Taylor & Francis
Taylor & Francis Group

Journal of the Air & Waste Management Association

ISSN: 1096-2247 (Print) 2162-2906 (Online) Journal homepage: http://www.tandfonline.com/loi/uawm20

Assessing the Impact of Differential Measurement Error on Estimates of Fine Particle Mortality

Timothy J. Carrothers & John S. Evans

To cite this article: Timothy J. Carrothers & John S. Evans (2000) Assessing the Impact of Differential Measurement Error on Estimates of Fine Particle Mortality, Journal of the Air & Waste Management Association, 50:1, 65-74, DOI: 10.1080/10473289.2000.10463988

To link to this article: http://dx.doi.org/10.1080/10473289.2000.10463988

Published online: 27 Dec 2011.

Submit your article to this journal

Article views: 84

View related articles

Citing articles: 7 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=uawm20

Download by: [University of North Texas] **Date:** 06 January 2017, At: 21:58

ISSN 1047-3289 *J. Air & Waste Manage. Assoc.* **50**:65-74

Copyright 2000 Air & Waste Management Association

Assessing the Impact of Differential Measurement Error on Estimates of Fine Particle Mortality

Timothy J. Carrothers and John S. Evans
Harvard School of Public Health, Boston, Massachusetts

ABSTRACT

In air pollution epidemiology, error in measurements of correlated pollutants has been advanced as a reason to distrust regressions that find statistically significant weak associations. Much of the related debate in the literature and elswhere has been qualitative. To promote quantitative evaluation of such errors, this paper develops an air pollution time-series model based on correlations among unit-normal variables. Assuming there are no other sources of bias present, the model shows the expected amount of relative bias in the regression coefficients of a bivariate regression of coarse and fine particulate matter measurements on daily mortality. The model only requires information on instrumental error and spatial variability, along with the observed regression coefficients and information on the true fine-course correlation. Analytical results show that if one pollutant is truly more harmful than the other, then it must be measured more precisely than the other in order not to bias the ratio of the fine and course regression coefficients. Utilizing published data, a case study of the Harvard Six-Cities study illustrates use of the model and emphasizes the need for data on spatial variability across the study area. Current epidemiology time-series regressions can use this model to address the general concern of correlated pollutants with differing measurement errors.

IMPLICATIONS

Ambient air pollution regulations must be based on the best available science, including a thorough treatment of statistical issues in air pollution epidemiology. One such issue, differences in measurement error between fine and coarse particle exposures, may yield biased and misleading estimates of toxicity. The quantitative model developed in this paper estimates the expected bias of particle toxicities as a function of the true correlation of fine and coarse particle exposures and the error with which each is measured. Application of this model to ongoing studies should help determine whether standards should focus on PM_{10}, $PM_{2.5}$, or both.

INTRODUCTION

Few scientists will debate the human health effects of extremely high air pollution episodes such as the London fogs of the 1950s.[1,2] By contrast, there is great debate regarding the effects of current U.S. air pollution levels.[3-5] Many scientists continue to question the scientific basis for the U.S. Environmental Protection Agency's (EPA) recent revisions of primary ambient air quality standards for particulate matter and ozone.[4,6,7] Some of the main issues raised include exposure errors, residual confounding, and correlated pollutants.[3,4,8,9] The uncertainties associated with these issues are rarely quantified, much less resolved, in the usual qualitative arguments set forth in the current epidemiological literature. This paper attempts to better quantify the uncertainty involved with one particular argument, the concern over correlated pollutants with different levels of exposure error.

The paper begins with a review of the relevant statistical and epidemiological background, followed by a discussion of what constitutes exposure error for regression analysis in the "time-series" study design of air pollution epidemiology. Next, a bivariate analytical model is developed to quantitatively predict the expected bias in regression coefficients. After a brief discussion of how published values are used as inputs for the model, a case study using the Harvard Six-Cities data set is presented.

BACKGROUND

Exposure Terminology

Airborne particles are classified in three ways: mode, cut point, and dosimetry.[10] Classification by mode is based on whether the particles are predominantly created by the accumulation of smaller particles from combustion sources or photochemical processes (i.e., "fine mode") or by mechanical grinding processes (i.e., "coarse mode"). Classification by cut point is based on the 50% cut point (e.g., 2.5 µm and 10 µm, of a standardized measuring device). Classification by dosimetry is based on how far particles penetrate into the human respiratory system.

The terms "fine" and "coarse," originally associated with "mode," are now also used to describe classification by cut point. $PM_{2.5}$ is often called the "fine" fraction of PM_{10}, while the $PM_{10-2.5}$ fraction of PM_{10} is termed the "coarse" fraction. This is an inexact classification, since the size modes overlap between 1- and 3-µm diameter, but it is now used in the exposure and health effects literature.[10] For the purposes of this paper, we will use this classification, denoting $PM_{2.5}$ measurements as *F* and $PM_{10-2.5}$ as *C*, to represent the approximate fractions of PM_{10} that are fine and coarse mode particles, respectively.

Statistical Terminology

The terms "measurement error" or "errors in variables" refer to any differences between the true value of a variable *x* and the observed or measured variable *z*. Depending on their relation to the outcome variable *y*, the errors will be "differential" or "non-differential." Differential errors, or errors that are related to the outcome variable, can be sources of sizable bias, and proper study design is needed to avoid them.[11] Formally, non-differential errors imply Pr (y | x, z) = Pr (y | x) (i.e., the measurement contains no information about the outcome above and beyond the information contained in the true value of *x*).[12] Non-differential errors can also induce several problems, including bias, but analytic techniques may be developed to address these problems.[12-15] This paper will only address non-differential errors.

The covariance of random variables *A* and *B* is defined by $Cov(A,B) = E[(A - \mu_A)(B - \mu_B)]$. The correlation of *A* and *B* is defined as $Corr(A,B) = Cov(A,B) / ((Var\ A)(Var\ B))^{0.5}$.[16] By itself, correlation among predictor variables in a regression hinders precise estimates of individual coefficients but will not hinder estimation of their joint effect.[13] "Confounding" is defined as the bias introduced to the observed coefficient of *A* by a variable *B* that is correlated with *A*, is an independent predictor for the outcome variable, *and* is either unaccounted for or imperfectly measured in the regression.[11]

In this paper, the regression contains predictor variables *F* and *C*, representing $PM_{2.5}$ and $PM_{10-2.5}$ as defined above. It is assumed that the coefficients of *F* or *C* are not confounded by other variables. As such, only the joint effect of measurement error in *F*, measurement error in *C*, and the correlation of *F* and *C* on the observed regression coefficients for *F* and *C* is addressed. Although confounding by additional pollutants or weather may be a serious additional concern,[3] it is not included in this analysis. For even in the absence of confounding by other variables, the joint effect of measurement error and correlation may preclude valid conclusions about *F* and *C* in certain epidemiology studies.[17]

Statistical Background

Quantitative approaches to the effects of measurement error in regression analysis trace back three decades to Cochran's 1968 *Technometrics* paper,[18] in which Cochran wrote on the calculation of expected values of true variable *x*, given observed variable *z*. This is known as the "classical error" model, where the measurement model $Pr(z|x)$ and the true distribution $Pr(x)$ are specified so that $E[x|z]$ may be calculated. An idealization of the classical error model is measuring the length of a line with a ruler. The expectation of the observed value is the true value (i.e., E[z|x] = x), and the standard deviation of the measurement from the truth is determined by the quality of the ruler and skill of the operator.

In the univariate classical error model, non-differential exposure error results in attenuation of both the exposure-response relationship and its statistical significance.[13,19] It may also obscure a true threshold relationship.[20] In a model with more than one predictor variable, the problem is more complicated.[21] Bias of the regression coefficients may occur in either direction due to the interplay of correlation and measurement errors. Measurement error can also create or hide effect modification.[19]

In some situations, it may be more appropriate to specify $Pr(x|z)$. Here, the classical error model does not apply. The idealization for this situation is a machine that delivers doses *x* based on a dial-setting *z*. If the machine has the additional property E[x|z] = z, a "Berkson error" model applies. An example of this occurs when all members of a group, each with true exposure x_i, are assigned the group mean exposure of *z*. In the Berkson error model, the effects of measurement error are the same as the classical model with one important exception: if the true dose-response is linear, the estimate of the slope parameter is unbiased.[19]

The univariate case of measurement error has been well studied for a model with a linear exposure-response relationship.[19,20] Recent publications have reviewed the current methods for understanding and correcting the effects of measurement.[12,19] Thomas et al. state that although qualitative conclusions are often made in this area, quantitative treatments and conclusions are quite rare.[12] Some recent publications have shown applications to the air pollution context.[8,15,22]

RELEVANCE TO AIR POLLUTION EPIDEMIOLOGY

In air pollution epidemiology, it is often infeasible to use personal monitors to precisely determine average human exposure to ambient pollutants. Instead, measurements made at central site monitoring stations serve as proxies (i.e., imperfect measurements of the true quantities). If proxies for two or more correlated pollutants are included in a multivariate regression, it may be difficult to separate their

effects on the dependent variable if the pollutants are measured with differing amounts of error. In cases where the true correlation of the pollutants or the difference in measurement error is significant, variables with smaller measurement error may absorb a non-trivial portion of the effect actually attributable to the other pollutant.[17] In the extreme, when the proxy for one variable is a better measure of the other variable than its own proxy, a benign pollutant could be found harmful while the truly harmful pollutant is found benign.[23]

Several other experts have recently raised this issue as a major concern for inferences regarding the true impact of air pollution. Moolgavkar and Leubeck[4] concluded that the contributions of each constituent of air pollution could not be determined, in part due to statistical shortcomings such as this. Dr. Anne Smith's testimony before Congress focused primarily upon the extreme case. She argued that air pollutants with relatively smaller exposure error (i.e., fine particulate matter and gases) could be wrongly implicated for the effects of air pollutants with relatively larger exposure error (i.e., coarse particulate matter).[23]

The Health Effects Institute (HEI) initiated a re-analysis of the Philadelphia time-series data in part to examine the effects of multiple pollutants. In their Phase I.B report, members of the HEI review team recognized that their results did not consider the possibility of measurement errors that differed across pollutants. They stated further, "Simple generalizations cannot be offered concerning the consequences of measurement error and residual confounding in complex data sets with multiple, correlated variables." The HEI oversight committee indicated in the same report that Phase II of the project, currently in progress, would attempt to develop models to address exposure errors and their effects on daily time-series studies.[24]

Other experts are less convinced that this issue presents such a major hurdle to inference, given the large body of air pollution epidemiology research.[25,26] They argue that studies have been conducted under a variety of study designs, air pollution levels, air pollutant correlations, study populations, and weather conditions with relatively robust results for the effects of particulate matter. In particular, the proxies used for ambient particulate matter correlate far better with fine particles from ambient sources than coarse particles from ambient sources, particles generated indoors, or personal cloud particles.[27,28] They maintain that this strongly bolsters the case that ambient fine particles are responsible for the effects seen in the studies.

EXPOSURE ERROR

Which Errors Matter?

Before one can assess the effects of exposure error in these studies, one must determine what constitutes exposure error. For a study intended to estimate the increased risk of health effects due to exposure to airborne pollutants, a general definition of exposure error would be any differences between the proxy measurement used in the regression and an individual's actual exposure. Several sources of difference are possible, including instrumental/analytical error, spatial variability, temporal variability, and variability in human time-activity patterns. Lipfert[7] provides a discussion of each of these sources of error. However, not all of these sources of difference will necessarily introduce bias into a linear dose-response relation, as some of them may follow a Berkson error model. Schwartz[29] has used the following equation for time-series studies:

$$x_{it} = z_t + (x_t - z_t) + (x_{it} - x_t) \qquad (1)$$

where x_{it} is the ith person's total exposure on day t, x_t is the mean personal total exposure on day t,[30] and z_t is the central site monitor total exposure on day t.

The second term, $(x_t - z_t)$, the difference between the mean personal and monitor exposure, can introduce bias in the regression results. This relation reflects both instrumental/analytical error and spatial variability in exposure. The third term, $(x_{it} - x_t)$, the difference between a specific person's exposure and the mean personal exposure, follows a Berkson error structure and will not introduce bias in the regression if the dose-response function is linear. This third term reflects temporal variability in exposure, variability across buildings in building penetration rates, and variability across individuals in time-activity patterns. If one is concerned only about the potential for bias of the regression coefficients, then one should focus exclusively on the types of errors that can introduce bias. For this paper, a linear dose-response is assumed, thus enabling the Berkson simplification.

Which Exposure Does the Proxy Estimate?

Personal exposure to particles can be broken down by source of generation: ambient (e.g., industrial and automotive emissions and suspended soils), indoor (e.g., cooking and passive smoke), and personal cloud (e.g., smoking and particles resuspended by personal activity). Although indoor and personal cloud exposures are the dominant source of personal exposure to particles, EPA only has the authority to regulate ambient air pollution. If exposures to these other sources of particles are confounding the observed association between mortality and ambient pollution, EPA's regulations could be misguided. However, such confounding is possible only if exposure to indoor-generated and personal cloud particles is correlated with exposure to ambient-generated particles in the context of the study design (e.g., time-series) concerned.

Results from the Particle Total Exposure Assessment Methodology (PTEAM) study indicate that daily exposures to indoor-generated and personal cloud-generated particles are not correlated with ambient-generated particles.[28,31] These

results imply that ambient central site measurements should only be thought of as proxies for average exposure to *ambient-generated* particles. Although future exposure studies are needed to reinforce the PTEAM results, it is assumed for this paper that this implication holds true.

In this section, it has been shown that information on instrumental/analytical error and spatial variability may be all that is required to assess the potential bias due to differential measurement error in regression coefficients from time-series studies. In the next section, a simplified model is introduced for determining the expected bias from the interplay of measurement error and correlation. Without loss of generality, the model is developed with correlated unit normal variables.

REGRESSION SIMULATION MODEL

Description

The model simulates a time-series study design that constitutes the vast majority of recent air pollution epidemiology studies.[4] In this design, the relation between pollution variables and daily death counts is studied for a specified geographical area. In the simulation model, pollutants *F* and *C* represent the true average population exposure of the fine and coarse fractions of PM_{10}. Daily observations of each pollutant are assumed to be unit normal variables, and *F* and *C* are assumed to be positively correlated.

Twenty-four-hr averaged pollution measurements at centrally located ambient monitors are used as proxies for average population exposure to ambient particles. Measurements of *F* and *C* are denoted as *F'* and *C'*, respectively. Without loss of generality, these are also unit normal variables. The measurement error between pollutant measurements and true pollutant exposure is represented as a positive correlation between 0 and 1. For instance, Corr(F, F') = 0.5 would indicate that one-fourth (i.e., $(0.5)^2$) of the variation in the measurement of *F* is determined by the variation in *F*, and the remaining three-fourths is determined by random error.

The model assumes that the total measurement errors of *F* and *C* are uncorrelated on a day-to-day basis. This simplification is justified by spatial variability's dominance of instrumental error in determination of the total error. Although instrumental errors for *F* and *C* may well be correlated,[32] there is little evidence that day-to-day changes in the spatial pattern of *F* and *C* across the city are correlated.

The dependent variable *Y*, the portion of daily mortality attributable to these two pollutants, is assumed to be a linear function of fine and coarse particle levels, as follows:

$$Y = \beta_F * F + \beta_C * C \qquad (2)$$

where β_F and β_C are chosen a priori to reflect various possible true impacts of fine and coarse particles on mortality. The regression performed in epidemiology is a function of measurements of the pollution levels,

$$Y = \left(\hat{\beta}_F * F'\right) + \left(\hat{\beta}_C * C'\right) + \varepsilon \qquad (3)$$

where $\hat{\beta}_F$ and $\hat{\beta}_C$ are estimators of β_F and β_C. The results of the regression are influenced by the measurement error in *F* and *C*, as well as the underlying true correlation of *F* and *C*.

With unit normal variables, the ordinary least-squares estimate of the coefficients is given by

$$\hat{\beta} = (X'X)^{-1}\, X'Y \qquad (4)$$

$$\text{where } X = \begin{bmatrix} F_1' & C_1' \\ F_2' & C_2' \\ \vdots & \vdots \\ F_n' & C_n' \end{bmatrix} \quad Y = \begin{bmatrix} \beta_F * F_1 + \beta_C * C_1 \\ \beta_F * F_2 + \beta_C * C_2 \\ \vdots \\ \beta_F * F_n + \beta_C * C_n \end{bmatrix} \text{ and } \hat{\beta} = \begin{bmatrix} \hat{\beta}_{F'} \\ \hat{\beta}_{C'} \end{bmatrix}$$

Taking expectations for the regression coefficients, we obtain

$$E\left[\hat{\beta}_{F'}\right] = \frac{\left[\beta_F * Cov(F',F) + \beta_C * Cov(F',C)\right] - \left\{\frac{Cov(C',F')}{V_{C'}} * \left[\beta_C * Cov(C',C) + \beta_F * Cov(C',F)\right]\right\}}{V_{F'} - \frac{Cov(C',F')^2}{V_{C'}}} \qquad (5)$$

$$E\left[\hat{\beta}_{C'}\right] = \frac{\left[\beta_C * Cov(C',C) + \beta_F * Cov(C',F)\right] - \left\{\frac{Cov(C',F')}{V_{F'}} * \left[\beta_F * Cov(F',F) + \beta_C * Cov(F',C)\right]\right\}}{V_{C'} - \frac{Cov(C',F')^2}{V_{F'}}} \qquad (6)$$

which reflect the dependence of the coefficients on the fine and coarse particle measurement error, Corr(F,F') and Corr(C, C'), respectively, and the true correlation between fine and coarse particles, Corr(F,C).

A convenient summary of the relative amount of expected bias in the estimates of the regression coefficients is given by "bias ratio," *B*,

$$B = \frac{\left\{ E\left[\hat{\beta}_{F'}\right] \Big/ E\left[\hat{\beta}_{C'}\right] \right\}}{\left\{ \beta_F \Big/ \beta_C \right\}} \qquad (7)$$

where high values of *B* indicate overestimation of *F*'s coefficient relative to *C*'s coefficient.

Discussion of Results

For the sake of brevity, results are presented for three illustrative, but plausible, cases, $(\beta_F/\beta_C) = \{1, 3, 10\}$. These results are summarized in a series of three figures: Figure 1 shows results for equally toxic pollutants (i.e., $(\beta_F/\beta_C) = 1$); Figure 2 shows results for the case where one pollutant is somewhat more harmful than the other (i.e., $(\beta_F/\beta_C) = 3$); and Figure 3 shows results for the case where one pollutant is substantially more harmful than the other (i.e., $(\beta_F/\beta_C) = 10$).

All of the graphs have the same structure. The degree of fine particle measurement error, denoted as Corr(F,F'), is given on the *x*-axis. The degree of coarse particle measurement error, denoted as Corr(C,C'), is given on the *y*-axis. The contours in the body of the graph show values of the expected bias, *B*, in the ratio of the estimated fine/coarse regression coefficients. Because the expected bias is a function of the true correlation between fine and coarse particle exposures, each figure includes three graphs, each corresponding to a different correlation. Values of 0.25, 0.50, and 0.75 are shown to illustrate a range of possible true correlations between fine and coarse particle exposures.

For the case in which the pollutants are equally toxic (see Figure 1), there is no bias in the ratio of the estimated coefficients (i.e., B = 1) as long as the pollutants are measured with equal precision. However, if they are measured with unequal precision, the ratio of the estimated coefficients will be biased and therefore potentially misleading. For example, if fine particles are better measured than coarse, then it will appear that fine particles are more toxic than coarse (i.e., B > 1), and conversely, if coarse particles are better measured than fine, it will appear as if coarse particles are more toxic than fine (i.e., B < 1). The degree of bias depends on the degree of relative measurement error between fine and coarse particle measurements and on the true correlation between fine and coarse particle exposures. As the true correlation between fine and coarse particle exposures increases, the impact of any imbalance in the relative measurement error increases (i.e., the value of *B* moves further away from 1).

When one pollutant (e.g., fine particles) is more toxic than the other (e.g., coarse particles), as shown in Figures 2 and 3, exposures to the more toxic pollutant must be measured with less error in order for the ratio of estimated coefficients to be unbiased. This need for different measurement error increases as the underlying difference in true toxicity increases (Figure 1a->2a->3a). Within each figure, the potential for bias increases as the true correlation between fine and coarse particle exposure increases (Figure 3a->3b->3c), and accordingly, the need for even more precise measurement of fine particles grows.

In contrast to Figure 1, if the pollutants in Figures 2 and 3 are measured with equal precision, the ratio of the estimated fine coefficient to the estimated coarse coefficient will be biased downward, resulting in a relative overestimation of the less toxic pollutant. Again, the degree of bias increases with higher correlations between fine and coarse particle exposure.

This rather counterintuitive result can be explained by examining the terms in the equations for $E[\beta_{F'}]$ and $E[\beta_{C'}]$. Since the results are symmetrical, only the result for $E[\beta_{F'}]$ will be explained. The numerator consists of two contributions and two subtractions, while the denominator is a normalizing constant. The first contribution is the "true" contribution from the true value of $\beta_{F'}$ diluted by measurement error in *F*. The second contribution is the "false" contribution from the true value of β_C, where the full effect of *C* is diluted once by the correlation of the true quantities and then by the measurement error in *F*. The two subtraction terms are the corresponding contributions to $E[\beta_{C'}]$, except for dilution by the ratio of the correlation of the measurements to the variance of *C'*. When *F* is more potent than *C*, and both *F* and *C* are measured with the same error, the "false" contribution from *C* to *F* will not make up for the corresponding subtraction from *F* to *C*. Thus, the stronger pollutant must be measured more accurately than the weaker pollutant to balance out the unequal transfers.

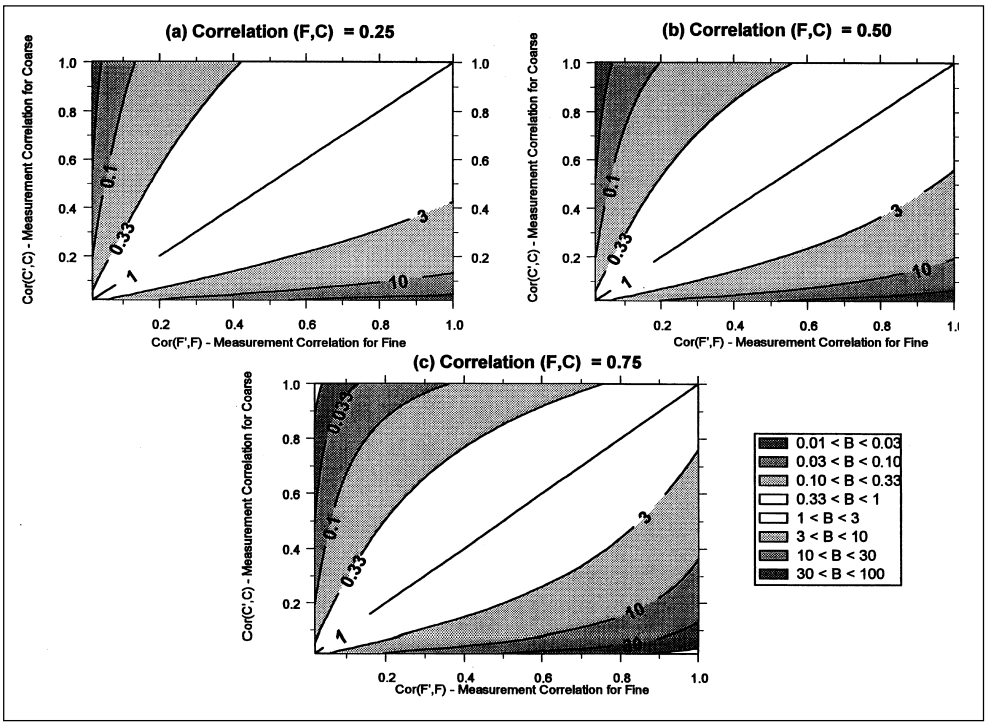

Figure 1. Bias contours for equally toxic pollutants (Bf/Bc = 1).

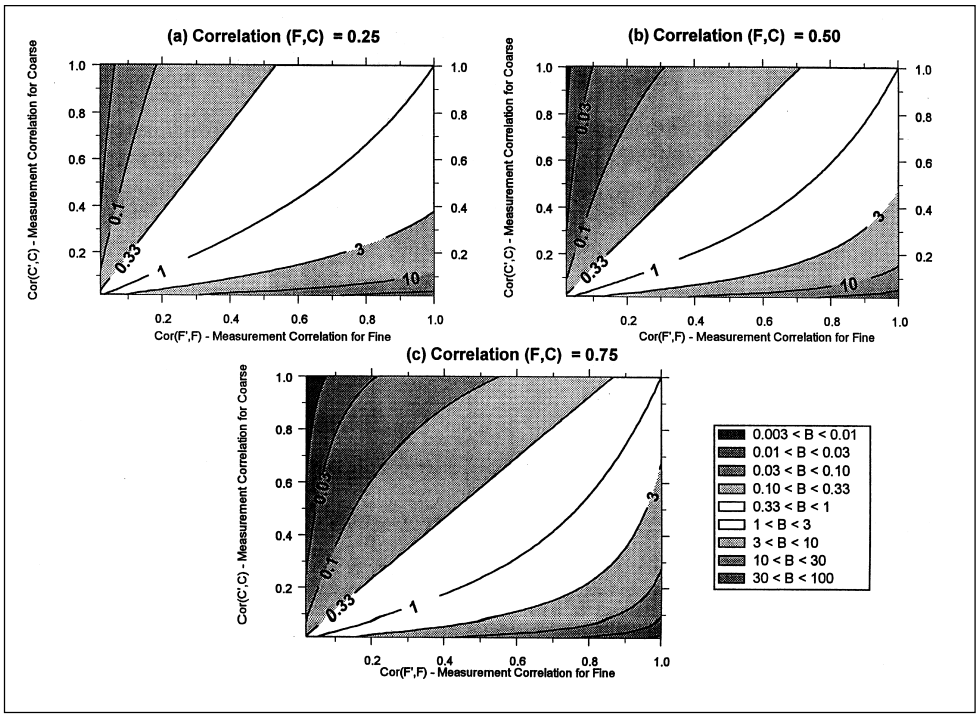

Figure 2. Bias contours for case of fine three-fold as toxic as coarse (Bf/Bc = 3).

The model also helps to show the implausibility of the extreme case (i.e., *F* is falsely implicated while *C* is falsely exonerated). Simplifying eqs 5 and 6 by substituting $ß_F = 0$, it is apparent that in order for *F*, harmless in this case, to be found merely as *equally harmful* as *C*, Cov(F',C) must equal Cov(C',C). In other words, *F'* must measure *C* as well as *C'* measures *C*. This requires Cov(C,F) to be greater than Cov(C,C') by a factor of 1/Cov(F',F). In order for *F* to be found much more harmful than *C* (the extreme case), *F'* would have to measure *C much better* than *C'* measures *C*.

CASE STUDY—HARVARD SIX-CITIES STUDY

While the theoretical results could be explored in more detail, the true usefulness of this model comes with its application to actual data sets. Below, the model developed above is applied to two cities from the "time-series" portion of the Harvard Six-Cities Study (SCS).[26] This study was chosen for several reasons. First, it was the first large-scale "time-series" study of the association of ambient particle concentrations with increased daily mortality that used both measurements of both ambient fine ($PM_{2.5}$) mass and coarse ($PM_{10-2.5}$) mass. Second, it was integral to the evidence cited by EPA in support of the introduction of a $PM_{2.5}$ standard.[33] Third, the study was specifically mentioned in previous qualitative discussions of measurement error and correlation.[23]

This case study analyzes Boston, MA, a city with an observed fine:coarse coefficient ratio of 11 to 1, and Knoxville, TN, a city with an observed fine:coarse coefficient ratio of 1.4 to 1. In the first step, the parameters needed for the model are derived from the published SCS literature. In the second step, these parameters are used in the model to assess the possibility that the published coefficients are misleading due to correlation and differences in measurement error. In the third step, a sensitivity analysis is conducted using a range of measurement correlations.

Estimation of Parameters

Since exposure studies do not directly report the parameters used in our simulation model, the information reported must be translated into the correlation parameters used by the model. Information on the fine and coarse regression coefficients, as well as the correlation between fine and coarse measurements, were reported in the SCS "time-series" publication.[26] Unfortunately, there was little attention devoted to the spatial variability of fine and coarse particles in the SCS.[34] Consequently, information on spatial variability from the Metropolitan Aerosol Acidity Characterization Study (MAACS),[27,35] a study of four large East Coast cities, is used to estimate spatial variability. Since the MAACS and SCS only overlap for one city (Boston), a range of spatial variabilities for fine and coarse particles is analyzed for Knoxville in the sensitivity analysis.

Table 1 presents the estimates of the simulation input parameters for Boston and Knoxville. The appendix presents a detailed description of how these are estimated from reported values of instrumental error, spatial variability, and measured central site concentrations.

Estimation of Expected Bias

The first six rows in Table 1 present the information available. The first three rows contain the estimated fine particle measurement, coarse particle measurement, and true fine particle:coarse particle correlations. The *observed* fine:coarse ratio is based on the observed coefficients in the SCS.[26] Since the model yields an expected *observed* fine:coarse ratio, given a true fine:coarse ratio and the three correlations, eqs 5–7 are used to calculate the *true* fine coarse ratio in each city based on the information available. The last row of Table 1 presents these calculated true fine:coarse ratios for Boston and Knoxville.

The calculated true fine: coarse ratio for Boston is 26:1, whereas the calculated true fine:coarse ratio for Knoxville is 0.9:1. In the case of Boston, the very high actual observed fine:coarse ratio, combined with the given measurement correlations, leads to the conclusion that the true ratio is even higher than the observed ratio. For Boston, this provides evidence against the hypothesis that the coefficient of the better measured pollutant is absorbing some of the coefficient of the other pollutant. For Knoxville, however, the calculated true fine:coarse ratio is markedly less than the actual observed fine:coarse ratio. This indicates that the better-measured pollutant, fine particles, may be overestimated relative to coarse particles due to the interplay of measurement error and correlation.

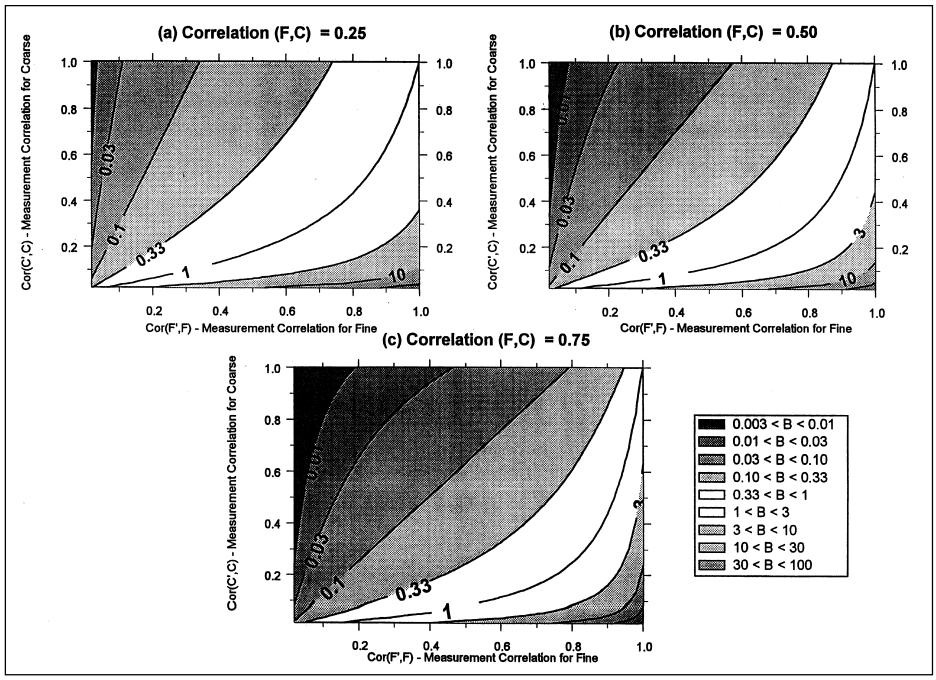

Figure 3. Bias contours for case of fine ten-fold as toxic as coarse (Bf/Bc = 10).

Sensitivity Analysis for Case Study

Since a proxy was used for spatial variability in Knoxville, it is important to consider a broader range of values than the central estimates used above. As shown in the appendix, the MAACS data translated to a range of measurement correlations between 0.85 and 0.95 for fine particles and 0.65 to 0.85 for coarse particles. The sensitivity analysis explores two cases: (1) smallest plausible difference in measurement error (i.e., Corr(F',F) = Corr(C',C) = 0.85) and (2) largest plausible difference in measurement error (i.e., Corr(F',F) = 0.95, Corr(C',C) = 0.65). These cases yield calculated true ratios of 1.65 and 0.48, respectively. Therefore, it is likely that the observed ratio for Knoxville is an overestimate of the relative toxicity of fine versus coarse particles. While the best estimate indicates that the toxicities were relatively equal (0.9:1), it is possible that the coarse particles could be twice as toxic as the fine particles under the boundaries of the model's assumptions.

LIMITATIONS

While this approach appears promising in diagnosing bias for a bivariate regression, there are a number of reasons to use caution in application. Analysis of this type is limited by the quality of the central estimates of coefficients. If there are other significant causes of bias affecting the coefficients, this analysis may hold little quantitative value. However, one can still qualitatively predict the potential for bias from the issue addressed in this paper simply from estimating the correlations. If the fine-coarse true correlation is low, or if the difference in the measurement correlations is small, there is less chance for bias in the observed coefficients due to the issue discussed herein.

In the case study, data from the MAACS study are used to estimate the bias in the SCS data set. Ideally, one should have data relevant to the time-period and city at hand. Rough estimates from other cities may lead to spurious results. The sensitivity analysis showed the importance of finding the correct estimate of measurement error for each pollutant. This sensitivity increases as the true correlation of the variables increases.

This case study assumes that exposures to other pollutants (i.e., gaseous pollutants, indoor-generated particles, etc.) are not correlated with exposures to ambient particles. If there are other covariates, the analysis must be expanded

Table 1. Case study parameters and results.

	Notation	City: Boston	City: Knoxville
Fine Particle Measurement Correlation	Cor(F', F)	0.85	0.90
Coarse Particle Measurement Correlation	Cor(C', C)	0.65	0.75
Estimated True Fine:Coarse Correlation	Cor(F,C)	0.28	0.54
Observed Fine Coefficient	$(\beta_F)_{OBS}$	2.2	1.4
Observed Coarse Coefficient	$(\beta_C)_{OBS}$	0.2	1.0
Observed Fine:Coarse Coefficient Ratio	$(\beta_F/\beta_C)_{OBS}$	**11**	**1.4**
Calculated True Fine Coefficient	(β_F)	2.6	1.23
Calculated True Coarse Coefficient	(β_C)	0.1	1.35
Calculated True Fine:Coarse Coefficient Ratio	(β_F/β_C)	**26**	**0.9**

to a multivariate setting. Although the computational requirements will rise substantially, the multivariate solution will be of the same general form as the bivariate solution shown in eqs 5 and 6.

The model assumed that the measurement errors for F and C are uncorrelated. This was based on the assumption that the spatial variability error is much larger than the instrumental error and the assumption that the day-to-day changes in the spatial pattern of F and C across the city are uncorrelated. If either of these assumptions fails, then the regression model should be adjusted to account for this additional correlation in the terms.

Finally, errors of the Berkson type can induce bias if the dose-response is non-linear. In the non-linear case, errors concerning the difference between the average individual and the ith individual cannot be ignored. This is a relatively unexplored area of research, and future work should specifically address errors in exposure assessment under non-linear dose-response.

CONCLUSIONS

This paper introduced a tool for predicting the amount of bias present in the ratio of fine and coarse particle regression coefficients due to the interaction of correlation and measurement error. The model showed that the amount of bias depended on several variables: the true correlation of fine and coarse particle exposures, the measurement errors for fine and coarse particle exposures, and the underlying true ratio of the fine particle toxicity to coarse particle toxicity. Analysis of the model proved that all of these variables must be discussed and analyzed before making any broad conclusions regarding bias. For instance, it is inadequate to state that differences in measurement error among fine and coarse particles will lead to false negative findings for coarse particles. If the underlying true ratio of the fine and coarse particle toxicities is large (i.e., greater than 3:1), fine particle exposure must be measured significantly more precisely in order not to *underestimate* the ratio of fine particle toxicity versus coarse particle toxicity.

Given the assumptions discussed previously, the case study applied the model to two cities from the SCS. The case study showed it was unlikely that coarse particle toxicity was underestimated relative to fine particle toxicity in Boston, where the observed $F:C$ coefficient ratio was 11:1. However, the case study also showed that it was possible that coarse particle toxicity was underestimated relative to fine particle toxicity in Knoxville. These results were subject to several limitations, in particular the lack of matching spatial variability data. Although the MAACS data set may make a reasonable proxy, possible conclusions across the sensitivity analysis for Knoxville varied widely. This result underscores the need for spatial variability data (i.e., several ambient monitors situated across a metropolitan area for a period of many months). Without these data, no definitive conclusions can be made regarding the possibility of bias due to differences in measurement error among correlated pollutants.

As ongoing studies in epidemiology and exposure assessment create a larger database of time-series analyses with concurrent measurements of the fine and coarse particle exposures, the quantitative tool developed in this paper will be useful in assessing the possibility of bias from the interplay of measurement error and correlation. Further work in quantitative methods should be able to better resolve the many other debates regarding the interpretation of air pollution epidemiology.

ACKNOWLEDGMENTS

Both authors were supported by the Harvard Center for Risk Analysis and the Department of Environmental Health, Harvard School for Public Health. Carrothers was also supported by a National Science Foundation Graduate Research Fellowship. The authors wish to thank David Mage, John Spengler, and the peer reviewers for their helpful comments.

REFERENCES

1. Ito, K.; Thurston, G.D.; Hayes, C.; Lippmann, M. *Arch. Env. Health* **1993**, *48*, 213-220.
2. Logan, W.P.D. *Lancet* **1953**, *1*, 336-338.
3. Vedal, S. *J. Air & Waste Manage. Assoc.* **1997**, *47*, 551-581.
4. Moolgavkar, S.H.; Luebeck, E.G. *Epidemiology* **1996**, *7*, 420-428.
5. Wilson, R.; Spengler, J. *Particles in Our Air: Concentrations and Health Effects*; Harvard University Press: Cambridge, MA, 1996.
6. Lipfert, F.W. *Science* **1997**, *278*, 19-20.
7. Lipfert, F.W.; Wyzga, R.E. *J. Air & Waste Manage. Assoc.* **1997**, *47*, 517-523.
8. Lipfert, F.W.; Wyzga, R.E. *J. Air & Waste Manage. Assoc.* **1995**, *45*, 949-966.
9. Gamble, J.F.; Lewis, R.J. *Environ. Health Perspect.* **1996**, *104*, 838-850.
10. *Air Quality Criteria for Particulate Matter*, Volume I; EPA/600/P-95/001aF; U.S. Environmental Protection Agency: Research Triangle Park, NC, 1996.
11. Hennekens, C.H.; Buring, J.E. *Epidemiology in Medicine*; Little, Brown, and Company: Boston, MA, 1987.
12. Thomas, D.; Stram, D.; Dwyer, J. *Annual Review of Public Health* **1993**, *14*, 69-93.
13. Armstrong, B.G. *Amer. J. Epi.* **1990**, *132*, 1176-1184.
14. Savitz, D.A.; Baron, A.E. *Amer. J. Epi.* **1989**, *129*, 1062-1071.
15. Navidi, W.; Lurmann, F. *J. Expos. Anal. Environ. Epidem.* **1995**. *5*, 111-124.
16. Casella, G.; Berger, R.L. *Statistical Inference*; Duxbury: Belmont, CA, 1990.
17. Lipfert, F.W.; Wyzga, R.E. *Inhalation Toxicology* **1995**, 7, 671-689.
18. Cochran, W.C. *Technometrics* **1968**, 10, 637-644.
19. Fuller, W.A. *Measurement Error Models*; John Wiley & Sons: New York, 1987.
20. Yoshimura, I. *Environ. Health Perspect.* **1990**, 87, 173-178.
21. Klepper, S.; Leamer, E.E. *Econometrics* **1984**, 52, 163-183.
22. White, W.H. *J. Air & Waste Manage. Assoc.* **1998**, *48*, 454-458.
23. Smith, A.E. Written statement of testimony before *U.S. Senate Subcommittee on Clean Air, Wetlands, Private Property, and Nuclear Safety*; February 5, 1997; U.S. Government Printing Office: Washington, DC.
24. Samet, J.M.; Zeger, S.L.; Kelsall, J.E.; Xu, J.; Kalkstein, L.S. The Phase I.B Report of the Particle Epidemiology Evaluation Project; Health Effects Institute: Boston, MA, 1997; p 26.
25. Özkaynak, H; Spengler, J.D. *Particles in Our Air, Concentrations and Health Effects*, ch. 4; Harvard University Press: Cambridge, MA, 1996.
26. Schwartz, J.; Dockery, D.W.; Neas, L.M. *J. Air & Waste Manage. Assoc.* **1996**, 46, 927-939.
27. Wilson, W.E.; Suh, H.H. *J. Air & Waste Manage. Assoc.* **1997**, 47, 1238-1249.
28. Mage, D.T.; Wilson, W.E.; Hasselblad, V.; Grant, L. *J. Air & Waste Manage. Assoc.* **1999**, *49*, 1280-1291.

29. Schwartz, J. Boston, MA. Personal communication, August 1998.
30. Alternatively, one can replace the population-mean exposure x_t with the baseline risk-weighted population mean and proceed with essentially the same argument. This allows for the possibility that personal exposures are influenced by activities correlated with the underlying risk of death. This would not be expected to significantly affect the underlying argument.
31. Özkaynak, H.; Xue, J.; Spengler, J.D.; Wallace, L.; Pellizzari, E.; Jenkins, P. *J. Exposure Anal. Environ. Epidem.* **1996**, 1, 57-78.
32. Evans, J.E.; Ryan, P.B. *Aerosol Sci. & Technol.* **1983**, 2, 531-536.
33. *Review of the National Ambient Air Quality Standards for Particulate Matter*; OAQPS Staff Paper; EPA/452/P-96/013; U.S. Environmental Protection Agency: Research Triangle Park, NC, 1996.
34. Spengler, J. Boston, MA. Personal communication, October 1998.
35. Bahadori, T. Sc. D. Thesis, Harvard School of Public Health, Boston, MA, 1997.
36. Duan, N. *Envir. Int.* **1982,** 8, 305-309.
37. Mathsoft. *S-Plus 4 User's Guide;* Data Analysis Producs Division, Seattle, WA, 1997.
38. California Air Resources Board. *Activity Patterns of California Residents;* A6-177-33; Sacramento, CA, 1991.
39. Koontz, M.D. *Estimation of Distributions for Residential Air Exchange Rates: Final Report*; U.S. Environmental Protection Agency, Office of Pollution Prevention and Toxics: Washington, DC, 1995.

APPENDIX

Estimating True *F-C* Correlation

The first step in estimating parameters for the case study is to estimate the true *F-C* daily correlation. The SCS measured both fine and coarse particles and reported correlations between these daily averaged measurements. However, since measurements of both types of particles involve some amount of instrumental error, the true correlation between true concentrations is actually somewhat higher. If we assume that the size of instrumental error is not strongly dependent on the overall mass of the measurement, we can estimate the true correlation via the method outlined below.

The Relation of the Additive Error Model to Bivariate Normality. First, since instrumental errors are often reported in an additive fashion (i.e., "error was estimated at ± 2 μg/m^3"), we must transform these numbers into correlations for use in our simulation model. The usual model for measurement error is an additive model: Z = X + E. If *X* and *E* are independent normal variables, then

$$\mathrm{var}(Z) = \mathrm{var}(X) + \mathrm{var}(E) \qquad \text{(A-1)}$$

and

$$\rho = corr(X,Z) = \left[\sqrt{\left(1+\frac{\mathrm{var}(E)}{\mathrm{var}(X)}\right)}\right]^{-1}. \qquad \text{(A-2)}$$

Conversely, if the variance of *Z* and the correlation ρ are known, one can calculate the variances of *X* and *E* using eq 7 and

$$\mathrm{var}(E) = \mathrm{var}(X) * \left(\frac{1}{\rho^2} - 1\right) \qquad \text{(A-3)}$$

(Note that if random variables *X* and *Y* have a bivariate normal distribution with correlation ρ, location and scale transforms will not affect ρ. This allows us to construct our model with unit normal variables without loss of generality.)

Calculation of True F-C *Correlation.* This example uses the Philadelphia MAACS results. For coarse particles, instrumental error was ~ 2 μg/m^3 and the standard deviation of the samples was ~ 4.5 μg/m^3.[35] This would correspond to var(Z) = 20, var(E) = 4, and var(X) = 16 by subtraction. By (eq A-2), ρ_{inst} = 0.89, which is the correlation of daily 24-hr average coarse measurements with the true 24-hr average concentration. Similarly, with fine particles, we find ρ_{inst} = 0.92, with instrumental error ~ 4 μg/m^3 and the standard deviation of the samples ~ 10 μg/m^3.

With this estimate, we back-calculate an expected true *F-C* correlation, undiluted by instrumental error. Under our assumptions of bivariate normality, if the reported correlation between two measured quantities is denoted as Corr(X', Y'), then the correlation between the true quantities is

$$\mathrm{corr}(X,Y) = \frac{\mathrm{corr}(X',Y')}{\mathrm{corr}(X,X') * \mathrm{corr}(Y,Y')} \qquad \text{(A-4)}$$

This implies that a measured correlation between fine and coarse particles of 0.29 should be adjusted upward to 0.29 / (0.92*0.89) = 0.35. This is likely a small adjustment relative to the overall uncertainties involved, but making no adjustment would incorrectly assume that no instrumental error was present.

Estimating Correlations of Central Site and Average Individual Exposure

The first step in the case study determined the true extent of correlation between the two co-pollutants. In the next step, we determine the amount of error present due to use of the central site monitor. Here, we find the correlation between the central site data and the true average individual exposure to each pollutant. Equation 1 implies that reported spatial variability and instrumental error are the only two pieces of information necessary to know this correlation. To illustrate this result, we conducted a simulation via a microenvironmental model of human exposure within a city.

The Microenvironmental Model. A commonly used model for estimating personal exposure is the microenvironmental model.[36] Average individual exposure to ambient particles is primarily composed of indoor exposures (at home and work/school) to ambient particles that have penetrated indoors and remained suspended. Therefore, a simple version of a microenvironmental model for an individual would be

$$E_{tot} = t_h f_h A_h + t_w f_w A_w + t_o A_o \qquad \text{(A-5)}$$

where A_k represents the ambient concentrations in the vicinity of the *k*th microenvironment, t_k is the time spent

in the kth microenvironment, and f_k is the penetration fraction. All of these are variable quantities that can be represented as probabilistic parameters in a simulation study using S-PLUS.[37] A similar study, focused on ozone, has been conducted by Navidi and Lurmann.[15]

Data Inputs and Procedure for Simulation. In this simulation, we created a hypothetical city divided into six sections, where each section had an ambient monitor. We chose the number six to be consistent with the average number of monitors in the MAACS study cities. We then generated a 300-day series of concentrations for each sector, with concentrations in each sector correlated by an average of ρ with the other sectors. Next, we randomly assigned 1000 people to a home and office location and computed their exposure using empirical distributions for time-activity patterns and building penetration rates.[38,39] Then, for each day, we found the average individual exposure. This allowed us to compute the correlation of the average individual exposure with each of the monitors, any of which could be the "central-site" monitor.

Results of Simulation. This simulation was performed for several different values of ρ, the assumed true spatial correlation. After performing this simulation for 1000 runs, it was apparent that the expected correlation was simply the square root of ρ, the true spatial variability. This finding agreed with eq 1, as predicted.

Estimated Correlations for SCS Analysis. Unfortunately, the SCS had limited spatial variability data for fine particles and none for coarse particles, so we had to use the spatial variability reported in the MAACS study as a rough substitute. The average fine particle spatial variability in these four cities ranged from 0.75 in Nashville and Boston to 0.85 in Washington and Philadelphia. The average coarse particle spatial variability ranged from 0.40 in Boston to 0.67 in Philadelphia and Washington. Again, we must adjust these estimates upward to account for expected instrumental error. After adjusting for instrumental error and taking the square root of the result, we estimated the correlation between the central site and average individual to be in the range of 0.85 to 0.95 for fine particles and 0.65 to 0.85 for coarse particles. Since Boston is in both studies, we can use its MAACS result directly.

About the Authors

Timothy J. Carrothers is a doctoral candidate in the Harvard School of Public Health Environmental Science and Risk Management Program in Boston, MA (665 Huntington Avenue, Suite 1308-B, Boston, MA 02115). John S. Evans is a senior lecturer and director of the Harvard School of Public Health Environmental Science and Risk Management Program (718 Huntington Avenue, Boston, MA 02115).