REVIEW OF EPIDEMIOLOGICAL EVIDENCE OF HEALTH EFFECTS OF PARTICULATE AIR POLLUTION

C. Arden Pope, III[1]

Douglas W. Dockery[2,3]

Joel Schwartz[2]

For submission to:

Inhalation Toxicology

Issue reporting the Colloquium on Particulate Air Pollution

April 28, 1994

Abbreviated title: **HEALTH EFFECTS OF PARTICULATE POLLUTION**

[1]Department of Economics, Brigham Young University, 142 FOB, Provo, UT 84602.

[2]Environmental Epidemiology Program, Department of Environmental Health, Harvard School of Public Health, 665 Huntington Ave., Boston, MA 02115.

[3]The Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, 180 Longwood Ave, Boston MA 02115.

ABSTRACT

This paper summarizes epidemiological evidence of health effects of particulate air pollution. Acute exposure to elevated levels of particulate air pollution has been associated with increased cardio-pulmonary mortality, increased hospitalization for respiratory disease, exacerbation of asthma, increased incidence and duration of respiratory symptoms, declines in lung function, and restricted activity. Small deficits in lung function, higher risk of chronic respiratory disease and symptoms, and increased mortality has also been associated with chronic exposure to respirable particulate air pollution. Health effects have been observed at levels common to many U.S. cities and at levels below current U.S. National Ambient Air Quality Standards. Although the biological mechanisms involved are poorly understood, recent epidemiological evidence supports the hypothesis that respirable particulate air pollution is an important risk factor for respiratory disease and cardio-pulmonary mortality.

INTRODUCTION

Numerous time-series studies have observed associations between particulate air pollution and various human health endpoints including: 1) incidence and duration of respiratory symptoms; 2) lung function; 3) restricted activity; 4) hospitalization for respiratory disease; and 5) mortality (see Tables 1 and 2). Because these studies typically evaluated relatively short-term temporal relationships (usually 1 to 5 days), the observed pollution effects have often been interpreted as "acute health effects" or "short-term effects" of air pollution. Measures of lung function, incidence rates of respiratory symptoms, and mortality rates have also been compared across communities or neighborhoods with different levels of particulate air pollution (see Table 3). Given their cross-sectional design and the use of longer-term pollution data (usually one year or more), these studies are often interpreted as evaluating the chronic or cumulative effects of exposure.

The results of most of these studies have been published since the mid-1980's. Health effects have been observed at pollution concentration levels common to many U.S. cities. Because of the public health relevance of these findings, there has been considerable interest about potential biological mechanisms, constituents of particle pollution, potential interactions with other pollutants and risk factors, sources of these air pollutants, and how these findings should be incorporated into public policy (Utell, Samet 1993). To begin the colloquium on particulate air pollution and human mortality and morbidity, we reviewed the epidemiological evidence of human health effects of particulate air pollution. Three papers were presented that reviewed evidence of 1) acute mortality effects, 2) acute morbidity effects, and 3) chronic health effects. These presentations were highly related and their substance has been or is being published elsewhere (Dockery, Pope 1994; Pope, Bates, Raizenne 1994; Schwartz

1994e). The objective of this paper is to combine and briefly summarize the essence of these three presentations.

ACUTE MORTALITY

Studies of air pollution episodes

Several early studies focused on severe air pollution episodes. One of the first episodes that was well documented occurred in the narrow industrial valley of the Meuse River in Belgium (Firket 1931). On December 1-5, 1930, a severe fog/air pollution episode resulted in more than 60 deaths--a mortality rate of more than 10 times the normal. Several hundred others also developed respiratory illness often accompanied by cardiac insufficiency. Another severe air pollution episode occurred in Donora, Pennsylvania (also an industrial town situated in a valley) in October 26-31, 1948 (Ciocco 1961). Very large increases in respiratory morbidity were reported. Approximately 20 deaths that occurred during and immediately following the episode, in the population of only 14,000, were attributed to the pollution. The most dramatic of the severe air pollution episodes that have been well documented occurred in London, England (Logan 1953). For four or five days beginning on December 4, 1952, particulate and sulfur dioxide air pollution concentrations were greatly elevated in a stagnant air mass. The rapid increase in air pollution was followed by a rapid increase in deaths with a lead-lag time of only about one day. The number of deaths returned to near normal within three week after the episode. The sudden increase in death was due largely to increases in respiratory and cardiovascular mortality. Several studies have also reported associations between daily mortality in London and air pollution at more recent and much lower pollution levels (Ostro 1984; Schwartz, Marcus 1990; Ito et al. 1993). Elevated respiratory and cardiovascular disease morbidity and mortality have also reported

for a more moderate pollution episode in the North Rhine-Westfalia area of Germany (Wichmann et al. 1989).

Although these air pollution episode studies have been referred to often, they are reiterated here because they linked cardio-pulmonary disease mortality and morbidity with extremely elevated concentrations of particulate and/or sulfur oxide air pollution and suggest effects at much lower levels. Furthermore, although the biological mechanisms involved were poorly understood, there remained little disagreement that, at very high levels, ambient air pollution can be an important risk factor associated with early mortality.

Daily time-series mortality studies

A series of recent studies analyzed the temporal distribution of deaths and particulate air pollution at current levels. Reviews of many of these studies have been or are being published (Ostro 1993; Schwartz 1994e; Dockery, Pope 1994; Pope, Bates, Raizenne 1994). These studies observed changes in daily death counts associated with short-term changes in particulate air pollution. A particulate pollution threshold was not generally observed in these studies. The relative risk of mortality increased monotonically with particulate concentrations usually in a near linear fashion. Table 1 summarizes 10 daily time-series studies that have been conducted over the last three or four years. These studies have observed similar particulate air pollution effects in varied locations. Because various measurements of particulate pollution were used, precise comparisons between studies are difficult. But when particulate measurements are converted to PM_{10} (see Schwartz 1994e or Dockery, Pope 1994 for more detail), there is remarkable consistency in the estimated effect of PM_{10} across these studies. Effect estimates range between .7% and 1.6% increase in daily mortality for each 10 $\mu g/m^3$ increase in PM_{10} concentration with a weighted mean of about 1.0%.

Five of the studies also provided a breakdown of mortality by broad cause-of-death categories. Particulate air pollution generally had the largest effect on respiratory disease deaths, but also often had a significant effect on cardiovascular disease mortality. A recent examination of cardiovascular deaths in Philadelphia reported that on days with high particulate air pollution there was a substantial increase in respiratory factors as contributing causes for deaths with cardiovascular disease reported as the underlying cause of death (Schwartz 1994c).

There are two principle concerns about these daily time-series mortality studies: 1) effects may be due to artifacts of the regression modeling techniques and 2) there may be confounding due to inadequate control of seasonal factors, epidemics, other long wavelength trends, weather variables, or other pollutants.

Mortality is classically modeled as following a Poisson process which can generally be described as a process that generates independent and random occurrences across time or space. If time is divided into discrete periods such as 24-hour periods, or days, the daily death counts theoretically would be distributed as a Poisson distribution. Therefore, most of the studies have used Poisson regression techniques to estimate pollution effects while controlling for other potential risk factors. For large mean daily deaths (more than 10 or 15), Gaussian (or normal) regression models are reasonable because, when mean counts are large, the Gaussian distribution is a close approximation of the Poisson distribution. However, even when the number of daily deaths is large, the variance of the data is not constant as assumed using normal ordinary least squares regression. Therefore, Poisson regression may be preferred. Nonetheless, Gaussian regression models and various robust regression approaches produce estimates of associations which are nearly identical. (For examples see Saldiva et al 1994; Schwartz 1994e; Schwartz 1994f).

While it can be argued that individual studies of daily mortality and particulate air pollution may be confounded by unmeasured or inadequately modelled covariates, taken together it is unlikely that such confounding could be consistently acting in all these studies. Daily changes in weather are expected to be associated with both air pollution and mortality, that is they are potential confounders. All the recent studies attempted to control for the effects of weather. Almost all of the studies allowed for non-linear relationships with weather factors in the analysis. The estimated particulate pollution effects are not sensitive to the inclusion of seasonal controls using approaches such as including seasonal indicator variables, including sine and cosine terms, using non-parametric smoothing techniques, or prefiltering the data. Similar particulate pollution effects have been observed in studies from both warm and cold climates, from both dry and humid locations, and from areas where particulate concentrations peak in both the summer and winter. For a comparison of these studies see Schwartz (1994e).

These studies further suggest that the observed particulate pollution effects are not confounded by associations with SO_2, or O_3. Effects observed in locations with winter peaking of particulate concentrations, when ozone concentrations are low and not a potential confounding factor are similar to effects in summer peaking areas. Most of the studies examined SO_2. The relationship between mortality and particulates was generally independent of SO_2, while the SO_2 relationship disappeared when particles were considered. In the Utah Valley (Pope, 1992) and Santa Clara (Fairley, 1990), SO_2 concentrations were low, yet the estimated PM_{10} effects were similar to other effects observed in study areas with higher SO_2 concentrations.

ACUTE MORBIDITY

If daily mortality is associated with daily particulate pollution levels, then associations with morbidity health endpoints should also be observed (Bates, 1992). In fact, associations with acute exposure to particulate air pollution have also been observed for various morbidity health endpoints including: increased health care visits for respiratory illness, exacerbations of asthma, increased incidence and duration of respiratory symptoms, and declines in measures of lung function. A brief summary of selected acute morbidity studies is given in Table 2.

Hospital Usage

Pope (1989) reported the results of a unique natural experiment that occurred in Utah Valley. During the winter of 1986-87 a labor dispute resulted in a closure of the local steel mill, the largest single source of particulate air pollution in the valley. During this winter PM_{10} concentrations averaged 51 $\mu g/m^3$ with a high of 113 $\mu g/m^3$ compared with a mean of 90 $\mu g/m^3$ and a high of 365 $\mu g/m^3$ in the previous winter. During this winter, children's hospital admissions for respiratory disease dropped by over 50 percent, compared with adjacent years. Regression analysis estimated a 4.2% decrease in asthma and bronchitis admission of children associated with a 10 $\mu g/m^3$ decrease in two-month mean PM_{10}. Subsequent research indicated that declines in bronchitis and asthma admissions during the winter when the steel mill was closed were not observed in neighboring Utah counties (Pope 1991).

Two studies have found that increased rates of respiratory hospital admissions in Southern Ontario are associated with increased sulfate and ozone concentrations (Bates and Sizto 1987 and 1989; and Burnett et al. 1994). The authors have suggested a link to acid aerosol concentrations, but alternatively, these associations may be attributable to particle

mass concentrations which were not specifically considered in the analysis. Thurston and colleagues have reported similar associations in Toronto, Ontario (Thurston, Ito, and Lippmann 1993) and for several cities in New York state (Thurston et al. 1992). The focus of these studies was the effects of acid aerosols, but estimates were reported for various measures of particle exposure. Taken together, these studies found an increase in hospital admissions for respiratory diagnoses ranging from approximately 0.8% to 3.4% for each 10 $\mu g/m^3$ increase in daily mean PM_{10}.

Several studies have also analyzed emergency department visits and found them to be associated with particulate air pollution. For example, particulate air pollution was associated with emergency department visits for asthma in Seattle (Schwartz et al. 1993), emergency department visits for chronic obstructive pulmonary disease in Barcelona (Sunyer et al. 1991; Sunyer et al. 1993), and emergency department visits in Steubenville, Ohio (Samet et al. 1981). Based on the reported results of these studies the estimated percent increase in emergency visits associated with a 10 $\mu g/m^3$ increase in PM_{10} range from 0.5% to 3.4% with a weighted mean of approximately 1.0%.

Given the associations between particulate pollution and emergency department visits, it would be expected that associations would also be observed for outpatient visits for respiratory illness. There is less availability of outpatient data to researchers than there is for hospitalization data. However, based on a limited study of outpatient data from a clinic in Salt Lake City, Utah, Lutz (1983) reported that strong positive associations were observed between weekly particulate pollution levels and the percentage of patients with a diagnosis of respiratory tract or cardiac illnesses.

Exacerbation of asthma

Evidence from the hospital admissions and emergency visit studies suggests that exposures to particulate air pollution may be directly associated with asthma attacks. Several investigators have considered less severe asthmatic attacks as reported by panels of asthma patients. Studies of asthmatics in the Los Angeles area (Whittemore and Korn 1980), The Netherlands (Roemer, Hoek, and Brunekreef 1993) and in Denver, Colorado (Ostro et al. 1991) reported asthmatic attacks associated with particle exposures. Bronchodilator use has also been evaluated as a measure of exacerbation of asthma in a panel of asthmatics in The Netherlands (Roemer, Hoek, Brunekreef 1993) and in Utah Valley (Pope et al. 1991). Based on the reported results of these studies the estimated percent increase in asthma attacks or use of bronchodilator associated with a 10 $\mu g/m^3$ increase in PM_{10} range from 1.1% to 12% with a weighted mean of approximately 3.0%.

Respiratory symptoms

The use of daily diaries to record respiratory symptoms is an inexpensive method of evaluating acute changes in respiratory health status. In a commonly applied study design, panels of schoolchildren recorded the presence of specific respiratory symptoms in these daily diaries. Symptom reports are often aggregated into upper respiratory symptoms (including such symptoms as runny or stuffy nose, sinusitis, sore throat, wet cough, head cold, hay fever, and burning or red eyes) and lower respiratory symptoms (including wheezing, dry cough, phlegm, shortness of breath, and chest discomfort or pain). In addition, cough, the most frequently reported symptom, is often analyzed separately.

Studies of upper and lower respiratory symptoms had been conducted in Utah Valley (Pope et al. 1991; Pope and Dockery 1992), The Netherlands (Hoek and Brunekreef 1993, 1994), a study of six U.S. cities (Schwartz et al 1994), and Southern California (Ostro et al.

1993). Very small, often statistically insignificant associations between particulate pollution and upper respiratory symptoms were observed. The association with lower respiratory disease was larger and usually statistically significant. Based on these studies, the estimated percent increase in reported lower respiratory symptoms associated with a 10 $\mu g/m^3$ increase in PM_{10} was as high as 15% but with a weighted mean of approximately 3.0%. Cough was analyzed separately in three of these studies as well as in another diary study in The Netherlands (Roemer, Hoek, and Brunekreef 1993), a study of two Swiss cities (Braun-Fahrlander et al. 1992), and a study in Uniontown, PA (Neas et al. 1992). Cough was typically associated with particulate pollution with the estimated percent increase in reported cough associated with a 10 $\mu g/m^3$ increase in PM_{10} as high as 28% but with a weighted mean of approximately 1.2%.

Lung function

Several studies have observed negative associations between particulate pollution and lung function. Panels of elementary school children in Steubenville, Ohio had their lung function measured weekly before, during, and after particulate and sulfur oxide episodes during four periods in 1978 through 1980 (Dockery et al. 1982). Declines in Forced Expired Volume in three-quarters of a second ($FEV_{.75}$) were observed following these episodes. Deficits in $FEV_{.75}$ remained for up to two weeks following the episodes. In a re-analysis of the Steubenville data, Brunekreef et al. (1991) found the strongest association with the mean particle concentrations over the previous five days. Similar decreases in Forced Expired Volume in one second (FEV_1) were observed in school children following a particulate and sulfur oxide pollution episode in January 1985 in The Netherlands (Dassen et al. 1986). Subsequent studies of panels of school children in The Netherlands with weekly lung function

measurements (Hoek, Brunekreef 1993, 1994) have also shown decreased FEV_1 associated with daily PM_{10} concentrations. Lagged effects of up to seven days were observed.

Recently Pope and Kanner (1993) analyzed repeated FEV_1 measurements in a panel of adults with chronic obstructive pulmonary disease who were participating in the Lung Health Study. Measurements were taken 10 to 90 days apart. FEV_1 levels were associated with PM_{10} concentrations. Koenig and colleagues (1993) studied the lung function (FVC and FEV_1) of children in Seattle, WA with relatively low particulate air pollution levels. Lung function declines were associated with fine particulate air pollution for asthmatic children, but not for non-asthmatic children. Overall, these studies generally observed a decrease of less than 0.35% in FEV_1 associated with each 10 $\mu g/m^3$ increase in daily mean PM_{10}.

Several studies have used peak flow measurements as an indicator of acute changes in lung function including studies in The Netherlands (Hoek, Brunekreef 1993, 1994; Roemer, Hoek, Brunekreef 1993), Utah Valley (Pope et al. 1991; Pope and Dockery 1992), and Uniontown, Pennsylvania (Neas et al 1992). In these studies a 10 $\mu g/m^3$ increase in PM_{10} was associated with a 0.04% to 0.25% decrease in peak flow measurements. As with FEV_1, the strongest associations with peak flow were often with particulate pollution over the previous several days.

Restricted activity

Associations between particulate air pollution and more general measures of acute disease have also been observed. For example, Ostro (1983, 1987, and 1990) and Ostro and Rothschild (1989) evaluated the timing of restricted activity days of U.S. adult workers. Restricted activity due to respiratory morbidity was consistently associated with particulate pollution. Morbidity was often more strongly associated with the fine, respirable, or sulfate component of particulate pollution. Ransom and Pope (1992) reported similar associations

between PM_{10} and grade school absences of children in Utah Valley, Utah. Lagged pollution effects of up to several weeks were observed for both restricted activity in adults and in school absences.

CHRONIC HEALTH EFFECTS

Early cross-sectional studies suggested chronic health effects of particulate air pollution. For example, Martin (1964) reported that in the Greater London region overall annual respiratory mortality (as opposed to episodic mortality) was significantly related to smoke (or particulate pollution) levels. Holland & Reid (1965) made a cross-sectional comparison of British male postal employees in London and in smaller country towns, where levels of SO_2 and particulate pollution were about half those in the metropolis. Accounting for cigarette smoking levels, significant decrements of FEV_1 in London employees compared to those in the provinces were reported. More recent cross-sectional studies of air pollution have also reported associations between particulate pollution and respiratory symptoms, lung function and mortality rates (Table 3).

Population-based (ecologic) mortality studies

Since the early 1970s, several studies have evaluated the mortality effects of exposure to particulate air pollution using population-based (ecologic) cross-sectional study designs (Chappie and Lave, 1982). These studies observed that on the average, mortality rates tend to be higher in cities with higher fine or sulfate particulate pollution levels than those with lower levels. Formal regression modeling techniques to evaluate cross-sectional differences in air pollution and mortality and to control for other ecologic variables was used. In an attempt to control for other risk factors, population average values for demographic variables and other factors such as smoking rates, education levels, income levels, poverty rates,

housing density and others were often included in the regression models. The basic conclusions from the population-based cross-sectional studies include: 1) Mortality rates are associated with air pollution; 2) Mortality rates are most strongly associated with fine or sulfate particulate matter; 3) An average mortality effect of 3 to 9 percent of total mortality can be estimated.

Although these population-based cross-sectional studies suggest that air pollution contributes to human mortality, the studies have been largely discounted for several reasons. One reason seems to simply be that the size of the association, if taken literally, suggests that as much as 3 to 9 percent of urban mortality in the U.S. is associated with particulate air pollution. Given that air pollution levels in the U.S. on the average are considered relatively low, such a large mortality effect seems implausible.

A more explicit limitation of the cross-sectional population-based studies is a prevailing concern that the observed association was due to confounding. Because of their ecologic design, these population-based cross sectional studies could not directly control for individual differences in cigarette smoking and other risk factors. They could only try to control for them by using population-based averages--making potential confounding a concern. Furthermore, the strength of the relationship between mortality and particulate pollution was often sensitive to model specification, socio-economic, demographic, and other risk factors included in the analysis, and the choice of study areas included in the analysis.

Chronic differences in lung function

There have been several recent studies that have evaluated associations between measures of lung function (FVC, FEV_1, PEF) and particulate pollution levels. These studies include analysis of children's lung function data from the Harvard six-city study, analysis of data from both the first and second National Health and Nutrition Examination Surveys

(NHANES I and NHANES II), and analysis of children's lung function from 24 U.S. cities. Unlike the population-based mortality studies, each of these studies had information on individual persons in the samples. The effects of air pollution on lung function was estimated after adjusting for individual differences in age, race, sex, height, weight and controlling for smoking or restricting the analysis to never smokers.

All of these studies observed small negative associations between lung function and particulate air pollution. In the six-cities study, which had the least statistical power, the association was very weak and statistically insignificant. In each of the other studies, the association was small, but statistically significant. The results suggest that a 10 $\mu g/m^3$ positive difference in PM_{10} was typically associated with less than a 2 percent decline in lung function. However, lung function measures have been shown to be important measures of health with remarkable predictive capacity for survival (Bates 1989). Furthermore, as reported in the 24-city study, the risk of relatively large deficits in lung function (15% or more) were much higher in the more polluted cities, suggesting detrimental effects of respirable particulates or particulate acidity on normal lung growth and development.

Chronic Respiratory Symptoms and Disease

There have also been several recent studies that have evaluated associations between particulate air pollution and chronic respiratory symptoms and disease. These studies include the Harvard six-cities study, analysis of COPD symptoms of never smoking Seventh-Day Adventists in California, analysis of data from the 1979 version of the U.S. National Health Interview Survey (NHIS), analysis using data from the first National Health and Nutrition Examination Survey (NHANES I), and analysis of the symptoms data from the 24-cities study. The effects of air pollution on respiratory disease or symptoms were estimated while adjusting for individual differences in various other risk factors.

In all of these studies, statistically significant associations were observed between particulate air pollution and respiratory symptoms. Particulate air pollution was most consistently associated with bronchitic symptoms. The results suggest that a 10 $\mu g/m^3$ increase in PM_{10} was typically associated with a 10 to 25 percent increase in bronchitis or chronic cough.

Prospective cohort mortality studies

Well designed prospective cohort mortality studies can provide some of the most compelling evidence about health effects because they can directly control for individual differences in age, sex, cigarette smoking, and other risk factors. Because prospective cohort studies of mortality involve collecting large amounts of information on a large number of people and following them prospectively for long periods of time, they are costly and time consuming. Currently there are only a few large prospective cohort studies that evaluate mortality effects of air pollution at levels common to U.S. urban areas.

One of these prospective cohort studies involved a 14 to 16 year prospective follow-up of 8,111 adults living in six U.S. cities (Dockery et al. 1993) TSP, PM_{10}, $PM_{2.5}$, SO_4, H^+, SO_2, NO_2, and O_3 levels were monitored. The data were then analyzed using survival analysis, including multivariate Cox proportional hazards regression modeling. Although TSP concentrations dropped over the study period, fine particulate and sulfate pollution concentrations were relatively constant. The most polluted city was Steubenville, the least polluted cities were Topeka and Portage. The Cox proportional hazards model was used to estimate adjusted mortality-rate ratios. Mortality risks were most strongly associated with cigarette smoking, but after controlling for individual differences in age, sex, cigarette smoking, body mass index, education, and occupational exposure, differences in relative mortality risks across the six cities were strongly associated with differences in pollution

levels in those cities. Associations between mortality risk and air pollution were stronger for respirable particles and sulfates, as measured by PM_{10}, $PM_{2.5}$, and SO_4, than for TSP, SO_2, acidity (H^+) or ozone. The association between mortality risk and fine particulate air pollution was consistent, nearly linear, with no apparent "no effects" threshold level. The adjusted mortality-rate ratio for the most polluted of the cities compared with the least polluted was 1.26 (95% CI, 1.08 to 1.47). Fine particulate pollution was associated with cardiopulmonary mortality and lung cancer mortality (not statistically significant) but not with mortality due to other causes analyzed as a group.

Similar results were observed in a second, larger prospective cohort study (Pope et al., 1994). This study linked ambient air pollution data from 151 U.S. metropolitan areas in 1980 with individual risk factor data and 8-year vital status data for approximately 500,000 adults drawn from the American Cancer Society (ACS) Cancer Prevention Study II (CPS-II). Adjusted mortality relative risk ratios were estimated using multiple regression analysis based on the Cox proportional hazards model. Adjusted risk ratios were calculated and reported for a difference in air pollution equal to the range of pollution exposure that existed across the metropolitan areas.

An association between mortality and air pollution was observed after adjusting for age, sex, race, cigarette smoking, pipe and cigar smoking, exposure to passive cigarette smoke, occupational exposure, education, body mass index, and alcohol use. Adjusted mortality-rate ratios (and 95% confidence intervals) of all-cause mortality for the most polluted areas compared with the least polluted equalled 1.15 (1.09-1.22) and 1.17 (1.09-1.26) when using sulfate and fine particulate measures respectively. For all-cause, cardiopulmonary, and lung cancer mortality, the associations with sulfate particles were highly statistically significant ($p < 0.001$). For all-cause and cardiopulmonary mortality,

significant associations were also found using fine particulate matter as the index of air pollution. Mortality due to other causes was not significantly associated with pollution levels. The association between air pollution and all-cause and cardiopulmonary mortality was observed for both men and women and for smokers and non-smokers.

SUMMARY AND CONCLUSIONS

Numerous studies have reported that acute exposure to elevated levels of particulate air pollution is associated with changes in various human health endpoints including: 1) increased mortality, especially cardio-pulmonary mortality; 2) increased hospitalization for respiratory disease; 3) exacerbation of asthma; 4) increased incidence and duration of respiratory symptoms; 5) declines in lung function; and 6) restricted activity. Acute health effects were observed at levels common to many U.S. cities including levels well below current U.S. National Ambient Air Quality Standards. The studies generally observed that health effects increase monotonically with pollution levels, often with a near linear dose-response relationship. Furthermore these studies suggests a coherence or cascade of associations across various health endpoints (Bates 1992).

Studies have also observed human health effects associated with chronic exposure to respirable particulate air pollution. Population-based (ecologic) cross-sectional studies that evaluated spatial distributions of mortality and air pollution have observed associations between mortality and sulfate or fine particulate pollution. These population-based studies have been criticized partly because they were not able to control directly for individual differences in cigarette smoking and other risk factors. Recent studies that could adjust for individual differences in other risk factors, however, have observed that long-term exposure to respirable particulate air pollution was associated with small deficits in lung function and

higher risk of chronic respiratory disease and symptoms. Furthermore, two recent prospective cohort studies have observed increased mortality risks associated with air pollution--even after directly controlling for individual differences in age, sex, race, cigarette smoking, and other risk factors.

Although the biological linkages remain poorly understood, the results of the acute and chronic exposure studies are complementary. In all epidemiologic studies there is the concern that the observed association is due to confounding, that is, that it results from a risk factor that is correlated with both exposure and mortality but is not adequately controlled for in the study design and analysis. Important potential confounders in cross-sectional studies such as unaccounted for differences in occupational exposure or socio-economic variables, are not likely to be confounders in daily time-series studies because such factors are unlikely to change daily in correlation with air pollution levels. The fact that daily time-series studies and cross-sectional studies observe qualitatively coherent associations between respirable particulate pollution and mortality, further supports the hypothesis that this pollution is an important risk factor for respiratory disease and cardio-pulmonary mortality.

Table 1. Summary of selected acute (daily time-series) mortality studies.

Study Area & Period	Reference	Farticulate Measure	Mean[a] PM_{10} ($\mu g/m^3$)	% increase in Mortality per 10 $\mu g/m^3$ increase in PM_{10} (95% CI)		
				Total	Respiratory	Cardiovascular
Santa Clara, CA 1980-82, 84-86	Fairley 1990	Coefficient of Haze	35	0.8 (0.2, 1.5)	3.5 (1.5, 5.6)	0.8 (0.1, 1.6)
Philadelphia PA 1973-80	Schwartz & Dockery 1992	TSP (2-day mean)	40	1.2 (0.7, 1.7)	3.3 (0.1, 6.6)	1.7 (1.0, 2.4)
Utah Valley, UT 1985-89	Pope et al 1992	PM_{10} (5-day mean)	47	1.5 (0.9, 2.1)	3.7 (0.7, 6.7)	1.8 (0.4, 3.3)
Birmingham, AL 1985-88	Schwartz 1993	PM_{10} (3-day mean)	48	1.0 (0.2, 1.9)	1.5 (-5.8, 9.4)	1.6 (-0.5, 3.7)
Cincinnati, OH 1977-82	Schwartz 1994	TSP	42	1.1 (0.5, 1.7)	2.7 (-0.9, 6.6)	1.4 (0.5, 2.4)
St. Louis, MO 1985-86	Dockery et al 1992	PM_{10} (Prev. Day)	28	1.5 (0.1, 2.9)	NA	NA
Kingston, TN 1985-86	Dockery et al 1992	PM_{10} (Prev. Day)	30	1.6 (-1.3, 4.6)	NA	NA
Detroit, MI 1973-82	Schwartz 1991	TSP	48	1.0 (0.5, 1.6)	NA	NA
Steubenville, OH 1974-84	Schwartz & Dockery 1992	TSP	61	0.7 (0.4, 1.0)	NA	NA
Sao Paulo, Brazil 1990-91	Saldiva et al. 1994	PM_{10} (2-day mean)	82	1.2[b] (0.7, 1.7)	NA	NA

[a]Conversions to PM_{10} assumed that: $PM_{10} = 0.55 * TSP$ and $PM_{10} = CoH/0.55$ (for more detail on the calculations see Dockery and Pope 1994).
[b]In the Sao Paulo study mortality counts were only for the elderly (65+ years of age).

Table 2. Summary of acute morbidity studies.

Health Indicator	Reference	Brief summary of findings
Hospital admissions (respiratory illness)	Pope 1989 and 1991 Thurston et al. 1992 Thurston, Ito, Lippmann 1993 Burnett et al. 1994 Schwartz 1994a Schwartz 1994d	Respiratory admissions, including asthma, were associated with particulate air pollution. The estimated percent increase in admissions associated with a 10 $\mu g/m^3$ increase in PM_{10} studies typically ranged from 0.8 to 3.4 percent.
Emergency visits (respiratory illness)	Samet et al. 1981 Sunyer et al. 1993 Schwartz et al. 1993	Respiratory emergency department visits were associated with particulate pollution. The estimated percent increase associated with a 10 $\mu g/m^3$ increase in PM_{10} ranged from 0.5 to 3.4 percent.
Exacerbation of asthma	Whittemore, Korn 1980 Ostro et al. 1991 Pope et al. 1991 Roemer, Hoek, Brunekreef 1993	Exacerbation of asthma, as measured by attacks, bronchodilator use, emergency visits, and hospital admission, was associated with particulate pollution. The estimated percent increase associated with a 10 $\mu g/m^3$ increase in PM_{10} was typically around 2 or 3 percent but ranged from 1.1 to 12.0 percent.
Respiratory symptoms	Pope et al. 1991 Pope, Dockery 1992 Braun-Fahrlander et al. 1992 Neas et al. 1992 Schwartz et al. 1992 Hoek, Brunekreef 1993, 1994 Ostro et al. 1993 Roemer, Hoek, Brunekreef 1993	Respiratory symptoms have been associated with particulate pollution. The estimated percent increase in lower respiratory symptoms and cough associated with a 10 $\mu g/m^3$ increase in PM_{10} was typically around 1 to 3 percent but had a wide range from 0 to 28 percent.
Lung Function	Dockery et al. 1982 Pope et al. 1991 Neas et al. 1992 Pope, Dockery 1992 Hoek, Brunekreef 1993, 1994 Pope, Kanner 1993 Roemer, Hoek, Brunekreef 1993 Koenig et al. 1993	Small, but statistically significant declines in lung function have been associated with elevated particulate air pollution. The estimated percent decline associated with a 10 $\mu g/m^3$ increase in PM_{10} was typically less that 0.5 percent.
Restricted activity	Ostro 1983, 1987, 1990 Ostro, Rothschild 1989 Ransom, Pope 1992	Restricted activity days of adult workers and school absences of grade school children were significantly associated with fine or inhalable particulate pollution.

Table 3. Summary of selected chronic mortality and morbidity studies.

Health Indicator	Study design	Reference	Brief summary of findings
Mortality rates	Population-based cross-sectional mortality studies	Lave, Seskin 1977 Chappie, Lave 1982 Lipfert 1984 Evans et al. 1984 Ozkaynak, Thurston 1987 Lipfert et al. 1988 Archer 1990 Bobak, Leon 1992	Statistical associations between mortality and fine or sulfate particulate pollution were observed. An average mortality effect of 3 to 9 percent of total mortality could be estimated from some of these studies; however, the strength of the effect was sensitive to model specification, choice of social, demographic, and other variables included in the models, and the choice of study areas used in the analysis.
Lung Function	Analysis of community ambient air pollution data with individual health and risk data from national surveys or collected cohorts	Dockery et al 1989 Schwartz 1989 Chestnut et al 1991 Raizenne et al 1993	Small, but generally statistically significant, negative associations between lung function and various measures of particulate pollution were observed. A 10 $\mu g/m^3$ positive difference in PM_{10} was typically associated with less than a 2 percent decline in lung function.
Respiratory symptoms	Analysis of community ambient air pollution data with individual health and risk data from national surveys or collected cohorts	Euler et al 1987 Dockery et al 1989 Portney & Mullahy 1990 Schwartz 1993 Dockery et al 1993	Chronic cough, bronchitis and chest illness (but not asthma) were associated with various measures of particulate air pollution. A 10 $\mu g/m^3$ increase in PM_{10} was typically associated with a 10 to 25 percent increase in bronchitis or chronic cough.
Mortality (survival times)	Prospective cohort mortality studies that linked community based air pollution data with individual risk factor and survival data	Dockery et al 1993 Pope et al. 1994	After controlling for individual differences in age, sex, cigarette smoking, and other risk factors, fine and/or sulfate particulate pollution was associated with mortality (mostly cardio-pulmonary mortality).

REFERENCES

Archer VE. Air pollution and fatal lung disease in three Utah Counties. Arch Environ Health 1990;45:325-34.

Bates DV. The health effects of air pollution. J Respir Dis 1980;1:29-37.

Bates DV. Respiratory Function in Disease, 3rd ed. Philadelphia, Pa: W.B. Saunders; 1989.

Bates DV. Health indices of the adverse effects of air pollution: The question of coherence. Environ Res 1992;59:336-49.

Bates DV, Baker-Anderson M, Sitzo R. Asthma attack periodicity: a study of hospital emergency visits in Vancouver. Environ Res 1990;51:51-70.

Bates DV, Sitzo R. Air pollution and hospital admissions in southern Ontario: the acid summer haze effect. Environ Res 1987;43:317-31.

Bates DV, Sitzo R. The Ontario air pollution study: identification of the causative agent. Environ Health Perspect 1989;79:69-72.

Bobak B, Leon DA. Air pollution and infant mortality in the Czech Republic, 1986-1988. Lancet 1992;340:1010-14.

Braun-Fahrlander C, Ackermann-Liebrich U, Schwartz J, Gnehm HP, Rutishauser M, Wanner HU. Air pollution and respiratory symptoms in preschool children. Am Rev Respir Dis 1992;145:42-7.

Brunekreef B, Kinney PL, Ware JH, Docker D, Speizer FE, et al. Sensitive subgroups and normal variation in pulmonary function response to air pollution episodes. Environ Health Perspect 1991;90:189-93.

Burnett RT, Dales RE, Raizenne ME, et al. Effects of low ambient levels of ozone and sulfates on the frequency of respiratory admissions to Ontario hospitals. Environ Res 1994 (in press).

Chappie M, Lave L. The health effects of air pollution: a reanalysis. J Urban Econ 1982;12:346-76.

Chestnut LG, Schwartz J, Savitz DA, Burchfiel CM. Pulmonary function and ambient particulate matter: epidemiological evidence from NHANES I. Arch Environ Health 1991;46:135-44.

Ciocco A, Thompson DJ. A Follow-up on Donora ten years after: methodology and findings. Am J Public Health 1961; 51:155-64.

Dassen W, Brunekreef B, Hoek G, et al. Decline in children's pulmonary function during an air pollution episode. J Air Pollut Control Assoc 1986;36:1223-27.

Delfino RJ, Becklake MR, Hanley JA. Reliability of hospital data for population-based studies of air pollution. Arch Environ Health 1993;48:140-46.

Dockery DW, Damokash AI, Neas LM, Raizenne M, Spengler JD, Koutrakis P, Ware JH, Speizer FE. Health effects of acid aerosols on North American children: Respiratory symptoms and aillness. Presented at the American Lung Association/American Thoracic Society International Conference, San Francisco, California, May 16-19, 1993.

Dockery DW, Pope CA III. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 1994;15:107-32.

Dockery DW, Pope CA III, Xu X, et al. Mortality risks of air pollution: a prospective cohort study. N Engl J Med 1993;329:1753-59.

Dockery DW, Schwartz J, Spengler JD. Air Pollution and daily mortality: associations with particulates and acid aerosols. Environ Res 1992;59:362-73.

Dockery DW, Speizer FE, Stram DO, Ware JH, Spengler JD, Ferris BG Jr. Effects of inhalable particles on respiratory health of children. Am Rev Respir Dis 1989;139:587-94.

Dockery DW, Ware JH, Ferris BG Jr, Speizer FE, Cook NR, Herman SM. Change in pulmonary function in children associated with air pollution episodes. J Air Pollut Control Assoc 1982;32:937-42.

Ellison JM, Waller RE. A review of sulphur oxides and particulate matter as air pollutants with particular reference to effects on health in the United Kingdom. Environ Res 1978;16:302-25.

Euler GL, Abbey DE, Magie AR, Hodgkin JE. Chronic obstructive pulmonary disease symptom effects of long-term cumulative exposure to ambient levels of total suspended particulates and sulfur dioxide in California Seventh-Day Adventist Residents. Arch Environ Health 1987;42:213-22.

Evans JS, Tosteson T, Kinney PL. Cross-sectional mortality studies and air pollution risk assessment. Environ Int 1984;10:55-83.

Fairley D. The relationship of daily mortality to suspended particulates in Santa Clara County, 1980-1986. Environ Health Perspect 1990;89:159-68.

Firket J. The cause of the symptoms found in the Meuse Valley during the fog of December, 1930. Bull Acad Roy Med Belg 1931;11:683-741.

Hoek G, Brunekreef B. Acute effects of a winter air pollution episode on pulmonary function and respiratory symptoms of children. Arch Environ Health 1993;48:328-35.

Hoek G, Brunekreef B. Effects of low level winter air pollution concentrations on respiratory health of Dutch children. Environ Res 1994 (in press).

Holland WW, Bennett AE, Cameron IR, et al. Health effects of particulate pollution: reappraising the evidence. Am J Epidemiol 1979;110:525-659.

Holland WW, Reid DD. The urban factor in chronic bronchitis. Lancet 1965;1:445-48;

Hill, Sir Austin B. The environment and disease: association or causation? Proc Roy Soc Med 1965;58:295-300.

Ito K, Thurston GD, Hayes C, Lippman M. Associations of London, England, daily mortality with particulate matter, sulfur dioxide, and acidic aerosol pollution. Arch Environ Health 1993;48:213-20.

Johnson KG, Gideon RA, Loftsgaarden DO. Montana air pollution study: children's health effects. J Off Stat 1990;5:391-408.

Johnson KG, Loftsgaarden DO, Gideon RA. The effects of Mount St. Helens volcanic ash on the pulmonary function of 120 elementary school children. Am Rev Respir Dis 1982;126:1066-69.

Kinney PL, Ozkaynak H. Associations of daily mortality and air pollution in Los Angeles County. Environ Res 1991;54:99-120.

Kinney PL, Ware JH, Spengler JD, Dockery DW, Speizer FE, Ferris BG Jr. Short-term pulmonary function change in association with ozone levels. Am Rev Respir Dis 1989;139:56-61.

Koenig JQ, Larson TV, Hanley QS, Rebolledo V, Dumler K, et al. Pulmonary function changes in children associated with fine particulate matter. Environ Res 1993;63:26-38.

Krupnick AJ, Harrington W, Ostro B. Ambient ozone and acute health effects: evidence from daily data. J Environ Econ Manage 1990;18:1-8.

Lave LB, Seskin EP. Air pollution and human health. Science 1970;169:723-33.

Lebowitz MD, Holberg CJ, Boyer B, Hayes C. Respiratory symptoms and peak flow associated with indoor and outdoor air pollutants in the southwest. J Air Pollut Control Assoc 1985;35:1154-58.

Lipfert FW. Air pollution and mortality: specification searches using SMSA-based data. J Environ Econ Manage 1984;11:208-243.

Lipfert FW, Hammerstrom T. Temporal patterns in air pollution and hospital admissions. Environ Res 1992;59:374-99.

Lipfert FW, Malone RG, Daum ML, Mendell NR, Yang CC. A Statistical Study of the Macroepidemiology of Air Pollution and Total Mortality. Prepared for the Office of Environmental Analysis United States Department of Energy. 1988.

Logan WPD. Mortality in London fog incident. Lancet 1953; 1:336-38.

Lutz LJ. Health effects of air pollution measured by outpatient visits. J Fam Pract 1983;16:307-13.

Martin AE. Mortality and morbidity statistics and air pollution. Proc Roy Soc Med 1964;57:969-75.

Mazumdar S, Sussman N. Relationships of air pollution to health: results from the Pittsburgh study. Arch Environ Health 1983;38:17-24.

Mendelsohn R, Orcutt G. An empirical analysis of air pollution dose-response curves. J Environ Econ Manage 1979;6:85-106.

Neas LM, Dockery DW, Spengler JD, Speizer FE, Tollerud DJ. The association of ambient air pollution with twice daily peak expiratory flow measurements in children. Am Rev Respir Dis 1992;145(4):A429.

Ostro BD. The effects of air pollution on work loss and morbidity. J Environ Econ Manage 1983;10:371-82.

Ostro BD. A search for a threshold in the relationship of air pollution to mortality: a reanalysis of data on London winters. Environ Health Perspect 1984;58:397-99.

Ostro BD. Air pollution and morbidity revisited: a specification test. J Environ Econ Manage 1987;14:87-98.

Ostro BD. Associations between morbidity and alternative measures of particulate matter. Risk Analysis 1990;10:421-27.

Ostro BD. The association of air pollution and mortality: examining the case for inference. Arch Environ Health 1993; 48:336-42.

Ostro BD, Lipsett MJ, Mann JK, Krupnick A, Harrington W. Air pollution and respiratory morbidity among adults in Southern California. Am J Epidemiol 1993;137:691-700.

Ostro BD, Lipsett MJ, Wiener MB, Selner JC. Asthmatic response to airborne acid aerosols. Am J Pub Health 1991;81:694-702.

Ostro BD, Rothschild S. Air pollution and acute respiratory morbidity: an observational study of multiple pollutants. Environ Res 1989;50:238-47.

Ozkaynak H, Spengler JD. Analysis of health effects resulting from population exposures to acid precipitation precursors. Environ Health Perspect 1985;63:45-55.

Ozkaynak H, Thurston GD. Associations between 1980 U.S. mortality rates and alternative measures of airborne particle concentration. Risk Analysis 1987;7:449-61.

Ponka A. Asthma and low level air pollution in Helsinki. Arch Environ Health 1991;46:262-70.

Pope CA III. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. Am J Pub Health 1989;79:623-28.

Pope CA III. Respiratory hospital admissions associated with PM_{10} pollution in Utah, Salt Lake, and Cache Valleys. Arch Environ Health 1991;46:90-7.

Pope CA III, Bates DV, Raizenne ME. Health effects of particulate air pollution: time for reassessment? Am J Pub Health 1994 (in review).

Pope CA III, Dockery DW. Acute health effects of PM_{10} pollution on symptomatic and asymptomatic children. Am Rev Respir Dis 1992;145:1123-28.

Pope CA III, Dockery DW, Spengler JD, Raizenne ME. Respiratory health and PM_{10} pollution: a daily time series analysis. Am Rev Respir Dis 1991;144:668-74.

Pope CA III, Kanner RE. Acute effects of PM_{10} pollution on pulmonary function of smokers with mild to moderate chronic obstructive pulmonary disease. Am Rev Respir Dis 1993;147;1336-40.

Pope CA III, Schwartz J, Ransom MR. Daily mortality and PM_{10} pollution in Utah Valley. Arch Environ Health 1992;47:211-17.

Pope CA III, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. JAMA 1994 (in review).

Portney PR, Mullahy J. Urban air quality and chronic respiratory disease. Regional Sci Urban Econ 1990;20:407-18.

Raizenne ME, Burnett RT, Stern B, Franklin CA, Spengler JD. Acute lung function responses to ambient acid aerosol exposures in children. Environ Health Perspect 1989;79:179-85.

Ransom MR, Pope CA III. Elementary school absences and PM_{10} pollution in Utah Valley. Environ Res 1992;58:204-19.

Roemer W, Hoek G, Brunekreef B. Effect of ambient winter air pollution on respiratory health of children with chronic respiratory symptoms. Am Rev Respir Dis 1993;147:118-24.

Saldiva PHN, Pope CA III, Schwartz J, Dockery DW, Lichtenfels AJ, et al. Air pollution and mortality in elderly people: a time series study in Sao Paulo, Brazil. Arch Environ Health 1994 (in review).

Samet JM, Speizer FE, Bishop Y, Spengler JD, Ferris BG Jr. The relationship between air pollution and emergency room visits in an industrial community. J Air Pollut Control Assoc 1981;31:236-40.

Schwartz J. Lung function and chronic exposure to air pollution: a cross-sectional analysis of NHANES II. Environ Res 1989;50:309-21.

Schwartz J. Particulate air pollution and daily mortality in Detroit. Environ Res 1991;56:204-13.

Schwartz J. Air pollution and daily mortality in Birmingham, AL. Am J Epidemiol 1993;137:1136-47.

Schwartz J. Air pollution and hospital admissions for the elderly in Birmingham, Al. Am J Epidemiol 1994a (in press).

Schwartz J. Particulate air pollution and chronic respiratory disease. Environ Res 1993;62:7-13.

Schwartz J. What are people dying of on high air pollution days? Environ Res 1994c;64:26-35.

Schwartz J. PM_{10}, ozone, and hospital admissions for the elderly in Minneapolis-St.Paul. Arch Environ Health 1994d (in press).

Schwartz J. Air pollution and daily mortality: a review and meta-analysis. Environ Res 1994e;64:36-52.

Schwartz J. Total suspended particulate matter and daily mortality in Cincinnati, Ohio. Environ Res 1994f (in press).

Schwartz J, Dockery DW. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 1992;145:600-04.

Schwartz J, Dockery DW. Particulate air pollution and daily mortality in Steubenville, Ohio. Am J Epidemiol 1992;135:12-19.

Schwartz J, Dockery DW, Neas LM, et al. Acute effects of summer air pollution on respiratory symptom reporting in children. Am Rev Respir Dis 1994 (in press).

Schwartz J, Marcus A. Mortality and air pollution in London: a time series analysis. Am J Epidemiol 1990;131:185-94.

Schwartz J, Slater D, Larson TV, Pierson WE, Koenig JQ. Particulate air pollution and hospital emergency visits for asthma in Seattle. Am Rev Resp Dis 1993;147:826-31.

Schwartz J, Spix C, Wichmann HE, Malin E. Air pollution and acute respiratory illness in five German communities. Environ Res 1991;56:1-14.

Shumway RH, Azari AS, Pawitan Y. Modeling mortality fluctuations in Los Angeles as functions of pollution and weather effects. Environ Res 1988;45:224-41.

Shy CM. Epidemiologic evidence and the United States air quality standards. Am J Epidemiol 1979;110:661-71.

Speizer FE. Studies of acid aerosols in six cities and in a new multi-city investigation: design issues. Environ Health Perspect 1989;79:61-7.

Sunyer J, Anto JM, Murillo C, Saez M. Effects of urban pollution on emergency room admissions for chronic obstructive pulmonary disease. Am J Epidemiol 1991;134:277-89.

Sunyer J, Saez M, Murillo C, Castellsague J, Martinez F, Anto JM. Air pollution and emergency room admissions for chronic obstructive pulmonary disease: A 5-year study. Am J Epidemiol 1993;137:701-5.

Thurston GD, Ito K, Kinney PL, Lippmann M. A multi-year study of air pollution and respiratory hospital admissions in three New York state metropolitan areas: results for 1988 and 1989 summers. J Expos Anal Environ Epidemiol 1992;2:429-50.

Thurston GD, Ito K, Lippmann M. The role of particulate mass vs. acidity in the sulfate-respiratory hospital admissions association. Preprint #93-11.03. Presented at Annu. Meet. Air Waste Manage. Assoc., 86th, June 13-18. Denver, CO.

Utell MJ, Samet JM. Particulate air pollution and health. New evidence on an old problem. Am Rev Respir Dis 1993;147:1334-35.

Vedal S, Manna B. Adverse respiratory health effects of ambient inhalable particle exposure. Paper 91-171.3, Air & Waste Management Association, 1991.

Ware JH, Ferris BG Jr, Dockery DW, Spengler JD, Stram DO, Speizer FE. Effects of ambient sulfur oxides and suspended particles on respiratory health of preadolescent children. Am Rev Respir Dis 1986;133:834-42.

Ware JH, Thibodeau LA, Speizer FE, Colome S, Ferris BG Jr. Assessment of the health effects of atmospheric sulfur oxides and particulate matter: evidence from observational studies. Environ Health Perspect 1981;41:255-76.

Whittemore AS, Korn EL. Asthma and air pollution in the Los Angeles Area. Am J Pub Health 1980;70:687-96.

Wichmann HE, Mueller W, Allhoff P, et al. Health effects during a smog episode in West Germany in 1985. Environ Health Perspect 1989;79:89-99.

Wyzga RE. The effect of air pollution upon mortality: a consideration of distributed lag models. J Am Stat Assoc 1978;73:463-72.

Xu X, Dockery DW, Wang L. Effects of air pollution on adult pulmonary function. Arch Environ Health 1991;46:198-206.

Long-Term Ambient Concentrations of Particulates and Oxidants and Development of Chronic Disease in a Cohort of Non-Smoking California Residents.

David E. Abbey
Michael D. Lebowitz
Paul K. Mills
Floyd F. Petersen
W. Lawrence Beeson
Raoul J. Burchette.

Correspondence to:
David E. Abbey, Ph.D.
Loma Linda University
CHR, Evans Hall, Rm 204
Loma Linda, CA 92350

Acknowledgements: We wish to acknowledge the helpful assistance of the California Air Resources Board staff, Dane Westerdahl, Andy Alexis and others; also the assistance of Beverley Stocker and Nona Grove in conducting the word processing for the manuscript. Supported by Grant #A933-160.

The statements and conclusions in this report are those of the authors and not necessarily those of the California Air Resources Board.

ABSTRACT

A cohort of 6340 non-smoking California Seventh-day Adventists (SDAs) who had resided within 5 miles of their present residence for the past 10 years have been followed since 1977 for: incidence of cancer and myocardial infarction (M.I.) through 1982, development of definite symptoms of, and increasing severity of, airway obstructive disease (AOD), chronic bronchitis, and asthma through 1987, and all natural cause mortality through 1987. Cumulative ambient concentrations of specific pollutants have been estimated for study participants from 1967 to 1987 by interpolating monthly statistics from statewide air monitoring stations to zip codes of residence and work location. Statistics include excess concentrations and exceedance frequencies above a number of cutoffs as well as mean ambient concentration and mean ambient concentration adjusted for time spent indoors. Indoor sources or nitrogen (NO_2), and of particulate pollution such as environmental tobacco smoke, both at home and at work, as well as occupational dusts and fumes, have been adjusted for in multivariate statistical models. Particulates included total suspended particulates (TSP), monitored from 1973-1987; inhalable particulates less than 10 microns in diameter (PM10), estimated from site/seasonal specific regressions on TSP for 1973-1987; fine particulates less than 2.5 microns in diameter estimated from airport visibility data for 1967-1987; and suspended sulfates (SO_4), monitored from 1977-1987. A direct measure of visibility, and gaseous pollutants--ozone, sulfur dioxide (SO_2), and (NO_2) monitored from 1973-1987 were also included in analyses.

No statistically significant associations between any of the disease outcomes studied and NO_2 or SO_2 were found in this cohort. None of the pollutants studied showed statistically

significant associations with all natural cause mortality or incidence of all malignant neoplasms in males. Statistically significant associations were observed between elevated ambient concentrations of one or more particulate pollutants and each of the other disease outcomes. In addition, ozone was significantly associated with increasing severity of asthma, and with the development of asthma in males. Multipollutant analyses indicated that none of the associations between particulate pollutants and disease outcomes were due to correlations with gaseous pollutants studied except possibly for PM2.5 and increasing severity of asthma, which could be due to a correlation with ozone. Observed associations between disease outcomes and PM2.5 or PM10 could be biased towards the null because of increased measurement error due to their indirect methods of estimation.

INTRODUCTION

In recent years prospective studies have begun to study the association of long-term ambient concentrations of air pollutants and development of chronic disease. (Comstock, 1973; Rokaw, 1980; Detels, 1987; Dockery, 1989; EPA, 1984, 1986; Ferris, 1973,1976; Jedrychowsky, 1988; Krzyzanowski, 1990; Lebowitz, 1993; Van der Landa, 1981; Carrozzi, 1993). Most studies have dealt with respiratory disease, though Dockery et al, (1993) studied all natural cause mortality. Few studies have been able to address the effects of many different pollutants on a wide range of disease outcomes. A cohort of 6,340 non-smoking Seventh-day Adventists (SDAs) who had resided within 5 miles of their 1976 residence for the past 10 years have been followed since 1977 to ascertain the incidence of all malignant neoplasms and myocardial infarction (MI) through 1982, the development of definite symptoms of airway obstructive disease (AOD), chronic bronchitis, and asthma through 1987, and all natural cause mortality through 1986. Cumulative ambient concentrations of five specific particulate air, and three gaseous, pollutants have been estimated for study participants from 1967 to 1987 by interpolating monthly statistics from air monitoring stations statewide to zip codes of residence and work location. Previous papers (Abbey, 1989, 1991a,b; 1993a,b,c; 1994a,b,c; Euler, 1987; 1988, Hodgkin, 1984; Mills, 1991; 1993a,b;) have reported on associations between ambient concentrations of air pollutants and development of different chronic diseases in this cohort and compared the findings with other studies. The effects of occupational exposures (Greer, 1993) and environmental tobacco smoke (Robbins, 1993) have also been addressed.

The purpose of this paper is to first give an overview of the pollutant specific analyses reported in previous papers and then report on multipollutant analyses which were conducted to determine which pollutant was most strongly associated with each disease, and which pollutants might only be indirectly associated with the disease because of a high correlation with the most strongly related pollutant. This was termed a surrogate relationship.

Epidemiological Methods

In April, 1977, 6,340 non-smoking, non-Hispanic white Seventh-day Adventist residents of California, who were members of the National Cancer Institute funded Adventist Health Study, were enrolled in a study to ascertain long-term chronic health effects of ambient air pollutants. To be included in the study participants must have lived for 10 years or longer within five miles of their present neighborhood and live in one of three air basins: San Francisco, South Coast (Los Angeles and Eastward), San Diego or be selected by a random sample from the rest of California. Age in 1977 ranged from 27 to 95, 64% were female. More descriptive characteristics are given in previous papers (Euler, 1987; Mills, 1991; Abbey, 1993). Incidence of cancer and myocardial infarction (MI) was ascertained on the cohort from 1977 through 1982. Subjects completed a standardized respiratory symptoms questionnaire (now part of the American Thoracic Society Questionnaire) in April, 1977 and, those still surviving, again in April, 1987. Three thousand, nine hundred and fourteen study participants who completed both the 1977 and 1987 respiratory symptoms questionnaires are termed the Respiratory Symptoms Cohort. Computer algorithms were used in 1977 and again in 1987 to classify each individual as to whether or not they had definite, possible, or no symptoms of

AOD, chronic bronchitis, or asthma (Hodgkin, 1984). In addition to this, computer algorithms scored the severity of symptoms for each of these respiratory complexes so that change in severity of symptoms could be measured between 1977 and 1987. Complete details of the algorithms used for scoring and classification are given by Abbey, (1994c).

Methods for Estimation of Ambient Air Pollution Concentrations

Ambient air pollution concentrations were estimated for study participants since 1966 by interpolating monthly statistics from the statewide network of air monitoring stations (Abbey, 1989; Abbey, 1991b). Monthly statistics included mean concentrations of air pollutants, exceedance frequencies, and excess concentrations above a number of cutoffs. Exceedance frequencies were defined as the number of days on which particulate pollutants exceeded different cutoff levels or number of hours in which gaseous pollutants exceeded different cutoffs. Excess concentration statistics were the sum of all concentrations above a cutoff. Correlations between exceedance frequency and excess concentration statistics were close to one, hence, results are summarized only for mean concentration and exceedance frequency statistics. The cutoffs used for each pollutant were: total suspended particulates (TSP) - 60, 75, 100, 150, and 200 $\mu g/m^3$; particulates less than 10 microns in diameter (PM10) - 40, 50, 60, 80, and 100 $\mu g/m^3$; particulates less than 2.5 microns in diameter (PM2.5) - 20, 30, and 40 $\mu g/m^3$; suspended sulfates (SO_4) - 6, 9, 12, and 15 $\mu g/m^3$; Ozone - 10, 12, 15, 20, and 25 pphm; nitrogen dioxide (NO_2) - 5, 15, 20, and 25 pphm; sulfur dioxide (SO_2) - 2, 4, 5, 8, and 14 pphm. The monthly statistics were interpolated from the network of air monitoring stations from which data is collected by the California Air Resources Board (CARB) to zip code centroids of

both work and home location of study participants (Abbey, 1991b). Interpolations utilized meteorological and geographic barriers. Study participants were assumed to be at work locations on working days, between the hours of 8:00 A.M., and 5:00 P.M., otherwise at home locations.

Pollutants for which estimates were available for 1966 - 1987 were -- TSP, ozone, SO_2, NO_2. Suspended sulfates -- SO_4 were available from 1977-1987. Since insufficient monitored PM10 data existed prior to 1985, it was indirectly estimated from TSP using site and seasonal specific regression equations (Abbey, 1994a). Though the indirect estimates tended to underestimate monitored values, the correlation between indirect estimates of PM10 interpolated to stations and monitored values at these stations was 0.86 when both were cumulated over the two year period of 1986 and 1987. Though not monitored on a statewide basis, sufficient PM2.5 data existed between 1979 and 1986 to allow the formation of regional and seasonal specific regression equations relating PM2.5 to airport visibility measures for nine airports scattered throughout California. These regression equations were used to estimate ambient concentrations of PM2.5 for 1966-1986 for approximately half of the cohort who lived in the vicinity of these nine airports (Abbey, 1994b). The multiple R^2 for the site/seasonal specific regression equations ranged from 0.28 to 0.83 with the average being 0.58. In addition to relating health effects to estimated ambient concentrations of PM2.5, derived from visibility, an average visibility measure was also calculated and related directly to health effects.

For each of the ambient pollution estimates except NO_2, an adjustment factor for indoor infiltration of ambient concentrations was obtained from work by Winer (1989). Using estimated hours spent indoors by study participants obtained by season in 1977 and again in 1987, these

adjustment factors were used to adjust ambient mean concentrations for time spent indoors. For NO_2 regression estimates of personal exposure were formed which used lifestyle and housing characteristics, indoor sources and ambient concentrations (Abbey, 1993c). Sensitivity analyses conducted using adjusted ambient mean concentrations concurred with those using unadjusted ambient mean concentrations. Table 1 gives correlations between estimated mean 1977-1987 concentrations of ambient pollutants for members of the Respiratory Symptoms Subcohort.

Statistical Methods

For development of AOD, chronic bronchitis, and asthma, separate multiple logistic regression models were used for studying associations between long-term cumulative ambient concentrations of air pollution and new cases of definite symptoms in 1987 among individuals who did not have definite symptoms in 1977. For each disease individuals having possible symptoms of that disease in 1977 were included in such analyses and a covariate was included to indicate this. Sensitivity analyses were conducted excluding this covariate. For SO_4 which was monitored only between 1977 and 1987, this covariate was not included in the main analyses but a sensitivity analyses was conducted including it.

Multiple linear regressions were used to study associations between change in the severity score for each of the three diseases and long-term ambient concentrations of air pollution. All individuals were included in these analyses with a positive value of change in score indicating an increase in severity of symptoms. For each disease, the 1977 symptoms severity score for that disease was used as a covariate in analyses with the exception, again, of SO_4.

For both multiple logistic and multiple linear regressions, gender, age, and education were forced into all models. Stepwise selection procedures were used to select statistically significant covariates ($p < .05$) from a large number of candidate covariates which included - years smoked prior to 1977,[1] years lived with a smoker, years worked with a smoker, years of dust exposure at work for models relating health effects to particulate pollutants, years of fume exposure at work for models relating health effects to gaseous pollutants, frequency of childhood colds, and definite symptoms of AOD before age of 16. Unless otherwise stated the term "statistically significant" will refer to the 2-tailed .05 level of statistical significance. The .05 level of statistical significance has not been used as an absolute criteria. Strong disease/pollutant associations which did not achieve the .05 level of statistical significance (possibly due to a small number of incident cases and lack of statistical power) will be described in the results below.

For incidence of MI, cancer, and all natural cause mortality Cox proportional hazards regression models were used. Further details on the epidemiological and statistical methods used are given by Abbey (1993a) and Mills (1991).

For each health outcome, multipollutant analyses compared different pollutants on the strength of association achieved by each pollutant in multivariate health effects regression models. The scale/invariant measure of strength of association used was the estimated regression coefficient for the pollutant divided by its standard error, or equivalently the computed level of statistical significance of the regression coefficient. Comparisons between pollutants were made for the mean concentration index as well as the most statistically significant of the exceedance frequency or indices. For many outcomes only one or two pollutants were associated with the

health outcome and it was clear which pollutant showed the strongest relationship. For other pollutants stepwise regression procedures were used to allow mean concentrations of the different pollutants to compete for entry in the model, if mean concentration for one of the pollutants was significantly associated with the health outcome. If this was not the case the most statistically significant of the exceedance frequency statistics for each pollutant were allowed to compete for entry in the model. In both situations gender, age, and education, and the covariates found to be significant from the single pollutant analyses were first forced into the model; then the two pollutants were allowed to compete for entry. Comparisons between TSP, ozone, SO_4, and PM10 were made on the entire cohort. Comparisons with each of these pollutants and PM2.5 were restricted to the subcohort living in the vicinity of airports for which PM2.5 estimates were available.

RESULTS

Table 2 gives relative risks and 95% confidence intervals (CIs) for selected increments of ambient concentrations. Increments for either an exceedance frequency index or mean concentration were chosen to be within the range experienced by the cohort. For TSP, PM10, SO_4, and ozone, the increment was expressed in terms of the most statistically significant of the various indices -- mean concentration or exceedance frequencies above the different cutoffs. For PM2.5 the increment was expressed in terms of mean concentrations since it was felt that exceedance frequencies were not reliable because of discreetness due to estimating PM2.5 from visibility. Airport visibility is measured using sightings of markers at a few discreet distances from airport control towers. Some cells in Table 2 contain the entry "not done." Possible

associations between SO_4 and cancer incidence were not analyzed since SO_4 was monitored only since 1977 and it was felt that this did not allow sufficient lag time for cancer incidence. Statistical analyses to study associations of PM2.5 and visibility with respiratory cancer incidence and incidence of MI were not conducted due to an insufficient number of study participants living in the vicinity of airports for which airport visibility measures and PM2.5 estimates were available.

For the prospective data collected since 1977, no statistically significant associations were seen between any of the health outcomes studied and SO_2 or NO_2 (Abbey, 1991a; 1993a,c Mills, 1993c). This result remained true when ambient concentrations were adjusted to reflect time spent indoors and when the regression estimates of personal exposure for NO_2 were used. No associations were observed between long term cumulations of any of the ambient pollutants studied and incidence of all malignant neoplasms in males, or all natural cause mortality. Relative risks and confidence intervals for these two outcomes as associated with elevated TSP and ozone have been reported previously. (Abbey, 1991a; 1993c, Mills 1993c) Relative risks for these two health outcomes for all pollutants studied were close to, or less than, one. The null relative risk of one was well within all 95% confidence intervals. Incidence of MI was significantly associated only with exceedance frequencies above 6 $\mu g/m^3$ of SO_4 (Abbey, 1993b). No associations were observed between visibility and health outcomes except for mean visibility and increasing severity of asthma symptoms and mean visibility and incidence of all malignant neoplasms in females. These observed associations with visibility, however, were not as

statistically significant as those with PM2.5 and were likely the result of the correlation between PM2.5 and visibility.

Table 3 gives the cutoffs for which exceedance frequencies show statistical significance and the results of the multipollutant analyses. Pollutants in parenthesis in Table 3 indicated that the observed association with a health outcome could be due instead to an association of the pollutant in parenthesis and the health outcome, as the pollutant in parenthesis showed a stronger association.

TSP showed the strongest associations of any of the pollutants with development of AOD, and chronic bronchitis, as well as increasing severity of symptoms for AOD, and chronic bronchitis. For asthma, TSP, ozone, and SO_4, were close competitors, (see ozone and SO_4 sections below). TSP was statistically significantly associated with incidence of all malignant neoplasms in females, but not males. When analyses were restricted to the airport sub-cohort for which estimates of PM2.5 were available, TSP remained statistically significantly associated with incidence of all malignant neoplasms in females but the association was not as strong as for PM2.5. From Table 3 it can been seen that ambient concentrations of TSP in excess of 60 $\mu g/m^3$ were associated with increasing severity of AOD symptoms. Ambient concentrations of TSP in excess of 100 $\mu g/m^3$ were associated with many of the health outcomes.

Ozone was highly significantly associated with development of asthma in males but not females. This relationship does not appear to be a surrogate relationship. The relative risk for a 1 pphm increase in mean concentration of ozone was 3.12, (95% CI: 1.61, 5.85), (Greer, 1993). In non-gender specific analyses exceedance above 10 pphm of ozone showed a trend

association ($p = .056$) with development of asthma. Ozone was more strongly associated with development of asthma than TSP when mean concentrations of the two pollutants were allowed to compete for entry in the models but was less strongly related when the most significant exceedance frequency statistic of each pollutant was used, i.e., TSP in excess of 200 $\mu g/m^3$ showed a stronger association with development of asthma than ozone in excess of 10 pphm. Mean concentration of ozone and exceedance frequencies for the two lowest cutoffs, 10 pphm and 12 pphm, show statistical significance when associated with increasing severity of asthma. For increasing severity of asthma, stepwise linear regression showed that mean concentration of ozone enters in preference to mean concentration of TSP. However, in another stepwise multiple linear regression; the F to enter for exceedance frequencies in excess of 200 $\mu g/m^3$ TSP was tied with the F to enter for exceedance frequencies in excess of 10 pphm of ozone. Ozone showed a strong trend association with respiratory cancer incidence. Due to the small number of incident cases, only 17, this association was not statistically significant. Future research will investigate lower cutoffs of ozone as well as 8-hour averages.

$\underline{SO_4}$ was statistically significantly associated with development of asthma and this association was stronger than that for TSP or ozone when the covariate, "possible symptoms in 1977" was not included in the model. Analyses based on the entire respiratory symptoms cohort indicated that the relationship between SO_4 and increasing severity of AOD symptoms may instead be due to TSP and the correlation between SO_4 and TSP. When analyses were restricted to the airport sub-cohort, however, SO_4 was more strongly associated with increasing severity of AOD symptoms than any of the other pollutants including TSP.

PM10 cannot be fairly compared with other pollutants due to the fact that PM10 was indirectly estimated from TSP. PM10 showed strong statistically significant associations with development of AOD and chronic bronchitis which were second only to those observed for TSP. If PM10 were not indirectly estimated from TSP, it may have shown the strongest relationships. PM10 showed a statistically significant association with increasing severity of asthma symptoms, as did ozone. Multipollutant analyses of PM10 versus ozone were not conducted because of PM10 being indirectly estimated from TSP. Ambient concentrations in excess of 40 $\mu g/m^3$ of PM10 showed statistical significance with increasing severity of symptoms and ambient concentration in excess of 80 $\mu g/m^3$ showed statistical significance for development of symptoms.

PM2.5 was indirectly estimated from visibility. Thus comparisons of strengths of association with other pollutants cannot be regarded as conclusive. Exceedance frequencies in excess of 20 $\mu g/m^3$ of PM2.5 were statistically significantly associated with development of chronic bronchitis. The strength of the association for mean concentration of PM2.5 was not as high as TSP or PM10 when analyses were restricted to the airport subcohort for which PM2.5 estimates were available. PM2.5 was significantly associated with increasing severity of AOD symptoms but this association was not as strong as that for TSP or SO_4 when analyses for these two pollutants were restricted to the airport subcohort. SO_4 showed the strongest association for the airport subcohort. PM2.5 was significantly associated with increasing severity of chronic bronchitis, however this association was not as strong as for TSP when analyses were restricted to the airport subcohort. PM2.5 was significantly related to increasing severity of asthma

symptoms. However, PM10 and ozone both showed stronger associations with this health outcome when analyses were restricted to the airport subcohort. PM2.5 was significantly associated with incidence of all malignant neoplasms in females. This association was stronger than for TSP when analyses were restricted to the airport subcohort. Ambient mean concentrations of PM2.5 as well as exceedance frequencies above 20 $\mu g/m^3$ were significantly associated with health outcomes.

DISCUSSION AND CONCLUSIONS

The limitations of this study have been thoroughly discussed in previous papers (Mills, 1991; Abbey, 1993a; 1993b; 1993c; Greer, 1993). Space does not permit reiteration of these limitations. The effects of TSP, PM10, and PM2.5 cannot be truly separated in this study due to the fact that the ambient concentrations of PM10 and PM2.5 were indirectly estimated whereas those for TSP were monitored, and also because of the high correlations between them (see Table 1). The increased measurement error for PM10 and PM2.5 due to indirect estimation could bias observed associations thus making it impossible to make fair comparisons with other pollutants.

Because of the large number of associations studied, some of our statistically significant associations could have occurred by chance. On the other hand air pollutant/disease associations that really exist may have failed to show statistical significance in this study due to lack of statistical power. This could be true for diseases such as incidence or respiratory cancer, where the size of our cohort and length of follow-up are inadequate to result in a sufficient number of incident cases. For the above reasons the results displayed in Table 3 should be interpreted with

caution. The primary purpose of Table 3 is to show that statistically significant associations between disease outcomes and particulate pollutants have been observed in this cohort at ambient levels in excess of ambient concentrations which frequently occur in urban areas of California, and that these associations appear not to be surrogate relationships solely due to correlations with gaseous pollutants, which were monitored during the time period of this study.

Estimated ambient concentrations of pollutants for study participants were averaged over different time periods before being used in analyses. The stepwise selection procedures used in forming the regression models were allowed to select the most statistically significant time period. The previous papers from this study show which time period for each pollutant was most significantly related to each health outcome. Except for SO_4 and PM2.5 the time period 1973-1977 was used as a surrogate for the time period 1966-1977 as many more monitoring stations were available for the latter time period. SO_4 was monitored only since 1977. PM2.5 was estimated from airport visibility data since 1966. Sensitivity analyses showed that results using either time period agreed closely. Cox proportional hazards regression models, which were used for incidence and mortality, where date of event was available, used an average of the pollutant from 1973 through data of event or risk set (time dependent models) as well as the average from 1973 through March 1977, (fixed time models). Sensitivity analyses indicated concurrence between the two types of models.

Our respiratory symptoms complex of chronic bronchitis included chronic productive cough (with sputum) as well as cough only. Development of definite symptoms of chronic bronchitis was statistically significantly associated with elevated ambient concentrations of TSP,

PM10, and PM2.5. Separate analyses which were conducted for development of chronic productive cough as associated with each of these particulates showed that the level of statistical significance remained the same or increased while that for development of definite symptoms of "cough only" decreased to a little below the (.05) level for statistical significance. For chronic productive cough the regression coefficient increased and for "cough only" the regression coefficient decreased as compared to the respective regression coefficients for development of overall symptoms of chronic bronchitis. This was true for each of TSP, PM10, and PM2.5.

No associations were observed between SO_2 and any of the health outcomes from the prospective data. An older phase of the study had shown associations between SO_2 and prevalence of definite AOD symptoms in 1977 (Euler, 1987), but multipollutant analyses indicated that the relationship may have been due to TSP (Euler, 1988). The lack of association between SO_2 and any of the prospective disease outcomes studied may be due to lower levels of SO_2 experienced by our cohort than in other studies where health effects have been observed (Mills, 1991; Abbey, 1993a). The lack of association between NO_2 and health effects seen in adults in this study does not negate associations seen in children (Abbey, 1993c; Shy, 1970; Dawson, 1979; Love, 1982; Neas, 1991). Our finding of no association between all natural cause mortality and long term concentrations for any of the ambient pollutants studied is inconsistent with an association seen between mortality and particulates in the Six Cities Study by Dockery (1993). Survival curves for their cohort indicated the greatest differences between

cities in years 8-16 of follow-up. We only had 10 years of mortality follow-up available for present analyses on our cohort. This may explain the discrepancy.

We observed an elevated risk of all malignant neoplasms in females but not in males as associated with TSP and PM2.5. A similar result has been observed in an epidemiological study in Israel (Biger, personal communication, 1991). Experiments in rats have shown that particulate matter in diesel exhaust fumes have a greater impact on lung tumor formation in females than males especially at the highest concentration levels (Mauderly, 1987). Fine particulates often contain the most toxic compounds (e.g., trace metals, acid sulfates, organics, etc.) (Özkaynak, 1987). Females have been shown to have a greater deposition fraction than males of fine particles in the lungs (Kim, 1994).

The increased risk of respiratory cancer associated with elevated ozone in our cohort is consistent with some animal studies. Dillon (1992) showed that ozone was mutagenic in Salmonella. Ozone has been shown to be carcinogenic in mice though not in rats. (Hassett, 1985; Last, 1987; NTP Technical Report, 1993). Our findings of elevated risks of asthma and respiratory cancer as associated with ozone is consistent with a significantly increased standardized incidence ratio of lung cancer observed in asthmatics in Finland (Vesterinin, 1993), and a very slightly elevated, but not statistically significant standardized morbidity ratio found for respiratory tract cancer in asthmatics in Sweden (Källén, 1993).

Ozone was significantly associated with development of asthma in males but not in females. T-test between males and females in our cohort showed that males were outdoors significantly more than females during the high ozone season of June through September (Greer,

1993). Ozone is highly volatile and not stable in the indoor environment. This may explain the gender discrepancy with respect to development of asthma.

We found ozone to be associated with asthma and asthmatic symptoms but not chronic bronchitis. This pattern of relationships has been observed in other studies (von Mutius et al., 1992; Viegi et al., 1991; Krzyzanowski et al, 1990; Lippmann, 1989). Chronic bronchitis and chronic obstructive pulmonary disease (COPD) appear to be associated much more with reducing type (PM-SO_x) atmospheres (Sunyer et al., 1993; Lebowitz, 1983; Ware, 1980; ATS, 1978; Colley & Holland, 1967). Our findings that TSP and PM10 are more strongly associated with development of and an increasing severity of symptoms of overall AOD and chronic bronchitis than other pollutants are consistent with these findings. However, in our study SO_4 was strongly associated with development of asthma as well as increasing severity of AOD.

Lung function parameters such as post bronchodilator response as well as lability indices from home peak flow diary data have recently been collected on 1,500 members of our cohort. An additional 10 years of cancer incidence data is being collected on the entire cohort as well as an additional six years of mortality. Ambient air pollution indices are being updated on the cohort through 1993. These updated indices will include monitored PM10, available on a statewide basis since 1987. New indices for ozone are being computed which include 8 hour averages and exceedance frequencies and excess concentrations for two new lower cutoffs, 6 and 8 pphm. Analyses of these new data may provide answers to some of the questions raised from current findings.

REFERENCES

Abbey, D.E., Euler, G.L., Moore, J.K., Petersen, F., Hodgkin, J.E., and Magie, A.R. (1989), "Applications of a Method for Setting Air Quality Standards Based on Epidemiological Data," *Journal of Air & Waste Management Association*, 39(4), 437-445.

Abbey, D.E., Mills, P.K., Petersen, F., Beeson, W.L. (1991a), "Long Term Ambient Concentrations of Total Suspended Particulates and Oxidants as Related to Incidence of Chronic Disease in California Seventh-day Adventists," *Environmental Health Perspectives*, 94, 43-50.

Abbey, D.E., Moore, J., Petersen, F., Beeson, W.L. (1991), "Estimating Cumulative Ambient Concentrations of Air Pollutants: Description and Precision of Methods Used for an Epidemiological Study," *Archives of Environmental Health*, 46(5), 281-287.

Abbey, D.E., Petersen, F.F., Mills, P.K., Beeson, W.L. (1993a), "Long-Term Ambient Concentrations of Total Suspended Particulates, Ozone, and Sulfur Dioxide and Respiratory Symptoms in a Non-Smoking Population," *Archives of Environmental Health*, 48(1), 33-46.

Abbey, D.E., Petersen, F.F., Mills, P.K., Kittle, L. (1993b), "Chronic Respiratory Disease Associated With Long Term Ambient Concentrations of Sulfates and Other Air Pollutants," *Journal of Exposure Analysis and Environmental Epidemiology*, 3(S1), 99-115.

Abbey, D.E., Colome, S.D., Mills, P.K., Burchette, R, Beeson, W.L., Tian, Y. (1993c), "Chronic Disease Associated With Long Term Concentrations of Nitrogen Dioxide," *Journal of Exposure Analysis and Environmental Epidemiology*, 3(2), 181-202.

Abbey, D.E., Hwang, B.L., Burchette, R.J. (1994a), "Long-Term Ambient Concentrations of PM10 and Airway Obstructive Disease," *Archives Environmental Health*, In press

Abbey, D.E., Ostro, B.E., Petersen, F., Burchette, R.J., VanCurren, T., Fraser, G. (1994b), "Chronic Respiratory Disease Associated With Long-Term Ambient Concentrations of Fine Particulates Less than 2.5 Microns in Diameter (PM2.5) and Other Air Pollutants," *Archives Environmental Health*, Submitted for publication.

Abbey, D.E. (March, 1994c) "Incidence of Respiratory Symptoms and Chronic Disease in a Non-Smoking Population as a Function of Long-Term Cumulative Exposure to Ambient Air Pollutants (AHSMOG Study)," *CARB Final Report*, Contract Number A933-160.

ATS. (1978) "Statement on the Health Effects of Air Pollution," *American Thoracic Society*, New York.

Carrozzi, L. (1993), "Longitudinal Analysis of Respiratory Symptoms in a General Population Sample of Northern Italy," *European Respiratory Journal*, 6(S17), 408s.

Colley, J.R.T., Holland, W.W. (1967), "Social and Environmental Factors in Respiratory Diseases," *Archives of Environmental Health*, 14, 157.

Comstock, G.W., Stone, R.W., Sakai, Y., Matsuya, T., Tonascia, J.A. (1973), "Respiratory Findings and Urban Living," *Archives of Environmental Health*, 27, 143-150.

Dawson, S.V., and Schenker, M.B. (1979), "Health Effects of Inhalation of Ambient Concentrations of Nitrogen Dioxide," *American Review of Respiratory Disease*, 120, 281-292.

Detels, R., Taskin, D.P., Sayre, J.W., Rokaw, S.N., Coulson, A.H., Massey, F.J., Wegiman, D.H. (1987), The UCLA Population Studies of Chronic Obstructive Respiratory Disease. "Lung Function Changes Associated with Chronic Exposure to Photochemical Oxidants, a Cohort Study Among Never-Smokers," *Chest*, 92, 54-603.

Dillon, D., Combes, R., McConville, M., Zeiger, E. (1992), "Ozone is mutagenic in Salmonella," *Environmental Mol Mutagen*, 19, 331-337.

Dockery, D.W., Speizer, F.E., Stram, D.O., Ware, J.H., Spengler, J.D., Ferris, B.G. (1989), "Effects of Inhalable Particles on Respiratory Health of Children," *American Review of Respiratory Disease*, 139, 587-594.

Dockery, D.W., Pope, A. III., Xiping, X., Spengler, J.D., Ware, J.H., Fay, M.A., et al. (1993), "An Association Between Air Pollution and Mortality in Six U.S. Cities," *New England Journal of Medicine*, 329(24), 1753-1759.

EPA (1984), "Ozone and Other Photochemical Oxidants," *V.5, Air Quality Criteria*, Research Triangle Park, NC.

EPA (1982), "Particulate Matter and Sulfur Oxides," *Air Quality Criteria*, Research Triangle Park, NC. 1982; 2nd Addendum.

Euler, G.L., Abbey, D.E., Magie, A.R., Hodgkin, J.E. (July/August,1987), "Chronic Obstructive Pulmonary Disease Symptom Effects of Long-Term Cumulative Exposure

to Ambient Levels of Total Suspended Particulates and Sulfur Dioxide in California Seventh-Day Adventist Residents," *Archives of Environmental Health*, 42(4), 213-222.

Euler, G.L., Abbey, D.E., Hodgkin, J.E., Magie, A.R. (July/August, 1988), "Chronic Obstructive Pulmonary Disease Symptom Effects of Long-Term Cumulative Exposure to Ambient Levels of Total Oxidants and Nitrogen Dioxide in California Seventh-Day Adventist Residents," *Archives of Environmental Health*, 43(4), 279-285.

Ferris, B.G., Higgins, I.T.T., Higgins, M.W., Peters, J.M. (1973), "Sulfur Oxides and Suspended Particulates: Possible Effects of Chronic Exposure," *Archives of Environmental Health*, 27, 179.

Ferris, B.G., Chen, H., Pulea, S., Murphy, R.L. (1976), "Chronic Non-Specific Respiratory Disease in Berlin, New Hampshire, 1967-73," *American Review of Respiratory Disease*, 113, 475.

Greer, J.R., Abbey, D.E., Burchette, R.J. (1993), "Asthma Related to Occupational and Ambient Air Pollutants in Nonsmokers," *Journal of Occupational and Environmental Medicine*, 35(9), 909-915.

Hassett, C., Mustafa, M.G., Coulson, W.F., and Elashoff, R.M. (1985), "Murine Lung Carcinogenesis Following Exposure to Ambient Ozone Concentrations," *Journal of National Cancer Institute*, 75, 771-777.

Hodgkin, J.E., Abbey, D.E., Euler, G., Magie, A.R. (December, 1984), "COPD Prevalence in Nonsmokers in High and Low Photochemical Air Pollution Areas," *Chest*, 86, 830-838.

Jedrychowsky, W., Krzyzanowski, M. (1988), "Ventilatory Lung Function and Chronic Chest Symptoms Among the Inhabitants of Urban Areas With Various Levels of Acid Aerosols: Prospective Study in Cracow," *Environmental Health Perspectives*, 79, 101-107.

Källén, B., Gunnarskog, J., Conradson, T-B. (1993), "Cancer Risk in Asthmatic Subjects Selected from Hospital Discharge Registry," *European Respiratory Journal*, 6, 694-697.

Kim, C.S., Hu, S.C., Dewitt, P., and Gerrity, T.R. (January 24-25, 1994), "Assessment of Local Deposition of Inhaled Particles in Human Lungs," *Colloquium on Particulate Air Pollution*, University of California, Irvine. Session P2.6 Presentation.

Krzyzanowski, M., Camilli, A.E., Lebowitz, M.D. (1990), "Relationships Between Pulmonary Function and Changes in Chronic Respiratory Symptoms - Comparison of Tucson and Cracow Longitudinal Studies," *Chest*, 98, 62-70.

Last, J.A., Warren, D.L., Pecquet-Goad, E., Witschi, H. (1987), "Modification by Ozone of Lung Tumour Development in Mice," *Journal of National Cancer Institute*, 78, 149-154.

Lebowitz, M.D. (1983), "Utilization of Data from Human Population Studies for Setting Air Quality Standards: Evaluation of Issues," *Environmental Health Perspectives*, 52, 193-205.

Lebowtiz, M.D., (1993), "Pulmonary Responses to Multipollutant Airborne Particulate Matter and Other Contaminants," in *Prevention of Respiratory Diseases*, eds. A. Hirah, et al. New York: Marcell Dekker, pp. 209-223.

Lippmann, M. (1989), "Health Effects of Ozone. A Critical Review," *Journal of Air Pollution Control Association*, 39, 672-95.

Love, G.J., Lan, S., Shy, C.M., Riggan, W.B. (1982), "Acute Respiratory Illness in Families Exposed to Nitrogen Dioxide Ambient Air Pollution in Chattanooga, Tennessee," *Archives of Environmental Health*, 37, 75-80.

Mauderly, J.L., Jones, R.K., Griffith, W.C., Henderson, R.F., McClellan, R.O. (1987), "Diesel Exhaust is a Pulmonary Carcinogen in Rats Exposed Chronically by Inhalation," *Fundamental and Applied Toxicology*, 9, 208-221.

Mills, P.K., Abbey, D.E., Beeson, W.L., Petersen, F. (1991), "Ambient Air Pollution and Cancer in California Seventh-day Adventists," *Archives of Environmental Health*, 46(5), 271-280.

Mills, P.K., Abbey, D.E., Beeson, W.L., Petersen, F. (1993a) "Long Term Ambient Air Pollution and Chronic Disease in California Seventh-day Adventists. II: SO_2 and SO_4," *Environ Research*, Submitted for publication.

Mills, P.K., Abbey, D.E., Beeson, W.L., and Petersen, F. (1993b), "Long Term Ambient Levels of Ambient Air Pollution and Chronic Disease in California Seventh-day Adventists. I. PM2.5 Results," *Environmental Research*, Submitted for publication.

Neas, L.M., Dockery, D.W., Ware, J.H., Spengler, J.D., Speizer, F.E., Ferris, B.J., Jr. (1991), "Association of Indoor Nitrogen Dioxide with Respiratory Symptoms and Pulmonary Function in Children," *American Journal of Epidemiology*, 134(2), 204-219.

NTP Technical Report. (1993), "Toxicology and Carcinogenesis Studies of Ozone and Ozone/NNK in F344/N Rats and B6C3F$_1$ Mice," *NTP TR 440, U.S. Dept of Health & Human Services*, NIH Publication No. 94-3371.

Özkaynak, H., Thurston, G.D. (1987). "Associations Between 1980 U.S. Mortality Rates and Alternative Measures of Airborne Particle Concentration," *Risk Analysis*, 7, 449-60.

Robbins, A.S., Abbey, D.E. (1993), "Passive Smoking and Chronic Respiratory Disease Symptoms in Non-Smoking Adults," *International Journal of Epidemiology*, 22, 809-817.

Rokaw, S.N., Detels, R., Coulson, A.H., Sayre, J.W., Taskin, D.P., Allwright, S.S., Massey, F.J. (1980), "The UCLA Population Studies of Chronic Obstructive Pulmonary Disease. "Comparison of pulmonary function in three communities exposed to photo-chemical oxidants, multiple primary pollutants or minimal pollutants," *Chest*, 78, 252-262.

Sawicki, F., Lawrence, P.S. (1977), "Chronic Non-Specific Respiratory Disease in the City of Cracow," *National Institute of Hygiene*, Warsaw, Poland.

Shy, C.M., Creason, J.P., Pearlman, M.E., McClain, K.E., Benson, F.B., Young, M.M. (1970), "The Chattanooga School Study: Effects of Community Exposure to Nitrogen Dioxide. II. Incidence of Acuter Respiratory Disease," *Journal of Air Pollution Control Association*, 20, 582-288.

Sunyer, J., Saez, M., Murillo, C., et al. (1993), "Air Pollution and Emergency Room Admissions for COPD: A five Year Study," *American Journal of Epidemiology*, 1273, 701-05.

Van der Landa, R. (1981), "Longterm Exposure to Air Pollution and Declines in Lung Function," *Chest*, 80s, 23-26.

Vesterinen, E., Pukkala, E., Timonen, T., Aromaa, A. (1993), "Cancer Incidence Among 78,000 Asthmatic Patients, *International Journal of Epidemiology*, 22, 976-982.

Viegi, G., Paoletti, P., Carrozzi, L., Guintini, C., Lebowitz, M.D. (1991), "Prevalence Rates of Respiratory Symptoms in Italian General Population Samples Exposed to Different Levels of Air Pollution, *Environmental Health Perspectives*, 94, 95-99.

von Mutius, E., Fritzsch, C., Weiland, S.K., et al. (1992), Prevalence of Asthma and Allergic Disorders Among Children in United Germany: A Descriptive Comparison," *British Medical Journal*, 305, 1395-1399.

Ware, J. (1980), "Assessment of the Health Effects of SO_x and PM: Analysis of the Exposure-Response Relationship," *U.S. EPA*, Research Triangle Park, NC.

Winer, A.M., Lurman, R.W., Coyner, L.A., Colome, S.D., Poe, M.P. (1989), "Characterization of Air Pollutant Exposures in the California South Coast Air Basin: Applications of a New Regional Human Exposure (REHEX) Model," Final Report. Contract No. TSA 106-01-88. *Statewide Air Pollution Research Center*. University of California, Riverside, California.

Footnote (page #8)

[1] Many SDAs join the church later in life and have smoked before becoming an SDA; any individuals who reported smoking since 1977 were excluded.

Table 1 Pairwise Correlations of Estimated Mean 1977-1987 Concentrations of Ambient Pollutants for Members of the AHSMOG Respiratory Symptoms Cohort (n = 3914)[a]

	TSP	PM10[b]	PM2.5[c]	SO_4	OZ	SO_2	NO_2
TSP	1	.95	.86	.69	.72	.61	.46
PM10[b]	. . .	1	.89	.72	.76	.64	.52
PM2.5[c]	. . .	. . .	1	.30	.62	.47	.25
SO_4	. . .	. . .	. . .	1	.57	.60	.63
OZ	. . .	. . .	. . .	. . .	1	.38	.36
SO_2	. . .	. . .	. . .	. . .	. .	1	.85
NO_2	. . .	. . .	. . .	. . .	. . .	. . .	1

[a] Correlations between PM2.5 and other ambient concentrations are computed only for 1977-1986 for individuals living in the vicinity of nine California airports, n = 1868. Correlations of other pollutants are computed for the entire cohort, n = 3914, April 1977 - April, 1987.

[b] PM10 was indirectly estimated from TSP using site/seasonal regression equations.

[c] PM2.5 was indirectly estimated from airport visibility data.

Table 2 Relative Risks Estimated from Multivariate Models and 95% Confidence Intervals for Selected Increments of Particulate Pollutants and Ozone

		POLLUTANT/INCREMENT					
DISEASE	n[a]	TSP 42 days/yr above 200 $\mu g/m^3$	PM10 42 days/yr above 100 $\mu g/m^3$	SO_4 Av. Annual increase of 7 $\mu g/m^3$	Ozone 500 hr/yr above 10 pphm	PM2.5 Average annual increase of 45 $\mu g/m^3$	n[b]
DEVELOPMENT OF:							
AOD[c]	272	1.36 (1.11, 1.66)	1.17 (1.02, 1.33)	1.43 (0.88, 2.26)	1.04 (0.86, 1.25)	1.46 (0.84, 2.46)	135
Chronic Bronchitis[d]	234	1.33 (1.07, 1.65)	1.17 (1.01, 1.35)	0.96 (0.58, 1.55)	1.02 (0.83, 1.25)	1.81 (0.98, 3.25)	117
Asthma[e]	80	1.74 (1.11, 2.71)	1.30 (0.97, 1.73)	2.85 (1.03, 7.40)	1.35 (0.93, 1.96)	1.41 (0.47, 4.06)	40
INCIDENCE OF ALL MALIGNANT NEOPLASMS Females[f]	175	1.37 (1.05, 1.80)	1.15 (0.97, 1.38)	Not Done	1.03 (0.81, 1.32)	2.01 (1.05, 3.86)	103
RESPIRATORY CANCER INCIDENCE [f]	17	1.72 (0.81, 3.65)	1.46 (0.88, 2.40)	Not Done	2.25 (0.96, 5.31)	Not done	

[a] n = number of new cases for entire cohort for whom estimated TSP, PM10, SO_4, and Ozone were available.
[b] n = number of new cases for subcohort for whom estimated PM2.5 data was available.
[c] Variables included in the model for definite AOD symptoms (besides age) are education, sex, childhood colds, childhood AOD, possible symptoms in 1977, years smoked, years lived with a smoker, and years worked with a smoker.
[d] Variables included in the model for definite bronchitis symptoms (besides age) are education, sex, childhood colds, possible symptoms in 1977, years smoked, and years lived with a smoker.
[e] Variables included in the model for definite asthma (besides age) are education, sex, possible symptoms in 1977, and years worked with a smoker.
[f] Variables included as covariates in the Cox model for cancer (besides age) are total years of past smoking and education attainment.

Table 3 **Statistically Significant ($p<.05$) Associations of Ambient Pollutants, as Measured by Mean Concentration and Exceedance Frequency Statistics, with Disease Outcomes for the AHSMOG Study (Association with health outcome may be due to correlation with other pollutant, if it is written in cell)**

DISEASE		New Cases 1977-1987	TSP	PM10	PM2.5	SO_4	Ozone
DEVELOPMENT OF:	AOD	272	μ[a], 100+[b]	80+ (TSP)	None[c]	None	None
	Chronic Bronchitis	234	μ, 100+	100+ (TSP)	20 (TSP)	None	None
	Asthma	80	150+	None	None	μ	[d]
INCREASING SYMPTOMS SEVERITY OF:	AOD		μ, 60+	μ, 40+ (TSP)	μ, 20+ (TSP)	6 (TSP)	None
	Chronic Bronchitis		150+ (PM2.5)	None	μ, 20	None	None
	Asthma		100+	μ, 50+	μ, 40 (TSP)	None	μ, 10, 12
INCIDENCE OF ALL MALIGNANT NEOPLASMS	Males	115	None	None	None	None	None
	Females	175	100+	None	μ, 20+	Not Done	None
RESPIRATORY CANCER INCIDENCE		17	None	None	Not Done	Not Done	10[d]

[a] μ = ambient mean concentration

[b] 100+, etc. exceedance frequencies of 100 and higher are statistically significant, $p<.05$. Units are in $\mu g/m^3$ for particulate pollutants - TSP, SO_4, PM10, PM2.5; otherwise units are pphm. The symbol 100 without the "+" would indicate that only exceedance frequencies for the cutoff of 100 were statistically significant.

[c] None = No statistically significant association with mean concentration or any exceedance frequency statistic.

[d] For ozone and asthma a trend association ($p=.056$) was noted for exceedance frequencies of 10 pphm and higher and a strong statistically significant association between mean concentration and development of asthma in males but not females.

Air Pollution and Asthma Exacerbations Among African-American Children in Los Angeles

Bart D. Ostro, Ph.D.,[1] Michael J. Lipsett, M.D.,J.D.,[1] Jennifer K. Mann, MPH,[1] Hazel Braxton-Owens, MPH[2] and Mary C. White, Sc.D.[3]

1 Office of Environmental Health Hazard Assessment
California Environmental Protection Agency
Berkeley, CA

2 California Public Health Foundation
Berkeley, CA

3 Centers for Disease Control and Prevention
National Center for Environmental Health
Atlanta, GA

Running Head: Air Pollution and Asthma

Address correspondence to:

Bart Ostro, Ph.D., Chief
Air Pollution Epidemiology Unit
California Environmental Protection Agency
2151 Berkeley Way, Annex 11
Berkeley, CA 94704

Acknowledgement: This project was supported by Cooperative Agreement #U60/CCU907015-01 from the Centers for Disease Control and Prevention, National Center for Environmental Health. The authors also would like to express appreciation to the following: Paul Papanek, M.D. and Amy Wohl, Ph.D. from the County of Los Angeles, Department of Health Services; James Rifenberg of the Centers for Disease Control and Prevention; the South Coast Air Quality Management District; the American Lung Association, Los Angeles County Affiliate; physicians, nurses, and other staff at the recruitment sites including Lawrence Robinson, M.D. (King Drew Medical Center), Joel Korotzer, M.D., Mark Haendel, M.D., and Mary Deutsche, M.D. (Kaiser West Los Angeles and Kaiser Inglewood), and Warren Richards, M.D. (Childrens Hospital Los Angeles); individuals from the recruitment area who helped plan the recruitment and questionnaire development; and the children and families involved in the study.

Abstract

Clinical and epidemiologic evidence suggests that particulate matter and ozone are associated with exacerbations of asthma. African-American children, who have experienced a marked increase in asthma morbidity and mortality during the 1980s, may represent a particularly sensitive subgroup. In order to examine potential effects of air pollution on exacerbations of asthma, a panel of 83 African-American children, aged 7 to 12, were recruited from four allergy and pediatrics clinics in central Los Angeles and two asthma camps in the summer of 1992. Their socioeconomic status varied greatly. Daily data on asthma symptoms, medication use, and peak flows were recorded for three months and examined in conjunction with data on PM10, ozone, nitrogen dioxide, sulfur dioxide, pollens, molds, and meteorologic factors. Using multiple logistic regression analysis corrected for autocorrelation, the daily probability of shortness of breath was shown to be associated with both ozone and PM10 concentrations. The effect of particles on shortness of breath was greater among children with moderate or severe asthma. These findings were confirmed by an individual-level analysis that took full advantage of the panel study design.

Introduction

Since the late 1970s, rates of asthma morbidity and mortality have increased strikingly in the United States, with African-Americans being disproportionately affected (Sly 1988; Evans et al. 1987; Weiss et al. 1990). Although asthma prevalence is only somewhat greater among African Americans than among whites, rates of morbidity and mortality are markedly higher. Part of the racial differences in rates are likely due to conditions associated with poverty (Wissow 1988). Various social and environmental factors, including exposure to indoor and outdoor pollutants and allergens, have been postulated as explanations of these asthma trends (Weiss, 1993). To date, however, little research has been conducted to determine the effects of environmental factors, such as air pollution, on asthma among African Americans.

Epidemiologic evidence indicates that ozone and particulate matter are associated with exacerbations of asthma (Schwartz et al., 1993; Thurston et al., 1992; Pope, 1991). Most epidemiologic studies have relied on measurement of airborne particles other than PM10 (particulate matter less than 10 microns in diameter), the current metric for the National Ambient Air Quality Standard for particulate matter in the United States. In most areas, particulate matter is measured only every sixth day. The lack of daily data on ambient levels of PM10 has impeded epidemiologic research, while the complex and heterogeneous nature of particulate matter has discouraged clinical efforts to study its effects in the laboratory. Although results of several epidemiologic studies suggest that ambient ozone is associated with exacerbations of asthma, the results from controlled exposure studies are mixed, especially at low effective doses (Kreit et al., 1989; Koenig et al., 1985; Linn et al., 1978).

In this paper, we report results of a pilot study of the respiratory health of a sentinel group of asthmatic children in relation to several ambient pollutants in Southern California. The results suggest that relatively low ambient levels of both PM10 and ozone are associated with shortness of breath among this population of African-American children.

Data and Methodology

The Study Population

The study sample consisted of African-American children, ages 7 to 12, with confirmed asthma, most of whom lived in the central and south-central sections of Los Angeles. Subjects were required to have had asthmatic symptoms requiring medication, in the absence of a respiratory infection, during the six months preceding study recruitment. Subjects were recruited during an office visit or by telephone, from allergy, pediatrics, and urgent care clinics at four large health care facilities (Kaiser West Los Angles, Kaiser Inglewood, Children's Hospital of Los Angeles and King Drew Medical Center), and two asthma camps run by the American Lung Association and by the Asthma and Allergy Foundation. Such disparate recruitment sources ensured variation in household income and socioeconomic status, access to medical care, and asthma severity. The locations of the recruitment sites are displayed in Figure 1.

After recruitment and screening, 109 children were eligible for participation. Most lived in Los Angeles or Inglewood, within 10 miles of downtown Los Angeles, where one fixed-site monitor was located. Recruitment began in June 1992, and 13 weeks of daily data collection began the first week of August 1992. We chose this study period so as to include both the high summer concentrations of particulate matter and ozone and the high early fall concentrations of pollen. Informed consent was obtained from both children and their parents, and nominal incentives were provided for each completed diary. Parents and children were told that they were participating in an investigation of environmental factors affecting asthma, but not that the principal variables of interest were air pollutants.

Health Data

Data collected from the participants included (1) answers to an intake questionnaire completed by the parents, providing information on the children's medical histories, asthma severity and known triggers, allergic status, medical management, family structure and demographics, and potential indoor exposures; and (2) a 13-week daily diary recording asthma symptoms, medication use, respiratory infections, peak expiratory flows, and physician visits. Interviewers administered questionnaires to the parent or legal guardian of each child within a few weeks of the

child's clinic visit or telephone contact. All participants were carefully instructed on the use of the daily-symptom diary and received a return visit approximately 2 weeks after the start of the study. To enhance compliance, the interviewers made additional home visits, and the parents or care-givers were telephoned weekly and asked about the children's respiratory status. Of the 109 eligible children, 91 completed an intake questionnaire, and 87 began the daily diaries. Of these 87, 69 (79%) completed at least 6 weeks of daily diaries, and 55 (63%) completed at least 9 weeks of daily diaries. Data from four participants were later dropped because of concerns about the accuracy of their responses, leaving a final sample size of 83. Comparison of the group that responded for more than 9 weeks with the group that completed less than 9 weeks of diaries showed no difference in the rate of reported symptoms during the study. However, the latter group (< 9 weeks) had less education and income and reported more symptoms when asthma history was reported in the intake questionnaire.

Table 1 summarizes the age, household income, and asthma history and status of the study population. Households with a wide range of income and education were represented in the study. The subjective evaluation of the parents and our own evaluation in which we used criteria established by the National Institutes of Health's National Asthma Education Project (based on medication use, restrictions in activity, symptoms and use of medical services) indicated that about 70% of the sample had moderate or severe asthma. Almost half of the sample were using anti-inflammatory medicine, and a quarter were using inhaled steroids.

This initial analysis focused on the daily reporting of respiratory symptoms including cough, shortness of breath, and wheeze. During the study period, the daily prevalences of cough, shortness of breath, and wheeze were 22, 5.6, and 22 percent, respectively. Since enrollment was staggered during the first week of the study, data from the first 5 days were deleted from the analysis.

Covariates

Daily data for ambient air pollution, meteorologic factors, pollens, and molds were collected during the 13-week study period. Specifically, daily measurements of ozone, PM10, nitrogen dioxide, and sulfur dioxide

were made at three fixed-site monitors (see Figure 1) in downtown Los Angeles, Lynwood, and Pico Rivera. In addition to its routine fixed-site monitoring, the South Coast Air Quality Management District also collected daily PM10 data at the downtown Los Angeles site (using a beta attenuation monitor), which was the closest fixed-site monitor to the recruitment locations. For all four of the above pollutants, both hourly and 24-hour average data were collected daily. Total suspended particulates (TSP), sulfates, and nitrates were collected every sixth day at the Los Angeles site. Daily values of the gaseous pollutants were collected at all three sites, and daily PM10 values were also collected at the Lynwood site with a TEOM monitor. Because most of the study population resided in the central Los Angeles area, data from the downtown Los Angeles monitor were used in the initial analysis. Sensitivity analysis was conducted to examine the impact of averaging the results from all three monitors and of assigning each participant to the monitor closest to his or her residence. Although there is substantial air quality variability within the Los Angeles air basin, with lower ozone and particle levels near the coast, the centralized recruitment ensured a sample that could be reasonably represented by the three monitoring stations. However, a few children recruited from the asthma camps lived a greater distance from downtown Los Angeles.

Pollution concentrations during the study period were relatively moderate; the 1-hour maximum ozone level at the central Los Angeles station averaged 0.08 ppm and the mean 24-hour PM10 level was 56 $\mu g/m^3$. Levels of PM10 and ozone were correlated ($r = 0.50$), and the maximum daily temperature was correlated with the ozone level ($r = 0.35$) but not with the PM10 level ($r = .06$). Figure 2 depicts the temporal pattern of ozone and PM10. Of note, PM10 concentrations peaked during the last third of the study, and ozone concentrations were highest at the beginning of the study and declined thereafter.

The daily meteorologic data, collected at the Los Angeles Airport and obtained from the National Climatic Center, included maximum and minimum temperature, humidity, precipitation, and dewpoint. Data on pollen and molds were collected at a private allergy clinic in Los Angeles, several miles west of the monitor site, and were examined both as aggregates (e.g., levels of total pollen) and in disaggregated form (e.g.,

levels of pollen from grasses, trees and weeds, and specific molds, such as alternaria and cladosporium). Levels of pollens and molds exhibited very distinct peaks and were not correlated with either PM10 or ozone concentrations. Temperature, humidity, and levels of pollens and molds were considered both as continuous variables and as binary variables indicating extremes. Specifically, we examined the effect of a given variable being in the top 25, 10, and 5 percent of all such measurements. As indicated in Figure 2, the levels of molds (and pollen, not displayed) peaked at the end of the study period. In general, there appeared to be only minimal confounding among the pollutants, pollens, and molds. Table 2 details the distribution of the levels of air pollutants, pollens, molds, and three of the weather variables.

Statistical Methods

In a panel design study each person can serve as his or her own control. The data set therefore will provide substantial degrees of freedom and a reduced potential for confounding and covariation among explanatory variables.

Three different levels of analysis were conducted for each of the six binary response variables indicating daily symptoms: the prevalence (i.e., the likelihood of a symptom on a given day) of cough, shortness of breath and wheeze, and the likelihood of an incident attack, for any of these three asthma symptoms, that might extend over several days. In the first analysis, qualitative in nature, the group rate of symptoms (e.g., the proportion of the sample reporting cough on a given day) was analyzed. This was undertaken in order to detect general patterns in the data, to determine the form of the dose-response relationship, to narrow down the list of candidate pollution and meteorologic variables, to examine patterns of residuals for autocorrelation, and to search for influential data points.

Next, all of the individual observations were pooled and examined by logistic regressions based on the generalized estimating equations (GEE) proposed by Liang and Zeger (1986). This method generates robust estimators regardless of the specification of the covariance matrix and corrects for the repeated measure design of the study. Thus, the standard error of the regression estimate is adjusted for the fact that responses

from any one person are likely to be correlated. The pooling of individuals allows for stratification of the sample and also for the examination of individual variables such as age, sex, household income, and education, severity, and the occurrence of a respiratory infection. Therefore, in this phase, both the full sample and a subset of those children who had reported moderate to severe asthma were examined.

Finally, in a third level of analysis, individual regressions were run in order to allow subjects to serve entirely as their own controls. The individual regression results were then aggregated by weighting each estimate by the inverse of its regression variance, a technique suggested by Whittemore and Korn (1980). This technique fully exploits the advantages of the panel study design. Because each person is followed over a short period of time, when most potentially confounding factors (e.g., age, family structure, access to medical care, family income) are constant, the analysis can isolate the impacts of variables that change daily, such as levels of indoor and outdoor air pollution, meteorologic factors, pollen counts, and respiratory infections. In addition, this technique allows investigators to examine individual sensitivity to air pollution exposure. The disadvantage of this technique is that subjects for whom the initial individual-level logistic analysis fails to converge (because of too few sample days or reported symptoms) are automatically omitted from subsequent analyses.

For the group-level and pooled analyses, a stepwise regression was first run for all covariates other than the pollutant level. For the group analysis, these variables included levels of all of the pollens and molds, meteorologic factors, study day, and indicator variables for weekend (versus weekday) and month. Maximum temperature, humidity, pollens and molds were examined by using both continuous variables and binary variables indicating peaks (levels at the 75th, 90th and 95th percentiles). For the pooled analysis, the covariates examined included age, sex, previous day with cold, and socioeconomic status (income level). For the pooled and individual-level analyses, contemporaneous values and 1-, 2- and 3-day lags of the pollution and meteorologic variables were considered. Covariates were examined in the full regression model (with air pollution) if their associated t-statistics, prior to the introduction of an air pollution term, were equal to or greater than one. To assess

the sensitivity of the results to alternative regression specifications, we reported the odds ratios and confidence intervals as additional variables were added to the model. Potential interactions among the included variables were also examined.

RESULTS

1. Pooled analysis

With the exception of the reporting of a symptom or a cold on the previous day, most of the variables were not associated with the reporting of asthma symptoms. Of the pollutants, PM10 and ozone were associated with shortness of breath but not with any other symptom. Accordingly, shortness of breath was the only symptom examined in subsequent models. The generalized estimating equations included a one-period autocorrelation correction that effectively reduced the serial correlation of the residuals. Table 3 shows the association of shortness of breath with PM10 and ozone concentrations as additional covariates are added to the model specification. As explanatory variables, the basic model includes a binary variable indicating the presence of shortness of breath on the previous day, a binary variable for respiratory infection on the previous day, and one air pollution variable, either PM10 or ozone. Shortness of breath had a statistically significant association with both ozone (OR = 1.40, 95%CI = 1.14 - 1.72 evaluated at the mean ozone level) and PM10 (OR = 1.58, 95%CI = 1.05 - 2.38 evaluated at the mean PM10 level). Having a respiratory infection on the previous day was statistically associated with shortness of breath ($p < 0.05$). Transforming the pollution variables into log terms only slightly improved the association of shortness of breath with PM10 and had no impact on its association with ozone. Contemporaneous exposures of ozone and PM10 were more associated with symptoms than were the lagged versions of these variables. Therefore, in subsequent analyses, we used untransformed same-day exposures of pollution in the regression specifications. When children were assigned exposures based on either the average of the three fixed-site monitors or the monitor closest to their residence, the significance of the associations with PM10 and ozone did not change.

When continuous values of temperature, pollens and molds were added to the model, only temperature was associated with shortness of breath (p

= .10). This association was unchanged when temperature extremes (i.e., days with the highest 10% maximum temperatures, greater than 82° F) were added to the model. Although the binary variables representing the extreme values for molds were not statistically significant, a binary variable indicating the 5% of days with highest levels of molds (i.e., days on which more than 2,000 spores per cubic meter were recorded) was kept in subsequent models because the t-statistic was close to one. The inclusion of these variables did not affect either the magnitude or significance of the coefficient for PM10. The statistical significance of ozone dropped slightly ($p < 0.05$) probably because of the covariation between ozone and temperature. When we added the variable indicating a previous day with a cold back into the model, the pollution results were not altered, although in the specification with ozone, temperature was now more significant ($p < 0.05$). There did not appear to be any interactions between pollutants or between pollution and either temperature or molds. Among the individual demographic variables, the frequency of symptom reporting did not vary on the basis of sex, but frequency did increase as age and family income increased.

For the subsample of the population with moderate or severe asthma, shortness of breath was associated with both PM10 and ozone but the magnitude of the effect was similar to the group as a whole. Finally, the beginning of an asthma episode was associated with ozone ($p < 0.05$) (OR = 1.60, 95%CI = 1.01-2.58) but not PM10 ($p > 0.10$) (OR = 1.06,95%CI = 0.61-1.84).

2. Individual-level analysis

As expected, many children were excluded from this part of the analysis because of low symptom rates or too few days of response, leaving only 41 subjects. Again, no association was found between any time-varying factors, including air pollution, and either cough or wheeze. However, shortness of breath was associated with both PM10 ($p < 0.01$) (OR = 2.42, 95%CI = 1.76-4.68) and ozone ($p = 0.06$) (OR= 1.51, 95%CI = 0.98-2.32). Also, the beginning of an incident attack (or episode) of shortness of breath was associated with ozone, but not PM10. Finally, to compare the models, we reran the pooled analysis using the same subsample of 41 subjects and found both ozone and PM10 to be associated with

shortness of breath. The odds ratios for this subset were higher than for the group as a whole (Table 3).

Discussion

There are relatively few epidemiologic investigations of the relationship between air pollution and asthma. Results from previous epidemiologic studies suggest that ambient particulate matter or ozone exacerbate asthma symptoms (Whittemore and Korn 1980; Ostro et al., 1992; Pope et al. 1993) and may result in increases in hospital admissions or emergency room visits (White et al., 1994; Thurston et al., 1992; Bates et al. 1987, 1990). Our study suggests that both PM10 and ozone are associated with increased daily reporting of shortness of breath. These effects appear to occur at the moderately elevated levels of pollution experienced in central and south-central Los Angeles in the late summer and early fall of 1992.

It is difficult to attribute these effects solely to one pollutant since concentrations of PM10 and ozone were correlated. However, if the effects are caused primarily by higher levels of the pollutants, both of the pollutants may have some effect, since the peak concentrations of the pollutants occurred at different times during the study. This possibility is supported by the observation that different relative effects of the pollutants appear to have occurred when those asthmatics with more symptoms were examined. Specifically, results of the individual-level and pooled analysis of the subsample of 41 children, suggest that the asthmatics in the study with more baseline symptoms have a heightened effect (relative to the total sample) from exposure to PM10 (Table 3); the effect of ozone on this group is generally similar to that observed for the entire sample.

This association does not appear to be a result of confounding by other time-variant factors such as meteorologic conditions, pollens or molds, since these factors did not exhibit significant univariate associations with asthma symptoms in our sample. Only on days of extremely high temperatures and high levels of molds were the relationships between these factors and shortness of breath even close to statistical significance. The potential for confounding was further reduced through the use of the individual-level analysis. Using this

approach, in which each individual served as his or her own control, no other factor is likely to be correlated with both outdoor air pollution and shortness of breath. Although this approach is based on a more sensitive subgroup of the sample (those reporting enough symptom days or total days so that the individual logistic regressions could converge) it suggests that there are some children who are especially sensitive to air pollution, especially PM10.

The biologic validity of the consistent statistical relationship between PM10, ozone, and reported shortness of breath would be strengthened by finding similar associations with wheezing. However, individual perceptions of shortness of breath may occur in relation to a sensation of chest tightness with or without audible wheezing. This seems to be the case with our study population: reporting a shortness or breath and chest tightness were strongly associated ($r = 0.42$; $p < 0.0001$).

In general, the impact of air pollution on this population of African-American children in Los Angeles with asthma appears significant. At the mean levels of pollution, we detected a 50 percent increase in the reporting of shortness of breath. For a 10 $\mu g/m^3$ change in PM10, this amounts to about a 9 percent increase in symptoms.

E. References

Bates DV, Baker-Anderson M, Sizto R (1990) Asthma attack periodicity: a study of hospital emergency visits in Vancouver. Environ Res 51:51-70.

Bates DV and Sizto R (1987) Air pollution and hospital admissions in Southern Ontario: the acid summer haze effect. Environ Res 43:317-331.

Gergen PJ, Weiss KB (1990) Changing patterns of asthma hospitalization among children: 1979 to 1987. JAMA 264:1688-1692.

Koenig JQ, Covert DS, Morgan MS et al. (1985) Acute effects of 0.12 ppm ozone or 0.12 ppm nitrogen dioxide on pulmonary function in healthy and asthmatic adolescents. Am Rev Respir Dis 132:648-651.

Kreit JW, Gross KB, Moore TB, Lorenzen TJ, D'Arcy J Eschenbacher WL. (1989) Ozone-induced changes in pulmonary function and bronchial responsiveness in asthmatics. J Appl Physiol 66:217-222.

Linn WS, Buckley RD, Spier CD et al. (1978) Health effects of ozone exposure in asthmatics. Am Rev Respir Dis 117:835-843.

Ostro BD, Lipsett MJ, Wiener MB, Selner JC (1991) Asthmatic responses to airborne acid aerosols. Am J Public Health;81:694-702.

Pope CA III, Dockery DW, Spengler JD, Razienne ME (1991) Respiratory health and PM_{10} pollution: a daily time series analysis. Am Rev Respir Dis 144:688-674.

Pope CA III (1991) Respiratory hospital admissions associated with PM10 pollution in Utah, Salt Lake and Cache Valleys. Arch Environ Health 46:90-97.

Schwartz J, Slater D, Larson T, et al. (1993) Particulate air pollution and hospital emergency room visits for asthma in Seattle. Am Rev Respir Dis 147:826-831.

Sly RM (1988) Mortality from asthma in children 1979-1984. Ann Allergy 60:433-443.

Thurston GP, Ito K, Kinney P, Lippmann M (1992) A multi-year study of air pollution and respiratory hospital admissions in three New York State metropolitan areas: Results for 1988 and 1989 summer. J Expos Anal Environ Epidemiol 2:429-450.

Weiss KB, Wagener DK (1990) Changing patterns of asthma mortality. Identifying target populations at high risk. JAMA 264:1683-1687.

White MC, Etzel RA, Wilcox WD, Lloyd C (1994) Exacerbations of children's asthma and ozone pollution in Atlanta. Environ Res (in press).

Wissow LS, Gittelsohn AM, Szklo M, Starfield B, Mussman M (1988) Poverty, race, and hospitalization for childhood asthma. Am J Pub Health 78:777-

782.

Whittemore AS, Korn EL (1980) Asthma and air pollution in the Los Angeles area. Am J Public Health 70:687-696.

Zeger S, Liang K (1979) Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42:121-130.

Table 1. Demographic and Socioeconomic Characteristics of Children who Completed Diaries (n = 83).

Characteristics	Categories	Percent (%)
Age Distribution (years)	7-8	33.7
	9-10	37.3
	11-12	28.9
Sex	Female	32.5
	Male	67.5
Annual Household Income	< $10,000	20.8
	$10-25,000	24.7
	$25-40,000	24.7
	$40-70,000	23.4
	> $70,000	6.5
Asthma Severity	Mild	28.9
	Moderate	53.0
	Severe	18.1
Use Inhaled Steroids	Yes	27.7
	No	72.3
Use any Anti-inflammatory Medicine	Yes	45.8
	No	54.2
Recruitment Site	Kaiser West L.A.	18.1
	Kaiser Inglewood	28.9
	King-Drew Med. Center	9.6
	Children's Hospital, L.A.	8.4
	Asthma Camps	34.9

Table 2. Descriptive Statistics for Meteorologic Conditions, Air Pollution, Pollens and Molds in Los Angeles, August to November, 1992.

Variable	Mean	S.D.	Range
Ozone (1-hour maximum, pphm)	8.02	3.58	1.00-16.00
PM10 (24-hour average, $\mu g/m^3$)	55.87	15.22	19.63-101.42
SO_2 (24-hour maximum, pphm)	0.114	0.143	0.00-0.61
NO_2 (1-hour maximum, pphm)	9.78	5.99	2.0-31.0
Maxtemp (°F)	75.77	4.84	67.00-90.00
Relative humidity (%)	86.41	7.80	40.00-97.00
Rainfall (inches)	0.009	0.038	0.00-0.25
Pollens (grains/m^3)	12.35	13.94	0.00-86.00
Molds (spores/m^3)	760.22	743.27	135-4,682

Table 3. Regression Coefficients for Shortness of Breath Using Pooled and Individual-level Models (odds ratio, 95% confidence interval evaluated at the mean pollution level, and associated p-value).

Model Specification	PM10	Ozone
pollutant + cold&	1.58 (1.05, 2.38) < 0.05	1.40 (1.14, 1.72) <0.001
pollutant +temp90 + mold95	1.60 (1.10, 2.32) < 0.05	1.34 (1.01, 1.77) < 0.05
pollutants + cold + temp90 + mold95	1.60 (1.07, 2.37) < 0.05	1.36 (1.02,1.83) 0.05
pollutant (moderate or severe asthma)	1.64 (1.01, 2.65) <0.05	1.47 (0.96,2.24) <0.10
pollutant + cold +temp90 + mold95 (n=41)§	1.82 (1.23, 2.67) < 0.10	1.46 (1.02, 2.08) <0.05
pollutant+cold (n=41)#	2.42 (1.76,4.68) < 0.01	1.51 (0.98,2.32) <0.05

& respiratory infection on previous day;
§ sample restricted to those in the individual-level analysis ;
# using results and weights of individual-level analysis;
temp90 = top 10 percent of all daily maximum temperatures; mold95 = top 5 percent of all daily mold readings.

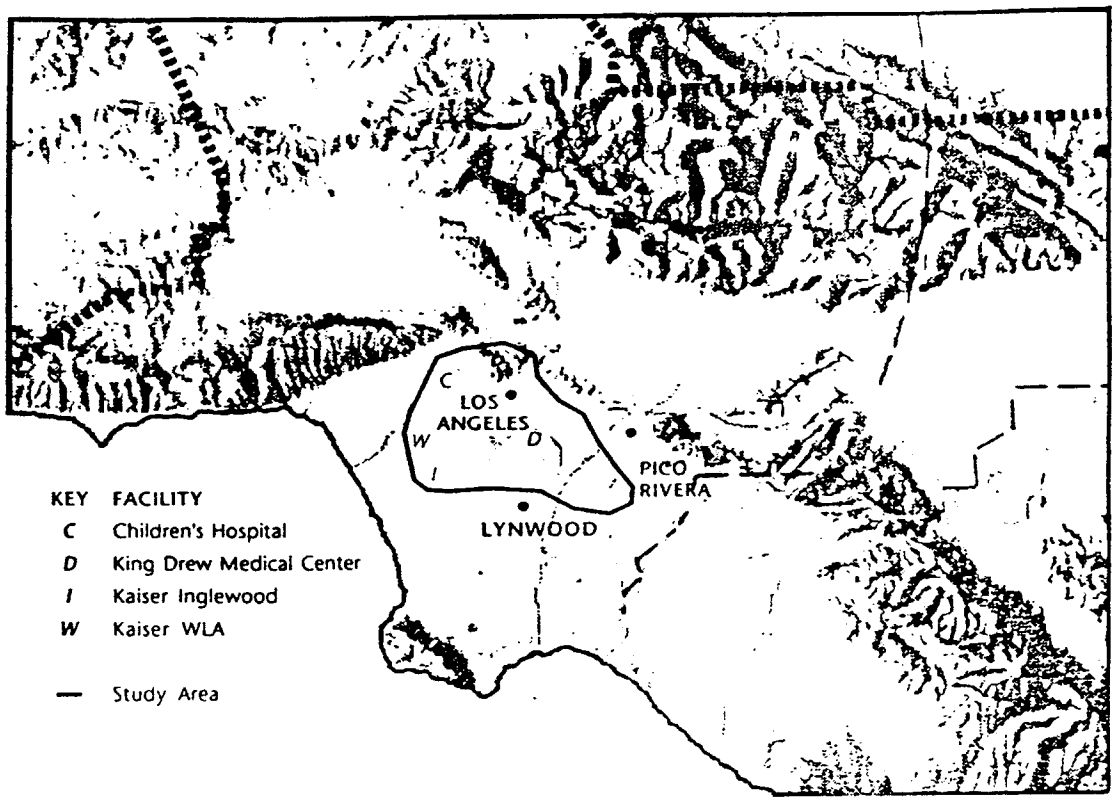

Figure 1. Locations of facilities used for recruitment of subjects and cities in which monitors were located.

Figure 2. Ambient Data During Study Period.

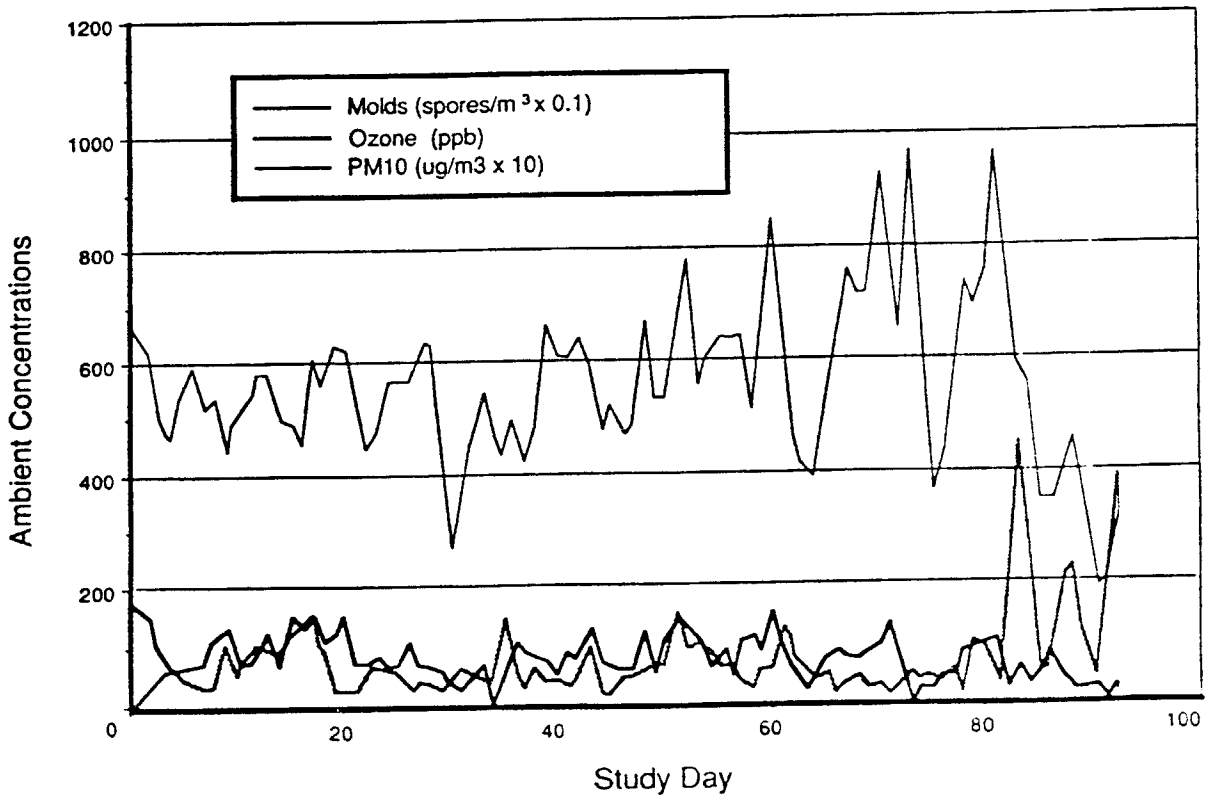

Figure 2. Ambient data during study period.

Excess Mortality and Exposure to PM_{10} Air Pollution: Additional Analyses from Utah

Joseph L. Lyon, M.D., M.P.H.

Motomi Mori, Ph.D.

Renlu Gao, M.Stat

Department of Family and Preventive Medicine, University of Utah School of Medicine, Salt Lake City, Utah

Address Correspondence to:

Joseph L. Lyon, M.D., M.P.H.

Deparatment of Family and Preventive Medicine

University of Utah School of Medicine

50 North Medical Drive, Room 1C26

Salt Lake City, UT 84132

ABSTRACT

We examined the association between exposure to PM_{10} air pollution and daily mortality in Utah County, Utah for 1985-92. We confirmed the previous finding that exposure to fine particulate air pollution (particulate diameter of ≤ 10 μm) in amounts of 50+ $\mu g/m^3$ increased daily mortality by 3%. The potential importance of this observation led us to test the association more rigorously by assessing rate ratios (RR) of PM_{10} for year, season, and location at time of death. For individual years there was no statistically significant association between increased mortality and exposure to PM_{10} air pollution. The strongest mortality effect was seen in the spring, not the winter. The largest numeric contribution to excess mortality was from individuals age 75+ dying in hospital, and the largest RR was for individuals ages 15-59 dying at home primarily of cancer. These findings do not support a causal association between exposure to PM_{10} air pollution and daily mortality.

INTRODUCTION

Daily exposure to fine particulate air pollution of average diameter 10 μ or less (PM_{10}) has been associated with increased mortality within a few days of exposure after controlling for relevant weather variables (Schwartz, 1991; Schwartz and Dockery, 1992a; Schwartz and Dockery, 1992b; Schwartz, 1993; and Pope et al., 1992). The excess mortality was greatest among those dying from cardiopulmonary diseases, and was present at levels of air pollution below the current National Ambient Air Quality Standard. These studies have assumed uniform exposure to air pollution within a city or county, regardless of where the death occurred. However, many people die in hospitals or nursing homes rather than at home, and it is questionable if these individuals are uniformly exposed to the same level of PM_{10}. None of the previous studies have determined whether the association differed by location at time of death, or was uniform across seasons and the years studied. The failure to find a consistent mortality effect across time or by location would suggest that the excess mortality attributed to air pollution was the result of confounding by other factors such as weather, or effect modification by an unknown factor.

Because of the potential impact of this association on the public health, this study tested the association between exposure to PM_{10} air pollution and excess mortality more rigorously than previous studies. The study objectives were: (1) to examine the excess mortality associated with exposure to fine particulate air pollution by age groups and location at time of death; (2) to examine the causes of death associated with PM_{10} exposure by age

group, location at time of death, and by year and season; (3) and to examine the annual and seasonal variation of the PM_{10} effect on daily mortality.

METHOD

Information on deaths was obtained from the death certificates filed with the State of Utah for April 1985-December 1992. A computerized file of all death certificates in Utah during these years was prepared by the State Registrar of Vital Statistics. The file contained the following information: a unique file number, date of birth, gender, date of death, cause of death (ICD-9 code), county of residence at time of death, county of death, and location at time of death (hospital, nursing home or other). Deaths due to an accidental cause (ICD-9 codes 800 and above) were excluded. Only those who died in their County of residence were used in the analysis. The 1985-1992 age- and sex-specific populations for each county were estimated by linear interpolations from the 1980 to 1990 U.S. census. The number of live births for 1985-92 was used for those under age 1.

Data on air pollution for Utah County were obtained from a daily monitoring station maintained by the State of Utah since April 7, 1985. Exposure to PM_{10} air pollution was defined as exposure to particles with an average diameter of 10 μm^3 during any 24 hour period using the U.S. Environmental Protection Agency method. Five-day moving average (i.e., average of PM_{10} observed on the current day and four previous days) was used as a measure of PM_{10} exposure since it was shown to be associated with increased mortality in the previous study conducted in Utah County (Pope et al., 1992). Data on minimum daily temperature were obtained from the National Weather Service. The highest levels of PM_{10} in Utah

County always occurred during the winter, and were always associated with low temperature and high humidity.

Age groups are categorized as <1, 1-14, 15-59, 60-74, and 75+. Location at time of death was defined as a hospital death if it occurred in one of the four hospitals serving the county. A nursing home death if it occurred in one of the nursing homes located in the county, and "home" if the death occurred outside of a nursing home or hospital. Based on the death certificate, 98.2% of deaths outside a hospital or nursing home occurred at home.

Since the number of deaths ranged from 0 to 10 with the mean of 2.52 and the variance of 2.68, the data were analyzed using a Poisson regression model. The generalized estimating equation (GEE) approach, previously used in the analysis of daily mortality and air pollution (e.g., Schwartz et al., 1992), was not used since there is very little correlation among death counts. For example, Lag 1 to 60 autocorrelations were all in the range of -0.05 and 0.05. Consequently, daily death counts were assumed to be independent and were analyzed using the standard Poisson regression program available in EGRET (SERC 1991). Daily minimum temperature was entered as a continuous variable, while PM_{10} was entered as a dichotomous variable using 50 $\mu g/m^3$ as a cutoff (i.e., 1 if above 50 and 0 otherwise). Justification for this dichotomy comes from the dose-response relationship (see Figure 1). A linear time trend was included as a continuous variable to adjust for an increase or decrease in background mortality rates over the years.

RESULTS

There were 7,125 deaths that met the study criteria. The number of deaths and the monthly levels of PM_{10} are shown in Figure 2. Coefficients obtained from the Poisson models for Utah County by Pope et al., (1992) are compared to those obtained from this study in Table 1.

Rate ratios for PM_{10} exposures of 50 $\mu g/m^3$ or greater compared to exposure of less than 50 $\mu g/m^3$ were calculated from the Poisson model for each location at the time of death, and for six age groups at each location. The cutoff level of 50+ $\mu g/m^3$ was chosen after plotting excess mortality by level of exposure. There was a break point at 50 $\mu g/m^3$ indicating a threshold effect (see Figure 1).

The rate ratios for place of death by age are shown in Table 2. For those dying at home the largest RRs were found for those ages <1 to 59. The commonest cause of death of those <1 year old (n=44) was sudden infant death syndrome, and for those in the age 15-59 year old category (n=285), cancer. RRs for deaths occurring in hospital were highest for those age 75+ (n=1585). For deaths in extended care facilities the largest RR was for the age <1, but this is based on two deaths, while those age 60-74 (n=248) contributed the most cases to the excess mortality.

The excess mortality associated with exposure to PM_{10} by calendar year is given in Table 3. The RR was greatest in 1985 and 1990, intermediate in 1986 and 1989, minimal in 1988 and 1992, and absent in 1987 and 1991. None of the RR's was significantly different from all other years combined.

Variations in PM_{10} effect by season are present in Table 4. The largest number of excess deaths occurred during spring (April, May, June), followed by winter (January, February, March).

Causes of death were divided into three broad categories (respiratory, cardiovascular and other), and RRs of $PM_{10} \geq 50\ \mu g/m^3$ were calculated for each cause as well as by year of death and season of death (see Tables 3 and 4). Overall the RRs for cardiovascular deaths were similar to the RRs for pulmonary deaths; however, this varied within each year. The RR for respiratory deaths was highest in 1986 and 1987 when PM_{10} levels were quite low, was below 1 in 1988, a year with high PM_{10}, and was 1.0 in 1990, the year when PM_{10} levels were the highest. The PM_{10} effect for respiratory causes of death was strongest in the spring with an RR of 1.47, while that for cardiovascular causes was greatest in the winter (RR=1.22). When the data were examined by season and location at time of death, the strongest PM_{10} effect was found for deaths at locations other than hospital or nursing home (RR=1.33 for spring, and 1.21 for winter), followed by deaths occurring in hospital (RR=1.09 for spring and 1.05 for winter). (See Table 5.)

DISCUSSION AND CONCLUSION

This study examined the association between PM_{10} and daily mortality in greater detail than previously published studies to assess the causality of the association. The lack of a consistent association for any calendar year, and in the lower RR for the winter season calls into question the causality of the association.

This study confirmed the findings previously reported by Pope et al. (1992) of excess mortality in Utah County associated with fine particulate air pollution. The coefficients obtained from for this study are in agreement with those of the previous study, but the standard errors are slightly bigger. Problems in the calculation of standard error in such studies have been reported previously (Mori et al., 1994).

No previous study has examined the PM_{10} effect by the location at time of death. Our findings suggest an effect for those age <1-59 who died at home (n=366), but these deaths were principally due to "other" than respiratory causes such as sudden infant death syndrome and cancer. The largest overall contribution to the excess deaths was from those age 75+ who died in the hospital (n=1585). There is no information on PM_{10} levels in the hospital compared to PM_{10} levels at home or outdoors. Hospital ventilation systems usually draw in outside air, condition it, and then pass it to the wards. The airflow is generally from patient rooms to the corridors. No information is available for PM_{10} levels in intensive care units, patients on respirators, in isolettes or receiving

supplemental oxygen, making it difficult to draw conclusions regarding the effect of outside PM_{10} levels on hospital mortality.

Excess mortality was not present every year, and excess deaths from respiratory disease, the cause believed to be most strongly influenced by PM_{10} exposure, followed a pattern independent of overall excess mortality by year. Further, this study found no effect of PM_{10} exposure at levels of less than 50 $\mu g/m^3$ contrary to findings from previous studies (Schwartz, 1993).

The lack of an association in certain years raises a question about the causal nature of PM_{10} air pollution and subsequent mortality. The lack of association (RR=1.0) in 1987 might be explained by a "threshold effect" where so many days of high PM_{10} exposure is necessary before a mortality effect occurs. PM_{10} levels in 1987 were the lowest of all the years studied. However, this explanation cannot account for the lack of association in 1991 (RR=1.0) and 1992 (RR=1.01) when PM_{10} levels were as high as any recorded in 1985, the year with the highest RR (RR=1.20). Nor can this explanation account for the RR in 1990 (RR=1.20), a year that had the second lowest PM_{10} levels after 1987.

The RRs for season are also perplexing. The highest levels of PM_{10} were reported in the winter (January and February), but the strongest effect occurs during the spring (RR=1.13), followed by winter (RR=1.08). There is no effect seen in the six months from July to December (RR=0.90 and 0.97). When RRs were calculated for each month the strongest effect occurred in May.

The cause of death was assigned by the State of Utah, and no effort was made to check its accuracy or recode the death

certificates. The location of death is recorded on the death certificate, and this was used to determine location at time of death. For Utah County, 50% of all deaths occur in the hospital, an additional 25% in a nursing home, and the remainder at home. Using Medicare data for those 65+, half of these individuals had been in the hospital for five or more days.

In conclusion, a more rigorous examination of the association between exposure to PM_{10} air pollution and daily mortality found inconsistencies in the association suggesting that it is not causal in nature. These inconsistencies included years when the association was not present, and a season effect in spring rather than the winter. This pattern was not related to the level of PM_{10} exposure in any of the years. The strongest association was found in the spring months (April-June), not the winter months when Utah Valley routinely experiences the highest levels of PM_{10} exposure. The effect was strongest in those at home dying of causes other than cardiopulmonary. These findings suggest that the association between PM_{10} air pollution and acute mortality is not causal, and is likely related to a confounding factor not adequately controlled in previous studies.

BIBLIOGRAPHY

Mori M., Lyon J. L., Carter R. A., Self, S. G. 1994. Evaluation of statistical method for assessing association between daily mortality and air pollution. Submitted.

Pope C. A., Schwartz J., Ransom, M. R. 1992. Daily mortality and PM_{10} pollution in Utah Valley. Arch Environ Health. 47:211-217.

Pope C. A., Dockery, D.W., Spengler, J.D., Raizenne, M.E. 1991. Respiratory health and PM_{10} pollution: a daily time series. Am Rev Respir Dis. 144:668-74.

Statistics and Epidemiology Research Corporation (SERC). EGRET. Seattle, Washington, 1991.

Schwartz J., Marcus A. 1990. Mortality and air pollution in London: A time series analysis. Am J Epidemiol. 131:185-195.

Schwartz, J. 1991. Air pollution and daily mortality in Detroit. Environ Res. 56:204-13.

Schwartz, J., Dockery, D.W. 1992a. Particulate air pollution and daily mortality in Steubenville, Ohio. Am J Epidemiol. 135: 12-19.

Schwartz, J., Dockery, D.W. 1992b. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis. 145:600-4.

Schwartz, J. 1993. Air Pollution and daily mortality in Birmingham, Al. Am J Epidemiol. 137: 1136-47.

Table 1. Comparisons of the Results Reported by Pope et al. (1992) and those of the Present Study: Coefficients and Standard Errors from Poisson Regression[1]

Variable	Pope et al. 1985 - 1989 Coefficients (Std. Error)	Lyon et al. 1985 - 1989 Coefficients (Std. Error)	Lyon et al. 1985 - 1992 Coefficients (Std. Error)
Low temp. (< 20 °F)	0.1851 (0.0495)	0.1143 (0.0717)	0.1619 (0.0541)
Low temp. (20-29 °F)	0.3427 (0.0357)	0.2403 (0.0619)	0.2185 (0.0475)
Low temp. (30-39 °F)	0.1967 (0.0390)	0.1556 (0.0591)	0.1698 (0.0445)
Low temp. (40-49 °F)	0.2442 (0.0247)	0.1892 (0.0583)	0.1740 (0.0450)
Low temp. (50-59 °F)	0.1279 (0.0332)	0.0763 (0.0601)	0.0809 (0.0456)
Time	-0.0147 (0.0118)	-0.0192 (0.0113)	0.0014 (0.0053)
PM_{10} ($\mu g/m^3$)	0.00147 (0.00031)	0.0016 (0.00052)	0.0012 (0.00042)

[1]Relative humidity was included as a covariate in the study by Pope et al. (1992) but not in the present study. PM_{10} was treated as a continuous variable for comparitive purposes.

Table 2. Risk Ratio of $PM_{10} \geq 50\ \mu g/m^3$ by Age Group and Location at Time of Death[2]

	Hospital		Nursing Home		Home + Other		Total	
Age Group	N	RR (95% C.I.)	N	RR (95% C.I.)	N	RR (95% C.I.)	N	RR (95% C.I.)
< 1	171	0.957(0.658,1.392)	2	3.719(0.140,98.67)	44	1.335(0.655,2.717)	217	1.039(0.747,1.443)
1 -14	69	1.078(0.573,2.027)	9	0.963(0.189,4.913)	36	1.270(0.546,2.954)	114	1.129(0.696,1.831)
15 - 24	30	0.425(0.133,1.348)	3	-	11	2.543(0.670,9.642)	44	0.727(0.320,1.654)
25 - 59	436	0.881(0.687,1.131)	53	1.046(0.489,2.233)	274	1.313(0.975,1.768)	763	1.035(0.860,1.245)
60 - 74	1154	0.906(0.777,1.055)	248	1.179(0.859,1.618)	549	1.005(0.806,1.254)	1951	0.965(0.859,1.085)
75 +	1585	1.158(1.022,1.312)	1297	1.000(0.868,1.152)	1154	1.058(0.911,1.229)	4036	1.078(0.995,1.166)
Total	3445	1.015(0.931,1.106)	1612	1.028(0.9057,1.166)	2068	1.092(0.976,1.220)	7125	1.039(0.979,1.104)

[2]From April 7, 1985 - December 31, 1992. Estimated by Poisson regression model adjusted for year and temperature.

Table 3. Risk Ratio of $PM_{10} \geq 50\ \mu g/m^3$ by Cause of Death and Year[3]

Year	Respiratory N	Respiratory RR (95% C.I.)	Cardiovascular N	Cardiovascular RR (95% C.I.)	Other N	Other RR (95% C.I.)	Total N	Total RR (95% C.I.)
1985	66	0.911(0.498,1.666)	330	1.355(1.062,1.730)	238	1.060(0.780,1.438)	634	1.192(0.994,1.429)
1986	96	1.301(0.818,2.066)	418	1.155(0.921,1.449)	396	0.977(0.766,1.246)	910	1.092(0.934,1.277)
1987	83	1.457(0.865,2.451)	406	0.966(0.752,1.241)	400	0.921(0.711,1.194)	889	0.987(0.833,1.170)
1988	89	0.765(0.472,1.243)	430	1.158(0.938,1.43)	382	0.832(0.660,1.051)	901	0.970(0.836,1.126)
1989	88	0.932(0.529,1.640)	376	1.030(0.780,1.360)	425	1.000(0.762,1.313)	889	1.006(0.837,1.209)
1990	89	1.026(0.538,1.956)	413	1.310(0.979,1.751)	433	1.126(0.838,1.513)	935	1.195(0.981,1.455)
1991	104	0.789(0.448,1.390)	424	1.082(0.824,1.421)	459	0.958(0.728,1.26)	987	0.996(0.830,1.196)
1992	94	1.111(0.634,1.947)	414	1.023(0.777,1.347)	472	0.968(0.747,1.257)	980	1.005(0.840,1.202)
Total	709	1.034(0.855,1.248)	3211	1.134(1.039,1.238)	3205	0.947(0.864,1.039)	7125	1.039(0.978,1.104)

[3]From April 7, 1985 - December 31, 1992. Estimated by Poisson regression model adjusted by age and temperature.

Table 4. Risk Ratio of $PM_{10} \geq 50\ \mu g/m^3$ by Cause of Death and Season[4]

Season	Respiratory N	Respiratory RR (95% C.I.)	Cardiovascular N	Cardiovascular RR (95% C.I.)	Other N	Other RR (95% C.I.)	Total N	Total RR (95% C.I.)
Winter	201	1.141(0.8382,1.554)	904	1.215(1.049,1.408)	771	0.930(0.799,1.083)	1876	1.078(0.975,1.191)
Spring	186	1.471(0.874,2.477)	774	1.116(0.843,1.477)	822	1.072(0.810,1.418)	1782	1.131(0.940,1.361)
Summer	155	0.669(0.344,1.303)	716	1.029(0.795,1.332)	776	0.821(0.615,1.097)	1647	0.904(0.752,1.087)
Fall	167	0.826(0.521,1.308)	817	0.899(0.734,1.10)	836	1.063(0.881,1.281)	1820	0.966(0.847,1.102)
Total	709	1.034(0.855,1.248)	3211	1.134(1.039,1.238)	3205	0.947(0.864,1.039)	7125	1.039(0.978,1.104)

[4]From April 7, 1985 - December 31, 1992. Estimated by Poisson regression model adjusted for year, age and temperature.

Table 5. Risk Ratio of $PM_{10} \geq 50\ \mu g/m^3$ by Location of Death and Season[5]

Season	Hospital N	Hospital RR (95% C.I.)	Nursing Home N	Nursing Home RR (95% C.I.)	Home + Other N	Home + Other RR (95% C.I.)	Total N	Total RR (95% C.I.)
Winter	911	1.048(0.908,1.209)	451	1.004(0.820,1.228)	514	1.207(0.994,1.465)	1876	1.078(0.975,1.191)
Spring	863	1.085(0.832,1.415)	394	0.969(0.627,1.479)	525	1.329(0.964,1.830)	1782	1.131(0.940,1.361)
Summer	752	0.811(0.615,1.069)	372	1.044(0.708,1.537)	523	0.955(0.692,1.318)	1647	0.904(0.752,1.087)
Fall	919	0.975(0.810,1.174)	395	0.955(0.720,1.267)	506	0.960(0.749,1.230)	1820	0.966(0.847,1.102)
Total	3445	1.015(0.930,1.106)	1612	1.028(0.905,1.166)	2068	1.092(0.976,1.220)	7125	1.039(0.978,1.104)

[5]From April 7, 1985 - December 31, 1992. Estimated by Poisson regression model adjusted for year, age and temperature.

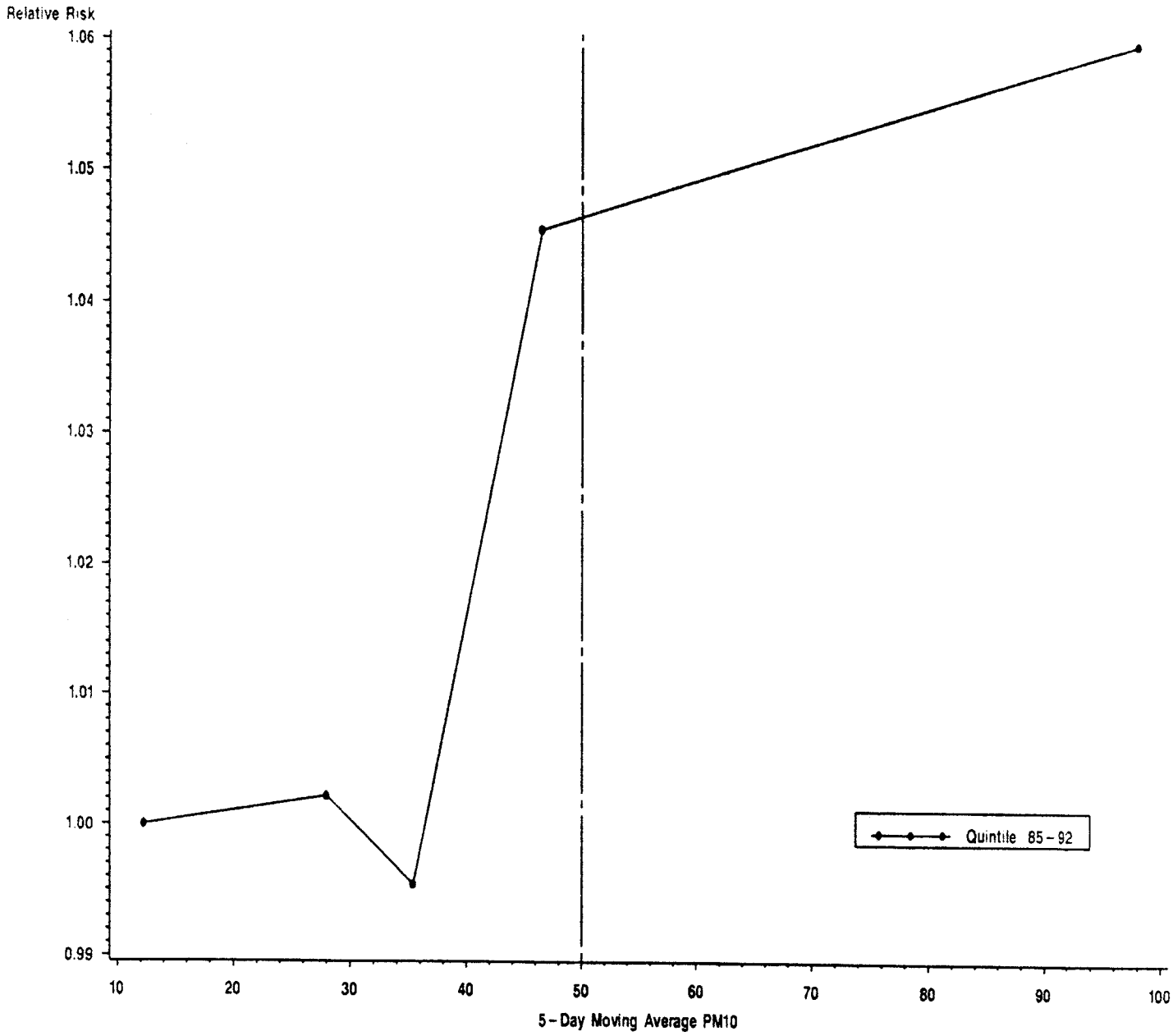

Figure 1 Excess Mortality by Each Quantile of PM_{10} Exposure: Dose-Response Function of PM_{10}

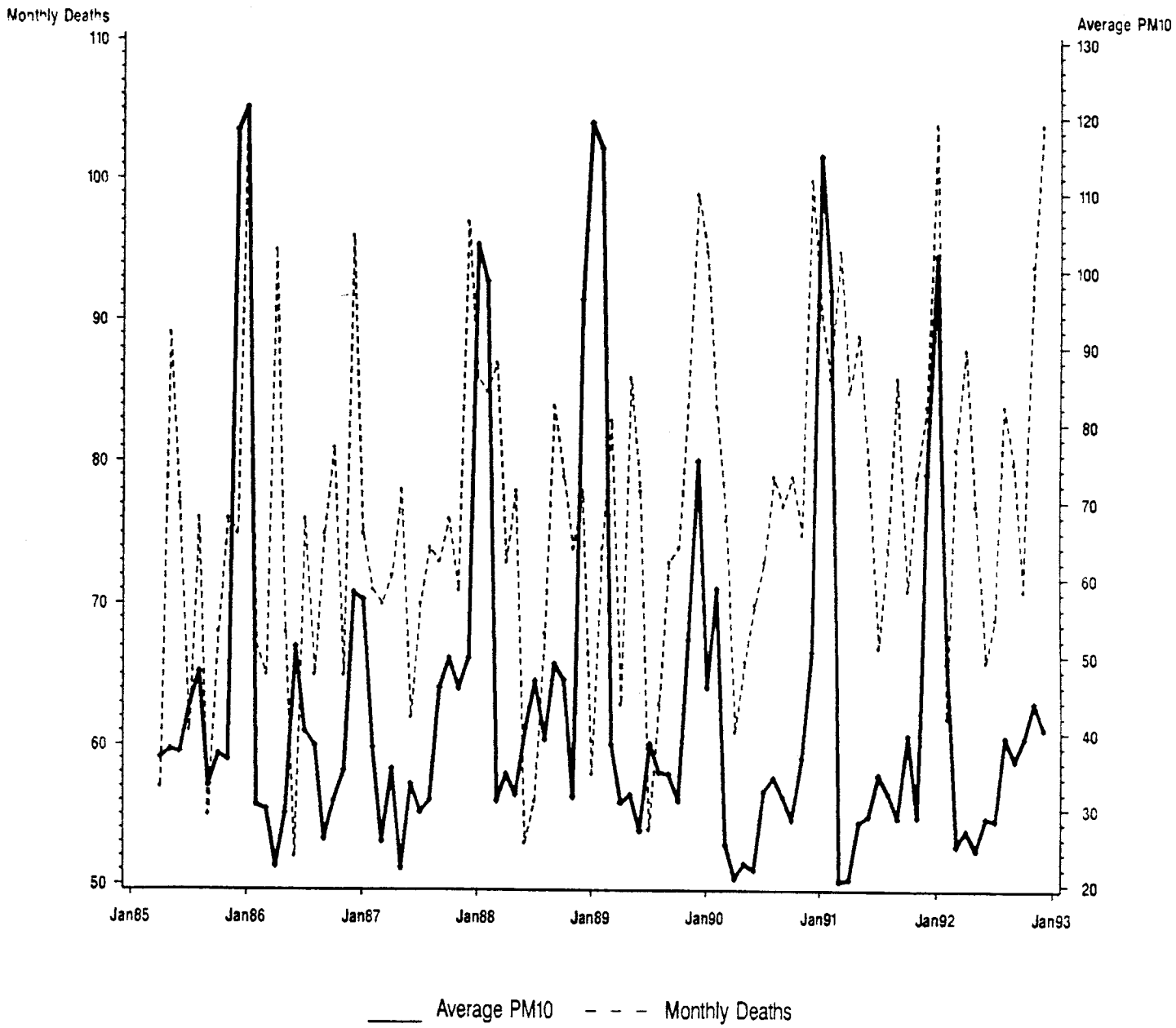

Figure 2 Monthly Deaths and Average PM_{10} Level

PARTICULATE AIR POLLUTION, SULFUR DIOXIDE AND DAILY MORTALITY: A REANALYSIS OF THE STEUBENVILLE DATA

by

Suresh H. Moolgavkar[1], E. Georg Luebeck[1], Thomas A. Hall[2] and Elizabeth L. Anderson[2]

[1] Fred Hutchinson Cancer Research Center

[2] Sciences International, Inc.

Abbreviated Title: Air Pollution and Daily Mortality.

Correspondence to: Suresh H. Moolgavkar
Fred Hutchinson Cancer Research Center, MP 665
1124 Columbia Street
Seattle, WA 98104

Acknowledgements: Research supported by the American Iron and Steel Institute. We are grateful to Sander Greenland for a careful reading of an earlier version of this paper.

INTRODUCTION: Recent studies have reported associations between low levels of air pollution and daily mortality (Schwartz 1991, Schwartz and Dockery 1992a, Schwartz and Dockery 1992b, Schwartz 1993). Taken together, the results of these studies have been interpreted by this group of investigators as providing evidence for a causal role of low levels of air pollution, and, in particular, particulate air pollution in acute mortality. Because the effects, if any, of low levels of air pollution on daily mortality, are small, the results of any single analysis must be interpreted within the broad context of the results from all studies. Robustness of results and consistency across studies are of great importance. According to Schwartz and colleagues the results of many analyses consistently point to a role for particulate air pollution in mortality even after adjusting for the effects of sulfur dioxide and weather. Further the magnitude of the estimated effect is, in their view, remarkably consistent from study to study (Schwartz 1993). We are currently undertaking a systematic re-analysis of the relationship between air pollution and daily mortality in several standard statistical metropolitan areas (SMSA). In this paper, we report on our analysis of air pollution and daily mortality in the Steubenville SMSA.

In our re-analysis of the Steubenville data, we first attempted to replicate the findings of previous analyses (Schwartz and Dockery, 1992a). In order to address the serial correlation of mortality counts from day to day, Schwartz and Dockery used generalized estimating equations (GEE) in their analyses. This approach allows the explicit consideration of correlations in the data by way of a working covariance matrix. However, as discussed below, we found no evidence of correlations among daily mortality counts. As a consequence, we felt that it was unnecessary to use methods such as generalized estimating equations (GEE) for the analysis of these data. Furthermore, we are of the opinion that the sample size is far too small for implementation of the GEE approach, and that straight Poisson regression under the assumption of independent daily mortality counts is preferable as a method of analysis. There are other problems, as well, in the use

of the GEE method for analysis of the Steubenville data. These are discussed in the body of the paper.

While processing the raw data, we realized that there are two distinct daily mortality data series and two distinct sulfur dioxide series that can be used for the analyses. Thus, four distinct mortality-sulfur dioxide combinations are available for analyses. We found that the results of the analyses depend rather strongly on which particular combination of data is used. The series of sulfur dioxide measurements is particularly important. When TSP (total suspended particulates) was the sole air pollution variable included in the analysis, it was significantly associated with mortality on the subsequent day. When the sulfur dioxide measurements used by Schwartz and Dockery were included in the regression, we found, in contrast to their results, that the estimate of the TSP effect was substantially decreased and became statistically insignificant. Finally, when the mortality data were analyzed by season, the effect of particulates was greatly attenuated when the same sulfur dioxide measurements were simultaneously considered in the regression.

MATERIALS AND METHODS:

Because of the sheer volume and complexity of the data, we were concerned that errors could occur while processing the data prior to analysis. We tried to duplicate as closely as possible the data reported in Schwartz and Dockery, but were not entirely successful in doing so, as is discussed below.

Mortality, pollution and weather data: The Steubenville standard metropolitan statistical area consists of the counties of Brooke and Hancock in West Virginia and Jefferson in Ohio. Daily deaths in this SMSA were obtained from the mortality tapes of the National

Center for Health Statistics for the years 1974 to 1984 inclusive. Deaths due to accidents were excluded. There are two mortality data sets that can be analyzed. The first, which we call 'total' mortality, consists of all deaths (excluding accidents) occurring in the SMSA. The other, which we call 'restricted' mortality, consists of all deaths (again excluding accidents) occurring in the SMSA among individuals who are classified as being residents of the SMSA. It is clear that the restricted mortality data form a subset of the total mortality data. It is not quite clear from their paper which mortality data were used by Schwartz and Dockery in their analysis. We were told by Schwartz (personal communication) that they used restricted mortality. When we compared the distribution of mortality counts with the distribution reported in their paper it was clear that neither mortality data set was identical to the data analyzed by them. However, the distribution of the restricted mortality was closer to the distribution reported in that paper. See figure 1.

Air pollution data were obtained from the Aerometric Information Retrieval Service of the Environmental Protection Agency. Table 1 shows the percentile distribution of the TSP readings in the Steubenville area during the period of observation. It is clear from that table that this distribution is in good agreement with that reported in Schwartz and Dockery. Between 1974 and 1978, sulfur dioxide measurements were made in two distinct ways. A continuous monitoring system, which reported hourly measurements was introduced in 1974. In addition, measurements were also available from an integrated gas bubbler sampler, which was run until 1978. Thus, the sulfur dioxide measurement series over the period covered by the analysis consists of either the continuous or the gas bubbler measurements prior to and including part of 1978, and continuous measurements thereafter. We label the two resulting series of measurements as C-C (to indicate that the continuous monitor was used for the entire period 1974-1984) and GB-C (to indicate that the gas bubbler was used from 1974 to 1978 and the continuous monitor thereafter). Because of problems with the newly introduced continuous monitoring system, Schwartz

and Dockery thought it preferable to use the GB-C series in their analyses. As can be seen from table 1, the percentile distributions of both sulfur dioxide series correspond quite closely to the distribution reported by them, with the C-C series distribution being virtually identical to the distribution in their paper. We used both the GB-C series and the C-C series in our analyses.

Information on daily weather was obtained from the National Oceanic and Atmospheric Administration weather station at the Pittsburgh airport. Summary statistics for the weather variables were quite similar to those reported by Schwartz and Dockery.

Methods of analysis: The data were analyzed using the method of Poisson regression. The regression model included an intercept term and additional terms for the weather and pollution related variables of interest. The parameters of the model were estimated by maximizing the likelihood. Because daily mortality counts from day to day are likely to be correlated, the method of Poisson regression may not yield consistent estimates of variances. A recent method, introduced by Liang and Zeger (1986) can then, in principle, be used for estimating the parameters and obtaining robust standard errors for these estimates. This, so called generalized estimating equation (GEE) approach, introduces a working correlation matrix to describe the correlations among the observations. Liang and Zeger showed that, providing the sample size is large enough, the choice of the working correlation matrix does not affect the consistency of the estimates. We examined the auto-correlation function of the mortality data in Steubenville and found, to our surprise, that there was no evidence that mortality counts on any two days were correlated (see figure 2). Thus, we believe that it is unnecessary to use the GEE approach. In fact, the GEE approach may yield misleading results for this data set, as is discussed below. However, in an effort to confirm published results we used GEE with various working covariance matrices in addition to straight Poisson regression for the analyses of the Steubenville

data. In order to compensate for possible extra-Poisson variation in the data, we computed a 'sandwich' estimator (see expression 7 in Zeger and Liang, (1986)) for the variance when using Poisson regression. We found, however, that this 'robust' estimate of variance was comparable to the 'naive' estimate of the variance obtained from Poisson regression.

Finally, a few words on the use of GEE for the analyses of mortality counts are in order because this method has now been applied to many different data sets. This approach was introduced for the analyses of correlated longitudinal data. Correlations in the data are addressed by way of a working covariance matrix, and if the sample size is large enough (i.e. asymptotically) the method yields consistent estimates of the parameters and the standard errors even if the choice of the working covariance matrix was not particularly accurate to begin with. However, for any given (finite) data set a reasonable choice of the working covariance matrix is always to be preferred to one that is clearly inappropriate. In the context of correlated mortality counts, sample size refers, not to the size of the population, but to the number of independent years of observation. Schwartz and Dockery used an exchangeable covariance structure for their analysis. This choice of working correlation matrix is clearly inappropriate because it assumes, for example, that mortality counts on January 1 and December 31 of a given year are correlated to the same extent as the mortality counts on January 1 and January 2 of that year, and that mortality counts on December 31 of that year and January 1 of the next year are not correlated. In fact, if full year mortality is to be analyzed, then mortality over the entire period of observation should be considered to be a single time series. There is little justification for treating single years as independent units of observation. If seasonal mortality is analyzed, on the other hand, then the assumption of independence would appear to be justified: mortality in the summer of 1979 can be considered to be independent of the mortality of the summer of 1978. But even if analysis is performed by season, there is a sample size problem in the use of GEE with these data. For the asymptotic results to hold, the number

of independent years of observation must be large. With only 11 years of observation in Steubenville, the sample size is probably inadequate for application of the GEE approach. We also note that the application of the GEE approach has heretofore been restricted to the analysis of correlated longitudinal data with relatively few observations in each individual longitudinal series. The fact that in the mortality data we have approximately 90 observations in each series (when analysis is done by season) probably makes the sample size problem even worse.

RESULTS: We note first that the various methods of parameter estimation (Poisson or GEE with various working covariance matrices) gave rather similar estimates of the parameters, but quite different standard errors and thus different confidence intervals. The results, both parameter estimates and standard errors, were found to depend upon which particular mortality and sulfur dioxide data sets were used in the analysis. Because of the pressures of space, we are unable to report all our results here. Instead we will focus on reporting the results of the Poisson analysis, and point out how the results differ when GEE is used instead. The results of Poisson regression are presented in table 2 and figures 3 and 4. Table 2 and figure 3 report the results of using the GB-C sulfur dioxide series, which, in the opinion of Schwartz and Dockery, is the desirable series of measurements to use in the analyses. Figure 4 reports the results of using the C-C sulfur dioxide series.

We first report on our analysis of the full-year data using the covariates used by Schwartz and Dockery in their analysis (table 2). In our analysis we define winter as comprising the months of December, January and February and spring as March, April and May. It is not clear how Schwartz and Dockery defined these seasons. When TSP was the only pollution covariate used in the analysis, it was significantly associated with mortality on the next day with a relative risk per $100\mu g/m^3$ = 1.029, 95% CI=(1.007, 1.051). This

result is in good agreement with the result reported in Schwartz and Dockery (1992a, their table 2).

We found that the effect of TSP is greatly attenuated when sulfur dioxide is included in the model. With full mortality, we find, as shown in table 2 and figure 3 that there is a substantial attenuation of the TSP effect when both TSP and sulfur dioxide are included in the model, the estimated parameter for TSP decreasing from 0.029 (s.e = 0.011) to 0.018 (s.e. = 0.013). If instead of using the full mortality data, we use restricted mortality (used by Schwartz and Dockery in their analysis) and the GB-C sulfur dioxide series, we get essentially the same results. The attenuation of the TSP coefficient is not as large, however (the estimated parameter drops from 0.031, s.e. = 0.012, to 0.025, s.e = 0.014). This is in contrast to the conclusion of Schwartz and Dockery who say, "When both sulfur dioxide and particulates were included in the model simultaneously, particulates remained significant ($\beta = 0.000300 \pm 0.000128$), with little reduction in its estimated coefficient, while sulfur dioxide was insignificant ($\beta = 0.0059 \pm 0.0048$) with a substantial attenuation in the estimated effect." If, however, GEE with an exchangeable correlation structure is used with restricted mortality and the GB-C sulfur dioxide series, both TSP (estimated coefficient = 0.026; s.e. = 0.019) and sulfur dioxide (estimated coefficient = 0.042; s.e. = 0.071) are statistically insignificant. When the C-C sulfur dioxide series is used, we can reproduce the finding of Schwartz and Dockery (TSP is significantly associated with mortality and sulfur dioxide is not), as is shown in figure 4. In summary, we can reproduce Schwartz and Dockery's results only if we use the C-C sulfur dioxide series, a series that they explicitly state they did not use because of measurement error problems.

We also analyzed the mortality data by season. We defined the seasons as follows: Winter: December, January, February; Spring: March, April, May; Summer: June, July,

August; Fall: September, October, November. The results of the analysis using both full and restricted mortality and the GB-C sulfur dioxide series are shown graphically in figure 3. It is quite clear from this figure that the introduction of sulfur dioxide into the model attenuates the coefficient for TSP, and vice versa. We also found a significant effect of temperature in Fall and in Winter, mortality being highest on the coldest days (lowest quintile of temperature). In Summer, mortality is highest on hot days (highest quintile of temperature), however the coefficient is not statistically significant. With the C-C series of sulfur dioxide measurements, the effect of TSP is stronger, whether or not sulfur dioxide is included in the model. Sulfur dioxide, however, appears to have a protective effect in Summer (statistically significant if GEE is used). Thus, the use of the C-C series does lead to some rather unexpected results.

A comparison of figures 3 and 4 shows that the choice of sulfur dioxide series used in the analyses makes a rather large difference to the results. The use of the C-C series leads to results that are much closer to those obtained by Schwartz and Dockery.

DISCUSSION: Regression analyses of the mortality data reveal that by far the most important term in the regression is the intercept. Thus, the weather and pollution variables explain only a very small fraction of the daily deaths. In view of the smallness of the effects we are trying to detect, robustness of results is of great importance. The Steubenville results are not robust, however. This is particularly disturbing in view of the fact that the restricted mortality is a subset of the total mortality data set and differs from it only to the extent of about 0.4 deaths per day on average. Similarly, the GB-C and C-C sulfur dioxide series are identical between mid-1978 and 1984, and as table 1 indicates the distribution of the two series is similar. This fact suggests that the results of the analyses are rather sensitive to measurement errors. The results depend, as well, on whether

Poisson regression or GEE is used for the analysis. For the reasons discussed above, we do not believe that GEE should be used for analyses of these data.

Taken together, the results of the Steubenville analyses are consistent with weak associations between weather and air pollution and mortality. However, in view of the instability of the results, consistency with the results of analyses in other cities must be sought before any firm conclusions can be drawn. Certainly, it would be premature on the basis of these data to single out any single component of air pollution as being responsible. In this report we have considered only two of the six criteria pollutants and have shown that introduction of the second pollutant into a regression model results in an attenuation of the coefficient of the first. There is little *a priori* reason for excluding the other criteria pollutants from the analysis. Future research efforts should be directed at investigating combinations of pollutants.

The role of weather needs to be better investigated as well. Our results suggest that pollutants and season may interact. Kalkstein (1991) has suggested that synoptic categories of weather, which take into account many more weather variables than considered in current analyses, may be more strongly associated with mortality than single weather variables. We are currently investigating these ideas in Steubenville. In particular it would be of great interest to consider synoptic categories for weather together with the criteria pollutants in a single regression model. Such an analysis should ideally be performed in a metropolitan area with a large number of daily deaths, and with many years of observation. If daily mortality counts are correlated, then both the GEE approach and bootstrap methods should be used for the construction of confidence intervals.

Finally, although we attempted to assemble and analyze a data set identical to the data analyzed by Schwartz and Dockery, we were not entirely successful in our efforts.

Although we went to the same primary data sources, we were unable to duplicate exactly either the mortality or the pollution data sets. If the differences in our results are due to the small differences in the data analyzed, they reinforce a point made earlier, namely that the results of these analyses are not robust to small perturbations of the data.

REFERENCES:

Kalkstein LS: A new approach to evaluate the impact of climate on human mortality. Environmental Health Perspectives 96: 145-150, 1991.

Liang KY, Zeger SL: Longitudinal data analysis using generalized linear models. Biometrika 73: 13-22, 1986.

Schwartz J: Particulate air pollution and daily mortality in Detroit. Environmental Research 56: 204-213, 1991.

Schwartz J: Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol 137: 1136-1147, 1993.

Schwartz J, Dockery DW: Particulate air pollution and daily mortality in Steubenville, Ohio. Am J Epidemiol 135: 12-19, 1992a.

Schwartz J, Dockery DW: In creased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145: 600-604, 1992b.

Zeger SL, Liang KY: Longitudinal data analysis for discrete and continuous models. Biometrics 42: 121-130, 1986.

Table 1: Percentile distributions of sulfur dioxide and total suspended particulates (TSP). Refer to the text for definitions of GB-C and C-C. S - D refers to the distributions reported in Schwartz and Dockery (1992).

SULFUR DIOXIDE (ppb) PERCENTILE DISTRIBUTION

	10	25	50	75	90	Mean
GB -C	7	13	23	39	58	28.8
C - C	7	12	23	38	56	28.6
S - D	7	12	23	38	55	28

TSP (micrograms per cubic meter) PERCENTILE DISTRIBUTION

	10	25	50	75	90	Mean
This study	38	58	93	144	212	113
S - D	36	56	91	139	209	111

Table 2: Results of Poisson regression analysis of the Steubenville total mortality data set using the covariates used in a previous analysis (Schwartz and Dockery, 1992a). Indicator variables for winter and spring (defined in text). Hot and humid defined as 24-hour mean temperature > 21.1°C and mean dew point temperature > 18.3 ° C. Coefficient for TSP represents effect of an increase of 100μg/m^3 in the previous day's mean. Coefficient for SO_2 represents effect of an increase of 100ppb in previous day's mean. Individual years were controlled for in the analysis by indicator variables.

VARIABLE	COEFFICIENT (S.E)	RELATIVE RISK (95% C.I.)
Intercept	1.174 (0.032)	
TSP	0.018 (0.013)	1.018 (0.992 - 1.044)
SO_2	0.072 (0.048)	1.072 (0.978 - 1.181)
Winter	0.080 (0.022)	1.080 (1.038 - 1.130)
Spring	0.020 (0.021)	1.020 (0.979 - 1.063)
Hot and humid	0.023 (0.042)	1.023 (0.942 - 1.111)

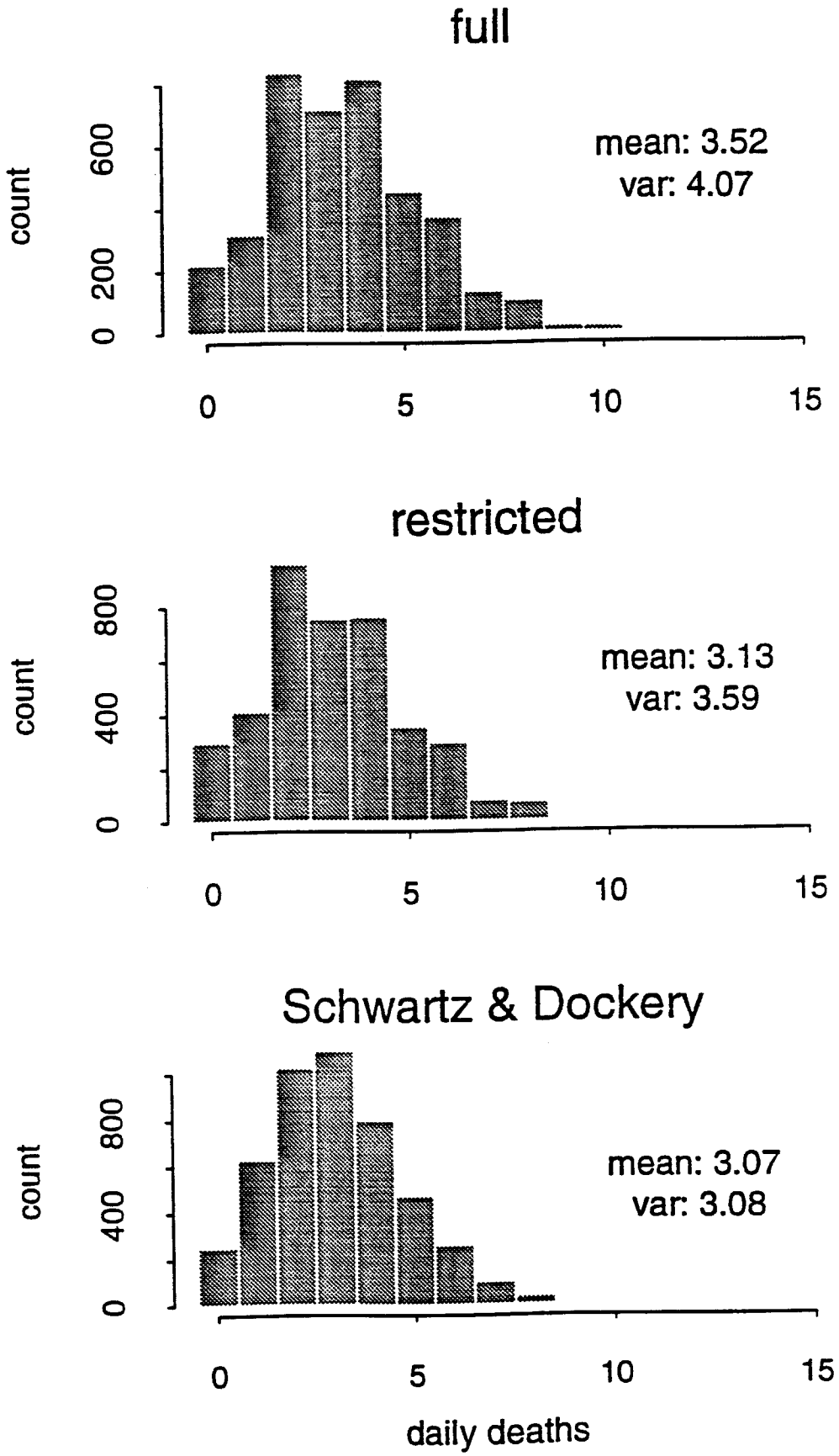

Figure 1: Distribution of daily mortality in the Steubenville SMSA between 1974 and 1984. For definitions of 'full' and 'restricted' see text.

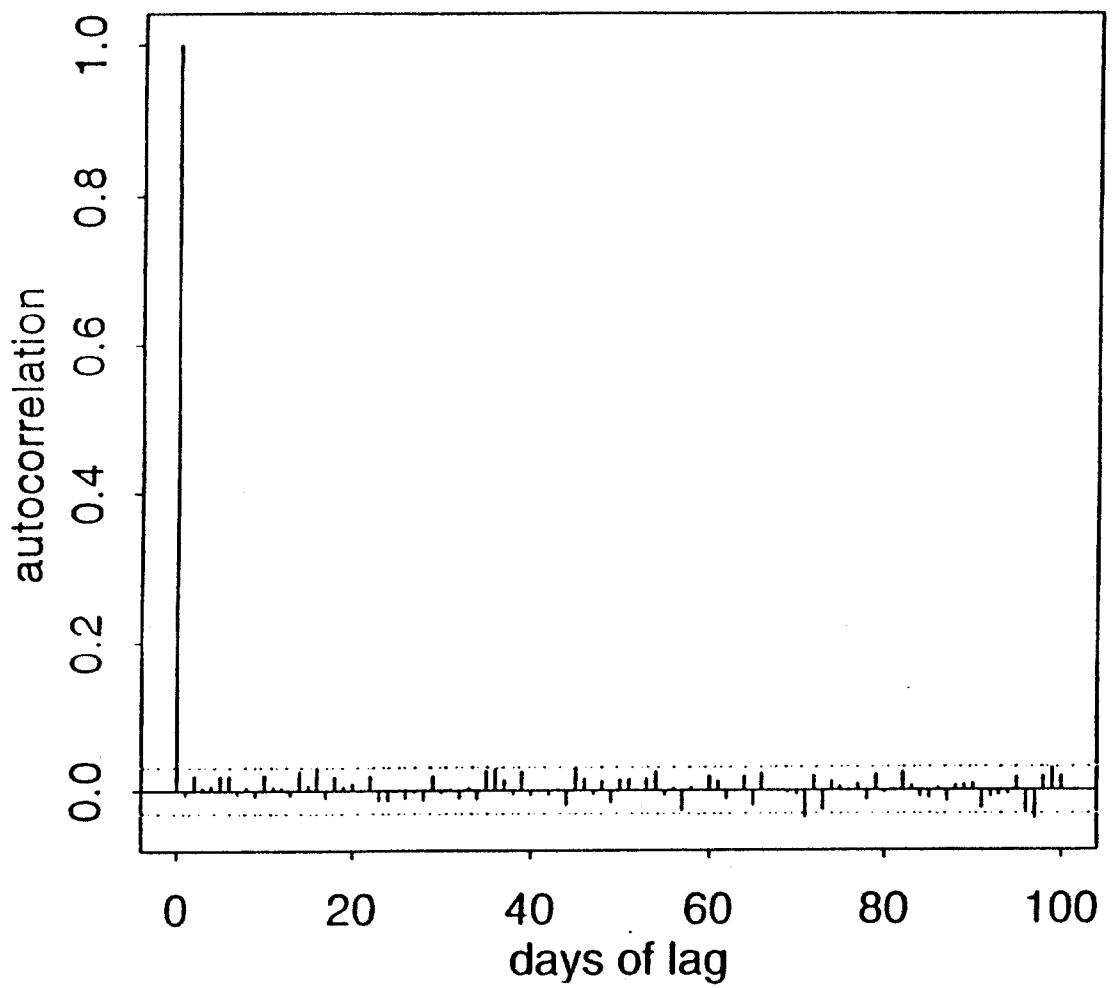

Figure 2: Autocorrelation function of the daily total mortality series in Steubenville. Correlation between mortality counts on days separated by various lag times is shown. Thus, correlation for a lag of 0 is 1, indicating that the mortality count on any given day is perfectly correlated with itself The small negative correlation at lag 1 indicates that the mortality counts on any two consecutive days are estimated to have a small negative correlation. The dotted lines represent the 95% confidence intervals for the correlations

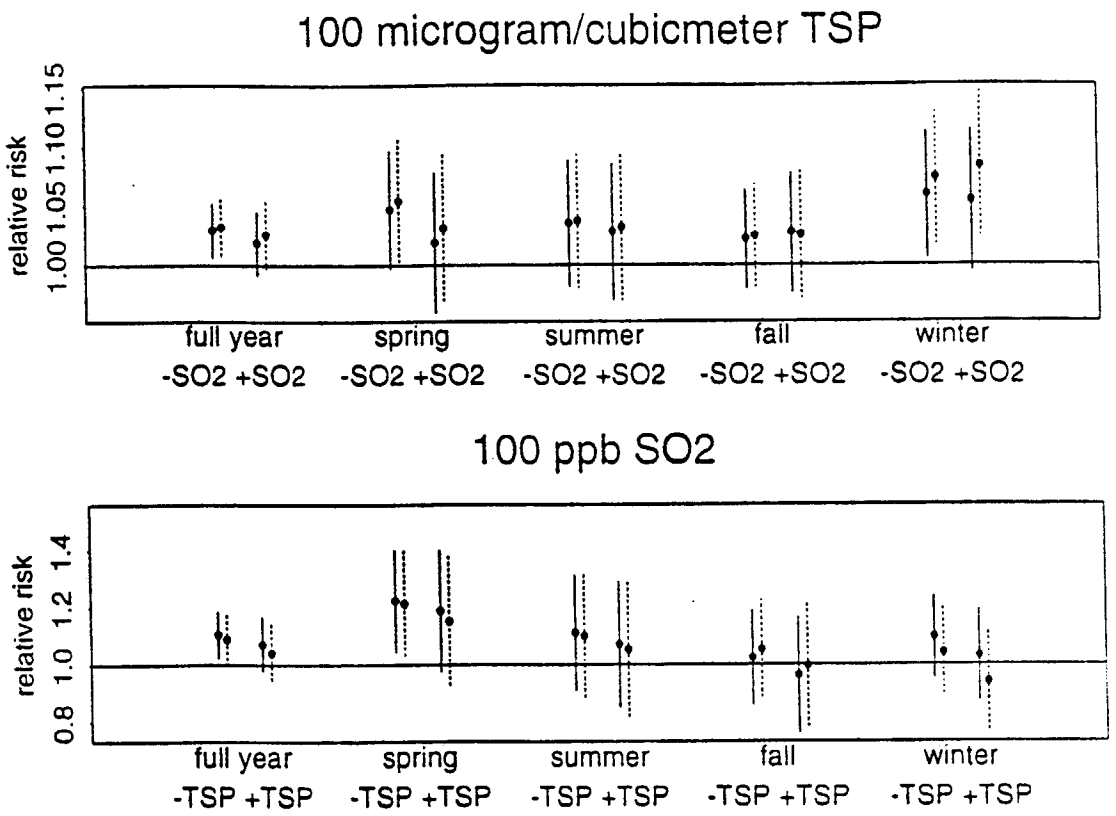

Figure 3: Relative risks of mortality associated with exposure to TSP and sulfur dioxide using the GB-C series of measurements. The relative risks are derived by fitting Poisson regression models to the data. For full year mortality, the covariates described in table 2 are included in the analysis For analysis by season, in addition to the pollution parameters, quintiles of temperature are included and individual years are controlled for by indicator variables. In all analyses, the solid line represents results from full mortality, the dotted line results from restricted mortality. Top panel shows relative risk of mortality associated with 100μg/m^3 of TSP, together with the 95% confidence interval. For full year and seasonal analyses, the relative risk with TSP as the only pollution variable in the regression is shown on the left and labelled -SO_2; the effect of introducing sulfur dioxide into the regression on the relative risk associated with TSP is shown on the right and labelled +SO_2 Lower panel shows the relative risk of mortality associated with 100 parts per billion sulfur dioxide. For full year and seasonal analyses, the relative risk with sulfur dioxide as the only pollution variable in the regression is shown on the left and labelled -TSP; the effect of introducing TSP on the relative risk associated with sulfur dioxide is shown on the right and labelled +TSP.

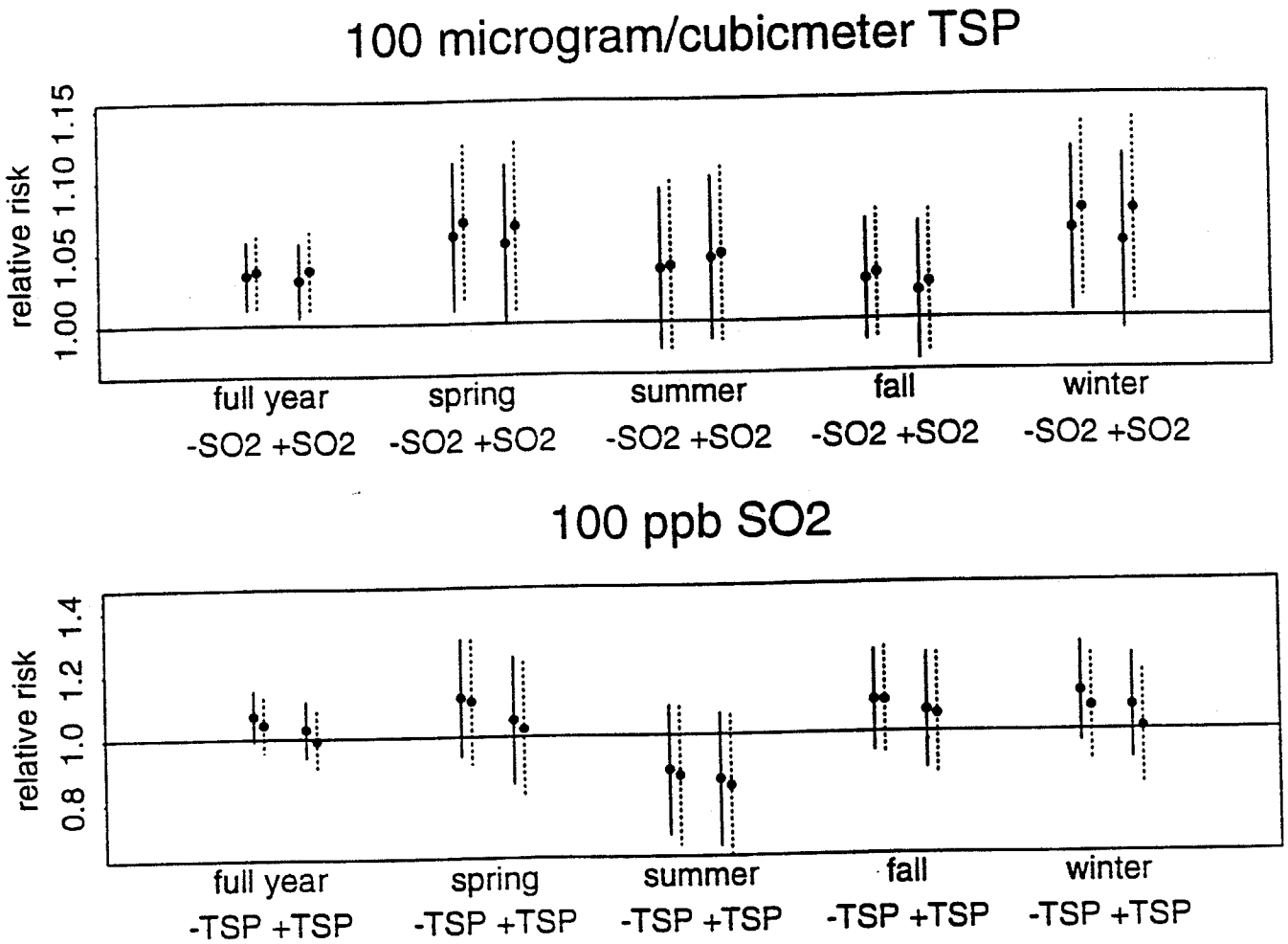

Figure 4: Same as figure 3, except that the C-C series of sulfur dioxide measurements is used.

DAILY MORTALITY ANALYSIS BY USING DIFFERENT REGRESSION MODELS IN PHILADELPHIA COUNTY, 1973-1990

by Yuanzhang Li and H. Daniel Roth

Yuanzhang Li is senior statistician at Roth Associates, Inc.

H. Daniel Roth is president and founder of Roth Associates, Inc.

ABBREVIATED TITLE: PHILADELPHIA MORTALITY

Please address all correspondence to:

Yuanzhang Li, Roth Associates, Inc., 6115 Executive Blvd., Rockville, MD 20852; telephone 301-770-4405; fax 301-770-9248

ABSTRACT

Earlier investigators have found that in the city of Philadelphia there exists an association between daily levels of particulate matter and mortality (Schwartz and Dockery, 1992a). To more fully explore this result, we have analyzed the daily mortality data in Philadelphia using a host of independent variables including air quality indicators such as TSP, SO_2, and O_3, and weather factors such as temperature, humidity, barometric pressure, and precipitation for the years 1973-1990. A wide variety of appropriate statistical models were applied to the data, including Poisson, multiple, and autoregression models. Our major findings on the association between mortality and daily pollution levels are that: (1) the estimates of this association are related to the selection of the type of statistical model and the variables included in the model; and (2) this relationship differs across age groups, causes of death, and seasons of the year. These results are not consistent with a unique point estimate of the association between air quality and mortality.

INTRODUCTION

Several studies have examined the association between daily mortality and air pollution in a host of cities across the United states and Europe (e.g. Wyzga, 1978; Ostro, 1984; Hatzakis et al., 1986; Ozkaynak and Thurston, 1987; Shumway et al., 1988; Fairley, 1990; Kalkstein, 1991; Kinney and Ozkaynak, 1991; Schwartz, 1991 and 1993; Schwartz and Dockery, 1992a and 1992b; and Pope et al., 1992.) The studies have differed by the data considered in the analysis as well as by the statistical models used to examine the data.

For example, some studies considered a full complement of air quality variables, whereas others just took into account a single air pollutant such as total suspended particulates. Also some studies examined only total deaths, whereas others looked at specific causes of death, paying particular attention to those more likely to be associated with air pollution such as chronic obstructive pulmonary disease, cardiovascular disease, and lung cancer. In addition, some studies used multiple linear regression models, others used Poisson models, still others used autoregression models, etc.

In this paper we analyze the Philadelphia data from 1973-90, paying particular attention to daily particulate matter levels but also taking into account levels of sulfur dioxide and ozone.

Philadelphia was analyzed because: (1) the city has a large enough count of daily deaths for a meaningful daily mortality analysis, and (2) daily data on particulate matter and other critical pollutants are available for analysis from sources such as EPA.

Other investigators who have analyzed Philadelphia data are Wyzga (1978) and Schwartz and Dockery (1992a). Our analysis of the Philadelphia data differs from these earlier analyses in several ways. First, we consider a greater number of years of data. Wyzga analyzed the data from 1957 to 1966, and Schwartz and Dockery from 1973 to 1980, as compared to our analysis which considers the data from 1973 to 1990. Also our analysis examines a greater complement of variables. Schwartz and Dockery examined sulfur dioxide and particulate matter; Wyzga looked at only average coefficient of haze (COH). As stated above, we investigated sulfur dioxide, ozone, and particulate matter. Finally, we subjected the data to a larger battery of statistical analyses than were applied by previous investigators in the literature. Wyzga analyzed the data using a regression model; Schwartz and Dockery used the Poisson model.

DATA

Data on mortality which included the age, the race, and the gender of the descendant and the place of death, the date of death, and the cause of death by ICD code (International Classification of

Diseases, Revision 9) were obtained for the years 1973 to 1990 from the Department of Health in the Commonwealth of Pennsylvania. For further analysis, the data were broken down by age (65 and older, and under 65), and by the following disease categories: cancer (ICD9 140-208), chronic obstructive pulmonary disease (COPD) (ICD9 490-496), cardiovascular disease (CVD) (ICD9 390-448), pneumonia (ICD9 480-486 and 507), and other diseases. Accidental deaths (ICD9 ≥800) were excluded from all analyses.

Weather data, which included maximum daily temperature (mxtp) (1973-1990), minimum temperature (tmin) (1973-1990), mean temperature (mntp) (1984-1990), daily precipitation -- rainfall and melted frozen precipitation (prcp) (1973-1990), average daily dewpoint (dewp) (1973-1990), barometric pressure (pres) (1973-1990), minimum relative humidity (1984-1990), maximum relative humidity (1984-1990), and average relative humidity (mnhumid) (1973-1983) were obtained from the National Climatic Data Center, Asheville, North Carolina. The missing years of data on mean temperature (1973-1983) did not pose a problem because this variable was not used in any of our main analyses; it was considered only when replicating the works of others. The missing data on mean humidity (i.e. from 1984-1990) were imputed by averaging the daily maximum and minimum levels of relative humidity.

The air pollution data, which were obtained from the National Air Data Branch, United States Environmental Protection Agency,

included: daily SO_2 (part of 1973, 1974-1990), daily O_3 (part of 1973, 1974-1990), and total suspended particulates (TSP) (1973-1990). The few days for which there were missing data were excluded from the analysis. The distributions of TSP, SO_2, O_3, maximum temperature, mean humid, dewpoint, and daily mortality by age are summarized in Table 1. Distributions of daily mortality by cause of death are given in Table 2. The correlations among daily mortality for people 65 and older, previous day's air pollutants, and previous day's weather factors are given in Table 3.

Several observations can be made from examining the distributions of the variables in the study (Table 1). To begin with, the levels of particulate matter in Philadelphia probably were in compliance with or close to attainment of the National Ambient Air Quality Standards for particulate matter. The 95 percentile level of TSP levels was 120 $\mu g/m^3$, which is well below the annual National Ambient Air Quality Standard of 260 $\mu g/m^3$ up until 1987. Furthermore, a more detailed examination of the data shows that almost all of the higher levels in the study came from the period of the 1970s. For example, the number of days with TSP levels over 150 $\mu g/m^3$ is 33 from 1973-1974, 29 from 1975-1979, and only 5 from 1980-1990.

Another item to observe is that the daily mortality levels in Philadelphia are not distributed according to the Poisson distribution, an important point because several investigators have based their analyses on the validity of this assumption. If the data are Poisson distributed, then the variance of the distribution should equal the mean. For the Philadelphia data, this does not seem to be the case for individuals 65 and older with the variance of daily deaths equaling 51 (7.13^2, Table 1) as compared to a mean of 35. Likewise, for individuals under 65, the variance is 25 compared to a mean of 19.

It should also be observed that there were many days in which there were no deaths due to chronic obstructive pulmonary disease and pneumonia (Table 2). For this reason, in examining these data we used a square root transformation instead of a logarithmic transformation in some of our analyses using the Poisson model. (For a more complete discussion of models used in the analysis, see below).

Much can also be gleaned from a correlation analysis of the data (Table 3). To begin with, some weather variables were more highly correlated with daily mortality than they were with TSP. For example, the correlation between maximum temperature and mortality was -0.28 compared to 0.05 which was the correlation between TSP and mortality. Another weather variable highly correlated with mortality was dewpoint (-0.26).

The problem of separating out the pollution effects from the weather effects was further complicated by the fact that some weather variables were highly correlated with each other as well as with pollutant variables. For instance, maximum temperature and dewpoint were correlated at the 0.92 level; dewpoint and mean humidity at the 0.54 level; TSP and maximum temperature at 0.26; and ozone with maximum temperature at 0.64.

Finally, the three pollutants considered in the study were also highly correlated with each other. TSP and sulfur dioxide were correlated at the 0.57 level; ozone and sulfur dioxide were correlated at the -0.26 level; and ozone and TSP at 0.18. All of these levels were significant at the .01 level or lower.

METHODS

Several statistical models exist for analyzing the association between daily mortality and air pollution. Ostro (1984) and Kinney and Ozkaynak (1991) applied a regression model to deviations of moving averages of daily mortality. Hatzakis et al. (1986) used a trigonometric model first to fit the monthly average daily mortality, and used a regression model on the residuals. Shumway et al. (1988) used a regression model on total mortality and a non-linear regression on smoothed mortality. Fairley (1990) used the Poisson-square root regression models on daily mortality. Finally, Schwartz (1991, 1993), Schwartz and Dockery (1992a,b), Dockery et

al. (1992), and Pope et al. (1992) used variations of the Poisson-log regression model. Also different investigators considered different combinations of air quality and meteorological data (e.g., one to five days' weather, the previous day's pollution, or the average of two to five days' pollution, etc.)

Each of these models used by others is designed to smooth out aberrations (i.e., factors unrelated to pollution) in changes of daily mortality. For example, the trigonometric approach fits the data to periodic functions that are known to characterize many patterns in nature. On the other hand, the deviation from the moving average model is a well-known technique designed to accomplish the same thing. The objective of our analyses was to apply a spectrum of statistical models used by other investigators to determine whether they would yield consistent results. Unfortunately, the results generated by these different models frequently differ, and there is no solid justification for accepting the finding from one model over another. Since the results about the association between daily mortality and air pollution are not consistent, the t-values are compared in this study. The discussion of insignificant or inconsistent β-coefficients is meaningless.

A description of the models we used is given in Table 4. In summary, we used Poisson regression, multiple, and autoregression models, and used two different formulations for the Poisson model:

a logarithmic transformation and a square root transformation. To adjust for additional time trends in the data, we used deviations from moving averages (7, 15, and 29 days), deviations from averages of that day of the year, and deviations from trigonometric curves.

In the model, the current and previous day's pollutant variables and several weather variables (such as temperature, dew-point, and relative humidity) were always considered to be continuous. On the other hand, following the lead of other investigators (e.g. Schwartz and Dockery), we defined the indicated variables such as "cold day," days in which the maximum temperature was in the lower 5% percentile (or less than 31 degrees; see Table 1). "Hot days," "humid days," and "hot-humid days" were defined in a similar manner. Other indicated variables were year, which was incorporated to adjust for long-term changes in daily mortality, and season. In some analyses a continuous variable -- "day of year" -- was used to adjust for time trends within a year. The data were also stratified by season and analyzed separately.

Lastly, following the lead of previous studies, in some analyses the average of two days' air quality was used, and in other analyses just the previous day's levels were considered. Instead of using "barometric pressure" as a variable directly, the increase (or decrease) in barometric pressure from the previous day (δ-barometric) was used because the δ-barometric pressure, which is significant in many models, was found to be more highly correlated

with daily deaths than the individual daily pressure values themselves. For the same reason, we used the difference in precipitation.

FINDINGS

The Influence of Season

The significance of the association between pollutants and mortality varied by age, by season, and by the number of pollutants in the model. The results from the Poisson regression with logarithm transformation analysis for the group 65 and older and for the group under 65 are given in Table 5. For the group 65 and older, the t-value of TSP ranged from -0.02 (in winter, when TSP and SO_2 are incorporated into the model) to 2.83 (also in winter), (see Table 5). TSP was a statistically significant predictor of mortality only in the winter and only in a model where no other pollutants were considered; the associated coefficient is 0.0005.

Regarding the other seasons, there is no indication of a TSP effect. The t-values in the summer ranged from 0.48 to 1.72; in the spring from 0.60 to 0.98; and in the fall from 0.93 to 0.99. TSP was never a significant predictor of mortality in the under-65 age group. Similarly, there was much variation in the predictive power of other pollutants in the seasonal analyses.

Analysis by Cause of Death

Total suspended particulates (TSP) was never a statistically significant predictor of mortality in any of our analyses by cause of death (Table 6). In the analysis of individuals 65 and older, the t-values of TSP were always positive for cancer and pneumonia; the t-values of TSP for CVD were consistently negative; and the t-values of TSP for COPD were mixed. The t-value of TSP for COPD was positive in the models where two or more pollutants were considered, but it was negative in the model where TSP alone was analyzed. In the analysis of the under-65 age group, the t-values of TSP for pneumonia and other disease were always negative, and the t-values of TSP for cancer, CVD, and COPD were mixed.

Averaging Time Considered in the Analysis

Different investigators have used different pollution level averaging assumptions. For example, in Schwartz (1991) daily mortality was compared to previous day's pollution levels; in Schwartz and Dockery (1992a) the mean of pollution in the previous day and the current day was used as the indicator; in Schwartz (1993) the mean of the current and the previous two days was used; in Pope et al. (1992) the mean of the current day and the previous four days was used. In our analysis of the data, we looked at the average from 1 to 4 days (Table 7).

In the analysis, the coefficients and their associated t-values are highest in the model using the mean of two days as the indicator. In the analysis where two days' weather is incorporated, the t-value of TSP is 2.21, 1.31 in the one-day analysis, and 0.38 or 1.02 in the three-day analyses. In addition, the t-values are also much higher in the two-day analysis even when one or two pollutants are considered. We do not know which pollutant averaging time is most desirable, but it is important to consider averaging time when comparing results across studies.

The Influence of Statistical Models

As we discussed earlier, a host of statistical models were applied to the data (Table 4). Also, three sets of analyses were examined: TSP, SO_2, and O_3 were incorporated into the model; TSP and SO_2 were used; and TSP alone was considered. The resulting t-values from these analyses (in the case where two days' weather variables were incorporated into the model) are given in Table 8 (65 and older) and Table 9 (under 65). Analyses of these tables show that:

1. The significance of a pollutant predictor of daily mortality changes from model to model. For example, in the group 65 and older, the t-value of TSP varies from 0.1 to 6.6.

2. In the models containing TSP only, the t-value of TSP is much higher than that in the models containing TSP and SO_2, or TSP, SO_2, and O_3 simultaneously. This phenomenon occurs because all pollutants are strongly confounded. When TSP, SO_2, and O_3 are analyzed simultaneously, the association between the mortality and TSP is generally not significant.
3. In the analyses of the group 65 and older where TSP is the only pollutant, TSP is significant in most cases. This may be due to the fact that TSP is highly correlated with other pollutants.
4. When all three pollutants are in the model, the significance of O_3 is stronger than that of SO_2, and the significance of SO_2 is stronger than that of TSP.
5. When O_3 is excluded from the model, the significance of TSP is stronger than that of SO_2 in some cases, and weaker in others.
6. In the under-65 group, the effects of pollutants are much less significant, and in many cases they are negative.

It should be observed that the models employed had relatively low predictive power as indicated by the R^2, the coefficient of the multiple determination. This statistic measures the proportion of the reduction of total variance in daily mortality by applying a regression model.

The R^2 of Poisson models with two days' weather was 0.15-0.16 for people 65 and older, and 0.17-0.18 for people under 65. Eliminating the pollutant variables from the analysis reduces the R^2 to 0.14. In an analysis with the three pollutants alone, the R^2 was 0.04 for people 65 and older, and 0.06 for people under 65. If a two-stage regression analysis is used (i.e., an examination of the residuals after forcing in non-pollutant variables), the R^2 is less than 0.005 for both age groups. Given the uncertainty of the model, it is unclear how much can be explained by pollution.

The Influence of Interaction

By using the Poisson-log model with two days' weather and the two-way interaction terms between pollutants and weather factors, we found that many interaction terms were significant. For example, for the model including TSP, SO_2, and O_3 simultaneously, or for the model including both TSP and SO_2, the interaction terms TSP with season, TSP with mxtp, TSP with δ-pressure, SO_2 with season, SO_2 with mxtp, and SO_2 with δ-pressure are significant ($p<0.05$). This reinforces our finding that the effect of pollutants on mortality is dependent on season and weather.

DISCUSSION

By all accounts, determining the extent to which TSP is associated with daily mortality is a complex problem. In our

analysis of eighteen years of Philadelphia data, the results varied by season, the cause of death, the number of pollutants in the model, the model used to analyze the data, and whether or not interactions between pollution variables and weather were considered.

Even in the models with the greatest explanatory power, 18% of variations in daily mortality could be explained, with weather variables and seasonality being the strongest predictors. In analyses using weather alone, about 14% of variations in daily mortality could be explained, but in models using pollution alone, less than 5% could be explained. Also in a two-stage analysis where weather variables are forced into the model and pollution is left to explain the residual, pollution accounts for less than 1% of the variation.

Further complicating the issue is the fact that the pollution variables and the weather variables are all highly correlated with each other, making it difficult to separate out the effect of one pollution variable from another and the effect of weather from pollution. Thus it is difficult to say with any degree of certainty whether those significant associations between TSP and daily mortality observed in our analyses are real or surrogate behavior for some other pollutants or weather variables.

Second, the analysis by season showed no consistent significant association between TSP and mortality. For older people (i.e., 65+), this relationship was significant only in winter in the case where other pollutants were not considered in the equation. In contrast, for younger people (i.e., under 65), the relationship was significant only in the spring in the case where other pollutants were not considered. In both cases (i.e., for older and younger subjects), TSP was insignificant where more than one pollutant was taken into account.

Third, analyses broken down by causes of death indicate that if anything, TSP is negatively correlated with CVD (for the 65 and older group), the disease category with the greatest number of deaths in our analysis. In addition, TSP is negatively correlated with other causes of death, but the associations are never statistically significant. Caution should be followed in overinterpreting the death by cause results because cause of death data are ultimately based on death certificates -- information that can be notoriously misleading.

In summation, a comprehensive analysis of the Philadelphia data does not point to a clear association between daily mortality and particulate matter or to any other pollutant. For almost every result suggesting a positive association between particulate levels

and daily mortality, there are negative or nonsignificant results pointing the other way. It would therefore be incorrect to conclude from a single regression analysis of the Philadelphia data that "X" number of deaths can be assigned to each incremental increase of particulate matter (or any other type of pollutant).

Our analysis also raises some important issues regarding the appropriate methods for analyzing air pollution daily mortality data in the future. First, a determination should be made about the most desirable method for analyzing such data. Following the lead of others, we have used Poisson models, regression models, and autoregressive models, and arrived at results ranging from positive to negative associations. Each of the models that we used has its strengths and weaknesses. In the future, more work should be done to develop optimum models for analyzing mortality data and new approaches should be explored.

Second, a determination should be made about the number of days of pollution and weather measurements to be taken into account in the analysis. As was shown in Table 8, the significance of the predictive power of TSP on mortality is highly dependent on the number of pollution days in the model.

Our analyses clearly indicate that results are data- and model-dependent. Moreover, it is unclear which statistical model generates the most definitive results. Results vary by the model

choice, pollution variable choice, season, and cause. A simple conclusion on the association between daily mortality and air pollution based on a specified model might be biased.

ACKNOWLEDGEMENT

The authors thank the Electric Power Research Institute for their support of this study.

BIBLIOGRAPHY

Dockery, D.W.; Schwartz, J.; Spengler, J.D. *Air pollution and daily mortality: associations with particulates and acid aerosols*. Environ. Res. 56:362-373, 1992.

Fairley, D. *The relationship of daily mortality to suspended particulates in Santa Clara County, 1980-1986*. Environ. Health Perspect. 89:159-168, 1990.

Hatzakis, A.; Katsouyanni, K.; Kalandidi, A., et al. *Short term effects of air pollution on mortality in Athens*. Int. J. Epidemiol. 15:73-81, 1986.

Kalkstein, Laurence S. *A new approach to evaluate the impact of climate on human mortality*. Env. Health Persp. 96:145-150, 1991.

Kinney, Patrick L.; Ozkaynak, Haluk. *Associations of daily mortality and air pollution in Los Angeles County*. Env. Res. 54:99-120, 1991.

Ostro, Bart. *A search for a threshold in the relationship of air pollution to mortality: a reanalysis of data on London winters*. Env. Health Persp. 58:397-399, 1984.

Ozkaynak, Haluk; Thurston, George D. *Associations between 1980 U.S. mortality rates and alternative measures of airborne particle concentration*. Risk Analysis 7(4):449-460, 1987.

Pope, C.A.; Schwartz, J.; Ransom, M. *Daily mortality and PM_{10} pollution in Utah Valley*. Arch. Environ. Health 42:211-217, 1992.

Schwartz, J. *Air pollution and daily mortality in Birmingham, Alabama*. Am. J. Epid. 137(10), 1993.

Schwartz, J. *Particulate air pollution and daily mortality in Detroit*. Environ. Res. 56:204-213, 1991.

Schwartz, J.; Dockery, D.W. *Increased mortality in Philadelphia associated with daily air pollution concentrations*. Am. Rev. Respir. Dis. 145:600-604, 1992a.

Schwartz, J.; Dockery, D.W. *Particulate air pollution and daily mortality in Steubenville, Ohio*. Am. J. Epidemiol. 135:12-19, 1992b.

Schwartz, J.; Spix, C.; Wichmann, H.E., et al. *Air pollution and acute respiratory illness in five German communities*. Environ. Res. 56:1-14, 1991.

Shumway, R.H.; Azari, A.S.; Pawitan, Y. *Modeling mortality fluctuations in Los Angeles as functions of pollution and weather effects*. Env. Res. 45:224-241, 1988.

Wyzga, R.E. *The effect of air pollution upon mortality: a consideration of distributed lag models*. JASA 73(363):463-472, September 1978.

TABLE 1

DISTRIBUTION OF DAILY MORTALITY, POLLUTANTS*, AND WEATHER FACTORS (1973-1990)

	5%	10%	25%	50%	75%	90%	95%	MEAN	STANDARD DEVIATION
TSP (μg)	32	37	48	64	83	105	120	68.51	27.91
SO_2 (ppb)	4.10	5.89	8.93	13.91	21.80	31.83	39.46	16.98	11.87
O_3 (ppb)	1.82	3.54	8.33	17.08	28.43	39.36	46.83	19.77	14.35
MXTP °F	31	37	48	63	78	84	86	64.06	18.88
MNHUMID	43.75	48.88	56.88	66.00	74.17	80.38	83.29	67.01	13.41
DEWP °F	10.00	16.75	28.50	44.25	59.17	66.42	70.17	42.90	19.04
DAILY DEATH <65	12	13	16	19	22	26	28	19.06	5.03
DAILY DEATH ≥65	24	26	30	35	40	45	48	35.15	7.13

* The air quality data is the average of values across Philadelphia.

TABLE 2

DISTRIBUTION OF CAUSE-SPECIFIED DAILY MORTALITY COUNTS
IN PHILADELPHIA COUNTY, 1973-1990

CAUSE	MORTALITY, DEATHS/DAY*	5%	10%	25%	50%	75%	90%	95%	Mean	Std
Cancer	Age <65 years	2	3	4	5	7	9	10	5.66	2.42
	Age ≥65 years	3	4	6	5	10	12	13	7.76	2.91
CVD	Age <65 years	2	2	3	5	7	8	10	5.22	2.43
	Age ≥65 years	10	12	14	17	21	24	26	17.58	4.76
Pneumonia	Age <65 years	0	0	0	0	1	1	2	0.38	0.64
	Age ≥65 years	0	0	0	1	2	3	4	1.31	1.23
COPD	Age <65 years	0	0	0	0	0	1	1	0.23	0.48
	Age ≥65 years	0	0	0	0	1	2	3	0.75	0.96
Other	Age <65 years	3	4	5	7	9	12	13	7.58	3.03
	Age ≥65 years	3	4	6	8	10	12	13	7.75	3.04

TABLE 3

PEARSON CORRELATION COEFFICIENTS OF THE PREVIOUS DAY'S AIR FACTORS AND DAILY MORTALITY (AGE ≥65) (1973-1990)

	SO_2	O_3	TSP	MXTP	DEWP	MNHUMID
DEATH*	0.179+	-0.143+	0.051+	-0.283+	-0.263+	-0.005
SO_2		-0.260+	0.566+	-0.248+	-0.264+	-0.035+
O_3			0.179+	0.635+	0.488+	-0.161+
TSP				0.262+	0.211+	0.046+
MXTP					0.921+	0.201+
DEWP						0.535+

* Death: daily mortality for persons age ≥65.

\+ Statistically significant (p-value = <0.01)

TABLE 4

LIST OF MODEL TYPE

NAME	DESIGNATION	CITATION	FORMULA
Poisson-log	PL	Y_i is daily mortality.	$\mathrm{Log}(E(Y_i)) = X_i\beta$
Poisson-square root	PSR	Y_i is daily mortality	$\sqrt{E(y_i)} = X_i\beta$
Auto-regression	AR(p)	R_i could be log $(E(Y_i))$ or others; p is the minimum variable linear predictor of the error term given the available past values of $n_{i-1}, \ldots, n_{i-p}$.	$R_i = X_i\beta + n_i\|_{i-1}$ where $n_i\|_{i-1} = -\alpha_1 n_{i-1} - \ldots - \alpha_p n_{i-p}$
Regression of residual from trigonometric fit	RFT	Y_m is month-average of daily death; $W = 2\pi/360$; t is days in the middle of the month; Ψ, μ are parameters.	$E(Y_m) = \mu + R\,\mathrm{Cos}(Wt + \Psi)$ $E(Y_i - Y_m) = X_i\beta$
Regression of deviation from that day of the year	RDY	γ_i is the difference between Y_i and $\bar{Y}_i$, where $\bar{Y}_i$ is the average daily death on the same day of the year.	$E(\gamma_i) = X_i\beta$
Regression of deviation from a moving average	RDM(n)	γ_{mi} is the difference between Y_i and $\bar{Y}_{mi}$, where $\bar{Y}_{mi}$ is the average of the n days daily death with center at ith day.	$E(\gamma_{mi}) = X_i\beta$

TABLE 5

THE t-VALUES* IN THE SEASONAL ANALYSIS
MODEL: POISSON-LOG
FACTORS: YEAR, DAY OF SEASON, TWO DAYS' WEATHER

	AGE						
	>65				<65		
SEASON	TSP	SO_2	O_3	**	TSP	SO_2	O_3
Fall	0.99	0.49	1.29		-1.16	2.75	-0.15
Winter	0.05	2.55	0.92		0.65	0.45	0.00
Spring	0.60	0.84	2.73		0.76	-0.26	-0.38
Summer	0.48	-0.53	2.67		0.05	2.56	-0.34
Fall	0.93	0.45			-0.76	2.09	
Winter	-0.02	2.68			0.01	1.02	
Spring	0.51	0.58			0.77	-0.23	
Summer	1.59	-0.94			0.79	1.92	
Fall	1.19				0.68		
Winter	2.83				0.96		
Spring	0.98				2.04		
Summer	1.72				1.89		

* All t-values in this analysis could be considered subject to standard normal distribution because the degree of freedom is higher than thousands.

** Blank means the pollution is not considered in the model.

TABLE 6

THE t-VALUES OF POLLUTANTS IN THE ANALYSIS BY CAUSE
MODEL: POISSON-SQUARE ROOT
FACTORS: YEAR, DAY OF SEASON, TWO DAYS' WEATHER

	AGE						
	>65				<65		
CAUSE	TSP	SO_2	O_3	**	TSP	SO_2	O_3
Cancer	0.67	-0.73	0.20		-0.88	1.36	1.79
CVD	-1.74	1.73	3.97		-0.46	1.37	2.32
Pneumonia	0.88	-0.40	0.65		-1.19	1.44	-0.03
COPD	0.08	-0.64	0.06		0.14	-0.69	-1.62
Other	-0.07	1.82	1.45		-0.56	-0.18	1.21
Cancer	0.93	-1.06			-0.48	0.76	
CVD	-0.74	0.56			0.12	0.65	
Pneumonia	0.61	0.53			-1.18	1.41	
COPD	0.23	-0.66			-0.07	-0.63	
Other	-0.04	1.53			-0.60	0.05	
Cancer	0.39				0.14		
CVD	-0.24				1.36		
Pneumonia	1.12				-0.82		
COPD	-0.37				-0.65		
Other	0.89				-0.36		

* All t-values in this analysis could be considered subject to standard normal distribution because the degree of freedom is higher than thousands.

** Blank means the pollution is not considered in the model.

TABLE 7

t-VALUES OF POLLUTANT INDICATORS GENERATED
BY DIFFERENT AVERAGING TIME ANALYSES
MODEL: POISSON-LOG
AGE ≥65

MODEL	SIMULTANEOUS ANALYSIS			TSP ONLY
	TSP	SO_2	O_3	
MP1(2)*	1.31	3.49	3.81	4.24
MP2(2)	2.21	3.59	4.34	5.97
MP3(2)	0.38	1.98	3.07	3.0
MP4(2)	0.20	1.56	2.35	2.44
MP3(3)	1.02	1.72	3.96	3.45
MP4(4)	1.19	1.38	3.82	3.46

* MPx(y) stands for mean pollution of current and prior (x-1) days with y days' weather. For example, MP1(2) is prior day's pollution with 2 days' weather.

TABLE 8

t-VALUES OF POLLUTANTS IN DIFFERENT STATISTICAL MODELS: AGE ≥65

	POLLUTANTS IN THE MODEL					
Model	TSP	SO_2	O_3	TSP	SO_2	TSP
PL	1.2	3.4	3.8	1.9	2.4	4.0
PSR	1.3	3.5	3.7	2.0	2.6	4.3
AR(6)	2.4	3.8	5.8	3.6	2.7	6.6
RFT	0.8	3.0	3.2	1.5	1.9	3.1
RDY	1.4	1.6	3.1	2.0	0.8	3.1
RDM(7)*	0.1	1.6	1.7	0.3	1.0	1.0
RDM(15)	0.6	2.2	2.8	0.9	1.6	2.4
RDM(29)	1.1	2.1	3.3	1.7	1.1	3.2

* 7, 15, 29 = 7, 15, or 29-day moving average.

TABLE 9

t-VALUES OF POLLUTANTS IN DIFFERENT STATISTICAL MODELS: AGE <65

	POLLUTANTS IN THE MODEL					
Model	TSP	SO_2	O_3	TSP	SO_2	TSP
PL	-0.04	2.33	1.67	0.09	1.96	1.92
PSR	0.04	2.75	1.73	0.23	2.30	2.43
AR6	1.11	3.00	-1.06	-1.90	-0.94	-0.66
RFT	-0.77	2.45	1.52	-0.50	1.89	1.33
RDY	-0.15	1.38	1.40	0.02	0.96	1.32
RDM(7)*	-1.34	2.10	1.65	-1.55	1.97	-0.06
RDM(15)	-0.55	1.69	1.97	-0.61	1.50	0.73
RDM(29)	1.09	2.11	-1.73	-4.89	4.46	-2.25

* 7, 15, 29 = 7, 15, or 29-day moving average.

A Sensitivity Analysis of Mortality/PM_{10} Associations in Los Angeles

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

New York University Medical Center
Nelson Institute of Environmental Medicine
Long Meadow Road
Tuxedo, NY

Running title: Mortality/PM_{10} Sensitivity

Corresponding Author: Patrick L. Kinney, Sc.D.
Nelson Institute of Environmental Medicine
Long Meadow Road
Tuxedo, NY 10987

ABSTRACT

Several recent studies have reported statistically significant and quantitatively similar associations between daily mortality and PM_{10} or 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. The present paper 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-hr average PM_{10}, daily 1-hr maximum O_3 and carbon monoxide, maximum daily temperature, and mean daily relative humidity. Analyses reported here were limited to the subset of days on which PM_{10} data were available (every 6th 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 $\mu g/m^3$ 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 US.

INTRODUCTION

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 μm) or TSP (total suspended particulate mass concentration), in metropolitan areas (Schwartz, 1991; Schwartz and Dockery, 1992a, 1992b; Dockery, Schwartz, and Spengler, 1992; Pope, Schwartz, and Ransom, 1992; Schwartz, 1993). It is remarkable that consistent results have been obtained in diverse cities that vary in population, weather patterns, and levels of co-pollutants. 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).

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 $\mu g/m^3$ 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 volume (Thurston and Kinney, 1994).

There are four broad issues that arise in the analysis of multi-year 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 US. 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 over-control 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 (O_3) and carbon monoxide (CO) (Kinney and Ozkaynak, 1991; Hexter and 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 which 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 paper evaluates the sensitivity of mortality/PM_{10} relationships over a range of statistical approaches to the four issues introduced above. 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 which occurred in Los Angeles County in the period January 1, 1985 to December 31, 1990 were obtained from National Center for Health Statistics death certificate tapes. Deaths due to accidents and suicides and non-resident deaths were excluded from the total counts, yielding the daily death count variable used in the analyses reported here. PM_{10}, O_3, and CO data collected in Los Angeles county were obtained in digital form from the US. Environmental Protection Agency's Aerometric Information and Retrieval System (AIRS). 24-hr average PM_{10} concentrations, collected every 6 days, were taken from 4 monitoring sites. Daily maximum 1-hr O_3 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_3, 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 one-day lag of O_3 was used based on previous analysis of a 10-year record in Los Angeles (Kinney and 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 month, 6 months, 1 year, and 2 years; the second used 10 sine and cosine waves, with periods of 1 month, 2 months, 3 months, 6 months, 9 months, 1 year, 15 months, 18 months, 21 months, and 2 years); season-specific regressions, as above, 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. Multi-period sine and cosine waves should provide more complete control. One measure of the adequacy of cyclic controls is the Durbin-Watson (DW) statistic (Durbin and 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 multi-day (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 evaluate this "pre-filtering" method, because the available PM_{10} data were restricted to every 6th day. Each of the sensitivity analyses utilized a basic model that involved regressing ln(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(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 O_3 (lagged one day) and CO alone, and in combination with PM_{10}, as predictors of ln(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 (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 described above 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 $\mu g/m^3$ increase in PM_{10}. Note that 100 $\mu g/m^3$ represents the upper 95th percentile of PM_{10} concentrations in this data set. Relative risks for O_3 and CO likewise were computed at the upper 95th percentiles: 143 ppb for 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-month "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 to 1.05. The confidence intervals were similar across the three models, with the lower bounds just touching or spanning a relative risk of 1.00 (i.e., p value in the neighborhood of 0.05). Similar relative risks were observed for analyses restricted to the winter or summer seasons, although the confidence intervals increased due to the reduction in sample size. Season-specific regressions that also included 4

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, three yielded DW statistics above 1.8: the model that included 10 sine and cosine variables, and the two 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 1 and Figure 2), although the PM_{10} relative risk, and statistical significance, was reduced somewhat with the more extensive weather controls.

Regression models that included alternative pollutants (PM_{10}, lag 1 O_3, or CO) all yielded significant (or nearly so) relative risks for the individual pollutants (Table 1 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_3 in the model, the PM_{10} relative risk was essentially unchanged (RR=1.05), while the O_3 relative risk dropped to one. This suggests that the O_3 effect on mortality, if any, is weaker than that of PM_{10}. The correlation of the slope estimates for O_3 and PM_{10} was -0.5, indicating a substantial collinearity in these two pollutants when included simultaneously in the model. In contrast to the situation for O_3, the relative risks for <u>both</u> 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 (Table 1 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-year data set from Los Angeles county. We found that the estimated proportional increase in daily mortality associated with a 100 $\mu g/m^3$ 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 papers 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 multi-day 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 multi-day 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 LA 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 0.05 level (i.e., the 95% confidence intervals often included unity). Non-significance 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 ($p=0.04$). The data set analyzed in this paper (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 6th 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, Schwartz, and Ransom, 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 can not 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, Schwartz, and Spengler, 1992) and two have reported significant associations in Los Angeles and New York City (Kinney and 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 (1992) reported strong associations involving CO in an analysis of a 10-year 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 US. 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.

ACKNOWLEDGMENTS

The authors wish to thank Esther Christie for her contribution to the development of the air pollution data base. This work was supported by NIEHS Grant # ES05711.

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. *Biometrica* 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. Resp. 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: is the model the message. *Inhalation Tox.* (current issue).

Table 1. Descriptive Statistics, Overall and by Season

Period	Statistic	Deaths (per day)	PM_{10} ($\mu g/m^3$)	O_3 (ppb)	CO (ppm)	T (Deg. F)	RH (%)
All	Mean	153	58	70	4.7	70	70
	Std. Dev.	20	23	41	2.9	7	16
	Range	113-224	15-177	3-201	1-13	54-98	14-97
Nov-Feb	Mean	169	61	36	7.4	67	63
	Std. Dev.	20	29	20	2.7	7	20
	Range	132-224	15-177	3-100	2-13	54-88	14-97
Jun-Sep	Mean	143	60	101	2.8	73	77
	Std. Dev.	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	.99-1.09
	4 sine/cosine waves	1.04	.99-1.09
	10 sine/cosine waves	1.05	1.00-1.11
	winter only	1.03	.95-1.11
	summer only	1.06	.96-1.15
	winter + 4 sine/cosine	1.04	.96-1.12
	summer + 4 sine/cosine	1.10	1.00-1.21
Weather	no control	1.05	1.00-1.10
	temp. + RH	1.05	1.00-1.11
	extensive weather model	1.04	.99-1.08
Other Pollutants	PM_{10} only	1.05	1.00-1.11
	O_3 only	1.02[a]	1.00-1.05
	CO only	1.07[b]	1.01-1.13
	$PM_{10} + O_3$ / PM_{10}	1.05	1.00-1.11
	$PM_{10} + O_3$ / O_3	1.00[a]	.94-1.06
	PM_{10} + CO / PM_{10}	1.04	.98-1.09
	PM_{10} + CO / CO	1.05[b]	.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

[a] Relative risk for O_3.

[b] Relative risk for CO.

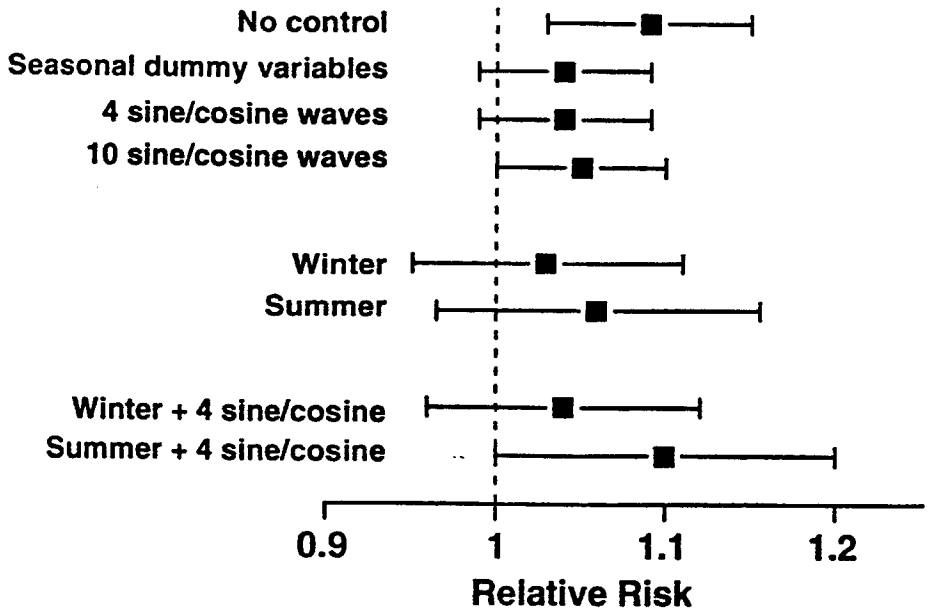

Figure 1. PM_{10} relative risks and 95% confidence intervals obtained with alternative methods to control for temporal cycles. All models included same-day temperature and relative humidity as covariates. Relative risks were computed for a 100 $\mu g/m^3$ increase in PM_{10} concentration.

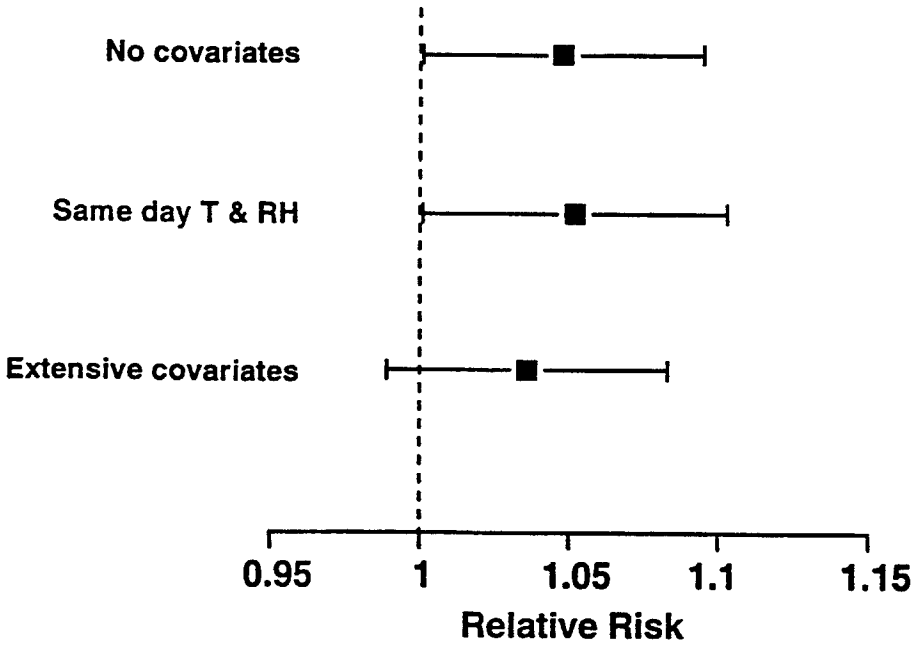

Figure 2. PM_{10} relative risks and 95% confidence intervals obtained with varying levels of weather controls. All models included 10 sine/cosine functions as covariates. Relative risks were computed for a 100 $\mu g/m^3$ increase in PM_{10} concentration.

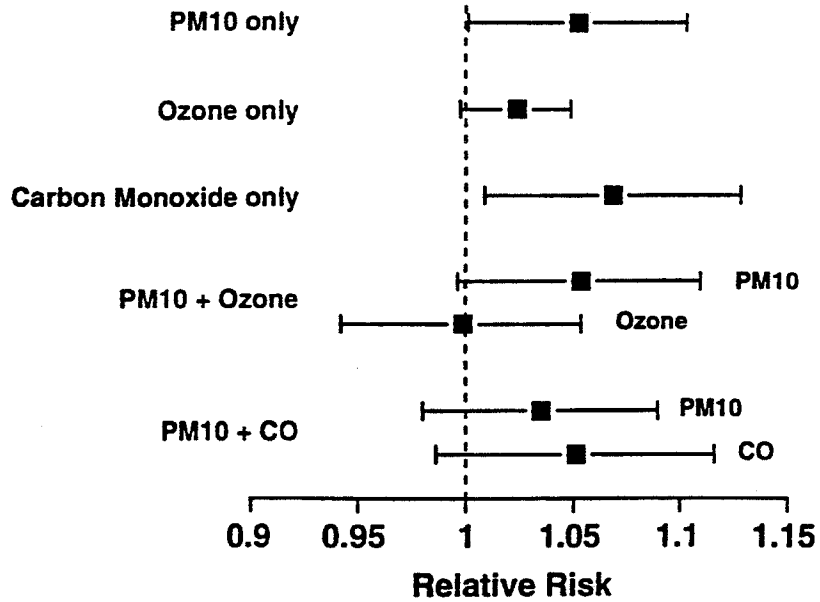

Figure 3. Relative risks and 95% confidence intervals for PM_{10}, O_3, and CO in univariate and bivariate regressions with mortality. All models included same-day temperature and relative humidity, and 10 sine/cosine functions as covariates. Relative risks were computed for a 100 $\mu g/m^3$ increase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

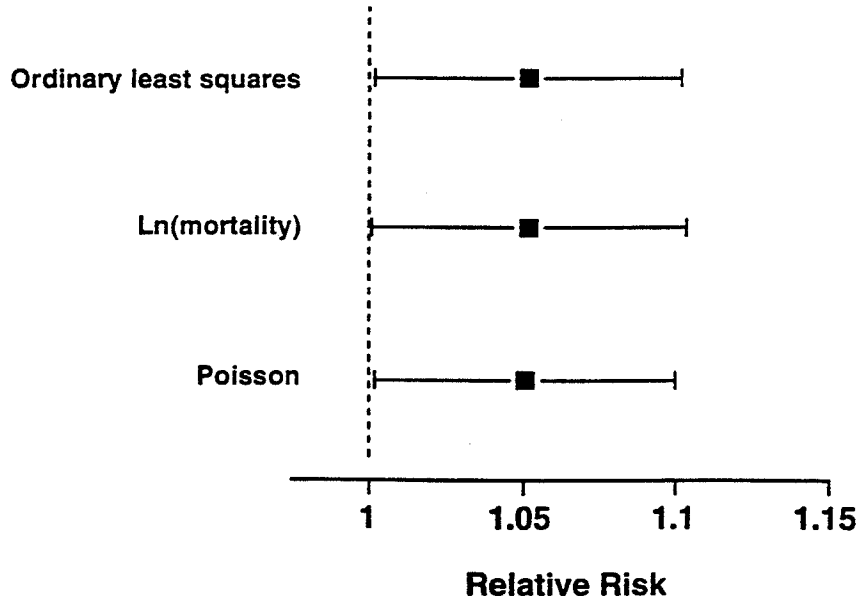

Figure 4. PM_{10} relative risks and 95% confidence intervals obtained with alternative regression models. All models included same-day temperature and relative humidity, and 10 sine/cosine functions as covariates. Relative risks were computed for a 100 $\mu g/m^3$ increase in PM_{10} concentration.

for submission to Inhalation Toxicology

Uncertainties in Identifying "Responsible" Pollutants in Observational Epidemiology Studies

Frederick W. Lipfert*
Environmental Consultant
23 Carll Court
Northport, NY 11768

Ronald E. Wyzga
Electric Power Research Institute
P.O. Box 10412
Palo, Alto, CA 94303

*person to whom correspondence should be addressed

Abstract

Studies of community air pollution must deal with a complex mixture of substances, for which the available data on concentrations and their distributions vary greatly in completeness and accuracy. The monitoring database available for some pollutants (such as suspended particulate matter) far exceeds that available for others (such as carbon monoxide or nitrogen dioxide) in terms of spatial and temporal coverage. Little or no routine monitoring data are available on aeroallergens or on particles classified by size and chemistry, for example. In addition, the relationships between outdoor air concentrations and personal exposures vary by chemical species. This paper addresses the concern that the availability and quality of observed data may limit the validity of the conclusions that can be derived from retrospective studies.

The basic assumptions of multiple regression analysis, the statistical tool most commonly used to study the effects of air quality on health, are reviewed. We show by data simulation and by numerical experiments with mortality and air quality data from Philadelphia that differences in the reliability of exposure estimates can be critical in the implied relationships between correlated variables in multiple (joint) regressions. Further, measurement error obscures the true degree of collinearity that may actually be present. Finally, we consider how nonlinear transformations can affect judgments about the relative importance of the variables considered. While models based on linear pollution relationships may be facile and may be convenient in characterizing effects, we have no assurance that they are in fact correct. Resolution of these issues will require better population-based air quality monitoring data as well as laboratory studies appropriate to characterizing the nature of the implied biological responses to the mixtures and concentrations that currently comprise community air quality.

This research was sponsored by the Electric Power Research Institute.

INTRODUCTION

Background. The observational or "ecological" epidemiological study is one study design that can be used to study irreversible endpoints such as mortality or hospitalization. At current air quality levels in the United States, even the severest air pollution health effects are considered "weak" from a statistical perspective (Wynder, 1987), because they generally involve individual risk ratios less than 1.5. Thus, strict adherence to statistical requirements is required in order to derive even valid qualitative conclusions. However, because large populations may be exposed to the risks associated with community air pollution, the findings of such observational studies may be important from the perspectives of public health and environmental policy.

Observational studies are intended to deduce relationships between population health responses and environmental characteristics. There are often two inter-related objectives: to identify associations between specific diseases and specific agents, and to define the concomitant dose-response functions and/or "safe" concentration levels. "Safe" levels may be used to establish ambient air quality standards, and dose-response functions are needed to estimate the economic or health benefits that might accrue from new policies to control emissions. In general, there is no longer interest in a study outcome that only shows that air pollution may be "harmful to health." Such harm has been unequivocably established by the air pollution episodes of past decades (Lipfert, 1994a) and by the many laboratory experiments that have followed. Current interest centers on the details of relationships between air pollution and biological responses; this requires separating the effects of collinear pollutants and determining the functional forms of dose-response functions, including whether no-effect thresholds are present. At the risk of pedantry, we first review some of the fundamental concepts involved in deducing

such relationships from experimental data. Much of the following material and notation follows that of Snedecor and Cochran (1967).

METHODOLOGICAL ISSUES

Statistical Methods. Statisticians have devised many ways to study relationships embodied in a data set. All of them involve comparing a calculated statistic with a range of values that would be expected to occur just to chance (random) variation. If the statistic falls outside of this range, it is said to be "significant" and we then look for explanations for this behavior. The limits of chance variation are established from the distributional properties associated with the statistic, such as those of the Poisson or Gaussian (normal) distributions. It follows that certain assumptions must be met for such comparisons to be valid. Sometimes one or more variables must be mathematically transformed in order to meet these requirements.

For large cities, the data describing community health responses to air pollution may be treated as quasi-continuous, rather than as categorical. The most common measures of association are correlation and regression. Tests of significance with respect to the **existence** of a relationship, using either (bivariate) regression or correlation, require only that the dependent variable be normally distributed. This test compares either the regression slope (b) or the correlation coefficient (r) with zero. When r is not zero and we wish to establish its confidence limits (for the purpose of model comparison, for example), both variables must be normally distributed.

However, the bivariate situation is rare in studies of practical interest, in part because the effects of interest are weak, and thus there are always intervening variables that could confound the outcome if not controlled. As a result, multiple regression is usually required, with

the model

$$Y = b_0 + b_1x_1 + b_2x_2 \ldots\ldots + b_nx_n + e \qquad [1]$$

where the x_i are error-free "independent" variables and e is an independent error term which is normally distributed with mean 0 and variance sigma^2 (note to typesetter: sigma = Greek sigma; e = epsilon). If these conditions are met, estimates of the b_i are unbiased and represent the effect of each variable with the others "controlled for" or held constant. In addition to the requirement of independence of the error term, the multiple regression model [1] requires that the Y values be normally distributed for each combination of the values of the x_i. If the true relationship between Y and some x_i is not linear, but say, exponential or logarithmic, then imposition of a linear relationship will lead to a correlation between e and x_i, which is a violation of the basic assumptions.

There are cases in which two pollutants are so highly correlated that separation of their effects by multiple regression is impractical. SO_2 and smoke in London constitute one such instance (r = 0.89); SO_2 and TSP are also highly correlated in Philadelphia (r=0.64). The common source for these pollutants in winter is probably space heating. Slopes may be calculated for one pollutant with the other held constant by defining specific subsets of the data, but in general these will not be the "best unbiased estimates" for the entire data set.

Development of Models. It is ineluctable that no one can claim a "true" model relating endpoints such as mortality or hospitalization to community air pollution. We "know" or strongly suspect on physiological grounds that certain other factors are important in a given community and must be controlled for, including age, sex, seasonal cycles, heat stress, infectious

epidemics. For comparisons between communities, we must add factors that describe lifestyle and basic health status differences, including smoking habits, education, diet, exercise, occupational exposure, selective migration, among others. However, there are also certain non-physiological factors that have been shown to contribute, even though biological explanations may be lacking; some of these are variations according to the day of the week, changes in meteorology, and even proximity to the decedent's birthday (Blakeslee, 1992) and cultural beliefs (Blakeslee, 1993). Differences in the "pace of life" have been implicated in differences in health status among communities (Levine, 1990).

In addition, if the effects of air pollution are to be considered on a physiological (rather than psychological) basis, we must estimate actual exposures to pollution. This exposure will vary with the daily activity patterns of the subjects and the nature of all the micro-environments to which they are exposed each day. The only data we have on air pollution across an entire community are the measurements routinely collected from fixed-base monitoring stations, which will have varying relationships to these microenvironments depending on the pollutant, time of day, and the age and health status of the subject, among other factors. The relationships between air monitoring data and the actual exposures of subjects constitute potential sources of bias and of errors in the independent variables. The effects of such errors will be addressed subsequently.

Because we cannot account for all factors that may deterministically affect health outcomes, our model is not in fact described by Eq. [1], but by

$$Y = b_0 + b_1x_1 + b_2x_2 \ldots\ldots + b_nx_n + b_kx_k + e' \qquad [2]$$

where the x_k represent factors that cannot be measured or for which suitable data are not avail-

able, and e' represents that portion of Y that is distributed independently of $x_1....x_n$ and x_k. By omitting the x_k which describe part of the "true" relationship, we find that we have not estimated the values of the $b_1....b_n$ that are needed, but instead the value of

$$b_1 + b_k b_{ik} ..... + b_n + b_k b_{nk} \qquad [3]$$

where b_{ik} is the regression coefficient of x_k on x_i. Relation [3] describes a situation that is often referred to as "confounding." The fact that the relationships of most interest here are usually weaker than the relationships among independent variables can make the potential errors involved in [3] important.* However, note that according to [3], for a variable to confound, it must be correlated with both health outcome (b_k) and pollution (b_{ik}). Since most pollutants are correlated with one another, and usually to a degree stronger than their correlations with health, **any (unmeasured) pollutant with a known or suspected effect on health qualifies as a potential confounder with respect to the attempt to estimate effects specific to any other pollutant.** Also, any criteria pollutant must be deemed *a priori* to have the potential for adverse health effects.

Lack of independence of the x_1 is the rule rather than the exception in observational studies. Health is affected by weather and by pollution; pollution affects weather and vice versa, and there are exogenous factors (such as day-of-week activity and emission patterns) that affect both health and pollution. Given this state of affairs, the errors involved in selecting an

---

For example, in Los Angeles (Kinney and Ozkaynak, 1991) after adjustment for seasonality, the maximum correlation between any of four air pollutants and daily mortality was 0.13 while the correlations among the pollutants were in the range 0.43 to 0.88.

---

inappropriate model can be much larger than the statistical confidence limits derived from a particular model (Lipfert, 1994b), especially for large data sets. The prudent investigator will thus compare results obtained from a range of models, variables, and transformations for each new application.

Study Designs and Collinearity. The most common study designs involve either following a fixed population over time (the time-series study) or contrasting many populations over space at a fixed time (the cross-sectional study). In both designs, collinearity among pollutants is to be expected and has been observed. With time-series studies, all pollutants tend to respond to the same weather perturbations, since emissions tend to be steady over time (secondary pollutants such as ozone may be an exception). With cross-sectional studies, commonality of pollutants emitted by the same types of sources creates collinearity. In both cases, the use of a limited set of fixed air monitoring stations can involve sufficient uncertainties so as to mask the true degree of exposure collinearity and thus to make it impossible to determine the "responsible" pollutants with certainty. In all cases, the relationships among pollutants (and between pollutants and weather variables for time-series studies) are stronger than the relationships between pollution and health, which can make control of collinear variables crucial.

Effects of Measurement Errors. Actual (personal) exposures to air pollution are impractical to measure directly in populations large enough to have adequate statistical power in observational studies. The difference between measured concentrations and actual exposures may be expected to vary from day to day and by pollutant species. Some pollutants are distributed more uniformly throughout a community than others, and some penetrate into indoor environments more readily than others. Thus, we expect that errors in estimating exposure will vary by pol-

lutant species. Actual exposure estimates require not only accurate data on local outdoor pollution levels, but also information on penetration into indoor spaces and on the effects of indoor sources of air pollution. These uncertainties add to those of the representativeness of the fixed monitor data *per se*.

The effect of error in the independent variable of a bivariate regression or correlation is to bias the regression or correlation coefficient and its significance toward the null (Snedecor and Cochran, 1967). It thus follows that measurement error serves to obscure the extent of the actual (structural) collinearity that may be present among pollutants in ambient air. The effect of collinearity among independent variables entered simultaneously into a multiple regression is to inflate the variance of the regression coefficients (Theil, 1978), with attendant reductions in statistical significance. In a multiple regression in which the various independent variables may have different magnitudes of measurement errors, the results may be less predictable (Fleiss and Shrout, 1977).

An often overlooked aspect of this question is the effect of differences in measurement reliability on the partitioning of response among collinear pollutants. The case of two competing pollutants with different measurement errors is discussed below in conjunction with results obtained with simulation techniques and from Philadelphia mortality data.

Pollutant Metrics and Functional Forms. The various components of the urban air pollution mix are often measured on different time scales; gaseous species and selected particulate measures (coefficient of haze and beta-gauge particle measures) are generally reported in the United States as hourly averages, and some of the ambient standards are posed in terms of maximum hourly averages. The more common particulate measures (TSP, PM_{10}, SO_4^{2-}, etc.) are reported as 24-hr or daily averages. There are substantial differences in diurnal profiles among species,

because of the combined effects of emission profiles and of meteorological effects. Differences in temporal profiles will be manifested as differences in spatial profiles as winds pass through an urban area. It is thus to be expected that errors in estimating personal exposure will vary with the averaging time of the pollutant, given that most people are outdoors for only a small part of the day. In addition, there will be less spatial variability as averaging times increase, and thus a fixed monitor can do a better job of representing exposures over a wide area if longer averaging times are used. As an example, comparisons of bivariate correlations of respiratory hospital admissions with four different pollutants are given in Figure 1, for peak hour and 24-hr average metrics (Lipfert and Hammerstrom, 1992). In all cases, the results for the 24-hr averages were higher, suggesting that this metric is more appropriate in a linear model.

The selection of the pollutant metric (i.e., averaging time) to be used to identify responsible pollutants in an observational study should not be confused with deriving information about the actual effects of each pollutant, which may indeed be nonlinear. In such cases, for which only peak exposures may be important, nonlinear dose-response models should be investigated to find the optimum distributions of model residuals. Because the regression fits for alternative models are often essentially identical, exogenous information may be required to select the most appropriate models.

Studies of mortality in Philadelphia (Li and Roth, 1994) show that the mortality shares assigned to competing pollutants may also be influenced by the selection of lag times for each pollutant. In general, there are no *a priori* guidelines for the appropriate lag structures, except that the outcome must lag the environmental insult in a causal model. We may expect to find different lags for different pollutants with respect to different health endpoints.

Interpretation of Regression Coefficients. If in model [1], we find a statistically significant relationship between Y and a specific x_j (which implies that the regression assumptions have been observed), we may conclude either:

1. the indicated value of b_j will accurately reflect the change in Y due to an incremental change in x_j (alone), or
2. The indicated value of b_j will reflect the change in Y due to an incremental change in x_j and the change in Y due to all other variables with which x_j is correlated.

If the relationship of interest fails the test of significance, we may conclude either:

1. there is no relationship between the two variables, or
2. the model selected was inappropriate to the true relationship between the two variables (which generally results in violating the regression assumptions), or
3. measurement errors obscured the true relationship.

A further important corollary is that the level of statistical significance should not be regarded as a measure of importance of the effect it presumably portrays; measurement error and collinearity can also affect significance levels. A better measure of importance is the relative magnitude of the incremental effect on Y.

Summary of Regression Model Requirements. In summary, we find the following requirements for model [1] to produce reliable (unbiased) estimates of the relationship between health outcomes and pollution:

1. Confounding variables must be accounted for. If the model does not include all intercorrelated pollutants known to affect health, the regression coefficient for any one pollutant must be regarded as also including the effects of the omitted components of the pollutant mixture. This caution also applies to non-pollutant variables which may be correlated with pollution, such as ambient temperature.
2. For the pollutants for which measurements are available, comparisons among species must be made in such a way that the errors in exposure estimates are approximately equal. This requires a common basis for averaging time, for example. If data from a network of monitors are averaged, roughly the same number of stations should be used on all days and for all species.
3. The model must be appropriate to the physiological scenario. If peak exposures are thought to be more important than average exposures, nonlinear pollutant transformations should be investigated. If it is reasonable to expect a lag between exposure and response, a range of different lag structures should be investigated, including effects accumulated over the lag period and the possibilities of different lag structures for different pollutants, diseases, and age groups. If chronic effects are suspected, allowance must be made for the latency period required for development of the disease in question.
4. Depending on the outcomes of the preceding steps, an iterative procedure may be required to identify the most important pollutants and to estimate each of their separate contributions to the health outcome with an optimum model.
5. The regression model assumptions must be observed. This may include transforming the dependent variable to achieve normality and checking the regression residuals for heteroscedasticity.

USE OF SIMULATION TECHNIQUES TO EXPLORE THE SIMULTANEOUS EFFECTS OF MEASUREMENT ERROR AND COLLINEARITY IN TIME SERIES ANALYSIS

This section explores the difficulties involved in interpreting regression results for two collinear pollutants having differing measurement errors, making use of rudimentary simulation techniques. The general problem of errors in measurement has a long history in statistics, but most of the attention has been given to the bivariate case and to special conditions involving correlated errors; see Cochran (1972) or Chai (1971), for example. Simulation can thus be useful in this application, since theoretical solutions are lacking.

Quasi-random variables were generated using the random number function of QuattroPro* and near-normal distributions were derived by adding independent sets of random values. A time-series of 101 days was simulated and 20 trials were performed in order to generate a range of results that could be plotted individually for detailed examination. This type of rudimentary simulation should be considered a demonstration pilot for a full-scale Monte Carlo simulation.

The Simulation Model. We postulate a causal mortality model in which daily deaths are proportional to concentrations of two pollutants:

$$D = B_0 + B_1P_1 + B_2P_2 + e_D \qquad [4]$$

(note to typesetter: B denotes Greek beta, e denotes epsilon.)

---

*registered trade mark of Borland International, Inc.

We solve this relationship for the regression coefficients and their standard errors using ordinary least-squares methods. In the simulations performed, the *a priori* assumed values of B_1 and B_2 are unity. We assume that the pollutants P_1 and P_2 emanate from a common source or group of sources and thus are highly correlated in ambient air, even though they may not be measured at the same times and places. However, the relationships between monitored concentrations (the observed values in this case) and actual population exposure differ; additional (random) variance is present in the distribution of the values of P_2. We designate the unobservable value of this pollutant as P_2 (true) and the actual measurements as P_2 (measured). Thus P_2 meas. = P_2 true + measurement error (e_m) and our operational regression model is

$$D = B_0 + B_1P_1 + B_2'(P_2 + e_m) + e_D \qquad [5]$$

where B_2' is the regression coefficient as determined in the presence of measurement error. In this sense, the measurement error term e_m may be thought of as the differential sum (P_2 vs. P_1) of instrument and analytical error, spatial representativeness of the monitor network, effects of missing data, differences in lag structure of response, and differences in personal exposure relative to monitored concentrations. The magnitudes (but not the existence) of most of these effects are unknown.

Simulations were performed for 7 different combinations (referred to hereafter as "cases") of relative measurement error and assumed collinearity between P_1 and the true value of P_2, using 20 trials each. Each trial produces statistics for 5 different regressions, as discussed below. Crossplots are presented across regressions, with each trial constituting an "observation." Twenty trials were deemed sufficient to define average regression coefficients and *t* values, but are undoubtedly insufficient to completely define their distributions. The first of these simula-

tions ("baseline") is described in detail, including plots of the results. Summary statistics are then given for the remaining 6 simulation cases.

Simulation Results for the Baseline Case. The means and standard deviations for the baseline simulation are given in Table 1. The additional exposure measurement error for P_2 corresponds to an increase in variance of 56%; no bias was added to P_2 by this procedure. The correlations between P_1 and P_2 (true) and P_2 (measured) were 0.96 and 0.75, respectively. The reduced correlation between P_1 and the "true" P_2 could be interpreted as reflecting measurement error in P_1 or perhaps daily variations in pollutant mix. Since air pollution is often distributed lognormally, one might assume that the logarithms of concentrations have been simulated rather than the actual values; this assumption entails no loss in generality. Subsequent Monte Carlo simulations using lognormal distributions produced essentially the same results.

For each simulation trial, five regressions were performed with respect to simulated mortality, as follows. Regressions 1, 2, 3: each of the three pollutant variables (P_1, P_2 (true), and P_2 (measured)) was regressed separately against the simulated mortality variable. Regression 4: joint regression on P_1 and the "true" value of P_2. Regression 5: joint regression on P_1 and the measured value of P_2.

The results for the 3 separate (bivariate) regressions are plotted in Figure 2, in terms of the regression coefficient t values. As expected, we see that P_1 and P_2 are essentially equivalent when the "true" P_2 values are used, but that use of the simulated P_2 measurements results in lower t values for P_2. Note that the parameters of this simulation were chosen to yield highly significant regression coefficients for bivariate regressions.

The average regression statistics for the 20 trials are given in Table 2 for the five different regression formats. Note that only Regression No. 4 provides the "correct" coefficients,

which were defined as 1.0 in the formulation of the model; however, because of the collinearity, these coefficients are not significant and thus would likely be disregarded if they occurred with real data (in the absence of a priori information on the "true" model). In addition, the regression coefficients derived in the bivariate case (Regressions No. 1 and 2, one pollutant at a time) are inflated to essentially twice the expected value, since the effect of both pollutants is picked up by the single pollutant entered into the model. With measurement error, the coefficient for P_2(meas) is biased low with respect to that for P_2(true), as expected.

We see from Table 2 that the regression coefficient for P_1 is essentially unchanged from the bivariate case when regressed jointly with P_2 in the presence of measurement error (Regression No. 5), while the joint coefficient for P_2 has become nil, even though the basic underlying relationship assigns equal weight to both P_1 and P_2. Thus the "standard" multiple regression procedure yields an inflated coefficient for the pollutant with the lower measurement error and essentially zero for the pollutant with higher measurement error. These results confirm Cochran's earlier (1970) prediction, which was also based on numerical examples, that "interpretation of the B_i' as if they were B_1 can be quite misleading...." (typesetter: set Greek beta for B).

Figure 3 compares the *t* values for P_1 for the two joint regressions. When regressed jointly with the "true" P_2, P_1 loses significance, on average, because of collinearity. However, when P_1 is regressed jointly with P_2 in the presence of increased measurement error, P_1 loses less significance and its regression coefficients are virtually unchanged from the bivariate case (Figure 4). Note that all but one of the P_1 regression coefficients in Figure 4 exceeds the *a priori* or "true" value (1.0).

The corresponding *t*-value information for P_2 is given in Figure 5. In both cases, P_2 lost significance in most of the 20 trials. A plot of the joint regression coefficients for P_1 versus P_2

showed a negative relationship (Figure 6); a negative coefficient for one variable contributes to a more strongly positive value for the other.

Sensitivity Simulations. Additional simulation cases were performed in the same way to explore the robustness of these findings. These results (cases 2-6) are given in Table 3 along with the baseline result from Table 2 (case 1), in terms of the coefficients from Regression No. 5, the joint regression of mortality on P_1 and P_2(meas). There were little or no changes in the bivariate regression coefficients for these simulations.

Note that the regression coefficients tend to sum to the correct value of 2.0 in all 7 cases, but that the proper split (1:1) is approached only in simulation case 5. In this case, the introduced measurement error is small and the regression on P_1 and P_2(meas) is essentially the same as on P_1 and P_2(true). Note also that some of the simulation cases feature very modest amounts of pollutant collinearity and that biased coefficients still result from the joint regressions. Additional sensitivity runs (not shown) established that modest levels (relative to the error in P_2) of measurement error in P_1 had no effect on these findings.

NUMERICAL EXPERIMENTS WITH 1973-80 PHILADELPHIA MORTALITY DATA

Data on mortality, weather, and air pollution (TSP, SO_2, and ozone) as measured in Philadelphia (Li and Roth, 1994) were used to further explore the ramifications of measurement errors in stepwise multiple regressions, but based on real instead of simulated data. The hypothesis here is that taking indoor/outdoor exposure relationships into account will add to the existing variance of the pollution terms, but in unknown ways. We created modified but non-

negative SO_2 and TSP variables by adding arbitrary amounts to the variance of their logarithms (using a normally distributed random variable with mean=0), and then transforming back into linear coordinates. The regression t-values were used to describe the resulting changes to the model, as above.

We used the 65-and-over age group for these sensitivity studies and examined the first regression step at which a pollutant was selected to enter the model. For this base case, with no additional error, SO_2 would normally enter first since its t-value is slightly higher than that of TSP (7.95 vs. 7.65). After SO_2 enters, the significance of TSP is sharply reduced because of the collinearity, but TSP enters on a subsequent step after other seasonal and weather variables. When the variance of SO_2 is artificially inflated, the situation quickly reverses, and a slight increase in SO_2 error can reverse the order of variable entry. When the variance of TSP is inflated instead, reversal of the order of entry does not take place (because SO_2 was preferred initially), and the rate of decrease of t for TSP as its variance was inflated was more gradual, perhaps because there was less daily variation in TSP to begin with. The behavior of models for daily mortality and for its logarithm were quite similar; the other variables included in the model were a linear time variable, daily mean temperature and dew point, the daily change in barometric pressure, and 24-hr average ozone concentrations. The contributions of ozone were not substantially affected by inflating the variance of TSP or SO_2. Similar results were obtained for the under-65 age group, except that the baseline case t values were lower, and as a result, only the initial entry pollutant remained in the model at the final step.

These data are seen to support the findings of the simulation study, in that they illustrate the important role that measurement error can play in partitioning the effects of correlated "independent" variables in joint regressions.

WHAT IS THE CORRECT SHAPE OF THE DOSE-RESPONSE FUNCTION?

Most models employ linear specifications for the pollution terms; use of the logarithm of response in a log-linear model is tantamount to assuming an exponential form for the pollutant dose-response function. Table 4 presents bivariate correlations between six different pollutants and summer respiratory hospital admissions in Southern Ontario (Wyzga and Lipfert, 1993) for different functional forms. These results are for the entire year, adjusted for day of week and seasonality and lagged 3 days. The table shows that selection of the most "important" pollutant cannot be made independently of knowledge of the shape of the dose-response function. The differences in linear models are small, but we see that coefficient of haze (COH), SO_2, and SO_4^{2-} show slightly higher associations when concave-downward functions are used (square root, log), while ozone and to a lesser extent TSP seem to point towards concave upward functions (square, exponential). The latter are more in keeping with the notion of pollution thresholds or log-linear models. NO_2 is seen to be indeterminant in this regard.

Table 5 compares beta coefficients (because they are nondimensional) for the logarithms of daily mortality in Philadelphia (1973-80), for the same functional forms. Data are shown for the 3 pollutants available in this study, taken jointly and separately, for 2 age groups. In general, functional form makes more difference in joint regressions than in separate regressions. We see some commonality with Table 4, in that SO_2 performs better with square root and log models (this was also the case over a much larger range of SO_2 values in the former East Germany [Spix et al., 1993]), and TSP performs better with the exponential and square functions. Ozone is seen to be indeterminate in joint regressions, but to indicate concave-upward functions in separate regressions.

We conclude from these data that it is necessary to know the appropriate form of a

dose-response function before the "best" pollutant can be selected with confidence.

CONCLUDING DISCUSSION

This paper has reviewed the basic assumptions of multiple regression analysis, which are undoubtedly familiar to most readers and certainly to all practitioners. However, it is also common to overlook some of these concepts in the search to derive meaning from data. Certainly, peak-hour concentrations have been used with linear model specifications, nonlinear responses are rarely examined, and the effects of measurement error have been largely ignored.

We show by simulation and by numerical experiments that differences in the reliability of exposure estimates can be critical in the relationships between correlated variables in multiple (joint) regressions and thus in the identification of the "best" variables. Thus, epidemiological studies of community air pollution should recognize these inherent limitations in formulating their conclusions. Assignment of the total effect to the most statistically significant pollutant under such conditions appears to be fraught with the potential for error. Joint regressions of correlated variables will yield unbiased estimates of the "true" effects only when their relative levels of measurement (or exposure) error are similar. Further, even if the relative measurement errors are similar, if collinearity is severe, neither variable will appear to be significant even though their true effects are real. However, unless it can be shown (exogenously) that the effects of the omitted pollutant(s) are negligible, regressing each pollutant variable individually under these conditions will almost certainly overstate its effect. A research or pollution control program based on such findings would also appear to have a high potential for misdirection.

Finally, nonlinear transformations and the choice of pollutant metric can affect judgments about the relative importance of competing pollutants. While models with linear pollution

effects may be facile and appeal to the need for convenient ways to characterize effects, we have no assurance that they are in fact correct.

None of these statistical and practical difficulties should be interpreted as diminishing the potential importance of the subject matter. The number of diverse studies that show associations between community air pollution and hospitalization or mortality is impressive (Lipfert, 1993, 1994). The existence of such effects seems highly probable; what remain uncertain are detailed descriptions and physiological mechanisms, as well as verification by long-term studies of pollution abatement. The one conclusion upon which agreement seems certain is that better estimates of actual exposures are essential and that population-based air monitoring data should be given a higher priority.

ACKNOWLEDGMENT

This research was supported by the Electric Power Research Institute under RP-3253. Figure 1 was reproduced with permission of Academic Press, Inc. Figures 2-6 and portions of the accompanying text first appeared in Air Pollution and Community Health and are used here with permission of Van Nostrand Reinhold, Inc. We thank 3 anonymous reviewers for helpful suggestions.

REFERENCES

Blakeslee, S., Birthdays: a Matter of Life and Death, New York Times, Sept. 22, 1992.

Blakeslee, S. (1993), Beliefs Reported to Shorten Life, New York Times, Nov. 6, 1993.

Chai, J.J. (1971), Correlated Measurement Errors and the Least Squares Estimators of the Regression Coefficient, J. Am. Stat. Ass. 66:478-83.

Cochran, W.G. (1970), Some Effects of Errors of Measurement on Multiple Correlation, J. Am. Stat. Ass. 65:22-34.

Cochran, W.G. (1972), Some Effects of Errors of Measurement on Linear Regression, in Proc. 6th Berkeley Symp. on Mathematical Statistics and Probability, ed. by J. Neyman, U. California Press, Berkeley, CA.

Fleiss, J.L., and Shrout, P.E. (1977), The Effects of Measurement Errors on Some Multivariate Procedures, Am.J.Public Health 67:1188-1191.

Kinney, P.L., and H. Ozkaynak, Associations of Daily Mortality and Air Pollution in Los Angeles County, Env.Res. 54:99-120 (1991).

Levine, R.V. (1990), The Pace of Life, Amer. Scientist 78:450-9.

Li, Y. and Roth, H.D. (1994), Daily mortality analysis by using different regression models in Philadelphia County, presented at the Colloquium on Particulate Air Pollution and Human Mortality and Morbidity, Irvine, CA, January 24-5, 1994.

Lipfert, F.W. (1993), A Critical Review of Studies of the Association Between Demands for

Hospital Services and Air Pollution, Environmental Health Perspectives Suppl. 101 (Suppl 2):229-268.

Lipfert, F.W. (1994a), Air Pollution and Community Health, Van Nostrand Reinhold, New York.

Lipfert, F.W. (1994b), Filter artifacts associated with particulate measurements: Recent evidence and effects on statistical relationship, submitted to Atm. Envir.

Lipfert, F.W., and Hammerstrom, T. (1992) Temporal patterns in air pollution and hospital admissions, Environmental Research 59:374-399.

Snedecor, G.W., and W.G. Cochran, Statistical Methods, 6th ed., The Iowa State University Press, Ames, Iowa. pp 164-166 (1967).

Spix, C., Heinrich, J., Dockery, D., Schwartz, J., Volksch, G., Schwinkowski, K., Coolen, C., and Wichmann, H.E. (1993), Air Pollution and Daily Mortality in Erfurt, East Germany, 1980-1989, Envir. Health Perspect. 101:518-27.

Theil, H. (1978), Introduction to Econometrics, Prentice-Hall, Englewood Cliffs, NJ. p. 137.

Wynder, E.L. (1987), Workshop on Guidelines to the Epidemiology of Weak Associations, Introduction. Preventive Medicine 16:139-141. See also Wynder, E.L. (1990), Epidemiological Issues in Weak Associations, Int.J.Epidemiology 19 (Suppl 1):S5-S7.

Wyzga, R.E., and Lipfert, F.W. 1993. Uncertainties in identifying responsible pollutants in environmental epidemiology. Presented at the 9th Congress on Aerosols in Medicine, Garmisch-Partenkirchen, Germany, April 1993.

Table 1 - Means and Standard Deviations (Baseline)

Variable	Mean	Std. Dev.
D	38.6	4.7
P_1	6.1	0.93
P_2 (true)	7.6	0.94
P_2 (meas)	7.6	1.18

Table 2 Baseline Regression Statistics

	-------- separate --------			-------- joint --------	
Regress.	# 1	2	3	4	5
Statistic	D vs P_1	D vs P_2true	D vs P_2meas	D vs $P_{1,2}$true	D vs $P_{1,2}$meas
B_1	1.88			1.06	1.88
sigmaB_1	0.45			2.1	0.78
t	4.47			0.6	2.94
B_2		1.81	1.07	0.85	-0.01
sigmaB_2		0.46	0.27	2.1	0.5
t		4.46	3.21	0.48	0.0
R^2	0.17	0.17	0.098	0.17	0.185
Std error of est.	4.18	4.18	4.37	4.18	4.17

note to typesetter: set Greek sigma for "sigma"

Table 3 Summary of Simulation Results for the Joint Regression of Mortality on P_1 and P_2meas.

Case #	Pollutant Correlations		Average Regression Coefficients	
	P_1,P_2true	P_1,P_2meas	P_1	P_2meas
1.	0.96	0.75	1.88	-0.01
2.	0.84	0.75	1.71	0.33
3.	0.90	0.84	1.75	0.36
4.	0.96	0.91	2.04	-0.08
5.	0.72	0.65	1.08	0.87
6.	0.66	0.29	1.76	0.18
7.	0.50	0.27	1.52	0.56

Table 4 - Bivariate Correlations of Respiratory Hospital Admissions in Southern Ontario, 1979-85, Adjusted for Season and day of Week, with Various Air Pollutants and Functional Forms of the Pollutant Dose-Response Function

	Function				
Pollutant	linear	exponential	square	square root	natural log
NO_2	0.038	0.024	0.038	0.037	0.033
COH	0.034	0.032	0.030	0.036	0.037
$SO_4^=$	0.034	0.027	0.029	0.035	0.033
O_3	0.029	0.053	0.042	0.020	0.008
SO_2	0.026	0.011	0.017	0.029	0.032
TSP	0.021	0.018	0.020	0.018	0.012

Table 5 - Beta Regression Coefficients for Philadelphia (Logarithm) Mortality and Air Pollution in Joint and Separate Regressions for Various Functional Forms (n=2459).

A. Deaths at Ages < 65

		LINEAR	EXP	SQUARE	SQ ROOT	LOG
joint regr	TSP	0.0375	*0.0509*	*0.0525*	0.0311	0.0359
	SO_2	*0.049*	0.0417	0.0389	*0.0551*	*0.0522*
	O_3	0.022	0.0182	0.012	0.0317	0.022
sep regr	TSP	0.068	*0.075*	*0.074*	0.065	0.062
	SO_2	*0.071*	0.071	0.066	*0.072*	*0.07*
	O_3	0.021	0.021	0.021	0.019	0.005

B. Deaths at Ages 65+

		LINEAR	EXP	SQUARE	SQ ROOT	LOG
joint regr	TSP	0.0736	*0.0832*	*0.0967*	0.0629	0.0566
	SO_2	*0.0835*	0.0731	0.0493	*0.0991*	*0.104*
	O_3	0.0688	0.0684	0.0619	0.0704	0.0795
sep regr	TSP	0.133	*0.134*	*0.134*	0.128	0.0121
	SO_2	*0.14*	*0.135*	0.117	*0.145*	*0.14*
	O_3	0.077	0.082	0.086	0.053	0.05

Covariates include date, dew point, hot days, mean temperature, and change in barometric pressure.

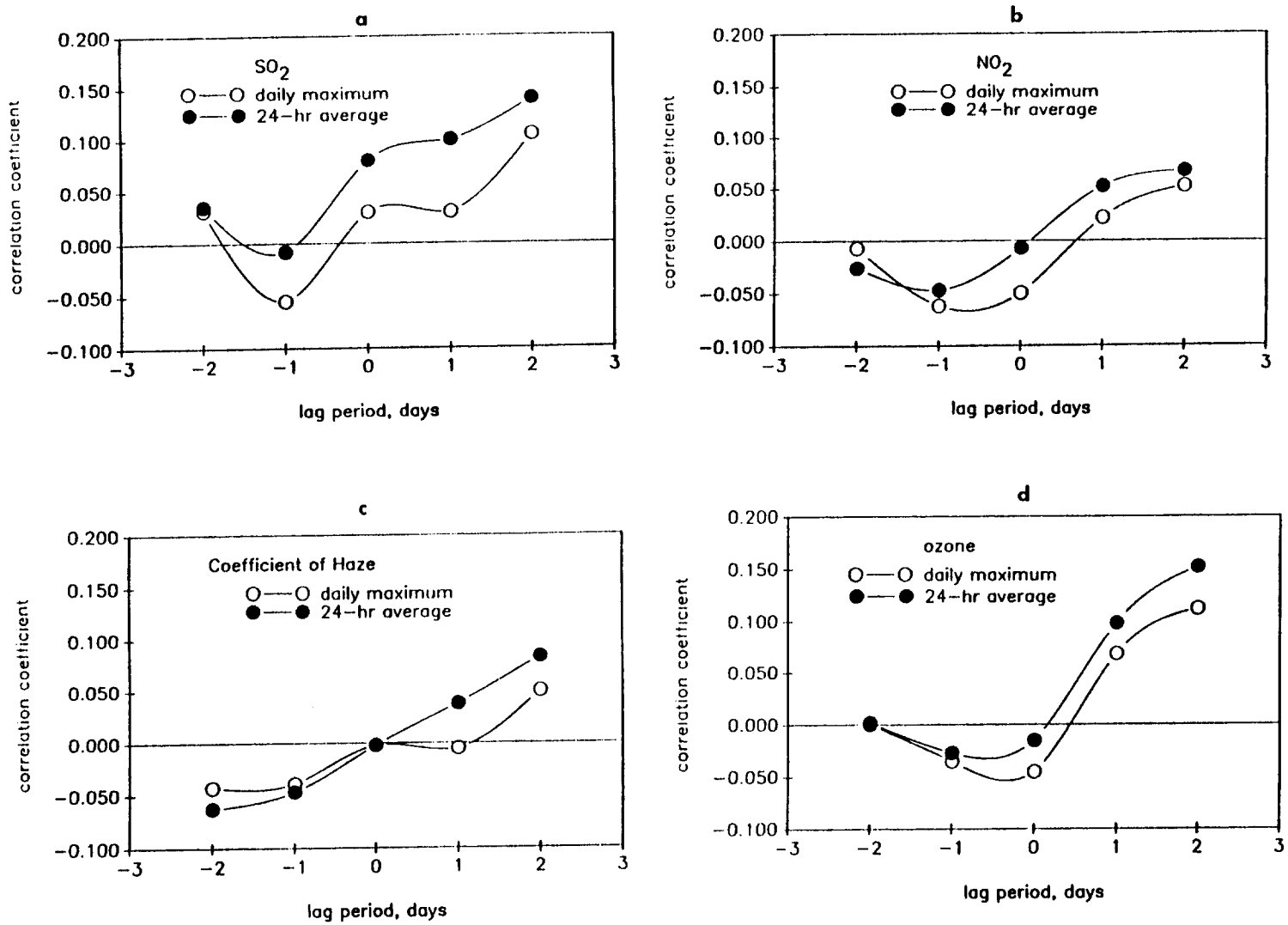

Figure 1. Bivariate correlations between July-August respiratory admissions in Southern Ontario (1979-85) and various air pollutants, comparing 24-hr averages with peak-hour data. (a) SO_2. (b) NO_2. (c) COH. (d) ozone. Confidence limits are for the nominal sample size of 360. Source: Lipfert and Hammerstrom (1992)

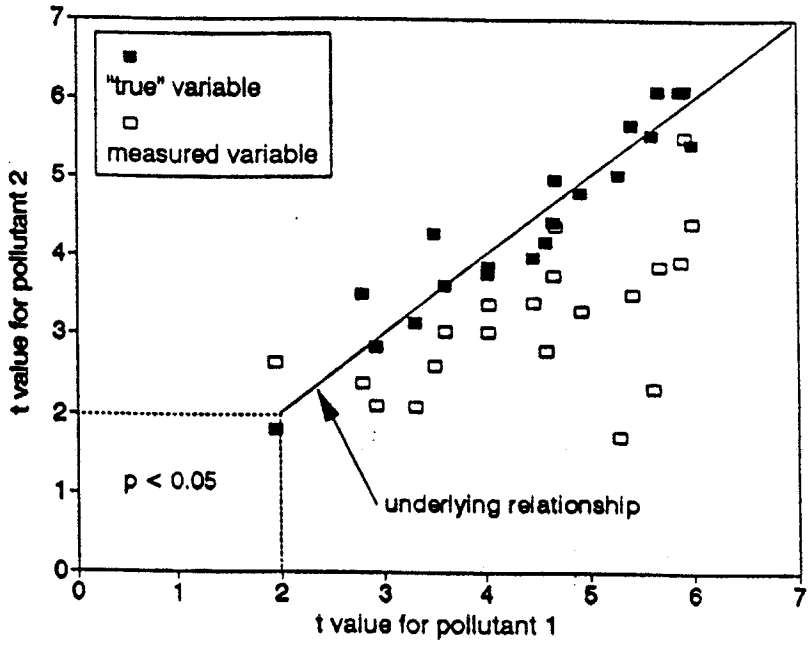

Figure 2. Relationships among the coefficient t values for the 20 trials of the baseline simulation, bivariate regressions. T values from regressions 2 and 3 are plotted against the t value for regression 1, for each trial. Source: Lipfert, 1994a.

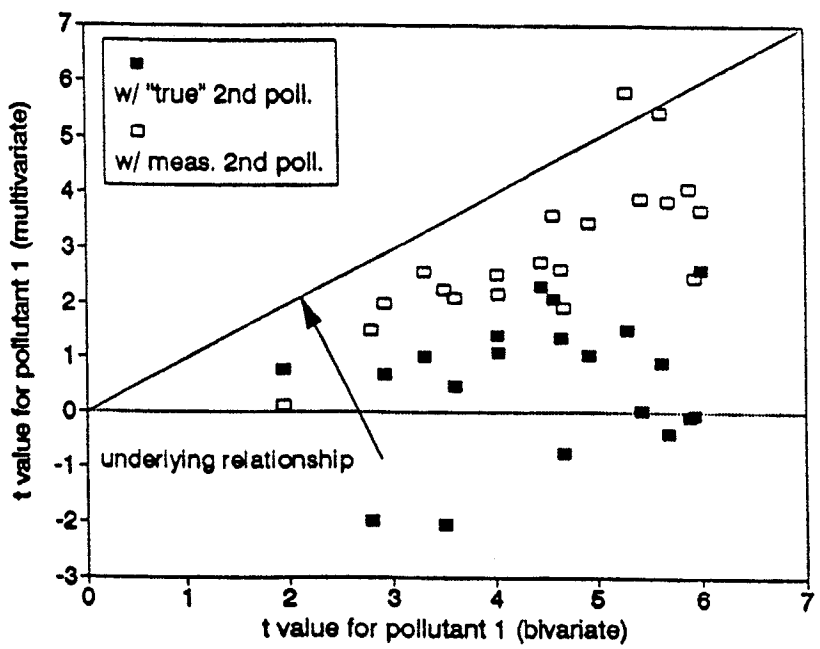

Figure 3. Relationships among the P_1 t values for the 20 trials of the baseline simulation. T values from regressions 4 and 5 are plotted against the t value for regression 1, for each trial Source: Lipfert, 1994a.

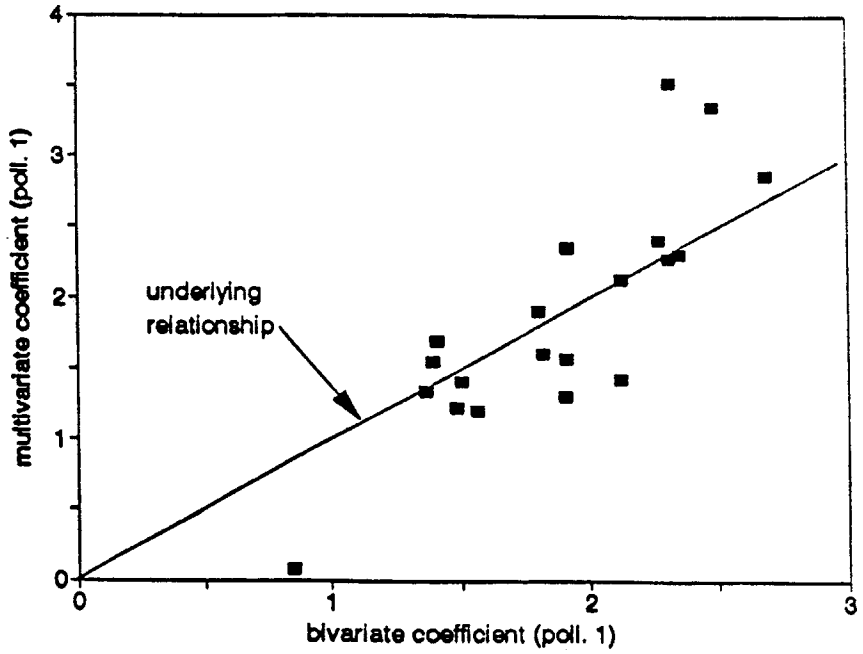

Figure 4. Relationships among the regression coefficients for P_1 for the baseline simulation. Coefficients from regression 5 (joint with P_2(meas) are plotted against the (bivariate) coefficient value for regression 1, for each trial. Source: Lipfert, 1994a.

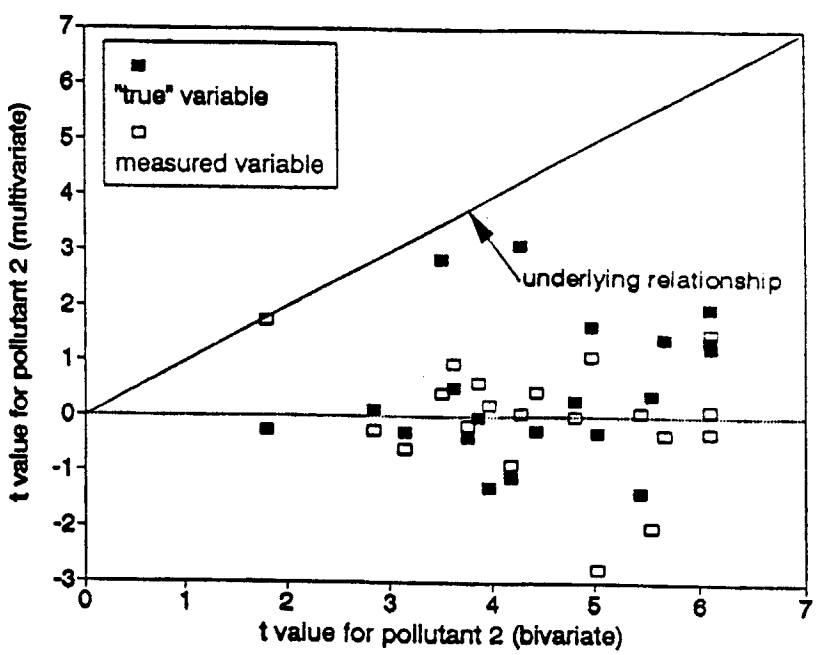

Figure 5. Relationships among the P_2 t values for the 20 trials of the baseline simulation. T values from regressions 4 and 5 are plotted against the t value for regression 2, for each trial. Source: Lipfert, 1994a.

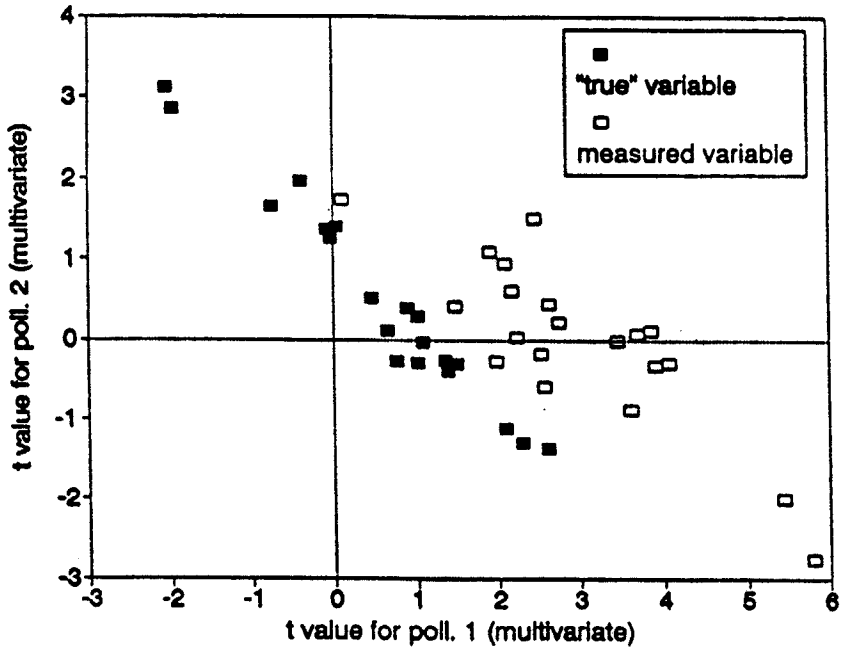

Figure 6. Relationships among the coefficient t values for the 20 trials of the baseline simulation. T values for P_1 from regressions 4 and 5 are plotted against the t value for for P_2, for each trial. Source: Lipfert, 1994a.