason et al. (2001), who found a relationship at lag one day for HF and LF. These results also agree with the panel study analysis conducted by Luttman-Gibson et al. (2006) which found associations for non-smoking seniors between HRV at lag 1 day and SDNN, r-MSSD, LF and HF. We also found an effect on LF/HF due to total personal $PM_{2.5}$ at shorter lags (lag 0 days) and longer lags (moving averages of days 0-1 and days 0-2).

The PEM apportionment modeling results of Hopke et al. (2003), which break personal $PM_{2.5}$ exposure into internal and external sources, allow characterization of the relative influences of ambient and non-ambient personal $PM_{2.5}$ exposure on HRV. In Wilson and Brauer (2006)'s examination of panel study data from Vancouver, Canada, a method based on the mass balance equation was developed to estimate separately the ambient and non-ambient components of personal $PM_{2.5}$ exposure. Wilson and Brauer (2006) report that for some health effects, resolution of total personal $PM_{2.5}$ exposure into its ambient and non-ambient parts showed that the ambient component was significantly associated with health effects. We also found that under the Bayesian hierarchical model, for a majority of the measured health effects and lags considered, the effect of internal non-ambient sources of $PM_{2.5}$ on HRV was minimal. Thus, our further modeling omitted personal exposure to non-ambient sources of $PM_{2.5}$.

The PEM apportionment data were then used to assess the relative importance of external sources of soil, sulfate and unknown source categories to HRV. A consistent effect of soil at short time scales (lag 0) was found across all five HRV measures, with an additional effect found for LF/HF at a longer lag (moving average of days 0 and 1). An effect of sulfate on

HRV was seen for HF and r-MSSD at the moving average of lags 0 and 1 day. An analysis of $PM_{2.5}$ data from the Harvard Six Cities study by Laden et al. (2000) did not find a similar association of crustal material with mortality.

Consideration of the ambient site apportionment data of Hopke et al. (2003) indicated effects of nitrate on HRV at lags 1 day, and moving averages of days 0 and 1 and days 0-2 for all but the ratio LF/HF. Sulfate had an effect on HRV at lag 1 day for four HRV measures (HF, LF, r-MSSD, SD of NN) and for LF/HF at a moving average of days 0-2. This is consistent with Luttman-Gibson et al. (2006)'s analysis of a panel study of 32 senior adults, which found a significant association between SO_4^{-2} and HRV at lag one day. In their analysis of the relationship between daily mortality in Phoenix, AZ and apportioned $PM_{2.5}$ using a number of methods, Mar et al. (2006) found secondary sulfate and traffic to have the largest cardiovascular mortality effect size.

In the analyses presented here, we fit separate models for each $PM_{2.5}$ component contribution, using diffuse priors to represent prior uncertainty about modeled $PM_{2.5}$ component concentrations. A more robust approach to propagating uncertainty in estimated source contributions through to estimation of health effects would be to jointly fit receptor models and health effects models. Such an approach is pursued by Nikolov et al. (2007), who used data from a concentrator study investigating the relationship between ST-segment, a cardiovascular outcome, and major sources of $PM_{2.5}$ in Boston; in this study a Bayesian structural equation approach was used to jointly fit a multivariate receptor model and health outcome model.

5 References

References

Cavallari, J., Fang, S., Eisen, E., Schwartz, J., Hauser, R., Herrick, R., and Christiani, D. (2008), "Time course of heart rate variability decline following particulate matter exposures in an occupational cohort," *Inhal Tox*, 20, 415–422.

Creason, J., Neas, L., Walsh, D., Williams, R., Sheldon, L., Liao, D., and Shy, C. (2001), "Particulate matter and heart rate variability among elderly retirees: the Baltimore 1998 PM study," *J Expo Anal Environ Epidemiol*, 11, 1–7.

Dekker, J., Schouten, E., Klootwijk, P., Pool, J., Swenne, C., and Kromhout, D. (1997), "Heart Rate Variability from Short Electrocardiographic Recordings Predicts Mortality from All Causes in Middle-aged and Elderly Men: The Zutphen Study," *Am J Epidemiol*, 145, 899–908.

Devlin, R., Ghio, A., Kehrl, H., Sanders, G., and Cascio, W. (2003), "Elderly humans exposed to concentrated air pollution particles have decreased heart rate variability," *Eur Respir J*, 21, 76s–80s.

Diggle, P. (1988), "An approach to the analysis of repeated measurements," *Biometrics*, 44, 959–971.

Dockery, D. (2001), "Epidemiologic evidence of cardiovascular effects of particulate air pollution," *Environ Health Perspect*, 109, 483–486.

Dockery, D., Pope, C., Xu, X., Spengler, J., Ware, J., Fay, M., Ferris, B., and Speizer, F. (1993), "An association between air pollution and mortality in six US cities," *New Engl J Med*, 329, 1753–1759.

Finlayson-Pitts, B. and Pitts, J. (1999), *Chemistry of the upper and lower atmosphere: Theory, experiments and applications*, San Diego, CA: Academic Press, Inc.

Gelman, A. (2006), "Prior distributions for variance parameters in hierarchical models," *Bayesian Analysis*, 1, 515–533.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2003), *Bayesian Data Analysis, Second Edition*, New York, NY: Chapman and Hall.

Gelman, A. and Hill, J. (2007), *Data analysis using regression and multilevel/hierarchical models*, New York, NY: Cambridge University Press.

Godleski, J., Verrier, R., Koutrakis, P., and Catalano, P. (2000), "Mechanisms of Morbidity and Mortality from Exposure to Ambient Air Particles," Tech. Rep. 91, Health Effects Institute, Cambridge, MA.

Hastie, T. and Tibshirani, R. (1990), *Generalized additive models*, Chapman and Hall.

Holloman, C., Bortnik, S., Morara, M., Strauss, W., and Calder, C. (2004), "A Bayesian Hierarchical Approach for Relating PM2.5 Exposure to Cardiovascular Mortality in North Carolina," *Environ Health Perspect*, 112, 1282 – 1288.

Hopke, P., Ramadan, Z., Paatero, P., Norris, G., Landis, M., Williams, R., and Lewis, C. (2003), "Receptor modeling of ambient and personal exposure samples: the 1998 Baltimore Particulate Matter Epidemiology-Exposure Study," *Atmos Environ*, 37, 3289–3302.

Huang, Y., Dominici, F., and Bell, M. (2005), "Bayesian hierarchical distributed lag models for summer ozone exposure and cardio-respiratory mortality," *Environmetrics*, 16, 547 – 562.

Jones, G., Haran, M., Caffo, B., and Neath, R. (2006), "Fixed-width output analysis for Markov chain Monte Carlo," *J Am Stat Assoc*, 101, 1537–1547.

La Rovere, M., Bigger, J., Marcus, F., Mortara, A., and Schwartz, P. (1998), "Baroreflex sensitivity and heart-rate variability in prediction of total cardiac mortality after myocardial infarction," *Lancet*, 351, 478–484.

Laden, F., Neas, L., Dockery, D., and Schwartz, J. (2000), "Association of Fine Particulate Matter from Different Sources with Daily Mortality in Six US Cities," *Environ Health Perspect*, 108, 941–947.

Landis, M., Norris, G., Williams, R., and Weinstein, J. (2001), "Personal exposures to $PM_{2.5}$ mass and trace elements in Baltimore, MD, USA," *Atmos Environ*, 35, 6511–6524.

Liao, D., Creason, J., Shy, C., Williams, R., Watts, R., and Zweidinger, R. (1999), "Daily variation of particulate air pollution and poor cardiac autonomic control in the elderly," *Environ Health Perspect*, 107, 521–525.

Luttman-Gibson, H., Coull, B., Dockery, D., Ebelt, S., Schwartz, J., Stone, P., Suh, H., and Gold, D. (2006), "Short-term effects of air pollution on heart rate variability in senior adults in Steubenville, Ohio," *J Occup Environ Med*, 48, 780–788.

Magari, S., Hauser, R., Schwartz, J., Williams, P., Smith, T., and Christiani, D. (2001), "Association of heart rate variability with occupational and environmental exposure to particulate air pollution," *Circulation*, 104, 986–991.

Mar, T., Ito, K., Koenig, J., Larson, T., Eatough, D., Henry, R., Kim, E., Laden, F., Lall, R., Neas, L., Stolzel, M., Paatero, P., Hopke, P., and Thurston, G. (2006), "PM source apportionment and health effects. 3. Investigation of inter-method variations in associations between estimated source contributions of PM2.5 and daily mortality in Phoenix, AZ," *J Expo Anal Environ Epidemiol*, 16, 311–320.

McBride, S., Williams, R., and Creason, J. (2007), "Bayesian Hierarchical Modeling of Personal Exposure to Particulate Matter," *Atmos Environ*, 41, 6143–6155.

Nikolov, M., Coull, B., Catalano, P., and Godleski, J. (2007), "An informative Bayesian structural equation model to assess source-specific health effects of air pollution," *Biostatistics*, 8, 609–624.

Ntzoufras, I. (2009), *Bayesian Modeling using WinBUGS*, Wiley.

O'Neill, M., Zanobetti, A., and Schwartz, J. (2003), "Modifiers of the temperature and mortality association in seven US cities," *Am J Epidemiol*, 157, 1074–1082.

Özkaynak, H., Xue, J., Spengler, J., Wallace, L., Pellizzari, E., and Jenkins, P. (1996), "Personal exposure to airborne particles and metals: Results from the Particle PTEAM study in Riverside, California," *J Expo Anal Environ Epidemiol*, 6, 57–78.

Paatero, P. (1997), "A weighted non-negative least squares algorithm for three-way PARAFAC factor analysis," *Chemometr Intell Lab*, 37, 223–242.

— (1999), "The multilinear engine - a table-driven least squares program for solving multilinear problems, including the n-way parallel factor analysis model," *J Comput Graph Stat*, 8, 854–888.

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

Pope, C., Hansen, M., Long, R., Nielsen, K., Eatough, N., Wilson, W., and Eatough, D. (2004), "Ambient particulate air pollution, heart rate variability, and blood markers of inflammation in a panel of elderly subjects," *Environ Health Perspect*, 112, 339–345.

Richardson, S. and Best, N. (2003), "Bayesian hierarchical models in ecological studies of health environment effects," *Environmetrics*, 14, 129–147.

Sarnat, J., Marmur, A., Klein, M., Kim, E., Russell, A., Sarnat, S., Mulholland, J., Hopke, P., and Tolbert, P. (2008), "Fine particle sources and cardiorespiratory morbidity: An application of chemical mass balance and factor analytical source-apportionment methods," *Environ Health Perspect*, 116, 459–466.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003), *WinBUGS Version 1.4.1*, MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK, http://www.mrc-bsu.cam.ac.uk/bugs.

Strand, M., Vedal, S., Rodes, C., Dutton, S., Gelfand, E., and Rabinovitch, N. (2006), "Estimating effects of ambient $PM_{2.5}$ exposure on health using $PM_{2.5}$ component measurements and regression calibration," *J Expo Sci Environ Epidemiol*, 16, 30–38.

Wallace, L. and Williams, R. (2005), "Use of personal-indoor-outdoor sulfur concentrations to estimate the infiltration factor and outdoor exposure factor for individual homes and persons," *Environ Sci Technol*, 39, 1707–1714.

Wallace, W., Williams, R., Suggs, J., and Jones, P. (2006), "Estimating Contributions of Outdoor Fine Particles to Indoor Concentrations and Personal Exposures: Effects of Household Characteristics and Personal Activities," Tech. Rep. ORD Report (APM 214), EPA/600/R-023, US Environmental Protection Agency, Washington, D.C.

Williams, R., Suggs, J., Creason, J., Rodes, C., Lawless, P., Kwok, R., Zweidinger, R., and Sheldon, L. (2000a), "The 1998 Baltimore particulate matter epidemiology-exposure study: Part 2. Personal exposure assessment associated with an elderly study population," *J Expo Anal Environ Epidemiol*, 10, 533–543.

Williams, R., Suggs, J., Zweidinger, R., Evans, G., Creason, J., Kwok, R., Rodes, C., Lawless, P., and Sheldon, L. (2000b), "The 1998 Baltimore particulate matter epidemiology-exposure study: Part 1. Comparison of ambient, residential outdoor, indoor and apartment particulate matter monitoring," *J Expo Anal Environ Epidemiol*, 10, 518–532.

Wilson, W. and Brauer, M. (2006), "Estimation of ambient and non-ambient components of particulate matter exposure from a personal monitoring panel study," *J Expo Sci Environ Epidemiol*, 16, 264–74.

A WinBUGS code

```
model;
{
## model for first 119 obs with both HRV and PM personal mmts
## separate loop for first observation since this is a lag 1 model and
## an observation at time 0 is not available; see defn of errorfirst.
for (obsnum in 1:1){
## measurement error model for observed HRV observations
     Z[obsnum]~dnorm(H[obsnum],tau.H);
     ## regression equation for unknown HRV
     H[obsnum] <- beta[Person[obsnum]] + Hcoef[1]+ Hcoef[2]*age[obsnum] +
       Hcoef[3]*cvcompro[obsnum] + Hcoef[4]*sex[obsnum] + Hcoef[5]*
      M.P.tot.lag1[obsnum] + Tcoef[1]*a.temp1[day.num[obsnum]] +
      Tcoef[2]*a.temp2[day.num[obsnum]] + Tcoef[3]*
      a.temp3[day.num[obsnum]] + Tcoef[4]*a.temp4[day.num[obsnum]]+
      error[obsnum];
     ## first observation of error is set to be random and issame for
     ## all individuals
     error[obsnum]<-errorfirst;
     ## measurement error model for observed PM2.5
     Y.P.lag1[obsnum]~dnorm(M.P.tot.lag1[obsnum],tau.M.P.tot);
     ## unknown total personal PM2.5 exposure as sum of ambient and
     ## non-ambient components
      M.P.tot.lag1[obsnum]<-M.P.NA.lag1 + M.P.A.lag1[obsnum];
      ## unknown personal PM2.5 of ambient origin as product of infiltration
      ## and unknown ambient PM2.5
     M.P.A.lag1[obsnum]<-gamma1*PMC[day.num[obsnum]];
    }

    ## loop for remaining 118 observations for which both HRV and PM2.5 are
    ## available (same code as above except for defn of "error[obsnum]")
    for (obsnum in 2:119){
    Z[obsnum]~dnorm(H[obsnum],tau.H);
      H[obsnum] <- beta[Person[obsnum]] + Hcoef[1]+ Hcoef[2]*age[obsnum] +
      Hcoef[3]*cvcompro[obsnum] + Hcoef[4]*sex[obsnum] + Hcoef[5]*
      M.P.tot.lag1[obsnum]+ Tcoef[1]*a.temp1[day.num[obsnum]] + Tcoef[2]*
      a.temp2[day.num[obsnum]] + Tcoef[3]*a.temp3[day.num[obsnum]] +
       Tcoef[4]*a.temp4[day.num[obsnum]] + error[obsnum];
      ##Definition of error term in terms of previous observations
      ##Indicator variable "Frst" indicates whether this observation
```

```
      ##is the first in the time series for a given individual.
        error[obsnum]<-Frst[obsnum]*errorfirst+(1-Frst[obsnum])*
        (exp(-alpha*delta[obsnum])*error[obsnum-1]);
    Y.P.lag1[obsnum]~dnorm(M.P.tot.lag1[obsnum],tau.M.P.tot);
     M.P.tot.lag1[obsnum]<-M.P.NA.lag1 + M.P.A.lag1[obsnum];
     M.P.A.lag1[obsnum]<-gamma1*PMC[day.num[obsnum]];
    }
    ### model for remaining measurements with HRV only, no personal PM mmts
    for (obsnum in 120:658){
    Z[obsnum]~dnorm(H[obsnum],tau.H);
     H[obsnum] <- beta[Person[obsnum]] + Hcoef[1]+ Hcoef[2]*age[obsnum] +
      Hcoef[3]*cvcompro[obsnum] + Hcoef[4]*sex[obsnum] +
      Hcoef[5]*M.P.tot.lag1[obsnum] + Tcoef[1]*a.temp1[day.num[obsnum]]+
      Tcoef[2]*a.temp2[day.num[obsnum]] + Tcoef[3] *
      a.temp3[day.num[obsnum]] + Tcoef[4]*a.temp4[day.num[obsnum]] +
      error[obsnum];
        error[obsnum]<-Frst[obsnum]*errorfirst+(1-Frst[obsnum])*
        (exp(-alpha*delta[obsnum])*error[obsnum-1]);
    M.P.tot.lag1[obsnum]<-M.P.NA.lag1 + M.P.A.lag1[obsnum];
       M.P.A.lag1[obsnum]<-gamma1*PMC[day.num[obsnum]];
    }

    ## random subject specific effect
for (i in 1:56){
beta[i]~dnorm(0.0,tau.b);
}

## non-ambient PM2.5 personal
M.P.NA.lag1~dnorm(mu.MPNA,tau.MPNA);

## personal sulfur measurements are X.Ind
## gamma1 is the infiltration factor
for (i in 1:N.sulfur){
X.Ind[i]~dnorm(Sulfur.Ind[i],tau.meas.sulfurInd);
Sulfur.Ind[i]<-gamma1*SulfurC[day.num.sulfur[i]] + errgamma1;
}

## submodel for outdoor sulfur; mean is linear function of meteorology
## outdoor sulfur measurements are X.C
## separate models for day 1 and days 2-29 since for day 1
## a mean of sulfur at day 0 is needed; this is set to be random
for (t in 1:1){
```

```
X.C[t]~dnorm(SulfurC[t],tau.meas.sulfurC);
SulfurC[t]~dnorm(mu.sulfurC[t],tau.mean.sulfurC);
mu.sulfurC[t] <- Scoef[1] + Scoef[2]*humid[t] + Scoef[3]*humid.lag1[t] +
 Scoef[4]*vws[t] + Scoef[5]*vws.lag1[t] + Scoef[6]*temp[t] + Scoef[7]*
 temp.lag1[t] + rho.sulfurc * mu.sulfurc.t0 + eps.sulfur;
}
for (t in 2:29){
X.C[t]~dnorm(SulfurC[t],tau.meas.sulfurC);
SulfurC[t]~dnorm(mu.sulfurC[t],tau.mean.sulfurC);
mu.sulfurC[t] <- Scoef[1] + Scoef[2]*humid[t] + Scoef[3]*humid.lag1[t] +
 Scoef[4]*vws[t] + Scoef[5]*vws.lag1[t] + Scoef[6]*temp[t] + Scoef[7]*
 temp.lag1[t] + rho.sulfurc * mu.sulfurC[t-1] + eps.sulfur;
}

## submodel for outdoor PM2.5 observations
## outdoor PM2.5 measurements are Y.C
for (t in 1:1){ ## for PM2.5
Y.C[t] ~  dnorm(PMC[t],tau.meas.pmc);
PMC[t] ~ dnorm(mu.pmc[t],tau.mean.pmc);
mu.pmc[t] <-  PMCcoef[1] + PMCcoef[2]*humid[t] + PMCcoef[3]*humid.lag1[t
] + PMCcoef[4]*vws[t] + PMCcoef[5]*vws.lag1[t]+PMCcoef[6]*dow[t]+ rho.
pmc * mu.pmc.t0 + eps.pmc;
}
for (t in 2:29){
Y.C[t] ~  dnorm(PMC[t],tau.meas.pmc);
PMC[t] ~ dnorm(mu.pmc[t],tau.mean.pmc);
mu.pmc[t] <-  PMCcoef[1] + PMCcoef[2]*humid[t] + PMCcoef[3]*humid.lag1[t
] + PMCcoef[4]*vws[t] + PMCcoef[5]*vws.lag1[t]+PMCcoef[6]*dow[t]+ rho.
pmc * mu.pmc[t-1] + eps.pmc;
}
###############
### PRIORS ###
###############

## coefficients in regression equation for HRV
for (i in 1:5) {
Hcoef[i]~dnorm(0.0,1.0E-2);
}
## error term in regression equation for HRV
alpha~dunif(0.0,20.0);
tau.H <- pow(sigma.tau.H,-2);
sigma.tau.H ~ dunif(0.0,100.0);
```

```
1
2  ## scale parameter for random subject effect
3  tau.b <- pow(sigma.tau.b,-2);
4  sigma.tau.b ~ dunif(0.0,100.0);
5  errorfirst~dnorm(0,1.0E-2);
6
7  ## coefficients in cubic splines for temperature
8      for (i in 1:4){
9    Tcoef[i]~dnorm(0.0,1.0E-2);
10   }
11
12   ## scale parameter for measured personal sulfur
13 tau.meas.sulfurInd<-pow(sigma.tau.meas.sulfurInd,-2);
14 sigma.tau.meas.sulfurInd ~ dunif(0.0, 100.0);
15
16 ## scale parameter for total personal PM2.5 exposure
17 tau.M.P.tot<-pow(sigma.tau.M.P.tot,-2);
18 sigma.tau.M.P.tot ~ dunif(0.0, 100.0);
19
20 ## parameters in non-ambient personal PM2.5 exposure
21 mu.MPNA~dnorm(5.0,1.0E-2);
22 tau.MPNA<-pow(sigma.tau.MPNA,-2);
23 sigma.tau.MPNA ~ dunif(0.0, 100.0);
24
25 ## infiltration factor
26 gamma1~dbeta(1.0,1.0);
27 errgamma1~dnorm(0.0,1.0E-2);
28
29 ## priors for terms in model for outdoor sulfur
30 eps.sulfur~dnorm(0.0,1.0E-2);
31 rho.sulfurc~dunif(-1,1);
32 mu.sulfurc.t0~dnorm(2.3,1.0E-2);
33 for (i in 1:7){
34 Scoef[i]~dnorm(0.0,1.0E-2);
35 }
36 tau.meas.sulfurC<-pow(sigma.tau.meas.sulfurC,-2);
37 sigma.tau.meas.sulfurC ~ dunif(0.0,100.0);
38 tau.mean.sulfurC<-pow(sigma.tau.mean.sulfurC,-2);
39 sigma.tau.mean.sulfurC ~ dunif(0.0, 100.0);
40
41 ## priors for terms in model for outdoor PM2.5
42 eps.pmc ~ dnorm(0.0,1.0E-2);
```

```
rho.pmc ~ dunif(-1,1);
mu.pmc.t0 ~ dnorm(18,1.0E-2);
for (i in 1:6){
PMCcoef[i]~ dnorm(0.0, 1.0E-2);
}
tau.meas.pmc<-pow(sigma.tau.meas.pmc,-2);
sigma.tau.meas.pmc~dunif(0.0,100.0);
tau.mean.pmc<-pow(sigma.tau.mean.pmc,-2);
sigma.tau.mean.pmc~dunif(0.0,100.0);

}

```

List of Figures

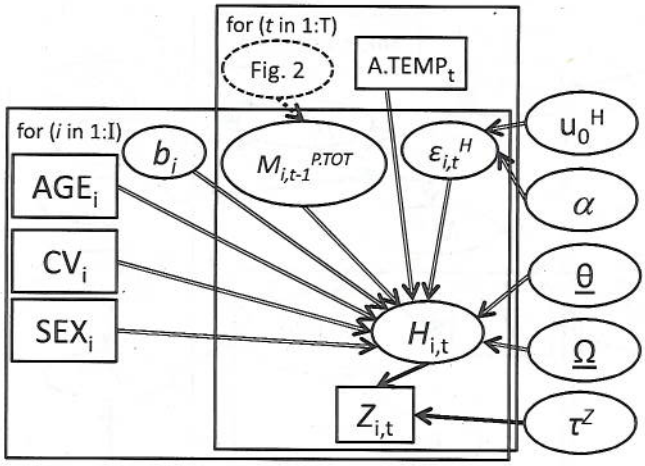

Figure 1: Model for the unknown mean health effect ($H_{i,t}$) for subject i on day t as a function of unknown total personal exposure to $PM_{2.5}$ on day $t-1$ ($M_{i,t-1}^{P.TOT}$), fixed subject-specific covariates (age, AGE_i; cardiovascular status, CV_i; sex, SEX_i), apparent temperature ($A.TEMP_t$,), random subject effects (b_i) and correlated error terms ($\varepsilon_{i,t}^H$).

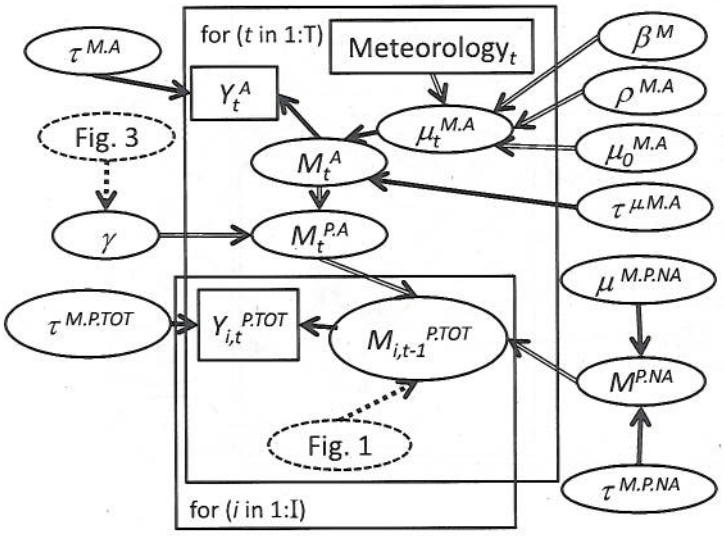

Figure 2: Model for unknown total personal exposure to $PM_{2.5}$ for subject i on day $t-1$ as a function of its personal ambient ($M_t^{P.A}$) and personal non-ambient ($M_t^{P.NA}$) components, where the mean of the ambient component, $\mu_t^{M.A}$, is a linear combination of meteorology and error terms.

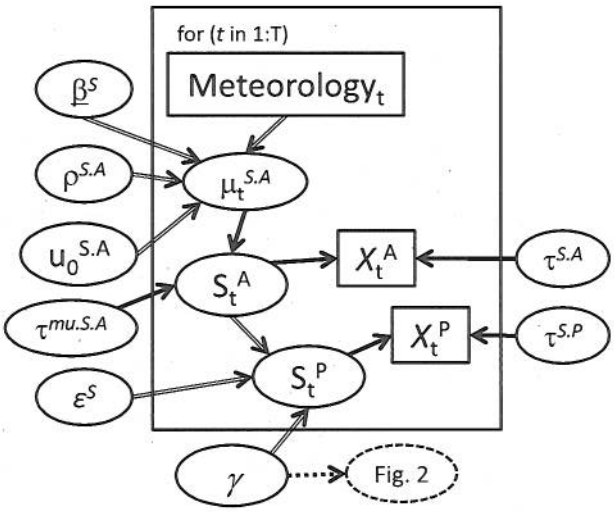

Figure 3: Model for unknown personal sulfur exposure (S_t^P), as a function of unknown ambient sulfur concentrations (S_t^A) and an indoor infiltration factor, γ.

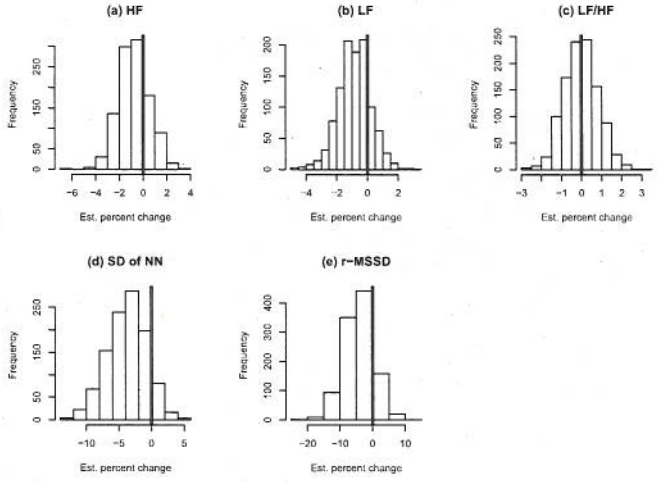

Figure 4: Histograms of posterior percentage change in HRV associated with a 6.5 μg/m^3(1 SD) increase in personal total $PM_{2.5}$ exposure at lag one day.

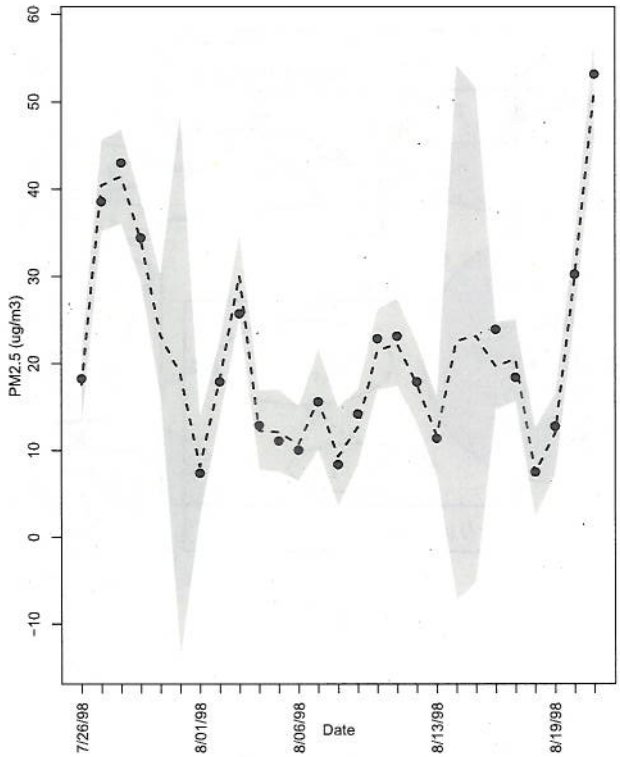

Figure 5: Time series of measured $PM_{2.5}$ at ambient site (dots) coplotted with 95% posterior intervals for $PM_{2.5}$ at ambient site (grey bands) and modeled mean $PM_{2.5}$ at ambient site (dotted line).

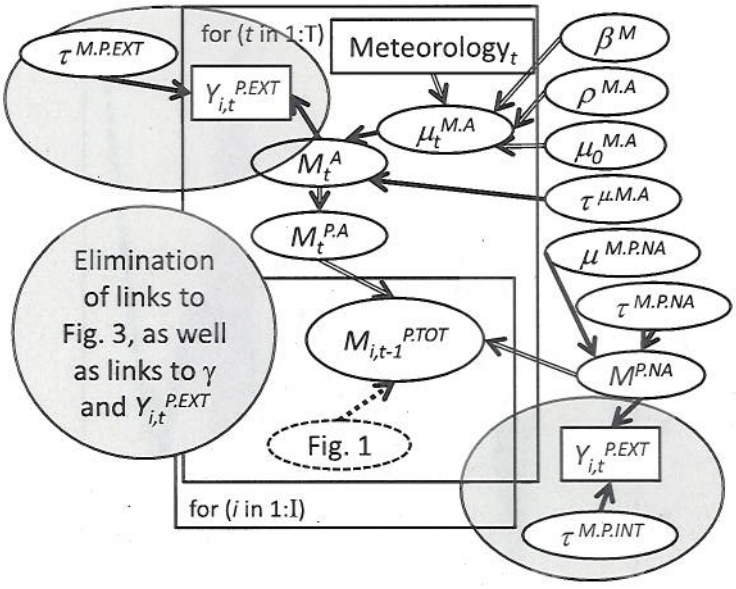

Figure 6: Modifications of Figure 2, shown circled in grey, to accommodate the personal PEM source apportionment results of Hopke et al. (2003), which define internal ($Y_{i,t}^{P.INT}$) and external ($Y_{i,t}^{P.EXT}$) factors.

List of Tables

Table 1: Summary statistics for heart rate variability (HRV) measures for 56 subjects

Heart rate variability measure	Average	Median	Std. Dev.	5%	95%
High freq. HRV (HF)	3.1	2.9	0.8	2.0	4.5
Low freq. HRV (LF)	3.1	2.9	0.7	2.1	4.5
LF/HF	1.0	1.0	0.1	0.8	1.3
SDNN	43.8	29.5	38.9	12.5	112.1
r-MSSD	52.4	28.7	59.7	6.9	162.5

Table 2: Source contribution estimates (%) identified by Hopke et al. (2003) for PEM personal indoor samples

	Source					
	External			Internal		
Statistic	Sulfate	Soil	Unknown	Gypsum	Activity	Personal Care
Average	46.3	13.6	2.8	36.0	0.7	0.4
Std. Dev.	23.0	9.3	1.9	23.0	0.8	0.9
Median	45.7	11.6	2.1	35.0	0.4	0.2
5%	8.2	3.3	0.9	0.0	0.0	0.0
95%	82.8	30.1	6.3	75	1.8	1.6

Table 3: Source contribution estimates ($\mu g/m^3$) identified by Hopke et al. (2003) for PEM personal indoor samples

	Source					
	External			Internal		
Statistic	Sulfate	Soil	Unknown	Gypsum	Activity	Personal Care
Average	6.16	0.30	1.56	0.07	4.54	0.05
Std. Dev.	4.60	0.19	0.99	0.10	3.39	0.08
Median	4.66	0.28	1.34	0.05	4.44	0.02
5%	0.71	0.10	0.34	0.00	0.00	0.00
95%	15.12	0.53	3.41	0.23	10.43	0.20

Table 4: Source contribution estimates (%) identified by Hopke et al. (2003) for VAPS samples at the ambient site

Statistic	Nitrate	Sulfate	Organic Carbon	Motor Vehicles
Average	20.7	53.7	8.8	16.8
Std. Dev.	9.2	18.4	3.7	14.0
Median	19.0	58.0	8.0	13.0
5%	8.6	20.3	3.3	0.0
95%	36.0	78.0	14.0	44.0

Table 5: Source contribution estimates ($\mu g/m^3$) identified by Hopke et al. (2003) for VAPS samples at the ambient site

Statistic	Nitrate	Sulfate	Organic Carbon	Motor Vehicles
Average	3.98	12.05	1.43	2.51
Std. Dev.	2.72	9.57	0.14	1.73
Median	3.13	10.02	1.39	2.92
5%	1.58	2.22	1.24	0.00
95%	10.43	32.87	1.69	5.14

Table 6: Posterior estimates of percentage change in HRV associated with a 6.5 $\mu g/m^3$(1 SD) increase in personal total $PM_{2.5}$ exposure at lag one day.

HRV Response	Posterior Mean	95% Posterior Interval	Post. Prob. that $\theta_4 < 0$
HF	-0.77	(-3.36, 1.80)	0.72
LF	-0.90	(-3.00, 1.27)	0.78
LF/HF	-0.00	(-1.65, 1.66)	0.49
SD of NN	-3.79	(-9.94, 2.67)	0.89
r-MSSD	-3.93	(-11.98, 6.16)	0.80

Table 7: Posterior estimates of percentage change in the ratio of low to high frequency HRV associated with a 6.5 $\mu g/m^3$(1 SD) increase in personal total $PM_{2.5}$ exposure at four time lag structures.

Personal total $PM_{2.5}$ Lag Structure	Posterior Mean	95% Posterior Interval	Post. Prob. that $\theta_4 < 0$
0 days	-0.69	(-3.09, 1.89)	0.71
1 day	0.00	(-1.65, 1.66)	0.49
Mov. avg., days 0-1	-0.90	(-3.25, 1.42)	0.77
Mov. avg., days 0-2	-1.48	(-3.60, 0.89)	0.90

Table 8: Posterior estimates of percentage change in HRV associated with a 4 $\mu g/m^3$(1 SD) increase in personal $PM_{2.5}$ exposure due to exterior sources. Also shown are posterior probabilities that coefficients of external personal and internal personal $PM_{2.5}$ concentrations are negative.

HRV measure	Time lag	Post. Mean	95% Post. Interval	Post. Prob. Ext. Coef. < 0	Post. Prob. Int. Coef. < 0
HF	0 days	-0.14	(-3.47, 3.45)	0.52	0.53
	1 day	-1.51	(-4.06, 1.08)	0.88	0.53
	Mov. avg., days 0-1	-1.74	(-4.51, 1.05)	0.89	0.53
	Mov. avg., days 0-2	0.00	(-2.47, 2.48)	0.51	0.54
LF	0 days	-0.89	(-4.51, 2.61)	0.69	0.52
	1 day	-2.19	(-4.82, 0.41)	0.95	0.53
	Mov. avg., days 0-1	-2.26	(-4.81, 0.21)	0.97	0.54
	Mov. avg., days 0-2	-1.45	(-3.63, 0.57)	0.92	0.56
LF/HF	0 days	-0.78	(-3.48, 2.08)	0.70	0.53
	1 day	-0.48	(-2.8, 2.03)	0.65	0.50
	Mov. avg., days 0-1	-0.40	(-2.85, 1.93)	0.63	0.51
	Mov. avg., days 0-2	-1.48	(-3.5, 0.52)	0.92	0.51
SD of NN	0 days	-4.58	(-14.16, 6.56)	0.81	0.64
	1 day	-5.70	(-13.29, 2.27)	0.91	0.65
	Mov. avg., days 0-1	-5.27	(-12.98, 2.54)	0.90	0.65
	Mov. avg., days 0-2	-3.53	(-10.77, 3.98)	0.82	0.65
r-MSSD	0 days	-6.86	(-21.55, 10.14)	0.77	0.70
	1 day	-9.06	(-19.58, 4.05)	0.93	0.67
	Mov. avg., days 0-1	-9.18	(-20.29, 2.73)	0.94	0.67
	Mov. avg., days 0-2	-5.43	(-16.43, 7.62)	0.83	0.68

Table 9: Posterior estimates of percentage change in HRV associated with a 1 $\mu g/m^3$increase in personal exposure to $PM_{2.5}$ components based on personal PEM apportionment. Also shown are posterior probabilities that coefficients of personal $PM_{2.5}$ component concentrations are negative. Of 60 possible models (5 HRV measures, 4 time lags, 3 $PM_{2.5}$ components), the 11 models shown have posterior probabilities greater than 70% that coefficients of personal $PM_{2.5}$ component concentrations are negative.

$PM_{2.5}$ Component	HRV Measure	Time Lag	Posterior Mean	95% Post. Interval	Post. Prob. Coef. < 0
Soil	HF	0 days	-50.83	(-86.59, 90.52)	0.91
	LF	0 days	-45.16	(-78.74, 23.57)	0.94
	LF/HF	0 days	-15.77	(-59.12, 93.68)	0.73
		Mov. avg., days 0-1	-26.46	(-71.00, 122.92)	0.76
	r-MSSD	0 days	-73.76	(-99.49, 1796.60)	0.80
	SD of NN	0 days	-63.11	(-96.59, 238.97)	0.84
Sulfate	HF	Mov. avg., days 0-1	-2.55	(-7.95, 2.81)	0.86
	r-MSSD	Mov. avg., days 0-1	-8.38	(-28.32, 12.64)	0.80
Unknown	LF/HF	0 days	-6.46	(-19.04, 9.57)	0.86
		1 day	-8.37	(-25.15, 13.40)	0.85
		Mov. avg., days 0-2	-15.59	(-34.42, 9.63)	0.93

Table 10: Posterior estimates of percentage change in HRV associated with a 3 $\mu g/m^3$increase in personal exposure to $PM_{2.5}$ components based on ambient site apportionment (VAPS). Also shown are posterior probabilities that coefficients of personal $PM_{2.5}$ component concentrations are negative. Of 80 possible models (5 HRV measures, 4 time lags, 4 $PM_{2.5}$ components), the 17 models shown have posterior probabilities greater than 75% that coefficients of personal $PM_{2.5}$ component concentrations are negative.

$PM_{2.5}$ Component	HRV Measure	Time Lag	Posterior Mean	95% Post. Interval	Post. Prob. Coef. < 0
Nitrate	HF	1 day	-10.23	(-21.20, -0.84)	0.98
		Mov. avg., days 0-1	-11.95	(-26.78, -0.98)	0.98
		Mov. avg., days 0-2	-12.11	(-27.02, 9.73)	0.96
	LF	1 day	-7.88	(-22.22, 4.64)	0.92
		Mov. avg., days 0-1	-12.14	(-25.94, 2.07)	0.97
		Mov. avg., days 0-2	-11.94	(-26.87, 9.86)	0.97
	r-MSSD	1 day	-34.92	(-65.48, -3.09)	0.98
		Mov. avg., days 0-1	-32.16	(-66.74, 8.48)	0.96
		Mov. avg., days 0-2	-37.41	(-70.47, 2.21)	0.97
	SD OF NN	1 day	-19.67	(-44.09, 8.56)	0.94
		Mov. avg., days 0-1	-19.43	(-46.93, 30.27)	0.89
		Mov. avg., days 0-2	-21.15	(-47.46, 24.31)	0.91
Sulfate	HF	1 day	-0.76	(-2.21, 0.98)	0.82
	LF	1 day	-0.74	(-2.12, 0.61)	0.85
	LF/HF	Mov. avg., days 0-2	-0.87	(-2.59, 0.77)	0.85
	r-MSSD	1 day	-2.72	(-8.70, 3.41)	0.83
	SD of NN	1 day	-3.49	(-7.34, 0.33)	0.97