The Spatial Association between Community Air Pollution and Mortality: A New Method of Analyzing Correlated Geographic Cohort Data

Richard Burnett,[1,2,7] Renjun Ma,[2] Michael Jerrett,[3] Mark S. Goldberg,[4,5] Sabit Cakmak,[1] C. Arden Pope III,[6] and Daniel Krewski[2,7]

[1]Healthy Environments and Consumer Safety Branch, Health Canada, Tunney's Pasture, Ontario, Ottawa, Canada; [2]Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ontario, Ottawa, Canada; [3]School of Geography and Geology and Institute of Environment and Health, McMaster University, Hamilton, Ontario, Canada; [4]Department of Medicine, McGill University, Montreal, Quebec, Canada; [5]Joint Departments of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada; [6]Economics Department, Brigham Young University, Provo, Utah, USA; [7]Institute of Population Health, University of Ottawa, Ottawa, Ontario, Canada

We present a new statistical model for linking spatial variation in ambient air pollution to mortality. The model incorporates risk factors measured at the individual level, such as smoking, and at the spatial level, such as air pollution. We demonstrate that the spatial autocorrelation in community mortality rates, an indication of not fully characterizing potentially confounding risk factors to the air pollution–mortality association, can be accounted for through the inclusion of location in the model assessing the effects of air pollution on mortality. Our methods are illustrated with an analysis of the American Cancer Society cohort to determine whether all cause mortality is associated with concentrations of sulfate particles. The relative risk associated with a 4.2 µg/m^3 interquartile range of sulfate distribution for all causes of death was 1.051 (95% confidence interval 1.036–1.066) based on the Cox proportional hazards survival model, assuming subjects were statistically independent. Inclusion of community-based random effects yielded a relative risk of 1.055 (1.033, 1.077), which represented a doubling in the residual variance compared to that estimated by the Cox model. Residuals from the random-effects model displayed strong evidence of spatial autocorrelation ($p = 0.0052$). Further inclusion of a location surface reduced the sulfate relative risk and the evidence for autocorrelation as the complexity of the location surface increased, with a range in relative risks of 1.055–1.035. We conclude that these data display both extravariation and spatial autocorrelation, characteristics not captured by the Cox survival model. Failure to account for extravariation and spatial autocorrelation can lead to an understatement of the uncertainty of the air pollution association with mortality. *Key words:* air pollution, cohort, epidemiology, mortality, spatial regression, sulfate particles, survival. — *Environ Health Perspect* 109(suppl 3):375–380 (2001). http://ehpnet1.niehs.nih.gov/docs/2001/suppl-3/375-380burnett/abstract.html

In 1997, the U.S. Environmental Protection Agency (U.S. EPA) promulgated new regulations for fine particulate matter in ambient air. This decision was based, in part, on the evidence that Americans had an increased risk of cardiopulmonary mortality if they lived in areas with elevated ambient fine particles compared to individuals who resided in less-polluted areas. Two of the key studies considered by the U.S. EPA in this regard were those of Dockery and colleagues (*1*), who used data from the Harvard Six Cities Study, and Pope and colleagues (*2*) who used data obtained from the American Cancer Society (ACS) Cancer Prevention II Study (*3*). A number of criticisms of these two studies (*4–6*) have been largely addressed in an extensive reanalysis (*7*) conducted at the request of the Health Effects Institute in Cambridge, Massachusetts, USA.

In both of these cohort studies (*1,2*), subjects were enrolled from communities with various levels of outdoor air pollution. Subject-specific information on factors such as age, gender, race, tobacco use, alcohol consumption, occupational exposures, and education were collected by the use of an interview and questionnaire. Subjects were followed over time to assess changes in their vital status. Air pollution was measured by fixed-site monitors either prior to enrollment or during follow-up or both. The standard Cox proportional hazards regression survival model (*8*) was used to assess associations between mortality rates and community-based average ambient air pollution while controlling for individual risk factors such as age, gender, race, tobacco use, education, cigarette smoking, body mass index, and occupational exposures. In both of these studies (*1,2*), statistically significant associations were found between mortality rates and particulate air pollution, as measured by fine and sulfate particulate concentrations.

The standard Cox proportional hazard model used in these two studies to relate longevity to exposure assumed that event information (time of death or censoring due to end of study or loss to follow-up) was statistically independent among subjects after controlling for available information on subject-specific mortality risk factors. Such an approach results in at least two somewhat related concerns. First, health responses can cluster by location (*9*). Clustering will induce a positive correlation of the response of subjects in the same location and thus suggests that there are one or more unmeasured or inadequately modeled risk factors specific to the location itself. Failure to account for this clustering can lead to an understatement of the uncertainty in these estimates (*10,11*).

Second, responses of subjects living in communities close together may be more similar than responses of subjects living in cities farther apart after controlling for subject-specific risk factors. Failure to account for this type of spatial autocorrelation can also lead to an understatement of the uncertainty of the effect estimates (*12,13*). Furthermore, if this spatial autocorrelation is due to missing or systematically mismeasured risk factors also spatially autocorrelated, then the estimates of the effect of air pollution on mortality could be biased. The direction and size of the bias will depend upon the direction and degree of correlation between the missing risk factors, air pollution, and mortality.

In this article we present a new statistical approach to deal with these two related methodologic concerns. We present a spatial random-effects survival model that links spatial variation in concentrations of ambient air pollution to longevity of cohort subjects after controlling for temporal effects and individual risk factors for mortality. We used data from

the original ACS study (*2*) to demonstrate the impact of modeling random location effects and spatial autocorrelation on the estimated air pollution–mortality association and estimates of uncertainty. These results are compared with those obtained using standard methods of survival analysis assuming statistical independence among subjects.

Data and Methods

In this section we review the data used for our analysis, and we present the statistical model used to assess the association between air pollution and mortality.

The American Cancer Society Study of Air Pollution and Mortality

Volunteers of the ACS enrolled over 1.2 million people in September 1982 throughout the United States. Information on history of disease, demographic characteristics, and mortality risk factors was obtained from respondents. Vital status was monitored through the end of 1989.

We obtained information on particulate sulfate levels from the Aerometric Information Retrieval System (AIRS) (*http://www.epa.gov/ air/data/info.html*) and the Inhalable Particle Network (IPN) for 1980 and 1981 for 144 metropolitan statistical areas (MSAs) in which ACS subjects were enrolled. Sulfates are secondarily formed particulate aerosols originating from sulfur dioxide emissions and are a major component of fine particulate matter. The sulfate data from AIRS were collected using glass-fiber filters, which react in the presence of sulfur dioxide and artifactually inflate the sulfate concentration. The sulfate data obtained from the IPN used Teflon filters, which are not subject to this artifact problem. Both monitoring networks were operating in 41 MSAs. We calibrated the AIRS sulfate data to the IPN sulfate data using six linear regression models, with separate calibrations for three regions of the county and two periods (April–September and October–March) (*7*). We used six calibration equations because sulfur dioxide concentrations vary both regionally and seasonally in the United States. Estimates of exposure were obtained by averaging all available sulfate data from all monitors located in an MSA for the years 1980 and 1981, inclusive.

We examined the association between concentrations of sulfate particles and longevity in 144 MSAs for white members of the ACS cohort, totaling 509,292 subjects. The mean age at enrollment was 56.7 years. Five percent of the subjects were younger than 40 years of age and 5% were older than 75 years of age; 56.3% of the subjects were women. During the course of the 7 years of follow-up, 39,474 subjects (7.8%) died. The mean concentration of sulfate particles across all 144 cities, corrected for the sulfur dioxide artifact, was 6.4 μg/m³, with a minimum value of 1.4 μg/m³, an interquartile range of 4.2 μg/m³, and a maximum value of 15.6 μg/m³.

Statistical Model

The model is formulated in two stages. In stage one, survival data were modeled by covariates at the individual level and indicator functions for each community, using the Cox proportional hazards model (*8*). The community-specific indicator functions represent the logarithm of the relative risk of death in a specific community compared to an arbitrarily defined reference community. Sulfate pollution was not induced at this stage.

Indicator functions for community are defined with respect to a reference community (in our case Greenville, South Carolina, as it had a sulfate concentration near the mean value). A limitation of this procedure is that the uncertainty of the estimate of the reference community is not defined. Because these values are based on comparisons with the same reference community, they are correlated. This correlation increases the estimated uncertainty in the community-specific log-relative risks. The induced correlation can be removed by methods developed by Easton and colleagues (*14*). This procedure eliminates the included correlation between the estimates of the community-specific log-relative risks and estimates the uncertainty for the reference community. Outputs from stage one are estimates of the community-specific logarithm of the relative risks adjusted for mortality risk factors other than sulfate pollution, denoted by $\{\delta(s), s = 1,...,S\}$, where s denotes a point in Cartesian (x,y) space representing the location of one of the S communities under study. Additional output from this stage is the variance–covariance matrix of the $\delta(s)$, denoted by V, which describes the uncertainty in the adjusted estimates of the community-specific log-relative risks.

In stage two, estimates of adjusted community-specific log-relative risks were related to levels of sulfate pollution levels using a linear random-effects regression model (*15*). We also included a two-dimensional term to account for spatial trends, denoted by $\Im$, and assigned a random effect to each community. These spatial random effects are shared by all individuals within the community and reflect the difference between the observed and predicted values from our statistical model. We assume that the random effects have zero expectation, variance $\theta > 0$, and correlation matrix Ω. Residual variation at the spatial level suggests that there is some unexplained (unmeasured or incorrectly specified) risk factors for mortality.

Evidence of spatial autocorrelation in the residuals of the model may indicate the need to account for additional risk factors, which may potentially exert a confounding effect on the air pollution–mortality association. An alternate approach to accounting for autocorrelation by modeling additional risk factor information is to filter out spatial contiguous variation by including a term that represents spatial trends $\{\Im(s)\}$. The total impact of these potentially numerous unmeasured risk factors may vary in a relatively smooth manner over space, and thus spatial detrending can remove autocorrelation between geographic areas. In this approach, location and other covariates such as air pollution, which also vary in space, compete in the regression model to predict mortality. Thus, the regression coefficients give the effect of these variables adjusted for each other. This approach is analogous to that used in time–series studies of mortality and air pollution in which temporal trends in daily mortality rates are jointly modeled with air pollution levels (*16*).

Here, $\delta(s)$ has expectation

$$\mu(s) = \Im(s) + \beta^T Z(s)$$

and variance–covariance matrix

$$\Sigma = \theta \, \Omega + V,$$

where $Z(s)$ is a vector of covariates defined at the community level. In our example we restricted the set of these spatial covariates to sulfate pollution. If the number of subjects and deaths in each community is large, as is assumed here, the $\{\delta(s)\}$ have approximately a multivariate normal distribution with mean vector μ and variance–covariance matrix Σ. The unknown parameter vector, β, and the variance of the random effects, θ, can be estimated by maximum likelihood methods (*15*) with the log-likelihood function minimized by a Fisher's scoring algorithm (*15*). The elements of V are assumed to be known in this stage. We considered nonparametric smoothed estimates of $\Im$ using the robust locally weighted regression (LOESS) smoother (*17*) within the generalized additive model (GAM) framework (*18*). The complexity of the location surface was controlled by a span parameter, which is the proportion of the data used for the local regression. Increasing the span increases the smoothness of the estimated surface. Estimates of β are interpreted as the logarithm of the relative risk of mortality for a unit change in sulfate pollution.

The approach described above yields unbiased and fully efficient estimates of the unknown parameters within a generalized estimating equation framework (*19*) if there is in fact no spatial autocorrelation in the random effects. We have developed a simple method to judiciously select the appropriate span in the LOESS smoother in order to

minimize the autocorrelation structure of the random effects. We do this by plotting a smoothed representation of the correlation of the standardized residuals versus the distance between communities. The span is decreased until there is no apparent association between the correlation of the standardized residuals and distance. We also tested the standardized residuals for white noise using Bartlett's test (20). Larger p-values based on this test indicate less statistical evidence of a pattern in the residuals. The computer program to implement the estimation procedure was written in SAS (21) and in S-PLUS (22).

We used the following method to visualize the spatial distribution of mortality and sulfate particles. We first generated sulfate and mortality surfaces using the GAM option in S-PLUS, then exported the grid of predicted values into ARC/INFO (23) and ArcView (24) for cartographic presentation. For sulfate concentrations we regressed the readings from fixed site monitors in the 144 cities where data were available on the geographic (x,y) coordinates for each city using a LOESS-smoothed representation of the location surface. We experimented with spans ranging from 25 to 50% of the data. A 40% span was used in our final model because it illustrated regional variability without producing too many local ridges. Predicted values for this model were created using a 10-km^2 grid. The matrix of predicted values was then exported into ARC/INFO 7.2.1, where a new coverage was created using the GRID module. We resampled the surface at a 2.5-km grid cell resolution to smooth the contours for final presentation. After resampling, the grid coverage was exported into ArcView 3.2, where the surface was classified using eight contours representing 1-μg intervals (Figure 1). The highest and lowest categories were classified at larger intervals because of the relatively small geographic areas covered by the extreme portions of the range.

We also regressed the community-specific–adjusted mortality log-relative risks $\{\delta(s)\}$ onto the (x,y) coordinates of the 144 MSAs with a nonparametric smoothed location surface, $\Im$, excluding spatial covariate information such as air pollution (i.e., $Z(s) = 0$) using our spatial random-effects model. A LOESS span of 35% produced a surface that depicted regional variability without too much local ridging. Generation of grid values and the cartographic preparation followed the same procedure as with the sulfate surface. We mapped the exponents of the $\{\delta(s)\}$ to depict the relative risk of death in each community compared to the reference community of Greenville. Classification here was based on intervals of 0.024 to show subtle differences in the spatial distribution of mortality (Figure 2).

Results

The first step in the analysis was to use the Cox survival regression modeling procedure in SAS to identify all relevant individual risk factors associated with mortality. This model also included indicator functions for the 144 communities. The baseline hazard function was stratified by 5-year age groups and gender. The following risk factors were selected: quadratic polynomial of the number of cigarettes smoked and years smoking for current and former smokers; age starting smoking for current and former smokers; consumption of beer, wine, and liquor; quadratic polynomial of body mass index (square of height divided by weight); educational attainment (less than high school, high school, or more than high school); marital status (single, married, other); passive exposure to tobacco smoke; and regular exposure to some air toxics in work or daily
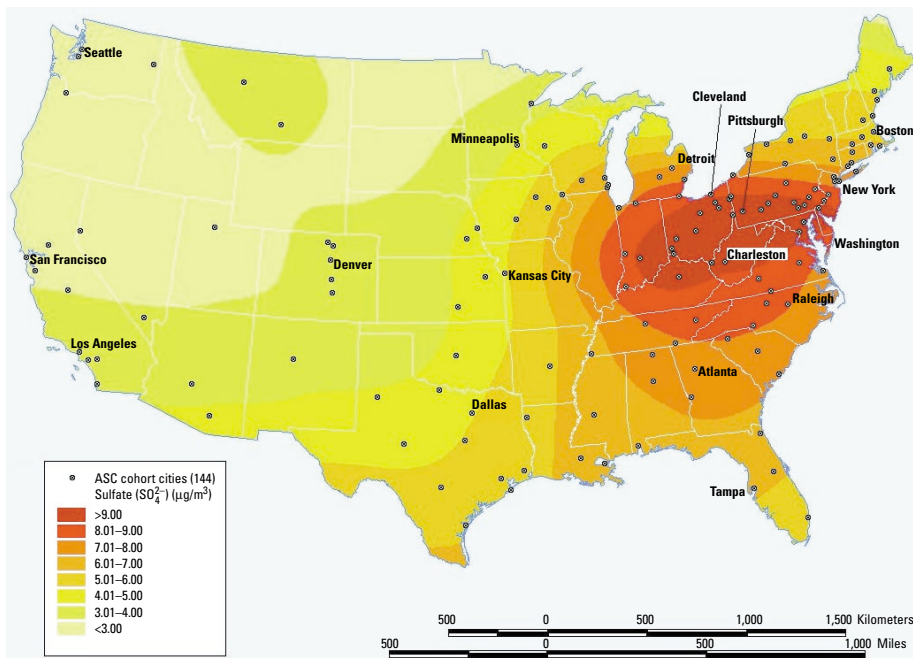

Figure 1. Spatial distribution of sulfate concentrations based on 1980–1981 average values in 144 metropolitan statistical areas. See text for details on construction of map.
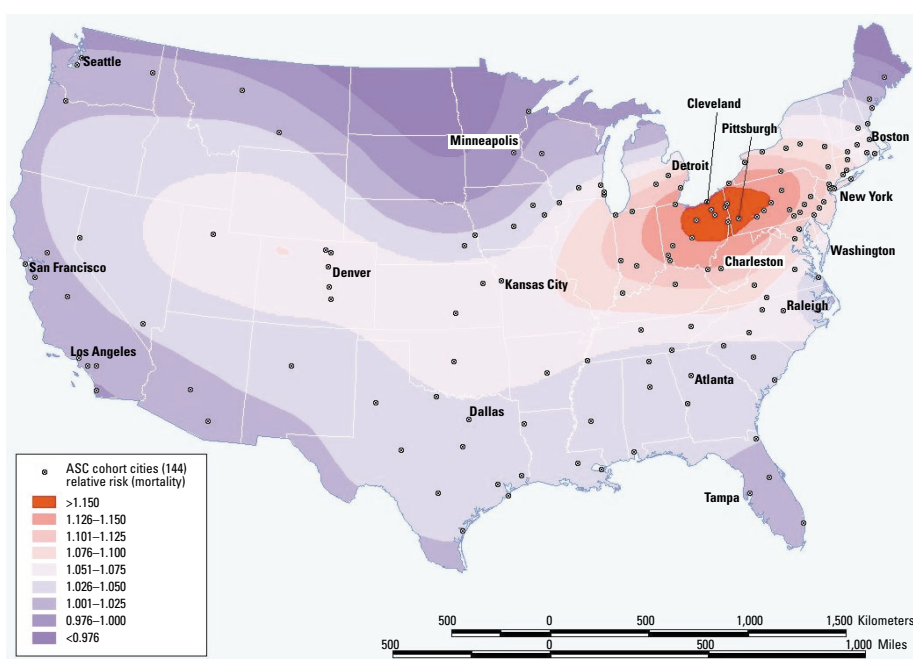

Figure 2. Spatial distribution of 144 metropolitan statistical area–specific relative risks of mortality compared to Greenville, South Carolina, adjusted for individual risk factors. See details in text for construction of map.

life (asbestos, chemicals/acids/solvents, coal or stone dusts, coal tar/pitch/asphalt, diesel engine exhaust, or formaldehyde).

Adjusted relative risks for mortality were elevated in a wide region of Lake Erie from the Lower Ohio River Valley to southern New England (Figure 2). Lower risks were observed in northern border states, west of the Great Lakes, in the Deep South, Southwest, and along the West Coast. Moderately elevated rates were present in the middle portion of the country from the Middle Mississippi Valley to the Great Basin. There is a corresponding elevation in concentrations of sulfate particles in the Ohio Valley region, with much lower concentrations west of the Mississippi (Figure 1). However, sulfates were also elevated in the southern portion of the eastern megalopolis conurbation from New York City to Norfolk, Virginia, and to a lesser extent in the Deep South, forming a different pattern than that found with mortality.

We then fit our spatial random-effects model, with no spatial predictors, and determined the standardized residuals from this model. A smoothed representation of the association between the autocorrelation of these standardized residuals and distance is presented graphically in Figure 3A. Autocorrelation declines with distance for communities within 1,000 km and for communities 2,500–4,500 km apart. The inclusion of sulfate particulate matter into the model dampened but did not eliminate the autocorrelations (Figure 3B). Further inclusion of a nonparametrically estimated location surface using LOESS spans of 80, 60, 40, and 20% (Figure 3C–F, respectively) reduced the autocorrelation. There was no apparent autocorrelaton pattern with distance using a span of 20%. (Estimates of starting values for θ were negative for spans less than 20%, indicating the spatial surface was overfitting the data.)

The association between sulfate concentrations and mortality (β) decreased as the complexity of the location surface modeling increased, or the span of the LOESS smoother decreased (Table 1). The residual variation between mortality rates, θ, also decreased with increasing location surface modeling complexity, indicating that the location surface is explaining differences in mortality among communities. The statistical evidence that residuals from these models are not autocorrelated increased with increasing complexity of the location surface on the basis of the *p*-value for Bartlett's test (*20*). This result demonstrates that the location surface is removing structure in the residuals.

Strong statistical evidence supported the assumption of additional variation in mortality between communities (i.e., θ > 0) on the basis of the likelihood-ratio test comparing models 3 and 4 (*p* < 0.0001). Increasing the

complexity of the location surface (i.e., decreasing the span) was also justified in terms of improving the model's fit to the data (*p* < 0.0001) (likelihood-ratio tests comparing model 4 to models 5–8, respectively).

Our spatial random-effects model implicitly assumes that the association between sulfates and mortality is linear on the logarithmic scale. We visually examined this assumption by plotting a cubic-spline representation of the association for location

surface models with a span of 20% (Figure 4). The association between sulfates and mortality appears nearly linear, with little statistical evidence of a departure from linearity (*p* = 0.9620).

Discussion and Conclusions

In previous studies using longitudinal cohort designs, statistically significant associations between mortality and particulate air pollution as measured by fine or sulfate particles
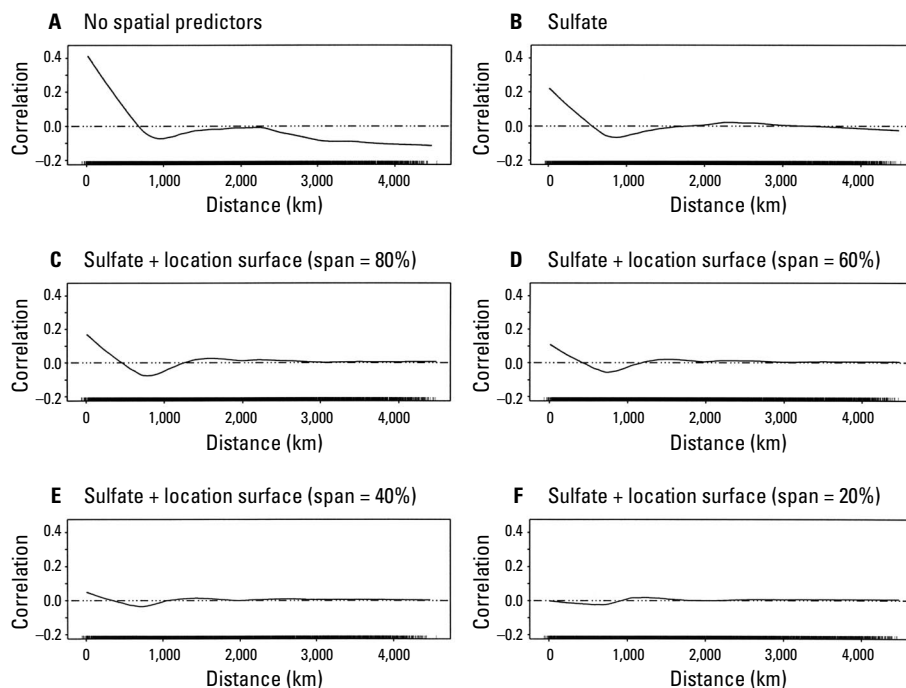
Figure 3. Smoothed representation using LOESS with a span of 50% of the association between the correlation of standardized residuals from our spatial random-effects regression model by distance between locations with no covariates (*A*), sulfate only (*B*), and sulfate and location with smoothing parameters of 80, 60, 40, and 20% (*C–F*, respectively). Horizontal line indicates zero values.

Table 1. Sulfate effect, β, by model type and span of LOESS smoother of the location surface.

Model type (model number)	Span (%)	Regression coefficient (β) (standard error)	Relative risk[a] (95% confidence interval)	Random-effects variance (θ)	−2lnL (df[b])	Bartlett's test *p*-value
Cox (1)	NA	0.0118 (0.00177)	1.051 (1.036, 1.066)	NA	NA	NA
Random-effects Cox (2)	NA	0.0125 (0.00252)	1.055 (1.033, 1.077)	0.0027	NA	NA
Spatial (3)	NA	NA	NA	0.0040	−492.52 (0)	
Spatial (4)	NA	0.0127 (0.00252)	1.055 (1.033, 1.077)	0.0027	−515.81 (0)	0.0052
Spatial (5)	80	0.0106 (0.00277)	1.046 (1.022, 1.070)	0.0021	−524.87 (3.2)	0.0297
Spatial (6)	60	0.0093 (0.00261)	1.040 (1.018, 1.062)	0.0016	−537.73 (5.5)	0.0559
Spatial (7)	40	0.0085 (0.00245)	1.036 (1.016, 1.058)	0.0010	−551.09 (8.9)	0.1450
Spatial (8)	20	0.0081 (0.00219)	1.035 (1.016, 1.053)	0.0003	−570.36 (18.5)	0.3272

NA, not applicable. [a]Relative risk evaluated at interquartile range of sulfate concentrations (4.2 μg/m³). [b]Approximate degrees of freedom of LOESS smoothed location surface.

have been observed (*1,2,25*). This article directly addresses two related concerns about these studies. First, the data were analyzed using the standard Cox proportional hazard survival model, with the implicit assumption that the observations were statistically independent after controlling for available risk factors measured on the individual (*8*). If the assumption of statistical independence is not valid, the uncertainty in the estimates of the association between air pollution and mortality may be understated (*9–13*). Second, missing or systematically mismeasured risk factors that may be correlated with air pollution could confound the pollution–mortality association (*4–6*).

Regarding the first concern, our spatial random-effects model provides more accurate estimates of the uncertainty of estimates of effect. On the basis of the analysis of the ACS data, although our model gave sulfate mortality estimates similar to those of the standard Cox model, the standard errors of these estimates were somewhat higher than those from the standard Cox model (Table 1), and the extra variation in the data was double that predicted by the Cox model.

With regard to the second concern, we observed a pattern of spatial autocorrelation in mortality that cannot be fully explained by ambient particulate sulfate concentrations, even after controlling for a host of risk factors measured at the individual level. We also found that the association between air pollution and mortality was somewhat sensitive to the specification of the complexity of the location surface, with more complex surface specifications resulting in lower estimates of the sulfate effect. These results suggest that there may be some confounding due to missing or systematically mismeasured risk factors that are also spatially correlated with pollution.

One approach to this potential confounding problem is to model additional spatially

Figure 4. Cubic-spline representation of the association between particulate sulfate concentrations (μg/m³) and community-specific log-relative risks for mortality adjusted for individual level covariates and a smoothed representation of location (LOESS span of 20%). Dashed lines represent 95% confidence intervals, with tick marks indicating sulfate values.

distributed (ecologic) risk factor data (*7*). However, the inclusion of available ecologic information may not be sufficient to account for all the spatial autocorrelation in the model residuals (*7*). We have shown that smoothed estimates of the location surface can remove residual spatial autocorrelation, circumventing the need to model ecologic covariates with the sole purpose of accounting for spatial autocorrelation.

We note, however, that ecologic risk factors could further explain residual variation in mortality between communities not captured by air pollution and the location surface and could potentially confound the air pollution-mortality association. Thus, removal of spatial autocorrelation does not guarantee that there are not additional risk factors that are potential predictors of mortality or confounders to the air pollution–mortality association.

We did not use our spatial random-effects survival model on the six cities data (*1*) because of the limited number of locations sampled. On the basis of the ACS data, positive autocorrelation was limited to communities less than 750 km apart. The six communities examined in the Six Cities Study were roughly this distance apart. There would likely be little need to consider spatial autocorrelation in this study if their spatial mortality pattern were similar to that observed in the ACS cohort.

This article represents an extension of the modeling approach used in our reanalysis (*7*). There, spatial autocorrelation was modeled by including regional indicator variables in the random-effects model. This procedure was ad hoc in that there was no unique method for defining regions. Our new approach of using location surface models allows the data to determine the form and extent of the spatial adjustment. We also removed spatial autocorrelation by prefiltering both the $\delta(s)$ and the sulfate data for spatial trends using a spatial moving average function. The radius of data inclusion for the moving average term was based on the distance beyond which no evidence of spatial autocorrelation was graphically apparent. However, this approach does not ensure that all the spatial autocorrelation was removed in the model residuals for all distances less than the selected value. Furthermore, the number of communities comprising the moving spatial average varies in space, yielding spatially filtered estimates with varying uncertainty. Finally, all evidence of associations between air pollution and mortality at the regional scale are removed using the prefiltering approach. Our new modeling method allows air pollution to compete with location to predict mortality. Evidence of regional-scale associations between mortality and air pollution will be captured with this new approach.

The observed association may be attenuated because measures of air pollution are known to misrepresent personal exposure and may not represent the average of personal exposure for all cohort members within a community. In addition, because location is measured very precisely, further bias could occur because the effect of a variable measured with large error (i.e., air pollution) can be transferred to another variable measured with small error (i.e., location) (*26*).

We have developed an alternative method for statistical estimation and inference for our spatial random-effects survival model in which we exploited the fact that the partial likelihood function used for parameter estimation in the independent observation Cox model can be written in terms of a Poisson likelihood. We have shown that our model can also be written as a random-effects Poisson likelihood (*27*). We then applied the estimation methods of Ma (*28*) for random-effects Poisson models to our suitably transformed model.

We then analyzed the ACS data with this alternative approach without any surface modeling. Here, the sulfate effect was 0.0125 (standard error of 0.00252), and the estimate of the random-effects variance was 0.0027, values nearly identical to our two-stage estimation procedure (Table 1). The close correspondence with the two approaches is likely due to the relatively large number of deaths per location (average of 274 deaths per MSA).

We found that the estimates of the association between the individual risk factors and mortality and their estimates of uncertainty were nearly identical in the Cox survival model and the random-effects Cox survival model, thus validating the use of the Cox model to identify the set of individual risk factors for mortality.

There is a substantial computational advantage to decomposing the estimation procedure into two stages. If there are a few deaths per community, the $\{\delta(s)\}$ will be poorly characterized (*15*). Areas in which no deaths occurred must be removed from the analysis, a limitation not inherent with the Cox random-effects modeling approach. A limitation of the Cox random-effects survival model is the intensiveness of computer resources. For example, for the ACS study this approach took 37 hr of computing time on a Sun Microsystem Ultra Enterprise 450 computer (Palo Alto, CA, USA). In contrast, our two-stage modeling approach took only a few minutes.

We assumed that the survival experience of subjects within an MSA were independent. There could be clustering of responses within these large metropolitan areas (i.e., county, city, neighborhood). If such clustering exists, our estimates of uncertainty in the $\{\delta(s)\}$, *V*,
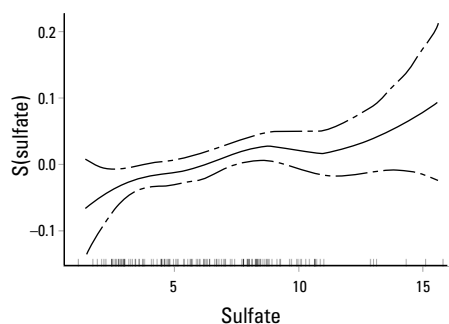
will be biased downward, resulting in inflated estimates of the between-community variance, θ. We recommend that spatial random-effects survival models be developed that incorporate several levels of clustering.

REFERENCES AND NOTES

1. Dockery DW, Pope CA III, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG, Speizer FE. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759 (1993).

2. Pope CA, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. Am J Respir Crit Care Med 151:669–674 (1995).

3. Thun MJ, Day-Lally CA, Calle EE, Flanders WD, Health CW. Excess mortality among cigarette smokers: changes in a 20-year interval. Am J Public Health 85:1223–1230 (1995).

4. Gamble JF. $PM_{2.5}$ and mortality in long-term prospective cohort studies: cause-effect or statistical associations? Environ Health Perspect 106:535–549 (1998).

5. Lipfert FW. Estimating air pollution-mortality risks from cross-sectional studies: propspective vs. ecologic study designs. In: Particulate Matter: Health and Regulatory Issues. Air and Waste Management Association Publication VIP 49. Proceedings of the International Specialty Conference, 4-6 April 1995, Pittsburgh, Pennsylvania. Pittsburgh, PA:Air and Waste Management Assoc., 1995;78–102.

6. Moolgavkar SH. Air pollution and mortality [Letter]. N Engl J Med 330:1237 (1994).

7. Health Effects Institute. Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality: A Special Report of the Institute's Particle Epidemiology Reanalysis Project. Cambridge, MA:Health Effects Institute, 2000.

8. Cox DR. Regression models and life-tables. J R Stat Soc Series B 34:187–202 (1972).

9. Ware JH, Stram DO. Statistical issues in epidemiologic studies of the health effects of ambient air pollution. Can J Stat 16:5–13 (1988).

10. Miron J. Spatial autocorrelation in regression analysis: a beginner's guide. In: Spatial Statistics and Models (Gaile GL, Willmott CJ, eds). Boston:D. Reidel Publishing, 1984.

11. Griffin DA, Doyle PG, Wheeler DC, Johnson DL. A tale of two swaths: urban childhood blood-lead levels across Syracuse, New York. Ann Assoc Am Geogr 88:640–645 (1988).

12. Matheron G. Principles of geostatistics. Ecol Geol 58:1246–1266 (1963).

13. Goodchild MF. Spatial Autocorrelation. Norwich:Geo Books, 1986.

14. Easton DF, Peto J, Babiker GAG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. Stat Med 10:1025–1035 (1991).

15. Burnett RT, Ross WH, Krewski D. Non-linear mixed regression models. Environmetrics 6:85–99 (1995).

16. Cakmak S, Burnett R, Krewski D. Adjusting for temporal variation in the analysis of parallel time series of health and environmental variables. J Expos Anal Environ Epidemiol 129–144 (1998).

17. Cleveland WS, Devlin SJ. Robust locally-weighted regression and smoothing scatterplots. J Am Stat Assoc 74:829–836 (1988).

18. Hastie T, Tibshirani R. Generalized Additive Models. London: Chapman and Hall, 1990.

19. Zeger SL, Liang K-Y, Albert PS. Models for longitudinal data: a generalized estimation equation approach. Biometrics 73:13–22 (1985).

20. Priestly MB. Spectral Analysis and Time Series. London: Academic Press, 1981.

21. SAS PROC PHREG, SAS/STAT Software: Changes and Enhancements through Release 6.12. Cary, NC:SAS Institute Inc., 1997.

22. S+SpatialStats: User's Manual for Windows and UNIX. Seattle, WA:MathSoft, 1997.

23. Understanding GIS: The ARC/INFO Method. Redlands, CA:Environmental Systems Research Institute, 1997.

24. Getting to Know ArcView GIS. Redlands, CA:Environmental Systems Research Institute, 1998

25. Abbey DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Beeson LW, Yang JX. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. Am J Respir Crit Care Med 159:373–382 (1999).

26. Zidek JV, Wong H, Le ND, Burnett R. Causality, measurement error and multicollinearity in epidemiology. Environmetrics 7:441–451 (1996).

27. Ma R, Krewski D, Burnett R. Random Effects Cox Models: a Poisson Modelling Approach. Tech Rep no 338. Ottawa, Ontario, Canada:Laboratory for Research in Statistics and Probability, Carleton University, 2000.

28. Ma R. An Orthodox BLUP Approach to Generalized Linear Mixed Models [PhD Thesis]. Vancouver, British Columbia, Canada:The University of British Columbia, 1999.