*Journal of Exposure Science and Environmental Epidemiology* (2007) 17, S2–S10
© 2007 Nature Publishing Group All rights reserved 1559-0631/07/$30.00
www.nature.com/jes

# Effects of exposure measurement error on particle matter epidemiology: a simulation using data from a panel study in Baltimore, MD

JOEL SCHWARTZ[a], JEREMY A. SARNAT[b], BRENT A. COULL[c] AND WILLIAM E. WILSON[d]

[a]*Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA*
[b]*Department of Environmental Health, Rollins School of Public Health of Emory University, Atlanta, Georgia, USA*
[c]*Department of Biostatistics, Harvard School of Public Health, Massachusetts, USA*
[d]*National Center for Environmental Assessment, US Enviornmental Protection Agency Research Triangle Park, North Carolina, USA*

Ascertaining the true risk associated with exposure to particulate matter (PM) is difficult, given the fact that pollutant components are frequently correlated with each other and with other gaseous pollutants; relationships between ambient concentrations and personal exposures are often not well understood; and PM, unlike its gaseous co-pollutants, does not represent a single chemical. In order to examine differences between observed *versus* true health risk estimates from epidemiologic studies, we conducted a simulation using data from a recent multi-pollutant exposure assessment study in Baltimore, MD. The objectives of the simulation were twofold: (a) to estimate the distribution of personal air pollutant exposures one might expect to observe within a population, given the corresponding ambient concentrations found in that location and; (b) using an assumed true health risk with exposure to one pollutant, to estimate the distribution of health risk estimates likely to be observed in an epidemiologic study using ambient pollutant concentrations as a surrogate of exposure as compared with actual personal pollutant exposures. Results from the simulations showed that $PM_{2.5}$ was the only pollutant where a true association with its total personal exposures resulted in a significant observed association with its ambient concentrations. The simulated results also showed that true health risks associated with personal exposure to $O_3$ and $NO_2$ would result in no significant observed associations with any of their respective ambient concentrations. Conversely, a true association with $PM_{2.5}$ would result in a significant, observed association with $NO_2$ ($\beta = 0.0115$, 95% confidence interval (CI): 0.0056, 0.0185) and a true association with exposure to $SO_4^{2-}$ would result in an observed significant association with $O_3$ ($\beta = 0.0035$, 95% CI: 0.0021, 0.0051) given the covariance of the ambient pollutant concentrations. The results provide an indication that, in Baltimore during this study period, ambient gaseous concentrations may not have been adequate surrogates for corresponding personal gaseous exposures to allow the question to be investigated using central site monitors. Alternatively, the findings may suggest that in some locations, observed associations with the gaseous pollutants should be interpreted with caution, as they may be reflecting associations with PM or one of its chemical components.
*Journal of Exposure Science and Environmental Epidemiology* (2007) **17,** S2–S10; doi:10.1038/sj.jes.7500619

**Keywords:** *particles, air pollution, mortality, time series, measurement error.*

## Introduction

Daily changes in ambient particulate matter (PM) concentrations have been associated with daily changes in deaths (Schwartz and Dockery, 1992; Katsouyanni et al., 1997), cardiovascular and pulmonary hospital admissions (Schwartz et al., 1999; Le Tertre et al., 2002), lung disease (Schwartz, 1994; Atkinson et al., 2001) and intermediary endpoints that serve as risk factors for such adverse health outcomes (Salvi et al., 1999; Gold et al., 2000; Schwartz, 2001). Some have questioned the validity of these time-series results based on the contention that ambient PM concentrations are poor surrogates for corresponding personal PM exposures (Gamble and Nicolich, 2000), whereby the degree to which that measurement error may impact health risk estimates depends upon the difference between the ambient PM concentration and the corresponding personal exposure. In the National Research Council's Report "Research Priorities for Airborne Particulate Matter: I. Immediate Priorities and Long-Range Research Portfolio" (National Research Council, 1998), the Committee on Research Priorities recommended that EPA specifically address measurement error issues in PM health effects studies.

Zeger et al. (2000) developed a statistical framework for studying the effects of this type of measurement error on health effect estimates. A key finding in this paper is that previous discussions of exposure measurement error often ignored the difference between classical and Berkson measurement error. The authors discuss three types of exposure measurement error in longitudinal studies of air pollution and health, such as those cited above. The first error is derived from the difference between the daily personal exposures of each individual and the daily

1. Address all correspondence to: Dr. Joel Schwartz, Department of Environmental Health, Harvard School of Public Health, 401 Park Dr., suite 415 W, PO Box 15677, Boston, MA 2215, USA.
Tel.: +1 617 384 8752. Fax: +1 617 384 8745.
E-mail: jschwrtz@hsph.harvard.edu
Received 7 September 2007; accepted 12 September 2007

community-average personal exposure. The second error stems from the difference between the daily community-average personal exposure and the true ambient concentration. Finally, the third component of error is from the difference between the measured and the true ambient concentration (i.e., measurement error). Zeger and co-workers argue that the first and third types of errors are of the Berkson type. Since Berkson measurement error is such that the true exposure is distributed around the measured value, this form of error will not lead to biased health effect estimates for a given pollutant, but does make the association less significant since relying on the mismeasured surrogate is not as informative as having true exposure. The second type of error is termed classical and will bias the relative risk. Upward bias can occur in rare settings where the ratio of the amount of measurement error in PM to the measurement error in the other pollutants is high and the correlations among the pollutants are extreme. Hence, the authors concluded that observed PM health risk estimates were likely not biased away from 0 by measurement error.

The second type of error will be influenced by pollution from indoor sources and by the difference between the ambient concentration and the ambient exposure (the ambient component of personal exposure). Since indoor-generated sources are generally uncorrelated with the ambient concentrations, the indoor-generated component of personal exposure will not affect the relative risk due to ambient pollution. The difference between ambient concentration and ambient exposure leads to a scaling factor. Consider the case where personal exposure to $PM_{2.5}$ is perfectly correlated with (no measurement error), but always half of a corresponding ambient $PM_{2.5}$ concentration (i.e., a slope of 0.5 between the personal exposure and ambient concentration). A 1-$\mu$g/m$^3$ increase in ambient $PM_{2.5}$, for example, is associated with a 0.5-$\mu$g/m$^3$ increase in personal exposure to $PM_{2.5}$. Therefore, a $\beta^A$ coefficient of 0.05 for exposure to ambient $PM_{2.5}$ as a predictor of health would result in a $\beta^C$ coefficient of 0.025 for ambient concentrations as a predictor of health. This result refers to the proportionality of coefficients discussed by Zeger et al. (2000). That is, this difference in coefficients is not bias due to measurement error, but rather the realization that a given change in concentration produces only half that change in exposure. Hence, $\beta^C$ would be the true coefficient for changes in ambient concentrations. More generally, in the absence of measurement error, we would expect to observe a coefficient for ambient concentrations related to health that was the product of the coefficient relating exposure to health with the coefficient relating concentration to exposure. Zeger et al. (2000) denoted this scaling factor (the ratio of personal exposure to an ambient pollutant to the ambient concentration) as "$\alpha$". In addition to the effect of this scaling factor, the $\beta$ obtained in an epidemiologic analysis using ambient concentrations will be reduced if the exposure to ambient $PM_{2.5}$ is not perfectly correlated with the ambient $PM_{2.5}$ concentration.

Finally, Zeger et al. (2000) noted that it is currently difficult to quantify the effects of exposure measurement error in PM epidemiology, due to the lack of data on both (1) the magnitude of these errors and (2) how the components of error covary across pollutants. As a result, the authors called for studies that collect daily measurements of personal exposure and ambient levels for multiple pollutants for each person (Zeger et al., 2000).

The development of novel multiple pollutant personal samplers (MPPS) has provided air pollution researchers with a tool for examining these issues using a human panel study design (Demokritou et al., 2001). The results from the current analysis represent the first effort to simulate particulate and gaseous pollutant correlations and co-variances using empirical data and assess their impact on epidemiological observations. Specifically, we used data from a longitudinal exposure assessment panel study conducted in Baltimore, MD to obtain realistic estimates of the effects of measurement error in PM epidemiology. Study subjects wore the MPPS that simultaneously measured personal exposures for multiple particulate and gaseous pollutants, including ozone ($O_3$), and nitrogen dioxide ($NO_2$). We also obtained corresponding ambient concentrations of these pollutants for use in the current analysis. We then performed simulation analyses to generate the distribution of the health effect estimates for different pollutants when the measurement errors and cross-pollutant correlations match those observed in the Baltimore data, for a variety of assumed true health effect associations with exposures to each of the pollutants. The resulting framework builds upon that of Zeger et al. (2000) by allowing the covariances among personal and ambient exposures to vary across subjects, and by explicitly incorporating observed measurement error in the gaseous pollutants.

## Data and methods

### *Sources of Data on Associations Among Pollutants*

The initial stage of the simulation consisted of collecting distributions of associations between personal pollutant exposures and corresponding ambient pollutant concentrations measured at fixed-site monitors, which are traditionally used in air pollution epidemiology as surrogates of personal exposure. We used data from a panel study conducted by Sarnat et al. (2001), which examined the associations between personal exposures to multiple pollutants as well as corresponding ambient concentrations in Baltimore, MD.

In the Baltimore panel study (Sarnat et al., 2001), the investigators measured personal $PM_{2.5}$, $PM_{10}$, $O_3$, and $NO_2$ exposures and corresponding ambient concentrations for 56 subjects living in Baltimore, MD (20 healthy senior adults, 15 adults with COPD, 21 children). All the subjects were non-smokers and lived in residences with non-smokers.

Subject selection was not random and was not intended to be representative of the population-at-large, but did broadly sample potentially sensitive populations of interest in air pollution epidemiology, such as children and senior adults.

Investigators measured the subjects for 12 consecutive days (henceforth, sampling session) in each of one or two seasons (summer and winter), with the exception of children in the summer, who were measured for 8 consecutive days. Of the 56 subjects, 10 senior adults and four children participated during both the summer and winter sampling seasons, which resulted in 70 individual sampling sessions. A total of 800 person-days of personal exposure data were collected for the following pollutants: particles with aerodynamic diameters less than 2.5 $\mu$m ($PM_{2.5}$), particles with aerodynamic diameters less than 10 $\mu$m ($PM_{10}$), sulfate ($SO_4^{2-}$), ozone ($O_3$), and nitrogen dioxide ($NO_2$). $PM_{2.5}$ and $PM_{10}$ were collected using Personal Environmental Monitors with 37-mm Teflon filters (37 mm Teflo™; Gelman Sciences). Particulate $SO_4^{2-}$ levels were obtained by analyzing the aqueous extract of the $PM_{2.5}$ filters using ion chromatography, $O_3$, and $NO_2$ concentrations were measured using passive Ogawa samplers (Koutrakis et al., 1994; Ogawa & Company, 1998).

Twenty-four-hour integrated ambient $PM_{2.5}$ concentrations were measured using Harvard Impactors operated at a centrally located site. Ambient $O_3$ and $NO_2$ data were obtained from local stationary ambient monitoring sites operated by the Maryland Department of the Environment (MDE), and USEPA. $O_3$ and $NO_2$ were measured using UV photometric analyzers and chemiluminiscence monitors, respectively. In cases where pollutant concentrations were measured at multiple sites, concentrations were averaged across the sites.

Standard QA/QC procedures were followed for this study. Collected data were assessed for bias, precision, and completeness. Detailed descriptions of the sampling system, precision, accuracy, and LOD information and study design have been provided elsewhere (Sarnat et al., 2000; Demokritou et al., 2001; Koutrakis, 2005, no. 266).

### *Analytical Approach*

The objectives of the simulation were twofold: (a) to estimate the distribution of personal air pollutant exposures one might expect to observe within a population, given the corresponding ambient concentrations found in that location, and (b) using an assumed true health risk with exposure to one pollutant to estimate the distribution of health risk estimates likely to be observed in an epidemiologic study using ambient pollutant concentrations as a surrogate of exposure as compared to actual personal pollutant exposures. In particular, we were interested in the possibility that a true association with exposure to pollutant "A" may, because of correlations among the ambient pollutant concentrations and exposures, result in an observed association with an ambient concentration of pollutant "B". The analytical approach used for the simulation can be described in the following six steps and is presented graphically in Figure 1.

(1) Using data from the Baltimore panel study (Sarnat et al., 2001), we calculated subject-specific covariance matrices, by

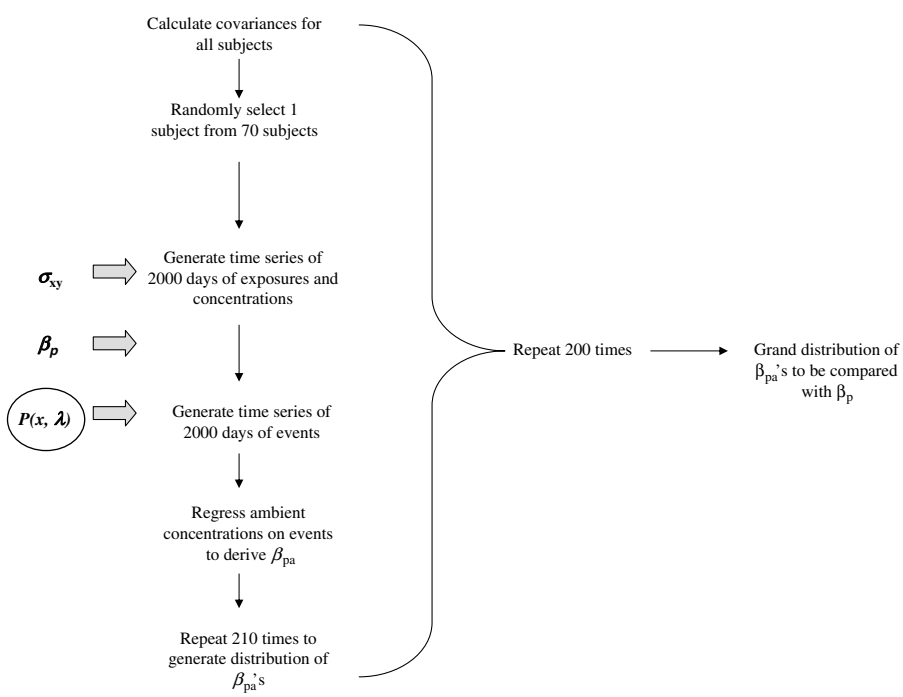


**Figure 1.** Schematic for the sampling scheme for the simulation analyses.

season, between the personal exposures and corresponding ambient concentrations for all of the measured pollutants. Importantly, the use of the covariance matrix allows us to capture not only the correlation between ambient concentrations of each pollutant and personal exposures to that same pollutant, but also the cross-correlations among the ambient pollutants and exposures (i.e., correlations between ambient concentrations of a given pollutant and personal exposures to a different pollutant).

(2) Next, we generated a time series consisting of 2000 days of ambient air pollution concentrations and personal exposures to $PM_{2.5}$, $O_3$, and $NO_2$ using a randomly selected covariance matrix from a given subject.

(3) We then assumed that a true health association, denoted as the coefficient $\beta_p = 0.05$, existed between personal exposures to one of the pollutants (i.e., $PM_{2.5}$, $O_3$, or $NO_2$) and an unspecified health event (either morbidity or mortality). We also assumed that exposures to the other two pollutants were not associated with any health risks. Assuming further that events are Poisson distributed about their true risk, we simulated the number of adverse events occurring each day over the 2000-day period by a Poisson distribution, with a daily mean equal to:

$$0.1 * e^{0.05 * X_{it}} \quad (1)$$

where $Xi_{it}$ is the personal exposure of the *i*th subject to the assumed causal pollutant on day *t*. This is equivalent to assuming that the exposure generated by this subject's covariance matrix was typical of the average covariance matrix of several thousand people in a population with an outcome such as a mortality or hospital admission.

(4) We regressed these events against each of the simulated ambient concentrations, producing a second estimated coefficient, $\beta_{pa}$, for each of the pollutants. $\beta_{pa}$ was compared with $\beta_p$ (i.e., 0.05 for the chosen pollutant), and to 0 for the other pollutants. To capture the range of possible covariances observed in Baltimore, this entire process was repeated 210 times (sampling with replacement). The mean of the $\beta_{pa}$'s was used as an indication of what might be expected for each ambient pollutant given the covariance between ambient concentrations and personal exposures and its distribution in Baltimore.

(5) We interpreted the mean of these 210 sets of coefficients to be comparable to what one might expect to find using Poisson regression analysis with ambient pollution as an exposure metric, in a city with 22 events per day, which is a reasonable count per day estimate for many observed pollution-related outcomes (Metzger et al., 2002). Multiple simulations were conducted to allow for each pollutant to have an assumed true association with a health risk.

(6) Finally, the entire process was repeated 200 times. The mean and empirical 95% confidence interval (CI) of the coefficients for the 200 replications was used to simulate the distribution of results that might be seen from a large multi-city study, where in each city we fit the Poisson regression of events for each air pollutant, under the assumption of a true association, with the chosen exposure. This was repeated assuming a true association with each exposure in turn.

(7) Results of the simulation are presented in Tables 1 and 2.

Table 1 gives the median $\beta$ for the association of ambient $PM_{2.5}$, ambient $O_3$, and ambient $NO_2$ with the health effect when the true association is with personal exposure to (1) total $PM_{2.5}$, (2) ambient $PM_{2.5}$, (3) sulfate, (4) $O_3$, and (5) $NO_2$. Since our interest is in the attenuation (or enhancement) due to measurement error, we report the reliability factor (median $\beta$/true $\beta$ [0.05]) (Fuller, 1987) for the pollutant that was assumed to be truly associated with the adverse health risk.

For the other pollutants, assumed not to be associated with the health outcome, the issue of reliability is irrelevant, and the results (median $\beta$ *versus* 0) provide an estimate of the deviation from the truth that might be expected, given measurement error and the correlations among the variables.

### *Additional Simulations*

We also conducted simulations using personal exposures to $SO_4^{2-}$, a component of $PM_{2.5}$ that is largely ambient in origin (Leaderer et al., 1999). Simulations including $SO_4^{2-}$ allowed us to examine what may happen with health risk estimates associated with $PM_{2.5}$ or its gaseous co-pollutants if the true health risk association was with a PM component. We

**Table 1.** Distribution of ambient concentrations and exposures.

| Variable | Mean | Standard deviation |
|---|---|---|
| *Personal exposure parameters* | | |
| Total personal $PM_{2.5}$ exposure | 20.9 $\mu g/m^3$ | 18.5 $\mu g/m^3$ |
| Personal exposure to $PM_{2.5}$ of ambient origin | 10.9 $\mu g/m^3$ | 6.5 $\mu g/m^3$ |
| Personal $SO_4^{2-}$ exposure | 3.1 $\mu g/m^3$ | 2.9 $\mu g/m^3$ |
| Personal $O_3$ exposure | 1.3 p.p.b. | 3.5 p.p.b. |
| Personal $NO_2$ exposure | 11.1 p.p.b. | 14.2 p.p.b. |
| Total personal CM exposure | 8.9 | 9.4 |
| Total personal $PM_{10}$ exposure | 30.7 | 22.5 |
| Personal exposure to CM of ambient origin | 1.8 $\mu g/m^3$ | 1.5 $\mu g/m^3$ |
| *Ambient concentration parameters* | | |
| Ambient $PM_{2.5}$ concentration | 21.2 $\mu g/m^3$ | 10.7 $\mu g/m^3$ |
| Ambient $O_3$ concentration | 23.8 p.p.b. | 12.7 p.p.b. |
| Ambient $NO_2$ concentration | 21.8 p.p.b. | 6.6 p.p.b. |
| Ambient $PM_{10}$ concentration | 30.2 $\mu g/m^3$ | 12.9 $\mu g/m^3$ |
| Ambient $SO_4^{2-}$ concentration | 7.4 | 6.4 |

CM, coarse-mode PM mass; PM, particulate matter.

**Table 2.** Distribution of coefficients observed in regressions against ambient pollutants, given a true health association with personal exposure, based on the covariance of personal and ambient exposures in Baltimore.

| Model scenario: | True association with personal exposure to total $PM_{2.5}$ | | |
|---|---|---|---|
| Observed association with: | Ambient $PM_{2.5}$ | Ambient $O_3$ | Ambient $NO_2$ |
| Median $\beta$ | 0.0143 | −0.0016 | 0.0115 |
| 95% CI | 0.0097, 0.0177 | −0.0056, 0.0025 | 0.0056, 0.0185 |
| Reliability factor | 0.286 | NA | NA |
| Model scenario: | True association with personal exposure to $PM_{2.5}$ of ambient origin | | |
| Observed association: | Ambient $PM_{2.5}$ | Ambient $O_3$ | Ambient $NO_2$ |
| Median $\beta$ | 0.0183 | −0.0037 | 0.0124 |
| 95% CI | 0.0168, 0.0200 | −0.0063, −0.0010 | 0.0085, 0.0173 |
| Reliability factor | 0.366 | NA | NA |
| Model scenario: | True association with personal exposure to particulate $SO_4^{2-}$ | | |
| Observed association with: | Ambient $PM_{2.5}$ | Ambient $O_3$ | Ambient $NO_2$ |
| Median $\beta$ | 0.0051 | 0.0035 | 0.0006 |
| 95% CI | 0.0036, 0.0065 | 0.0021, 0.0051 | −0.0014, 0.0028 |
| Reliability factor | 0.1 | NA | NA |
| Model scenario: | True association with personal exposure to $O_3$ | | |
| Observed association with: | Ambient $PM_{2.5}$ | Ambient $O_3$ | Ambient $NO_2$ |
| Median $\beta$ | 0.0014 | 0.0010 | 0.0009 |
| 95% CI | −0.0005, 0.0035 | −0.0007, 0.0024 | −0.0011, 0.0034 |
| Reliability factor | NA | 0.02 | NA |
| Model scenario: | True association with personal exposure to $NO_2$ | | |
| | Ambient $PM_{2.5}$ | Ambient $O_3$ | Ambient $NO_2$ |
| Median $\beta$ | 0.0015 | 0.0009 | 0.0010 |
| 95% CI | −0.0006, 0.0036 | −0.0008, 0.0025 | −0.0014, 0.0036 |
| Reliability factor | NA | NA | 0.02 |

CI, confidence interval; NA, not available; PM, particulate matter.
All simulations assume a true $\beta$-coefficient of 0.05.

conducted additional simulations including PM mass with an aerodynamic diameter between 2.5 and 10 $\mu$m, or coarse-mode PM mass (CM). CM was calculated as the difference between measured $PM_{10}$ and $PM_{2.5}$ exposures (for personal CM exposures) and concentrations (for ambient CM concentrations). We also examined what health risks might be found associated with concentrations of either CM or $PM_{2.5}$ if the true health risk association existed with exposure to the other pollutant.

As suggested by Wilson et al. (2000), the ratio of personal $SO_4^{2-}$ to ambient $SO_4^{2-}$ should be the same as the ratio for personal $PM_{2.5}$ to ambient $PM_{2.5}$, provided there are no indoor sources of $SO_4^{2-}$ and the $PM_{2.5}$ and $SO_4^{2-}$ have similar size distributions (Wilson et al., 2000). A recent study has shown that this technique may provide a good surrogate for exposure to ambient $PM_{2.5}$ mass in some US locations (Sarnat et al., 2002). Since the personal $PM_{2.5}$ exposures measured in our study included particles with indoor sources whose toxicity may differ from that of particles of outdoor origin, we used this technique to construct an estimate of personal exposure to ambient $PM_{2.5}$ using the equation:

$$\left(\frac{[SO_4^{2-}]_{\text{Personal}_{ij}}}{[SO_4^{2-}]_{\text{Ambient}_j}}\right) \cdot [PM_{2.5}]_{\text{Ambient Concentration}_j} = [PM_{2.5}]_{\text{Ambient Exposure}_{ij}} \quad (2)$$

where "Personal$_{ij}$" represents the personal exposure to $SO_4^{2-}$ for subject $i$ on day $j$, "Ambient Concentration$_j$" represents the ambient concentration measured at the stationary site on day $j$, and "Ambient Exposure$_{ij}$" represents the personal exposure to ambient $PM_{2.5}$ for subject $i$ on day $j$. Personal

exposure to the estimate of $PM_{2.5}$ of ambient origin was also included in the covariance matrixes, and the simulations were conducted with this variable.

An important feature of the simulations was that we separately calculated covariances from the winter and summer samples. Since the correlation between ambient $PM_{2.5}$ and ambient $O_3$, for example, is positive in the summer but negative in the winter, season-specific analyses seemed more reasonable than calculating one covariance for each subject and averaging over the two seasons.

Finally, because much of the published PM epidemiologic literature relates to $PM_{10}$ rather than $PM_{2.5}$, we also conducted simulations to examine what one might expect to find using ambient $PM_{10}$ concentrations as the exposure metric, assuming a true association with exposure to $PM_{2.5}$ of ambient origin, CM of ambient origin, or both. The estimates of exposure to CM and $PM_{10}$ of ambient origin were based on an approach by Wilson derived from the steady-state solution to the particle mass balance equation. A complete discussion of this method and its assumptions is described in Wilson and Brauer (2006) and Ebelt et al. (2005).

## Results

Table 1 shows descriptive statistics of the ambient concentrations and personal exposures for the subjects from the Baltimore panel study (Sarnat et al., 2001). Mean personal exposure to ambient particles ($10.9\,\mu g/m^3$) was approximately half the mean total personal $PM_{2.5}$ exposure ($20.9\,\mu g/m^3$). Mean personal $SO_4^{2-}$ exposure ($3.1\,\mu g/m^3$) comprised approximately one-seventh of the total personal $PM_{2.5}$ exposures and about one-third of the exposure to $PM_{2.5}$ of ambient origin. Results from the single-pollutant simulations showed that $PM_{2.5}$ was the only pollutant where a true association with its total personal exposures resulted in a significant observed association with its ambient concentrations ($\beta = 0.0143$, 95% CI: 0.0097, 0.0177; Table 2). Since the true, assumed coefficient was set at 0.05, the observed coefficient using ambient $PM_{2.5}$ concentrations is biased from 0.05 to 0.014 or biased low by roughly 70%. The bias is slightly lower ($\sim 60\%$) when the true association is assumed to be with personal exposure to $PM_{2.5}$ of ambient origin. In this case the association of the ambient concentration of $PM_{2.5}$ with the health risk is given by $\beta = 0.0183$, 95% CI: 0.0168, 0.0200. Similarly, a true association with personal exposure to $SO_4^{2-}$ resulted in a significant observed association with ambient $PM_{2.5}$ ($\beta = 0.0051$, 95% CI: 0.0036, 0.0065).

The simulated results also showed that true health risks associated with personal exposure to $O_3$ and $NO_2$ would result in no significant observed associations with any of their respective ambient concentrations. Conversely, a true association with $PM_{2.5}$ would result in a significant, observed association with $NO_2$ ($\beta = 0.0115$, 95% CI: 0.0056, 0.0185) and a true association with exposure to $SO_4^{2-}$ would result in an observed significant association with $O_3$ ($\beta = 0.0035$, 95% CI: 0.0021, 0.0051) given the covariance of the ambient pollutant concentrations.

Finally, the simulated findings showed that a true association with personal exposure to $PM_{2.5}$ of ambient origin would result in an observed association with both ambient $PM_{10}$ and CM ($\beta = 0.016$, 95% CI: 0.0147, 0.0174; $\beta = 0.0208$, 95% CI: 0.0164, 0.0257, respectively; Table 3). The relatively large coefficient for the association with ambient CM is expected due to the fact that the mean and standard deviation of CM of ambient origin is almost an order of magnitude lower than for $PM_{2.5}$ (Table 1). Thus, the modest correlation coefficient induces a large regression coefficient. Accordingly, the CI for this estimated coefficient is broader. Results for simulations where the reverse was true (i.e., personal exposure to CM of ambient origin was assumed to be the causal agent), showed that significant associations existed with ambient $PM_{10}$ and ambient CM but not with ambient $PM_{2.5}$. This would be expected because of the low correlation between ambient CM and ambient $PM_{2.5}$.

## Discussion

Epidemiologists have made substantial progress establishing exposure to PM as an environmental health concern and quantifying the magnitude of this risk. The questions for future PM research now focus on the specific component/s responsible for the observed effects. These questions are made more difficult due to the fact that (a) the components are frequently correlated with each other and with other gaseous pollutants, (b) relationships between ambient concentrations and personal exposures are often not well understood, and (c) that particulate matter, unlike its gaseous co-pollutants, does not represent a single chemical. Total personal exposures to $PM_{2.5}$ are comprised of particles from both indoor sources (e.g., cooking and cleaning) as well as particles from the outdoor sources (e.g., from traffic and power plants). Ambient $PM_{2.5}$ particles vary in their sources and chemical composition. This leads to several subsidiary questions. If true health effects are associated with exposures to $PM_{2.5}$ of ambient origin, as opposed to total personal $PM_{2.5}$ exposures, how would that affect the expected health risk estimates when using ambient concentrations in a health study? In addition, if observed health risk estimates were associated with only a component of the particles of ambient origin, such as $SO_4^{2-}$, how would that affect the associations when using ambient concentrations as the exposure metric? A goal of our analysis was to examine these questions in our simulations.

The pattern of results in these simulations based on the Baltimore panel study is quite informative. For example, our

**Table 3.** Distribution of coefficients observed in regressions against ambient particle measures, given a true association with personal exposure to either fine or coarse PM, based on the covariance of personal and ambient exposures in Baltimore.

| Model scenario: | True association with personal exposure to $PM_{2.5}$ of ambient origin | | |
|---|---|---|---|
| Observed association with: | Ambient $PM_{2.5}$ | Ambient CM | Ambient $PM_{10}$ |
| Median $\beta$ | 0.0183 | 0.0208 | 0.0160 |
| 95% CI | 0.0168, 0.0200 | 0.0164, 0.0257 | 0.0147, 0.0174 |
| Reliability factor | 0.366 | NA | NA |
| Model scenario: | True association with personal exposure to CM of ambient origin | | |
| Observed association with: | Ambient $PM_{2.5}$ | Ambient CM | Ambient $PM_{10}$ |
| Median $\beta$ | 0.0007 | 0.0050 | 0.0008 |
| 95% CI | −0.0004, 0.0018 | 0.0020, 0.0077 | 0.0000, 0.0017 |
| Reliability factor | NA | 0.1 | NA |
| Model scenario: | True association with personal exposure to both $PM_{2.5}$ and CM of ambient origin | | |
| Observed association with: | Ambient $PM_{2.5}$ | Ambient CM | Ambient $PM_{10}$ |
| Median $\beta$ | 0.0198 | 0.0239 | 0.0167 |
| 95% CI | 0.0185, 0.0211 | 0.0197, 0.0278 | 0.0155, 0.0180 |
| Reliability factor | NA | NA | 0.334 |

CI, confidence interval; CM, coarse-mode PM mass; NA, not available; PM, particulate matter.
All simulations assume a true coefficient of 0.05.

results suggested that if there were a true health risk associated with total personal $PM_{2.5}$ exposures, one would expect to see a corresponding significant association using ambient $PM_{2.5}$ concentrations, albeit biased low relative to the real association. Moreover, if a true causal association existed with personal exposure to $PM_{2.5}$ of ambient origin, similar results would be seen using ambient $PM_{2.5}$ concentrations, with a more modest reduction in the attenuation factor. However, in this latter case, the negative health risk estimate also observed to exist with ambient $O_3$ becomes stronger and more significant. Based on the regressions presented in Sarnat et al. (2001), this is likely driven by the negative correlations between ambient $O_3$ and ambient $PM_{2.5}$ concentrations during the winter in Baltimore. These associations would be expected to result in a stronger negative association between ambient ozone and daily events if the events were truly associated with the personal exposure to particles of ambient origin than if the events were associated with all particle exposure, since particles of indoor origin are not correlated with ambient ozone.

The results involving $SO_4^{2-}$ were particularly interesting and suggest that if only a small component of ambient $PM_{2.5}$, such as $SO_4^{2-}$, were responsible for an observed health effect, we would still expect to see an association with ambient $PM_{2.5}$. Obviously this result may differ for a component less well correlated with ambient $PM_{2.5}$. The magnitude of the association of health effects with the component can be calculated from the value of the association found with ambient $PM_{2.5}$ if the slope and coefficient of determination of the regression of $PM_{2.5}$ on the component are known.

In the eastern US, $SO_4^{2-}$ is present in greater concentrations during the summer, and produced by the same photochemistry and long-range transport that produces high summertime $O_3$ levels in Baltimore. Hence, $O_3$ is more strongly correlated with $SO_4^{2-}$ during the summer than with total ambient $PM_{2.5}$. If there were a true association just with the $SO_4^{2-}$ particles, our results indicated that one would also expect to see positive observed associations between $O_3$ and daily events.

These results suggest that in some locations, observed associations with $O_3$ should be interpreted with caution, as they may be reflecting associations with PM or one of its chemical components. In the Baltimore simulation, observed $O_3$ associations could have reflected true health association with $SO_4^{2-}$ particles, which are created by the same secondary photochemistry processes that produce $O_3$. This can be assessed using models that control for $SO_4^{2-}$ particles, but these data are not generally available. There is an added reason to interpret the $O_3$ health risk carefully, given that our simulations showed that a true association with personal exposure to $O_3$ did not result in an observed association with any ambient pollutant concentration.

Conversely, our results show that a true association with $PM_{2.5}$ of ambient origin would lead to a finding of associations with $NO_2$, whereas if the association were with $SO_4^{2-}$ exposure only, no $NO_2$ association would be seen. This is probably due to fact that $NO_2$ is more strongly correlated with PM from traffic sources and not secondary $SO_4^{2-}$. Were the true association only with the traffic particles, the pattern would likely be opposite, and $NO_2$ would remain significant,

while the association with $O_3$ would disappear or become negative.

Given these results, a two-pollutant model can be viewed as a form of a source apportionment model, since in some locations ambient $O_3$ is a better predictor of exposure to $SO_4^{2-}$ than of exposure to all particles (or to $O_3$) (Sarnat et al., 2001). Likewise, ambient $NO_2$ may be a better predictor of exposure to traffic particles than of exposure to all particles (or to $NO_2$). Therefore, in a model with both $O_3$ and $PM_{2.5}$, one might interpret an $O_3$ coefficient as representing the effect of secondary particles, and the coefficient of particles as representing the effect of $PM_{2.5}$ exposure that is independent of $SO_4^{2-}$, which will mostly be primary particles. In a two-pollutant model with $PM_{2.5}$ and $NO_2$, the interpretation would be reversed from above.

It should be strongly noted that we do not interpret the results of our simulation as indicative of the lack of adverse effects associated with exposures to either $O_3$ or the other gaseous pollutants. The results do indicate that in Baltimore during this study period, ambient $O_3$ concentrations may not have been adequate surrogates for corresponding personal $O_3$ exposures to allow the question to be investigated using ambient monitors. This may not be true in other locations or for other population cohorts. For example, recent analyses from our Boston exposure assessment study showed that ambient $O_3$ was significantly associated with corresponding personal exposures (Sarnat et al., 2005). Similarly, ambient $O_3$ concentrations were shown to be good surrogates for exposure to $O_3$ for children spending considerable time outdoors during summer camp (Brauer and Brook, 1997). Even in the summer camp studies, however, it may be difficult to separate $O_3$ effects from $SO_4^{2-}$ effects. Indeed, any study using ambient $O_3$ that did not consider ambient $SO_4^{2-}$ as an alternative explanation of any observed association will need to be interpreted with caution, particularly in the eastern US.

The results observed for models using $NO_2$ also suggest caution in interpreting significant associations with ambient $NO_2$ in epidemiology studies. A true association with exposure to $NO_2$ may not be expected to result in an observed association with ambient $NO_2$, whereas a true association with $PM_{2.5}$ exposure may result in an observed association with ambient $NO_2$. Once again it should be noted that, as in the case of $O_3$, the Sarnat et al. (2001) study observed that personal exposure to $NO_2$ was not well correlated with ambient concentration of $NO_2$ in Baltimore. For cities with exposure patterns like those of Baltimore, these results again suggest that associations with ambient $NO_2$ are much more likely to represent the results of a true association with exposure to $PM_{2.5}$ than with exposure to $NO_2$. This makes the use of two-pollutant models highly suspect in this case as well.

Finally, our results showed that a true association of health risk with either exposure to $PM_{2.5}$ of ambient origin or exposure to CM of ambient origin would also likely result in an observed significant association with $PM_{10}$. This is a useful finding in interpreting health risks associated with $PM_{10}$. Of equal interest was the result showing that a true association with exposure to $PM_{2.5}$ of ambient origin would result in an observed association with ambient CM, but a true association with CM of ambient origin would not result in an observed association with ambient $PM_{2.5}$. In fact, $PM_{2.5}$ was the only pollutant examined for which a false health association would not be observed, given a causal association with another pollutant. Also, because exposure to CM of ambient origin is correlated with ambient $PM_{2.5}$, but has a much lower mean, a true association with $PM_{2.5}$ can actually result in a larger regression coefficient for CM than for $PM_{2.5}$.

There are several limitations to our methodology, including the issue of differential measurement error among the personal samplers used in the Baltimore panel study. This is especially true for $O_3$, where a majority of personal exposure measurements were below the analytical limits of detection (Koutrakis, 2005). Our results indicated that ambient gas monitors were generally poor surrogates of personal gas exposures for a 24-h integrated period. It is likely that some of the simulation results for the gases may also be reflecting differential measurement error, (i.e., the personal exposure samplers may not have provided as accurate a measure of true personal exposure for the gases as they did for $PM_{2.5}$). Finally, it is worth emphasizing that our results are based on simulations from a relatively few number of individuals in a single location. Using results from exposure assessment studies conducted in other locations with different pollutant mixtures, meteorology and exposure factors will produce covariance matrices that may differ substantially from those observed in Baltimore during this study. We therefore believe that repeating this and similar simulations are necessary to validate the robustness of our reported findings and attain a clearer understanding of the effects of measurement error in air pollution epidemiologic findings.

## Acknowledgements

This work was supported in part by EPA Contracts 1C-R238-NASA with Joel Schwartz and EP07C000040 with Jeremy A. Sarnat and in part by EPA Grant R-832416.

## Disclaimer

This article has been reviewed by the U.S. Environmental Protection > Agency and cleared for publication. However, the views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the EPA.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

## References


Atkinson R.W., Anderson H.R., Sunyer J., Ayres J., Baccini M., Vonk J.M., Boumghar A., Forastiere F., Forsberg B., Touloumi G., Schwartz J., and Katsouyanni K. Acute effects of particulate air pollution on respiratory admissions—results from APHEA 2 project. *Am J Respir Crit Care Med* 2001: 164(10): 1860–1866.

Brauer M., and Brook J. Ozone personal exposures and health effects for selected groups residing in the Frazer Valley. *Atmos Environ* 1997: 31(14): 2113–2121.

Demokritou P., Kavouras I., Ferguson S., and Koutrakis P. Development and laboratory performance evaluation of a multipollutant sampler for simultaneous measurements for particulate and gaseous pollutants. *Aerosol Sci Technol* 2001: 35: 741–752.

Ebelt S.T., Wilson W.E., and Brauer M. Exposure to ambient and nonambient components of particulate matter—a comparison of health effects. *Epidemiology* 2005: 16(3): 396–405.

Fuller W.A. *Measurement Error Models*. John Wiley & Sons: New York, 1987.

Gamble J.F., and Nicolich M.J. Comparison of ambient PM risk with risks estimated from PM components of smoking and occupational exposures. *J Air Waste Manag Assoc* 2000: 50(8): 1514–1531.

Gold D.R., Litonjua A., Schwartz J., Lovett E., Larson A., Nearing B., Allen G., Verrier M., Cherry R., and Verrier R. Ambient pollution and heart rate variability. *Circulation* 2000: 101(11): 1267–1273.

Katsouyanni K., Touloumi G., Spix C., Schwartz J., Balducci F., Medina S., Rossi G., Wojtyniak B., Sunyer J., Bacharova L., Schouten J.P., Ponka A., and Anderson H.R. Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *Br Med J* 1997: 314(7095): 1658–1663.

Koutrakis P. *Characterization of Particulate and Gas Exposures of Sensitive Subpopulations Living in Baltimore and Boston*. Health Effects Institute: Boston, MA, 2005.

Koutrakis P., Wolfson J.M., Bunyaviroch A., and Froehlich S. A passive ozone sampler based on a reaction with nitrite. *Res Rep Health Eff Inst* 1994: (63): 19–47; discussion 67–82.

Le Tertre A., Medina S., Samoli E., Forsberg B., Michelozzi P., Boumghar A., Vonk J.M., Bellini A., Atkinson R., Ayres J.G., Sunyer J., Schwartz J., and Katsouyanni K. Short-term effects of particulate air pollution on cardiovascular diseases in eight European cities. *J Epidemiol Community Health* 2002: 56(10): 773–779.

Leaderer B.P., Naeher L., Jankun T., Balenger K., Holford T.R., Toth C., Sullivan J., Wolfson J.M., and Koutrakis P. Indoor, outdoor, and regional summer and winter concentrations of PM10, PM2.5, SO4(2)−, H+, NH4+, NO3−, NH3, and nitrous acid in homes with and without kerosene space heaters. *Environ Health Perspect* 1999: 107(3): 223–231.

Metzger K.B., Tolbert P.E., Klein M., Flanders W.D., Peel J.L., Mulholland J.A., and Langberg J.J. Ambient air pollution and arrhythmic events in patients with automatic implantable cardioverter defibrillators, Atlanta, 1993–2000 (ARIES/SOPHIA). *Epidemiology* 2002: 13(4): S125.

National Research Council. *Research Priorities for Airborne Particulate Matter. I. Immediate Priorities and a Long-Range Research Portfolio*. National Academy Press: Washington, DC, 1998.

Ogawa & Company. NO, NO2, NOx, and SO2 sampling protocol using the Ogawa sampler. Version 3. Pompano Beach, FL, 1998.

Salvi S., Blomberg A., Rudell B., Kelly F., Sandstrom T., Holgate S.T., and Frew A. Acute inflammatory responses in the airways and peripheral blood after short-term exposure to diesel exhaust in healthy human volunteers. *Am J Respir Crit Care Med* 1999: 159(3): 702–709.

Sarnat J.A., Brown K.W., Schwartz J., Coull B.A., and Koutrakis P. Ambient gas concentrations and personal particulate matter exposures—implications for studying the health effects of particles. *Epidemiology* 2005: 16(3): 385–395.

Sarnat J.A., Koutrakis P., and Suh H.H. Assessing the relationship between personal particulate and gaseous exposures of senior citizens living in Baltimore, MD. *J Air Waste Manag Assoc* 2000: 50(7): 1184–1198.

Sarnat J.A., Long C.M., Koutrakis P., Coull B.A., Schwartz J., and Suh H.H. Using sulfur as a tracer of outdoor fine particulate matter. *Environ Sci Technol* 2002: 36(24): 5305–5314.

Sarnat J.A., Schwartz J., Catalano P.J., and Suh H.H. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? *Environ Health Perspect* 2001: 109(10): 1053–1061.

Schwartz J., and Dockery D.W. Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am Rev Respir Dis* 1992: 145(3): 600–604.

Schwartz J., Norris G., Larson T., Sheppard L., Claiborne C., and Koenig J. Episodes of high coarse particle concentrations are not associated with increased mortality [see comments]. *Environ Health Perspect* 1999: 107(5): 339–342.

Schwartz J. Air pollution and blood markers of cardiovascular risk. *Environ Health Perspect* 2001: 109: 405–409.

Schwartz J. Air pollution and hospital admissions for the elderly in Detroit, Michigan. *Am J Respir Crit Care Med* 1994: 150(3): 648–655.

Wilson W.E., and Brauer M. Estimation of ambient and non-ambient components of particulate matter exposure from a personal monitoring panel study. *J Expo Sci Environ Epidemiol* 2006: 16(3): 264–274.

Wilson W.E., Mage D.T., and Grant L.D. Estimating separately personal exposure to ambient and nonambient particulate matter for epidemiology and risk assessment: why and how. *J Air Waste Manag Assoc* 2000: 50(7): 1167–1183.

Zeger S.L., Thomas D., Dominici F., Samet J.M., Schwartz J., Dockery D., and Cohen A. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 2000: 108(5): 419–426.