

Total Suspended Particulate Matter and Daily Mortality in Cincinnati, Ohio

Joel Schwartz

Environmental Epidemiology Program, Harvard School of Public Health, Boston, MA 02115 USA

The London smog disaster of December 1952 established that high levels of airborne particles and sulfur dioxide produced large increases in daily death rates¹. Since the London episode, time-series analyses have associated daily mortality with daily particle concentrations in London (2,3), New York (4,5), Santa Clara, California (6), Detroit (7), Steubenville, Ohio (8), Philadelphia (9), St. Louis and eastern Tennessee (10), Utah Valley (11), Minneapolis (12), and Birmingham, AL (13). Both the air pollution concentrations and mortality increases were lower than seen in London in 1952.

These studies have reported that the primary association was with particles and not sulfur dioxide, and none have found any evidence for a threshold. In Santa Clara (6) and the Utah Valley (11), coexposure with sulfur dioxide was essentially absent, avoiding any possible confounding with that pollutant. Clearly, airborne particles are associated with daily mortality independent of sulfur dioxide exposure. The regression coefficients reported in Santa Clara and Utah were similar to the regression coefficients in the other studies. This suggests that the other associations were due to the particle and not the sulfur dioxide exposure. In the other cities, the regression coefficients of particles in models containing sulfur dioxide were little changed from their values without simultaneous consideration of sulfur dioxide; they also remained statistically significant. In contrast, the coefficients of sulfur dioxide were generally much lower in models that controlled for particle exposure and were not statistically significant. This indicates that variations in particle concentrations independent of sulfur dioxide were associated with variations in daily mortality. Variations in sulfur dioxide concentrations that were independent of particle concentrations did not appear to be associated with variations in daily mortality in these studies.

The study in Philadelphia provided some more details on the specific causes of death that were elevated on high air pollution days. The relative risks were highest for pneumonia and cardiovascular disease. They were also highest for the elderly. The present analysis sought to examine the relationship between airborne particles and daily mortality in Cincinnati, Ohio. In particular, it sought to determine if the finding of higher relative risks for pneumo-

nia, cardiovascular disease, and the elderly can be replicated. In addition, this study used diagnostic plots and sensitivity analyses that have been used in few if any prior studies to ensure that the association is not confounded by inadequate control for weather.

Daily deaths of residents of Hamilton County, Ohio, the county containing Cincinnati (population 873,224 in 1980) were extracted from the detail mortality tapes of the National Center for Health Statistics for the calendar years 1977–82. Deaths from accidental causes (ICD 9 >800) and deaths occurring outside the county were excluded from the analysis. In addition to daily deaths, deaths of persons aged 65 or older and daily deaths from cardiovascular disease (ICD 390–448) and from pneumonia (ICD 480–486) were tabulated. These daily counts were matched to the 24-hr average (midnight to midnight) of total suspended particulates (TSP) from all the population-oriented monitors within the city. Pollution data were retrieved from the Environmental Protection Agency's aerometric data bank. For each day, TSP from all available monitors were averaged to produce an area wide mean. Daily monitoring for TSP began in 1977. Only one monitoring station was running every day in the city of Cincinnati that year. In 1979–81, two additional monitors were operated on a daily schedule. Several additional monitoring stations were operational on a one-day-in-six basis, and were included in the average when available. Data on daily mean temperature and daily mean humidity were obtained from the weather station at the Cincinnati airport.

Counts of rare events, such as dying on a particular day, are not normally distributed, and were modeled in Poisson regressions. Substantial seasonal and other long-term temporal patterns are often found in daily mortality data. While air pollution may be responsible for part of that pattern, other factors clearly play the dominant role. Temperature and humidity may be related to mortality in a nonlinear way. The analytic method was chosen to deal with these issues.

The basic approach taken in this analysis was to use dummy variables for each of the 72 months in the study period and for 8 categories of temperature and dew point temperature. Dummy variables for each month in the study controlled for seasonal

Recent studies have associated particulate air pollution with daily mortality in a number of U.S. communities. This study sought to replicate those analyses in Cincinnati, Ohio, and to test the strength of the association to the deletion of days with extreme weather conditions and to sensitivity analyses in the analytical approach. This study demonstrates the use of graphical diagnostic procedures to assure adequate control for season and weather and to confirm that the risk is particularly elevated in the elderly and for deaths from pneumonia and cardiovascular disease. Daily total suspended particulate (TSP) concentrations were available in Cincinnati from 1977 to 1982. They were matched to daily counts of nonaccidental deaths, temperature, and dew point temperature. Poisson regression analysis controlled for seasonal and monthly variations and potentially nonlinear relationships to temperature and humidity. TSP was associated with increased risk of mortality. The relative risk for a 100 $\mu\text{g}/\text{m}^3$ increase in TSP was 1.06 (95% CI = 1.03–1.10). The relative risk for the elderly was higher (1.09), as was the risk for pneumonia (1.18) and cardiovascular disease (1.08). The similarity to recently reported results in Philadelphia is striking. Given the consistent findings from multiple locations, the relationship should be considered causal. *Key words:* mortality, nonparametric smoothing, particles, Poisson regression, splines, time series, total suspended particulates. *Environ Health Perspect* 102:186–189(1994)

and monthly variations, including the possibility that seasonal peaks may vary in date and magnitude from year to year. Multiple categories of temperature and dew point temperature allow for a U-shaped relationship with daily deaths, or other nonlinearities, if present. A linear and quadratic time term were also included to capture any continuous temporal trend in mortality over the period.

Several alternative approaches were also examined to see how sensitive the air pollution results were to the method used to control for these potential confounders. The second approach replaced the dummy variables with natural splines. This modeled the dependence of mortality on time and temperature using cubic polynomials fit to intervals of time (or temperature). The cubic polynomials were subject to continuity constraints at the boundaries of the intervals. The number of intervals each variable was divided into was chosen by analysis of deviance. First an initial num-

Address correspondence to J. Schwartz, Environmental Epidemiology Program, Harvard School of Public Health, 665 Huntington Avenue, Boston MA 02115 USA.

Joel Schwartz is supported in part by a John D. and Catherine T. MacArthur Fellowship. Received 21 September 1993; accepted 28 December 1993.

ber of intervals was chosen. For example, the initial number of intervals chosen for time was 24, which means each interval was 6 months long, and was fit with a separate cubic polynomial. Then the number of intervals was increased (decreasing the number of months per interval in our example) until the improvement in model fit for the increased degrees of freedom used was not statistically significant. The spline approach has the capability of fitting dependencies that may be highly nonlinear and is therefore useful in assuring that any association with TSP is not confounded by inadequate control for weather or season.

Another approach to ensuring that any observed association with TSP is not due to a coincidence of high TSP days with extreme weather conditions is to repeat the analysis excluding such days. The basic analysis was repeated excluding days with 24-hr mean temperature below 10°F or above 80°F. These represent the coldest and warmest 2.5% of the days, respectively. In addition, the analysis was repeated using the minimum daily temperature instead of the average daily temperature as the control variable.

The final approach used nonparametric regression. The generalized additive model (14) allows a Poisson regression to be fit as a sum of nonparametric smooth functions of each of the predictor variables. A nonparametric smoother is a tool for summarizing the trend of a response measurement, Y , as a function of one or more predictor measurements, X_1, \dots, X_p . Most smoothers are generalizations of weighted moving averages. These predict the expected value of Y at X_k as the weighted mean of the actual values of the Y 's corresponding to all the X 's in a symmetric neighborhood around X_k . The weights decline with distance from the center of the neighborhood. The properties of such smoothers have been extensively discussed (15,16). Loess (17) was used in this analysis, which fit the counts of daily deaths to smooth functions of temperature, dew point temperature, day of study, and TSP. Loess uses a running regression rather than a running moving average, with weights that decline as the cube of the distance from the center of the neighborhood. The generalized additive model allows for a test of the improvement in fit compared to a linear no-thresh-

old model for air pollution. In addition to the statistical test, the smoothed covariate adjusted plot of daily deaths versus air pollution allows a direct graphical examination of the shape of the relationship between exposure and outcome.

Time series data are often serially correlated. Although this may be due to the weather, there may be serial correlation remaining in the residuals of the regression models. Ignoring such serial correlation can lead to biased hypothesis tests. All regression models were tested for serial correlation in the residuals, and if any was found, autoregressive Poisson models were estimated (18).

Table 1 shows the distribution of air pollution, temperature, dew point temperature, and daily mortality in Cincinnati during the years 1977–82. The distribution of counts of daily deaths is illustrated

Table 1. Distribution of variables in the analysis: daily mortality and total suspended particulates (TSP) in Cincinnati, 1977–1982

Variable	25th Percentile	Median	75th Percentile	Mean	SD
Daily deaths	17	20	24	21	4.92
TSP ($\mu\text{g}/\text{m}^3$)	53	71	93	76	30.8
Temperature, °F	37	55	70	52	19.6
Dew point temperature, °F	28	44	60	43	19
Daily deaths, ≥age 65	12	14	17	15	4.1
Pneumonia deaths	0	0	1	0.6	0.8
Cardiovascular deaths	9	11	13	11	3.5

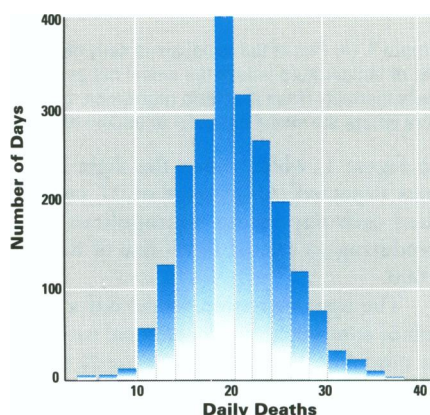


Figure 1. Histogram of daily deaths in Cincinnati, Ohio, during the study period, 1977–1982.

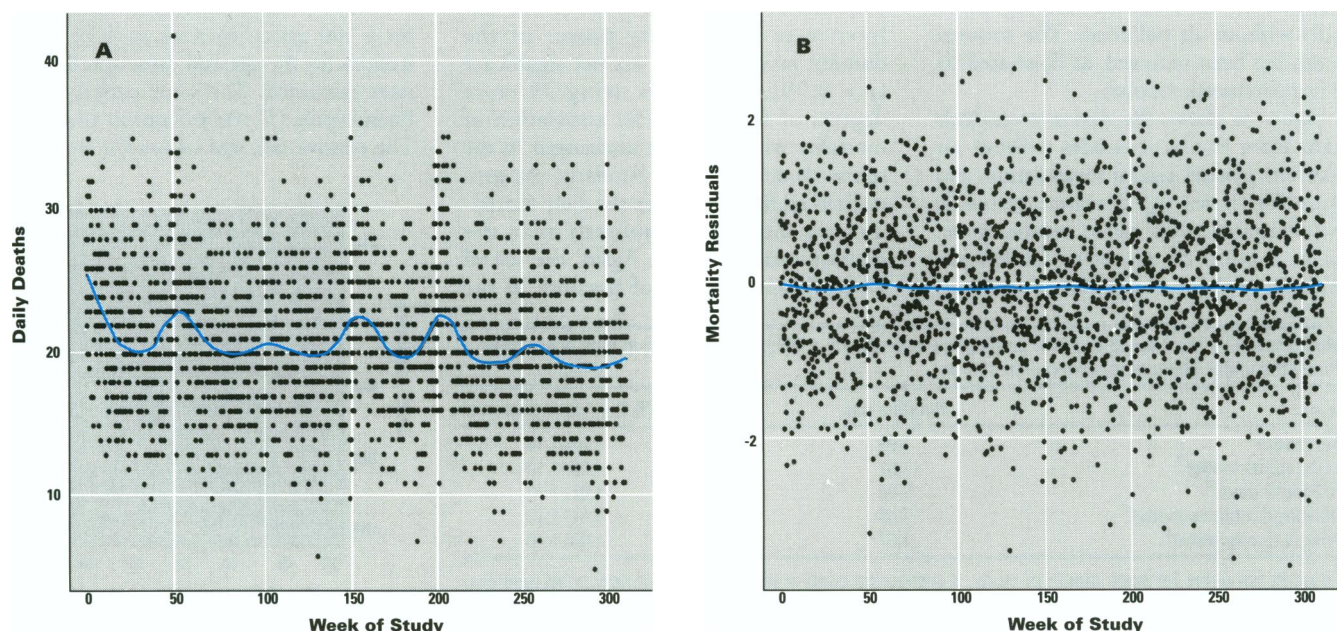


Figure 2. (A) Plot of daily deaths in Cincinnati, Ohio, versus time during the course of the study. Each point represents 1 of the 2191 days during the study period. The line shown is a nonparametric smoothed curve fit to the data and illustrates the seasonal variability in the data. (B) Plot of the residuals of daily mortality (from the basic regression model without total suspended particulates) versus time. There is no long-term pattern remaining in the data, as indicated by the nonparametric smooth.

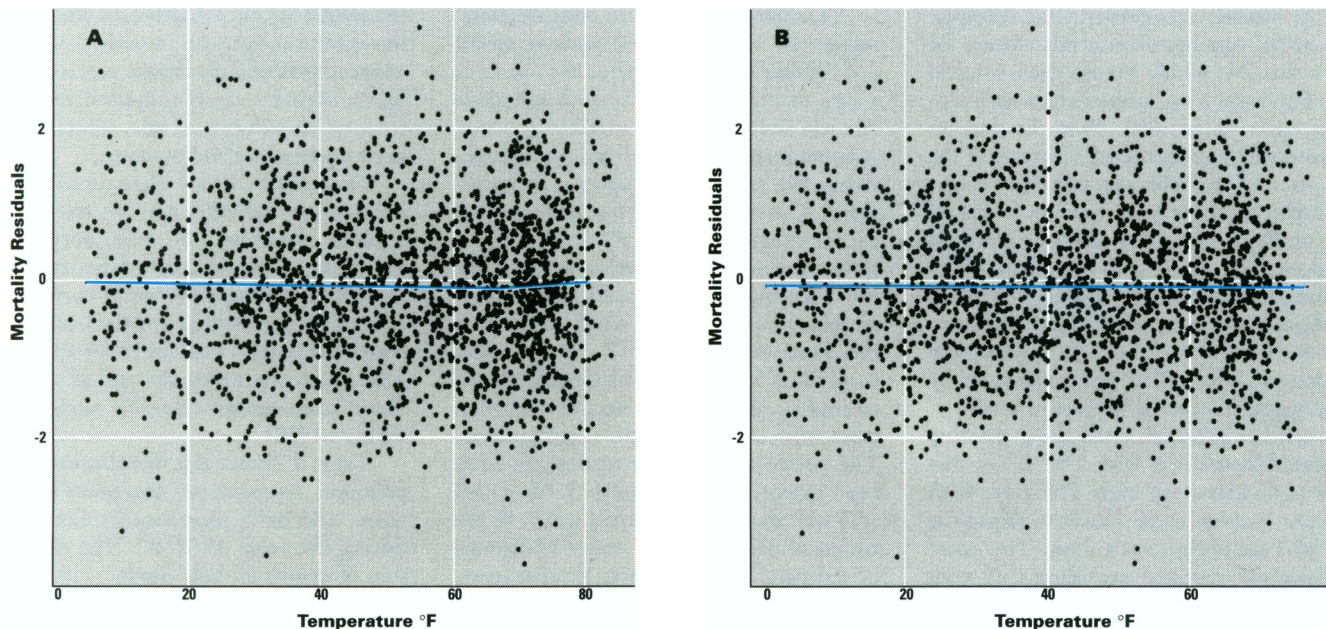


Figure 3. (A) Plot of the residuals of daily mortality (from the basic regression model without total suspended particulates) versus temperature. There is no interval of temperature where the model net over- or underestimates daily mortality, as indicated by the nonparametric smoothed curve. (B) Plot of the residuals of daily mortality (from the basic regression model without total suspended particulates) versus dew point temperature. There is no interval of dew point temperature where the model net over- or underestimates daily mortality, as indicated by the nonparametric smoothed curve.

in Figure 1, which shows the slight skewness expected. An average of 21 persons died each day in the metropolitan area, producing an annual death rate of 8.8 per 1000.

The basic regression model did a good job of controlling for the seasonal variation in mortality. For example, Figure 2a shows a plot of daily deaths by week of study. The line through the points is a nonparametric smooth. The strong seasonal variation is obvious. Figure 2b shows the deviance residuals from the basic regression model without air pollution. The seasonal pattern has been removed, as illustrated by the nonparametric smooth.

Figure 3a shows the deviance residuals (again from the basic model without air pollution) plotted against temperature, and Figure 3b shows the deviance residuals plotted against dew point temperature. Again, a good fit seems to have been achieved, with no temperature intervals

showing systematic over- or underestimation of mortality.

Table 2 shows the relative risk of (RR) death for a $100 \mu\text{g}/\text{m}^3$ increase in TSP from the basic model and the various sensitivity analyses. TSP was a significant risk factor for mortality in this analysis (RR = 1.06, 95% CI = 1.03–1.10). The use of natural splines resulted in a more parsimonious model, with 34 intervals of time and 2 intervals each of temperature and dew point temperature being fit with cubic polynomials within each interval. The difference in explanatory power of the dummy variable model was not significant ($p = 0.99$), despite its using 39 more degrees of freedom. The association of mortality with TSP was unchanged in the spline model (Table 2). Similarly, the generalized additive model fit the daily mortality data more parsimoniously than the dummy variable model. Again, the use of 42 additional degrees of freedom in the

basic model produced no significant improvement in fit ($p = 0.98$). Excluding the warmest and coldest days also had little effect on the association between TSP and daily deaths. The relative risk for TSP in that model was also similar to the basic model. The test for nonlinearity in the TSP association was not significant ($p = 0.20$). Figure 4 shows the nonparametric smoothed plot of daily deaths versus TSP in Cincinnati, after controlling for covariates.

Table 3 shows the relative risk of death for a $100 \mu\text{g}/\text{m}^3$ increase in TSP concentrations for the age and cause specific death rates examined. The same pattern seen in Philadelphia (9) was evident in Cincinnati. The relative risk was higher for the elderly

Table 2. Relative risk of death and 95% confidence intervals for a $100 \mu\text{g}/\text{m}^3$ increase in total suspended particulates under different modeling alternatives

Model	Relative risk	95% Confidence interval
Basic model ^a	1.06	1.03–1.10
Natural spline model ^b	1.07	1.03–1.10
No extreme days ^c	1.06	1.02–1.10
Generalized additive model ^d	1.08	1.04–1.13
Minimum temperature ^e	1.07	1.03–1.10

^a72 dummy variables for each month of study, 8 categories each of temperature and dew point temperature, and linear and quadratic time trend variables.

^bSeparate cubic polynomials fit for 34 intervals of time, and 2 intervals each of temperature and dew point temperature.

^cModel a, excluding days with temperature below 10°F or above 80°F.

^dNonparametric smooth functions of time, temperature, and dew point temperature.

^eModel 1, but with daily minimum temperature instead of daily average temperature.

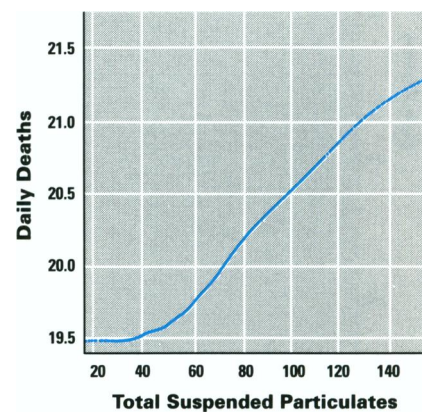


Figure 4. Nonparametric smooth curve of daily mortality versus total suspended particulate, after controlling for temperature, dew point temperature, and time in a generalized additive model. No evidence for a threshold is seen in the relationship.

Table 3. Relative risk of death and 95% confidence intervals for a 100 µg/m³ increase in total suspended particulates for all-cause mortality, and for selected subgroups by age and cause of death

Outcome	Relative risk	95% Confidence interval
All deaths	1.06	1.03–1.10
Deaths, age ≥65	1.09	1.05–1.14
Pneumonia deaths	1.16	0.95–1.42
Cardiovascular deaths	1.08	1.03–1.14

All models are after adjustment for covariates.

and for deaths from pneumonia and cardiovascular disease.

The association between mortality and TSP seen in Cincinnati was almost identical to that seen in Philadelphia. Moreover, the relationships between TSP and cause- and age-specific daily deaths in the two cities were remarkably similar. For comparison, in Philadelphia, the relative risk for a 100 µg/m³ increase in TSP was 1.07 (95% CI = 1.04–1.10) for all-cause mortality, 1.11 (95% CI = 0.97–1.27) for pneumonia deaths, 1.10 (95% CI = 1.06–1.14) for cardiovascular deaths, and 1.10 (95% CI = 1.06–1.13) for deaths in the elderly. The finding of such similar associations in different locations makes it less likely that the associations could be due to a confounder.

Smoking habits, occupational exposure, dietary patterns, and socioeconomic factors are not potential confounders of this association because their day-to-day variation is small and not correlated with air pollution. More likely, potential confounders are epidemics, other factors that produce the seasonal variation in mortality, and weather. In Cincinnati, daily mortality and respiratory epidemics both peaked in the winter. In contrast, winter was the season with the lowest TSP concentrations. Hence, these factors are unlikely sources of upward bias in the association. Furthermore, Figure 2 demonstrates the success of

the model in removing seasonal variation from the data. The results are not confounded by extreme weather conditions because the association is little changed when those days are excluded. In addition, the residual plots indicate that the basic model did not under- or overestimate the dependence on weather conditions in any temperature range. The sensitivity of the results to alternative ways of controlling for weather was slight. In the absence of a plausible candidate for a major confounder, the conclusion that the observed association is causal seems reasonable.

One way to indirectly assess the possibility of confounding by omitted factors is to examine the relationship in different locations, as the correlation between those factors and air pollution should differ among cities. If the association is due to an omitted or imperfectly controlled confounder, the induced regression coefficient for particulate air pollution will depend on the regression coefficient of the confounder with daily mortality and the regression coefficient of the confounder with particulate air pollution. If the latter varies widely but the coefficient of particulate air pollution with mortality does not vary widely, confounding by that factor is unlikely. The recent studies of particulate air pollution and daily mortality have all reported relatively similar effect sizes. This lends credence to their findings, since if the associations were due to confounding, we would expect a greater variability in the effect size estimates. Figure 5 shows the relative risk of death for a 100 µg/m³ increase in TSP concentration from the recent studies of air pollution and daily mortality. The consistency is remarkable.

Although the mechanism by which airborne particles exacerbate illnesses and increase their mortality rates is not understood, neither is the mechanism by which tobacco smoking increase the risk of death from myocardial infarctions. This fact has not prevented a conclusion being drawn from the strong epidemiologic information in the case of smoking and it should not impede a similar conclusion in the case of

respirable particles. Moreover, the London episode of 1952 provides ample demonstration of biological plausibility: it is clear that respirable particles increased mortality in that episode, although no mechanism was determined for that case, either.

REFERENCES

1. HMSO. Mortality and morbidity during the London fog of December 1952. Report no. 95. on Public Health and Medical Subjects, Her Majesty's Public Health Service. London: Her Majesty's Stationary Office, 1954.
2. Mazumdar S, Schimmel H, Higgins ITT. Relation of daily mortality to air pollution: an analysis of 14 London winters, 1958/59–1971/72. *Arch Environ Health* 37:213–220 (1982).
3. Schwartz J, Marcus A. Mortality and air pollution in London: a time series analysis. *Am J Epidemiol* 131:185–194(1990).
4. Glaser M, Greenburg L. Air pollution, mortality, and weather: New York City, 1960–1964. *Arch Environ Health* 22:334–343 (1971).
5. Schimmel H, Murawski TJ. The relation of air pollution to mortality. *J Occup Med* 18:316–333(1976).
6. Fairley D. The relationship of daily mortality to suspended particulates in Santa Clara County, 1980–1986. *Environ Health Perspect* 89:159–168(1990).
7. Schwartz J. Particulate air pollution and daily mortality in Detroit. *Environ Res* 56:204–213(1991).
8. Schwartz J, Dockery DW. Particulate air pollution and daily mortality in Steubenville, Ohio. *Am J Epidemiol* 135:12–20(1992).
9. Schwartz J, Dockery DW. Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am Rev Respir Dis* 145:600–604(1992).
10. Dockery DW, Schwartz J, Spengler JD. Air pollution and daily mortality: associations with particulates and acid aerosols. *Environ Res* 59:362–373(1992).
11. Pope CA, Schwartz J, Ransom M. Daily mortality and PM₁₀ pollution in Utah Valley. *Arch Environ Health* 42(3):211–217(1992).
12. Schwartz J. Particulate air pollution and daily mortality: a synthesis. *Public Health Rev* 19:39–60(1992).
13. Schwartz J. Air pollution and daily mortality in Birmingham, AL. *Am J Epidemiol* 137:1136–47(1993).
14. Härdle W. Smoothing techniques with implementation in S. New York. New York: Springer-Verlag, 1991.
15. Gasser T, Müller HG, Marmitsch V. Kernals for nonparametric curve estimation. *J R Stat Soc B* 47:238–52(1985).
16. Cleveland WS, Devlin SJ. Robust locally-weighted regression and smoothing scatterplots. *J Am Stat Assoc* 74:829–36(1988).
17. Hastie T, Tibshirani R. Generalized additive models. London:Chapman and Hall, 1990.
18. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 73:13–22(1986).

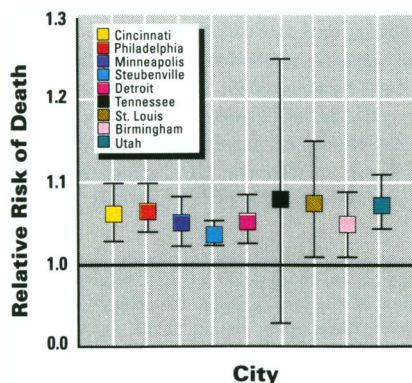


Figure 5. Relative risk of all-cause mortality from a 100 µg/m³ increase in daily total suspended particulates from studies in different cities.