



## A Sensitivity Analysis of Mortality/Pm-10 Associations in Los Angeles

Patrick L. Kinney, Kazuhiko Ito & George D. Thurston

To cite this article: Patrick L. Kinney, Kazuhiko Ito & George D. Thurston (1995) A Sensitivity Analysis of Mortality/Pm-10 Associations in Los Angeles, *Inhalation Toxicology*, 7:1, 59-69, DOI: [10.3109/08958379509014271](https://doi.org/10.3109/08958379509014271)

To link to this article: <https://doi.org/10.3109/08958379509014271>



Published online: 27 Sep 2008.



Submit your article to this journal [↗](#)



Article views: 34



View related articles [↗](#)



Citing articles: 47 View citing articles [↗](#)

## A SENSITIVITY ANALYSIS OF MORTALITY/PM-10 ASSOCIATIONS IN LOS ANGELES

Patrick L. Kinney, Kazuhiko Ito, George D. Thurston

New York University Medical Center, Nelson Institute  
of Environmental Medicine, Tuxedo, New York, USA

*Several recent studies have reported statistically significant and quantitatively similar associations between daily mortality and PM-10 or total suspended particulates (TSP). These results have raised questions of biological plausibility, as well as questions regarding the statistical methods employed, which are relatively new and not widely understood. This article evaluates the sensitivity of mortality/PM-10 results to a range of statistical methods in a newly developed data set from Los Angeles county for the period 1985–1990. Data reported here include total daily deaths (excluding accidents and suicides), 24-h average PM-10, daily 1-h maximum O<sub>3</sub> and carbon monoxide, maximum daily temperature, and mean daily relative humidity. Analyses were limited to the subset of days on which PM-10 data were available (every sixth day). Several alternative methods for addressing each of four issues were explored in this data set. These four issues were cyclic data variations, weather influences, other air pollutants, and the distribution of residuals. The associations between mortality and PM-10 concentrations, measured as relative risks associated with a 100- $\mu\text{g}/\text{m}^3$  increase in PM-10, were only mildly sensitive to the alternative statistical methods. In particular, no difference was observed between the results of ordinary least squares and Poisson models. We observed a relative risk of about 1.05, which is similar to, but somewhat smaller than, the mortality/PM-10 relative risks reported in recent studies. These new results add to the growing body of data suggesting that current levels of airborne particulate matter may contribute to excess deaths in the United States.*

There have been several studies published since 1990 that have reported statistically significant and quantitatively similar associations between daily numbers of deaths and gravimetric particulate matter concentrations, either PM-10 (mass concentrations of particles with aerodynamic diameters less than 10  $\mu\text{m}$ ) or TSP (total suspended particulate mass concentration), in metropolitan areas (Schwartz, 1991, 1993; Schwartz & Dockery, 1992a, 1992b; Dockery et al., 1992; Pope et al., 1992). It is remarkable that consistent results have been obtained in diverse cities that vary in population, weather patterns, and levels of copollutants. The quantitative consistency of the reported associations and their coherence with epidemiologic studies of morbidity outcomes have led Schwartz to conclude that the effects of low-level airborne particle exposures on mortality are likely to be causal (Schwartz, 1993).

The authors wish to thank Esther Christie for her contribution to the development of the air pollution database. This work was supported by NIEHS grant ES05711.

Address correspondence to Patrick L. Kinney, Columbia University School of Public Health, Division of Environmental Sciences, 60 Haven Avenue, B-1, New York, NY 10032.

This body of results has been viewed by many with skepticism. Primary among the concerns raised has been the lack of biological plausibility. It is difficult to understand how a  $10\text{-}\mu\text{g}/\text{m}^3$  increase in PM-10 levels could result in about a 1% increase in daily deaths in a city, as would be implied from the data. Further, there are no animal toxicology data for particulate matter that suggest significant biological mechanisms that could account for health effects at the low levels of exposure encountered in the population-based studies. The other principal concern has been that the results might be due to an artifact of the complex and specialized statistical methods utilized in most of the recent literature on this topic.

Time-series studies of daily mortality and air pollution use a variety of complex statistical methods designed to address the special character of serial data. Because many of the statistical methods employed are relatively new and specialized, they are fully understood by few analysts and even fewer users of time-series results. In addition, there is currently no firm consensus among analysts as to which statistical approaches are most valid for the analysis of the effects of environmental factors on daily mortality. A general discussion of these issues is presented elsewhere in this issue (Thurston & Kinney, 1994).

There are four broad issues that arise in the analysis of multiyear records of daily time-series data. The first is the need to control for cyclic data variations. It is usually the case that a substantial part of the variability over time in mortality counts, air pollution concentrations, and weather variables is related to temporal cycles, the most significant being the cycle of the seasons. Mortality rates in the United States are always highest in the winter and lowest in the summer. Air pollutant concentrations may peak in either the winter or summer, depending on source strength, meteorology, and atmospheric transformation processes. These cycles, if not addressed somehow in the analysis, can lead to biases in estimating the relationship between mortality and air pollution.

A second issue is the role of weather as an independent cause of death. It is generally agreed that heat spells and cold snaps can kill people, but the forms of the functional relationships are not well understood. Furthermore, weather and air pollution concentrations are usually related. As a result, the manner in which temperature is modeled will usually affect the mortality/pollution effect estimates. Both lack of control and overcontrol may yield biased pollution estimates, though in different directions.

A third, and related, issue is the complexity of the pollution exposures. Particulate matter is not the only air pollutant that may influence mortality. Past studies have implicated other pollutants, including ozone ( $\text{O}_3$ ) and carbon monoxide (CO) (Kinney & Ozkaynak, 1991; Hexter & Goldsmith, 1971), two pollutants with extensive research bases supporting biologic plausibility as well. Because of this, and because of the correlations among pollutants, the estimated particulate matter effect may vary depending on which other pollutants are included in the model.

A final issue relates to the distribution of the residual variability in mortality that remains after the model is fit. Because death is a discrete event, daily mortality counts will tend to have a Poisson distribution rather than the normal Gaussian distribution that is assumed in most standard analyses. It has been argued that methods that assume Poisson residuals are therefore necessary (Schwartz, 1993). On the other hand, when the mean is large, a Poisson process will be nearly Gaussian. It is not yet clear how much difference this issue makes in practice.

This article evaluates the sensitivity of mortality/PM-10 relationships over a range of statistical approaches to the four issues just introduced. Rather than attempting to dictate the single most valid set of methods, we pose the question, "how much difference does it make to the results when the full range of commonly used methods are tested?" The sensitivity analysis was performed using data from Los Angeles county for the period 1985–1990, the period during which PM-10 data first became available.

## METHODS

Daily counts of total deaths that occurred in Los Angeles County in the period 1 January 1985 to 31 December 1990 were obtained from National Center for Health Statistics death certificate tapes. Deaths due to accidents and suicides and nonresident deaths were excluded from the total counts, yielding the daily death count variable used in the analyses reported here. PM-10, O<sub>3</sub>, and CO data collected in Los Angeles county were obtained in digital form from the U.S. Environmental Protection Agency's Aerometric Information and Retrieval System (AIRS). The 24-h average PM-10 concentrations, collected every 6 days, were taken from 4 monitoring sites. Daily maximum 1-h O<sub>3</sub> and CO levels were obtained from 8 sites each. The multiple-site data for each pollutant were averaged after filling missing values using a multiple regression algorithm. Missing values that were filled in this way represented 7%, 3%, and 5% of the total numbers of observations for PM-10, O<sub>3</sub>, and CO, respectively. The resulting county-wide mean time series were the pollution variables used in all analyses. Meteorological data (maximum daily temperature and mean daily relative humidity) collected at Los Angeles International Airport were obtained from the National Climatic Center. The analyses reported here were focused on evaluating the sensitivity of mortality/PM-10 associations to a range of analytical methods. We therefore restricted our attention to the subset of days on which PM-10 was monitored ( $n = 364$ ). Exploratory cross-correlation analysis indicated that the mortality/PM-10 correlation was greatest on the same day; thus, same-day PM-10 was used in all the present analyses. The 1-day lag of O<sub>3</sub> was used based on previous analysis of a 10-yr record in Los Angeles (Kinney & Ozkaynak, 1991).

Several methods aimed at controlling for temporal cycles were evaluated: regression with dummy (i.e., indicator) variables for seasons as covariates;

restricting analysis to data from a particular season (i.e., winter or summer); regression with sine and cosine functions of time included as covariates (two levels of control were evaluated: the first used 4 sine and cosine waves, with periods of 1 mo, 6 mo, 1 yr, and 2 yr; the second used 10 sine and cosine waves, with periods of 1 mo, 2 mo, 3 mo, 6 mo, 9 mo, 1 yr, 15 mo, 18 mo, 21 mo, and 2 yr); and season-specific regressions, as previously, but with the 4 sine and cosine terms included as covariates. The use of seasonal dummy variables, or analyses restricted to a particular season, may not entirely control for cycle-related variations in the data, since these cycles can have substantial influence within seasons. Multiperiod sine and cosine waves should provide more complete control. One measure of the adequacy of cyclic controls is the Durbin–Watson (DW) statistic (Durbin & Watson, 1951), which is a measure of the first-order autocorrelation in the model residuals. With no autocorrelation, a DW statistic of 2 would be expected. With inadequate control of cycles, positive autocorrelation among residuals is likely, yielding a DW statistic between 0 and 2.

The subtraction of a multiday (e.g., 15-day) moving average from each data point prior to analysis has been used to remove cyclic influences in some previous studies. We were not able to evaluate this “prefiltering” method, because the available PM-10 data were restricted to every sixth day. Each of the sensitivity analyses utilized a basic model that involved regressing  $\ln(\text{mortality})$  on same-day PM-10, temperature, and relative humidity.

The sensitivity of PM-10 results to two levels of control for temperature and relative humidity was evaluated. The first model included only same-day temperature and relative humidity as covariates along with PM-10. The second model followed the two-stage approach described by Schwartz (1993), in which a set of several weather variables were first tested for association with mortality in a multiple regression model. Weather variables considered included same-day and up to 3-day lags of temperature, relative humidity, and indicator variables for hot days, humid days, and hot and humid days. Those weather variables with moderate associations with mortality (regression  $t$  ratio equal to or greater than 1) were included as covariates in a model that regressed  $\ln(\text{mortality})$  on same-day PM-10 and 10 sine/cosine functions for control of cycles.

To evaluate the sensitivity of PM-10 results to inclusion of other pollutants in the model, regressions were run that tested the effects of  $\text{O}_3$  (lagged 1 day) and CO alone, and in combination with PM-10, as predictors of  $\ln(\text{mortality})$ . Each regression model included same day weather control and 10 sine/cosine functions.

The final sensitivity analyses examined the influence of alternative assumptions about the residual distribution and functional form. Three models were compared: ordinary least squares (OLS, linear regression of mortality on one or more predictor variables), log-linear analysis (linear regression of the natural logarithm of mortality on one or more predictor

variables), and Poisson regression (this assumes the same functional form as the log-linear model, but with Poisson, rather than normally, distributed residuals). OLS and log-linear regressions were carried out using the GLM procedure of the Statistical Analysis System software package (SAS Institute, Cary, NC). Poisson regression was carried out using the SAS Nlin procedure, with code developed and kindly provided by Dr. Joel Schwartz.

Results from each of the analyses just described are reported in both tabular and graphical forms as relative risks and 95% confidence intervals. To be consistent with recent literature, PM-10 relative risks were computed for a 100- $\mu\text{g}/\text{m}^3$  increase in PM-10. Note that 100  $\mu\text{g}/\text{m}^3$  represents the upper 95th percentile of PM-10 concentrations in this data set. Relative risks for  $\text{O}_3$  and CO likewise were computed at the upper 95th percentiles: 143 ppb for  $\text{O}_3$  and 10 ppm for CO.

## RESULTS

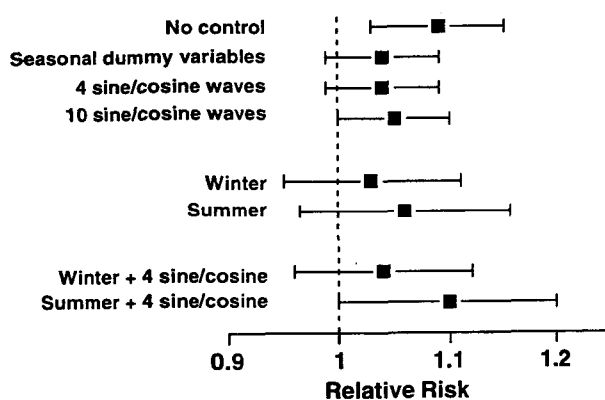
Table 1 presents summary statistics (mean, standard deviation, and range) for each of the variables analyzed, over the entire data set and within two 4-mo "seasons," November to February and June to September. Table 2 and Figure 1 present the results of the sensitivity analysis directed at the issue of seasonal cycles. In the absence of seasonal control, a relative risk of 1.09 (95% CI 1.03–1.15) for PM-10 was observed; however, this result is likely confounded by uncontrolled cyclic variations. The DW statistic for this model was 1.38, indicating positive first-order autocorrelation in the residuals. Among the models that used alternative methods of cycle controls, the use of seasonal dummy variables or two alternative sets of sine/cosine variables yielded relative risks of 1.04–1.05. The confidence intervals were similar across the 3 models, with the lower bounds just touching or spanning a relative risk of 1.00 (i.e.,  $p$  value in the neighborhood of .05). Similar relative risks were observed for analyses restricted to

**TABLE 1.** Descriptive Statistics, Overall and by Season

Period	Statistic	Deaths (per day)	PM-10 ( $\mu\text{g}/\text{m}^3$ )	$\text{O}_3$ (ppb)	CO (ppm)	Temperature (°F)	Relative humidity (%)
All	Mean	153	58	70	4.7	70	70
	SD	20	23	41	2.9	7	16
	Range	113–224	15–177	3–201	1–13	54–98	14–97
November–February	Mean	169	61	36	7.4	67	63
	SD	20	29	20	2.7	7	20
	Range	132–224	15–177	3–100	2–13	54–88	14–97
June–September	Mean	143	60	101	2.8	73	77
	SD	12	18	37	1.5	4	8
	Range	116–179	20–116	21–201	1–8	63–87	49–90

**TABLE 2.** Results of Sensitivity Analysis

Methodological issue	Method	PM-10 relative risk	95% Confidence interval
Temporal cycles	No control	1.09	1.03–1.15
	Seasonal dummies	1.04	0.99–1.09
	4 Sine/ cosine waves	1.04	0.99–1.09
	10 Sine/cosine waves	1.05	1.00–1.11
	Winter only	1.03	0.95–1.11
	Summer only	1.06	0.96–1.15
	Winter + 4 sine/cosine	1.04	0.96–1.12
	Summer + 4 sine/cosine	1.10	1.00–1.21
Weather	No control	1.05	1.00–1.10
	Temperature + relative humidity	1.05	1.00–1.11
	Extensive weather model	1.04	0.99–1.08
Other pollutants	PM-10 only	1.05	1.00–1.11
	O <sub>3</sub> only	1.02 <sup>a</sup>	1.00–1.05
	CO only	1.07 <sup>b</sup>	1.01–1.13
	PM-10 + O <sub>3</sub>	1.05	1.00–1.11
	PM-10 + O <sub>3</sub>	1.00 <sup>a</sup>	0.94–1.06
	PM-10 + CO	1.04	0.98–1.09
	PM-10 + CO	1.05 <sup>b</sup>	0.99–1.12
Model type	OLS	1.05	1.00–1.10
	Log-linear	1.05	1.00–1.11
	Poisson	1.05	1.00–1.10

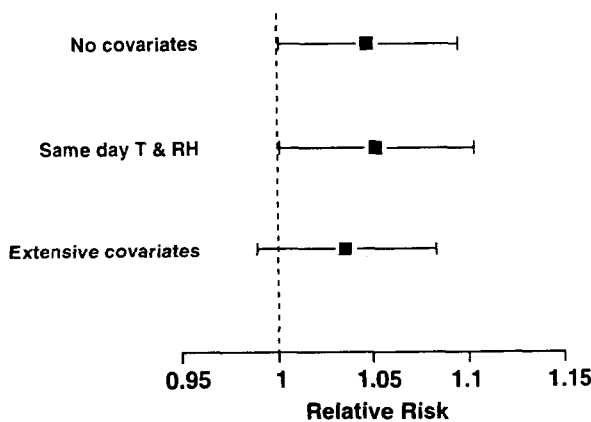
<sup>a</sup>Relative risk for O<sub>3</sub>.<sup>b</sup>Relative risk for CO.**FIGURE 1.** PM-10 relative risks and 95% confidence intervals obtained with alternative methods to control for temporal cycles. All models included same-day temperature and relative humidity as covariates. Relative risks were computed for a 100- $\mu\text{g}/\text{m}^3$  increase in PM-10 concentration.

the winter or summer seasons, although the confidence intervals increased due to the reduction in sample size. Season-specific regressions that also included four sine and cosine variables gave somewhat larger PM-10 relative risks. There was some indication from these latter results that the mortality/PM-10 association was stronger in summer than in winter. As a whole, however, the PM-10 relative risks exhibited a substantial consistency across a range of alternative methods aimed at controlling the cyclic behavior of the data.

Among the cyclic control methods we explored, 3 yielded DW statistics above 1.8: the model that included 10 sine and cosine variables, and the 2 models restricted to the summer season. On the basis of these results, we chose to employ the 10 sine/cosine model as the basic approach for cyclic control in all the models discussed below.

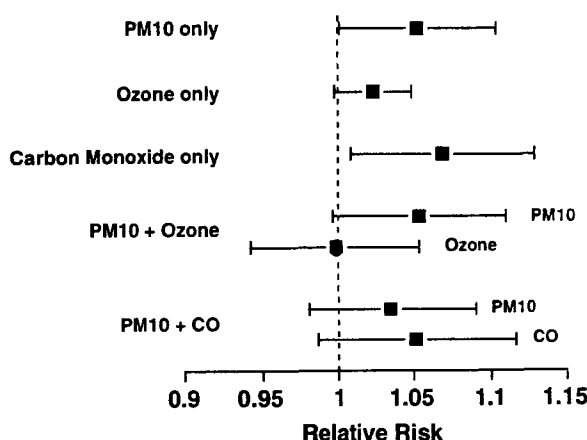
Regressions with varying levels of control for temperature and relative humidity (none; same-day; extensive lagged variables) all yielded similar results (Table 2 and Figure 2), although the PM-10 relative risk and statistical significance were reduced somewhat with the more extensive weather controls.

Regression models that included alternative pollutants (PM-10, lag 1 O<sub>3</sub>, or CO) all yielded significant (or nearly so) relative risks for the individual pollutants (Table 2 and Figure 3). In bivariate regressions (i.e., regressions including PM-10 and one other pollutant), results were more variable. With both PM-10 and O<sub>3</sub> in the model, the PM-10 relative risk was essentially unchanged (RR = 1.05), while the O<sub>3</sub> relative risk dropped to one. This suggests that the O<sub>3</sub> effect on mortality, if any, is weaker than that of PM-10. The correlation of the slope estimates for O<sub>3</sub> and PM-10 was  $-0.5$ , indicating a substantial collinearity in these 2 pollutants when included simultaneously in the model. In contrast to the situation for O<sub>3</sub>, the relative



**FIGURE 2.** PM-10 relative risks and 95% confidence intervals obtained with varying levels of weather controls. All models included 10 sine/cosine functions as covariates. Relative risks were computed for a 100- $\mu\text{g}/\text{m}^3$  increase in PM-10 concentration.

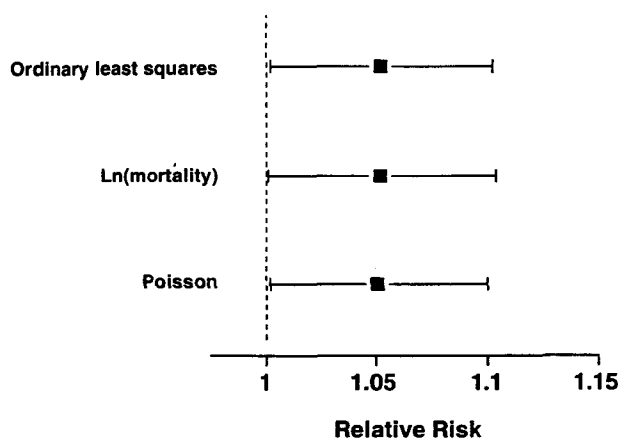




**FIGURE 3.** Relative risks and 95% confidence intervals for PM-10,  $O_3$ , and CO in univariate and bivariate regressions with mortality. All models included same-day temperature and relative humidity, and 10 sine/cosine functions as covariates. Relative risks were computed for a  $100\text{-}\mu\text{g}/\text{m}^3$  increase in PM-10 concentration, a 143-ppb increase in  $O_3$  concentration, and a 10-ppm increase in CO concentration.

risks for both PM-10 and CO dropped somewhat when both were included in the model, suggesting a similar strength of association with mortality for the two pollutants. There was a moderate collinearity in the slope estimates for these two pollutants (slope correlation =  $-0.4$ ). Overall, the range of PM-10 relative risks observed in these alternative models (1.03–1.05) again demonstrated a rather mild degree of sensitivity.

Finally, we evaluated the sensitivity of mortality/PM-10 associations under differing assumptions about functional form and residual distributions



**FIGURE 4.** PM-10 relative risks and 95% confidence intervals obtained with alternative regression models. All models included same-day temperature and relative humidity, and 10 sine/cosine functions as covariates. Relative risks were computed for a  $100\text{-}\mu\text{g}/\text{m}^3$  increase in PM-10 concentration.

(Table 2 and Figure 4). Three approaches were used: ordinary least squares (assumes a linear relationship and normally distributed errors with constant variance), log-linear regression (assumes an upward-curving relationship between mortality and PM-10 and normally distributed errors with variance that increases in proportion to level of mortality), and Poisson regression (same assumptions as log-linear except that errors are assumed to be Poisson distributed). No difference was observed in the PM-10 relative risks from the three models.

## DISCUSSION AND CONCLUSIONS

We have evaluated the sensitivity of daily mortality/PM-10 associations to a range of analytical methods in a newly developed 6-yr data set from Los Angeles county. We found that the estimated proportional increase in daily mortality associated with a  $100\text{-}\mu\text{g}/\text{m}^3$  increase in PM-10 concentration (or relative risk) fell generally between 1.03 and 1.05, regardless of the method used. These results indicate that sensitivity to methods was low in this data set.

These results for Los Angeles represent the first independent confirmation of the mortality/particulate matter associations reported for PM-10 and TSP in the recent series of articles by Schwartz and colleagues. The PM-10 relative risk (RR) we obtained for Los Angeles (1.05) is somewhat smaller than those reported previously in Utah Valley (RR = 1.16) by Pope and colleagues (1992), St. Louis, MO (RR = 1.16) by Dockery and colleagues (1992), and Birmingham, AL (RR = 1.11), by Schwartz (1993). Two of those studies used multiday averages of PM-10 as the exposure metric, which may have yielded a larger effect by picking up both same-day and lagged PM-10 effects. We were unable to evaluate multiday averages because of the limitations imposed by every-sixth-day measurement of PM-10 in Los Angeles. It is also possible that the lower RR in Los Angeles is due to unique features of the Los Angeles aerosol, such as lower acidity or a relatively large ratio of primary to secondary particles.

It should be noted that the PM-10 relative risks we estimated were often not statistically significant at the .05 level (i.e., the 95% confidence intervals often included unity). Nonsignificance tended to occur when statistical power was reduced by including several explanatory variables simultaneously in the models or when the data set was reduced in size for seasonal analysis. Note, however, that for the "basic" model that included PM-10, 10 sine/cosine waves for seasonal control, and same-day temperature and relative humidity for weather control, a statistically significant relative risk of 1.05 was obtained ( $p = .04$ ). The data set analyzed in this article ( $n = 364$ ) appears to be of borderline size for detecting statistically significant relations between mortality and PM-10 in Los Angeles County.

The generalizability of the sensitivity results presented here is not yet known. The data set was relatively small ( $n = 364$ ), and this was the first

reported time-series analysis based solely on PM-10 data collected every sixth day. In addition, it is possible that sensitivity to weather covariates in Los Angeles may be lower than would be seen in other climates where more pronounced swings in temperature are experienced (e.g., Chicago). On the other hand, recent sensitivity analyses carried out using daily time-series data from the Utah Valley (Pope et al., 1992) and Birmingham, AL (Schwartz, 1993), have indicated a similar degree of insensitivity to methods.

Ozone appeared to be associated with mortality in univariate regressions (with weather and cyclic controls). However, the association disappeared when PM-10 was added to the model. These results suggest that the relationship, if any, between  $O_3$  and mortality is weaker than that involving PM-10 and, further, that the univariate  $O_3$  effect may have been due to  $O_3$  acting as a PM-10 surrogate. The correlation between the PM-10 and  $O_3$  slope estimates ( $-0.5$ ) indicates that the two variables were not independent in this data set. Definitive conclusions regarding a possible role of  $O_3$  in daily mortality cannot be drawn from this small data set, nor from the existing literature. Among the several studies that have analyzed mortality in relation to  $O_3$ , two have reported no associations in Detroit, St. Louis, and Kingston/Harriman, TN (Schwartz, 1991; Dockery et al., 1992), and two have reported significant associations in Los Angeles and New York City (Kinney & Ozkaynak, 1991, 1992).

Results of the present study indicated that CO had an independent association with mortality that was of similar strength to that of PM-10. Kinney and Ozkaynak (1991) reported strong associations involving CO in an analysis of a 10-yr Los Angeles county data set. They also noted that there were high correlations between CO, nitrogen dioxide, and km (a filter tape soiling measure similar to British Smoke), all of which were likely to be markers for primary motor vehicle pollution in Los Angeles. Again, multi-pollutant analyses using larger data sets will be needed before more definitive conclusions can be drawn about the relative roles of PM-10 and CO as predictors of daily mortality.

These new results for Los Angeles County, taken together with other recently published time-series analyses, support the view that particulate matter air pollution may contribute to excess deaths in the United States. The challenges facing researchers in the future will be to develop a better understanding of the mechanisms by which this phenomenon is occurring and to more fully sort out the respective influences of PM-10, other pollutants, and weather in the epidemiologic associations. These issues are the subject of ongoing work at our institute.

## REFERENCES

- Dockery, D. W., Schwartz, J., and Spengler, J. D. 1992. Air pollution and daily mortality: Associations with particulates and acid aerosols. *Environ. Res.* 59:362-373.
- Durbin, J., and Watson, G. S. 1951. Testing for serial correlation in least squares regression II. *Biometrika* 38:159.

- Hexter, A. C., and Goldsmith, J. R. 1971. Carbon monoxide: Association of community air pollution with mortality. *Science* 172:265–266.
- Kinney, P. L., and Ozkaynak, H. 1991. Associations of daily mortality and air pollution in Los Angeles county. *Environ. Res.* 54:99–120.
- Kinney, P. L., and Ozkaynak, H. 1992. Associations between ozone and daily mortality in Los Angeles and New York City. *Am. Rev. Respir. Dis.* 145:A95.
- Pope, C. A. III, Schwartz, J., and Ransom, M. R. 1992. Daily mortality and PM-10 pollution in Utah Valley. *Arch. Environ. Health* 47:211–217.
- Schwartz, J. 1991. Particulate air pollution and daily mortality in Detroit. *Environ. Res.* 56:204–213.
- Schwartz, J. 1993. Air pollution and daily mortality in Birmingham, AL. *Am. J. Epidemiol.* 137:1136–1147.
- Schwartz, J., and Dockery, D. W. 1992a. Particulate air pollution and daily mortality in Steubenville, Ohio. *Am. J. Epidemiol.* 135:12–19.
- Schwartz, J., and Dockery, D. W. 1992b. Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am. Rev. Respir. Dis.* 145:600–604.
- Thurston, G. D., and Kinney, P. L. 1994. Air pollution epidemiology: Considerations in time-series modeling. *Inhal. Toxicol.* 7(1):71–83.