RESEARCH REPORT

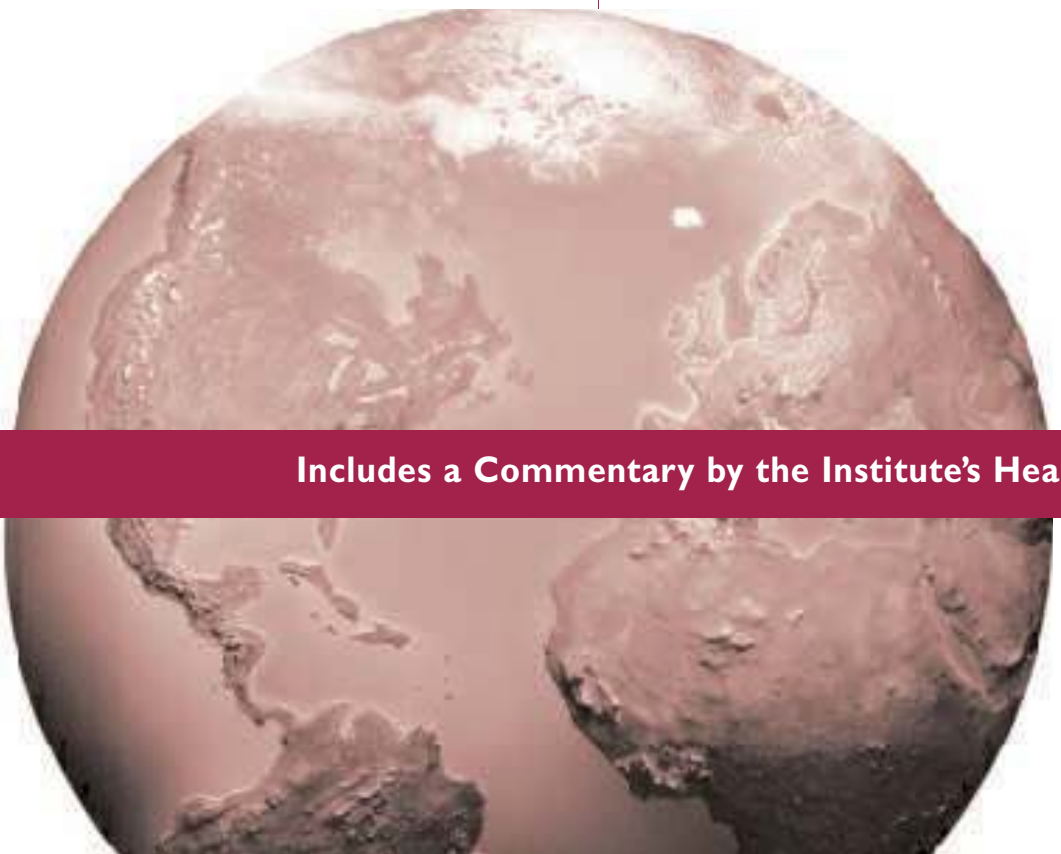
HEALTH EFFECTS INSTITUTE

Number 94, Part II
June 2000

Final Version

The National Morbidity, Mortality, and Air Pollution Study
Part II: Morbidity and Mortality from Air Pollution in the United States

Jonathan M Samet, Scott L Zeger, Francesca Dominici, Frank Curriero, Ivan Coursac, Douglas W Dockery, Joel Schwartz, and Antonella Zanobetti



Includes a Commentary by the Institute's Health Review Committee

HEALTH
EFFECTS
INSTITUTE

The Health Effects Institute, established in 1980, is an independent and unbiased source of information on the health effects of motor vehicle emissions. HEI supports research on all major pollutants, including regulated pollutants (such as carbon monoxide, ozone, nitrogen dioxide, and particulate matter) and unregulated pollutants (such as diesel engine exhaust, methanol, and aldehydes). To date, HEI has supported more than 200 projects at institutions in North America and Europe and has published over 100 research reports.

Typically, HEI receives half its funds from the US Environmental Protection Agency and half from 28 manufacturers and marketers of motor vehicles and engines in the United States. Occasionally, funds from other public and private organizations either support special projects or provide resources for a portion of an HEI study. Regardless of funding sources, HEI exercises complete autonomy in setting its research priorities and in reaching its conclusions. An independent Board of Directors governs HEI. The Institute's Health Research and Health Review Committees serve complementary scientific purposes and draw distinguished scientists as members. The results of HEI-funded studies are made available as Research Reports, which contain both the Investigators' Report and the Review Committee's evaluation of the work's scientific quality and regulatory relevance.

HEI STATEMENT

Synopsis of Research Report 94, Part II

The National Morbidity, Mortality, and Air Pollution Study: Morbidity and Mortality from Air Pollution in the United States

BACKGROUND

Epidemiologic time-series studies conducted in a number of cities have identified, in general, an association between daily changes in concentration of ambient particulate matter (PM) and daily number of deaths (mortality). Increased hospitalization (a measure of morbidity) among the elderly for specific causes has also been associated with PM. These studies have raised concerns about public health effects of particulate air pollution and have contributed to regulatory decisions in the United States. However, scientists have pointed out uncertainties that raise questions about the interpretation of these studies.

One limitation to previous time-series studies of PM and adverse health effects is that the evidence for an association is derived from studies conducted in single locations using diverse analytic methods. Statistical procedures have been used to combine the results of these single location studies in order to produce a summary estimate of the health effects of PM. Difficulties with this approach include the process by which cities were selected to be studied, the different analytic methods applied to each single study, and the variety of methods used to measure or account for variables included in the analysis. These individual studies were also not able to account for the effects of gaseous air pollutants in a systematic manner.

APPROACH

HEI funded the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) to characterize the effects of airborne particles less than 10 μm in aerodynamic diameter (PM_{10}) alone and in combination with gaseous air pollutants in a consistent way, in a large number of cities. The study was designed to select multiple locations based on the specific criteria of population size and availability of PM_{10} data from the US Environmental Protection Agency's Aerometric Information Retrieval System (AIRS) database, and to apply the same statistical procedures to all locations. Dr Jonathan Samet and his colleagues at Johns Hopkins University conducted a time-series study of mortality effects in large US cities representing various levels of PM_{10} and gaseous pollutants. In their analysis, the investigators first estimated risk in each city using the same method and then combined these results systematically to draw more information than any single city could provide. The 20 and 90 largest cities were analyzed for effects of PM_{10} and other pollutants on mortality; the 90 largest cities were analyzed for possible modification of PM_{10} effects among cities by factors other than air pollutants. Dr Samet's coinvestigators at Harvard University also applied a unified statistical method, although different from the one used in the mortality analysis, to 14 cities with daily PM_{10} data to examine effects on hospitalization among those 65 years of age or older.

RESULTS AND IMPLICATIONS

NMMAPS has made a substantial contribution in addressing major limitations of previous studies. The mortality analysis used one analytic approach to examine the PM_{10} effect in many cities that cover a wide geographic area and have varying levels of different air pollutants. The results of both the 20 cities and 90 cities analyses are generally consistent with an average approximate 0.5% increase in overall mortality for every 10 $\mu g/m^3$ increase in PM_{10} measured the day before death. This effect was slightly greater for deaths due to heart and lung disease than for total deaths. Effects of PM_{10} measured on the day of death or 2 days before did not vary substantially from one another for total or for heart and lung deaths. The

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI since 1996 and conducted by Dr Jonathan M Samet of Johns Hopkins University School of Public Health, Baltimore MD. The following Research Report contains a preface, the detailed Investigators' Report, Part II, and a Commentary on Part II prepared by a panel of the Institute's Health Review Committee.

PM_{10} effect on mortality also did not appear to be affected by other pollutants in the model.

Although individual estimates for each of the 90 cities varied, as expected, the strength of the analysis was in its ability to combine data from nearby cities in a particular region to estimate a PM_{10} effect. Combining the data in this systematic way provided additional statistical power to the analysis that is not available in single-city analyses. Some differences in PM_{10} effect on mortality were seen by region of the US: for the 90 cities, the largest effect was evident in the Northeast. The investigators did not identify any factor or factors that might explain these differences. This analysis is an important first step, and further evaluation of the reasons for these regional differences will advance our understanding of the association between PM_{10} and mortality. The heterogeneity of effect across cities offers the potential to identify factors that could influence the effects of PM_{10} on health and thus provide valuable insights into the mechanisms by which PM_{10} causes adverse health effects.

The morbidity analysis also used a unified analytic method to examine the association of PM_{10} with hospitalization of those 65 years of age or older in 14 cities with daily PM_{10} measurements. The results were consistent with an approximate 1% increase in admissions for cardiovascular disease and about a 2% increase in admissions for pneumonia and chronic obstructive pulmonary disease for each 10 $\mu g/m^3$ increase in PM_{10}. A greater estimate of effect on hospitalizations at lower concentrations (less than 50 $\mu g/m^3$) was found for the three diagnoses considered, but the meaning of these findings should await completion of concentration-response analyses for mortality now under way using data from 20 cities.

NMMAPS has made substantial contributions to our understanding of the relationship between exposure to PM_{10} and health effects. Further analyses in these databases of regional differences, the effects on morbidity and mortality combined, and concentration-response relationships will enhance our understanding.

Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.
The paper in this publication meets the minimum standard requirements of the ANSI Standard Z39.48-1984 (Permanence of Paper) effective with Report Number 21, December 1988, and with Report Numbers 25, 26, 32, 51, 65 Parts IV, VIII, and IX, and 91 excepted. These excepted Reports are printed on acid-free coated paper.

CONTENTS

Research Report 94

HEALTH EFFECTS INSTITUTE

The National Morbidity, Mortality, and Air Pollution Study Part II: Morbidity and Mortality from Air Pollution in the United States

Jonathan M Samet, Scott L Zeger, Francesca Dominici, Frank Curriero, Ivan Coursac, Douglas W Dockery, Joel Schwartz, and Antonella Zanobetti

Departments of Epidemiology and Biostatistics, Johns Hopkins University School of Public Health, Baltimore, Maryland, and Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts

HEI STATEMENT

This Statement, prepared by the Health Effects Institute, is a nontechnical summary of the Investigators' Report and the Health Review Committee's Commentary.

PREFACE

INVESTIGATORS' REPORT

When an HEI-funded study is completed, the investigators submit a final report. This Investigators' Report is first examined by three outside technical reviewers and a biostatistician. The report and the reviewers' comments are then evaluated by members of the HEI Health Review Committee, who had no role in selecting or managing the project. During the review process, the investigators have an opportunity to exchange comments with the Review Committee and, if necessary, revise the report.

Continued

COMMENTARY Health Review Committee

The Commentary about the Investigators' Report is prepared by the HEI Health Review Committee and staff. Its purpose is to place the study into a broader scientific context, to point out its strengths and limitations, and to discuss the remaining uncertainties and the implicatons of the findings for public health.

RELATED HEI PUBLICATIONS: PARTICULATE MATTER

Publishing History: This document was posted as a preprint on *www.healtheffects.org* and then finalized with revisions for printing.

Citation for whole report:

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DM, Schwartz J, Zanobetti A. 2000. The National Morbidity, Mortality and Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Health Effects Institute, Cambridge MA.

When citing a section of this report, refer to it as a chapter of this document.

PREFACE

ORIGINS AND OBJECTIVES

In 1996, HEI initiated the National Morbidity, Mortality, and Air Pollution Study (NMMAPS)*, based on the realization that a national study could address one of the major questions regarding air pollution and daily mortality: whether particulate air pollution is responsible for the associations between air pollution and daily mortality that have been observed in multiple studies, or whether these associations are due, in part or completely, to other air pollutants. This realization emerged both from the experience of the Particle Epidemiology Evaluation Project (PEEP), funded by HEI from 1994 to 1997, and from an evaluation of the literature at that point, which largely included studies of single cities.

PEEP was designed to (1) address the replicability and validity of key epidemiologic studies of particulate air pollution and daily mortality by conducting detailed reanalyses of selected data sets, and (2) explore in more extensive data sets some of the larger scientific and public health issues raised by the findings of these earlier epidemiologic studies. PEEP investigators, led by Drs Jonathan Samet and Scott Zeger, successfully replicated the numerical results of the earlier studies, including the previously reported associations between total suspended particles (TSP) and daily mortality in Philadelphia (Schwartz and Dockery 1992). More detailed analysis of the Philadelphia data led Samet and Zeger to conclude, however, that the associations with air pollution in that city could not be attributed to particulate air pollution alone. In its Commentary on the contributions and limitations of PEEP (Samet et al 1997), the Oversight Committee concluded: "Although individual air pollutants (TSP, SO_2, and ozone) are associated with increased daily mortality [in Philadelphia], the limitations of the … data make it impossible to establish that particulate air pollution alone is responsible for the widely observed associations between increased mortality and air pollution in that city. All we can conclude is that it appears to play a role. … Ultimately, it will require joint analyses of data from multiple cities with different copollutant correlations … to address further the role of multiple pollutants."

NMMAPS was also designed to address two additional issues that complicated interpretation of the results of daily mortality and air pollution studies considered in PEEP: the effect of measurement error in exposure estimates on relative risk estimates, and whether any effect of life-shortening (mortality displacement) associated with increased daily mortality can be removed from estimates of risk associated with air pollution. With regard to exposure measurement error, the Oversight Committee stated in its Commentary on PEEP that "Errors in exposure measurements as a result of using data provided by centrally located monitors rather than exposures or doses measured in individuals, could, in the context of complex multivariable models for daily mortality, affect the relative risk estimates in ways that are difficult to predict. The possibility of such errors are an important source of uncertainty about the true magnitude of the estimated effects of individual air pollutants on daily mortality." For this reason, the Oversight Committee recommended "… developing models to assess exposure measurement errors in daily time-series analyses, and applying those models to a national data set using more detailed exposure data, if available."

The extent of life-span reduction associated with pollution-related daily mortality in Philadelphia and other locales remained unclear. If such reductions were small, due mainly to the advancement of the date of death for frail individuals by a matter of days (mortality displacement), then the public health implications would be less profound. The Oversight Committee remarked that "Estimating the extent of life-shortening caused by short-term elevations in air pollution remains one of the most important tasks for future studies." Developing methods for addressing the questions of whether any excess daily mortality is associated with air pollution only, or of whether any association largely reflects short-term mortality displacement, became an important methodologic objective of NMMAPS.

To address these questions, NMMAPS had the following two broad objectives:

- To conduct a nationwide study of acute health effects of air pollution on morbidity and mortality. NMMAPS is based on data from the US national air monitoring network provided by the US Environmental Protection Agency's (EPA's) Aerometric Information Retrieval System (AIRS) database, which contains information on particulate matter less that 10 μg in aerodynamic diameter (PM_{10}) and other criteria pollutants from 1987 to 1994, as well as from information on health and the population from the National Centers for Health Statistics, the Health Care Financing Administration, and the US Census. NMMAPS evaluates two issues: (1) air pollution and daily mortality in the 20 and

* A list of abbreviations and other terms appears after the Investigators' Report.

90 largest US cities, and; (2) daily hospital admissions of the elderly (≥ 65 years old) in 14 US cities with daily measurements of PM_{10}. A combined analysis using daily mortality *and* hospital admissions in the same cities is planned.

- To develop the statistical and epidemiologic methods required for data analysis and interpretation of results from such an investigation. NMMAPS investigators have developed methods for combining the evidence across multiple locations and for assessing the impact of exposure misclassification on the estimated association between daily mortality and air pollution. They have also developed approaches that begin to answer the question of whether or not the excess daily mortality that has been associated with air pollution reflects only, or largely, small reductions in survival among frail individuals.

NMMAPS focuses on the acute health effects of particulate air pollution, measured as PM_{10}. Its design, however, was intended by the investigators also to provide a framework for the study of pollutants other than particles.

STUDY PARTICIPANTS AND CONDUCT

NMMAPS has been conducted by a team of investigators from the Johns Hopkins School of Public Health, led by Principal Investigator Jonathan Samet and including Drs Scott Zeger and Francesca Dominici. As discussed above, Samet and Zeger had conducted PEEP, from which NMMAPS developed. The Johns Hopkins investigators were responsible for the design and analysis of the mortality component of NMMAPS. They have worked in collaboration with Drs Douglas Dockery and Joel Schwartz of the Harvard School of Public Health on methods for addressing mortality displacement and measurement error. Dockery and Schwartz designed and conducted the morbidity analyses.

NMMAPS has been overseen by the same Oversight Committee that worked on PEEP, on HEI's behalf. This committee, chaired by Dr Gerald van Belle of the University of Washington, comprises leading experts in epidemiology, biostatistics, pulmonary medicine, and aerometric measurement. The Oversight Committee was responsible for working with the investigators to develop, and ultimately to approve the analytic plan that has guided NMMAPS from its inception.

As the analytic plan for NMMAPS was being developed, HEI sought the comments of a broad range of scientists and technical experts from industry, government, and public interest groups. To provide continuing updates on the progress of the study to these diverse groups, HEI has organized regular presentations of interim results at its Annual Conference (1997 to 1999), a symposium at the International Society for Environmental Epidemiology (September 1999), and briefings for HEI sponsors (July 1997, February 1998, and December 1998). Besides providing interested parties with up-to-date information on the progress of NMMAPS, these events provided HEI, the Oversight Committee, and the investigators with valuable comments and suggestions for their work.

REVIEW OF INVESTIGATORS' REPORT

All HEI reports are reviewed by the HEI Health Review Committee and external reviewers with relevant expertise as required by the subject matter of the report. NMMAPS was reviewed by a Panel that included members of the HEI Health Review Committee as well as several other individuals with expertise relevant to the methods and analyses in this report. The Panel also wrote the Commentaries for Part I and Part II of the NMMAPS report with input from the full HEI Health Review Committee, members of the NMMAPS Oversight Committee, and the HEI Research Committee.

STRUCTURE OF INVESTIGATORS' REPORT

The results of NMMAPS are presented as two reports. *Part I: Methods and Methodologic Issues* comprises a collection of methodologic papers on three topics: (1) measurement error in air pollution exposure, (2) mortality displacement, and (3) methods for combining the evidence in multiple locations using Bayesian hierarchical models. *Part II: Morbidity, Mortality, and Air Pollution in the United States* presents the results of analyses of daily mortality in the 20 and 90 largest US cities and in hospital admissions of the elderly (those 65 years old or older) in 14 US cities.

ACKNOWLEDGMENTS

HEI thanks the investigators and many other individuals whose contributions enhanced the quality of the NMMAPS project and this Research Report. Oversight of this complex project and evaluation of the findings would not have been possible without the help of the NMMAPS Oversight Committee, which included members of the HEI Research Committee, and the NMMAPS Review Panel, which included members of the HEI Review Committee. In particular, the Institute thanks Dr Aaron Cohen for his role

in assisting the Oversight Committee with management of this project, and Dr Diane Mundt for her role in assisting the Review Panel in developing the Commentary. The Review Panel gratefully acknowledges the cooperation of the investigators during the review process. Finally, the Institute appreciates the efforts of HEI's editorial and publications staff in preparing this Research Report.

INVESTIGATORS

Johns Hopkins School of Public Health

Jonathan M Samet (Principal Investigator)

Scott L Zeger

Francesca Dominici

Frank Curriero

Ivan Coursac

Harvard School of Public Health

Douglas W Dockery

Joel Schwartz

Antonella Zanobetti

HEI OVERSIGHT COMMITTEE

Gerald van Belle, Chair
University of Washington
HEI Research Committee

Glen R Cass
Georgia Institute of Technology
HEI Research Committee

Leon Gordis
Johns Hopkins University

Klea Katsouyanni
University of Athens

Thomas A Louis
University of Minnesota

Carl Shy
University of North Carolina

Duncan Thomas
University of Southern California

Mark J Utell
University of Rochester

HEI REVIEW COMMITTEE

Review Panel for NMMAPS

Ralph D'Agostino, Panel Chair
Boston University
HEI Review Committee

Ben Armstrong
London School of Hygiene and Tropical Medicine

Edo Pellizzari
Research Triangle Institute
HEI Review Committee

Mark D Schluchter
Case Western Reserve University

Daniel Tosteson
Harvard Medical School
HEI Review Committee

Sverre Vedal
University of British Columbia
HEI Review Committee

Mary White
Agency for Toxic Substances and Disease Registry

HEI PROJECT STAFF

Scientific Staff

Aaron Cohen
Senior Scientist (Research Project Manager)

Diane J Mundt
Staff Scientist (Review Project Manager)

Jane Warren
Director of Science

Publications Staff

John Abbott
Desktop Publishing Specialist

Julia F Campeti
Scientific Copy Editor

Sally Edwards
Managing Editor

Hope Steele
Consulting Editor

Quentin Sullivan
Consulting Proofreader

REFERENCES

Samet JM, Zeger DL, Kelsall JE, Xu J. 1997. Air pollution, weather, and mortality in Philadelphia 1973–1988. In: Particulate Air Pollution and Daily Mortality: Analyses of the Effects of Weather and Multiple Air Pollutants, The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

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

INVESTIGATORS' REPORT

The National Morbidity, Mortality, and Air Pollution Study Part II: Morbidity and Mortality from Air Pollution in the United States

Jonathan M Samet, Scott L Zeger, Francesca Dominici, Frank Curriero, Ivan Coursac, Douglas W Dockery, Joel Schwartz, and Antonella Zanobetti

OVERVIEW

PROJECT OBJECTIVES

The National Morbidity, Mortality, and Air Pollution Study (NMMAPS)* comprises analyses of air pollution and mortality and morbidity set in a national sampling frame created from the Aerometric Information Retrieval System (AIRS), the monitoring database of the US Environmental Protection Agency. The project is a collaboration between investigators at Johns Hopkins University School of Public Health (Drs Samet, Zeger, and Dominici) and Harvard School of Public Health (Drs Dockery, Schwartz, and Zanobetti). The project's overall objectives lie in the complementary domains of methods development and methods application.

This Report, Part II of NMMAPS, presents the findings on air pollution and morbidity and mortality in detail. For daily mortality, we have analyzed data for the 20 and 90 largest cities in the United States. (Throughout this Report, we refer to the selected study communities as *cities* although, because of the data structure, the counties making up the cities were the actual units of analysis.) Using a hierarchical modeling approach, we have assessed the mortality risks associated with particulate matter (PM) less than 10 μm in aerodynamic diameter (PM_{10}) and the other combustion-related criteria pollutants: ozone (O_3), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and carbon monoxide (CO). In the 20-city analysis, we provide detailed findings on the full set of pollutants, combining evidence across the cities to estimate the effects of the individual pollutants while controlling to the extent possible for the effects of the other pollutants. These analyses are then extended to the 90-city database, which we use to explore heterogeneity of effects across broad geographic regions and the determinants of the heterogeneity. The morbidity analysis, which uses hospitalization data from the Health Care Financing Administration (HCFA) for Medicare enrollees, addresses the association of hospitalization risk with PM_{10} and other pollutants in 14 cities. Hierarchical models are used to summarize the effects of air pollution on hospitalization risk. In a separate NMMAPS report, we will provide the findings of a planned joint analysis of morbidity and mortality.

Part I of the NMMAPS Report, *Methods and Methodologic Issues* (Samet et al 2000b), provides comprehensive descriptions of the methods used in the present report to summarize evidence on air pollution and mortality across multiple locations. It also presents a systematic analysis of the problem of measurement error in time-series studies of air pollution and proposes an approach to correcting for the consequences of measurement error, using data from studies with measurements for PM from both personal monitors and central sites. The report also describes 2 conceptually similar analytic approaches to evaluating the extent of associations found in daily time series on short-term mortality displacement (also termed *harvesting*). These methodologic topics, measurement error and mortality displacement, were selected for development because both were proposed as potentially severe limitations to interpretation of the time-series studies.

The objectives for developing specific methodologic components for NMMAPS are fivefold.

1. To develop semiautomated or automated approaches for database construction using databases of the EPA,

*A list of abbreviations and other terms appears at the end of the Investigators' Report.

This Investigators' Report is Part II of Health Effects Institute Research Report 94, which also includes a Commentary by the Health Review Committee and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr Jonathan M Samet, Department of Epidemiology, School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Ste W 6041, Baltimore MD 21205-2179.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

the National Center for Health Statistics (NCHS), the Health Care Financing Administration (HCFA), the Census Bureau, and the National Weather Service;

2. To develop and apply statistical methods for regression analyses of the multisite data and to develop spatial time-series methods to estimate spatial maps of the relative rates of mortality associated with air pollution, while accounting, as necessary, for the spatial and temporal correlations in the mortality data;
3. To develop and apply methods that adjust for smooth trends and seasonality on mortality caused by changing demographics and health behaviors, influenza epidemics, and other unidentified factors;
4. To examine the consequences of measurement error in the exposure variables for assessing pollutant-mortality associations; and
5. To examine the degree to which pollution-related mortality reduces years of life (mortality displacement).

The objectives for application of methods developed for NMMAPS are threefold.

1. To assess the relation between air pollution and mortality in the largest US cities monitored for PM_{10} from 1987 forward;
2. To assess the relation between air pollution and morbidity in selected US cities monitored for PM_{10} from 1987 forward; and
3. To conduct paired analyses of morbidity and mortality in the same locations.

The design for NMMAPS builds on prior work supported by the Health Effects Institute in the Particle Epidemiology Evaluation Project (PEEP) (Samet et al 1995, 1997). This project was initiated in 1994 with the objectives of validating the data and replicating the findings in several of the time-series studies of air pollution and mortality reported during the 1990s. In a second phase, PEEP addressed several methodologic issues. These included selecting the approach for controlling for potential confounding by weather (Samet et al 1998) and determining the sensitivity of findings to model-building strategies (Kelsall et al 1997; Samet et al 1997).

The present project, NMMAPS, evolved from PEEP. The objectives encompassed methodologic issues that were persistent sources of uncertainty in interpreting the epidemiologic evidence: mortality displacement and exposure measurement error. The plan for multicity analyses was prompted by questioning the rationale for the study locations previously selected and by the prospect of setting this concern aside with analyses conducted using a defined sampling frame. Additionally, advances in hardware and software made this type of analysis feasible. The NMMAPS project was initiated at the end of 1996, as PEEP was ending.

INTRODUCTION TO NMMAPS PART II

This report provides an integrated synthesis of the key findings of NMMAPS on air pollution and morbidity and mortality. The report begins by introducing the rationale for the multicity approach that is used in NMMAPS and briefly describing the statistical methods used to combine evidence across locations. The findings on mortality are then presented for the 2 databases: the 20 and 90 largest US cities. In the analysis of the 20 cities, the primary analytic thrust was toward estimating the overall effects of PM_{10} and other criteria air pollutants. We used the previously described Bayesian hierarchical model developed for this purpose (Dominici et al 2000; Samet et al 2000a). Air pollution–mortality associations are assessed within the individual cities with previously described methods (Kelsall et al 1997); the evidence is then combined across the cities using the model of Dominici and coworkers. We next provide the results of using a multistage, regional modeling approach for exploring spatial heterogeneity in the 90-city database. We also evaluate sociodemographic and other characteristics of the cities as determinants of heterogeneity in the effects of PM_{10}.

Hospitalization data are also analyzed by combining information across cities. For morbidity, the cities were selected with preference given to those 14 locations having the most abundant PM_{10} measurements. The within-cities time-series analyses are accomplished with a distributed lag approach developed by Schwartz (2000b). Evidence is then combined across locations using hierarchical methods common to meta-analysis. This approach also allows the examination of sociodemographic characteristics of the population as modifiers of the effect of PM_{10} on heart and lung disease. In addition, the assessment of confounding by other pollutants was done in the second stage of the hierarchical model.

BACKGROUND

NMMAPS was initiated to follow up on evidence from daily time-series studies that showed associations between mortality and morbidity—primarily hospitalization in the elderly—at levels of particulate air pollution found in many cities in the United States, Europe, and other developed countries. These observations, published in increasing numbers of reports from the early 1990s (Bascom et al 1996; Pope and Dockery 1999), have motivated reassessment of

air quality standards in many countries throughout the world, including the United States, the United Kingdom, and the countries in the European Union. The key findings on mortality have come from 2 types of epidemiologic studies: (1) time-series studies that address associations between levels of air pollutants and the daily numbers of deaths, both total and within specific categories of cause-of-death; and (2) prospective cohort studies that have assessed associations of estimates of longer-term air pollution exposure with mortality during follow-up of the participants (Pope et al 1995b; Dockery et al 1993). The evidence from these 2 types of studies is complementary. The time-series studies can be readily carried out using publicly available data; their findings provide a warning that air pollution may be adversely affecting public health. The findings of the cohort studies, showing associations between pollution exposure and long-term mortality, suggest that the findings of the time-series studies represent an effect beyond mortality displacement alone.

Although this evidence has motivated reevaluation of the health effects of air pollution, potential limitations of the time-series studies of both mortality and morbidity have been recognized (Lipfert and Wyzga 1997; Vedal 1997; EPA 1996b). Two of the limitations—bias from error in the exposure measures and lack of information on life-shortening arising from associations found on the daily timeframe for mortality—were addressed in NMMAPS Part I. Evidence from the time-series studies was also questioned because the study locations had been seemingly identified without a defined sampling plan and the ability of the findings to be generalized to other locations was uncertain. The data had also been analyzed using somewhat different models, with differing specifications of weather variables, of the lag intervals between the pollution measures, of the averaging time for the pollutants, and of the outcomes. Additionally, the ability of multivariable models to separate the effect of a single component of a pollutant mixture, such as PM_{10}, from the effects of other pollutants was questioned repeatedly (Lipfert and Wyzga 1997; Moolgavkar et al 1995). In its 1996 *Review of the National Ambient Air Quality Standards for Particulate Matter*, the EPA (1996a) acknowledged the limitations of the individual time-series studies while interpreting the pattern of effects across different locations with different air pollution mixtures as indicating adverse effects of PM. Others had similarly interpreted the evidence (Dockery and Pope 1994; Schwartz 1994).

NMMAPS addresses the limitations of evidence derived from time-series studies within single locations. The AIRS database offered a potential sampling frame for selecting study locations based on specific criteria, such as population size or availability of PM_{10} data (AIRS 1999). Additionally, in 1996 when NMMAPS began, PM_{10} data for the United States had been accumulating since 1987 and the monitoring data were sufficiently abundant to support time-series analyses for a number of cities. Software and hardware were no longer a barrier to the analysis of large databases, so feasibility was not an obstacle to carrying out NMMAPS.

A central objective of NMMAPS was to characterize the effects of PM_{10} and each of the other criteria pollutants alone and in combination. In the time-series studies within individual cities, this goal had been met analytically by using multivariable regression models that included PM_{10} and the other pollutants. Regression coefficients estimated from such models are often construed as providing "independent" effects of the various pollutants under the assumption that estimates of log-relative risks are approximately additive. Limitations to this interpretation of model findings are well known. The model, as a representation of underlying biological phenomena, may be misspecified and consequently misleading. In urban air, pollutant levels are typically correlated, often to a moderate degree, because of shared sources and a common relationship to weather, and the model may not be able to separate fully the effects of the pollutants in the model. Errors in the pollution estimates, possibly variable in extent across the pollutants, may complicate model interpretation (Zeger et al 1999a). Sample size may further limit interpretation of model findings within single cities because lengthy series of PM measurements are available for only a few locations and the EPA has required monitoring of PM_{10} every 6 days for regulatory purposes.

Information concerning the effect of a particular pollutant, such as PM_{10}, can also be gained by assessing the effect of the pollutant across different locations that have differing levels of other pollutants such as O_3, or different correlations of those other pollutants with PM_{10}. In fact, concern for PM as the pollutant causing the positive associations found in the daily time-series studies initially came from the coherence of the findings across multiple locations. For example, in the earliest reports of time-series studies during the 1990s, positive associations of PM with mortality were found in Philadelphia (SO_2, NO_2, CO, and O_3 present) (Schwartz and Dockery 1992); the Utah Valley (CO present) (Pope et al 1992); and Santa Clara (low levels of other pollutants) (Fairley 1990). The coherence of the evidence across locations was repeatedly cited in interpreting the evidence as supporting an adverse effect of PM on health (Dockery and Pope 1994; Schwartz 1994; EPA 1996b).

Multicity approaches have the additional advantage of facilitating evaluation of potential modifying factors. As for air pollution, variation across populations is evident in factors that may determine susceptibility, such as socioeconomic status and the underlying mortality rates for heart and lung disease.

Model-based methods offer a more valuable approach than informal syntheses for using the across-city information on the effects of pollutants and potential modifiers. In a 2-step modeling approach, within-city time-series regressions are fit in the first stage. In the second stage, levels of other pollutants or correlations among pollutants are included in the model to explore (1) modification of the effect of a single pollutant by another pollutant, and (2) heterogeneity in pollution effects across cities associated with demographic and other city-specific factors. This approach uses the variation in pollution mixture characteristics across locations in its second stage. For this 2-stage approach to be informative, data from a sufficient number of locations are needed, and there must be heterogeneity of air pollution mixtures across locations. In NMMAPS, we use the hierarchical model developed by Dominici and colleagues (2000) to implement this multistage approach for mortality.

In several reports (Dockery and Pope 1994; Schwartz 1994), meta-analysis was used to combine information across locations in order to derive a summary estimate of the effect of PM on daily mortality counts. The reports commented on the seeming consistency of effect estimates across locations, but the authors did not formally explore confounding or modification of the effect of PM by including other pollutants in a meta-regression.

Air Pollution and Health: A European Approach (APHEA) was a project designed to derive summary estimates of the effects of air pollutants using data from 15 European cities in 10 countries; its methods are similar in principle to those of NMMAPS (Katsouyanni et al 1996). The initial APHEA project included data on mortality and hospital emergency admissions for locations selected because of availability of data and a local team of investigators. The data for each city were analyzed by time-series regression according to a standard protocol, and the city-specific estimates were then combined using a meta-analytic regression approach. Multiple pollutants were considered in the second-stage regression, as were variables for meteorological factors, the accuracy of measurements, and the health of the population. Both fixed-effects and random-effects models were used. For example, an analysis of data from 12 cities addressed the effects of PM and SO_2 on mortality; models were fit to the estimates from the 12 cities that included the pollutants by themselves and together (Katsouyanni et al 1997). APHEA II, now in progress, extends the number of locations.

NMMAPS represents a further evolution of these multicity approaches. The cities for both the mortality and morbidity analyses were selected from a defined sampling frame, and the findings were pooled in a weighted fashion in a second stage of analysis, while controlling for the effects of other potentially relevant factors and pollutants. The sample-based selection of study communities has been extended to both morbidity and mortality, and further analyses will explore the relationship between the effect of air pollution on both morbidity and mortality.

In NMMAPS, the analyses of the hospitalization data have been carried out using distributed lag models (Schwartz 2000b), a further advance in time-series models applied to air pollution and health. Most prior studies had used ad hoc approaches to selecting lag intervals, sometimes exploring the sensitivity of findings to the particular choice of lag. The distributed lag approach allows an assumption-free exploration of the lag structure, and it also allows the lag structure to be constrained on an a priori basis. By combining information across cities, estimates of the effects of air pollutants over multiple days can be made with greater precision than using data from any single city and can be used to examine effect modification and confounding by other pollutants or socioeconomic factors.

MORTALITY

Two sets of mortality analyses are reported: the first is based on the 20 largest cities and the second on the 90 largest cities, including the 20 largest (Figure 1). The 90 cities were selected to encompass the largest locations with the largest number of deaths (Figures 2 and 3) and also with the expectation that monitoring data would be

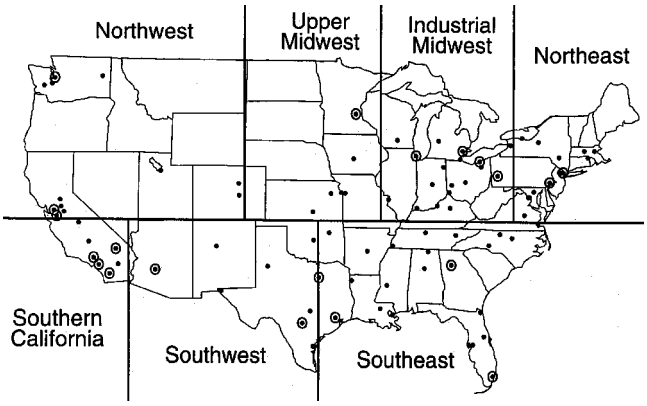


Figure 1. Map of the US showing the 90 cities (the 20 cities are circled) and the 7 regions considered in the geographic analysis.

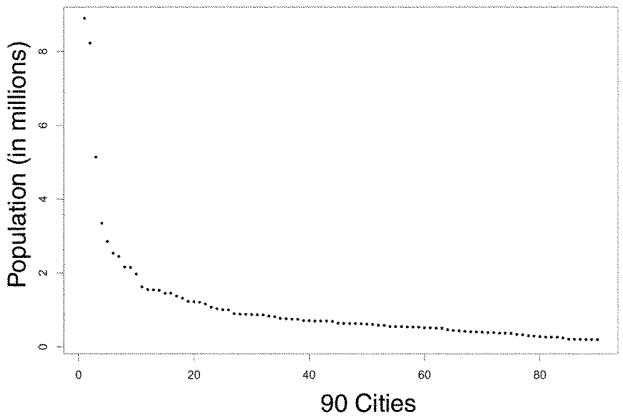


Figure 2. Populations for the 90 cities in the 1990 census.

available for multiple pollutants, including PM_{10} (Figures 4 through 8). These data were symmetrically analyzed to explore not only the effect of PM_{10} but also the effects of other pollutants. The data set was then expanded to include the largest 90 cities. This larger data set was used primarily to explore spatial variation in the effect of PM_{10} and determinants of heterogeneity of the effect. The analytic methods for the 20 cities were those developed by Dominici and colleagues (2000) and described fully in NMMAPS Part I.

Methods for the 90-cities analysis are novel. A 3-stage regional model was developed to estimate PM_{10} effects in several US regions. Estimation was performed using Markov chain Monte Carlo (MCMC) methods. Second-stage regressions to identify determinants of heterogeneity of the effects were performed by using a weighted regression method. The univariate effects of all pollutants for the 90 cities were estimated using the MCMC method. All additional data analyses focused on estimating adjusted effects of pollutants, and sensitivity analyses and other model checking used a random-effect weighted-average approach as in DerSimonian and Laird (1986). We found that the weighted-average approach and MCMC yielded very similar results (Figure 9), and for computational practicability we used the weighted-average method.

A multistage analysis was used because a uniform analytic approach was applied to each city, power is gained from the pooled sample, and heterogeneity can be explored across the locations. Hierarchical models offer a flexible approach to the analysis of multilevel data (Morris and Normand 1992; Lindley and Smith 1972). The hierarchical approach provides a unified framework for estimating individual pollutant effects for particular cities, covariate effects, and components of variation. This approach facilitates more precise estimation of relative rates within each city than can be accomplished by analysis for each city individually. In addition, multistage analyses allow estimation of overall or average pollution effects while taking into account the variability in the air pollution–mortality association across cities. For example, if there is substantial heterogeneity in the relative rates across cities, the estimate of the overall pollution effect will be less precise. The estimation of city-specific and overall relative rates, and of the within-cities and between-cities variability, is carried out with MCMC methods (Gelfand and Smith 1990). One useful feature of these methods is that they provide, in addition to the point estimate, an approximation of the entire posterior distribution of the unknown parameter.

METHODS

Data

We obtained data on mortality, weather, and air pollution for the selected metropolitan areas in the United States from publicly available data sources. In initially selecting the locations, we first listed the cities in rank order by population according to the 1990 Census. Because mortality data were available only at the county level, we then used the deaths for the counties comprising the selected cities. For some locations, it was possible to separate the city and the county deaths, as the city was a subunit of the county or distinct from the county: Cook County and Chicago, Los Angeles County and the city of Los Angeles, and Baltimore County and Baltimore City. The selected counties and associated cities, along with the abbreviations that are used in this report, are provided in Appendix A.

Daily mortality counts were obtained from the NCHS. After excluding deaths from external causes and deaths of nonresidents of the county, we classified the deaths by age group (< 65 years, 65–74 years, and ≥ 75 years) and by cause according to the ninth revision of the *International Classification of Diseases* (ICD-9): cardiac (390–448); respiratory, including chronic obstructive pulmonary disease (COPD) and related disorders (490–496); influenza (487); pneumonia (480–486, 507); and the other remaining diseases. Hourly temperature and dew point were available from the National Climatic Data Center, as assembled in the Earth-Info CD database. For analysis, we used the 24-hour mean temperature and mean dew point for each day.

The air pollution data were obtained from the AIRS database, a computerized repository of information about airborne pollution in the United States and various other countries. Its information includes air quality, emissions, compliance, and enforcement. The air pollution data are collected using reference methods established by the EPA

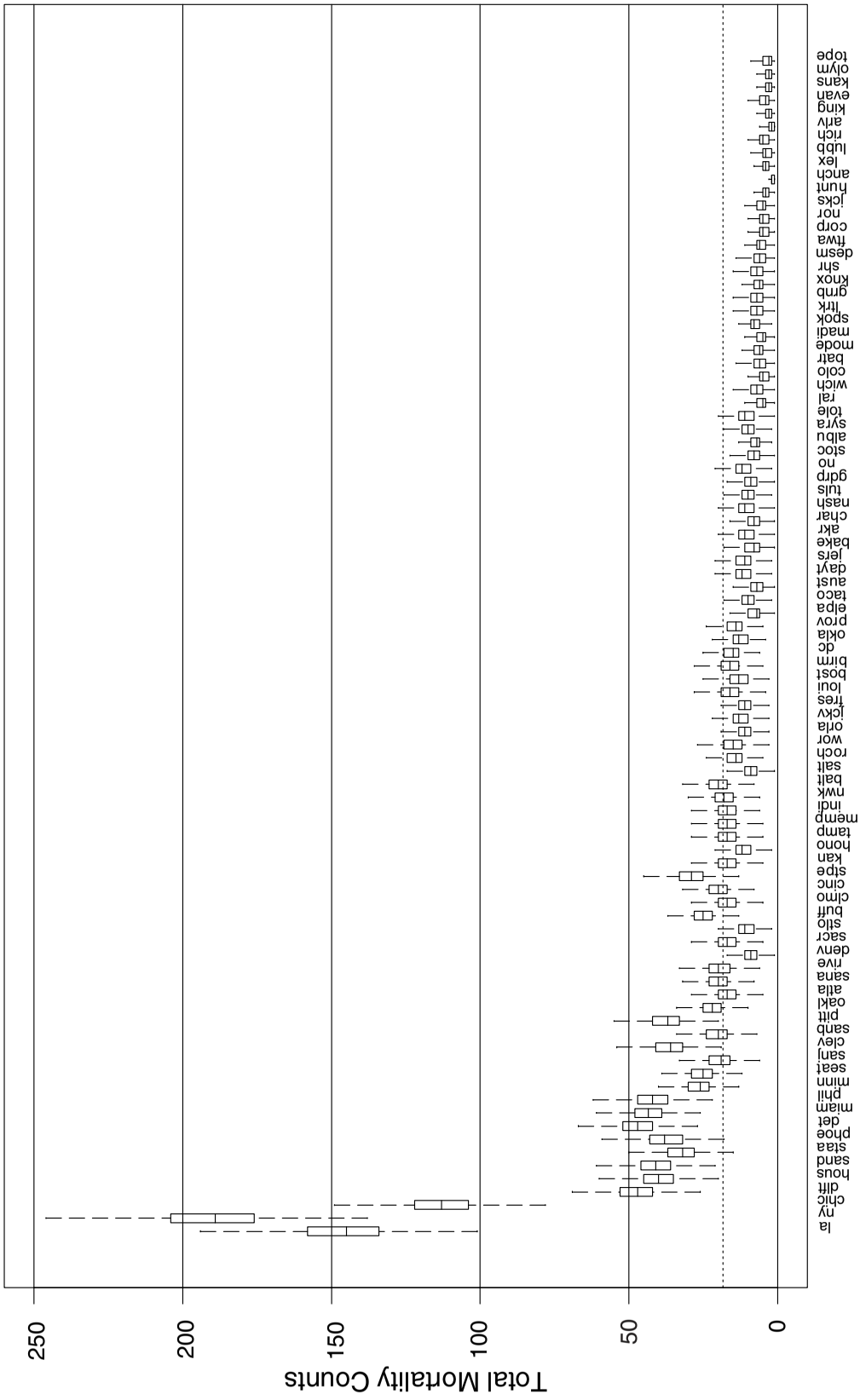


Figure 3. Daily mortality counts for the 90 cities, 1987–1994. Boxplots show the median and IQR (box) with the 1% to 99% range.

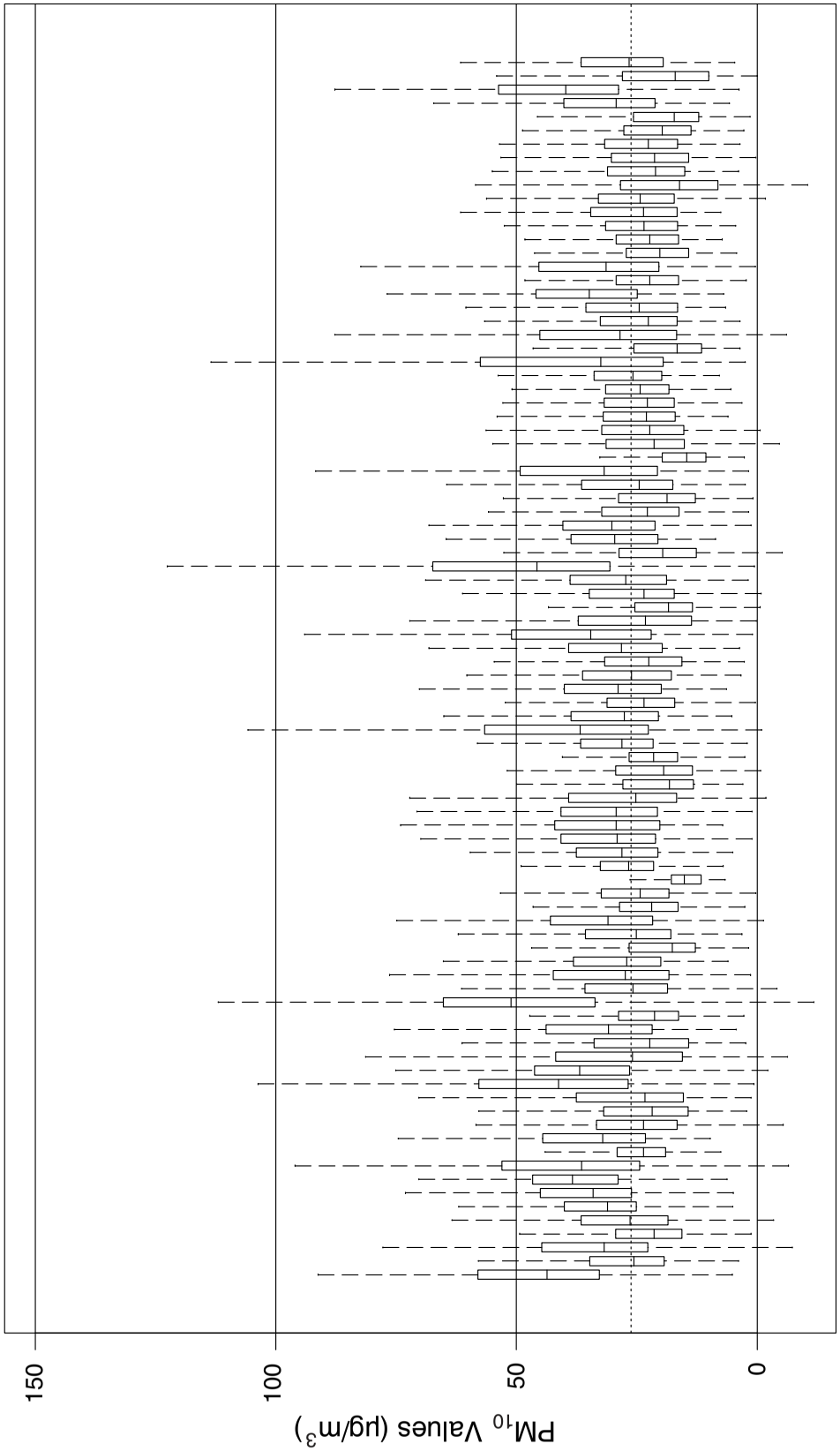


Figure 4. Daily concentrations of PM_{10} for the 90 cities, 1987–1994. Boxplots show the median and IQR (box) with the 1% to 99% range. (Pollution values are based on a 10% trimmed mean as described in Appendix E.)

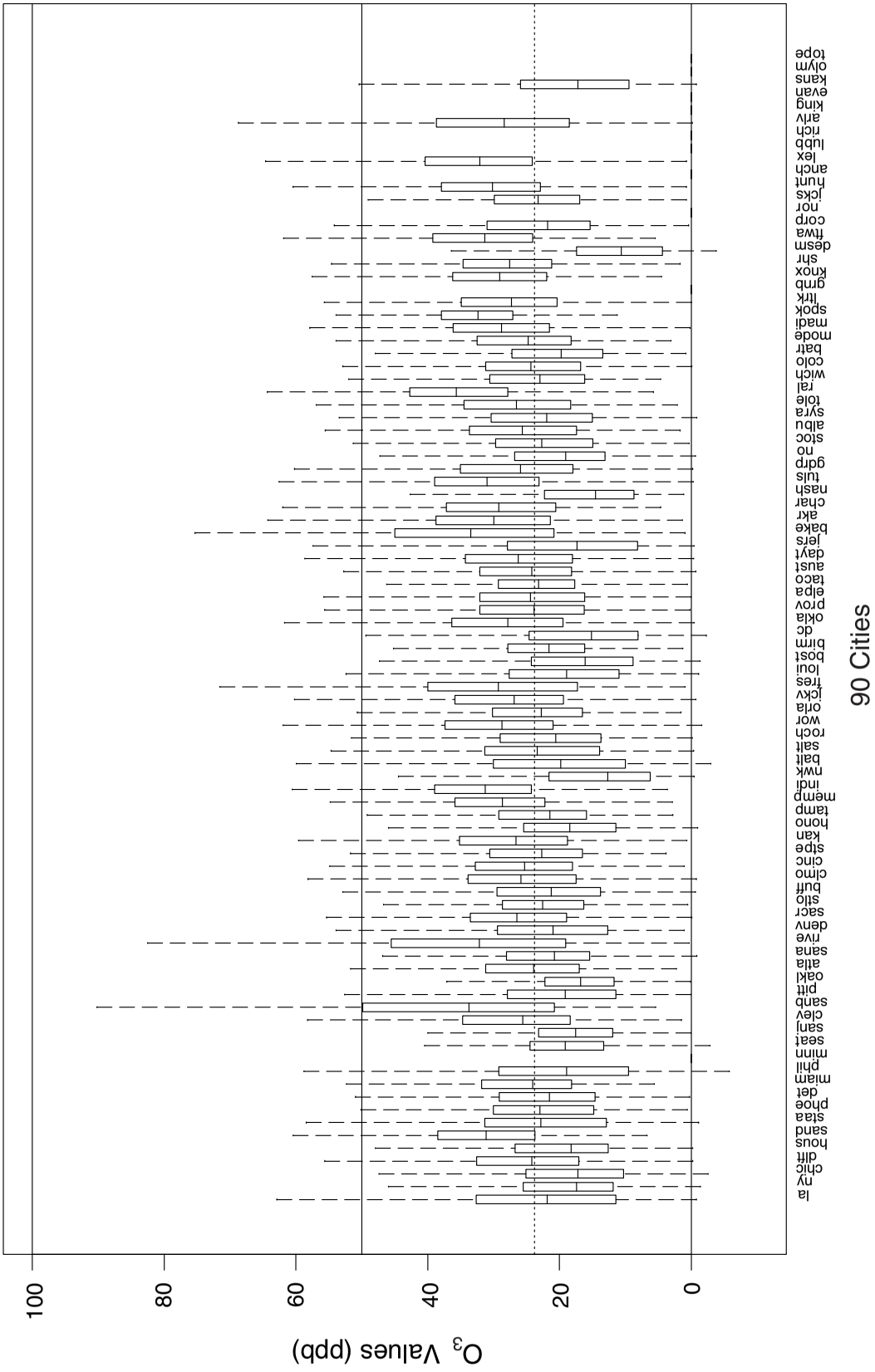


Figure 5. Daily mean concentrations of O_3 for the 90 cities, 1987–1994. Boxplots show the median and IQR (box) with the 1% to 99% range. (Pollution values are based on a 10% trimmed mean as described in Appendix E.)

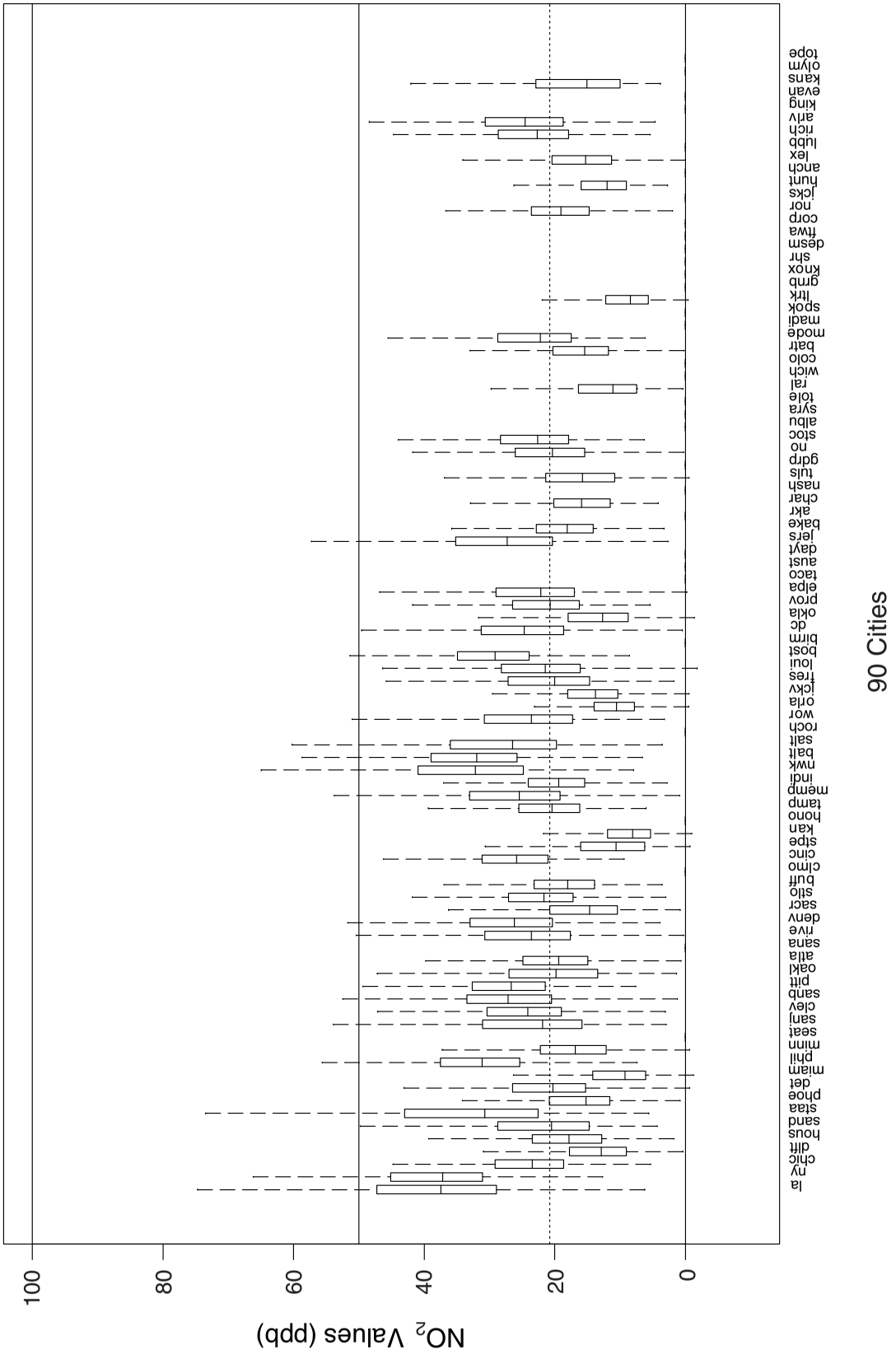


Figure 6. Daily mean concentrations of NO_2 for the 90 cities, 1987–1994. Boxplots show the median and IQR (box) with the 1% to 99% range. (Pollution values are based on a 10% trimmed mean as described in Appendix E.)

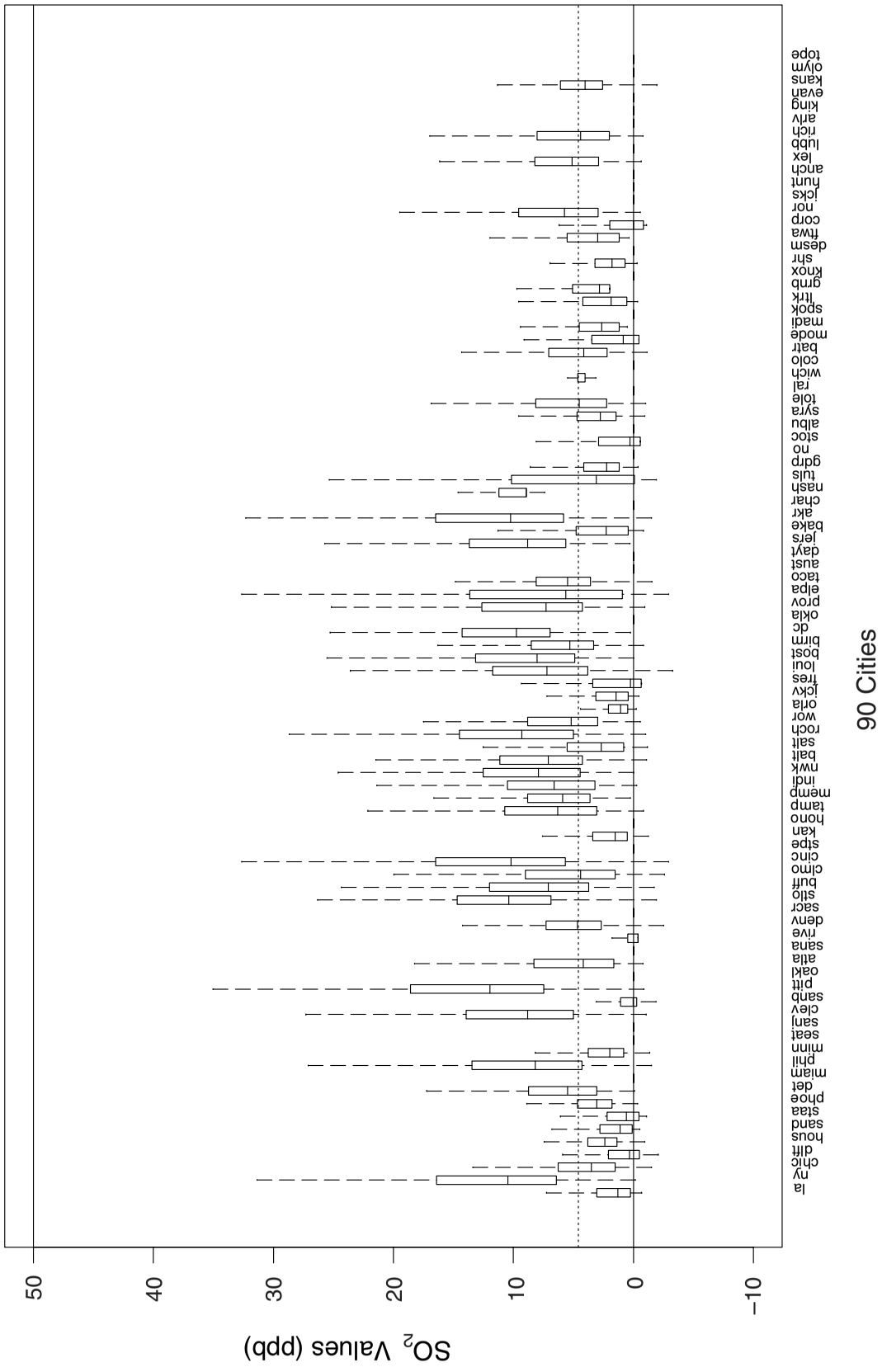


Figure 7. Daily mean concentrations of SO_2 for the 90 cities, 1987–1994. Boxplots show the median and IQR (box) with the 1% to 99% range. (Pollution values are based on a 10% trimmed mean as described in Appendix E.)

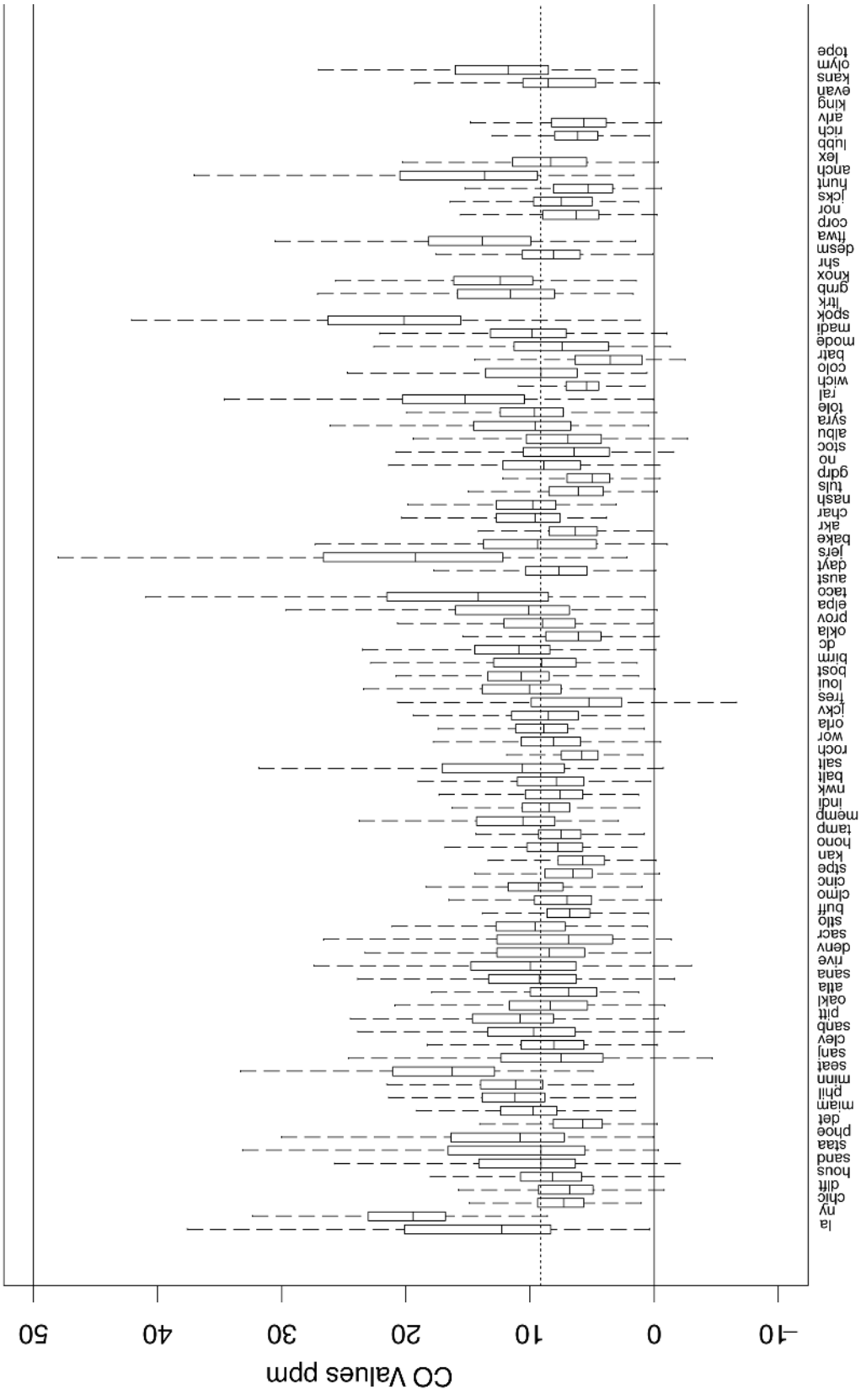


Figure 8. Boxplots of daily mean concentrations of CO for the 90 cities, 1987–1994. Boxplots show the median and IQR (box) with the 1% to 99% range. (Pollution values are based on a 10% trimmed mean as described in Appendix E.)

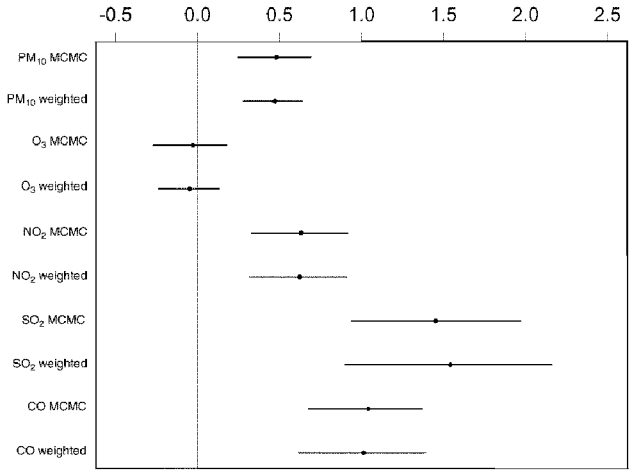


Figure 9. Comparison of univariate lag-1 results of all pollutants for the 90 cities, first pooled by the MCMC method, then by the weighted

(*www.epa.gov/airs/airs.html*). We downloaded all available data for PM_{10}, O_3, CO, SO_2, and NO_2. For the pollutants measured on an hourly basis, we calculated the 24-hour average. If multiple monitors were available for a metropolitan area, we averaged the data. To protect against the potential consequences of outlying values, a 10% trimmed mean was used to average across monitors, after correction for yearly averages for each monitor. Details of analysis methods are in NMMAPS Part I (Samet et al 2000b) and in Dominici and colleagues (2000).

While the 24-hour average O_3 value was used, the average concentration is tightly correlated with the maximum value. For example, for 1987-1994 for the cities of Chicago, Los Angeles, and New York, correlation coefficients of the maximum 24-hour O_3 value with the average 24-hour O_3 value were above 0.9 in each location.

Time-series plots of the raw data were generated for quality control purposes and inspected for unexplained extreme outlying values, for sudden changes in the day-to-day variation in the series indicative of possible reporting errors, and for missing data. Some exclusions were made based on this review, including several gaps in the mortality series for specific cities and for selected pollutants in specific cities when data were available for only a brief interval.

Data Analysis: 20 Cities

Full details of the analytic methods for the analyses of the 20 cities have been described in NMMAPS Part I (Samet et al 2000b) and are only briefly summarized here. In the first stage of the analysis, the basic model for each city is a log-linear generalized additive model that accounts for smooth longer-term fluctuations in mortality (Hastie and Tibshirani 1990), potentially confounding the pollution-mortality associations at the daily level. In the second stage of the analysis, the pollution-mortality associations in the individual locations are combined using hierarchical models (Morris and Normand 1992; Lindley and Smith 1972). We used the Bayesian hierarchical model developed by Dominici and colleagues (2000). The Data Analysis: 90 Cities section in this report provides further key details.

In the first-stage models, the outcome measure in the log-linear model is the observed daily mortality. The log-linear model allows the mortality counts to have variances that may exceed their means (ie, be overdispersed) and the overdispersion parameter is allowed to vary by location. To control for possible confounding by longer-term trends due to changes in population size and characteristics, health status, and health care, and to control for possible confounding by shorter-term factors such as seasonality and influenza epidemics, we introduced smooth functions of calendar time for each city, allowing 7 degrees of freedom (*df*) for each year when the mortality record was complete. We allowed these smoothing functions to vary by age group, adding a separate smooth function of time with 8 *df* for each age group. To control for the effect of weather on mortality, we also fit smooth functions of the same day's temperature and average temperature for the 3 previous days, along with comparable functions for dew point. Finally, we included indicator variables for the day of the week. These model specifications were based on extensive exploratory analyses that have been previously reported (Samet et al 1996; Kelsall et al 1997). The sensitivity of findings to key assumptions with regard to smoothing for time and to control of weather variables was explored (see Figure A.1).

As there were missing values for some variables on some days, we restricted analyses to days with no missing values for all covariates. More specifically, to estimate the main effect of PM_{10} on mortality, we restricted the analysis to days with data for PM_{10}. To estimate the effect of PM_{10} on mortality with adjustment for O_3, however, we restricted the analysis to days with no missing data for either pollutant. Consequently, because of missing data, the estimates of the main effects of PM_{10} were more precise than the estimates of the effects of PM_{10} adjusted by other pollutants. The model estimates the log-relative rate for each pollutant in the model (that is, a coefficient gives the percentage change in mortality per unit change in the pollutant). In this report, we express all results as the percentage change in mortality per 10-unit change in the pollutant. The model fitting also supplies the estimated covariance matrix for

log-relative risks, which is needed for the second stage of the analysis.

Because the PM_{10} data were often available only on every sixth day, it is difficult to estimate distributed lag models for the full set of cities. Hence, we restricted our attention to models in which pollutant levels on a single day were used to predict future mortality. We did, however, explore a range of lags for the pollutant variables, including the current day's pollution data, and 1-day and 2-day lags. In the second stage of the analysis, we further explored the consequences of varying the lag interval.

In our data set, for 3 of the 20 cities (Chicago, Minneapolis/St Paul, and Pittsburgh) we have at least 85% of possible daily PM_{10} values. For these 3 cities, we fit unconstrained distributed lag models as proposed by the Harvard group (see Appendix B).

We approached the assessment of the effects of individual pollutants by including multiple pollutants in models in a sequential fashion. The levels of combustion-related pollutants tend to be correlated, and consequently estimates from models that include multiple pollutants need to be interpreted with caution (Kelsall et al 1997). On the other hand, estimates from models with only a single pollutant may also capture some effect from other, uncontrolled pollutants. We initially carried out univariate analyses for PM_{10} and O_3, and then considered their effects in a bivariate model that included both pollutants. Because O_3 levels are substantially higher in the summer than in the winter, the univariate analyses for O_3 were repeated with stratification by season. We next explored the sensitivity of the estimates to the addition, one at a time, of the remaining pollutants—CO, SO_2, and NO_2—into trivariate models for air pollution. We further explored the effects of CO, SO_2, and NO_2 on mortality while including PM_{10} and O_3 in the models.

Data Analysis: 90 Cities

The analysis of the data from the 90 cities was directed at describing the heterogeneity of effects across the cities, at the determinants of heterogeneity, and at estimating regional effects. We used a new 3-stage regional model to estimate PM_{10} effects for multiple US regions and a weighted linear regression approach to identify determinants of heterogeneity of PM_{10} coefficients across locations.

To estimate regional effects, a 3-stage regional model was then applied to the data from the 90 cities. The 3 stages of the regional model describe (1) within-city variability, (2) within-region variability, and (3) between-regions variability. In the first stage, we estimated the log-relative mortality rate associated with PM_{10} and its standard error for each location using a semiparametric log-linear model, as in Dominici and colleagues (2000):

$$\log(\mu_t{}^c) = \beta_R^c \mathrm{PM}_{10t}^c + \textit{confounders} \qquad (1)$$

where β_R^c denotes the log-relative of mortality for location c in region R. The second stage of the model describes between-city variation in the true log-relative rates within region and estimates regional effects $\beta_R^c = \beta_R + \varepsilon_R^c$. The third stage describes between-region variation in the true regional coefficients and estimates the overall vector of regression coefficients and the overall variance of the regional coefficients $\beta_R = \beta^* + \varepsilon_R$.

Specifying dispersed but proper baseline conjugate prior distributions completes the Bayesian formulation. To approximate the posterior distribution of all the unknown parameters, we implement an MCMC algorithm with a block Gibbs sampler. The full conditional distributions were available in closed form. Their derivation was routine and is not detailed here.

Combining the data across cities requires an assumption concerning the extent of heterogeneity in the air pollution effect on mortality among the locations. The analyses of both the 20 and 90 cities shown to this point have assumed some heterogeneity (that is, the possibility of homogeneity has been excluded). We explored the consequences of this assumption by also considering a prior assumption that gave some weight to complete homogeneity.

This prior model, designated *model A* [inverse gamma(3,3) for σ^2; inverse gamma(3,1) for τ^2] assumes heterogeneity across cities and regions, possibly substantial in size, and excludes homogeneity. The effect of this prior assumption is twofold: the city-specific relative risk estimates draw more heavily on data from some cities and less heavily on data from other cities; and this model yields conservative confidence bands on the overall relative risk. To be conservative, this prior assumption was used in all analyses of the 20-cities data because only 20 cities does not provide a strong picture of the degree of heterogeneity in the data.

In the second approach, the alternative assumption in the 90-city analysis, we allowed for little or no heterogeneity as well as more substantial heterogeneity (model B) [half normal(0,1) for σ^2, half normal(0,1) for τ^2]. Under B, the data from the 90 cities were themselves used to inform the degree of heterogeneity assumed by the model. The resulting city-specific estimates draw on data for each city, but more heavily from other cities than in model A. This second approach provides more realistic, less conservative intervals for the overall effect. Comparison of the two sets of estimates shows, as anticipated, less homogeneity

within region for model B. Results are provided in Appendix C.

For the regression analyses on the predictor variables in Table 1, we used weighted linear regression. The weight for each city was the inverse of the sum of the statistical variance for its PM_{10} relative rate, determined from the first stage, and an estimate of the natural variability in true relative rates, determined from a full Bayesian analysis using MCMC. The best models were chosen to have the smallest C_p value (Mallows 1973). Mallows C_p statistic is used as a covariate selection criterion in linear regression analysis. Better-fitting models have small C_p values that are close to the number of predictors.

Data Analysis: A Weighted Second-Stage Regression Method

Multiple predictor variables were considered in the second stage of the analysis for the 90 cities. From the air pollution and mortality databases assembled for NMMAPS and from the 1990 CensusCD (*www.censuscd.com*), we constructed a data set of county-specific variables. We then organized the variables into 5 distinct groups for the purpose of variable selection: (1) mean levels of pollutants, temperature, and dew point; (2) mortality variables, including crude mortality rate over the time span of the analysis; (3) sociodemographic variables from the 1990 Census, including percentages of persons in poverty and persons lacking a high school degree; (4) several variables related to urbanization, including the percentage of the population classified as urban and the percentage of the population using public transportation; and (5) variables related to the measurement error of PM_{10} exposure levels. The 33 variables considered are listed in Table 1. Within each of these 5 categories, we examined pairwise scatterplots of the 90 PM_{10} coefficients with the variables. Based on the pattern of correlation, we limited the number of variables in each of the 5 subsets and created several summary variables. The resulting new set of 9 variables is identified in Table 1 with a footnote.

We fit weighted linear regression to estimate the effects of the county-specific variables on the PM_{10} coefficients. All predictors were centered with respect to their mean, so that the intercept can be interpreted as the log-relative rate of mortality for PM_{10} when the predictor is centered at its mean value.

We consider the following model:

$$\hat{\beta}^c = \alpha_0 + \sum_{j=1}^{k} \alpha_j (X_j^c - \bar{X}^c) + \varepsilon^c. \quad (2)$$

Table 1. City-Level Variables Evaluated for Second-Stage Models by Group

Mean levels of pollutants, temperature, and dew point
Effect of PM_{10} lag 1 on total mortality[b]
Mean of PM_{10} levels (PM_{10})[b]
Median PM_{10} cross correlation (MCC)[b]
Mean of O_3 levels (O_3)[b]
Mean of NO_2 levels (NO_2)[b]
Mean of SO_2 levels (SO_2)
Mean of CO levels (CO)
Mean temperature
Mean dew point

Total mortality rate[a]
Mean total mortality
Crude mortality rate (CMR)[b]
log(%RESP/%OTHER)
log(%CVD/%OTHER)
log(%<65/%>75)
log(%65–75/%>75)

Sociodemographic variables from 1990 US Census
% High school degree or above (%NoHS)[b]
% College educated
% Unemployed (%Unemp)
% Drive alone to work
% Public transportation (%PubTrans)
Mean travel time to work
Median household income (HsInc/1000)[b]
Median family income
Median per capita income
% Poverty (%Pov)

Variables related to urbanization
Total population
Area
% Urban population
% Drive alone to work
% Public transportation (%PubTrans)[b]
Mean travel time to work

Variables related to measurement error
Number of monitors
Median of all pairwise correlations between monitors (PC)

[a] Calculated as the crude mortality rate over the time span of the analysis.

[b] Variables selected for the second-stage analysis.

Estimates of $\hat{\alpha}_0$ and $\hat{\alpha}_j$ are found by weighted least squares regression with weights $\hat{w}_j^c$ based on a random-effects model, defined as

$$\hat{w}_j^c = \frac{1}{\hat{v}^c + \tau^2},$$

where $\hat{v}^c$ is the estimated statistical variance of $\hat{\beta}^c$ and τ^2 measures the heterogeneity of the true slopes across locations, which was fixed at the estimate obtained from the

MCMC analysis under the first prior model of substantial heterogeneity (model A).

Because all cities do not monitor all pollutants, the second-stage regressions must take account of missing explanatory variables, particularly mean NO_2, O_3, and the median cross-correlation among PM_{10} monitors. We therefore used multiple imputation (Little and Rubin 1987) to estimate the stage 2 regression in the presence of missing covariate information. Briefly, we created 5 complete data sets by simulating imputed values for the missing covariate from its posterior distribution, given all the other covariates and log-relative risk for that city. For each of the 5 complete data sets, we refit the stage 2 model, obtaining regression coefficients and their estimated covariance matrix. The final regression coefficients are the mean of the 5 sets of coefficients. The variances (standard deviations squared) are the sum of the average of the 5 statistical variances calculated from each complete data set plus the variance of the 5 sets of coefficients across the replications. By incorporating the variation in the coefficients across 5 distinct imputations of the missing covariates, we see that the standard errors in Table 5 reflect increased uncertainty from having partial covariate information for some cities.

All analyses were carried out using the statistical language S-plus.

RESULTS

City Characteristics

The selected cities represent nearly every region and state of the United States excepting only the sparsely populated states (Figure 1). Population ranged from nearly 9 million (Los Angeles) to 160,976 (Topeka) for the 90 cities and to 1,185,394 (San Antonio) for the 20 cities (Figure 2). The numbers of deaths followed a similarly broad range (Figure 3). Detailed tables concerning the selected counties are provided in Appendix A.

Box plots for the daily values of the 5 criteria pollutants considered in these analyses are represented in Figures 4 through 8. The extent of monitoring data available varied among the pollutants (see Appendix A). Mean daily values for PM_{10} ranged from about 20 $\mu g/m^3$ to near 50 $\mu g/m^3$. The present 24-hour standard of the EPA is 150 $\mu g/m^3$. For PM_{10}, substantial variation was evident from day to day (Figure 4); median values tended to drop with population size. Values for O_3 were also variable but had no relation to population size (Figure 5). Data were not as abundant for the remaining pollutants (Figures 6 through 8). Consequently, multipollutant models drew on different sets of locations, depending on the selected pollutants.

Numbers of PM_{10} monitors varied across the locations. To characterize the quality of PM_{10} exposure estimates better, we have calculated the median of all pairwise correlations among monitors within each location. Figure 10 shows the median correlations for each location. The numbers following city name denote number of monitors.

Results for PM_{10}

The initial univariate analysis for the 20 cities, conducted with prior assumption of substantial heterogeneity, indicated that the lag-1 PM_{10} concentration was positively associated with total mortality in most locations (Figure 11; Appendix A provides the coefficient values and 95% confidence intervals [CIs]). The estimated effects ranged from over 1% per 10 $\mu g/m^3$ in New York City, San Diego and Oakland to no effect or even a negative effect in Dallas/Fort Worth, Cleveland, and Atlanta.

We next explored the sensitivity of the findings for PM_{10} to the inclusion of other pollutants to the model. We initially included O_3 with PM_{10} to make the 2-pollutant or bivariate model and then added the other pollutants (NO_2, SO_2, or CO), yielding 3-variable or trivariate models.The associations of PM_{10} with mortality changed little with the addition of O_3 to the model or with the addition of a third pollutant in the trivariate models. The stability of the associations of PM_{10} with total mortality is shown in pairwise plots of the estimated relative rates (Figure 12). The points in these plots, with 1 point for each city, represent the values of the PM_{10} coefficients with (vertical axis) and without (horizontal axis) control for additional pollutants. The larger circle is plotted in correspondence to the pooled PM_{10} coefficients with and without adjustment. The sets of coefficients generally tracked along the line of identity for PM_{10}. This figure suggests that the effect of PM_{10} is robust to the inclusion of other pollutants.

A similar pattern of effects of PM_{10} was evident for the grouping of cardiovascular and respiratory deaths (Figure 13). The associations of PM_{10} with mortality from these causes, as for total mortality, were robust to the inclusion of other pollutants in the model. In general, PM_{10} was not associated with mortality for the grouping of other causes of death (Figure 14).

The pooled analysis of the 20-cities data confirmed the association of PM_{10} with total mortality and with cardiovascular and respiratory deaths. Figures 15 and 16 provide the marginal posterior distributions of the overall effect of PM_{10} on total mortality and cardiorespiratory deaths, respectively, at lags 0, 1, and 2.

The posterior distributions indicate the strength of evidence that the pollutant effect is greater than 0. The posterior probability that the overall effect is greater than 0

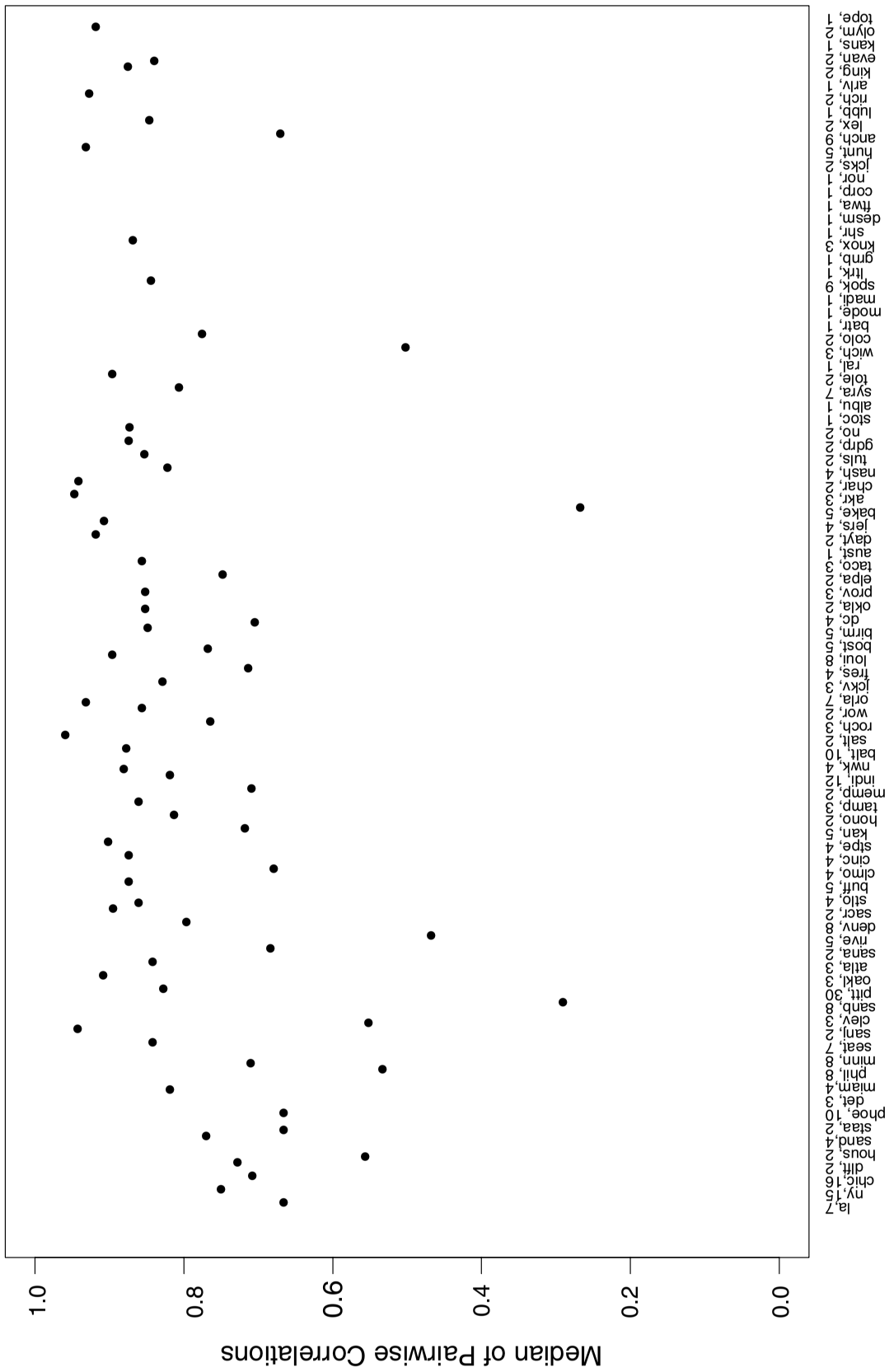


Figure 10. Median correlations for each of the 90 locations (and number of monitors). Aggregation of the measured PM_{10} levels across monitors is done by a 10% trimmed mean.

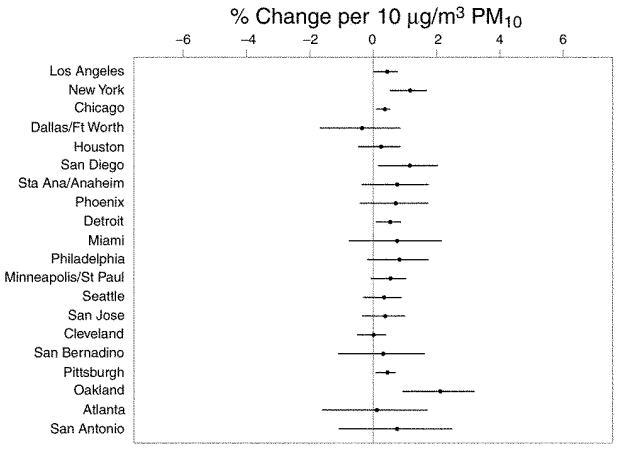


Figure 11. PM_{10} effect without adjustment for O_3 in the 20 cities. Each bar shows the regression coefficients and 95% CIs for total mortality.

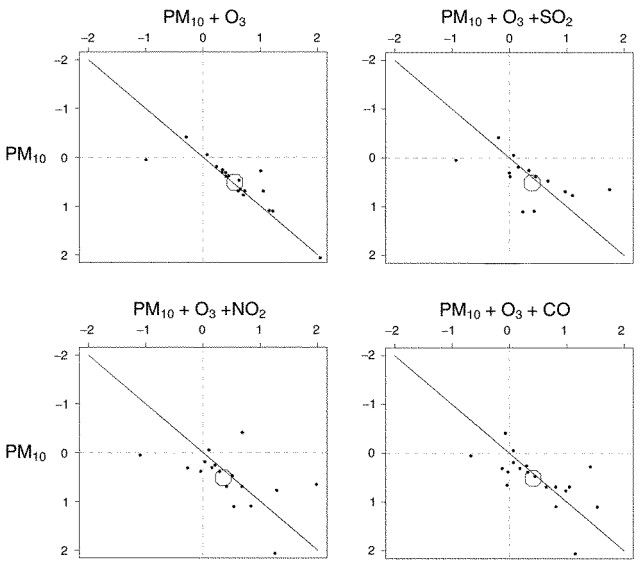


Figure 12. Pairwise plots of the estimated relative rates for effects of PM_{10} on mortality adjusted for O_3, O_3 and NO_2, O_3 and SO_2, and O_3 and CO for the 20 cities. The empty circles are plotted at the posterior means of the corresponding overall effects. Small dots are plotted at the MLE estimates of each city. The number of the dots may be smaller than 20 because of the missing data.

is given in the plot in the box. Note that these posterior distributions are conservatively wide because they are based on the prior model of substantial heterogeneity across cities. For total mortality, the distributions are shifted toward the right and centered on an effect of approximately 0.5% for each 10 µg/m^3 increase in PM_{10} at lag 1; the posterior distributions indicate that the effect is not likely to be due to chance. For cardiorespiratory deaths (Figure 16), the effect of PM_{10} is somewhat greater than it is for total mortality, and there is an even greater posterior probability that the effect is larger than 0. In contrast to the PM_{10} findings for cardiorespiratory deaths, there was less evidence for an effect of PM_{10} for noncardiorespiratory deaths (Figure 17). The posterior distributions for PM_{10} did not shift with inclusion of other pollutants, suggesting that the univariate findings were not affected by confounding by other pollutants (Figure 18).

Figure 19 shows the estimates of the PM_{10} effects for Chicago, Minneapolis/St Paul, and Pittsburgh under the unrestricted distributed lag model and estimates made using the current day's pollution data, or with 1-day and 2-day lags. The summary of the 7-day distributed lag coefficients was greater than each of the estimates based on a single day's value. The 14-day estimate was substantially lower than the 7-day estimate in Chicago and Minneapolis/St Paul.

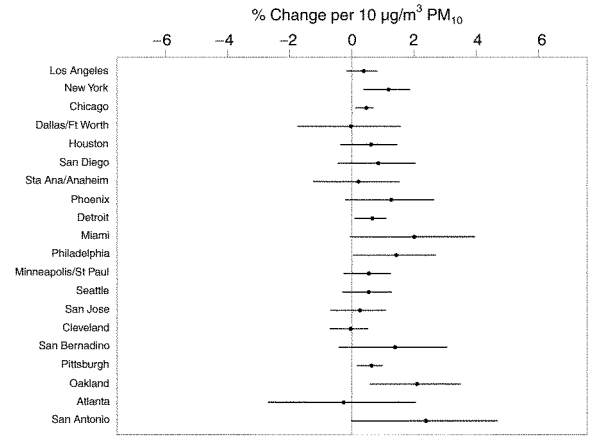


Figure 13. PM_{10} effect without adjustment for O_3. Regression coefficients and 95% CIs for cardiorespiratory mortality for the 20 cities.

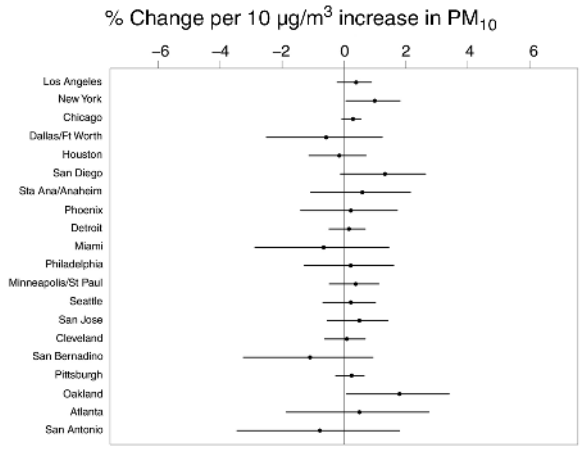


Figure 14. PM_{10} without adjustment for O_3. Regression coefficients and 95% CIs for other mortality for the 20 cities.

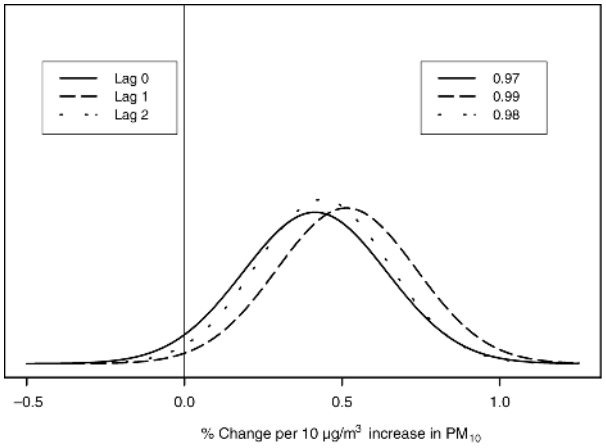


Figure 15. Marginal posterior distributions for effects of PM_{10} on all-cause mortality at lags 0, 1, and 2, without control for other pollutants, for the 20 cities. The box to the top right provides the posterior probabilities that the overall effects are greater than 0.

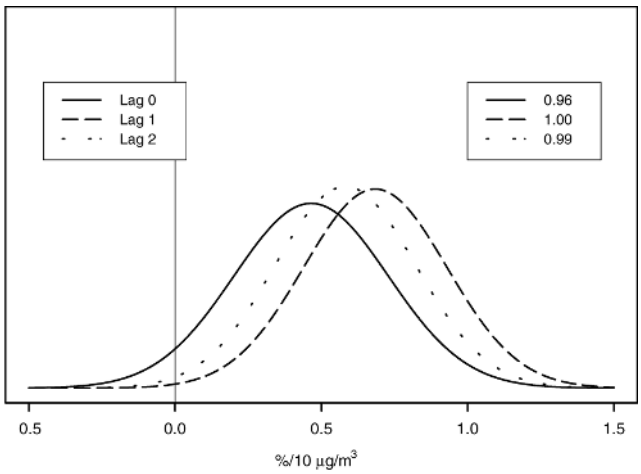


Figure 16. Marginal posterior distributions for effects of PM_{10} on cardio-respiratory mortality at lags 0, 1, and 2, without control for other pollutants, for the 20 cities. The box to the top right provides the posterior probabilities that the overall effects are greater than 0.

In the next phase of the analysis, we explored factors modifying the effect of PM_{10} in the 90 cities as determined using the prior assumption model of substantial heterogeneity. We first fit simple univariate weighted linear regression models to estimate the effects of the mean levels of pollutants and also of the sociodemographic variables on the PM_{10} coefficients (the weights were defined earlier). The results are summarized in Tables 2 and 3. For the variables considered, effects were weak and none were statistically significant.

We then fit multivariate weighted linear regression models to assess further factors that might predict the PM_{10} coefficients, using the variables previously selected (Table 1). More specifically, within each of the 5 groups of variables (Table 1) we identified the optimal subset of predictors by the value of the prediction error, as measured by the C_p statistic (Mallows 1973). In reviewing the optimal subset of models, we identified a final list of 8 predictors: mean PM_{10}, O_3, and NO_2 levels; fraction of persons without a high school degree (%NoHS); percentage of persons using public transportation (%PubTrans); household income (HsInc); crude mortality rates (CMR); and the median of all pairwise correlations across monitors (MCC).

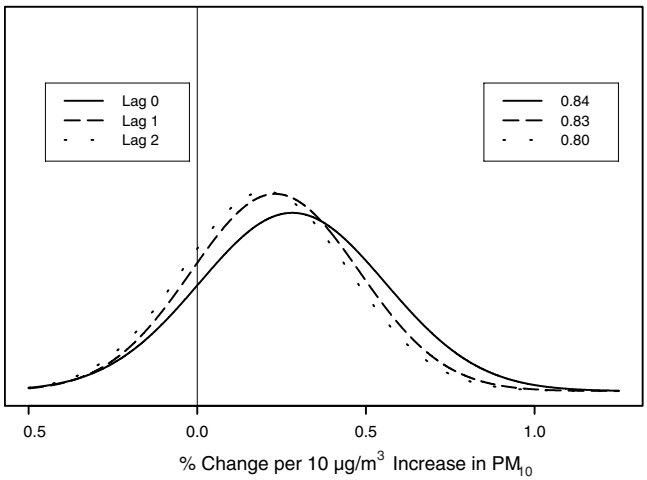


Figure 17. Marginal posterior distributions for effects of PM_{10} on other mortality at lags 0, 1, and 2, without control for other pollutants, for the 20 cities. The box to the top right provides the posterior probabilities that the overall effects are greater than 0.

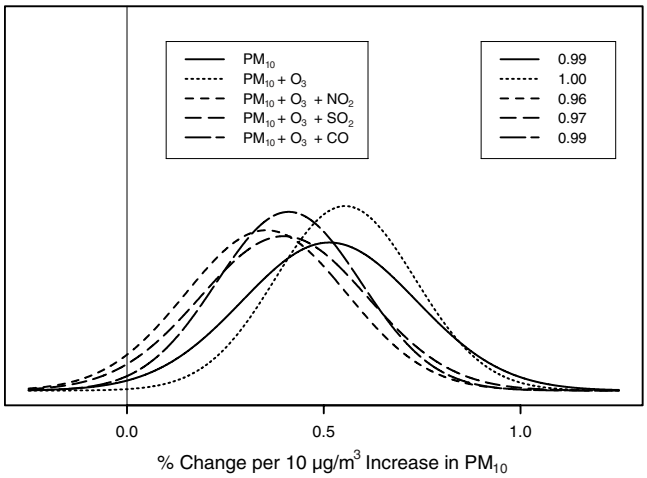


Figure 18. Marginal posterior distributions for effects of PM_{10} on total mortality at lag 1 with and without control for other pollutants, for the 20 cities. The box to the top right provides the posterior probabilities that the overall effects are greater than 0.

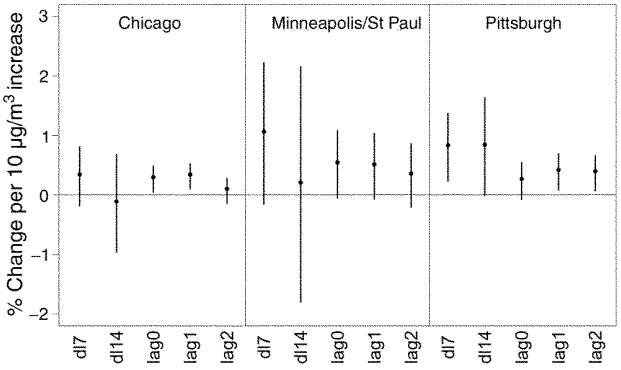


Figure 19. Summary estimates of the PM_{10} effects for Chicago, Minneapolis/St Paul, and Pittsburgh under the unrestricted distributed lag model for 7 days and 14 days as well as estimates made for lags 0, 1, and 2.

Table 2. Point Estimates of Weighted Linear Regression Model of PM_{10} Log-Relative Rates for 90 Cities Against Mean Levels of Pollutants in 90 Cities

Pollutant	Slope[a] (SD)	
PM_{10} (µg/m^3)	–0.009	(0.010)
O_3 (ppb)	0.008	(0.016)
NO_2 (ppb)	0.016	(0.012)
SO_2 (ppb)	0.023	(0.022)
CO (ppm)	0.0084	(0.020)

[a] The slope is the expected value of the PM_{10} log-relative rate per 1 unit change of mean level of pollutant (see Equation 2).

Figure 20 shows all pairwise scatterplots of city-specific coefficients for the effect of PM_{10} divided by the inverse of their standard deviations and the final list of predictors. We then found the best regressions among the identified subset of 9 exploratory variables by using the leaps function in S-plus (Weisberg 1985). The leaps function provides a way to select a few promising regressions (sets of exploratory variables) based on the prediction error as measured by the C_p statistic.

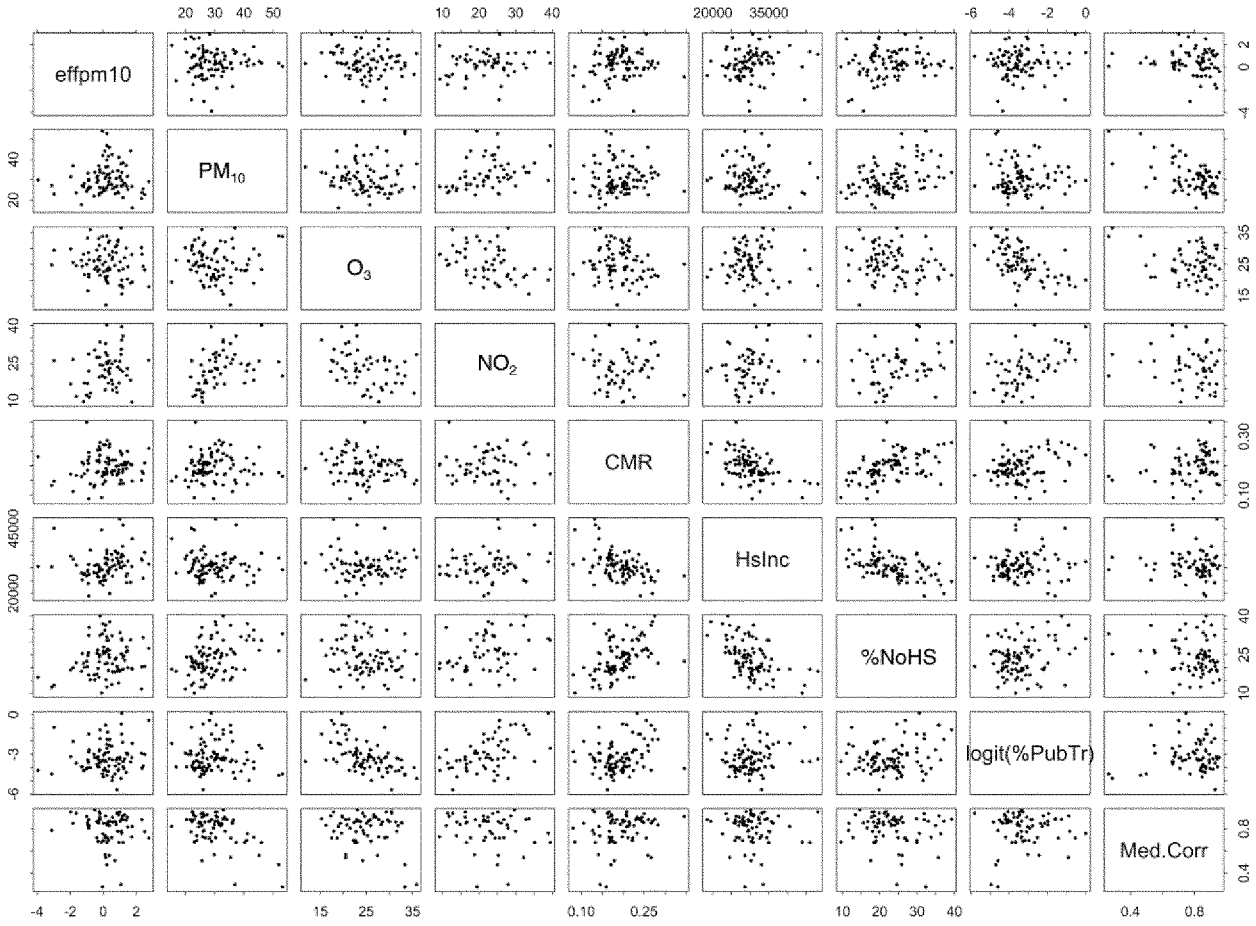


Figure 20. Pairwise scatterplots of city-specific coefficients for effect of PM_{10} (effpm10) with mean levels of PM_{10}, O_3, and NO_2, crude mortality rate (CMR), household income (HsInc), percentage of residents not graduating from high school (%NoHS), percentage using public transportation [logit (%PubTrans)] and median PM_{10} concentrations (Med.Corr) for the 90 cities. Numbers along the x-axis and y-axis represent the ranges of the variables.

Table 3. Point Estimates of Weighted Linear Regression Model of PM_{10} Log-Relative Rates for 90 Cities Against Sociodemographic Variables

Variable	Slope[a] (SD)	
HsInc/1000[b]	0.008	(0.015)
%NoHS[c]	0.011	(0.012)
%Unemp[d]	0.020	(0.038)
%Pov[e]	0.001	(0.017)

[a] The slope is the expected value of PM_{10} log-relative rate per 1 unit change of the level of the predictor variable.

[b] HsInc/1000 = Median income/1,000.

[c] %NoHS = % of adults without high school diploma.

[d] %Unemp = % unemployed.

[e] %Pov = % of households below poverty level.

Table 4 provides the best-fitting models by the C_p value according to the number of indicators in the model. We picked the best 1-variable, 2-variable, 3-variable and 4-variable models and the best model overall. Table 5 provides the results of these 5 models, each having the PM_{10} coefficient for total mortality and lag 1 as the outcome measure. The coefficients that are statistically significant in these 5 models are the mean levels of NO_2 in models 3 and 4 and the mean level of PM_{10} in model 5.

We next performed a spatial analysis of the PM_{10} coefficients by grouping the 90 counties into 7 geographical regions (Northwest, Upper Midwest, Industrial Midwest, Northeast, Southern California, Southwest, Southeast), following the stratification of the United States used in the 1996 *Review of the National Ambient Air Quality Standards for Particulate Matter* (EPA 1996a) (see Figure 1).

More specifically, we fit the 3-stage regression model by MCMC simulation with the index R denoting the 7 geographical regions. Figure 21 shows the maximum likelihood point estimates (open circles) and associated 95% CIs of the log-relative rates of mortality per 10 $\mu g/m^3$ increase in PM_{10} at lag 1 for each location. Estimates for the individual cities were made independently without borrowing information from the other cities. The 90 cities are grouped into the 7 regions in the figure. The bolded

Table 4. Best-Fitting Models as Determined by Smaller C_p Values with 1, 2, 3, 4, or 5 Predictors Chosen from Table 1

Number of Predictors	Predictor Variables	C_p
1	PM_{10}	10.8
	NO_2	11.5
	% Public transportation (%PubTrans)	11.9
	% No high school (%NoHS)	13.3
	Household income (HsInc)	13.4
	Median PM_{10} cross correlation (MCC)	13.6
	Crude mortality rate (CMR)	13.7
2	PM_{10}, %NoHS	8.4
	PM_{10}, NO_2	8.8
	PM_{10}, O_3	10.6
	PM_{10}, CMR	12.5
3	PM_{10}, O_3, NO_2	6.8
	PM_{10}, O_3, %NoHS	7.7
	PM_{10}, NO_2, %NoHS	8.4
	PM_{10}, HsInc, %NoHS	10.3
4	PM_{10}, O_3, NO_2, MCC	5.8
	PM_{10}, O_3, NO_2, %NoHS	6.4
	PM_{10}, O_3, %NoHS, %PubTrans	6.4
	PM_{10}, O_3, NO_2, %PubTrans	6.5
5	PM_{10}, O_3, NO_2, %NoHS, MCC	4.6
	PM_{10}, O_3, NO_2, %NoHS, %PubTrans	6.4
	PM_{10}, O_3, CMR, %NoHS, %PubTrans	7.2
	PM_{10}, O_3, NO_2, HsInc, %NoHS	8.1

Table 5. Results of Weighted Linear Regressions for Best Model (Smallest C_p) for 1 to 5 Predictor Variables Chosen from Table 4[a]

Model	PM_{10}	O_3	NO_2	%NoHS[b]	MCC[c]
1	0.009 (0.010)	—	—	—	—
2	−0.017 (0.011)	—	—	0.019 (0.012)	—
3	−0.021 (0.012)	0.020 (0.017)	0.025 (0.013)[d]	—	—
4	−0.025 (0.015)	0.020 (0.017)	0.026 (0.013)[d]	—	−0.33 (0.70)
5	−0.032 (0.016)[d]	0.023 (0.018)	0.024 (0.013)	0.059 (0.012)	−0.30 (0.70)

[a] Coefficient (SD).

[b] %NoHS = percentage without a high school diploma.

[c] MCC = median PM_{10} cross correlation.

[d] Coefficient is statistically significant ($P < 0.05$).

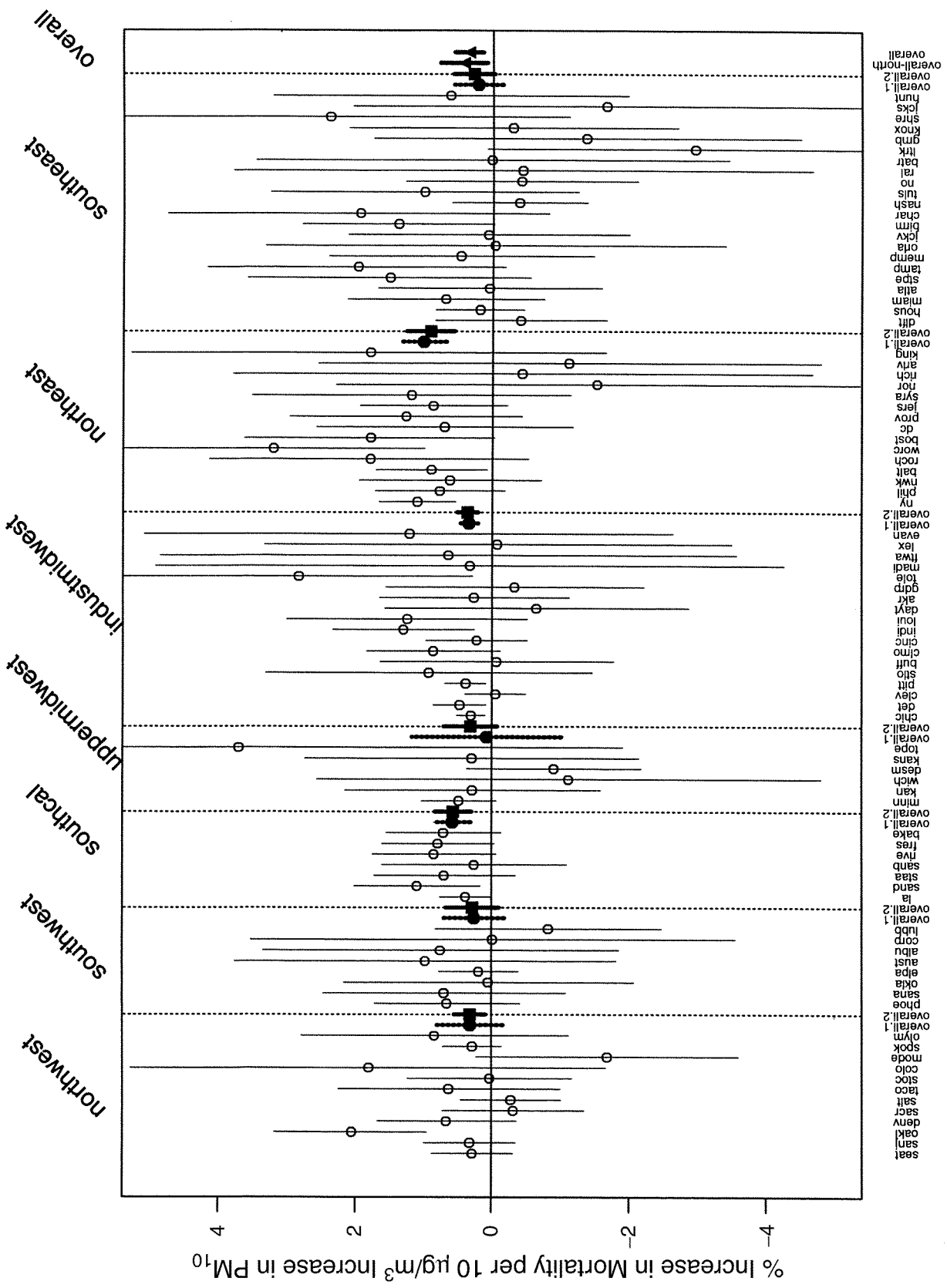


Figure 21. Effects of PM_{10} on total mortality and 95% CIs for each of the 90 cities, grouped by region.

segments represent the posterior means and 95% posterior intervals of the pooled regional effects under the more conservative prior A for the heterogeneity across both regions and cities within regions. The solid circle and the square denote the overall regional means without and with borrowing information from the other regions, respectively. At the extreme right, marked with triangles and bolded segments, are displayed the overall effects of PM_{10} for all cities, minus those in the Northeast, and the overall effect of PM_{10} for all cities. The effect of PM_{10} varied somewhat across the 7 regions (Figure 22). The effect of PM_{10} was estimated to be greatest in the Northeast, with a log-relative rate of 0.9 (95% CI, 0.58, 1.31). Appendix C explores the sensitivity of these findings to the prior assumption with regard to heterogeneity across cities and regions.

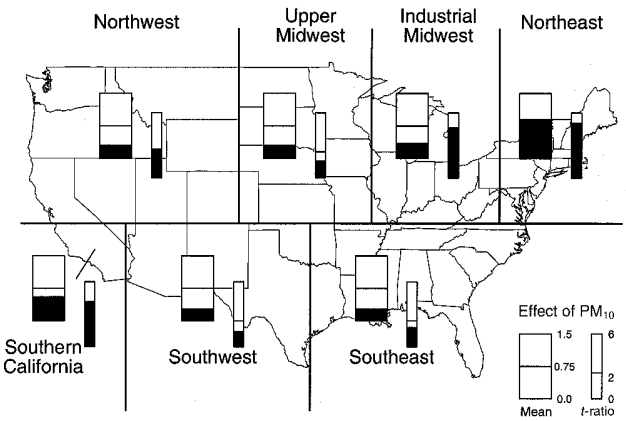


Figure 22. Posterior means/posterior standard error (*t* ratio) of regional effects of PM_{10} at lag 1.

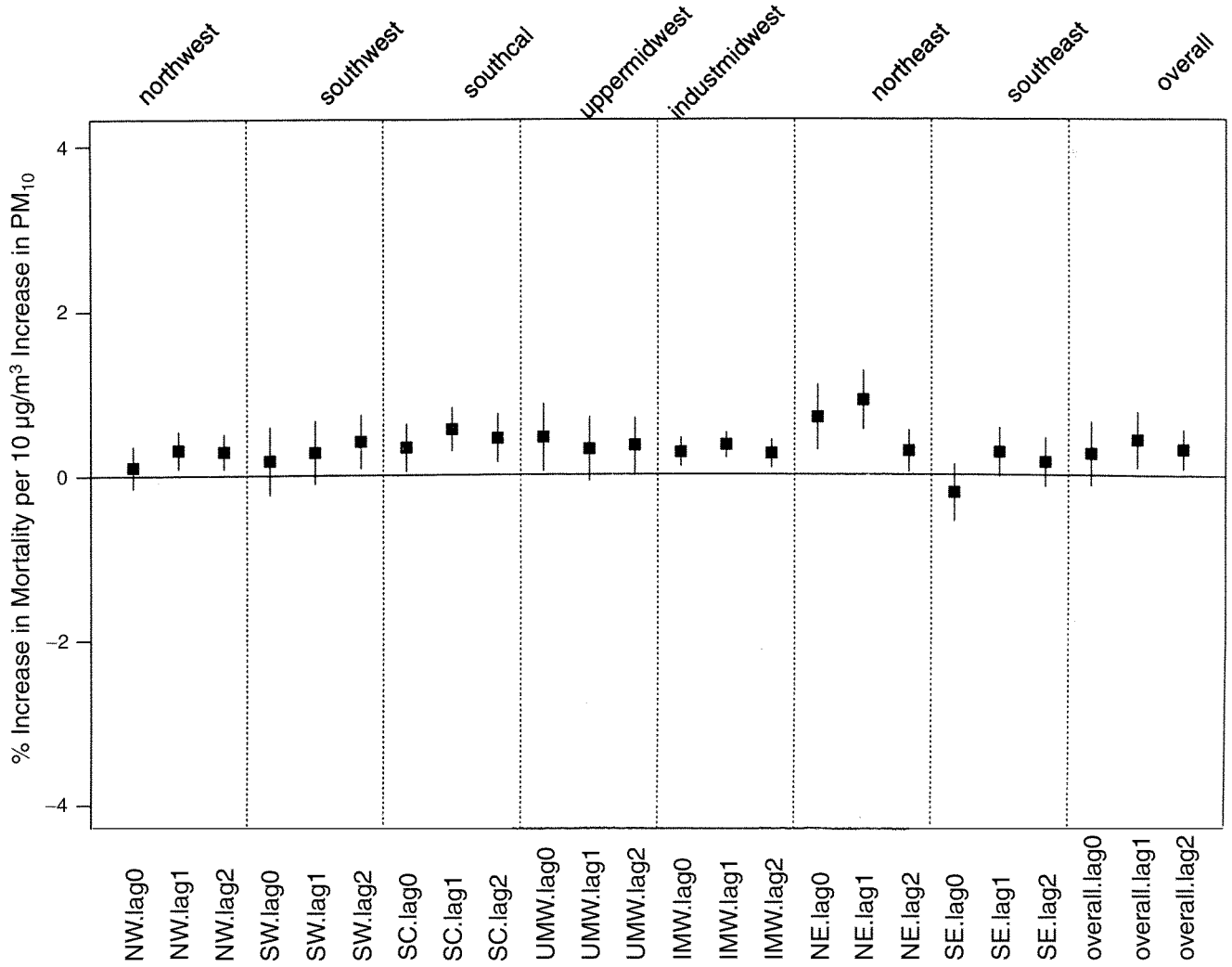


Figure 23. Posterior means and 95% posterior intervals of regional effects of PM_{10} at lags 0, 1, and 2 for the 90 cities. At the far right are the overall effects at lags 0, 1, and 2.

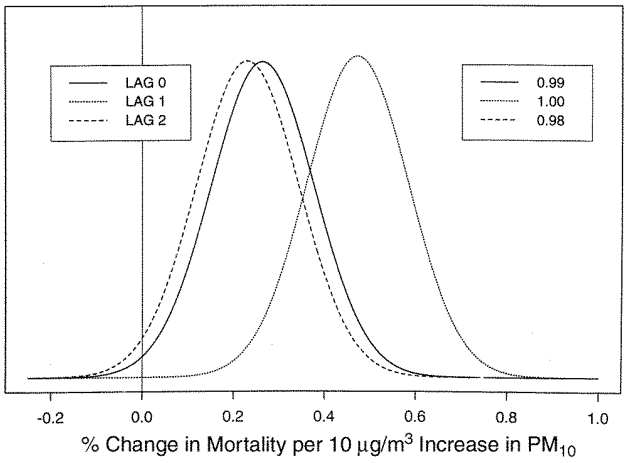


Figure 24. Marginal posterior distributions for effects of PM_{10} on all-cause mortality at lag 0, 1, and 2, without control for other pollutants, for the 90 cities. The box at the top right provides the posterior probabilities that the overall effects are greater than 0.

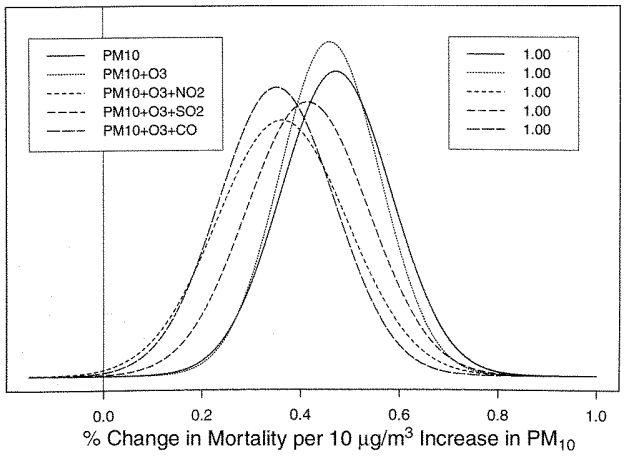


Figure 25. Marginal posterior distributions for effects of PM_{10} on total mortality at lag 1 with and without control for other polltants, for the 90 cities. The box at the top right provides the posterior probabilities that the overall effects are greater than 0.

Results for Other Pollutants

We further explored whether the regional effects of PM_{10} were sensitive to the lag specification and to adjustment of other pollutants. Figures 23 and 24 show Bayesian regional and overall estimates and 95% credible regions of the PM_{10} effects on total mortality at lags 0, 1, and 2, respectively. As for the 20 cities, the evidence suggests a positive effect and a small variation among the three lags. Figure 25 shows the posterior distribution of PM_{10} effect on total mortality for the 90 cities at lag 1 as well as the effect of PM_{10} adjusted for other pollutants. As for the 20 cities, the effect of PM_{10} changed little with control for the other pollutants.

Figures 26 through 29 provide the findings of models for the other pollutants (O_3, SO_2, NO_2, and CO) in 20 cities. For total mortality, each figure presents the regression coefficients and 95% CIs at lags 0, 1, and 2 for univariate, bivariate (adding PM_{10}), and trivariate (adding PM_{10} and 1 of the remaining pollutants). For O_3, there was little evidence of an effect except at lag 2, although the univariate estimate was negative at lag 2. For SO_2 (Figure 27), the univariate models had the highest coefficients and adjustment for other pollutants tended to reduce the effect, suggesting confounding. There was no consistent pattern of association for NO_2 (Figure 28). For CO, effects were positive and generally significant (Figure 29). Adjustment for other pollutants tended to reduce the effect.

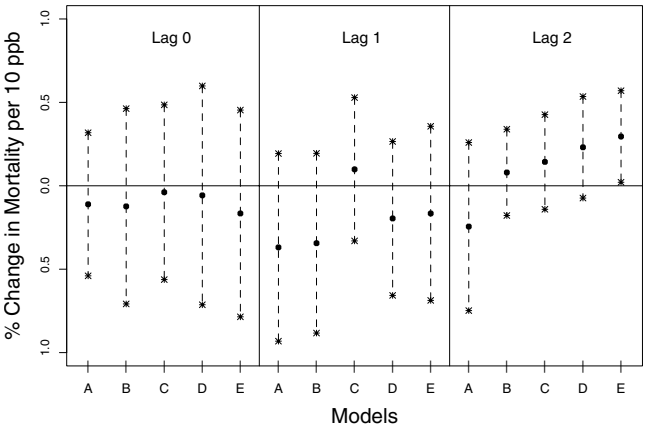


Figure 26. Weighted mean effects of O_3 on total mortality at lags 0, 1, and 2 in the 20 cities. Models A = O_3 alone; B = O_3 + PM_{10}; C = O_3 + PM_{10} + NO_2; D = O_3 + PM_{10} + SO_2; E = O_3 + PM_{10} + CO.

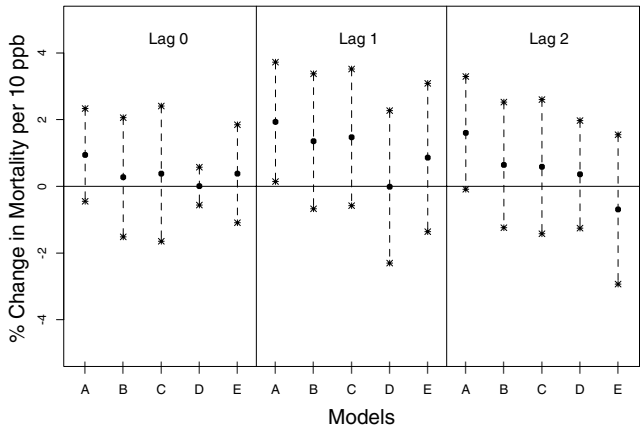


Figure 27. Weighted mean effects of SO_2 on total mortality at lags 0, 1, and 2 in the 20 cities. Models A = SO_2 alone; B = SO_2 + PM_{10}; C = SO_2 + PM_{10} + O_3; D = SO_2 + PM_{10} + NO_2; E = SO_2 + PM_{10} + CO.

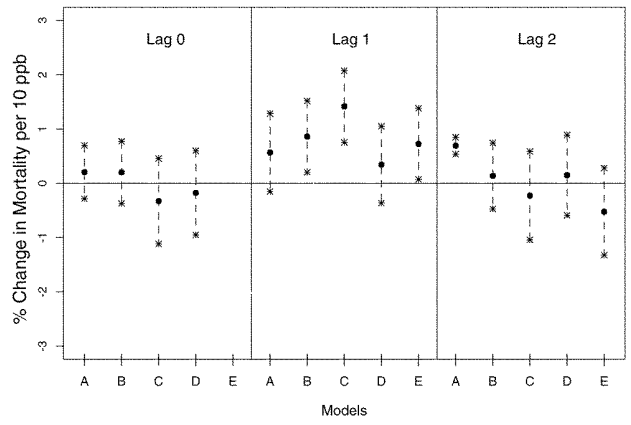


Figure 28. Weighted mean effects of NO_2 on total mortality at lags 0, 1, and 2 in the 20 cities. Models A = NO_2 alone; B = NO_2 + PM_{10}; C = NO_2 + PM_{10} + O_3; D = NO_2 + PM_{10} + SO_2; E = NO_2 + PM_{10} + CO.

The effect of O_3 was examined across the whole year and by season because of the generally higher levels during the summer. Overall, for the full year, the distributions for the effects of O_3 are centered on 0 and there is only an even chance that the effect is larger than 0 for both total and cardiorespiratory deaths. Because O_3 levels vary widely by season, we compared the effects of O_3 during the 3 hottest months (June, July, and August) and winter months (December, January, and February). With stratification, there was reasonably strong evidence of an effect of O_3 during the summer (posterior mean = 0.41 ppb; 95% CI [−0.20, 1.01]), but not during the winter (posterior mean = −1.86 ppb, 95% CI [−2.70, −0.96]) (Figure 30).

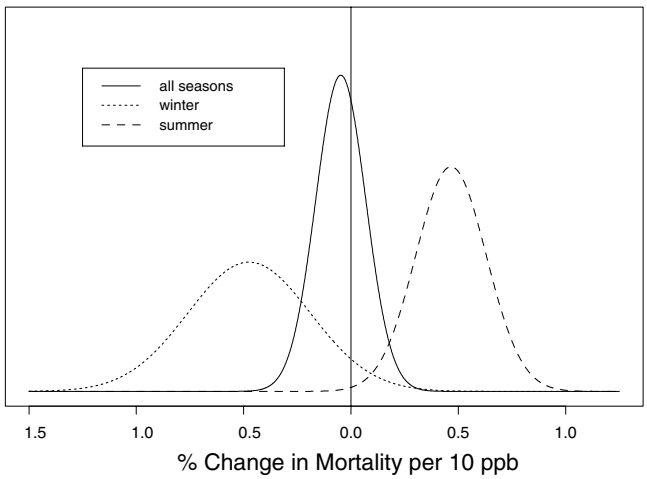


Figure 30. Marginal posterior distributions of the overall effects of O_3 at lag 0 for all seasons, Summer (June, July, August) and Winter (December, January, February) for the 90 cities.

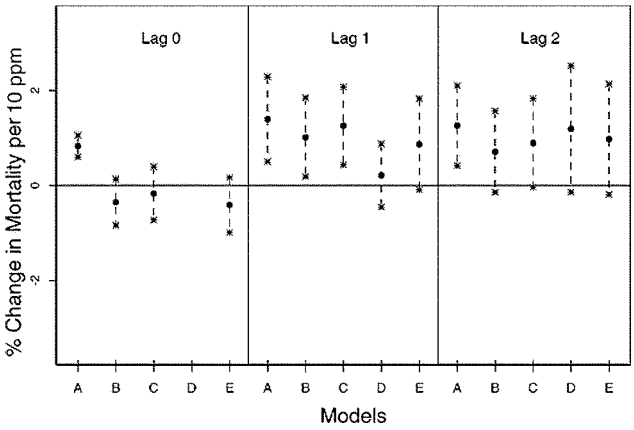


Figure 29. Weighted mean effects of CO on total mortality at lags 0, 1, and 2 in the 20 cities. Models A = CO alone; B = CO + PM_{10}; C = CO + PM_{10} + O_3; D = CO + PM_{10} + NO_2; E = CO + PM_{10} + SO_2. The weighted mean effect of CO on total mortality at lag 0 under model D is not reported because the confidence interval was too wide to be considered informative.

HOSPITALIZATION OF THE ELDERLY

In the last decade, many studies have assessed the association of air pollution with daily numbers of hospitalizations and emergency room admissions (reviewed by Pope and Dockery 1999). Although positive findings were reported for most of the studies, the study methods differed substantially. As for the NMMAPS mortality analysis, we applied a systematic 2-stage approach to data from 14 cities selected from the sampling frame afforded by the AIRS database. We used the distributed lag models (Appendix B) for this analysis; to gain the most precise estimates possible with these models, we restricted the analysis to those cities with daily PM_{10} measurements.

The distributed lag models differ conceptually from the ad hoc approaches based on best fit in individual cities, which have been generally used. These distributed lag models assume that the effect of PM_{10} on hospital admissions may be distributed over several days. To test this assumption, we included PM_{10} concentrations on the same and several prior days in the model to estimate the effect on each lag simultaneously. Estimates based on such approaches can be subject to substantial variability due to stochastic error. The multicity approach reduces stochastic variability by combining information from different locations. Most past studies have also used simple moving averages of pollution to assess whether the effect of air pollution on health persists for more than 1 day following exposure. The effect of air pollution might plausibly, however, diminish gradually over several days or be initially delayed, resulting in an estimated effect that is larger for air pollution exposures on the prior day than on the day of

detected response. In modeling time-series data, it is possible to include air pollution values on multiple days to estimate directly the effects at different lags, but this approach is limited in analyses of data from single cities because multicollinearity makes the estimated effects at different lags very imprecise. Although imprecise, the estimates are unbiased; hence a multiple-city analysis, which can average out the noise, makes the distributed lag approach feasible. We have used the distributed lag model to estimate the association between PM_{10} and hospital admission for heart and lung disease, including the distribution of effects over time.

METHODS

Cities

Cities were sorted by the number of days with PM_{10} observations for the period 1985 through 1994. The goal was to select cities with extended daily PM_{10} measurements for the period 1985 through 1994 (years of available hospital admission data) from cities distributed geographically across the United States. Eighteen cities met the following criteria for inclusion:

- At least 1,460 days (4 years) with PM_{10} measurements between 1985 and 1994.
- Daily PM_{10} measurements on at least 50% of days between the city-specific start and end of measurements.

To maintain geographic diversity, no more than 2 cities were included in each state. Cincinnati was excluded based on this criterion. Denver was excluded because for most of the years it did not have daily data, making it difficult to fit distributed lag models. In addition, in order to focus on combustion-related particles, Albuquerque (1,751 days of observations) was excluded because of the influence of wind-blown dust.

We chose the metropolitan county containing each city except for Minneapolis and St Paul, which were combined and analyzed as a single city, and Birmingham, which included some suburban counties to increase the population (Table 6). Thus, 14 cities were selected for inclusion: Birmingham, Boulder, Canton, Chicago, Colorado Springs, Detroit, Minneapolis/St Paul, Nashville, New Haven, Pittsburgh, Provo/Orem, Seattle, Spokane, and Youngstown.

Hospital Admissions Data

Daily counts of hospital admissions were extracted from HCFA using Medicare billing records. Data were requested for respiratory and cardiovascular admissions for the entire United States. These data, more than 75 million records, were received on 264 IBM 9-track tapes. Most of the data were received in Medical Provider Analysis and

Table 6. Cities with Daily Measurements of PM_{10} Included in Analysis, Counties Used to Define Cities, and Demographic Characteristics of Cities Based on 1990 Census

City	County	Population	> 65 Years (%)	College Educated (%)	Unemployed (%)	Poverty (%)	Nonwhite (%)
Birmingham	Blount, Jefferson, Shelby, St Clair, Walker	907,810	14.0	19.9	6.5	16.0	36.0
Boulder	Boulder	225,339	7.6	17.9	2.7	10.5	6.5
Canton	Stark	367,585	14.4	14.3	7.2	11.1	8.0
Chicago	Cook	5,105,067	12.4	22.8	8.0	14.2	37.0
Colorado Springs	El Paso	397,014	8.0	25.8	7.3	10.4	14.0
Detroit	Wayne	2,111,687	12.5	13.7	12.4	20.1	43.0
Minneapolis/St Paul	Hennepin, Ramsey	1,518,196	12.2	30.7	4.8	9.9	11.0
Nashville	Davidson	510,784	11.6	11.6	2.7	12.4	25.2
New Haven	New Haven	804,219	14.7	24.2	5.8	7.9	14.0
Pittsburgh	Allegheny	1,336,449	17.4	22.6	6.3	11.5	12.0
Provo/Orem	Utah	263,590	7.0	8.7	2.2	14.8	3.6
Seattle	King	1,507,319	11.1	32.8	4.1	8.0	15.0
Spokane	Spokane	361,364	13.3	20.6	7.3	13.7	5.0
Youngstown	Columbiana, Mahoning	373,082	16.4	5.9	3.9	15.7	12.2

Table 7. Mean Number of Admissions Per Person 65 Years of Age or Older for Cardiovascular Disease (CVD), Chronic Obstructive Pulmonary Disease (COPD), and Pneumonia[a]

City	Population 65+ yr	Mean Medicare Admissions per Day			Admissions per 10,000 Person-Years		
		CVD	COPD	Pneumonia	CVD	COPD	Pneumonia
Birmingham	119,400	17.28	1.63	5.33	528	50	163
Boulder	17,100	1.98	0.27	0.60	423	58	128
Canton	52,900	9.39	0.33	2.39	648	23	165
Chicago	633,000	102.10	7.82	26.67	589	45	154
Colorado Springs	31,800	2.92	0.42	0.98	335	48	112
Detroit	264,000	50.03	4.18	10.80	692	58	149
Minneapolis/St Paul	175,900	16.57	1.74	5.09	344	36	106
Nashville	59,300	9.23	1.02	2.79	568	63	172
New Haven	118,200	16.18	0.98	4.07	500	30	126
Pittsburgh	232,500	47.73	5.75	10.39	749	90	163
Provo/Orem	18,500	2.30	0.14	0.70	454	28	138
Seattle	167,300	16.81	1.68	4.42	367	37	96
Spokane	48,100	5.81	0.75	1.79	441	57	136
Youngstown	61,150	11.69	1.25	2.69	698	75	161

[a] Estimated admissions per 10,000 person-years based on 65+ population.

Review (MEDPAR) format—that is, 1 record for each admission. Data for 2 years, 1992 and 1993, were received as Standard Analytic Files (SAF), with 1 record for each billing. In the SAF format, a single admission may generate multiple records. The multiple SAF records were therefore combined into a single record for each admission.

For each county, day-specific counts of hospital admissions by diagnostic categories were calculated for cardiovascular disease (CVD) (ICD-9, 390–429); COPD (ICD-9, 490–492, 494–496); and pneumonia (ICD-9, 480–487), in persons aged 65 years and older.

Table 7 presents the average number of HCFA respiratory and cardiovascular hospital admissions per day for NMMAPS morbidity cities. For comparison across communities, admissions per 10,000 person-years were calculated.

We excluded days when the hospital admissions showed outliers in daily counts, defined as more than 4 times the interquartile range above the median for pneumonia. For COPD, the outliers were defined as the values that were 3 times the interquartile range above the median, or at least 100% higher than the mean of the nearby data. These errors can occur for clerical reasons; for example, records without the date of admission are coded to the first day of the month or year. Alternatively they may represent epidemics, of influenza for example. This exclusion eliminated a total of 2 days of data for CVD in all of the 14 cities, 44 days of data for pneumonia, and 13 days of data for COPD.

Table 8. Weather Stations and Mean Temperature, Relative Humidity, and Barometric Pressure by City

City	Weather Station	Weather		
		Temperature (°F)	Relative Humidity (%)	Barometic Pressure (inches H_2O)
Birmingham	Birmingham Airport	63	71	29.4
Boulder	Denver Stapleton Airport	52	53	24.7
Canton	Akron-Canton Airport	50	74	28.7
Chicago	Chicago O'Hare Airport	50	71	29.3
Colorado Springs	Colorado Springs Municipal Airport	49	52	24.0
Detroit	Detroit Metropolitan Airport	51	71	29.3
Minneapolis/St Paul	Minneapolis–St Paul Airport	47	68	29.1
Nashville	Nashville Metropolitan Airport	56	65	30.0
New Haven	Hartford Bradley Airport	52	67	29.8
Pittsburgh	Pittsburgh	52	70	28.8
Provo/Orem	Salt Lake City Airport	53	55	25.8
Seattle	Seattle-Tacoma Airport	53	75	29.6
Spokane	Spokane Airport	48	67	27.5
Youngstown	Youngstown Municipal Airport	50	74	29.0

Weather Data

Daily mean meteorologic measurements (temperature, barometric pressure, and relative humidity) were obtained from the NCDC Surface Airways CD (EarthInfo). The nearest National Weather Service station was selected for each city (Table 8).

Air Pollution Data

As for the mortality analyses, air pollution data for the years 1985 through 1994 were obtained from the EPA's AIRS database (Nehls and Akland 1973). We selected cities that had an extended period of daily PM_{10} measurements between 1984 and 1995 (Table 9). Many of the cities have

Table 9. PM_{10} Monitoring in 14 Cities: Start and End Dates, Number of Days and Percentage of Days with PM_{10} Samples, Mean and Maximum PM_{10}[a]

City	Number of Sites	Start Date	End Date	Number of Days	Number of Observations	% Observations	Mean ($\mu g/m^3$)	Max ($\mu g/m^3$)
Birmingham	7	1-Apr-87	31-Dec-93	2,467	2,417	98	34.8	124.8
Boulder	2	1-May-89	24-Dec-94	2,064	1,879	91	24.4	125.0
Canton	2	1-Jan-89	24-Dec-94	2,189	1,642	75	28.4	94.8
Chicago	6	1-Mar-88	24-Dec-94	2,547	2,354	92	36.4	144.7
Colorado Springs	4	1-Jul-87	24-Dec-94	2,734	2,427	89	26.9	147.2
Detroit	4	1-May-86	24-Dec-94	3,159	2,764	87	36.8	133.6
Minneapolis/St Paul	4	1-Apr-87	24-Dec-94	2,824	2,672	95	27.4	141.5
Nashville	6	1-Sep-89	24-Dec-94	1,941	1,588	82	31.6	128.0
New Haven	2	1-May-87	31-Dec-91	1,707	1,475	86	29.3	95.4
Pittsburgh	6	1-Jan-87	24-Dec-94	2,915	2,891	99	36.0	139.3
Provo/Orem	3	1-Apr-87	24-Dec-94	2,825	2,682	95	38.9	241.0
Seattle	4	1-Jan-86	24-Dec-94	3,280	3,195	97	31.0	145.9
Spokane	4	1-Oct-85	24-Dec-94	3,372	2,778	82	45.3	605.8
Youngstown	2	1-Jan-89	31-Dec-92	1,461	1,215	83	33.1	104.0

[a] Days above 150 $\mu g/m^3$ were excluded from analysis.

more than 1 monitoring location. Monitors were excluded if they had fewer than 300 observations within the period. To ensure that our exposure measure best represented general population exposure and not unique local conditions, monitors within the lowest 10th percentile of the correlation among monitors across all counties were excluded. Data from both population-oriented monitors and monitors sited for other purposes were reviewed. A complete list of the air monitors and the number of PM_{10} observations for each monitor is included in Appendix D (see Table D.1). Even with the selection of cities with a high frequency of PM_{10} sampling, some monitors within a selected county measured PM_{10} only during 1 day in 6. Concentrations measured by different monitors have different means and standard deviations. Data on gaseous air pollutants (SO_2, NO_2, CO and O_3) were also downloaded.

In calculating a daily PM_{10} value, we needed a scheme that did not introduce changes in exposure estimates from day to day because of the particular monitors contributing data. The annual mean was computed for each monitor for each year. This monitor-and-year-specific mean was subtracted from the daily measurements for that monitor, and the difference was divided by the monitor-specific standard deviation to produce a daily standardized deviation. The standardized deviations for all reporting monitors were averaged for each day. County-specific daily mean was calculated as the average standardized deviation multiplied by the standard deviation of all the centered measurements for the year and added to the annual average of all the monitors. We excluded days when air pollution exceeded the ambient air quality standard of 150 $\mu g/m^3$ PM_{10} in order to study the association at common concentrations and limit the influence of outlier days.

A data set of the city-specific 24-hour mean concentrations for criteria pollutants (PM_{10}, SO_2, and NO_2) and 1-hour maximum concentrations for O_3 and CO was created using the above method. For the continuously measured gaseous pollutants, monitor-specific mean and maximum values were calculated for each day having at least 18 hours of observations. An 8-hour (10 am to 5 pm) mean O_3 concentration was also calculated for these days.

County-specific daily means for each pollutant were calculated, adjusting for the expected value and variance of any missing monitors on a given day. Table 10 presents the city-specific averages of the daily mean concentrations for PM_{10}, SO_2, and NO_2 and the 1-hour maximum concentrations for O_3 and CO. As indicated by blanks in the table, not all criteria pollutants are measured in all of these communities.

Table 10. Average Concentrations of PM_{10}, SO_2, NO_2, O_3 and CO) in 14 Cities

	24-Hr Mean			1-Hr Max	
City	PM_{10} ($\mu g/m^3$)	SO_2 ($\mu g/m^3$)	NO_2 (ppb)	O_3 (ppb)	CO (ppm)
Birmingham	34.8	19.9	11.0	52.0	2.88
Boulder	24.4	—[a]	—	48.0	2.50
Canton	28.4	27.7	—	56.0	1.25
Chicago	36.4	14.7	26.0	39.0	2.02
Colorado Springs	26.9	7.5	—	44.0	3.81
Detroit	36.8	22.1	22.0	45.0	1.90
Minneapolis/ St Paul	27.4	9.8	20.0	—	3.11
Nashville	31.6	31.9	13.0	37.0	2.57
New Haven	29.3	28.6	28.0	54.0	3.39
Pittsburgh	36.0	45.6	27.0	44.0	2.36
Provo/Orem	38.9	—	24.0	60.0	4.14
Seattle	31.0	8.8	—	39.0	3.79
Spokane	45.3	8.8	—	45.0	5.37
Youngstown	33.1	39.7	18.0	52.0	0.99

[a] — = No measurement (or very limited measurement) taken.

Analytic Methods

In each city, the associations between hospital admissions and PM_{10} were investigated with a generalized additive robust Poisson regression model (Hastie and Tibshirani 1990), which we introduced to the study of Poisson time series (Schwartz 1993, 1994). In the generalized additive model, the outcome is assumed to depend on a sum of nonparametric smooth functions for each variable. This allows us to model the potential nonlinear dependence of daily admissions on weather and season better.

The model is of the form:

$$\log[E(Y_t)] = \alpha_0 + S_1(X_1) + \ldots + S_p(X_p) \quad (1)$$

where $E(Y_t)$ is the expected value of the daily count of admissions Y_t, and S_i are the smooth functions of the covariates X_i. The locally weighted smoother (LOESS) (Cleveland and Devlin 1988) was chosen as the smooth function. All nonparametric smoothing functions are characterized by a smoothing parameter, which determines the smoothness of the fit. LOESS is a generalization of a weighted moving average and estimates a smooth function by fitting a weighted regression within a moving window (or fraction of the data) centered about each value of the predictor variable. The weights are close to 1 for the central third of the window, and decline to 0 rapidly outside that range. Outside

of the window, the weights are all 0. A window of 200 days corresponds to a moving average of about 80 days. To control for weather variables and day of the week, we choose the smoothing parameter that minimized the Akaike Information Criterion (AIC) (Akaike 1973).

The control of seasonal patterns requires investigation of the residuals to ensure that seasonal patterns have been removed and to minimize autocorrelation of residuals. If omitted confounding variables have subseasonal patterns that are correlated with air pollution, removing those patterns may reduce confounding. Excessive filtering, however, reduces the power to find any association because air pollution typically varies with multiple-day patterns. In addition, it is well known in the filtering literature that overfiltering can produce high frequency ringing in the data that induces autocorrelation (Rabiner and Gold 1965) and may distort the relationship between air pollution and morbidity or mortality. To make this tradeoff, we chose the window size that minimized the autocorrelation of the residuals. This approach was used because each admission is an independent event, and autocorrelation in residuals indicates there are omitted, time-dependent covariates the variation of which may confound the effect of air pollution. If the autocorrelation is removed, remaining variation in omitted covariates has no systematic temporal pattern, and hence confounding is less likely (Rabiner and Gold 1965). We chose smoothing parameters in each city that met both objectives. In some cases, it was necessary to use autoregressive terms to eliminate serial correlation (Brumback et al 1999). This approach has been used in previous studies (Schwartz 1999).

We built a model for each city in order to allow for city-specific differences. The variables included in each model were season, weather variables (temperature, relative humidity, and barometric pressure) and day of the week. As weather and season vary across the cities, however, the smoothing parameter for each variable was optimized separately in each location.

We treated PM_{10} as a linear term in the analysis to assess how its effects were distributed over different lags and to allow the use of meta-analytic techniques to combine results across cities.

Some have argued that there are thresholds for the effects of air pollution and that no adverse responses occur on most days. To test this hypothesis, we repeated our analysis restricted to those days when PM_{10} was less than 50 $\mu g/m^3$, which is one third of the current National Ambient Air Quality Standard (NAAQS).

Distributed Lag Models

Distributed lag models were introduced by Almon (1965) and have been applied mainly in econometrics and social sciences. The models allow us to examine the possibility that air pollution can influence hospital admissions not only on the same day, but also on subsequent days.

Let $(Y_1, Z_1),\ldots,(Y_t, Z_t)$ denote a regression data set that is ordered with respect to time, where Z_t represents a daily pollution measure and Y_t represents daily counts of hospital admissions or of mortality. The unconstrained distributed lag model of order q is:

$$Y_t = \alpha + \beta_0 Z_t + \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q} + \varepsilon_t \qquad (2)$$

where ε_t are independent random variables with mean zero and constant variance. Hence, the outcome Y_t at time t may depend on the exposure (Z_t) measured not only on the current day but also on previous days. The overall impact of a unit change in exposure on 1 day is the sum of its impact on that day and its impacts on subsequent days—that is, $\beta_0 + \beta_1 + \ldots + \beta_q$. The problem is that Z_t is correlated with $Z_{t-1}, \ldots, Z_{t-q,}$ and the high degree of collinearity will result in unstable estimates of the β_j. Both the β_j and the sum of β_j, however, will be unbiased estimators of the effects at each lag and of the overall effects. Because the estimators are unbiased, combining results across cities will produce more stable unbiased estimates.

A 1-day exposure model can be seen as a constrained model, where $\beta_j = 0$ for $j = 1 \ldots q$. If we have no strong biological basis for that constraint, it is preferable to let the data inform us about the pattern of effect in time. The 1-day model may be an unreasonably strong constraint, which risks introducing bias. A more flexible constraint may reduce the variance of the individual β_j with less risk of bias. One common approach is to constrain the β_j to follow a flexible polynomial (Judge et al 1985; Pindyck and Rubinfeld 1998).

The polynomial distributed lag function constrains β_j:

$$\beta_j = \sum_{k=0}^{d} c_k j^k \qquad (3)$$

where d is the degree of the polynomial. For $d = 2$, we restrict the β_j to be a quadratic function of j:

$$\beta_j = c_0 + c_1 j + c_2 j^2 \qquad (4)$$

Rewriting the distributed lag in the context of a general additive model, including the covariates that can be fit as linear or smoothing function and applying a quadratic polynomial constraint for the β_j, for $j = 1 \ldots q$, we have:

$$\log(\mathrm{E}[Y_t]) = \alpha + \textit{covariates} + c_0 Z_t + (c_0 + c_1 + c_2)Z_{t-1} + (c_0 + 2c_1 + 2^2 c_2)Z_{t-1} + \ldots + (c_0 + qc_1 + q^2 c_2)Z_{t-q} \quad (5)$$

If we factor out each c_i, for $i = 0, 1, 2$,

$$\log(\mathrm{E}[Y_t]) = \alpha + \textit{covariates} + c_0(Z_t + Z_{t-1} + \ldots + Z_{t-q}) + c_1(Z_{t-1} + 2Z_{t-2} + \ldots + qZ_{t-q}) + c_2(Z_{t-1} + 2^2 Z_{t-2} + \ldots + q^2 Z_{t-q}), \quad (6)$$

which can be written as

$$\log(\mathrm{E}[Y_t]) = \alpha + \textit{covariates} + c_0 W_0 + c_1 W_1 + c_2 W_2. \quad (7)$$

The parameters c_i can be estimated from the regression in equation (7), and we can obtain β_j from equation (4).

The vector of β_j of dimension $(q \times 1)$, given by equation (4), can be expressed in matrix notation as:

$$\beta = \mathbf{a} * c \quad (8)$$

where c is the $([d + 1] \times 1)$ vector of the estimated coefficients from the above regression, $\mathbf{a}$ is the $(q \times [d + 1])$ matrix built from the index j.

Therefore the variance-covariance matrix of the vector β is

$$\mathrm{cov}(\beta) = \mathrm{var}(\mathbf{a} * c) = \mathbf{a}^t \,\mathrm{cov}(c)\, \mathbf{a} \quad (9)$$

and the standard error (se) is given by

$$\mathrm{se}(\beta) = \mathrm{sqrt}(\mathrm{diag}(\mathrm{cov}(\beta)) \quad (10)$$

In the distributed lag model, the overall effect of a unit increase in pollution is given by the sum of the effects on each day considered, that is $\beta_0 + \beta_1 + \ldots + \beta_q$; therefore, the standard error of the $\Sigma(\beta) = \beta^t v$, where v is a $(q \times 1)$ vector of 1 given by

$$\mathrm{sqrt}(\mathrm{var}(\Sigma\,\beta)) = \mathrm{sqrt}(\mathrm{var}(\beta^t v)) = \mathrm{sqrt}(v^t(\mathrm{cov}\beta)v) = \mathrm{sqrt}(v^t(\mathbf{a}^t\mathrm{cov}(c)\mathbf{a})v). \quad (11)$$

We have used the unconstrained model as our primary approach, relying on our ability to combine results across cities to cancel out noise and provide reasonable estimates. We have also used quadratic distributed lag models as a sensitivity analysis, however. To compare our results with those of the types of models that have been fit previously, we have used PM_{10} on the same day, the previous day, and also the mean of the same and previous day as exposure variables. We also repeated the analyses restricted to days when PM_{10} was less than 50 $\mu g/m^3$.

Second-Stage Assessment of Confounding and Effect Modification

In the second stage of the analysis, we used 2 different approaches: the inverse variance weighted averages to combine results across cities and a meta-regression to analyze confounding. The first is equivalent to a meta-regression with only an intercept.

The inverse variance weighted averages were computed for both the estimated overall effect (the sum of β_j) and for the effect of each lag. In the weighted average, the pooled summary of the effect is given by a weighted sum of the results divided by the sum of the weights, where the weights are the inverse of the variance of the effect. More formally, we assume that the effect of PM_{10} is $\hat{\beta}_i^* \sim N\,\mathrm{I}\,(\mu, V_i)$ and we estimate μ from the 14 city-specific β_i and their variances by computing an inverse variance weighted average. A combined random effects estimate was also computed, assuming

$$\hat{\beta}_i \sim \mathrm{N}(\beta, V_i)$$

and that the true city-specific effect is

$$\beta_i \sim \mathrm{N}(\beta, \Omega).$$

Therefore,

$$\hat{\beta}_i \sim \mathrm{N}(\beta, V_i + \Omega).$$

The between-city variance, Ω, is estimated as

$$\hat{\Omega} = \max(0, \mathrm{var}(\beta) - \bar{\hat{V}_i}).$$

The random effects estimate of the combined effect is then the weighted average with weights of $1/(V_i + \Omega)$.

We estimated the random component of the variance between the city-specific effect estimates as the total variance in β minus the within-city variance given the mean of the city-specific variances. A chi-squared test of heterogeneity between city-specific effect estimates was calculated following the methods of DerSimonian and Laird (1986).

To examine effect modification by socioeconomic status, for example, we can fit a weighted least-squares regression:

$$\hat{\beta}^*_i = \beta^* + \delta P_i + \varepsilon_i \quad (12)$$

where $\hat{\beta}^*_i$ is the estimated PM_{10} effect in city i, P_i is the city-specific socioeconomic status measure, and again the city-specific estimates are weighted by the inverse variance. The coefficient δ estimates change in the effect of PM_{10} for a unit increase in the poverty rate. We examined poverty rate and percentage of the population that was nonwhite as potential modifiers of the effect of PM_{10} on hospital admissions of the elderly using 1990 Census data for the cities included.

Confounding is usually examined by including potential confounders in the equivalent of the first stage of a hierarchical regression model. Since weather patterns tend to increase or decrease all pollutants in parallel, however, this approach risks substantial collinearity problems. Although levels of most pollutants typically vary together, the quantitative associations between pollutants vary substantially across locations. We have used this variation as a basis for examining confounding in the second stage of our analysis.

To illustrate this approach, suppose the true association between our outcome (4) and pollutant X_1 is:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon_t. \quad (13)$$

Assume X_1 is correlated with another pollutant X_2, which itself is not causal for Y. Quantifying the association between them as

$$X_1 = \gamma_0 + \gamma_1 X_2 + \varepsilon_t \quad (14)$$

and substituting (14) in (13), it follows that

$$Y = \beta_0 + \beta_1\gamma_0 + \beta_1\gamma_1 X_2 + \varepsilon_t, \quad (15)$$

and we see that the induced coefficient for the noncausal variable X_2 depends on γ_1, which is the slope of the relationship between X_1 and X_2. Thus, it is appropriate to extend our meta-regression approach to use the slope between pollutants as an explanatory factor in the second-stage model. That is,

$$\beta^*_i = \beta^* + \delta\gamma_i + \varepsilon_i \quad (16)$$

where γ_i is the slope between two pollutants, for example, SO_2 and PM_{10}. The parameter β^*, the intercept term in this regression, is the estimated effect of PM_{10} in a city where PM_{10} has no correlation with SO_2. This is the unconfounded estimated effect of PM_{10}.

Note that in NMMAPS mortality analysis, confounding by copollutants is addressed in the first stage of the model. In the morbidity analysis, we examined confounding in the second stage. In contrast, effect modification by the mean level of PM_{10} is not examined in the second stage in the morbidity analysis but in the first stage by restricting the analysis to days with less than 50 $\mu g/m^3$ PM_{10}. These differences will strengthen any common conclusions of the parallel analysis.

Table 11. Quartiles of HCFA Hospital Admissions of Persons 65 Years of Age or Older for Cardiovascular Disease (CVD), Chronic Obstructive Pulmonary Disease (COPD), and Pneumonia

	CVD			COPD			Pneumonia		
City	25%	50%	75%	25%	50%	75%	25%	50%	75%
Birmingham	14	17	21	1	1	2	3	5	7
Boulder	1	2	3	0	0	0	0	0	1
Canton	7	9	12	0	1	2	1	2	3
Chicago	86	103	117	4	7	11	20	25	31
Colorado Springs	2	3	4	0	0	1	0	1	2
Detroit	41	50	59	2	4	6	7	10	13
Minneapolis/St Paul	13	16	20	1	1	3	3	5	7
Nashville	7	9	12	0	1	2	1	2	4
New Haven	12	16	20	0	1	1	2	4	5
Pittsburgh	38	48	56	3	5	8	7	10	13
Provo/Orem	1	2	3	0	0	0	0	0	1
Seattle	13	17	20	1	1	2	3	4	6
Spokane	4	6	7	0	1	1	1	1	3
Youngstown	9	11	14	0	1	2	1	2	4

Table 12. Quartiles of Daily Air Pollution for Criteria Air Pollutants in 14 Cities

	PM_{10} (μg/m^3)			SO_2 (μg/m^3)			O_3 Max (ppb)			NO_2 (ppb)			CO Max (ppm)		
City	25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%	25%	50%	75%
Birmingham	19.9	30.6	46.2	11.2	16.7	25.0	39.1	50.2	63.9	—	—	—	1.7	2.5	3.6
Boulder	16.0	22.0	30.0	—[a]	—	—	35.0	47.0	60.0	—	—	—	1.4	1.9	3.1
Canton	19.2	25.6	34.5	13.9	23.4	36.5	42.1	54.6	68.6	0.0	0.0	0.0	0.7	1.0	1.5
Chicago	23.4	32.6	45.8	7.9	12.4	19.4	26.5	34.9	47.0	20.3	25.1	31.0	1.4	1.8	2.4
Colorado Springs	18.1	22.9	30.9	4.5	6.7	9.5	35.8	44.1	53.0	0.0	0.0	0.0	2.1	3.1	4.8
Detroit	20.7	32.0	48.9	12.0	19.5	29.4	30.9	42.3	56.7	16.1	21.3	27.0	1.1	1.6	2.4
Minneapolis/St Paul	16.8	24.1	34.7	3.2	7.6	14.4	—	—	—	14.4	19.2	25.0	2.2	2.9	3.7
Nashville	21.5	29.2	39.4	26.2	26.2	32.2	22.6	34.8	49.7	8.2	12.3	17.5	1.7	2.3	3.1
New Haven	17.3	26.0	37.5	13.2	20.6	36.4	35.7	47.0	63.4	21.3	27.1	33.2	2.2	3.0	4.1
Pittsburgh	19.5	30.5	46.9	25.3	39.3	58.8	28.8	39.3	54.6	20.7	26.1	31.9	1.4	2.0	2.9
Provo/Orem	21.3	30.3	44.1	—	—	—	53.0	60.0	68.0	16.3	21.3	28.3	—	—	—
Seattle	18.5	26.7	38.5	2.4	6.6	13.1	28.5	36.4	46.8	—	—	—	2.7	3.5	4.5
Spokane	23.0	36.2	56.5	0.0	3.3	12.0	38.0	44.0	52.0	—	—	—	3.6	4.8	6.5
Youngstown	21.9	29.4	40.5	25.3	36.8	49.4	36.5	50.0	68.0	12.6	16.8	21.9	0.5	1.0	1.5

[a] — = No measurement (or very limited measurement) taken.

RESULTS

Table 11 provides the 25th, 50th, and 75th percentiles of the city-specific daily counts of the Medicare hospital admissions for CVD, COPD, and pneumonia. Table 12 similarly shows the quartiles of the daily air pollution concentrations. Boulder had the lowest median PM_{10} concentration (22.0 μg/m^3) and Spokane the highest (36.2 μg/m^3).

Table 13 shows the correlations between PM_{10} concentrations and the weather variables. The correlations were modest, and, for temperature, included both positive and negative values. In 3 cities (Boulder, New Haven, and Spokane), PM_{10} was essentially uncorrelated with temperature. Table 13 also presents the correlations between daily PM_{10} concentrations and the other pollutants. Although PM_{10}

Table 13. Pearson Correlation Coefficients Between PM_{10} and Other Environmental Variables in 14 Cities

City	Temperature	Relative Humidity	SO_2	O_3 Max	CO Max	NO_2
Birmingham	0.26	–0.30	0.20	0.57	0.63	0.31
Boulder	–0.02	–0.24	—[a]	–0.01	0.46	—
Canton	0.42	–0.13	0.39	0.69	0.31	—
Chicago	0.38	–0.31	0.41	0.51	0.32	0.54
Colorado Springs	–0.34	–0.13	0.38	–0.26	0.48	—
Detroit	0.33	–0.13	0.46	0.48	0.33	0.55
Minneapolis/St Paul	0.29	–0.35	0.46	—	0.29	0.41
Nashville	0.18	–0.15	0.13	0.39	0.44	0.32
New Haven	0.08	–0.14	0.41	0.50	0.41	0.60
Pittsburgh	0.46	–0.23	0.51	0.54	0.54	0.65
Provo/Orem	–0.37	0.29	—	0.14	0.55	0.79
Seattle	–0.21	–0.09	0.80	–0.03	0.76	—
Spokane	–0.06	–0.15	0.82	0.14	0.48	—
Youngstown	0.43	–0.25	0.26	0.62	0.34	0.23

[a] — = No measurement (or very limited measurement) taken.

concentration was correlated positively with O_3 in some cities, ($r > 0.50$), there was no correlation in Boulder and Seattle and a negative correlation in Colorado Springs. PM_{10} was moderately correlated with SO_2 in most cities. There was a range of city-specific correlations between PM_{10} and CO. These differences in the pollution mixtures across the cities provide an opportunity to assess confounding and effect modification by these copollutants in the second-stage analysis.

Base Models

The city-specific base models were adjusted for season, mean temperature (same and previous day), relative humidity, and barometric pressure using LOESS smoothers, along with indicators for day of week and autoregressive terms. The city-specific LOESS spans and corresponding *df* for each parameter in the model, as well as the included autoregressive terms for the CVD, COPD, and pneumonia base models, are tabulated in Tables D.2, D.3, and D.4, respectively.

The period of study (Table 9) averaged 2,617 days (7.2 years) and ranged from 1,461 days (Youngstown) to 3,372 days (Spokane). The average spans for the LOESS smooth on season were equivalent to 260 days for CVD, 264 days for COPD, and 196 days for pneumonia. The fitted spans for temperature, relative humidity, and barometric pressure were all approximately 0.5. After adjusting for these parameters and day of week, adjustment for autoregression was required in 3 cities for CVD admissions, in 3 cities for COPD admissions, and in 7 cities for pneumonia admissions (see Tables D.2 through D.4).

Overall Effects of PM_{10}

Table 14 shows the combined overall estimate for the constrained (current day, previous day, 2-day mean, and quadratic distributed lag) and the unconstrained distributed lag model for a 10 $\mu g/m^3$ increase in PM_{10}. The city-specific PM_{10} effect estimates are presented in Tables D.5 through D.7. The effect size estimates for the 2-day mean and the quadratic distributed lag were quite similar to the effect estimate using the unconstrained lag model, and all 3 estimates were higher than the value from the constrained 1-day lag model. When the analysis using the 2-day mean of PM_{10} was repeated using only days with PM_{10} less than 50 $\mu g/m^3$, the effect size increased by 20% or more for all 3 outcomes. (City-specific analyses are presented in Table D.8.)

The city-specific estimates have considerable range (see Tables D.5 through D.7). For example, the unrestricted distributed lag estimate of the net effect on CVD admission of a 10 $\mu g/m^3$ increase in PM_{10} ranges from –1.2% in Canton to 2.2% in Colorado Springs. The fixed effect estimate across the 14 cities was 1.19% (95% CI, 0.97% to 1.41%) (Table 14). The random effect estimate (Table 14) was similar, 1.07%, although the 95% CI (0.67% to 1.46%) was wider. The random effect estimates were larger but similar to the fixed effect estimates for pneumonia and COPD admissions and again had wider confidence intervals (Table 14).

The test for heterogeneity between the city-specific unrestricted distributed lag estimate was statistically significant for pneumonia and COPD, but not for CVD (Table 15). In general, the heterogeneity of the city-specific effect estimates was greatest for COPD admissions, followed by pneumonia, and least for CVD. Recall that the

Table 14. Percent Increase in HCFA Hospital Admissions per 10 $\mu g/m^3$ Increase in PM_{10} in 14 Cities

	CVD		COPD		Pneumonia	
	% Increase	(95% CI)	% Increase	(95% CI)	% Increase	(95% CI)
Constrained lag models (Fixed Effect Estimates)						
One day mean[a]	1.07	(0.93, 1.22)	1.44	(1.00, 1.89)	1.57	(1.27, 1.87)
Previous day mean	0.68	(0.54, 0.81)	1.46	(1.03, 1.88)	1.31	(1.03, 1.58)
Two day mean[b]	1.17	(1.01, 1.33)	1.98	(1.49, 2.47)	1.98	(1.65, 2.31)
$PM_{10} < 50$ $\mu g/m^3$ (two day mean)[b]	1.47	(1.18, 1.76)	2.63	(1.71, 3.55)	2.84	(2.21, 3.48)
Quadratic distributed lag	1.18	(0.96, 1.39)	2.49	(1.78, 3.20)	1.68	(1.25, 2.11)
Unconstrained distributed lag						
Fixed effects estimate	1.19	(0.97, 1.41)	2.45	(1.75, 3.17)	1.90	(1.46, 2.34)
Random effects estimate	1.07	(0.67, 1.46)	2.88	(0.19, 5.64)	2.07	(0.94, 3.22)

[a] Lag 0.

[b] Mean of lag 0 and lag 1.

Table 15. Decomposition of Total Variance (× 1,000) Between City-Specific Estimates of Effect of 10 μg/m^3 Increase in PM_{10} into Estimated Within-City Variance and Random Variance for 14 Cities[a]

	PM_{10} Lag 0	PM_{10} Lag 1	PM_{10} Lag 0/1	Quadratic Distributed Lag	Unrestricted Distributed Lag	PM_{10} Lag 0/1 (< 50 μg/m^3)
CVD						
Total variance	0.00054	0.00025	0.00029	0.00104	0.00108	0.00090
Within variance	0.00032	0.00030	0.00035	0.00083	0.00084	0.00085
Random variance	0.00022	0.00000	0.00000	0.00021	0.00024	0.00005
Heterogeneity χ^2	23.3[b]	20.4	24.5[b]	17.5	16.8	33.6[b]
COPD						
Total variance	0.00527	0.00877	0.00352	0.02834	0.02668	0.01014
Within variance	0.00287	0.00264	0.00318	0.00703	0.00710	0.00887
Random variance	0.00240	0.00613	0.00034	0.02131	0.01959	0.00127
Heterogeneity χ^2	22.0	29.6[c]	22.9[b]	31.3[c]	30.5[c]	18.5
Pneumonia						
Total variance	0.00176	0.00092	0.00157	0.00586	0.00576	0.00685
Within variance	0.00123	0.00113	0.00133	0.00275	0.00278	0.00325
Random variance	0.00053	0.00000	0.00023	0.00310	0.00298	0.00360
Heterogeneity χ^2	18.7	21.3	26.8[b]	21.4	28.0[c]	41.1[c]

[a] Chi squared (χ^2) test for heterogeneity between city-specific effects.

[b] $P < 0.05$.

[c] $P < 0.01$.

fewest admissions were reported for COPD, followed by pneumonia, with the largest numbers for CVD admissions. The estimated random variance of the PM_{10} associations was similarly largest for COPD, followed by pneumonia, and smallest for CVD admissions (Table 15).

Distributed Lag Effects Over Time

Figure 31 shows the combined city estimate of the unrestricted distributed lag associations between PM_{10} and the 3 analyzed causes of admissions. Cardiovascular admissions show a higher effect at lag 0, dropping to a more modest effect at lags 1 and 2, and then oscillates around 0 (no effect) for longer lags.

For COPD admissions, the effect was similar for PM_{10} on the concurrent day and the previous day and dropped to around 0 at lag 2 and subsequent days. For pneumonia admissions, the effect decreased continuously for lags 0 through 2 and then oscillated about 0 for further lags. For comparison, Figure 31 also presents the combined constrained quadratic distributed lag associations for each hospital admissions diagnosis.

Table 16. Estimated Change in Effect of 10 μg/m^3 PM_{10} for 5 Percentage Point Increase in College Education, Unemployment, Poverty and Nonwhite Percentages of the Population for 14 Cities, by Diagnosis

	College Education		Unemployment		Poverty		Nonwhite	
	%	(95% CI)	%	(95% CI)	%	(95% CI)	%	(95% CI)
CVD	−0.00	(−0.21, 0.21)	0.38	(−0.02, 0.78)	0.16	(−0.16, 0.49)	0.06	(−0.02, 0.14)
COPD	0.55	(−0.32, 1.42)	−0.15	(−2.12, 1.86)	−0.84	(−2.21, 0.54)	−0.18	(−0.54, 0.19)
Pneumonia	0.38	(−0.03, 0.78)	−0.22	(−1.21, 0.78)	−0.55	(−1.24, 0.15)	−0.03	(−0.22, 0.16)

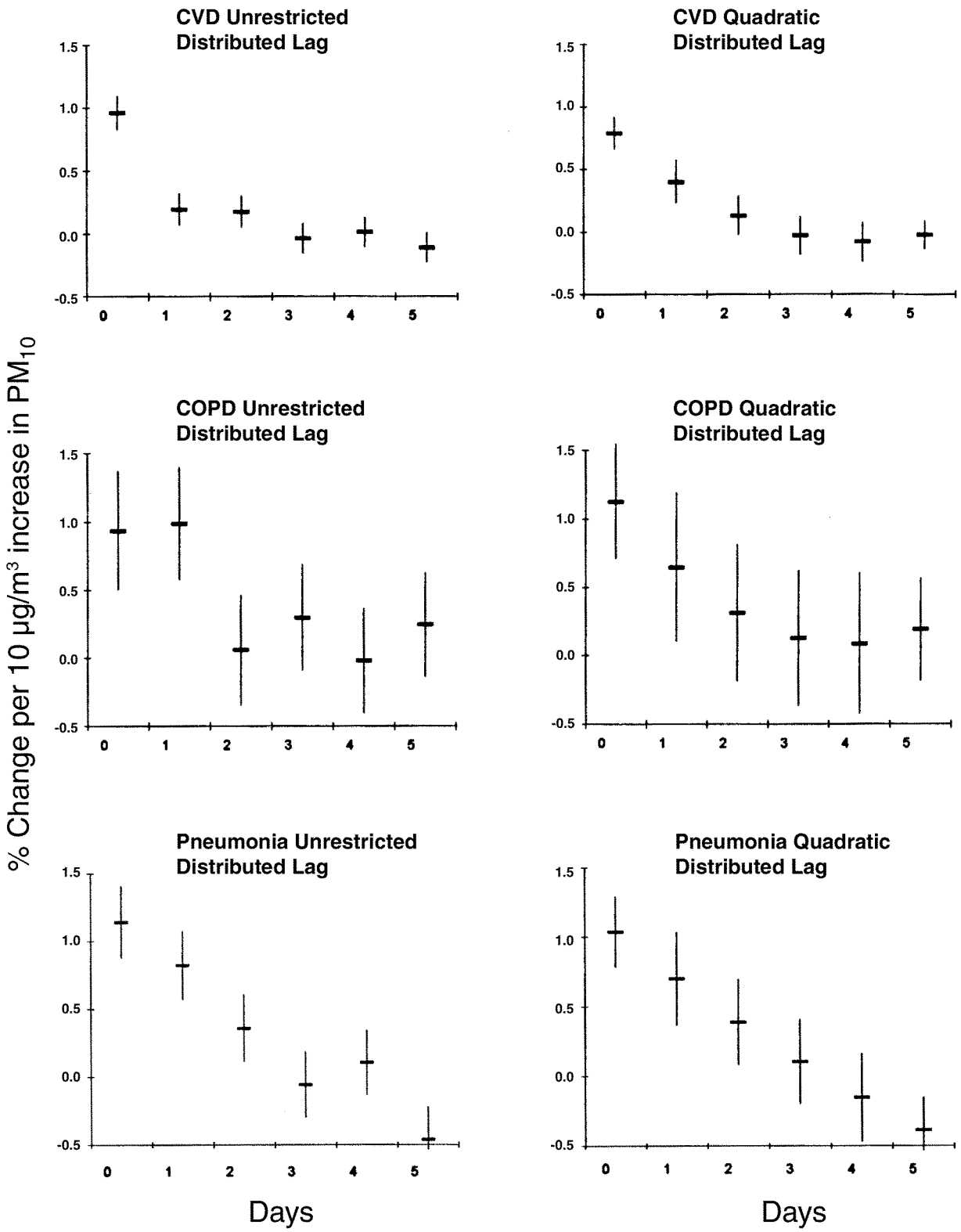


Figure 31. Combined estimates across the 14 cities of the percentage change in HCFA hospital admissions per 10 µg/m^3 PM_{10} by diagnosis for unrestricted and quadratic distributed lag models.

Second-Stage Models

Sociodemographic factors Neither the percentage of the population living in poverty nor the percentage of the population that was nonwhite was a significant modifier of the PM_{10} effect estimates in the selected 14 cities. Table 16 shows the change from the baseline PM_{10} effect size (as percent increase in admission per 10 μg/m^3 increase in concentration) associated with a 5 percentage point increase in each measure.

Hospitalization rates Neither hospital admission rates for CVD, pneumonia, or COPD were associated with modification of the PM_{10} effect estimates.

Weather In the meta-regression with weather variables, we found the coefficients for temperature and relative humidity were highly nonsignificant for all 3 outcomes (Figure 32). In addition, the intercept terms in these models were similar in magnitude to those in the baseline meta-analysis. Thus, temperature and relative humidity did not modify the PM_{10} effect estimates.

Copollutants Figure 33 shows, for CVD, COPD, and pneumonia admissions, the city-specific effects of PM_{10} plotted against the regression coefficients relating SO_2 and O_3 to PM_{10} in each city. We also considered the CVD regression coefficients for CO versus PM_{10}.

The plots of Figure 33 give a picture of the range of the results by city. The city-specific effect estimates ranged between –2% and 2% for each 10 μg/m^3 increase in PM_{10}. For pneumonia, estimated city-specific PM_{10} effects were between –2% and 9%. For COPD, the maximum city-specific effect was about 7%, except for an estimate of 19% in Boulder (95% CI, 8%, 31%). They also show the range of regression coefficients relating PM_{10} to the other pollutants. For O_3 they include both positive and negative slopes, with a wide range among the positive slopes. For SO_2 and CO, the slopes were always positive but varied by almost an order of magnitude.

As explained previously in the Methods section, Second-Stage Assessment of Confounding and Effect Modification, if the PM_{10} effect were due to confounding by other pollutants, the plots would show a significantly increasing trend with increasing slope

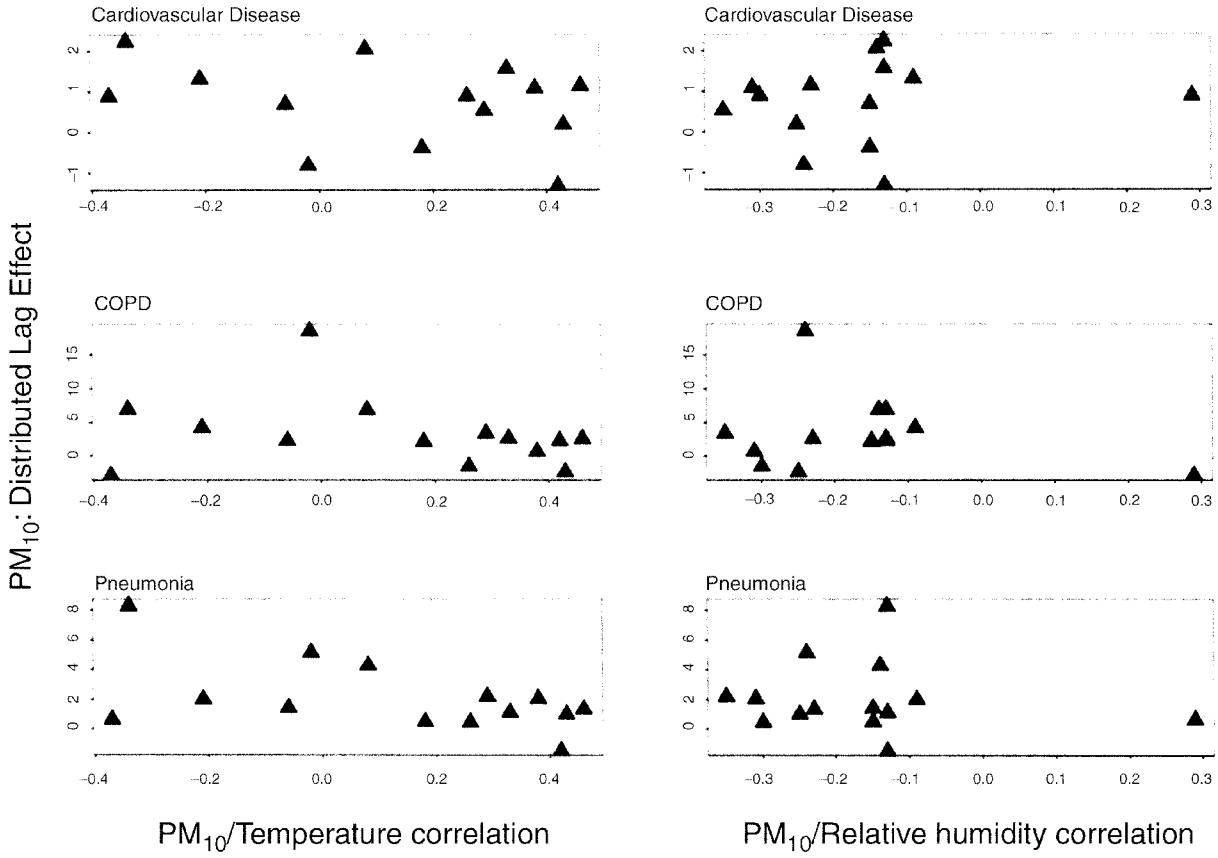


Figure 32. Effect size estimates of PM_{10} versus the correlation of PM_{10} with temperature and relative humidity for the three diseases.

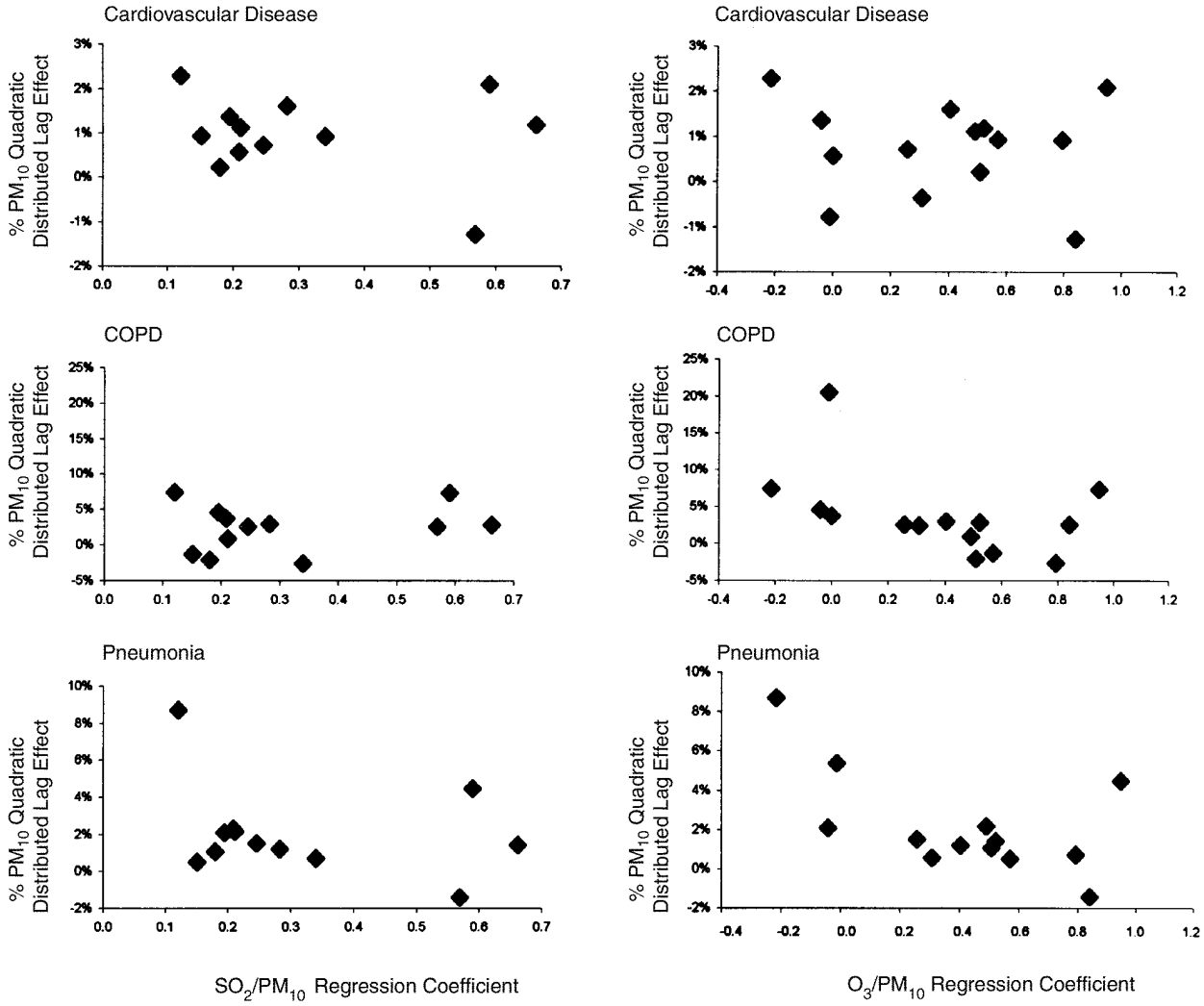


Figure 33. Effects of PM_{10} in each city plotted against the regression coefficients in relation to SO_2 and O_3 to PM_{10} for cardiovascular disease, COPD and pneumonia.

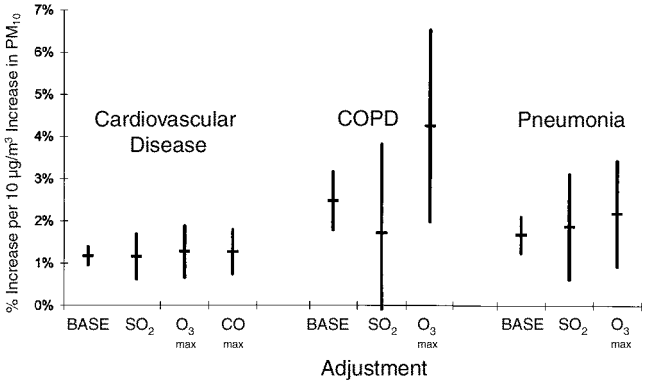


Figure 34. Meta-Regression adjustment for copollutants. Estimated effects of PM_{10} on cardiovascular disease, COPD, and pneumonia in each city without adjustment (base) or with adjustment for individual gaseous pollutants.

between the pollutants. Figure 33 shows little evidence for this pattern. These results are confirmed by the meta-regression estimates, shown in Figure 34. Here the baseline estimate is the result of a simple meta-analysis not incorporating any of the other pollutants. Plotted above each pollutant is the estimated intercept in the meta-regression of the PM_{10} coefficients against the slopes between that copollutant and PM_{10}. For all 3 outcomes, the results appear quite stable when controlling for confounding by gaseous pollutants.

DISCUSSION

NMMAPS Part II provides the results of 2 complementary analyses of air pollution and health in the United States: one on mortality in up to 90 cities and the other on

hospitalization in persons 65 years of age and older in 14 cities. The 2 analyses are parallel in their methods, both drawing on a national sampling frame for selection of study locations and both using 2-stage analytic approaches for combining evidence across locations while controlling for potential confounding and also evaluating effect modification. Although the details of the implementation of the statistical analyses were slightly different, reflecting different data sources and some differences in approach, the underlying statistical approaches are conceptually similar.

At the first stage, a time-series analysis is conducted within each of the locations to estimate the effect of PM_{10} and other pollutants. The distributed lag approach as applied to the hospitalization data is flexible in its assessment of the lag-response relationship, and its use provided new insights into the lag structure. Distributed lag models have the strength of not requiring selection of specific lag structures. For hospitalization, analyses using the distributed lag approach showed that the effect of a 24-hour increase in PM_{10} is spread over that day and several subsequent days, so that using only a single day's concentration underestimates the impact of PM_{10}. The lack of daily data for PM_{10} was a barrier to applying a similar approach to the mortality data, although a study of 10 cities (Appendix B) similarly shows that the effects of PM_{10} are spread over several days and are underestimated by a single-day model. An analysis of mortality data in this study for Chicago, Minneapolis/St Paul, and Pittsburgh (Figure 19) using distributed lag methods gave generally similar findings to those using specified lags. At the second stage of the analysis, the Bayesian hierarchical model used to combine the mortality data and the meta-analysis used for the morbidity data provide point estimates that can be adjusted for potential confounding factors.

Together, the 2 sets of analyses—that of mortality in 90 cities and of hospitalization in persons 65 years and older in 14 cities—provide new and strong evidence linking particulate air pollution to adverse health effects. In locations broadly representative of the United States, and, in fact, capturing a substantial proportion of the population, PM_{10} concentration was associated with both daily mortality counts and hospitalization of persons 65 years of age and older. For mortality, the effect of PM_{10} was greater for cardiorespiratory deaths, as would be expected if persons with advanced, chronic heart and lung disease are particularly susceptible to air pollution. For morbidity, effects were found in specific diagnostic groups (CVD, COPD, and pneumonia) included within the overall cardiopulmonary grouping used for mortality. The effects of PM_{10} on mortality and hospitalization persisted and were not substantively changed with control for other pollutants. This is noteworthy because the potential for confounding by other pollutants was examined using different approaches (first stage versus second stage) in the 2 sets of analyses. We found some evidence for modification of the PM_{10} effect by mean PM_{10} level, in a direction implying greater effect at lower concentrations. These results were supported by the morbidity analysis, which showed a higher slope when restricted to days with PM_{10} less than 50 $\mu g/m^3$. There was some evidence for heterogeneity of the effect of PM_{10} by geographic site, with greater effect in the Northeast.

The 20-city and 90-city analyses strongly support the evidence from prior studies of PM and mortality. These earlier studies, largely based on data from single cities, used a variety of measures of PM, including total suspended particles (TSP), British smoke, PM_{10}, and $PM_{2.5}$. The statistical methods applied to assess the pollution-mortality relationships were also heterogeneous among the studies; for example, there was no uniformity in the approaches used to control for temporally varying factors or for other pollutants. Nonetheless, using a weight-of-evidence approach, the EPA interpreted the study results as indicating a possibly causal association between particulate matter and air pollution (EPA 1996a).

In a meta-analysis of US studies published through 1993, Dockery and Pope (1994) estimated the effect of particulate air pollution on mortality as an increase of 1% for each 10 $\mu g/m^3$ increase in PM_{10}. In a subsequent update using reports published through 1995, there was little change in this estimate (Pope et al 1995a). Schwartz (1994) also carried out a meta-analysis of studies published through 1993, but included London and Minneapolis in addition to the 8 cities considered by Dockery and Pope. The resulting point estimate was 1.07% for each 10 $\mu g/m^3$ increase in PM_{10}. In the APHEA project, common analysis techniques were applied to data from 12 European cities. Summary measures were then estimated in a second step. For the 6 western European cities, mortality was estimated to increase by 0.4% for each 10 $\mu g/m^3$ increase in PM_{10}. In a recent reanalysis of the APHEA data using generalized additive models, Touloumi and colleagues (1996) report higher results, above 0.6% for each 10 $\mu g/m^3$ increase in PM_{10}. The estimate in our 20-city analysis, approximately 0.5% for each 10 $\mu g/m^3$ increase in PM_{10} at lag 1, is close to the APHEA project's estimate. The lower value, in comparison with the prior meta-analyses by Dockery and Pope and by Schwartz, may reflect differences in analysis techniques and the cities selected. The lower value may also reflect the fact that the prior studies used multiple-day averages of PM_{10}, whereas our study could use only a single day's exposure.

We did not find an effect of O_3 on total mortality across the full year (Figures 26 and 30), but an effect was observed after limiting the analysis to the summer when O_3 levels were highest. Other recent studies have generally shown associations of O_3 with mortality (Thurston and Ito 1999). In the APHEA study, 1-hour maximum O_3 levels were associated with daily numbers of deaths in 4 cities (London, Athens, Barcelona, and Paris), and a quantitatively similar effect was found in a group of 4 additional cities considered by the authors (Amsterdam, Basel, Geneva, and Zurich) (Touloumi et al 1997). For an increase of 50 $\mu g/m^3$ in the 1-hour maximum, the estimated relative risk was 1.029 (1.1% for each 10 ppb), using a random effects model for combining the city-specific data. Thurston and Ito (1999) pooled data from 15 studies and estimated the relative risk of death to be 1.036 for each 100 ppb (0.36% for each 10 ppb) increase in the daily 1-hour maximum. For the summer months, we estimated a comparable level of effect, 0.25% for each 10 ppb. The findings of these 3 analyses (APHEA, Thurston and Ito, and NMMAPS) provide consistent data that O_3 exposure is associated with increased mortality. The negative associations in the winter remain puzzling and may reflect some unmeasured confounding factor.

The effect of PM_{10} varied across the 90 cities and the pooled estimates by region for mortality were higher in the Northeast, Industrial Midwest, and Southern California (Figures 21 and 22). These regions tend to have higher sulfate concentrations than the other regions (EPA 1996a), and follow up on this observation with morbidity outcomes is warranted. The higher effect in the Northeast might also reflect a greater density of monitors and hence less downward bias from measurement error.

Although we have analyzed a large data set based on the 90 largest cities of the United States, limitations of the analyses need to be considered. Data on concentrations of $PM_{2.5}$, the respirable particles now regulated by the EPA, are not yet available, as a monitoring network for particles in this size range is only now being implemented. We used PM_{10} because it has been monitored since 1987; there is variation across the United States in the proportion of PM_{10} mass that is made up of $PM_{2.5}$. Additionally, for regulatory purposes, monitoring of PM_{10} is required only every 6 days, limiting the extent of available data. Concentration of PM_{10} measured at the outdoor monitors is a surrogate in the time-series analyses for the contribution of PM_{10} in ambient air to personal exposures of persons at risk for dying. Measurement error arising from the inherent assumption that outdoor concentration is valid for this purpose has been proposed as a substantial limitation of the time-series analyses (Lipfert and Wyzga 1997). Our review of this issue, reported in NMMAPS Part I, suggests that measurement error would generally tend to bias estimates of the effect of PM_{10} downward.

Both the mortality and morbidity analyses provided results for single cities. The point estimates of relative risk varied across the cities to a degree; exploration of the causes of this variation is a focus of the planned combined analyses. We caution against attempts to interpret estimates for any specific city, particularly if the goal is to gauge whether PM is having a greater or lesser effect than in other locations.

These daily time-series analyses do not address the extent of life-shortening associated with these daily associations. The finding that the association is strongest for cardiorespiratory causes of death is consistent with the hypothesis that persons made frail by advanced heart and lung disease are susceptible to the effects of air pollution. The findings from several epidemiologic studies of longer-term effects of air pollution on mortality suggest that air pollution may have a more severe effect than simply advancing death by a few days (Dockery et al 1993; Pope et al 1995b). Analyses of daily time-series data, conducted at different temporal frequencies, also indicate that the effect of air pollution may go beyond only shortening life by a few days (Azizi et al 1995; Schwartz 1999; Zeger et al 1999b). An appropriate next step is the application of the new analytic tools described by Zeger and colleagues and Schwartz to the multicity database.

The analyses of the HCFA data on hospitalization confirm the many previous reports, based primarily on single cities, that PM_{10} levels are associated with hospital admissions for heart and lung disease (Pope and Dockery 1999). In general, the effect size estimates in NMMAPS are consistent with the previous studies. This new analysis involves more years of follow up than most previous studies and includes 14 cities spread across the country, with a wide range of weather and copollutants. Burnett and colleagues (1995) have evaluated the association of cardiac and respiratory admissions to 168 acute care hospitals in Ontario, Canada, with daily levels of particulate sulfate. Spatial detail was lacking in that study, as there were only 9 monitors. Both categories of admissions were significantly associated with sulfate levels. In a subsequent report for 16 Canadian cities, Burnett and colleagues (1997) found that O_3 concentration was associated with respiratory hospitalizations from April through December. In the 11 cities with coefficient of haze data, their surrogate for particle levels was also associated with respiratory hospitalizations.

The substantial range of weather patterns in the 14 locations that are listed in Table 7 supports the conclusion that

the results were not confounded by inadequate control for weather. This was confirmed by the second-stage regressions, which showed that the PM_{10} coefficients were not influenced by the correlation between PM_{10} and weather.

For all 3 of the diagnostic categories, the effect of PM_{10} was shown to be spread over more than a single day in the distributed lag model. As a consequence, use of a single exposure day would underestimate the effect of PM_{10}, substantially in some instances (Table 14). Summaries of the health evidence for risk assessment and policy-making purposes need to take this downward bias into account. Most studies of air pollution have used multiday averages of pollutant concentrations, but some have not, and this will need to be taken into account in any future meta-analysis.

Overall, this study provides strong evidence of association between PM_{10} levels and exacerbation of chronic heart and lung disease sufficiently severe to warrant hospitalization. The association cannot be explained by confounding that is addressed in both stages of the analysis, although there is always the possibility of some residual confounding. Confounding by weather was considered above and can be set aside. Confounding by the gaseous pollutants (NO_2, SO_2, and CO) has been raised as a major concern in relation to the findings of prior studies (Lipfert 1997). We have addressed this possibility in the second stage of the regression modeling. We found that the effect size estimates for PM_{10} and hospital admissions for CVD, COPD, and pneumonia changed little after control for potential confounding by the gaseous air pollutants. The standard errors increased because our second-stage analysis had a limited sample size (14 points in a regression estimating an intercept and a slope), but overall the evidence for confounding was small.

We have not found evidence that key socioeconomic factors such as poverty and race are modifiers of the effect of PM_{10} on either mortality or hospitalization. The scale of Poisson models needs to be considered in interpreting the lack of effect modification. The Poisson models are relative risk models and inherently multiplicative. Hence, a given change in PM_{10} is associated with a given percent increase in the event rate. If a town with greater poverty or a larger percentage of nonwhites has a higher baseline event rate, then a 3% increase in the rate from baseline will be a greater increase (per 10,000 persons aged 65 and older) in that town than it would be in a town with a lower baseline rate.

Further, it is possible that effect modification does not occur. Alternatively, the medical conditions that predispose to higher risk may not be well captured by the socioeconomic factors recorded by the Census. More specific information on medical conditions, rather than on social factors, may be needed to explore effect modification, particularly in relation to personal susceptibility. Finally, we have used county-level data for these social factors because our admissions and mortality data are on the county level. The variation in socioeconomic status within the typical urban county, however, is usually considerably larger than the variation across counties. The sociodemographic factors considered in the second stages of the models may be too ecologically coarse to be meaningful. Future studies using finer geographical data may be able to detect effect modification.

Recent observational and experimental studies provide further insights into mechanisms by which PM could increase morbidity and mortality. In animal models, exposure to particles is associated with reduced heart rate variability and increased fibrinogen levels (Godleski et al 1997; Watkinson et al 1998a,b). These are risk factors for arrhythmias and ischemic events, which underlie many hospital admissions for heart disease. In human studies, exposure to airborne particles has been associated with increases in plasma viscosity (Peters et al 1997) and decreases in heart rate variability (Pope et al 1999), paralleling the animal study findings.

Animal models of COPD or chronic lung inflammation have been shown to have increased vulnerability to combustion particles in comparison with normal animals (Gilmour et al 1996; Li et al 1996; Pritchard et al 1996; Costa and Dreher 1997). Exposure of animals previously infected with strep pneumonia to concentrated air particles resulted in a doubling of streptococcal lung area involved with pneumonia and in bacterial burdens. Experimental influenza infections have been similarly shown to be exacerbated by air pollution (Clarke et al 1997; Zelikoff et al 1997).

In interpreting the associations of PM_{10} with mortality and morbidity, we find that the shape of the exposure-response relationship at typical ambient concentrations is a critical uncertainty. For assessing the impact of air pollution on public health, we need to know whether the associations are dominated by high pollution days or persist at the concentrations seen on most days. For mortality, the PM_{10} daily effect across the communities may be modified by mean PM_{10} level (Table 5). This finding implies that a tendency toward higher slopes at lower concentrations was seen and is sometimes significant.

When we restricted our analysis of hospitalization to days with concentrations below one-third of the current NAAQS, we still found a significant association between PM_{10} and hospital admissions for all 3 diagnostic categories. Moreover, the effect size increased by 20% or more

with this restricted range of concentration. This increase in effect size at lower concentrations has been noted previously in a mortality study and in the 90-city analysis (Schwartz et al 1996). For a significant association to persist and to be greater on days with PM_{10} levels below 50 $\mu g/m^3$, any threshold would have to be far below that level, probably down at background levels. These findings suggest that if the true concentration-response curve is curvilinear, the higher slopes occur at lower concentrations and no threshold exists.

The epidemiologic evidence on PM and mortality and morbidity has prompted the promulgation of a new standard in the United States and a rethinking of guidelines for PM in Europe. Our analyses provide evidence that air pollution with particles is still adversely affecting the public's health and strengthen the rationale for limiting concentrations of respirable particles in outdoor air.

ACKNOWLEDGMENTS

The investigators acknowledge Ms Jing Xu for her work in database development and programming, and Ms Dawn Kelly and Ms Charlotte Gerczak for their assistance in preparing this report.

REFERENCES

AIRS (Aerometric Information Retrieval System). US Environmental Protection Agency, Office of Air Quality Planning and Standards, AIRS Web site. Available at: *www.epa.gov/airs/airs.html*. Updated December 14, 1999.

Akaike H. 1973. Information theory and an extension of the maximum likelihood principal. In: Second International Symposium on Information Theory (Petrov BN, Csaki F, eds). Akademiai Kaiado, Budapest, Hungary.

Almon S. 1965. The distributed lag between capital appropriations and expenditures. Econometrica 33:178–196.

Azizi BH, Zulkifli HI, Kasim MS. 1995. Protective and risk factors for acute respiratory infections in hospitalized urban Malaysian children: A case-control study. Southeast Asian J Trop Med Public Health 26:280–285.

Bascom R, Bromberg PA, Costa DA, Devlin R, Dockery DW, Frampton MW, Lambert W, Samet JM, Speizer FE, Utell M. 1996. Health effects of outdoor air pollution. Part 2. Am J Respir Crit Care Med 153:477–498.

Brumback BA, Ryan LM, Schwartz J, Neas LM , Stark PC, Burge HA. 1999. Transitional regression models with application to environmental time series. J Am Stat Assoc (in press).

Burnett RT, Brook JR, Yung WT, Dales RE, Krewski D. 1997. Associations between ozone and hospitalization for respiratory diseases in 16 Canadian cities. Environ Res 72:24–31.

Burnett RT, Dales R, Krewski D, Vincent R , Dann T, Brook JR. 1995. Associations between ambient particulate sulfate and admissions to Ontario hospitals for cardiac and respiratory diseases. Am J Epidemiol 142:15–22.

Clarke RW, Hemenway DR, Frank R, Kleeberger SR, Longphre MV, Jakab GJ. 1997. Particle associated sulfate exposure enhances murine influenza mortality. Am J Respir Crit Care Med 155:A245.

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

Costa D, Dreher K. 1997. Bioavailable transition metals in particulate matter mediate cardiopulmonary injury in health and compromised animal models. Environ Health Perspect (Suppl 5) 105:1053–1060.

Denison D, Mallick B, Smith A. 1998. Automatic Bayesian curve fitting. J R Stat Soc Series B 60:333–350.

DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7:177–188.

Dockery DW, Pope CA III. 1994. Acute respiratory effects of particulate air pollution. Annu Rev Public Health 15:107–132.

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

Dominici F, Samet J, Xu J, Zeger S. 2000. Combining evidence on air pollution and daily mortality from the largest 20 US cities: A hierarchical modeling strategy. J R Stat Soc Series C (in press).

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

Gelfand AE, Smith AFM. 1990. Sampling based approaches to calculating marginal densities. J Am Stat Assoc 85:398–409.

Gilmour PS, Brown DM, Lindsay TG, Beswich PH, MacNee W, Donaldson K. 1996. Adverse health effects of

PM_{10} particles: Involvement of iron in generation of hydroxy radical. Occup Environ Med 53:817–822.

Godleski JJ, Lovett EG, Sioutas C, Killingsworth CR, Krishnamurthi GG, Hatch V, Wolfson M, Ferguson ST, Koutrakis P, Verrier RL. 1997. Impact of inhaled concentrated ambient air particles on canine electrocardiographic patterns (abstract). In: The Thirteenth Health Effects Institute Annual Conference Program. Health Effects Institute, Cambridge MA.

Hastie TJ, Tibshirani RJ, eds. 1990. Generalized Additive Models. Chapman & Hall, New York NY.

Judge GG, Griffiths WE, Hill RC, Lutkepohl H, Lee T-C. 1985. The Theory and Practice of Econometrics. John Wiley & Sons, New York NY.

Katsouyanni K, Schwartz J, Spix C, Touloumi G, Zmirou D, Zanobetti A, Wojtyniak B, Vonk JM, Tobias A, Ponka A, Medina S, Bacharova L, Anderson HR. 1996. Short-term effects of air pollution on health. A European approach using epidemiologic time series data: The APHEA protocol. J Epidemiol Community Health 50:S12–S18.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak B, Sunyer J, Bacharova L, Schouten JP, Ponka A, Anderson HR. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: Results from the APHEA project. Br Med J 314:1658–1663.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Li XY, Gilmour PS, Donaldson K, MacNee W . 1996. Free radical activity and proinflammatory effects of particulate air pollution (PM_{10}) in vivo and in vitro. Thorax 51:1216–1222.

Lindley DV, Smith AFM. 1972. Bayes estimates for the linear model (with discussion). J R Stat Soc 34:1–41.

Lipfert FW. 1997. Air pollution and human health: Perspectives for the '90s and beyond. Risk Anal 17:137–146.

Lipfert FW, Wyzga RE. 1997. Air pollution and mortality: The implications of uncertainties in regression modeling and exposure measurement. J Air Waste Manage Assoc 47:517–523.

Little RJA, Rubin DB. 1987. Statistical Analysis With Missing Data. John Wiley & Sons, New York NY.

Mallows CL. 1973. Some comments on C_p. Technometrics 15:661–675.

Moolgavkar SH, Luebeck EG, Hall TA, Anderson EL. 1995. Air pollution and daily mortality in Philadelphia. Epidemiology 6:476–484.

Morris CN, Normand S-L. 1992. Hierarchical models for combining information and for meta-analysis. In: Bayesian Statistics 4. Proceeding of the Fourth Valencia International Meeting (Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds). Oxford University Press, Oxford, England.

Nehls GJ, Akland GG. 1973. Procedures for handling aerometric data. J Air Pollut Control Assoc 23:180–184.

Peters A, Doring A, Wichmann HE, Koenig W. 1997. Increased plasma viscosity during an air pollution episode: A link to mortality? Lancet 349:1582–1587.

Pindyck RS, Rubinfeld DL. 1998. Econometric models and economic forecasting. Irvin–McGraw-Hill, New York NY.

Pope CA III, Dockery DW. 1999. Epidemiology of particle effects . In: Air Pollution and Health (Holgate ST, Samet JM, Koren HS, Maynard RL, eds). Academic Press, San Diego CA.

Pope CA III, Schwartz J, Ranson MR. 1992. Daily mortality and PM_{10} pollution in Utah Valley. Arch Environ Health 47:211–217.

Pope CA III, Dockery DW, Schwartz J. 1995a. Review of epidemiological evidence of health effects of particulate air pollution. Inhalation Toxicol 7:1–18.

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

Pope CA III, Verrier RL, Lovett EG, Larson AC, Raizenne ME, Kanner RE, Schwartz J, Villegas GM, Dockery DW. 1999. Heart rate variability associated with particulate air pollution. Am Heart J 138(5 Pt 1):890–899.

Pritchard RJ, Ghio AJ, Lehmann JR, Winsett DW, Tepper JS, Park P, Gilmour MI, Dreher KL, Costa DL. 1996. Oxidant generation and lung injury after particulate air pollutant exposure increase with the concentrations of associated metals. Inhalation Toxicol 8:457–477.

Rabiner LR, Gold B. 1965. Theory and application of digital signal processing. Prentice Hall, Englewood Cliffs NJ.

Samet J, Dominici F, Xu J, Zeger S. 2000a. Particulate Air Pollution and Mortality: Findings from 20 US Cities. Technical Report. Johns Hopkins University, Baltimore MD.

Samet JM, Zeger S, Dominici F, Dockery D , Schwartz J. 2000b. The National Morbidity, Mortality, and Air Pollution Study, Part I: Methods and Methodologic Issues. Health Effects Institute, Cambridge MA.

Samet JM, Zeger SL, Kelsall JE, Xu J. 1996. Weather, Air Pollution, and Mortality in Philadelphia 1973–1980. Interim Report to Particulate Air Pollution and Daily Mortality: Analysis of the Effects of Weather and Multiple Air Pollutants, The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge MA.

Schwartz J. 1994. Air pollution and daily mortality: A review and meta-analysis. Environ Res 64:36–52.

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

Schwartz J. 1999. Air pollution and hospital admissions for heart disease in eight US counties. Epidemiology 10:17–22.

Schwartz J. 2000a. Harvesting and long-term exposure effects in the relationship between air pollution and mortality. Am J Epidemiol 151:440–448.

Schwartz J. 2000b. The distributed lag between air pollution and daily deaths. Epidemiology 11:320–326.

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

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

Thurston GD, Ito K. 1999. Epidemiological studies of ozone exposure effects. In: Air Pollution and Health (Holgate ST, Samet JM, Koren HS, Maynard RL, eds). Academic Press, San Diego CA.

Touloumi G, Katsouyanni K, Zmirou D, Schwartz J, Spix C, Ponce de Leon A, Tobias A, Quenel P, Rabczenko D, Bacharova L, Bisanti L, Vonk JM, Ponka A. 1997. Short-term effects of ambient oxidant exposure on mortality: A combined analysis within the APHEA project. Am J Epidemiol 146:177–185.

Touloumi G, Samoli E, Katsouyanni K. 1996. Daily mortality and "winter type" air pollution in Athens, Greece—A time series analysis within the APHEA project. J Epidemiol Community Health (Suppl 1) 50:S47–S51.

US Environmental Protection Agency. 1996a. Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. EPA-452\R-96-013. Office of Air Quality Planning and Standards, Research Triangle Park NC.

US Environmental Protection Agency. 1996b. Air Quality Criteria for Particulate Matter. EPA/600/P-95/001aF. Office of Research and Development, Washington DC.

Vedal S. 1997. Ambient particles and health: Lines that divide. J Air Waste Manage Assoc 47:551–581.

Watkinson WP, Campen MJ, Costa DL. 1998a. Cardiac arrhythmia induction after exposure to residual oil fly ash particles in a rodent model of pulmonary hypertension. Toxicol Sci 41:209–216.

Watkinson WP, Campen MJ, Kodavanti UP, Ledbetter AD, Costa DL. 1998b. Effects of inhaled residual oil fly ash particles on electrocardiographic and thermoregulatory parameters in normal and compromised rats. Am J Respir Crit Care Med 157:A150.

Zeger S, Thomas D, Dominici F, Cohen A, Schwartz J, Dockery D, Samet J. 1999a. Exposure Measurement Error in Time-Series Studies of Air Pollution: Concepts and Consequences. Technical Report. Johns Hopkins University, Baltimore MD.

Zeger SL, Dominici F, Samet J. 1999b. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.

Zelikoff JT, Fang KG, Nadziejko LC, Chen LC, Gordon T, Schlesinger RB, Cohen MD. 1997. Immunomodulation as a mechanisms for PM_{10}-induced effects upon host mortality (abstract). In: The Thirteenth Health Effects Institute Annual Conference Program. Health Effects Institute, Cambridge MA.

APPENDIX A. Demographic and Pollution Data for 90-City Analysis

Table A.1. The 90 Cities and Their Included Counties by Population Size with Mean Daily Number of Deaths by Category, 1987–1994

City	Abbreviation	County	State	Population	Mean Daily Deaths		
					Total	CVD/ Respiratory Disease	Other
Los Angeles	la	Los Angeles	CA	8,863,164	148.1	87.0	61.1
New York	ny	Bronx, Kings, New York, Richmond, Queens, Westchester	NY	8,197,430	190.9	108.3	82.6
Chicago	chic	Cook	IL	5,105,067	113.9	62.0	51.9
Dallas/Fort Worth	dlft	Collin, Dallas, Rockwall, Tarrant	TX	3,312,553	47.9	26.0	21.9
Houston	hous	Harris	TX	2,818,199	39.9	20.0	19.8
San Diego	sand	San Diego	CA	2,498,016	41.6	22.6	19.0
Santa Ana/Anaheim	staa	Orange	CA	2,410,556	32.4	18.7	13.6
Phoenix	phoe	Maricopa	AZ	2,122,101	38.4	20.9	17.5
Detroit	det	Wayne	MI	2,111,687	46.9	26.5	20.4
Miami	miam	Dade	FL	1,937,094	43.8	23.6	20.2
Philadelphia	phil	Philadelphia	PA	1,585,577	42.3	21.5	20.8
Minneapolis/St Paul	minn	Hennepin, Ramsey	MN	1,518,196	26.3	13.9	12.4
Seattle	seat	King	WA	1,507,319	25.6	13.4	12.2
San Jose	sanj	Santa Clara	CA	1,497,577	19.7	10.7	9.0
Cleveland	clev	Cuyahoga	OH	1,412,140	36.5	20.1	16.4
San Bernardino	sanb	San Bernardino	CA	1,418,380	20.6	12.1	8.5
Pittsburgh	pitt	Allegheny	PA	1,336,449	37.6	21.0	16.9
Oakland	oakl	Alameda	CA	1,279,182	22.2	12.2	10.0
Atlanta	atla	Fulton, De Kalb	GA	1,194,788	17.5	8.8	8.7
San Antonio	sana	Bexar	TX	1,185,394	20.1	10.5	9.6
Riverside	rive	Riverside	CA	1,170,413	20.1	12.4	7.7
Denver	denv	Denver, Adams, Arapahoe	CO	1,124,159	9.1	5.0	4.1
Sacramento	sacr	Sacramento	CA	1,041,219	17.2	9.5	7.7
St Louis	stlo	St Louis City	MO	993,529	10.7	6.0	4.7
Buffalo	buff	Erie	NY	968,532	25.2	14.8	10.3
Columbus	clmo	Franklin	OH	961,437	16.8	8.9	7.9
Cincinnati	cinc	Hamilton	OH	866,228	19.9	11.0	8.9
St Petersburg	stpe	Pinellas	FL	851,659	29.3	17.7	11.6
Kansas City	kan	Clay, Jackson, Platte	MO	844,510	16.7	9.3	7.5
Honolulu	hono	Honolulu	HI	836,231	11.9	6.4	5.5
Tampa	tamp	Hillsborough	FL	834,054	16.9	9.1	7.8
Memphis	memp	Shelby	TN	826,330	17.5	9.7	7.7
Indianapolis	indi	Marion	IN	797,159	16.9	9.0	8.0
Newark	nwk	Essex	NJ	778,206	18.4	8.7	9.7
Baltimore	balt	Baltimore City	MD	736,014	20.2	9.8	10.4

(Table continues on next page.)

Table A.1, *continued*. The 90 Cities and their Included Counties by Population Size with Mean Daily Number of Deaths by Category, 1987–1994

City	Abbreviation	County	State	Population	Mean Daily Deaths		
					Total	CVD/ Respiratory Disease	Other Deaths
Salt Lake City	salt	Salt Lake	UT	725,956	9.3	4.9	4.4
Rochester	roch	Monroe	NY	713,968	14.6	7.9	6.7
Worcester	wor	Worcester	MA	709,705	15.2	8.2	6.9
Orlando	orla	Orange	FL	677,491	11.0	5.8	5.2
Jacksonville	jckv	Duval	FL	672,971	13.0	7.0	6.0
Fresno	fres	Fresno	CA	667,490	11.1	6.2	4.9
Louisville	loui	Jefferson	KY	664,937	16.3	8.8	7.5
Boston	bost	Suffolk	MA	663,906	13.2	6.5	6.7
Birmingham	birm	Jefferson	AL	651,525	16.2	8.5	7.7
Washington	dc	Washington DC	DC	606,900	15.5	7.0	8.5
Oklahoma City	okla	Oklahoma	OK	599,611	12.9	7.3	5.6
Providence	prov	Providence	RI	596,270	14.6	7.9	6.7
El Paso	elpa	El Paso	TX	591,610	7.7	3.8	3.9
Tacoma	taco	Pierce	WA	586,203	10.0	5.7	4.3
Austin	aust	Travis	TX	576,407	7.0	3.4	3.6
Dayton	dayt	Montgomery	OH	573,809	11.9	6.5	5.4
Jersey City	jers	Hudson	NJ	553,099	11.5	5.9	5.6
Bakersfield	bake	Kern	CA	543,477	8.6	5.0	3.6
Akron	akr	Summit	OH	514,990	10.7	5.8	4.9
Charlotte	char	Mecklenburg	NC	511,433	8.5	4.3	4.2
Nashville	nash	Davidson	TN	510,784	11.0	6.0	5.0
Tulsa	tuls	Tulsa	OK	503,341	10.0	5.8	4.2
Grand Rapids	gdrp	Kent	MI	500,631	8.7	4.9	3.8
New Orleans	no	Orleans	LA	496,938	12.0	5.9	6.1
Stockton	stoc	San Joaquin	CA	480,628	8.5	4.8	3.6
Albuquerque	albu	Bernalillo	NM	480,577	7.6	3.8	3.8
Syracuse	syra	Onondaga	NY	468,973	9.7	5.4	4.3
Toledo	tole	Lucas	OH	462,361	10.8	6.3	4.5
Raleigh	ral	Wake	NC	423,380	5.6	2.9	2.7
Wichita	wich	Sedwick	KS	403,662	7.2	4.0	3.3
Colorado Springs	colo	El Paso	CO	397,014	5.0	2.8	2.3
Baton Rouge	batr	East Baton Rouge	LA	380,105	6.3	3.4	3.0
Modesto	mode	Stanislaus	CA	370,522	6.6	3.8	2.8
Madison	madi	Dane	WI	367,085	5.3	2.9	2.4
Spokane	spok	Spokane	WA	361,364	7.8	4.5	3.3
Little Rock	ltrk	Pulaski	AR	349,660	7.0	3.7	3.3
Greensboro	grnb	Guilford	NC	347,420	6.9	3.8	3.1
Knoxville	knox	Knox	TN	335,749	6.7	3.5	3.1
Shreveport	shr	Bossier, Caddo	LA	334,341	6.8	3.7	3.1
Des Moines	desm	Polk	IA	327,140	6.1	3.4	2.6

(Table continues on next page.)

Table A.1, *continued*. The 90 Cities and their Included Counties by Population Size with Mean Daily Number of Deaths by Category, 1987–1994

City	Abbreviation	County	State	Population	Mean Daily Deaths		
					Total	CVD/ Respiratory Disease	Other
Fort Wayne	ftwa	Allen	IN	300,836	5.9	3.4	2.5
Corpus Christi	corp	Nueces	TX	291,145	4.9	2.5	2.4
Norfolk	nor	Norfolk	VA	261,229	4.8	2.6	2.2
Jackson	jcks	Hinds	MS	254,441	5.3	3.0	2.3
Huntsville	hunt	Madison	AL	238,912	3.9	2.2	1.7
Anchorage	anch	Anchorage	AK	226,338	1.9	0.8	1.1
Lexington	lex	Fayette	KY	225,366	4.1	2.1	2.0
Lubbock	lubb	Lubbock	TX	222,636	3.9	2.3	1.6
Richmond	rich	Richmond City	VA	203,056	5.1	2.7	2.4
Arlington	arlv	Arlington	VA	170,936	2.4	1.3	1.2
Kingston	king	Ulster	NY	165,304	3.0	1.8	1.2
Evansville	evan	Vanderburgh	IN	165,058	4.4	2.5	1.9
Kansas City	kans	Wyandotte	KS	161,993	3.2	1.8	1.4
Olympia	olym	Thurston	WA	161,238	2.8	1.5	1.3
Topeka	tope	Shawnee	KS	160,976	3.6	2.0	1.6

Table A.2. Regression Coefficients (95% CI) for Effect of Lag1 PM_{10} on Total Mortality, Cardiorespiratory and Other Deaths, Unadjusted for Other Pollutants, in 20 Cities

City	Total Mortality β (95% CI)	CVDRESP Mortality β (95% CI)	Other Diseases β (95% CI)
Los Angeles	0.38 (0.01, 0.76)	0.32 (–0.17, 0.81)	0.33 (–0.22, 0.88)
New York	1.11 (0.53, 1.68)	1.12 (0.38, 1.86)	0.93 (0.06, 1.81)
Chicago	0.31 (0.10, 0.52)	0.40 (0.12, 0.69)	0.23 (–0.07, 0.54)
Dallas/Ft Worth	–0.41 (–1.67, 0.85)	–0.09 (–1.74, 1.55)	–0.64 (–2.51, 1.23)
Houston	0.19 (–0.47, 0.84)	0.55 (–0.35, 1.45)	–0.21 (–1.13, 0.71)
San Diego	1.09 (0.16, 2.03)	0.80 (–0.44, 2.04)	1.26 (–0.11, 2.63)
Santa Ana/Anaheim	0.69 (–0.35, 1.74)	0.15 (–1.23, 1.53)	0.53 (–1.08, 2.14)
Phoenix	0.65 (–0.42, 1.72)	1.21 (–0.22, 2.63)	0.16 (–1.40, 1.72)
Detroit	0.48 (0.09, 0.86)	0.60 (0.09, 1.10)	0.10 (–0.48, 0.68)
Miami	0.69 (–0.76, 2.15)	1.95 (–0.05, 3.94)	–0.72 (–2.87, 1.44)
Philadelphia	0.77 (–0.19, 1.73)	1.37 (0.05, 2.69)	0.16 (–1.28, 1.60)
Minneapolis/St Paul	0.48 (–0.07, 1.03)	0.48 (–0.26, 1.23)	0.32 (–0.48, 1.12)
Seattle	0.28 (–0.31, 0.87)	0.48 (–0.30, 1.27)	0.16 (–0.68, 1.01)
San Jose	0.32 (–0.35, 0.98)	0.19 (–0.69, 1.07)	0.44 (–0.54, 1.42)
Cleveland	–0.05 (–0.50, 0.40)	–0.10 (–0.71, 0.50)	0.03 (–0.62, 0.68)
San Bernardino	0.26 (–1.10, 1.61)	1.33 (–0.41, 3.06)	–1.16 (–3.25, 0.93)
Pittsburgh	0.39 (0.09, 0.69)	0.57 (0.17, 0.97)	0.19 (–0.27, 0.65)
Oakland	2.06 (0.94, 3.18)	2.03 (0.58, 3.48)	1.73 (0.08, 3.39)
Atlanta	0.05 (–1.60, 1.70)	–0.32 (–2.69, 2.04)	0.44 (–1.86, 2.74)
San Antonio	0.69 (–1.08, 2.47)	2.32 (–0.02, 4.66)	–0.84 (–3.46, 1.79)

Table A.3. Mean Daily Pollution[a] by City, 1987–1994

City	PM_{10} μg/m^3	O_3 ppb	NO_2 ppb	SO_2 ppb	CO ppm
Los Angeles	46.0	22.8	39.4	1.9	1.51
New York	28.8	19.6	38.9	12.8	2.04
Chicago	35.6	18.6	24.3	4.6	0.79
Dallas/Ft Worth	23.8	25.3	13.8	1.1	0.74
Houston	30.0	20.5	18.8	2.8	0.89
San Diego	33.6	31.6	22.9	1.7	1.10
Santa Ana/Anaheim	37.4	23.0	35.1	1.3	1.23
Phoenix	40.3	22.5	16.6	3.5	1.27
Detroit	40.9	22.6	21.3	6.4	0.66
Miami	25.7	25.9	11.0	NA	1.06
Philadelphia	35.4	20.5	32.2	9.9	1.18
Minneapolis/St Paul	26.9	NA	17.6	2.6	1.18
Seattle	25.3	19.4	NA	NA	1.78
San Jose	30.4	17.9	25.1	NA	0.94
Cleveland	45.1	27.4	25.2	10.3	0.85
San Bernardino	37.0	35.9	27.9	0.7	1.03
Pittsburgh	31.6	20.7	27.6	14.2	1.22
Oakland	26.3	17.2	21.2	NA	0.91
Atlanta	36.1	25.1	26.0	6.0	0.89
San Antonio	23.8	22.2	NA	NA	1.01
Riverside	52.0	33.4	25.0	0.4	1.12
Denver	29.6	21.4	27.9	5.5	1.03
Sacramento	33.3	26.7	16.3	NA	0.94
St Louis	30.1	22.8	22.5	11.3	1.05
Buffalo	21.7	22.9	19.0	8.6	0.73
Columbus	29.0	26.0	NA	5.9	0.76
Cincinnati	34.2	25.8	26.7	11.9	1.00
St Petersburg	23.5	24.6	11.8	NA	0.71
Kansas City	25.9	27.6	9.2	2.4	0.62
Honolulu	15.3	18.9	NA	NA	0.83
Tampa	28.3	23.5	21.2	7.8	0.78
Memphis	30.3	29.0	26.8	6.8	1.19
Indianapolis	32.0	31.9	20.2	7.7	0.90
Newark	32.9	15.2	33.6	9.6	0.87
Baltimore	32.9	21.2	32.9	8.4	0.92
Salt Lake City	32.9	23.0	29.6	4.4	1.35
Rochester	21.9	22.7	NA	10.4	0.63
Worcester	22.2	30.0	25.2	6.7	0.89
Orlando	22.7	24.1	11.4	1.5	0.93
Jacksonville	29.9	28.2	14.8	2.2	0.92
Fresno	43.4	29.4	21.7	1.9	0.68
Louisville	30.8	19.8	22.4	8.4	1.12
Boston	26.0	17.9	29.9	10.0	1.13
Birmingham	31.2	22.4	NA	6.6	1.05
Washington DC	28.2	17.5	25.6	11.2	1.23

(Table continues in next column.)

Table A.3, *continued*.

City	PM_{10} μg/m^3	O_3 ppb	NO_2 ppb	SO_2 ppb	CO ppm
Oklahoma City	25.0	28.4	13.9	NA	0.71
Providence	30.9	25.4	21.9	9.5	1.00
El Paso	41.2	24.4	23.6	9.1	1.25
Tacoma	28.0	23.8	NA	6.5	1.66
Austin	21.1	25.5	NA	NA	NA
Dayton	27.4	26.6	NA	NA	0.82
Jersey City	30.5	19.7	28.7	10.7	2.01
Bakersfield	53.2	33.3	19.4	3.0	1.05
Akron	22.4	30.5	NA	12.0	0.70
Charlotte	30.7	29.3	16.2	NA	1.11
Nashville	32.4	16.2	NA	11.6	1.12
Tulsa	26.6	31.4	16.6	6.9	0.65
Grand Rapids	22.8	27.7	NA	3.0	0.57
New Orleans	29.0	20.5	21.3	NA	0.94
Stockton	39.0	22.6	24.2	1.7	0.82
Albuquerque	16.9	25.8	NA	NA	0.79
Syracuse	24.5	23.7	NA	3.6	1.17
Toledo	25.6	27.1	NA	5.9	1.03
Raleigh	25.6	35.4	12.7	NA	1.61
Wichita	25.6	24.2	NA	4.8	0.65
Colorado Springs	26.3	24.3	NA	NA	1.09
Baton Rouge	27.3	21.2	16.6	5.2	0.43
Modesto	41.7	26.1	24.2	1.9	0.91
Madison	19.9	29.7	NA	3.3	1.04
Spokane	36.0	32.6	NA	NA	2.19
Little Rock	25.8	27.7	9.3	2.6	NA
Greensboro	27.5	NA	NA	4.2	1.22
Knoxville	36.3	29.6	NA	NA	1.36
Shreveport	24.7	28.2	NA	2.3	NA
Des Moines	35.5	11.8	NA	NA	0.86
Fort Wayne	23.2	32.1	NA	4.0	1.44
Corpus Christi	24.7	23.9	NA	1.0	NA
Norfolk	26.0	NA	19.6	6.7	0.73
Jackon	26.4	23.9	NA	NA	0.79
Huntsville	26.0	30.4	12.9	NA	0.63
Anchorage	23.0	NA	NA	NA	1.61
Lexington	24.5	32.8	16.4	6.2	0.88
Lubbock	25.1	NA	NA	NA	NA
Richmond	25.4	NA	23.7	5.8	0.66
Arlington	22.0	29.0	25.5	NA	0.66
Kingston	20.4	NA	NA	NA	NA
Evansville	32.4	NA	NA	NA	NA
Kansas City	43.4	18.5	17.6	4.7	0.82
Olympia	22.7	NA	NA	NA	1.27
Topeka	29.0	NA	NA	NA	NA

[a] Pollution values are based on a 10% trimmed, mean as described in Appendix E.

Table A.4. Number of Days for Which Monitoring Is Available by Pollutant and Used for Mortality Analysis, 90 Cities, 1987–1994

City	PM_{10}	O_3	NO_2	SO_2	CO
Los Angeles	580	2,922	2,922	2,922	2,922
New York	489	2,922	2,493	2,920	2,920
Chicago	2,683	2,922	2,922	1,409	2,922
Dallas/Ft Worth	624	2,922	2,557	2,908	2,922
Houston	793	2,922	2,557	2,922	2,922
San Diego	521	2,922	2,922	2,922	2,922
Santa Ana/Anaheim	480	2,922	2,922	2,922	2,922
Phoenix	376	2,554	740	1,272	2,554
Detroit	1,348	1,861	2,686	2,922	2,922
Miami	484	2,882	2,863	0	2,919
Philadelphia	495	2,901	2,554	2,919	2,919
Minneapolis/St Paul	2,764	0	2,725	2,914	2,918
Seattle	2,205	1,820	0	0	2,922
San Jose	945	2,922	1,957	0	2,922
Cleveland	1,298	1,712	2,555	2,922	2,897
San Bernardino	538	2,922	2,922	2,922	2,922
Pittsburgh	2,899	2,883	2,537	2,922	2,920
Oakland	508	2,922	2,921	0	2,922
Atlanta	482	2,200	2,922	2,918	2,839
San Antonio	670	2,918	0	0	2,891
Riverside	545	2,922	2,904	2,908	2,921
Denver	1,645	2,922	2,484	2,860	2,922
Sacramento	488	2,922	2,916	0	2,922
St Louis	487	1,731	2,919	2,919	2,920
Buffalo	489	2,884	2,522	2,922	2,921
Columbus	1,564	1,494	0	964	2,557
Cincinnati	1,705	1,712	2,554	2,905	2,922
St Petersburg	367	2,920	2,235	0	2,922
Kansas City	670	2,856	2,922	1,094	2,922
Honolulu	415	1,681	0	0	2,919
Tampa	508	2,922	941	1,818	2,922
Memphis	480	1,707	2,254	2,823	2,922
Indianapolis	1,269	1,588	2,874	2,922	2,922
Newark	484	2,726	2,882	2,896	2,894
Baltimore	1,220	2,063	2,843	2,912	2,865
Salt Lake City	1,356	2,409	1,903	2,739	2,922
Rochester	486	2,886	0	2,921	2,921
Worcester	450	1,763	2,864	2,452	2,899
Orlando	421	2,920	2,024	2,878	2,921
Jacksonville	555	2,791	2,727	2,738	2,922
Fresno	517	2,922	2,922	2,398	2,922
Louisville	485	2,603	1,604	2,841	2,922
Boston	631	2,882	2,922	2,922	2,922
Birmingham	900	2,200	0	1,916	2,922
Washington DC	417	2,847	2,842	2,286	2,341

Table A.4, *continued*.

City	PM_{10}	O_3	NO_2	SO_2	CO
Oklahoma City	563	2,832	2,295	0	2,909
Providence	485	1,634	2,441	2,922	2,921
El Paso	2,587	2,922	2,472	2,906	2,922
Tacoma	482	1,601	0	2,756	2,766
Austin	646	2,909	0	0	0
Dayton	461	1,696	0	0	2,922
Jersey City	1,367	2,843	2,496	2,918	2,883
Bakersfield	550	2,557	2,557	2,557	2,659
Akron	1495	1,677	0	2,827	2,922
Charlotte	454	1,936	1,593	0	2,922
Nashville	1,989	2,861	0	2,619	2,771
Tulsa	411	2,834	2,462	2,426	2,836
Grand Rapids	777	1,615	0	2,907	2,903
New Orleans	531	2,889	2,879	0	2,922
Stockton	488	2,475	2,379	867	2,906
Albuquerque	1,200	2,922	0	0	2,922
Syracuse	485	2,864	0	2,857	2,908
Toledo	416	1,711	0	2,921	2,897
Raleigh	480	1,267	1,219	0	2,160
Wichita	366	2,913	0	1,423	2,922
Colorado Springs	481	2,920	0	0	2,922
Baton Rouge	474	2,922	2,880	2,891	2,888
Modesto	199	2,496	2,449	845	2,892
Madison	338	1,698	0	2,432	2,709
Spokane	2,393	974	0	0	2,922
Little Rock	516	2,922	2,921	2,908	0
Greensboro	445	0	0	1,077	1,855
Knoxville	577	1,679	0	0	2,511
Shreveport	349	2,922	0	2,881	0
Des Moines	1,334	2,782	0	0	2,825
Fort Wayne	336	1,587	0	1,219	1,822
Corpus Christi	613	2,919	0	2,920	0
Norfolk	474	0	1,787	2,148	2,921
Jackon	508	2,191	0	0	2,574
Huntsville	1,382	2,173	1,090	0	2,532
Anchorage	2,379	0	0	0	1,488
Lexington	816	1,709	2,871	2,906	2,865
Lubbock	1,306	0	0	0	0
Richmond	474	0	2,537	2,907	2,922
Arlington	313	1,705	2,306	0	2,896
Kingston	323	0	0	0	0
Evansville	404	0	0	0	0
Kansas City	551	2,890	324	2,909	2,775
Olympia	1,135	0	0	0	950
Topeka	269	0	0	0	0

Table A.5. Regression Coefficients (95% CI) for Effect of Lag1 PM_{10} on Total Mortality Adjusted for O_3, in 20 Cities

City	Estimate (95% CI)
Los Angeles	0.42 (0.05, 0.79)
New York	1.19 (0.59, 1.80)
Chicago	0.31 (0.10, 0.52)
Dallas/Ft Worth	–0.32 (–1.58, 0.94)
Houston	0.21 (–0.44, 0.87)
San Diego	1.13 (0.19, 2.06)
Santa Ana/Anaheim	1.03 (–0.01, 2.07)
Phoenix	0.63 (–0.47, 1.73)
Detroit	0.60 (0.14, 1.06)
Miami	0.59 (–0.90, 2.07)
Philadelphia	0.68 (–0.30, 1.66)
Minneapolis/St Paul	NA
Seattle	0.98 (0.00, 1.96)
San Jose	0.37 (–0.32, 1.06)
Cleveland	0.05 (–0.54, 0.64)
San Bernardino	0.31 (–1.04, 1.67)
Pittsburgh	0.38 (0.07, 0.68)
Oakland	2.02 (0.86, 3.17)
Atlanta	–1.02 (–3.04, 1.01)
San Antonio	0.71 (–1.07, 2.48)

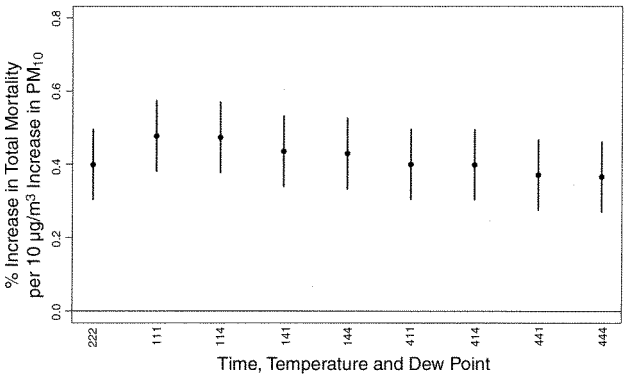


Figure A.1. Overall means and 95% CIs for the estimated effects of PM_{10} on total mortality at lag 1 for the 90 cities, under 9 different modeling assumptions. Each overall result was obtained using a random effects model. The nine labels on the x-axis indicate whether the degrees of freedom for the smooth function of time, temperature, and dew point, taken in that order, have been halved, doubled, or kept the same. Half is represented by a "1", double by a "4", and default by a "2".

Table A.6. Regression Coefficients[a] (95% CI) for Effect of Lag1 PM_{10} on Total Mortality Adjusted for O_3 and NO_2; O_3 and SO_2; O_3 and CO, in Selected Cities

City	O_3 and NO_2 Mean (95% CI)	O_3 and SO_2 Mean (95% CI)	O_3 and CO Mean (95% CI)
Los Angeles	–0.06 (–0.64, 0.51)	–0.01 (–0.47, 0.44)	–0.05 (–0.60, 0.50)
New York	0.52 (–0.34, 1.37)	0.21 (–0.88, 1.29)	1.51 (–0.65, 2.36)
Chicago	0.13 (–0.10, 0.37)	–0.03 (–0.34, 0.29)	0.16 (–0.06, 0.38)
Dallas/Ft Worth	0.66 (–0.75, 2.07)	–0.22 (–1.48, 1.05)	–0.09 (–1.36, 1.17)
Houston	0.01 (–0.67, 0.69)	0.13 (–0.54, 0.79)	0.04 (–0.66, 0.75)
San Diego	0.81 (–0.46, 2.08)	0.41 (–0.63, 1.45)	0.79 (–0.44, 2.01)
Santa Ana/Anaheim	0.38 (–1.12, 1.89)	0.95 (–0.22, 2.11)	1.02 (–0.41, 2.45)
Phoenix	1.95 (–0.39, 4.29)	1.72 (0.14, 3.31)	-0.07 (–1.42, 1.29)
Detroit	0.49 (–0.17, 1.14)	0.65 (0.12, 1.17)	0.42 (–0.10, 0.94)
Miami	0.65 (–0.86, 2.16)	NA	0.62 (–0.88, 2.12)
Philadelphia	1.26 (–0.02, 2.55)	1.07 (–0.31, 2.46)	0.96 (–0.16, 2.08)
Minneapolis/St Paul	NA	NA	NA
Seattle	NA	NA	1.39 (–0.06, 2.83)
San Jose	–0.29 (–1.73, 1.15)	NA	-0.15 (–1.46, 1.15)
Cleveland	0.07 (–0.54, 0.69)	0.04 (–0.55, 0.64)	0.04 (–0.55, 0.64)
San Bernardino	0.19 (–1.35, 1.72)	0.32 (–1.06, 1.69)	0.27 (–1.24, 1.79)
Pittsburgh	0.26 (–0.12, 0.65)	0.43 (0.05, 0.80)	0.30 (–0.07, 0.67)
Oakland	1.23 (–0.36, 2.83)	NA	1.12 (–0.39, 2.64)
Atlanta	–1.12 (–3.36, 1.11)	–0.95 (–3.15, 1.24)	–0.70 (–3.17, 1.17)
San Antonio	NA	NA	0.78 (–1.00, 2.56)

[a] Percentage of change per 10 µg/m^3 PM_{10}.

APPENDIX B. Distributed Lag Between Air Pollution and Daily Deaths*

ABSTRACT

Many studies have reported associations between air pollution and daily deaths. Those studies have not consistently specified the lag between exposure and response, although most have found associations that persisted for more than 1 day. A systematic approach to specifying the lag association would allow better comparison across sites and give insight into the nature of the relationship. To examine this question, I fit unconstrained and constrained distributed lag relationships to the association between daily deaths of persons aged 65 and older with PM_{10} in 10 US cities (New Haven, Birmingham, Pittsburgh, Detroit, Canton, Chicago, Minneapolis/St Paul, Colorado Springs, Spokane, and Seattle) that had daily monitoring for PM_{10}.

After control for temperature, humidity, barometric pressure, day of the week, and seasonal patterns, I found evidence in each city that the effect of a single day's exposure to PM_{10} was manifested across several days. Averaging over the 10 cities, the overall effect of an increase in exposure of 10 $\mu g/m^3$ on a single day was an estimated 1.4% increase in deaths (95% CI 1.15%, 1.68%) obtained when using a quadratic distributed lag model, and an estimated 1.3% increase (CI 1.04%, 1.56%), obtained when using an unconstrained distributed lag model. In contrast, constraining the model to assume that the effect occurs all in a single day resulted in an estimate of only 0.65% (CI 0.49%, 0.81%), indicating that this constraint leads to a substantially lower estimate of effect. Combining the estimated effect at each day's lag across the 10 cities showed that the effect was spread over several days and did not reach 0 until 5 days after the exposure. Given the distribution of sensitivities likely in the general population, this result is biologically plausible. I also found a protective effect of barometric pressure in all 10 locations.

INTRODUCTION

Numerous studies in the last decade have reported associations between day-to-day fluctuations in air pollution and day-to-day fluctuations in daily deaths in cities (Schwartz and Marcus 1990; Schwartz and Dockery 1992; Touloumi et al 1994; Hoek et al 1997; Katsouyanni et al 1997; Kelsall et al 1997). These associations have been seen with pollution levels on the same day and within 5 previous days (Pope et al 1992). The studies have been criticized, however, because the same lag relationship between air pollution and deaths was not used in all studies. Instead, the strongest single-day lag relationship with mortality has usually been fit in each study, leading to inconsistencies in the lag relationship specified. A more consistent approach would facilitate combining evidence across studies.

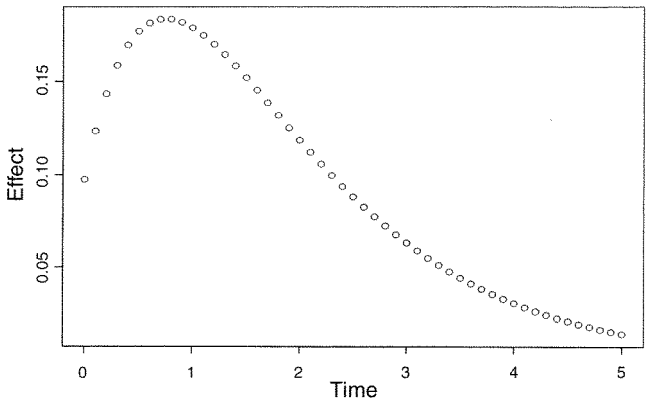


Figure B.1. Hypothesized curve showing the impact of an environmental toxin over time. The effect rises, and then falls, possibly with a long tail. The goal of this analysis is to identify the actual shape of the curve representing the time course of deaths after exposure to PM_{10}.

Most of the studies that considered associations with a multiday moving average found that a 2-day or 3-day moving average of air pollution has fit better than any single day's pollution or longer moving average (Schwartz and Dockery 1992; Katsouyanni et al 1997; Kelsall et al 1997). In some cases, longer (up to 5 days) moving averages have been fit. This suggests that the effect of an increase in pollution concentration on a single day is distributed across several subsequent days. Toxicologic data also suggest that effects of exposure may be seen over several subsequent days. For example, Clarke and coworkers (1999) reported changes in tidal volume of rats immediately after exposure to concentrated ambient air particles, and increases in inflammatory markers approximately 36 hours after exposure.

In general, we might suppose the distribution of effect of air pollution over time to look something like Figure B.1. If the time scale in this figure is short (hours, for example), then a single 24-hour average pollution concentration will be an acceptable proxy for the true relationship shown in the figure. If the time scale extends over several days, a 2-day moving average will probably be a reasonable proxy, and a single day's pollution will be a poorer surrogate. In both those examples, we are approximating the true shape of the relationship by a simple step function. If the relationship extends over a longer time scale, then the extent to

* Much of this appendix was taken from Schwartz J. 2000. The distributed lag between air pollution and daily deaths. Epidemiology 11:320–326.

which a 1-day or 2-day average concentration will represent the true relationship depends on the serial correlation in the air pollution measure. Air pollution concentration measurements taken within a few days of each other tend to be correlated. That correlation varies from city to city, however. Thus, some of the variation in the effect sizes and best-fitting lags of the reported associations among cities may be due to fitting a simplistic model for the distribution of the exposure response over time.

This issue is increasing in importance. Previous studies of the association between daily pollution and daily deaths have been done in cities with daily monitoring. Recent studies have begun to examine large numbers of cities in which airborne particles are monitored only 1 day in 6 (Dominici et al 1999). Two-day or three-day averages cannot be used in those cities, and it is important to quantify the impact of this limitation.

A systematic approach to investigating the distribution of effect over time offers the possibilities of explaining some of the variation, of testing the potential bias associated with the use of a single day's pollution, of providing estimates in individual locations that can more appropriately be combined, and of indicating what the nature of the lag structure between air pollution and daily deaths is. This study demonstrates the methodology for such an approach and applies it to a study of PM_{10} pollution and daily deaths in 10 cities across the United States.

DATA AND METHODS

To analyze effectively the distributed lag between PM_{10} and daily deaths, we need daily PM_{10} measurements. Most US cities measured PM_{10} only 1 day in 6, but a number of locations had monitors on a daily schedule. I selected 10 US cities with roughly daily PM_{10} monitoring to provide a reasonable number of locations for a combined analysis. The cities were New Haven, Birmingham, Pittsburgh, Detroit, Canton, Chicago, Minneapolis/St Paul, Colorado Springs, Spokane, and Seattle. Daily deaths in the metropolitan county containing each city were extracted from NCHS mortality tapes for the years 1986 through 1993. Deaths due to external causes (ICD-9 800–999) were excluded. Because a previous study suggested the seasonal pattern of mortality may differ by age (Samet et al 1998), this study is limited to deaths of persons aged 65 and older. Minneapolis and St Paul were combined and treated as a single city. Daily weather data were obtained from the nearest airport weather station, and daily concentrations of PM_{10} were obtained from the EPA's AIRS monitoring network.

The assignment of PM_{10} exposure raised a number of issues. Many of the locations have more than 1 monitoring location, but typically only 1 monitor operates on a daily basis, with the others operating every third or sixth day. If the monitors were simply averaged, the daily mean would jump on days when new monitors were included merely because their annual average differs from the monitoring station that operates on a daily basis.

The variance of PM_{10} measurements also can differ from monitoring location to location. Day-to-day changes in which monitors are in the daily average would also result in changes in the day-to-day exposure measure that represent not true changes in exposure but only changes in the sampling of monitors. To remove these influences, I used the following algorithm. The annual mean was computed for each monitor for each year and subtracted from the daily values of that monitor. I then standardized these daily deviances from each monitor's annual average by dividing by the standard deviation for that monitor. The daily standardized deviations for each monitor on each day were averaged, producing a daily averaged standardized deviation. I multiplied this by the standard deviation of all of the monitor readings for the entire year and added back in the annual average of all of the monitors.

Analytical Approach

For each city, a generalized additive Poisson regression was fit, modeling the logarithm of the expected value of daily deaths as a sum of smooth functions of the predictor variables (Hastie and Tibshirani 1990; Schwartz et al 1993). The generalized additive model allows regressions to include nonparametric smooth functions to model the potential nonlinear dependence of daily admissions on weather and season. It assumes that

$$\log(E(Y)) = \beta_0 \; S_1 \ldots \; S_p(X_p)$$

where Y is the daily count of deaths, $E(Y)$ is the expected value of that count, the X_i are the covariates and the S_i are the smooth (ie, continuously differentiable) functions for $i = 1, \ldots, p$. For the S_i, I used LOESS, a moving regression smoother (Cleveland and Devlin 1988). This approach is now standard in air pollution time-series studies (Schwartz 1994). For each covariate, it is necessary to choose a smoothing parameter that determines how smooth the function of that covariate should be. Three sets of predictor variables were used in this analysis: a smooth function of time to capture seasonal and other long-term trends in the data, weather and day-of-the-week variables to capture shorter-term potential confounding, and PM_{10}. The choice of smoothing parameter for each set of variables is described as follows.

The purpose of the smooth function of time is to remove the basic long-term pattern from the data. Seasonal patterns can vary greatly between Birmingham and Spokane, for example, and a separate smoothing parameter was chosen for each city to reduce the residuals of the regression to white noise (ie, to remove serial correlation) (Schwartz 1999). This approach was used because each death is an independent event, and autocorrelation in residuals indicates there are omitted, time-dependent covariates, the variation of which may confound air pollution. If the autocorrelation is removed, remaining variation in omitted covariates has no systematic temporal pattern, and hence confounding is less likely. Sometimes it was necessary to incorporate autoregressive terms to eliminate serial correlation from the residuals (Brumback et al 1999).

The other covariates were temperature, relative humidity, and barometric pressure on the same day; the previous day's temperature; and day of the week. To allow for city-specific differences, the smoothing parameters for these covariates were also optimized separately in each location. The criterion used was to choose the smoothing parameter for each variable that minimized AIC (Akaike 1973).

PM_{10} was treated as having a linear association in this analysis to facilitate the combination of coefficients across cities and the examination of lag structure. Robust regression was used to reduce sensitivity to outliers in the dependent variable. These regressions were done using the generalized additive model function in Splus, and M-estimation was the robust regression method. To reduce sensitivity to outliers in the pollution variable, analysis was restricted to days when PM_{10} levels were below 150 $\mu g/m^3$, the currently enforced ambient standard. This also ensures that the results are unambiguously relevant to questions of revision of those standards.

Distributed Lag Models

Distributed lag models have been used for decades in the social sciences (Judge et al 1980), and Pope and Schwartz (1996) recently described the use of this approach in epidemiology.

Motivation for the distributed lag model is the realization that air pollution can effect deaths occurring not only on the same day, but also on several subsequent days. The converse is therefore also true: Deaths today will depend on the same-day effect of today's pollution levels, the 1-day lag effects of yesterday's PM_{10} concentrations, and so on. Therefore, suppressing covariates and assuming Gaussian data for the moment, the unconstrained distributed lag model assumes

$$Y_t = \alpha + \beta_0 X_t + \ldots + \beta_q X_{t-q} + \varepsilon_1 \quad (1)$$

where X_{t-q} is the PM_{10} concentration q days before the deaths. The overall effect of a unit increase in air pollution on a single day is its impact on that day plus its impact on subsequent days. That is, it is the sum of $\beta + \ldots + \beta_q$ (Schwartz 1993). To see this more easily, note that equation (1) can be recast as:

$$Y_t = \alpha + \beta^*(w_0 X_t + \ldots + w_q X_{t-q}) + \varepsilon_t$$

where the w_i are individual weights that sum to 1, and β^* is $\beta_0 + \ldots + \beta_q$. That is, β^* is also interpretable as the marginal effect of a unit increase in a weighted average pollution variable. Since a unit increase in pollution on a single day increases the weighted average on all q subsequent days, the effect of that single day's increase will be $\beta^* w_i$ on each of the q subsequent days, or β^* overall.

Since there is substantial correlation between air pollution concentrations on days close together, the above regression will have a high degree of collinearity. This will result in unstable estimates of the individual β_q and hence poor estimates of the shape of Figure B.1. However, the sum of the individual β_q will be an unbiased estimate of the overall effect of a unit increase in pollution, albeit an inefficient one.

To gain more efficiency, and more insight into the shape of the distributed effect of air pollution over time, it is useful to constrain β_q. If this is done flexibly, substantial gains in reducing the noise of the unconstrained distributed lag model can be obtained with minimal bias. This approach has been widely applied in the social sciences, using Gaussian data. The most common approach is to constrain the shape of the variation of each β_q with lag number (ie, the shape of Figure B.1) to fit some polynomial function. That is, the polynomial distributed lag model (PDL(q, d) with q lags and degree d) is the model (1) above, subject to the restriction:

$$\beta_j = \sum_0^d \eta_k j^k \quad (2)$$

This approach originated with Almon (1965). Here I extend that model to the generalized additive model case. Assume that

$$\log(E(Y)) = \alpha_0 + S_1(X_1) + \ldots S_p(X_p) + \beta_0 Z_0 + \ldots \beta_p Z_p$$

where Z_0 is the exposure on the concurrent day, Z_1 on the previous day, and so on. If we impose the constraints in equation (2) and suppress the covariates, we can write this as

$$\log(E(Y)) = covariates + \eta_0 Z_0 + (\eta_0 + \eta_1 + 2\eta_2 + p\eta_p)Z_1 + (\eta_0 + p\eta_1 + p^2\eta_2 + p^d\eta_d)Z_1$$

We can rewrite this by collecting terms in each of the θs. This yields:

$$\log(E(Y)) = covariates + \eta_0(Z_0 + Z_1 + \ldots + Z_p) + \ldots + \eta_p(pZ_1 + p^2Z_2 + \ldots + p^dZ_d)$$

Hence, if we define $d + 1$ new variables W_d to be weighted sums of the exposure variable Z and its lags, with

$$W_d = pZ_1 + p^2Z_2 + \ldots + p^dZ_d$$

and

$$W_0 = Z_0 + Z_1 + \ldots + Z_p,$$

we can estimate the model

$$\log(E(Y)) = covariates + \eta_0 W_0 + \ldots + \eta_p W_p$$

and the coefficients of the Ws will be the parameters of the polynomial distributed lag.

Note that the use of a single day's exposure is also a constrained lag model. In that case, we are fitting equation (1), with the constraint that $\beta_1 = \beta_2 = \ldots = \beta_p = 0$. If we are not quite sure that the effects of pollution are limited to a single day, these constraints are much more restrictive than those of equation (2) and are much more likely to introduce bias into the estimated overall effect.

I have chosen a maximum lag of 5 days prior to the deaths for the air pollution variable. This is because published studies have shown that lags of more than a few days had little correlation with daily deaths, and because the goal of this analysis is to estimate the short-term effects of air pollution. Because the explanatory power of air pollution on daily deaths is modest, parsimony in the degree of the polynomial is necessary. Therefore, I have chosen a second-degree polynomial in this analysis. To test the sensitivity of the conclusions regarding the overall effect of air pollution on daily deaths to the degree of the polynomial, I have also fit the unconstrained distributed lag model in each city. To see if the traditional approaches of using a 1-day or 2-day moving average resulted in a downward bias in the estimated effects of PM_{10}, I have also fit those models in each city.

To combine results across cities, I used inverse variance weighted averages. For the distributed lag model, the effect at each day's lag (and its variance) was estimated from the parameters of the polynomial (ie, θ_0, θ_1, and θ_2) and their covariances, as was the overall effect and its variance.

RESULTS

Table B.1 shows the populations, mean daily deaths of people aged 65 and over, and means of the environmental variables in the 10 study locations. Table B.2 shows the correlation between PM_{10} and the weather variables. The correlation between PM_{10} and barometric pressure was quite small and mixed in sign (positive and negative). The correlation between PM_{10} and relative humidity was generally negative and moderately low. The correlation with temperature

Table B.1. Population and Mean Daily Values for Environmental Variables and Mean Daily Death Counts for Each City (1986–1994)

City	1990 Population	Deaths (ages 65+)	PM_{10} ($\mu g/m^3$)	Relative Humidity (%)	Barometric Pressure (inches H_2O)	Temperature (°F)	Days of PM_{10} Monitoring
New Haven	804,219	15.5	28.6	67.2	29.8	50.5	1,450
Birmingham	651, 525	13.8	34.8	70.5	29.4	62.4	2,485
Pittsburgh	1,336,449	32.8	36.4	69.3	28.7	52.1	2,920
Detroit	2,111,687	38.8	36.9	69.2	29.3	50.9	2,705
Canton	367,585	7.4	29.3	73.3	28.7	50.4	1,750
Chicago	5,105,067	90.2	36.5	70.9	29.3	50.3	2,149
Minneapolis/St Paul	1,518,196	24.5	27.5	68.7	29.1	46.3	2,489
Colorado Springs	397,014	3.9	27.1	52.6	24.0	48.9	2,344
Spokane	361,364	6.6	40.6	66.6	27.5	47.9	1,977
Seattle	1,507,319	21.2	32.5	75.5	29.6	52.5	2,915

Table B.2. Correlation Between PM_{10} and Other Environmental Variables

City	Barometric Pressure	Relative Humidity	Temperature
New Haven	0.11	–0.15	0.05
Birmingham	0.12	–0.30	0.26
Pittsburgh	0.14	–0.23	0.45
Detroit	–0.05	–0.14	0.37
Canton	0.15	–0.16	0.42
Chicago	–0.02	–0.30	0.36
Minneapolis/ St Paul	–0.03	–0.35	0.29
Colorado Springs	–0.01	–0.11	–0.34
Spokane	0.16	–0.19	–0.01
Seattle	0.24	–0.11	–0.22

varied considerably across the locations, ranging from –0.34 in Colorado Springs to 0.45 in Pittsburgh.

Weather Results

The most notable result using weather variables is the consistent association between barometric pressure and daily deaths. In every city, higher barometric pressure was associated with fewer deaths. In 6 of the cities, AIC (which trades off improvement in model fit versus the number of degrees of freedom that produced that improvement) was reduced by including this term in the model. A lower AIC is generally taken to indicate a better fitting model. In 3 smaller locations (Canton, Birmingham, and New Haven), AIC was lower without including barometric pressure; in Detroit, AIC was the same for both models. For comparison, temperature, a measure traditionally included in models relating air pollution to daily death, also improved (lowered) AIC in 6 out of the 10 locations, and relative humidity improved model fit in only 4 of the 10 locations. The association with barometric pressure was not linear—the protective effect tended to flatten out at high pressure. This is illustrated in Figure B.2, which shows the results for Minneapolis/St Paul. To illustrate the magnitude of the barometric pressure effect better, I also fit linear terms for barometric pressure and performed a meta-analysis of the linear coefficients. Days with high barometric pressure had lower deaths. A 0.25 inches H_2O increase in barometric pressure (which is approximately the mean interquartile range in the 10 cities) was associated with a 1.58% decrease in daily deaths (95% CI, 1.29%, 1.86%).

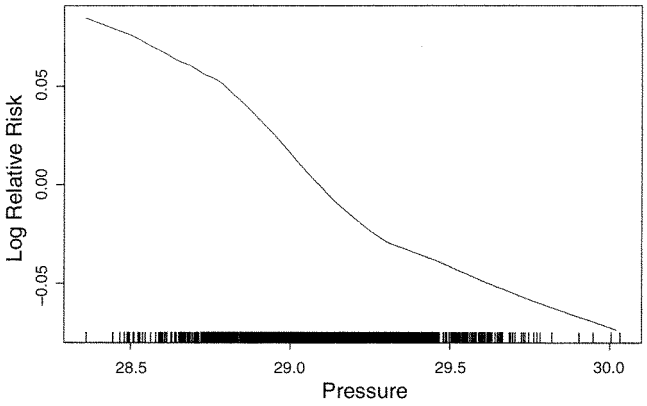


Figure B.2. Nonparametric smooth plot of the log relative risk of death versus pressure in Minneapolis/St Paul.

Table B.3. Estimated Percentage of Increase in Daily Deaths (Standard Errors) for a 10-$\mu g/m^3$ Increase in PM_{10} Under Each Distributed Lag Model

	Constrained Lag Models			
City	1-Day Mean (Lag 0)	2-Day Mean (Lag 0 and Lag 1)	Quadratic Distributed Lag	Unconstrained Lag Model
New Haven	0.62 (0.42)	1.69 (0.52)	1.85 (0.72)	1.80 (0.79)
Birmingham	–0.45 (0.31)	–0.02 (0.36)	0.36 (0.50)	0.34 (0.53)
Pittsburgh	0.58 (0.19)	1.02 (0.21)	0.89 (1.04)	1.00 (0.31)
Detroit	0.66 (0.19)	1.31 (0.21)	1.53 (0.32)	1.75 (0.30)
Canton	1.75 (0.83)	1.80 (0.78)	1.61 (1.25)	1.72 (1.36)
Chicago	0.85 (0.17)	1.18 (0.20)	0.98 (0.26)	0.91 (0.27)
Minneapolis/St Paul	1.43 (0.30)	1.84 (0.32)	2.08 (0.49)	2.01 (0.53)
Colorado Springs	0.16 (0.78)	1.15 (1.00)	1.94 (1.18)	1.75 (1.26)
Spokane	0.34 (0.30)	0.45 (0.31)	2.04 (0.34)	0.74 (0.43)
Seattle	0.70 (0.23)	0.65 (0.31)	1.46 (0.31)	1.46 (0.34)
Overall	0.65 (0.08)	1.05 (0.09)	1.41 (0.13)	1.29 (0.13)

PM_{10} Results

Table B.3 shows the estimated effect of a 10 μg/m^3 increase in PM_{10} in each city using the concurrent day's pollution, the 2-day moving average, the polynomial distributed lag model, and the unconstrained distributed lag model. The effect size estimates using the unconstrained distributed lag were quite similar to those estimated using the polynomial distributed lag model, suggesting the constraint introduced little bias. Both the distributed lag models showed substantially greater overall effects than models using only a single day's exposure and moderately larger effects than the 2-day average models.

The distributed lag model explained some of the heterogeneity among cities in their effect estimates. For example, using the 2-day average model, the variance of the estimated coefficients of PM_{10} across the 10 cities was 3.41×10^{-7}. The average within-city variance of the coefficients was 2.39×10^{-7}, suggesting a heterogeneity not attributable to sampling variability of 1.02×10^{-7}. The across-city variance of the estimated overall effect from the polynomial distributed lag model was 3.13×10^{-7}. This suggests that failure to account for the distributed lag properly accounted for about a quarter of the unexplained variation in effect size among the cities.

Figure B.3 shows the combined estimate of the distributed lag between air pollution and daily deaths. It clearly remains positive for several days before falling toward 0, which explains why the single day's pollutant model is not a good proxy for the overall effect of air pollution.

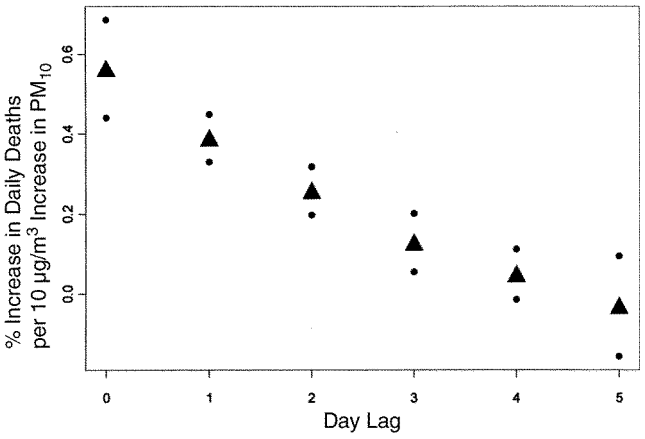


Figure B.3. Distribution over time of the increase in daily deaths (shown as percent increase) associated with a 10 μg/m^3 increase in PM_{10} on a single day. The curve comes from the polynomial distributed lag model, and shows that the effects are distributed over multiple days, falling to zero at about 4 to 5 days after exposure. (• = 45% CI; ▲ = estimate).

DISCUSSION

In every city in this study, evidence was seen that the effect of an incremental increase in particulate air pollution on a given day was spread over several succeeding days. First, the unconstrained distributed lag models in each location always showed greater total effects than the concurrent day models. In addition, Figure B.2, which combines data across cities, shows the effects of pollution were spread over multiple days and do not reach 0 until a lag of 5 days has occurred. In plots of the distributed lag in each city (not shown), the effect is spread over multiple days in each of the 10 locations.

These results are biologically plausible. Given a distribution of sensitivity to air pollution in the general population, which seems likely, and a distribution of severity of preexisting illness, one would expect some variation in the time between exposure and response. This seems even more likely for a summary measure such as all-cause mortality. This measurement mixes deaths from myocardial infarctions, which have been shown to be acutely triggered by immediate exposure to stress or certain activities, with deaths from respiratory disease, where it may take more time for an exacerbatory event to result in the cascade of biological responses leading to death (Mittleman et al 1995).

In a recent study of air pollution and daily deaths in Milan, Italy, causes of death were examined separately (Rossi et al 1999). Some causes were more strongly associated with particle exposure on the concurrent day, but other causes were more strongly associated with exposure 2 days before. The mixing of such deaths in all-cause mortality would naturally result in a distributed lag between exposure and response.

Support is also provided by the London smog episode of 1952 (Her Majesty's Public Health Service 1954). Air pollution concentrations shot up on December 5, 1952, and there was an immediate, same-day increase in deaths. The curve of increase and decrease of daily deaths in general, however, lagged behind the increase and decrease of air pollution, with peak deaths occurring a day or two after peak exposure. This suggests that there were substantial lagged effects.

There is also toxicologic support for effects that persist longer than a single day. In addition to the report of Clarke and colleagues (1999) cited earlier, Lay and coworkers (1999) have reported that particles instilled in the lung induced an inflammation that took up to 4 days after exposure to resolve.

The implications of this are evident in Table B.3. The estimated effect of a 10 μg/m^3 increase in PM_{10} on 1 day using the 1-day moving average model was only about 40% of the estimate using an unconstrained or constrained

model that estimates effects at longer lags. Notably, the unconstrained model makes no assumption about the shape of the distributed lag, or even the existence of effects at lags greater than 1 day. It merely allows for the possibility that they exist and estimates them based on the data. Hence, these results indicate that studies that rely on single-day exposure averages will, on average, substantially underestimate the effect of particle exposure. It is possible that this may not be true for other pollutants, but it would not be prudent to make that assumption when the methodology to test it is straightforward.

A 2-day moving average did substantially better than the 1-day average in estimating the effects of PM_{10}, but this average still underestimated the overall impacts by about 40%.

For other outcomes, such as cause-specific mortality or hospital admissions, the distribution of effect over time may have a different pattern. Separate evaluations, using distributed lag models, will be needed to assess the adequacy of simple single-day or multiple-day averages. For multiple time-series studies, the unconstrained lag model avoids any risk of bias. For single time-series studies, the polynomial distributed lag model appears to risk little bias and should be the method of choice. The advantage of these distributed lag models is that we do not need to leave the question of how the effects are distributed over time to chance. By fitting a model that allows but does not require the effect of pollution to be distributed over several days, we can make that question part of our investigation. By using the simple transformation shown in the Data and Methods section of this appendix, this approach can be implemented in any Poisson regression package. Hence, polynomial distributed lag models should become standard practice in air pollution epidemiology, unless there is clear biological reason for assuming that the response is limited to a single day. These models can be applied equally well to other acute triggers. For example, aeroallergen exposure and acute asthmatic response or triggers of acute myocardial infarction may represent areas where these models can be usefully applied.

The finding that barometric pressure is consistently associated with lower mortality is also of considerable interest. Previously, we have shown that barometric pressure was associated with oxygen saturation (Pope et al 1999). These changes might plausibly influence mortality risk; however, they were noted in a high altitude location (Provo/Orem). The findings in this study of an association even in cities closer to sea level, such as Minneapolis/St Paul (Figure B.2), suggests that further attention should be paid to this variable. Future studies of the effects of air pollution on morbidity and mortality should include barometric pressure in the list of covariates to be controlled.

REFERENCES

Akaike H. 1973. Information theory and an extension of the maximum likelihood principal. In: Second International Symposium on Information Theory. (Petrov BN, Csaki F, eds). Akademiai Kaiado, Budapest, Hungary.

Almon S. 1965. The distributed lag between capital appropriations and expenditures. Econometrica 33:178–196.

Brumback BA, Ryan LM, Schwartz J, Neas LM, Stark PC, Burge HA. 1999. Transitional regression models with application to environmental time series. J Am Stat Assoc (in press).

Clarke RW, Catalano PJ, Koutrakis P, Murthy GG, Sioutas C, Paulauskis J, Coull B, Ferguson S, Godleski JJ. 1999. Urban air particulate inhalation alters pulmonary function and induces pulmonary inflammation in a rodent model of chronic bronchitis. Inhalation Toxicol 11:637–656.

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

Dominici F, Samet J, Xu J, Zeger S. 2000. Combining evidence on air pollution and daily mortality from the largest 20 US cities: A hierarchical modeling strategy. J R Stat Soc Series C (in press).

Hastie T, Tibshirani R, eds. 1990. Generalized Additive Models. Chapman & Hall, New York NY.

Her Majesty's Public Health Service. 1954. Mortality and Morbidity During the London Fog of December 1952. Report Number 95 on Public Health and Medical Subjects. Her Majesty's Stationery Office, London, England.

Hoek G, Schwartz J, Groot B, Eilers P. 1997. Effects of ambient particulate matter and ozone on daily mortality in Rotterdam, The Netherlands. Arch Environ Health 52:455–463.

Judge GG, Griffiths WE, Hill RC, Lutkepohl H, Lee T-C. 1980. The Theory and Practice of Econometrics. John Wiley & Sons, New York NY.

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, Rossi G, Wojtyniak D, Sunyer J, Bacharova L, Schouten JP, Pontak A, Anderson HR. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: Results from time-series data from the APHEA project. Br Med J 314:1658–1663.

Kelsall JE, Samet JM, Zeger SL, Xu J. 1997. Air pollution and mortality in Philadelphia, 1974–1988. Am J Epidemiol 146:750–762.

Lay JC, Bennett WD, Ghio AJ, Bromberg PA, Costa DL, Kim CS, Koren HS, Devlin RB. 1999. Cellular and biochemical response of the human lung after intrapulmonary instillation of ferric oxide particles. Am J Respir Cell Mol Biol 20(4):631–642.

Mittleman MA, Maclure M, Sherwood JB, Mulry RP, Tofler GH, Jacobs SC, Friedman R, Benson H, Muller JE. 1995. Triggering of acute myocardial infarction onset by episodes of anger. Determinants of Myocardial Infarction Onset Study Investigators. Circulation 92:1720–1725.

Pope CA, Dockery DW, Kanner RE, Villegas GM, Schwartz J. 1999. Oxygen saturation, pulse rate, and particulate air pollution: A daily time series panel study. Am J Respir Crit Care Med 159:365–372.

Pope CA III, Schwartz J. 1996. Time series for the analysis of pulmonary health data. Am J Respir Crit Care Med 154:S229–S233.

Pope CA, Schwartz J, Ransom M. 1992. Daily mortality and PM_{10} pollution in Utah Valley. Arch Environ Health 42:211–217.

Rossi G, Vigotti MA, Zanobetti A, Repetto F, Gianelle V, Schwartz J. 1999. Air pollution and cause-specific mortality in Milan, Italy, 1980–1989. Arch Environ Health 54:158–164.

Samet J, Zeger S, Kelsall J, Xu J, Kalkstein L. 1998. Does weather confound or modify the association of particulate air pollution with mortality? An analysis of the Philadelphia data, 1973–1980. Environ Res 77(1):9–19.

Schwartz J. 1999. Air pollution and hospital admissions for heart disease in eight US counties. Epidemiology 10:17–22.

Schwartz J. 1994. Generalized additive models in epidemiology. International Biometric Society, Invited Papers. 17th International Biometric Conference, pp 55–80. International Biometric Society, Washington DC.

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

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

Schwartz J, Marcus A. 1990. Mortality and air pollution in London: A time series analysis. Am J Epidemiol 131:185–194.

Touloumi G, Pocock SJ, Katsouyanni K, Trichopolous D. 1994. Short-term effects of air pollution on daily mortality in Athens: A time series analysis. Int J Epidemiol 23:957–967.

APPENDIX C. Comparison of Heterogeneity Models

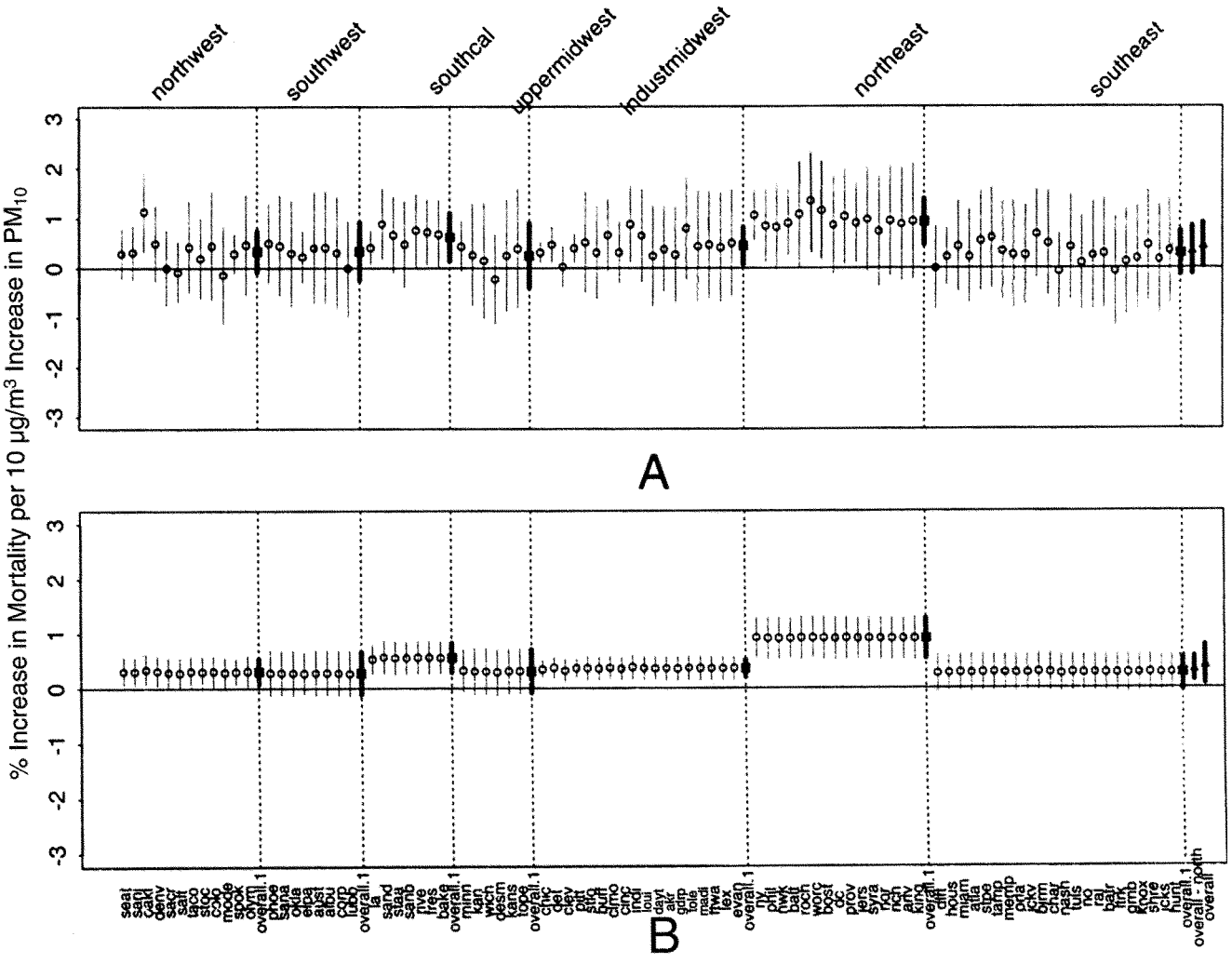


Figure C.1. City-specific posterior means and 95% posterior regions at lag 1 for each of 90 locations under both priors (models A, more heterogeneity; B, less or no heterogeneity). Also shown are the posterior estimates and 95% posterior regions at lag 1 for the regional and overall means. The estimates of overall PM_{10} relative risk are nearly identical for the two prior models [A:0.45 (0.03, 0.89); B: 0.44 (0.07, 0.76)]. The posterior regions are narrower and more realistic using model B. Note also that under prior B, the city and region-specific estimates are more similar to one another. This is because the posterior distributions for the heterogeneity parameters σ^2 and τ^2 are centered at medians 0.002 and 0.12, respectively, indicating a small degree of heterogeneity of effects within a region and a relatively moderate degree of heterogeneity of effects across regions. For example, a median value of $\sigma = \sqrt{0.002} = 0.045$ corresponds to 95% of cities within a region having the PM_{10} relative risks of $\pm 2 \times 0.045 = \pm 0.09$ or about $\pm$ 20% of the overall relative risk.

APPENDIX D. Monitoring Sites, Base Model Characteristics, and Results for Morbidity Analysis

Table D.1. Monitoring Sites

Location	PM Observations	Site Code	Site Address
Birmingham AL			
North Birmingham, Jefferson County	2,397	010730023 - 3	3009 28th St N
Wylam, Jefferson County	900	010732003 - 1	1242 Jersey St
Leeds, Jefferson County	872	010731010 - 1	201 Ashville Rd
Tarrant City, Jefferson County	462	010736002 - 2	Tarrant Elem Sch, 1269 Portland St
Birmingham, Jefferson County	421	010730026 - 1	Inglenook Elem Sch, 3937 44th Ave N
Bessemer, Jefferson County	419	010730002 - 1	1500 1st Ave N
Birmingham, Jefferson County	363	010730034 - 1	2301 11th Ave N
Boulder CO			
Longmont, Boulder County	2,136	080130003 - 2	3rd & Kimbark St
Longmont, Boulder County	383	080130003 - 2	3rd & Kimbark St
Canton OH			
Canton, Stark County	1,271	391511001 - 1	
Canton, Stark County	527	391510017 - 1	1330 Dueber Ave SW
Canton, Stark County	372	391510020 - 1	420 Market Ave N
Canton, Stark County	332	391510009 - 1	1901 Midway NE
Chicago IL			
Chicago, Cook County	2,173	170310022 - 1	3535 E 114th St
Mccook, Cook County	1,334	170311016 - 1	50th St & Glencoe
South Holland, Cook County	670	170313701 - 1	170th St & S Park Ave
Blue Island, Cook County	412	170312001 - 1	Eisenhower HS, 12700 Sacramento Ave
Cicero, Cook County	407	170316001 - 1	Roosevelt HS, 15th St & 50th Ave
Chicago, Cook County	404	170310014 - 1	Farr Dormitory, 3300 S Michigan Ave
Chicago, Cook County	401	170310049 - 1	Chicago Ave Pump Sta, 805 N Michigan
Lyons, Cook County	354	170311701 - 1	4043 Joliet Ave
Chicago, Cook County	350	170310060 - 1	13100 S Doty
Summit, Cook County	346	170313301 - 1	Graves Elem Sch, 60th St & 74th Ave
Colorado Springs CO			
Colorado Springs, El Paso County	2,412	080410008 - 2	3730 Meadowlands
Colorado Springs, El Paso County	433	080410010 - 2	701 N Circle (Service Center)
Colorado Springs, El Paso County	400	080410011 - 2	101 W Costillia
Colorado Springs, El Paso County	442	080410011 - 3	101 W Costillia
Detroit MI			
Dearborn, Wayne County	1,507	261630033 - 1	2842 Wyoming
Wayne County	910		
Detroit, Wayne County	571	261630015 - 1	6921 West Fort
Detroit, Wayne County	501	261630092 - 2	312 West End
Allen Park, Wayne County	472	261630001 - 1	14700 Goddard
River Rouge, Wayne County	448	261630005 - 1	315 Genesee
Wayne County	319		
Wayne County	301		

(Table continues next page)

Table D.1 (*continued*). Monitoring Sites

Location	PM Observations	Site Code	Site Address
Minneapolis/St Paul MN			
St Paul, Ramsey County	2,407	271230047 - 1	1303 Red Rock Rd
St Paul, Ramsey County	1,344	271230040 - 1	Wabasha & W 5th St
Minneapolis, Hennepin County	1,306	270530051 - 1	300 Nicollet Mall
Minneapolis, Hennepin County	1,249	270530053 - 1	300 Nicollet Mall
St Louis Park, Hennepin County	428	270532006 - 1	City Hall, 5005 Minntonka Boulevard
St Paul, Ramsey County	418	271230021 - 1	1038 Ross Ave
Richfield, Hennepin County	417	270533004 - 1	I35W & 70th Ave South
Minneapolis, Hennepin County	415	270531007 - 1	4646 Humboldt Ave N
Nashville TN			
Nashville-Davidson, Davidson County	2,076	470370006 - 1	8th Ave N on Roof of Housing Auth Bldg
Nashville-Davidson, Davidson County	301	470370006 - 2	8th Ave N on Roof of Housing Auth Bldg
Nashville-Davidson, Davidson County	300	470370011 - 1	1015 Trinity Lane
Nashville-Davidson, Davidson County	300	470370023 - 1	Lockeland School, 105 S 17th St
New Haven CT			
New Haven, New Haven County	1,620	090090018 - 1	Stiles St
Waterbury, New Haven County	437	090092123 - 1	Shed Meadow & Bank St
New Haven, New Haven County	383	090090011 - 1	Hamilton Ave
New Haven, New Haven County	380	090090013 - 1	Fire Headquarters, Grand Ave
Waterbury, New Haven County	380	090092123 - 2	Shed Meadow & Bank St
New Haven, New Haven County	378	090091123 - 1	715 State St
New Haven, New Haven County	374	090091123 - 2	715 State St
Waterbury, New Haven County	372	090093007 - 1	519 East Main St
Milford, New Haven County	355	090090010 - 1	Egan Center, Mathew St
Wallingford, New Haven County	346	090094006 - 1	45 S Main St
Meriden, New Haven County	336	090094002 - 1	Stoddard Bldg, 165 Miller St
Pittsburgh PA			
Liberty, Allegheny County	2,299	420030064 - 1	2743 Washington Blvd (Mckeesport)
Braddock, Allegheny County	1,449	420032001 - 1	St Thomas Sch, 1025 Braddock Ave
Pittsburgh, Allegheny County	1,111	420030027 - 2	3333 Forbes Ave
Clairton, Allegheny County	944	420033004 - 1	Greenway Alley
South Fayette, Allegheny County	873	420030067 - 1	Old Oakdale Rd
Glassport, Allegheny County	814	420033006 - 1	High St Water Tower
Braddock, Allegheny County	793	420032001 - 2	St Thomas Sch, 1025 Braddock Ave
Clairton, Allegheny County	705	420033007 - 1	Clairton Ed Ctr, 501 Waddel
Clairton, Allegheny County	643	420033008 - 1	
Allegheny County	640	420037003 - 1	
North Braddock, Allegheny County	610	420031301 - 1	600 Anderson St
Lincoln Boro, Allegheny County	538	420037004 - 1	Bellebridge Rd
Pittsburgh, Allegheny County	508	420030021 - 1	Gladstone High Sch
Pittsburgh, Allegheny County	387	420030031 - 1	1275 Bedford Ave
Avalon, Allegheny County	312	420030002 - 1	520 Orchard St
North Braddock, Allegheny County	305	420031301 - 2	600 Anderson St
Allegheny County	303	420030041 - 1	
Provo/Orem UT			
Lindon, Utah County	2,390	490494001 - 2	30 N Main St
Orem, Utah County	2,025	490495001 - 1	300 N 1200 West
Provo, Utah County	1,859	490490002 - 2	1355 N 200 West
Utah County	325	490495002 - 1	

(Table continues next page)

Table D.1 (*continued*). Monitoring Sites

Location	PM Observations	Site Code	Site Address
Seattle WA			
Seattle, King County	2,307	530330057 - 3	Duwamish Pump Sta, 4752 E Marginal Way S
Kent, King County	1,922	530332004 - 2	James St & Central Ave
Lake Forest Park, King County	1,679	530330086 - 1	17711 Ballinger Way NE
Seattle, King County	1,028	530330057 - 1	Duwamish Pump Sta, 4752 E Marginal Way S
Bellevue, King County	960	530330004 - 2	West & Wheeler Bldg, 504 Bellevue Way
Seattle, King County	396	530330068 - 3	South Park, 723 S Concord
Seattle, King County	373	530330066 - 2	Harbor Island, 3400 13th Ave SW
Seattle, King County	318	530330004 - 1	West & Wheeler Bldg, 504 Bellevue Way
Spokane WA			
Spokane, Spokane County	1,735	530630016 - 3	Crown Zellerbach, E 3530 Ferry
Spokane, Spokane County	935	530630016 - 1	Crown Zellerbach, E 3530 Ferry
Spokane County	620	530631014 - 2	
Spokane, Spokane County	593	530631015 - 1	Nazarene, N 9004 Country Homes Blvd
Spokane, Spokane County	382	530630036 - 2	Auto Glass, S 214 Post
Millwood, Spokane County	365	530632002 - 1	City Hall, E 9103 Frederick, Millwood
Spokane, Spokane County	356	530630016 - 4	Crown Zellerbach, E 3530 Ferry
Youngstown OH			
East Liverpool, Columbiana County	1587	390292001 - 1	East Liverpool City Hall 126 West 6th
Youngstown, Mahoning County	1236	390990006 - 1	Fire Station 5, Superior & Oakland
Youngstown, Mahoning County	1115	390990005 - 1	Fire Station 7, Elm & Madison
East Liverpool, Columbiana County	377	390290003 - 1	Eastside Firehouse, East Penna Ave

Table D.2. Base Model Characteristics for Cardiovascular Admissions

	Season		Temperature		Temperature Lag1		Relative Humidity		Barometric Pressure		Day of Week		
City	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	Auto-regressive Terms
Birmingham	220/2,467	20.1	0.5	3.3	0.6	2.7	0.6	2.5	0.5	3.4	0.5	6.0	none
Boulder	550/2,064	6.4	0.7	2.2	0.6	2.7	0.7	2.1	0.7	2.2	0.5	6.0	none
Canton	330/2,189	11.5	0.6	2.7	0.5	3.4	0.7	2.2	0.5	3.7	0.5	6.0	none
Chicago	235/2,547	19.6	0.5	3.3	0.7	2.2	0.7	2.2	0.5	3.6	0.5	6.0	ar1,ar2,ar3
Colorado Springs	500/2,734	9.7	0.5	3.4	0.6	2.6	0.6	2.6	0.5	3.2	0.5	6.0	none
Detroit	335/3,159	17.1	0.5	3.2	0.6	2.5	0.4	4.4	0.5	3.6	0.5	6.0	ar1,ar2,ar3
Minneapolis/St Paul	240/2,824	21.1	0.5	3.2	0.5	3.1	0.4	4.2	0.6	2.7	0.5	6.0	none
Nashville	900/1,941	3.6	0.7	2.1	0.7	2.1	0.5	3.3	0.7	2.3	0.5	6.0	none
New Haven	240/1,707	12.5	0.7	2.2	0.5	3.2	0.6	2.7	0.5	3.6	0.5	6.0	none
Pittsburgh	195/2,915	26.4	0.5	3.2	0.5	3.2	0.7	2.1	0.5	3.5	0.5	6.0	ar1
Provo/Orem	600/2,825	8.2	0.5	3.2	0.4	4.2	0.5	3.2	0.5	3.3	0.5	6.0	none
Seattle	210/3,280	27.5	0.5	3.4	0.6	2.6	0.7	2.2	0.5	3.5	0.5	6.0	none
Spokane	335/3,372	18.0	0.6	2.5	0.5	3.2	0.7	2.2	0.6	2.7	0.5	6.0	none
Youngstown	210/1,461	12.1	0.6	2.5	0.5	3.3	0.5	3.3	0.7	2.4	0.5	6.0	none

Table D.3. Base Model Characteristics for COPD Admissions

Cities	Season		Temperature		Temperature Lag 1		Relative Humidity		Barometric Pressure		Day of Week		Auto-regressive Terms
	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	
Birmingham	250/2,464	17.6	0.7	2.2	0.7	2.2	0.7	2.1	0.5	3.4	0.5	6.0	none
Boulder	365/2,064	10.1	0.5	3.3	0.7	2.2	0.7	2.2	0.7	2.2	0.5	6.0	none
Canton	380/2,184	10.2	0.6	2.7	0.5	3.5	0.7	2.2	0.6	2.9	0.5	6.0	none
Chicago	185/2,545	24.5	0.5	3.3	0.7	2.2	0.7	2.2	0.5	3.5	0.5	6.0	ar1,ar2,ar3
Colorado Springs	340/2,734	14.2	0.5	3.3	0.7	2.1	0.5	3.2	0.7	2.1	0.5	6.0	none
Detroit	285/3,158	20.5	0.5	3.2	0.6	2.6	0.6	2.6	0.5	3.4	0.5	6.0	ar1,ar2,ar3
Minneapolis/St Paul	210/2,824	23.7	0.6	2.5	0.5	3.1	0.5	3.3	0.6	2.7	0.5	6.0	none
Nashville	250/1,941	13.7	0.5	3.2	0.5	3.2	0.5	3.2	0.5	3.6	0.5	6.0	none
New Haven	550/1,707	5.6	0.5	3.3	0.7	2.2	0.5	3.2	0.5	3.7	0.5	6.0	none
Pittsburgh	160/2,912	32.5	0.6	2.6	0.6	1.6	0.7	2.2	0.0	3.5	0.5	6.0	none
Provo/Orem	600/2,825	8.3	0.9	1.5	0.9	1.5	0.7	2.1	0.9	1.5	0.5	6.0	none
Seattle	280/3,280	21.1	0.6	2.6	0.5	3.3	0.5	3.5	0.4	4.7	0.5	6.0	none
Spokane	360/3,370	16.7	0.6	2.5	0.6	2.5	0.7	2.2	0.6	2.7	0.5	6.0	none
Youngstown	380/1,461	6.6	0.7	2.1	0.9	1.5	0.4	4.5	0.6	3.0	0.5	6.0	none

Table D.4. Base Model Characteristics for Pneumonia Admissions

City	Season		Temperature		Temperature Lag1		Relative Humidity		Barometric Pressure		Day of Week		Auto-regressive Terms
	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	LOESS Span	*df*	
Birmingham	200/2,458	22.6	0.4	4.6	0.4	4.6	0.5	3.1	0.5	3.5	0.5	6.0	ar1,ar2,ar4,ar5,ar7
Boulder	350/2,064	10.5	0.4	4.5	0.5	3.5	0.5	3.2	0.5	3.3	0.6	2.9	none
Canton	230/2,180	16.9	0.5	3.4	0.5	3.5	0.7	2.2	0.6	2.9	0.5	6.0	none
Chicago	130/2,542	36.2	0.7	2.2	0.7	2.2	0.7	2.2	0.6	2.8	0.5	6.0	ar1,ar2,ar3
Colorado Springs	200/2,734	24.2	0.7	2.2	0.7	2.2	0.4	4.1	0.7	2.1	0.6	2.9	none
Detroit	260/3,156	23.0	0.5	3.2	0.4	4.2	0.7	2.1	0.5	3.6	0.5	6.0	ar1,ar2,ar3,ar4
Minneapolis/St Paul	175/2,824	28.1	0.6	2.5	0.5	3.2	0.5	3.3	0.6	2.8	0.5	6.0	ar1,ar2,ar3
Nashville	250/1,941	13.8	0.5	3.3	0.5	3.3	0.5	3.3	0.7	2.4	0.6	2.6	ar1,ar2
New Haven	170/1,702	18.0	0.5	3.3	0.4	4.4	0.6	2.7	0.4	4.9	0.5	6.0	ar1
Pittsburgh	145/2,906	36.5	0.4	4.2	0.4	4.1	0.7	2.2	0.4	5.0	0.5	6.0	ar1,ar2,ar3,ar4
Provo/Orem	270/2,825	19.2	0.9	1.5	0.9	1.5	0.6	2.5	0.5	3.4	0.5	6.0	none
Seattle	225/3,276	25.8	0.4	4.7	0.4	4.7	0.5	3.4	0.6	2.7	0.5	6.0	ar1,ar2,ar3
Spokane	225/3,369	26.7	0.6	2.5	0.6	2.5	0.7	2.2	0.6	2.7	0.5	6.0	none
Youngstown	190/1,461	13.5	0.4	4.2	0.5	3.3	0.4	4.4	0.4	5.3	0.5	6.0	none

Table D.5. City-Specific and Combined Results of Analysis of PM_{10} Association with HCFA Cardiovascular Admissions

	PM_{10} Lag 0			PM_{10} Lag 1			PM_{10} Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value
Birmingham	0.00063	0.00028	2.2	0.00003	0.00027	0.1	0.00042	0.00032	1.3	0.00093	0.00047	2.0	0.00100	0.00047	2.1
Boulder	0.00254	0.00126	2.0	–0.00022	0.00127	–0.2	0.00168	0.00139	1.2	–0.00078	0.00210	–0.4	–0.00104	0.00210	–0.5
Canton	0.00041	0.00071	0.6	0.00005	0.00067	0.1	0.00073	0.00068	1.1	–0.00128	0.00113	–1.1	–0.00122	0.00113	–1.1
Chicago	0.00111	0.00015	7.3	0.00074	0.00014	5.5	0.00131	0.00017	7.5	0.00112	0.00023	5.0	0.00111	0.00023	4.9
Colorado Springs	0.00030	0.00086	0.4	0.00133	0.00082	1.6	0.00116	0.00089	1.3	0.00226	0.00138	1.6	0.00219	0.00139	1.6
Detroit	0.00127	0.00017	7.3	0.00094	0.00016	6.0	0.00151	0.00018	8.2	0.00161	0.00023	6.9	0.00161	0.00023	6.9
Minneapolis/ St Paul	0.00108	0.00034	3.1	0.00022	0.00033	0.7	0.00073	0.00038	1.9	0.00058	0.00059	1.0	0.00063	0.00060	1.1
Nashville	–0.00023	0.00059	–0.4	–0.00001	0.00059	0.0	0.00020	0.00060	0.3	–0.00035	0.00103	–0.3	–0.00032	0.00103	–0.3
New Haven	0.00204	0.00040	5.1	0.00108	0.00040	2.7	0.00210	0.00042	5.0	0.00209	0.00071	2.9	0.00209	0.00071	2.9
Pittsburgh	0.00099	0.00015	6.5	0.00084	0.00014	5.9	0.00124	0.00017	7.3	0.00118	0.00025	4.8	0.00119	0.00025	4.8
Provo/Orem	0.00060	0.00055	1.1	0.00000	0.00050	0.0	0.00035	0.00057	0.6	0.00092	0.00068	1.4	0.00095	0.00068	1.4
Seattle	0.00130	0.00025	5.3	0.00056	0.00024	2.3	0.00107	0.00026	4.1	0.00136	0.00036	3.8	0.00141	0.00036	3.9
Spokane	0.00055	0.00032	1.7	0.00050	0.00029	1.7	0.00060	0.00033	1.8	0.00072	0.00038	1.9	0.00077	0.00038	2.0
Youngstown	0.00176	0.00062	2.8	–0.00009	0.00059	–0.2	0.00102	0.00062	1.6	0.00022	0.00103	0.2	0.00022	0.00104	0.2
Combined Results															
Fixed effects	0.00107	0.00007	14.6	0.00067	0.00007	9.8	0.00116	0.00008	14.6	0.00117	0.00011	10.7	0.00118	0.00011	10.8
Random effects	0.00102	0.00017	6.09	0.00067	0.00007	9.8	0.00116	0.00008	14.6	0.00104	0.00019	5.42	0.00106	0.00020	5.29

Table D.6. City-Specific and Combined Results of PM_{10} Analysis with HCFA Pneumonia Admissions

	PM_{10} Lag 0			PM_{10} Lag 1			PM_{10} Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value
Birmingham	0.00002	0.00056	0.0	0.00048	0.00053	0.9	0.00026	0.00063	0.4	0.00052	0.00083	0.6	0.00082	0.00088	0.9
Boulder	0.00544	0.00236	2.3	0.00023	0.00238	0.1	0.00349	0.00257	1.4	0.00523	0.00362	1.4	0.00488	0.00363	1.3
Canton	0.00130	0.00146	0.9	0.00030	0.00139	0.2	0.00112	0.00139	0.8	–0.00143	0.00227	–0.6	–0.00138	0.00228	–0.6
Chicago	0.00147	0.00030	4.9	0.00153	0.00028	5.5	0.00231	0.00034	6.8	0.00215	0.00045	4.8	0.00282	0.00045	6.2
Colorado Springs	0.00349	0.00147	2.4	0.00323	0.00140	2.3	0.00447	0.00152	2.9	0.00836	0.00225	3.7	0.00815	0.00225	3.6
Detroit	0.00209	0.00038	5.5	0.00200	0.00034	5.9	0.00272	0.00040	6.8	0.00119	0.00050	2.4	0.00109	0.00050	2.2
Minneapolis/ St Paul	0.00233	0.00069	3.4	0.00231	0.00065	3.6	0.00337	0.00076	4.4	0.00227	0.00112	2.0	0.00218	0.00189	1.2
Nashville	0.00093	0.00144	0.6	0.00066	0.00130	0.5	0.00113	0.00144	0.8	0.00056	0.00188	0.3	0.00061	0.00189	0.3
New Haven	0.00259	0.00083	3.1	0.00103	0.00077	1.3	0.00276	0.00089	3.1	0.00438	0.00142	3.1	0.00599	0.00145	4.1
Pittsburgh	0.00157	0.00032	4.9	0.00109	0.00030	3.6	0.00182	0.00036	5.0	0.00141	0.00054	2.6	0.00142	0.00054	2.6
Provo/Orem	0.00084	0.00098	0.9	–0.00017	0.00087	–0.2	0.00039	0.00099	0.4	0.00072	0.00111	0.6	0.00104	0.00112	0.9
Seattle	0.00109	0.00051	2.1	0.00136	0.00048	2.8	0.00145	0.00053	2.7	0.00208	0.00069	3.0	0.00198	0.00070	2.8
Spokane	0.00155	0.00061	2.6	–0.00003	0.00058	–0.1	0.00083	0.00062	1.3	0.00150	0.00068	2.2	0.00169	0.00069	2.5
Youngstown	0.00203	0.00136	1.5	0.00116	0.00131	0.9	0.00174	0.00137	1.3	0.00107	0.00220	0.5	0.00124	0.00221	0.6
Combined Results															
Fixed effects	0.00155	0.00015	10.4	0.00130	0.00014	9.3	0.00196	0.00016	12.0	0.00167	0.00022	7.7	0.00188	0.00022	8.6
Random effects	0.00160	0.00028	5.7	0.00130	0.00014	9.3	0.00189	0.00023	8.1	0.00186	0.00058	3.2	0.00205	0.00057	3.6

Table D.7. City-Specific and Combined Results of PM_{10} Analysis with HCFA Chronic Obstructive Pulmonary Disease Admissions

	PM_{10} Lag 0			PM_{10} Lag 1			PM_{10} Lag 0/1			Quadratic Distributed Lag			Unrestricted Distributed Lag		
City	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value
Birmingham	−0.00106	0.00093	−1.1	−0.00134	0.00092	−1.5	−0.00145	0.00107	−1.4	−0.00133	0.00160	−0.8	−0.00120	0.00161	−0.7
Boulder	0.00693	0.00339	2.0	0.00961	0.00327	2.9	0.01122	0.00371	3.0	0.01870	0.00560	3.3	0.01791	0.00563	3.2
Canton	0.00052	0.00190	0.3	0.00312	0.00175	1.8	0.00137	0.00182	0.8	0.00254	0.00302	0.8	0.00250	0.00303	0.8
Chicago	0.00081	0.00047	1.7	0.00141	0.00045	3.1	0.00187	0.00055	3.4	0.00091	0.00083	1.1	0.00089	0.00083	1.1
Colorado Springs	0.00219	0.00227	1.0	−0.00178	0.00233	−0.8	0.00078	0.00249	0.3	0.00718	0.00344	2.1	0.00780	0.00344	2.3
Detroit	0.00183	0.00055	3.3	0.00213	0.00050	4.3	0.00271	0.00058	4.7	0.00288	0.00078	3.7	0.00284	0.00078	3.7
Minneapolis/ St Paul	0.00246	0.00106	2.3	0.00241	0.00104	2.3	0.00351	0.00117	3.0	0.00364	0.00183	2.0	0.00346	0.00184	1.9
Nashville	0.00170	0.00183	0.9	0.00118	0.00187	0.6	0.00213	0.00189	1.1	0.00239	0.00306	0.8	0.00199	0.00307	0.6
New Haven	0.00437	0.00166	2.6	0.00217	0.00165	1.3	0.00370	0.00175	2.1	0.00710	0.00291	2.4	0.00701	0.00291	2.4
Pittsburgh	0.00191	0.00042	4.5	0.00182	0.00040	4.6	0.00250	0.00047	5.4	0.00280	0.00073	3.8	0.00263	0.00074	3.6
Provo/Orem	−0.00229	0.00248	−0.9	−0.00242	0.00223	−1.1	−0.00288	0.00253	−1.1	−0.00267	0.00289	−0.9	−0.00281	0.00293	−1.0
Seattle	0.00119	0.00078	1.5	0.00053	0.00076	0.7	0.00103	0.00082	1.2	0.00446	0.00112	4.0	0.00462	0.00113	4.1
Spokane	0.00214	0.00095	2.3	0.00144	0.00086	1.7	0.00197	0.00096	2.1	0.00251	0.00104	2.4	0.00261	0.00106	2.5
Youngstown	−0.00027	0.00185	−0.1	0.00150	0.00167	0.9	0.00141	0.00178	0.8	−0.00212	0.00313	−0.7	−0.00177	0.00315	−0.6
Combined Results															
Fixed effects	0.00143	0.00022	6.4	0.00144	0.00021	6.8	0.00196	0.00025	8.0	0.00246	0.00035	7.0	0.00242	0.00035	6.9
Random effects	0.00141	0.00054	2.6	0.00117	0.00076	1.5	0.00180	0.00033	5.5	0.00290	0.00137	2.1	0.00284	0.00134	2.1

Table D.8. City-Specific and Combined Results for 2-Day Mean PM_{10}, for Values Less Than 50 $\mu g/m^3$

	CVD			Pneumonia			COPD		
	β	SE	*t* value	β	SE	*t* value	β	SE	*t* value
Birmingham	0.00024	0.00058	0.41	−0.00282	0.00144	−1.96	−0.00340	0.00190	−1.79
Boulder	−0.00011	0.0018	−0.06	0.00319	0.00327	0.98	0.00818	0.004935	1.66
Canton	0.00081	0.00090	0.89	0.00139	0.00181	0.77	0.00073	0.00239	0.30
Chicago	0.00073	0.00032	2.30	0.00268	0.00062	4.33	0.00220	0.00100	2.21
Colorado Springs	0.00200	0.00143	1.40	0.00588	0.00255	2.31	0.00318	0.00398	0.80
Detroit	0.00231	0.00036	6.43	0.00326	0.00080	4.06	0.00238	0.00118	2.01
Minneapolis/St Paul	0.00099	0.00052	1.90	0.00486	0.00104	4.66	0.00339	0.00163	2.08
Nashville	0.0004	0.00087	0.46	−0.00112	0.00188	−0.60	0.00506	0.002650	1.91
New Haven	0.00338	0.00062	5.45	0.00446	0.00131	3.40	0.00002	0.00269	0.01
Pittsburgh	0.00211	0.00034	6.20	0.00389	0.00086	4.53	0.00327	0.00097	3.36
Provo/Orem	0.00149	0.00137	1.09	0.00207	0.00267	0.77	−0.00272	0.006268	−0.43
Seattle	0.00121	0.00048	2.54	0.00414	0.00105	3.94	0.00399	0.00147	2.71
Spokane	0.00058	0.00081	0.71	−0.00190	0.00147	−1.29	0.00387	0.00230	1.68
Youngstown	0.00154	0.00098	1.57	0.00327	0.00210	1.55	0.00646	0.002766	2.34
Combined Results									
Fixed effects	0.00146	0.00015	9.9	0.0028	0.00032	8.9	0.00259	0.000457	5.7
Random effects	0.00144	0.00017	8.7	0.00243	0.00065	3.7	0.00256	0.000599	4.3

APPENDIX E. Method for Averaging Pollutant Data

For each day, more than one monitor within a county may have measured a particular pollutant simultaneously. In order to obtain one estimate of all values collected on a given day and to exclude extreme values that might not be valid, we applied a trimmed mean method as follows.

From all available measurements of each day, we dropped values beyond the upper and lower tenth percentiles; when the values were from fewer than 10 monitors, we dropped the minimum and maximum values. With 1 monitor, the absolute value was used; for 2 monitors, the average. We calculated the trimmed mean value across all monitors. For individual monitor data, we first took the straight mean across all days for each monitor. We then took the daily residual value from the mean for each monitor. We next calculated the 10% trimmed mean of the residuals across all monitors and added back the trimmed mean of measurements. The time-series values for each pollutant in each county were then output to one vector.

ABOUT THE AUTHORS

Jonathan M Samet is Professor and Chairman of the Department of Epidemiology at Johns Hopkins School of Public Health. He received his medical degree from the University of Rochester School of Medicine and Dentistry and a master's degree in epidemiology from Harvard School of Public Health. His research has focused on the effects of environmental and occupational agents. He is chairman of the NRC's Committee on Research Priorities for Airborne Particulate Matter and was chairman of the NRC's Committee on Health Risk of Exposure to Radon (BEIR VI).

Francesca Dominici, Assistant Professor of Biostatistics at the Johns Hopkins School of Public Health, received her doctoral degree from the University of Padua. Her research interests include time-series analysis in environmental epidemiology, semiparametric Bayesian models for dose-response data, hierarchical models for combining information and for meta-analysis, and missing data. She is the first recipient of the Walter A Rosenblith New Investigator Award from the Health Effects Institute. Her work on air pollution and mortality was recently honored by an invitation to read her paper on this topic to the Royal Statistical Society in London.

Scott L Zeger is Professor and Chairman of the Department of Biostatistics at Johns Hopkins School of Public Health. He earned his doctorate in statistics from Princeton University. With Kung-Yee Liang, he developed the generalized estimating equation (GEE) approach to regression modeling with correlated data as occur in time-series or longitudinal studies. This work has been recognized with awards from both the American Statistical Association and American Public Health Association. Zeger's current research focuses on improving the design and analysis of data from prospective studies.

Joel Schwartz, Associate Professor of Environmental Epidemiology at Harvard School of Public Health and Harvard Medical School, received his doctorate from Brandeis University in theoretical physics and is a MacArthur Fellow. Schwartz has held appointments as a visiting scientist at the University of Basel, Switzerland, and at the University of Wupperal, Germany, and has served as a Senior Scientist at the US EPA. His research focuses on the health effects of air and water pollution, lead, and methodologic questions regarding the modeling of covariates in epidemiologic studies. Over the past decade, his studies have been instrumental in precipitating renewed interest in airborne particles.

Douglas W Dockery is Professor of Environmental Epidemiology at Harvard School of Public Health and Professor of Epidemiology at Harvard Medical School. He received his master's degree in meteorology from the Massachusetts Institute of Technology and his doctorate from Harvard School of Public Health in environmental health sciences. His research interests focus on the health effects of air pollution, asthma, lead toxicity, and water quality. He is currently president of the International Society of Environmental Epidemiology.

Frank Curriero received his PhD in statistics from Kansas State University. He is currently a postdoctoral fellow in the Biostatistics Department at Johns Hopkins School of Public Health. His research interests include spatial statistical analysis, interpretation and analysis of GIS-related public health data, and statistical computing.

Ivan Coursac graduated in statistics and economics from the National School of Statistics and Economics (ENSAE) in Paris, France. He is currently employed by PSA-Peugeot-Citroen. His research interests include air pollution, environmental health, and statistical computing.

Antonella Zanobetti received her PhD in 1999 in applied statistics from the University of Florence. She is currently working as a statistical analyst with Joel Schwartz in the Environmental Epidemiology Program at Harvard School of Public Health. Her field of research includes statistical analysis of environmental epidemiology data with particular interest in time-series analysis.

OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH

Daniels M, Dominici F, Samet JM, Zeger SL. 2000. Estimating PM_{10}-mortality dose-response curves and threshold levels: an analysis of daily time series for the 20 largest cities. Am J Epidemiol 152:397–406.

Dominici F, Samet J, Zeger S. 1999. Combining evidence on air pollution and daily mortality from the largest 20 US cities: A hierarchical modeling strategy. J R Stat Soc, Series A 163:263–302.

Dominici F, Zeger SL, Samet J. 2000. A measurement error for time-series studies of air pollution and mortality. Biostatistics 1:157–174.

Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. 2000. Fine particulate air pollution and mortality in 20 U.S. cities, 1987–1994. N Engl J Med 343:1742–1749.

Schwartz J. 2000. Harvesting and long-term exposure effects in the relationship between air pollution and mortality. Am J Epidemiol 15:440–448.

Schwartz J. 2000. The distributed lag between air pollution and daily deaths. Epidemiology 11:320–326.

Zeger S, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, Cohen A. 2000. Exposure measurement error in time-series studies of air pollution: Concepts and consequences. Environ Health Perspect 108:419–426.

Zeger SL, Dominici F, Samet J. 1999. Harvesting-resistant estimates of air pollution effects on mortality. Epidemiology 10:171–175.

ABBREVIATIONS AND OTHER TERMS

AIC	Akaike Information Criterion
AIRS	Aerometric Information Retrieval System
APHEA	Air Pollution and Health: A European Approach
CI	confidence interval
CO	carbon monoxide
COPD	chronic obstructive pulmonary disease
C_p	Mallows C_p statistic
CVD	cardiovascular disease
df	degrees of freedom
EPA	US Environmental Protection Agency
HCFA	Health Care Financing Administration
ICD-9	*International Classification of Diseases,* Ninth Revision
LOESS	locally weighted smoother
MCMC	Markov chain Monte Carlo
NAAQS	National Ambient Air Quality Standard
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
NO_2	nitrogen dioxide
O_3	ozone
PDL	polynomial distributed lag
PEEP	Particle Epidemiology Evaluation Project
PM	particulate matter
PM_{10}	particulate matter less than 10 μm in aerodynamic diameter
SAF	Standard Analytic Files
SO_2	sulfur dioxide

INTRODUCTION

Epidemiologic time-series studies conducted in a number of cities have found, in general, an association between daily changes in particulate matter (PM)* and daily number of deaths. Increased hospitalization of the elderly for specific diagnoses (a measure of morbidity) among the elderly also has been associated with ambient PM. These findings have raised concerns about public health effects of particulate air pollution and have contributed to decisions about regulating PM in the United States. However, scientists have pointed out a number of uncertainties that raise questions about interpretation of the results.

The *National Morbidity, Mortality, and Air Pollution Study* (NMMAPS)† was designed to improve our understanding of the association between PM less than 10 µm in aerodynamic diameter (PM_{10}) and adverse health effects, using a national database for the first time to address key uncertainties regarding these previous studies. NMMAPS Part I (Samet et al 2000) addresses methodologic issues, including uncertainties about exposure measurement error, the extent to which pollution-related mortality reduces life in the short term, and analysis of multisite data.

In NMMAPS Part I, the investigators developed a framework for considering exposure measurement error and, through an example, effectively showed that any bias introduced is likely to be toward the null (that is, the magnitude of the observed estimate of association is moved closer to an estimate of no association). Thus, the reported effect estimate will be lower than the true effect size. Second, the report presents two conceptually similar methods for assessing whether PM exposure moves death closer by more than a few days among susceptible individuals, a phenomenon referred to as *harvesting* or *mortality displacement*. Both methods showed an apparent effect of mortality beyond a few days, although the interpretation of that effect is not yet well understood. Finally, NMMAPS Part I developed multistage hierarchical methods to examine data from many cities. NMMAPS Part II uses these multicity analytic methods to assess the relationship between air pollution and mortality and morbidity, applied to the 20 largest US cities. Similar, but less complex hierarchical methods were applied to the 90 largest cities.

The analyses in NMMAPS II address questions about bias in selecting locations to study, differences in the statistical techniques applied, and adequacy of control for the effects of other pollutants on the associations between PM_{10} and morbidity and mortality. NMMAPS Part II overcame questions of bias in selection of cities to study (a criticism of previous studies conducted in individual cities) by having clear criteria for selection: population size and availability of data on PM_{10}. To address the concern that previous studies utilized different statistical techniques, NMMAPS Part II used one analytic approach to assess the PM_{10} effect on mortality and a similar approach to analyze the morbidity data. Because the cities included areas with varying levels of PM_{10} and other pollutants, the adequacy of controlling for these pollutants could be examined as well.

In NMMAPS Part II, the 20-city analysis examined the effects of copollutants in detail, and the analysis of 90 cities (which included the 20) explored possible heterogeneity (ie, variability) of the effect of PM_{10} on mortality among the cities. The PM_{10} effect on morbidity (hospitalizations of persons 65 years of age and older) for cardiovascular disease, chronic obstructive pulmonary disease (COPD), and pneumonia was explored in 14 cities; confounding by copollutants was also addressed.

OVERVIEW OF THE NATIONAL MORBIDITY, MORTALITY, AND AIR POLLUTION STUDY

The NMMAPS mortality analysis used the Aerometric Information Retrieval System (AIRS) of the US Environmental Protection Agency (EPA) to select cities meeting specific criteria of population size and availability of PM_{10} data. Daily mortality counts were obtained from the National Center for Health Statistics. The investigators estimated the average effect of PM_{10} on daily mortality in all 90 cities, variation in that effect, and the impact of copollutants on the estimated effect in 20 and 90 cities. Independent effects of the gaseous pollutants were also examined in detail for the 20 cities. Possible factors that might modify

* A list of abbreviations and other terms appears at the end of the Investigators' Report.

† Dr Jonathan Samet's investigation, *The National Morbidity, Mortality, and Air Pollution Study*, which will generate several reports, began in December 1996 and has cost about $700,000 to date. Part II of the Investigators' Report from Dr Samet and colleagues was received for review in October 1999. A revised report, received in January 2000, was accepted for publication in February 2000. During the review process, the HEI Review Committee and the investigators had the opportunity to exchange comments and to clarify issues in the Investigators' Report and in the Review Committee's Commentary.

the PM_{10} mortality effect estimates among cities, such as sociodemographic variables, were also considered.

A number of metaregression statistical methods, from simple to complex, are available for combining information in order to explore variations and effects (Normand 1999; Thompson and Sharp 1999). The hierarchical multistage modeling used in NMMAPS provided a flexible method to analyze multilevel data and a unified framework for combining results from individual cities and for examining the effect of other factors on the PM-mortality association. More complex methods were used to analyze the 20-city mortality data, and simpler methods were applied to the 90 cities. Data from the many cities were combined in a multistage or hierarchical statistical model. In the first stage, a time-series regression model for PM_{10} and mortality was fit for each city. This process included exploration of various days of effect (ie, relating mortality to the same day's exposure [lag 0], previous day's exposure [lag 1], etc) as well as the effects of adjusting for copollutants. In the second stage, the individual city estimates were combined. Covariables such as sociodemographic factors were added to the model to explore the extent to which they could explain the heterogeneity in pollutant effects across cities. A three-stage regional model using the 90 cities was also used to examine variability of PM_{10} effects on mortality across defined regions of the US.

The analysis of illnesses (morbidity) used the AIRS and Health Care Financing Administration (HCFA) databases to obtain air pollution and daily counts of urgent hospitalizations of those 65 years of age and older in 14 cities. Similar to the mortality analysis, the primary objective of the morbidity analysis was to improve on previous PM-hospitalization studies by analyzing many cities with one analytic method. Secondary objectives were to investigate consistency and determinants of variation in associations observed with control for gaseous pollutants and certain sociodemographic covariates.

The morbidity analysis examined the pattern of the effect of PM_{10} on increases in urgent hospitalizations using a distributed-lag approach, which enables the investigators to estimate the effects of PM_{10} over several days rather than for a specified day. This method was applied to 14 cities with daily PM_{10} data. For all analyses, three specific diagnoses most plausibly associated with PM_{10} were considered: cardiovascular disease, COPD, and pneumonia.

KEY RESULTS

MORTALITY

The investigators report an average approximate increase of 0.5% in total nonaccidental mortality per 10 $\mu g/m^3$ increase in PM_{10} concentration for both the 20 and 90 cities at 1 day after exposure (1-day lag). The PM_{10} effect was slightly greater for cardiorespiratory mortality than for total mortality at 1 day after exposure; effects at other times after exposure did not vary substantially from one another for total or cardiorespiratory mortality. For both the 20 and 90 cities, the association between PM_{10} and mortality did not appear to be sensitive to the inclusion of other pollutants in the model. Also, when other pollutants (sulfur dioxide [SO_2], nitrogen dioxide [NO_2], ozone [O_3], carbon monoxide [CO]) were considered for their independent association with mortality in 20 cities, some (CO, NO_2) showed associations, but these associations were not robust to the inclusion of PM_{10} in the model. The investigators also report that the effect of PM_{10} varied across regions in the US, with the largest effect observed in the Northeast.

MORBIDITY

The investigators found that the concentration of PM_{10} was positively associated overall with elderly hospital admissions in the 14 cities for the 3 diseases studied: cardiovascular disease, COPD, and pneumonia. Although nearly all cities showed a positive association between admissions and PM_{10}, the magnitude of the estimated effect on admissions varied considerably among cities.

On average, cardiovascular admissions increased by about 1% for every 10 $\mu g/m^3$ of PM_{10}, and pneumonia and COPD admissions rose by about 2% for the same increase in PM_{10}. Using the distributed lag model, the effect was distributed over several days; therefore selecting a specific day of exposure was not required. The investigators' report that the observed effects of PM_{10} were not confounded by the effects of other pollutants and were not associated with census measures of unemployment, poverty, nonwhite race, or college education in these cities.

The observed effect of PM_{10} concentration on hospital admissions persisted in analyses excluding days with PM_{10} concentrations above 50 $\mu g/m^3$: the effect estimate for each of the 3 diseases increased by about 20% over estimates that included the days with higher PM_{10} concentrations.

TECHNICAL EVALUATION

The Review Panel, comprised of members of the HEI Health Review Committee and outside experts, reviewed the NMMAPS Part II analyses in depth and agreed with the overall findings of both the mortality and morbidity analyses. Some additional figures and accompanying text were received prior to publication and were not reviewed by the panel (Figures 24, 25, and 32 and Appendix C). While the Panel welcomes these new analyses, time was inadequate to review them, and therefore the comments here do not address this material.

This project represents a major analytic effort to address comprehensively the association of PM and morbidity and mortality, and the Panel appreciates the extensive statistical computational expertise required. The mortality analysis applied a uniform statistical approach to many cities selected in an unbiased fashion to address issues and uncertainties that could not be assessed previously; this is therefore a major step forward in our understanding of PM_{10} effects on mortality.

The morbidity analysis used a method that obviated the need for preselecting a specific exposure lag period to examine PM_{10} effects, and consistently found an effect for the 3 diagnoses examined. This method appears to be a tool with analytic potential in this field.

In the course of reviewing NMMAPS Part II, the Panel identified several issues for discussion and offers some suggestions for improvements to consider for future analyses.

MORTALITY ANALYSIS

To investigate the PM_{10} effect on mortality, the authors used multistage Bayesian hierarchical analyses and simpler random-effects weighted regression analyses (DerSimonian and Laird 1986; Normand 1999; Thompson and Sharp 1999). Both are accepted statistical approaches, and both approaches can be categorized as random-effects (as compared to fixed-effects) meta-analytic techniques.

With both the weighted regression and Bayesian hierarchical methods, the first stage of modeling involved fitting individual time-series regressions to the air pollution and mortality data from each city (as described in NMMAPS Part I). Results from this stage are the estimated city-specific regression coefficients of mortality on the pollutant (eg, PM_{10}) and their standard errors. These results are then used as input to the stage 2 modeling based on either the Bayesian or weighted regression approach. A strength of this approach is that the investigators can systematically combine and draw statistical strength from each city's individual regression estimate in order to obtain regional and national estimates.

The Bayesian hierarchical approach was used in mortality analyses of the 20 cities and to a limited extent in the 90 cities. This method had specific advantages over the simple weighted regression used for most of the 90 cities and the morbidity analyses. Unlike the simple weighted regression approach, the Bayesian approach properly accounts for effect modifiers used as covariates in stage 2 when estimating the between-cities variance. In the 90-city analysis, however, variance between cities was estimated from a model with no effect modifiers. This estimate was then used for the subsequent weighted regression analyses regardless of how the effect modifiers actually might have reduced or explained the variance between cities.

Potential difficulties in implementing the Bayesian hierarchical approach include:

- its intensive computational demands,
- issues related to determining whether or not the Markov chain Monte Carlo (MCMC) simulations have properly converged, and
- the need to specify prior distributions for parameters in the model, which may influence the final results.

The intensive computation associated with MCMC was apparently the main reason why the simpler weighted regression methods were chosen for examination of effect modification in the 90 cities. This decision is supported by the demonstration (Figure 9 in the Investigators' Report) that the weighted regression and Bayesian hierarchical approaches provide similar point and interval estimates for the overall average effects of single pollutants at lag 1-day using the 90-city data. In NMMAPS Part I, the authors describe the statistical tools used to assess whether the MCMC simulations converged and also demonstrate that the prior distributions used are sufficiently uninformative as to have little impact on the overall results.

Multistage approaches vary in complexity, with the MCMC method used for the 20 cities being relatively complex and the method used for the 90 cities being relatively simple (Thompson and Sharp 1999). Such a large data set as used in NMMAPS presents a challenge to determine the appropriate level of modeling complexity to adequately reflect random variation while remaining simple and computationally tractable. Methods intermediate on this spectrum provide a compromise between complexity and simplicity. In NMMAPS, use of different methods probably would not have had a major impact on the results.

Both the 20- and 90-city analyses were conducted at lags of 0, 1, and 2 days for PM_{10} effects. The average approxi-

mate 0.5% increase in mortality for each 10 μg/m^3 increase in PM_{10} concentration for the 20-city analysis is consistent with the 90-city analyses, which appears to vary slightly over the lag periods analyzed. Additionally, assessment of socioeconomic factors in the 90 cities showed no evidence of effect modification. The 90-city analyses in Figures 24 and 25 were conducted after the Panel's full review of the report. In general, it appears that the conclusions concerning the estimated PM_{10} effect on mortality are consistent in both 20- and 90-city analyses.

MORBIDITY ANALYSIS

The analytic methods used to examine the PM_{10} effect on hospitalizations in the elderly were somewhat different from those used for the mortality analyses, but they did involve weighted multistage regression analysis. The main differences were the approaches to handling the time period of the PM_{10} effect and confounding by copollutants. The data were from 14 cities that had daily PM_{10} measurements.

Several statistical models were applied to the hospitalization data: each varied the approach to considering the time period of PM_{10} effect. The investigators examined effects for specified days, or lags, before hospitalization. They also used an *unconstrained distributed-lag model*, in which the overall effect of an increase in PM_{10} on a particular day is the sum of its impact on that day, and on each subsequent day considered for several days, which are estimated in models that include all PM_{10} lags simultaneously. With a distributed-lag method, investigators avoid (1) having to decide which day to use in estimating an effect or (2) having to interpret which lag is appropriate if findings are not the same for all lags. The investigators compared their results from the distributed-lag model to those from the other models, which include single day lag models (Table 14), and the quadratic distributed lag (Figure 28 and Table 15).

Results from both random and fixed effects models are also reported. However, because significant between-city heterogeneity was found, particularly for pneumonia and COPD, the random-effects estimates are preferred to the fixed effects estimates. The fixed effects estimates are appropriate only when there is no between-city variability; in other cases the confidence intervals based on the fixed effects will be too narrow. The PM_{10} effect on hospitalizations was clearly demonstrated for all three diseases. This effect persisted when days with PM_{10} above 50 μg/m^3 were excluded. However, it should be noted that only results of a fixed-effects analysis were reported for this analysis. Confidence intervals based on this analysis are likely too narrow; therefore, the observation that the effect of PM_{10} is greater when restricted to days less than 50 μg/m^3 should be viewed as suggestive, rather than definitive.

The evidence from the distributed-lag models shows that a day's admissions were affected by PM_{10} concentrations on more than one day prior to hospitalization (Figure 28). The Panel agrees with the NMMAPS investigators that using a specific lag time for exposure (for example, one day mean, or previous day mean) will underestimate the overall PM_{10} effect and should only be used to compare results with studies not using the distributed-lag approach. The similarity of the estimates of the quadratic versus unconstrained distributed-lag models suggests that the quadratic model provides a reasonable approximation to the lag-structure of the morbidity data for cardiovascular disease, COPD, and pneumonia.

The Panel views the distributed-lag approach as an advance in methods for handling the lag issue, but they are less confident in other aspects of the morbidity analysis. Effect-modification results in the morbidity analyses, given that only 14 cities could be included due to limited daily pollution data, are less convincing than analogous analyses for mortality. Although the PM_{10} effect was not modified by sociodemographic factors (Table 16), this should not be considered strong evidence against such modifying effects. The Panel agrees with the authors that associations of morbidity with air pollution might be useful to investigate in sociodemographic groups defined by finer geographic areas. The authors also note, importantly, that because the models implicitly assume that effects of air pollution multiply baseline rates of effect (admissions or mortality), the absence of modification of the pollution effect implies that the absolute additive excess rate of death due to air pollution is greater in groups with higher baseline death rates.

Further, the use of weighted stage 2 regression to assess effect modification is less sophisticated than the weighted regression methods used for the mortality analysis. The method of estimating the variance components of the model, which are then used to construct weights used for the regression may be statistically inefficient. Approaches such as maximum likelihood, restricted maximum likelihood, or Bayesian methods described by Thompson and Sharp (1999) might have been preferable although the extent to which differences in methods would impact the substantive results is difficult to assess.

Two additional methodologic issues arise as a result of using hospitalization data: when hospitalization occurs (that is, day of the week when people tend to be hospitalized), and comparison with a control diagnosis. Hospitalization data can present a challenge to the analyst for 2 reasons: weekend and holiday effects can be strong and

number of hospital admissions and air pollution concentrations are typically lower on weekends and holidays. Either of these factors might introduce a spurious effect if not adequately controlled for in the analyses. The investigators have addressed this problem by including only emergency admissions and by controlling for day of the week in the analysis. Holidays were not accounted for in the models, but given the relative rarity of holidays, it is not very likely that they would strongly confound the pollutant effect, although it is plausible.

The second issue involves including a control diagnosis in order to determine whether PM_{10} effects are specific to certain outcomes, which would strengthen the plausibility of these observed effects. If the effects are observed for cardiovascular disease, pneumonia, and COPD only, then the confidence should be greater than if effects are also observed for a control diagnosis such as gastrointestinal disorders. The findings from these 14 cities in NMMAPS would be strengthened if the PM_{10} effect were shown to be relatively specific for hypothesized hospitalization diagnoses.

COPOLLUTANTS AS CONFOUNDING VARIABLES

An important consideration in assessing the validity of the observed PM_{10} effects is whether they are due to PM_{10} itself or due to another air pollutant that is correlated with PM_{10}. That is, do effects of other pollutants confound the observed PM_{10} effect? The NMMAPS investigators took a commonly used approach to address this issue in the mortality analysis: does the addition of other air pollutant concentrations to the PM_{10} regression models result in any substantial change in the estimated PM_{10} effect? If the PM_{10} effect does not change, the other pollutants presumably have not confounded the observed PM_{10} effect.

The Panel identified a few issues related to possible confounding effects by copollutants although the probable impact of any of these was not considered to be sufficiently large to alter the observed PM_{10} effect. For example, when the investigators controlled for copollutants, they assumed the copollutants effect in the model to be linear. The use of more flexible smoothing terms might have allowed the investigators to explore whether the relationship was other than linear.

Another consideration is the impact of limiting assessment of the possible confounding effect to the relevant season for pollutants that have seasonal patterns. This assessment is complicated in these data because the seasonal effect of ozone, for example, is assumed to be somewhat different across the cities. Finally, a pollutant for which only inadequate data are available in the AIRS database, and which therefore could not be analyzed, might be responsible for the effects attributed to PM_{10}. Examples include sulfate or acid aerosols as specific components of the PM mixture.

Given these considerations, the Panel agrees that in the 20 cities no convincing evidence suggests that PM_{10} effects on mortality are changed by addition of either O_3, SO_2, NO_2, or CO concentrations to the models, suggesting that none of the other pollutants is responsible for the observed PM_{10} effects. Subsequent analyses by the investigators that appear to use similar techniques, controlled for gaseous pollutants in the 90 cities and did not show a confounding effect of the other pollutants.

In the morbidity analysis, when the investigators assessed the likelihood of confounding by other pollutants in stage 2 of the modeling, no evidence indicated that the PM_{10} effect on each diagnosis was confounded. This finding is similar to the finding in the mortality analysis, yet differences in the approach make it difficult to assess whether morbidity findings are as robust. While the approach used in the morbidity analysis is novel (comparing the PM_{10} regression coefficient with the regression coefficient between PM_{10} and the copollutants), the question arises as to the adequacy of statistical power for performing these analyses. Power may be low because the regression is fit to only 14 locations and in some cases 12 locations. Also, the power will be low when the regression coefficients between PM_{10} and the potentially confounding copollutants are similar across cities (Figure 29). These plots suggest that heterogeneity is sometimes due to only a few cities. In contrast, the mortality analysis examined confounding by copollutants in the first stage of modeling and demonstrated that it is possible to fit such models in spite of the correlation between copollutants. Application of the approach used in the mortality analysis to control for copollutants in future analyses of these morbidity data would be valuable for confirming the current findings.

INDEPENDENT EFFECTS OF COPOLLUTANTS

Although NMMAPS focuses on the effects of PM_{10}, examination of the independent effects of other pollutants is also warranted. Effects on daily mortality were found for most of the gaseous pollutants (SO_2, CO, NO_2) in the 20 cities although these effects were generally diminished when the model controlled for PM_{10} and other pollutants. In contrast, the PM_{10} effect did not appear to be affected by other pollutants in this model. An effect of each pollutant except ozone on mortality in the 90 cities is shown in Figure 9 of the Investigators' Report. The effect estimates and width of the confidence intervals represent a 10-unit change in concentration for each pollutant, each of which is measured in different units. This does not allow comparison of

the effect size and interval width across pollutants. A relatively strong and precise effect appears to be present for each of the gaseous pollutants in 90 cities in the analysis that assesses the effect of each pollutant alone. Therefore, the findings on the independent effects of the gaseous pollutants based on the 20 cities should be viewed as preliminary until such time as a 90-city analysis specifically controlling for PM_{10} and other pollutants is available.

The independent effect of the gaseous pollutants on hospitalizations was not assessed, but this information would also be of interest in any follow-up analyses.

REGIONAL VARIABILITY

The investigators conducted an important analysis drawing statistical strength from the individual 90-city estimates to examine the heterogeneity, or variation, in PM_{10} mortality effects across regions of the country. Differences in PM_{10} effect by region could provide additional insight into effects or features of the pollutant mix or the exposed population on PM_{10}-associated mortality. A visual inspection of the results appears to show the largest PM_{10} mortality effect in the Northeast with increased effects also apparent in the industrial Midwest and southern California. This limited NMMAPS assessment did not include a formal statistical test of heterogeneity and potential sources were not explored in depth. Material received by the Panel (Appendix C) prior to publication discusses some additional issues regarding heterogeneity; however, this analysis was not reviewed in detail by the Panel. The heterogeneity analysis has only scratched the surface of evaluating and understanding reasons for these apparent regional differences.

In the morbidity analysis, formal tests for heterogeneity (Table 14) indicated that statistically significant heterogeneity existed among cities for the effects of PM_{10} on COPD and pneumonia but not cardiovascular disease hospitalizations. Analyses conducted to examine modification of effect by copollutants (Figure 29) and sociodemographic factors (Table 16) did not identify any factors to explain the observed heterogeneity.

The Panel expressed concern about the emphasis on fixed effects estimates in the morbidity analyses, which underestimate uncertainty when the pollution effect varies among cities. Uncertainty in average pollution effects is thus greater than is suggested by the confidence intervals in Table 14. The methods used for the mortality analysis (Bayesian models and random effects regression) were random-effects models.

SENSITIVITY ANALYSES

One approach to enhance confidence in observational findings is to perform sensitivity analyses to determine whether findings are robust (ie, do not change substantially) across a range of reasonable analytic approaches. Examples of NMMAPS sensitivity analyses for mortality include: (a) those based on models using various lags, (b) those in which various degrees of smoothing were performed, (c) those based on models with and without control for other pollutants, and (d) those that used the 20 cities and the full set of 90 cities. These analyses suggest that the overall findings are not very sensitive to these analytic choices; thus we can have more confidence in the mortality results.

The sensitivity analyses are not as extensive for examining the PM_{10} effect on morbidity, and the investigators used a different time window across the 14 cities to control for temporal effects. In many of the cities, the window selected was considerably longer than 90 days, which has been suggested as the maximum for effective control of temporal effects (Cakmak et al 1998). Although the windows selected by the authors met their criterion of effective control (minimizing autocorrelation of residuals), spurious effects in the association of interest might be present due to inadequate control of longer-term temporal effects resulting from using relatively long windows. An analysis of PM_{10} morbidity effects using several time windows to adjust for temporal effects would have provided reassurance that these effects were adequately controlled. Sensitivity analyses for morbidity also included examination of the effects of different methods of specifying lags (such as by single day, quadratic distributed lag, and unrestricted distributed lag).

In a similar fashion, future analyses of both the mortality and morbidity data might include a seasonally stratified analysis (given the seasonal variability in pollutant concentrations, outcome measures, and potential confounding factors). Loss of statistical power due to the shorter periods of observation in any season should be only a minor issue, at least in the mortality data set.

AMBIENT POLLUTANT DATA

The investigators obtained air pollution data from the AIRS database, a computerized repository of information about air pollution in the United States. Operation of the monitoring equipment, collection and review of data, and assembly of the air quality database are the result of combined efforts of state and local environmental personnel in concert with EPA.

In NMMAPS, available ambient PM_{10}, O_3, CO, SO_2, and NO_2 data for 1987 to 1994 were used. In many of the cities, data were available from multiple pollutant monitors and were averaged across the monitors representing an area. Also, some pollutants that were measured on an hourly basis were averaged for 24 hours. To protect against the potential consequences of including outlying values, the investigators used an average of ambient data from which the highest 10% and lowest 10% of values were trimmed.

The monitoring sites used to collect the air quality data were established for the primary purpose of monitoring compliance with the National Ambient Air Quality Standards. Air pollution levels measured at a particular monitoring site may not represent population exposure for the entire county or urban area, however. Thus, the AIRS database here serves a secondary use and raises the issue of exposure measurement error and its potential effect on the results.

The framework for assessing exposure measurement error presented in NMMAPS Part I (Samet et al 2000) suggests that the precision and accuracy of the monitors make only small contributions to total exposure measurement error. The investigators conclude that the magnitude of the difference between the true ambient level and the average personal exposure is not known. Further, given that these errors are likely to be random in nature, the impact of measurement error on the findings reported in Part II using the AIRS database would most likely bias the effect estimates to the null. Thus estimates of positive findings in the analysis would tend to be lower than if fully accurate exposure measures were available.

In NMMAPS Part II, the investigators attempted to estimate the extent of measurement error by calculating the median of all pairwise correlations of PM_{10} measurements from the different monitors within each city or county. This factor was then included to investigate whether measurement error explained heterogeneity in the 90-city analysis. The Panel was concerned about the effectiveness of such an approach for several reasons. The median pairwise correlations, whether high or low, may simply reflect the spatial relationship and/or proximity of the monitors to each other and the prevailing annual wind direction. Also, median pairwise correlations based on annualized PM_{10} data could be misleading because the predominant seasonal wind direction can be quite different from the annualized wind direction. In light of the above, whether any correlation is particularly meaningful in assessing measurement error is unclear.

SUGGESTIONS FOR FUTURE ANALYSES OF NMMAPS DATA

Based on its review, the Panel suggests additional analyses of the NMMAPS data that would enhance the understanding to be gained from this substantial undertaking.

1. A more detailed analysis of the heterogeneity by region in the 90 cities seems warranted. Apparent differences in regional PM_{10} effect could be explored in a series of more in-depth analyses employing improved measures of city (or regional) measurement error, PM composition, data on copollutants mixes, and other data on population characteristics and susceptibilities.
2. Because the methods to examine PM effects were different for morbidity and mortality, questions may linger concerning the effect of analytic methods on the results. As the investigators have proposed, a detailed link of morbidity and mortality with the same statistical methods in the same cities would provide a powerful test of the consistency of findings presented in this report.
3. Examination of the independent effects of gaseous pollutants in the 90-city analyses, while controlling for PM_{10} and other pollutants, would be a natural extension of the 20-city analyses. The findings would likely confirm whether the gaseous pollutants also exert an independent effect on mortality.
4. Seasonally stratified analyses of morbidity and mortality would be beneficial because: (a) several factors influence population exposure to ambient air pollutants by season (such as air conditioning, open windows, time spent outdoors, and other factors); (b) sources of air pollution, including particles, often are seasonal (such as woodburning and other methods of home heating); (c) pollutants (such as the photo-oxidants, including ozone) require sunlight and other seasonal weather conditions in order to reach relevant concentrations; and (d) seasonal analyses do not require the analytic choices otherwise made to control for cyclical patterns. However, the Panel recognizes that seasonal analyses for more than one pollutant will be difficult given the differing seasonal patterns among the pollutants (eg, O_3 and PM_{10}).

DISCUSSION OF STUDY FINDINGS

The investigators use the term *strong* to characterize the degree to which their results provide evidence of an effect of increasing PM_{10} levels on morbidity and mortality. The Panel also concluded that the evidence for PM_{10} effects on both number of deaths and hospitalizations can be regarded as compelling and consistent. Planned paired analyses of morbidity and mortality using the same method, in the same locations, should enhance this finding with unique information regarding the comparability of PM_{10} effect on mortality and hospitalizations across cities.

The results relating to mortality and particulate air pollution also can be said to be strong in that they are robust: results were essentially the same regardless of the manner in which the statistical models were specified. The morbidity (hospitalizations) and PM_{10} findings were also positive for each of the diagnostic groups although they did not undergo the same degree of sensitivity analysis to assess robustness as the mortality analyses.

Another definition of strength relates to the size of the effect. Although the increase in risk shown in this report was small, a large number of people in the United States are exposed to PM_{10}. Therefore, the absolute public health impact of even a small relative increase in risk can be substantial. The strength of the potential public health impact, however, should not be confused with the size of the effect relative to background rates. The incremental increase in mortality related to a 10 $\mu g/m^3$ increase in PM_{10} was less than 1% for each lag examined in 20 and 90 cities. The incremental increase in risk for hospitalizations using a distributed lag for a 10 $\mu g/m^3$ increase in PM_{10} was somewhat larger, but still about 2% for respiratory admissions and about 1% for cardiovascular admissions.

The small magnitude of the increase in risk may limit application of this analytic approach to the surveillance of health effects of air pollution because mortality and hospitalization rates are influenced by many factors that may have greater impacts on health than PM_{10}. The health outcomes examined in NMMAPS are associated with a number of exposures other than air pollution, and a critical evaluation of whether surveillance can benefit from the NMMAPS analytic approach would be needed. Certainly the ability to assess the public health impact of changes in air quality over time would be desirable. The data sets utilized in NMMAPS will continue to be available and augmented, which will likely make them more attractive for these purposes in the future.

The heterogeneity of effect across cities offers the potential to identify factors that could influence the effect of PM_{10} on health and thus provide valuable insights into the mechanism by which PM_{10} causes adverse health effects. Evaluation of heterogeneity of PM_{10} effects among subgroups in the population would also be of interest but may require a different study design and collection of original data.

Finally, the investigators have reported that a larger effect per unit PM was found for all hospitalization diagnoses when analysis was restricted to lower PM_{10} concentrations. This finding suggests that the shape of the concentration-response plot is curvilinear and that no threshold exists. Since NMMAPS was not designed specifically to test this issue, the meaning of these findings should await completion of concentration-response analyses now under way.

CONCLUSIONS

The mortality analysis of 90 sampled cities, using the same methods and a unified approach to look at questions previously unexamined, is a major contribution beyond other attempts to combine data from many cities and understand the PM_{10}-mortality association. The results of both the 20-city and 90-city analyses are generally consistent with an average approximate 0.5% increase in overall mortality per 10 $\mu g/m^3$ increase in PM_{10}. Copollutants in the 20-city and 90-city analyses did not appear to affect the PM_{10}-mortality association. The difference, or heterogeneity, in PM_{10} effect by region in these results needs further elucidation.

The morbidity analysis also used a unified analytic method to provide a previously unavailable basis for examining the consistency of association of PM_{10} and hospital admissions and determinants of its variation. The results were consistent with an approximately 1% increase in cardiovascular admissions and about 2% increase for pneumonia and COPD per 10 $\mu g/m^3$ PM_{10}. The analysis demonstrated convincingly, using the distributed-lag approach, that numbers of hospital admissions were associated with pollution levels over several previous days. The distributed-lag approach will be an important tool in future analyses because it eliminates the need to choose a specific lag for analysis.

The observation that the PM_{10} effect on hospital admissions persisted at ambient concentrations below 50 $\mu g/m^3$ is important, but the greater effect per unit exposure at these lower concentrations should be considered suggestive rather than definitive. Given the uncertainties involved in calculating precise effect estimates with the models, the meaning of these preliminary findings should await completion of concentration-response analyses now under way. The PM_{10} effect did not appear to be confounded by

copollutants, but it is unclear whether the approach to assessing and controlling for confounding in the morbidity analysis had adequate statistical power for this purpose. No large modifying effects of sociodemographic factors were identified. Moderate effects cannot be ruled out because the power to detect effect modification was limited by the number of cities studied and because the available census data might not have adequately reflected the unknown vulnerabilities of the elderly in each city.

In conclusion, NMMAPS Part II presents results of a major study using new approaches to time-series analysis of PM_{10} health effects and has added substantially to our understanding of PM_{10} effects on mortality and morbidity. These findings also provide valuable suggestions for additional research that will further enhance our understanding of air pollution effects on health.

REFERENCES

Cakmak S, Burnett R, Krewsi D. 1998. Adjusting for temporal variation in the analysis of parallel time series of health and environmental variables. J Expo Anal Environ Epidemiol 8:129–144.

Der Simonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7:177–188.

Normand ST. 1999. Meta analysis: Formulating, evaluating, combining, and reporting. Stat Med 18:331–359.

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000. The National Morbidity, Mortality, and Air Pollution Study, Part I. Methods and Methodologic Issues. Research Report 94. Health Effects Institute, Cambridge MA.

Thompson SG, Sharp SJ. 1999. Explaining heterogeneity in meta-analysis: A comparison of methods. Stat Med 18:2693–2708.

RELATED HEI PUBLICATIONS: PARTICULATE MATTER

Report No*	Title	Principal Investigator	Publication Date
Research Reports			
40	Retention Modeling of Diesel Exhaust Particles in Rats and Humans	CP Yu	1991
45	The Effects of Exercise on Dose and Dose Distribution of Inhaled Automotive Pollutants	MT Kleinman	1991
68	Pulmonary Toxicity of Inhaled Diesel Exhaust and Carbon Black in Chronically Exposed Rats		
	Part I. Neoplastic and Nonneoplastic Lung Lesions	JL Mauderly	1995
	Part II. DNA Damage	K Randerath	1995
	Part III. Examination of Possible Target Genes	SA Belinsky	1995
75	Ozone Exposure and Daily Mortality in Mexico City: A Time-Series Analysis	DP Loomis	1996
83	Daily Changes in Oxygen Saturation and Pulse Rate Associated with Particulate Air Pollution and Barometric Pressure	DW Dockery	1999
86	Statistical Methods for Epidemiologic Studies of the Health Effects of Air Pollution	W Navidi	1999
91	Mechanisms of Morbidity and Mortality from Exposure to Ambient Air Particles	JJ Godleski	2000
93	Effects of Concentrated Ambient Particles in Rats and Hamsters: An Exploratory Study	T Gordon	2000
94	The National Morbidity, Mortality, and Air Pollution Study *Part I.* Methods and Methodologic Issues	JM Samet	2000
95	Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations	M Lippmann	2000
96	Acute Pulmonary Effects of Ultrafine Particles in Rats and Mice	G Oberdörster	2000
97	Identifying Subgroups of the General Population that May Be Susceptible to Short-Term Increases in Particulate Air Pollution: A Time-Series Study in Montreal, Quebec	MS Goldberg	2000
98	Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany *Part I.* Role of Particle Number and Particle Mass	H-E Wichmann	2000
99	A Case-Crossover Analysis of Fine Particulate Matter Air Pollution and Out-of-Hospital Sudden Cardiac Arrest	H Checkoway	2000
Special Reports			
	Particulate Air Pollution and Daily Mortality: The Phase I Report of the Particle Epidemiology Evaluation Project		
	Phase I.A. Replication and Validation of Selected Studies		1995
	Phase I.B. Analyses of the Effects of Weather and Multiple Air Pollutants		1997
	Reanalysis of the Harvard Six Cities Study and American Cancer Society Study of Particulate Air Pollution and Mortality: A special Report of the Institute's Particle Epidemiology Reanalysis Project		2000

** Reports published since 1990. Copies of these reports can be obtained by contacting the Health Effects Institute, 955 Massachusetts Avenue, Cambridge MA 02139. Phone (617) 876-6700 FAX (617) 876-6709 E-mail pubs@healtheffects.org www.healtheffects.org*

RELATED HEI PUBLICATIONS

Research Reports

Report No.	Title	Principal Investigator	Publication Date
HEI Communications			
5	Formation and Characterization of Particles: Report of the 1996 HEI Workshop		1997
8	The Health Effects of Fine Particles: Key Questions and the 2003 Review		1999
HEI Program Summaries			
	Research on Particulate Matter		1999

Copies of these reports can be obtained by contacting the Health Effects Institute, 955 Massachusetts Avenue, Cambridge, MA 02139. Phone (617) 876-6700 FAX (617) 876-6709 E-mail pubs@healtheffects.org www.healtheffects.org

HEALTH
EFFECTS
INSTITUTE

955 Massachusetts Avenue
Cambridge MA 02139 USA
+1-617-876-6700
www.healtheffects.org

RESEARCH
REPORT

Number 94, Part II
June 2000