RESEARCH REPORT

HEALTH EFFECTS INSTITUTE

Number 157
March 2011

Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Two Indian Cities

HEI Public Health and Air Pollution in Asia Program

Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Two Indian Cities

HEI Public Health and Air Pollution in Asia Program

with a Critique by the HEI Health Review Committee

Research Report 157
Health Effects Institute
Boston, Massachusetts

Trusted Science • Cleaner Air • Better Health

Publishing history: This document was posted at *www.healtheffects.org* in March 2011.

Citation for document:

Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Two Indian Cities. Research Report 157. Health Effects Institute, Boston, MA.

Citation for Part 1 only:

Balakrishnan K, Ganguli B, Ghosh S, Sankar S, Thanasekaraan V, Rayudu VN, Caussy H. 2011. Part 1. Short-term effects of air pollution on mortality: Results from a time-series analysis in Chennai, India. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Two Indian Cities. Research Report 157. Health Effects Institute, Boston, MA.

Citation for Part 2 only:

Rajarathnam U, Seghal M, Nairy S, Patnayak RC, Chhabra SK, Kilnani, Santhosh Ragavan KV. 2011. Part 2. Time-series study on air pollution and mortality in Delhi. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Two Indian Cities. Research Report 157. Health Effects Institute, Boston, MA.

 Asterisk Typographics, Barre, Vt., Compositor. Printed by Recycled Paper Printing, Boston, Mass. Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.

Cover paper: made with at least 55% recycled content, of which at least 30% is post-consumer waste; free of acid and elemental chlorine. Text paper: made with 100% post-consumer waste recycled content; acid free; no chlorine used in processing. The book is printed with soy-based inks and is of permanent archival quality.

CONTENTS

Research Report 157

Research Report 157

ABOUT HEI

The Health Effects Institute is a nonprofit corporation chartered in 1980 as an independent research organization to provide high-quality, impartial, and relevant science on the effects of air pollution on health. To accomplish its mission, the institute

- Identifies the highest-priority areas for health effects research;
- Competitively funds and oversees research projects;
- Provides intensive independent review of HEI-supported studies and related research;
- Integrates HEI's research results with those of other institutions into broader evaluations; and
- Communicates the results of HEI's research and analyses to public and private decision makers.

HEI receives half of its core funds from the U.S. Environmental Protection Agency and half from the worldwide motor vehicle industry. Frequently, other public and private organizations in the United States and around the world also support major projects or certain research programs. The Public Health and Air Pollution in Asia (PAPA) Program was initiated by the Health Effects Institute in part to support the Clean Air Initiative for Asian Cities (CAI-Asia), a partnership of the Asian Development Bank and the World Bank to inform regional decisions about improving air quality in Asia. Additional funding was obtained from the William and Flora Hewlett Foundation.

HEI has funded more than 280 research projects in North America, Europe, Asia, and Latin America, the results of which have informed decisions regarding carbon monoxide, air toxics, nitrogen oxides, diesel exhaust, ozone, particulate matter, and other pollutants. These results have appeared in the peer-reviewed literature and in more than 200 comprehensive reports published by HEI.

HEI's independent Board of Directors consists of leaders in science and policy who are committed to fostering the public–private partnership that is central to the organization. The Health Research Committee solicits input from HEI sponsors and other stakeholders and works with scientific staff to develop a Five-Year Strategic Plan, select research projects for funding, and oversee their conduct. The Health Review Committee, which has no role in selecting or overseeing studies, works with staff to evaluate and interpret the results of funded studies and related research.

All project results and accompanying comments by the Health Review Committee are widely disseminated through HEI's Web site (*www.healtheffects.org*), printed reports, newsletters and other publications, annual conferences, and presentations to legislative bodies and public agencies.

ABOUT THIS REPORT

Research Report 157, *Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Two Indian Cities*, presents two studies funded by the Health Effects Institute. This report contains these main elements:

The HEI Statement, prepared by staff at HEI, is a brief, nontechnical summary of the two studies and their findings; it also briefly describes the Health Review Committee's comments on the studies.

The Investigators' Reports on the two studies describe the scientific background, aims, methods, results, and conclusions of each of the studies.

The Critique is prepared by members of the Health Review Committee with the assistance of HEI staff; it places the study in a broader scientific context, points out their strengths and limitations, and discusses remaining uncertainties and implications of the findings for public health and future research.

The two studies contained in Research Report 157 have gone through HEI's rigorous review process. When an HEI-funded study is completed, the investigators submit a draft final report presenting the background and results of the study. This draft report is first examined by outside technical reviewers and a biostatistician. The report and the reviewers' comments are then evaluated by members of the Health Review Committee, an independent panel of distinguished scientists who have no involvement in selecting or overseeing HEI studies. During the review process, the investigators have an opportunity to exchange comments with the Review Committee and, as necessary, to revise their report. The Critique reflects the information provided in the final version of the Investigators' Reports.

PREFACE

Coordinated Time-Series Studies in Asian Cities

Exposure to outdoor air pollution is associated with short-term increases in daily mortality, higher rates of hospital admissions, increases in emergency room utilization, and exacerbation of chronic respiratory conditions in many parts of the world (WHO 2002). The World Health Organization estimates that air pollution contributes to approximately 800,000 deaths and 4.6 million lost life-years annually (WHO 2002).

Developing nations are particularly affected by air pollution; as many as two-thirds of the deaths and lost life-years associated with air pollution on a global scale occur in Asia (WHO 2002). To date, estimates of the health effects resulting from exposure to air pollution in Asia have relied largely on the extrapolation of results from research conducted outside Asia — primarily in Europe and North America (Cohen et al. 2004). However, the nature of the ambient air pollution mix in Asia, the high levels of pollutants in some parts of the continent, the environmental conditions, and the background health conditions of the population may all contribute to health outcomes that differ from those in Europe and North America.

To address some of the uncertainties in estimating the adverse health impact of air pollution in Asia, the Health Effects Institute, in partnership with the Clean Air Initiative for Asian Cities, initiated the Public Health and Air Pollution in Asia (PAPA) program in December 2002. The PAPA program has three major components:

1. assessment and review of existing science on the effects of exposure to air pollution in Asia;
2. initiation of significant new research in several large Asian cities, including major new epidemiologic studies on the health effects of air pollution; and
3. development of the scientific and technical capacity of a network of Asian investigators, including targeted opportunities for training in epidemiology and related areas.

HEI set up the International Scientific Oversight Committee (ISOC), chaired by Dr. Frank Speizer of the Harvard School of Public Health and comprising members of HEI's Research and Review Committees and experts from the United States and Asia, in order to provide expert scientific advice and oversight of the PAPA program. A full list of ISOC members is included at the end of this report.

PAPA FIRST-WAVE STUDIES

In 2003, HEI issued a Request for Information and Qualifications (RFIQ) for scientists interested in conducting epidemiologic studies of the health effects of air pollution in Asian cities. Through this RFIQ, HEI sought to ascertain potential investigators' qualifications and their access to appropriate study populations, pollutant monitoring data — such as levels of particulate matter (PM), carbon monoxide (CO), and ozone (O_3) — and health data (mortality and morbidity) in Asian cities. Thirty-two teams in eight Asian countries responded to the RFIQ.

ISOC evaluated all the proposals and decided to support a coordinated series of time-series studies in several Asian cities. In addition, ISOC decided to request applications from investigators who appeared to have both the best qualifications and quick access to the necessary information, allowing them to begin studies within a short period after receiving funding.

The PAPA program ultimately initiated four time-series studies of the health effects of air pollution in Bangkok, Hong Kong, Shanghai, and Wuhan. This was the first set of coordinated time-series studies ever undertaken in Asian cities and the first phase of an effort by ISOC to conduct a series of studies in Asian cities intended to deepen understanding of air pollution effects in local populations and inform extrapolation from the extensive body of existing science.

This "first wave" of PAPA studies comprised four time-series studies:

- **"Estimating the Effects of Air Pollution on Mortality in Bangkok, Thailand."** Dr. Nuntavarn Vichit-Vadakan at Thammasat University in Thailand and her team proposed to examine the effects of PM ≤ 10 μm in aerodynamic diameter (PM_{10}) and several gaseous pollutants — O_3, nitrogen dioxide (NO_2), nitric oxide (NO), and sulfur dioxide (SO_2) — on daily mortality for the years 1997 through 2003 and for all 50 districts of Bangkok, which had a population of 10.4 million in 2004.
- **"Interaction Between Air Pollution and Respiratory Viruses: Time-Series Study of Daily Mortality and Hospital Admissions in Hong Kong."** Dr. Chit-Ming Wong of the University of Hong Kong and his team proposed to examine the short-term effects of PM_{10}, NO_2, SO_2, and O_3 on mortality and hospital admissions over the period 1996 through 2002. The confounding and modifying effects of influenza epidemics were also to be assessed. The study included the whole Hong Kong population of 6.8 million.
- **"A Time-Series Study of Ambient Air Pollution and Daily Mortality in Shanghai, China."** Dr. Haidong Kan from the Fudan University School of Public Health in China and his team proposed to evaluate the association between mortality and major air pollutants (PM_{10}, SO_2, NO_2, and O_3), using daily data from 2001 through 2004. The target population was all residents living in the urban area of Shanghai, which covers nine districts and encompasses a population of more than 6 million.
- **"Association of Daily Mortality with Ambient Air Pollution, and Effect Modification by Extremely High Temperature in Wuhan, China."** Dr. Zhengmin Qian at the Pennsylvania State College of Medicine in Hershey, Pennsylvania, and his team proposed to determine whether daily variations in ambient PM_{10}, SO_2, NO_2, or O_3 concentrations in Wuhan (with 4.5 million permanent residents in the nine urban core districts) from July 1, 2000, through June 30, 2004, were associated with variations in daily mortality due to all natural causes and daily cause-specific mortality.

These four studies were recently published together as HEI Research Report 154, parts 1 through 4 (Kan et al. 2010; Qian et al. 2010; Vichit-Vadakan et al. 2010; Wong et al. 2010a).

A COORDINATED APPROACH TO ANALYSIS

Coordinated multicity studies currently provide the most definitive epidemiologic evidence of the effects of short-term exposure and, as a result, play a central role in health impact assessment and environmental policy. While robust and consistent results have been observed in Europe and North America (Samet et al. 2000; Katsouyanni et al. 2001), few coordinated, multicity time-series studies have been conducted elsewhere. The four PAPA studies conducted in China and Thailand are the first coordinated multicity analyses of air pollution and daily mortality in Asia.

The principal investigators developed a common set of criteria for the inclusion and analysis of data in each city; this common protocol was codified in a "Protocol for Coordinated Time-Series Studies of Daily Mortality in Asian Cities," which is included at the end of HEI Research Report 154. In addition, at the end of the four studies, the investigators, led by Dr. C.-M. Wong, undertook a combined analysis, incorporating data from all four cities. This combined analysis is included in HEI Research Report 154 as Part 5 (Wong et al. 2010b).

STUDIES IN INDIA

In recognition of the fact that India is a diverse, densely populated country where the burden of disease attributable to ambient air pollution is likely to be substantial, HEI issued a request for applications (RFA) for retrospective time-series studies of air pollution and mortality in Indian cities in the spring of 2004. This RFA was intended to facilitate a set of Indian time-series studies that would be an important addition to the group of studies in progress at that time in other Asian cities.

Proposed studies would utilize existing sources of data to explore how daily death rates change in relation to contemporaneous daily concentrations of air pollutants, while correcting for other risk factors. Investigators were requested to submit information related to the nature and availability of monitoring data for air pollution, including major air pollutants measured (e.g., PM, O_3, and CO) and visibility, as well as health data on measures of mortality (and morbidity, where available). In order to maximize the capabilities of researchers from different disciplines, it was strongly recommended that interdisciplinary teams of scientists apply together.

HEI subsequently extended the PAPA research program to include three studies of air pollution and mortality due to all natural causes in Chennai, Delhi, and Ludhiana. The Indian studies focused on the association between increased air pollution and all natural (nonaccidental) mortality from 2002 though 2004. The studies were as follows:

- **"Short-Term Effects of Air Pollution on Mortality: Results from a Time-Series Analysis in Chennai, India."** Dr. Kalpana Balakrishnan from Sri Ramachandra Medical College and Research Institute and her team explored the association between air pollution and all-natural-cause mortality in Chennai, a city in Southern India.
- **"Time-Series Study on Air Pollution and Mortality in Delhi."** Dr. Uma Rajarathnam at The Energy and Resources Institute and her team explored the association between air pollution and all-natural-cause mortality in Delhi, the nation's capital.
- **"A Time-Series Study on the Relation of Air Pollution and Mortality in Ludhiana City, India."** Dr. Rajesh Kumar of the Postgraduate Institute of Medical Education and Research in Chandigarh and his team proposed to explore the association between air pollution and all-natural-cause mortality in Ludhiana, an industrial city in northern India.

Because of key differences in the availability and completeness of data between the first four PAPA studies and the Indian studies, the common protocol developed for the first four studies was only partially applicable for use in the Indian context. Indeed, the Ludhiana study was terminated prematurely because of substantial limitations and uncertainties in the data that made it unclear whether an interpretable result would be possible. Drs. Balakrishnan and Rajarathnam and their coinvestigators developed city-specific approaches for using available air quality data to develop daily estimates of exposure.

The PAPA studies in representative cities in Thailand, China, and India were funded in order to bridge the gap between studies conducted in different localities with the goal of providing information to Asian policy makers. The PAPA studies were designed and conducted by local investigators in concert with local air pollution and public health officials and international experts. These studies explore key aspects of the epidemiology of exposure to air pollution in each location — issues of local as well as global relevance — including the effects of exposure at high concentrations and at high temperatures, the potential influence of influenza epidemics on the relationship between air pollution and health, and the ways in which social class might modify risks associated with air pollution.

REFERENCES

Cohen AJ, Anderson HR, Ostro B, Pandey KD, Krzyzanowski M, Künzli N, Gutschmidt K, Pope CA III, Romieu I, Samet JM, Smith KR. 2004. Urban air pollution. In: Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. (Ezzati M, Lopez AD, Rodgers A, Murray CJL, eds.). pp. 1153–1433. World Health Organization, Geneva, Switzerland.

Kan H, Chen B, Zhao N, London SJ, Song G, Chen G, Zhang Y, Jiang L. 2010. Part 1. A time-series study of ambient air pollution and daily mortality in Shanghai, China. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 12:521–531.

Qian Z, He Q, Lin H-M, Kong L, Zhou D, Liang S, Zhu Z, Liao D, Liu W, Bentley CM, Dan J, Wang B, Yang N, Xu S, Gong J, Wei H, Sun H, Qin Z. 2010. Part 2. Association of daily mortality with ambient air pollution, and effect modification by extremely high temperature in

Wuhan, China. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

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

Vichit-Vadakan N, Vajanapoom N, Ostro B. 2010. Part 3. Estimating the effects of air pollution on mortality in Bangkok, Thailand. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Wong C-M, Thach TQ, Chau PYK, Chan EKP, Chung RY-N, Ou C-Q, Yang L, Peiris JSM, Thomas GN, Lam T-H, Wong T-W, Hedley AJ. 2010a. Part 4. Interaction between air pollution and respiratory viruses: Time-series study of daily mortality and hospital admissions in Hong Kong. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Wong C-M on behalf of the PAPA teams: Bangkok, Hong Kong, Shanghai, and Wuhan. 2010b. Part 5. Public Health and Air Pollution in Asia (PAPA): A Combined Analysis of Four Studies of Air Pollution and Mortality. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

World Health Organization. 2002. World Health Report 2002: Reducing Risk, Promoting Healthy Life. World Health Organization. Geneva, Switzerland. Available from *www.who.int/whr/2002/en/*.

HEI STATEMENT

Synopsis of Research Report 157

Air Pollution and Mortality in Two Indian Cities

BACKGROUND

Epidemiologic time-series studies have provided useful information about the association between short-term exposure to air pollution and health outcomes in single geographic locations. However, it is often challenging to compare results across locations because of differences among the methods used to collect and analyze the data. Recent multicity time-series studies have been conducted in cities in Canada, Europe, and the United States using similar analytic approaches, but considerable uncertainties remain regarding the extrapolation of this evidence to developing countries, including large populations in Asia that are exposed to relatively high concentrations of air pollution.

Given the extensive gaps in knowledge on the health effects of air pollution in Asia, HEI funded several coordinated time-series studies of air pollution and health in Asia under the Public Health and Air Pollution in Asia (PAPA) Program. The purpose of the program was to collect information that would be relevant to local populations, with the added goal of supporting capacity building in the region. After a first wave of four studies in China and Thailand was underway (research that was recently published as HEI Research Report 154), HEI's International Scientific Oversight Committee selected three additional investigator teams to conduct time-series studies of air pollution and mortality in Indian cities: Chennai, Delhi, and Ludhiana. The investigators coordinated their approaches by adapting the common protocol of the first-wave PAPA studies; differences in data availability and completeness in the Indian cities prompted the investigators to develop city-specific approaches. Substantial deficiencies in the available air pollution data in Ludhiana prevented that study from being completed, although recently those investigators estimated mortality risks using visibility as a surrogate for air pollution in that city.

APPROACH

The investigator teams in Chennai and Delhi conducted time-series studies on the relationship between daily all-cause mortality and daily concentrations of particulate matter less than or equal to 10 µm in aerodynamic diameter (PM_{10}) for the period between January 1, 2002, and December 31, 2004. Mortality data were provided by local registries of births and deaths in each of the two cities and were coded by trained medical professionals to exclude deaths from non-natural causes. Pollutant data for nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and PM_{10} were provided by the local government agencies in each city and met local quality control and assurance standards. Investigators initially followed an independent, standardized procedure to ensure both the completeness and representativeness of the average daily exposure of the population. Subsequently, the investigators developed alternative exposure models for their specific location using a novel zonal approach in Chennai and centering techniques in Delhi. Gaseous pollutants were included only in the Delhi analyses. Data from 5 and 10 air quality monitoring stations were included in Chennai and Delhi, respectively.

The teams in Chennai and Delhi used a generalized additive modeling approach to obtain the excess risk of daily mortality associated with daily increases in pollutant concentrations. The investigators fitted quasi-Poisson regression models to the data and carried out sensitivity analyses, such as the inclusion of various degrees of freedom for model parameters, different temperature lags, and alternative exposure models.

This Statement, prepared by the Health Effects Institute, summarizes two research projects funded by HEI. The Chennai study was conducted by Dr. Kalpana Balakrishnan at Sri Ramachandra University, Porur, Chennai, India, and colleagues. The Delhi Study was conducted by Dr. Uma Rajarathnam at The Energy and Resources Institute, New Delhi, India, and colleagues. Research Report 157 contains both of the detailed Investigators' Reports and a Critique of the studies prepared by the Institute's Health Review Committee.

RESULTS

Using the core zonal model, the Chennai investigators reported an increase in the relative risk (RR) for nonaccidental, all-cause mortality of 1.004 (95% confidence interval [CI] = 1.002 to 1.007) per 10-μg/m^3 increase in PM_{10} concentration on the previous day. Alternative, nonzonal exposure series using single- or multiple-monitor models did not appreciably change the RR for mortality. Additional sensitivity analyses of the core zonal model results in Chennai showed that there was little variation in RR between males and females, but that the RRs for the age groups 5–44 years and 45–64 years were slightly higher than for the younger or older age groups. The RR for all-cause mortality associated with PM_{10} was slightly lower at lag 0 (same day) than at lag 1 (previous day) and was elevated at 2- or 3-day lags. No change in RR was found with a 7-day distributed lag for temperature and relative humidity compared with the core model. Varying the degrees of freedom for time, temperature, and relative humidity, stratifying by season, or excluding outliers in the exposure series did not appreciably change the RRs.

Using the core model with centered air quality data, the Delhi investigators reported an increase in the RR for nonaccidental, all-cause mortality of 1.0015 (95% CI = 1.0007 to 1.0023) per 10-μg/m^3 increase in PM_{10} concentration on the same day. The RR was slightly lower at 1 to 3 day lags and slightly higher at a cumulative lag of 0–1 days and when PM_{10} concentrations exceeding 400 μg/m^3 were excluded from the analysis. The investigators also reported an increase in RR for daily mortality of 1.0084 (95% CI = 1.0029 to 1.014) associated with a 10-μg/m^3 increase in NO_2 concentration; there was no evidence of an association between SO_2 concentration and mortality. Using two- and three-pollutant models, the RRs associated with PM_{10} and NO_2 concentrations were slightly attenuated.

Additional sensitivity analyses of the core model results in Delhi showed that the RR associated with PM_{10} was lower in males than females; the RR for the age group 5–44 years was higher than for the other age groups. There was considerable variability in RRs associated with PM_{10} among individual monitoring stations. However, the RRs for analyses that included the data from all 10 stations were similar to those for analyses that excluded the data from the one station with continuous data (which reported consistently higher pollutant concentrations). Using different lags for temperature did not appreciably change the RR, except for a decrease at longer cumulative lags (0–7 or 8–14 days).

TECHNICAL EVALUATION

The two time-series studies of air pollution and daily mortality in Chennai, India, and Delhi, India, have provided useful additional information on air pollution and health outcomes in developing countries. Results from Chennai (0.4% increase in risk per 10-μg/m^3 increase in PM_{10} concentration) and Delhi (0.15% increase in risk per 10-μg/m^3 increase in PM_{10} concentration) suggest a generally similar risk of mortality associated with PM_{10} exposure compared with the first four PAPA studies, as well as with multicity studies conducted in South Korea, Japan, Europe, and North America. The association of mortality with exposure to NO_2 observed in Delhi was also similar to values reported from other studies in Asia. The associations were fairly consistent in spite of the fact that concentrations of criteria air pollutants were substantially higher than those observed in the United States and Europe.

In its independent assessment of the Chennai study, the HEI Health Review Committee thought that the investigators had applied an innovative approach for exposure estimation in Chennai, assigning nearest ambient monitor values to each grid cell in the 10 zones to address heterogeneity in PM_{10} concentrations among monitors. Because the results from the core zonal model and alternative exposure series were fairly consistent, the Committee thought it would be unlikely that a different approach to exposure estimation would have yielded very different risk estimates, adding to confidence in the risk estimates. However, the Committee identified several caveats regarding the zonal approach. For example, the zonal model appears to ignore missing data, which may lead to spurious jumps in exposure estimates when data from one monitor are missing for a period; there may be important spatiotemporal variation even within zones; and monitors are not weighted in proportion to the population, because the population is unlikely to be uniformly distributed across grids.

In its evaluation of the Chennai mortality data analysis model, the Committee noted that the investigators used a relatively new, ambitious approach to select degrees of freedom for smoothing of time and meteorology covariates. This approach minimizes PM_{10} coefficient variability and bias using a resampling method. Whether this goal is achievable is uncertain and is the topic of ongoing HEI-funded work by James M. Robins and colleagues. The Committee commented that the degrees of freedom selected for temperature and relative humidity seemed high, which may indicate that those variables are strong predictors of pollution concentrations. The Committee noted that it is not advisable to use different degrees of smoothing for individual monitors, in particular because of the high amount of missing data at a given site. However, because the final degrees of freedom were relatively similar across the different approaches, it is unlikely that the amount of temporal smoothing would substantially affect the results.

In its assessment of the Delhi study, the Committee thought that the use of a centering approach was appropriate for estimating exposure in the context of missing data, although other approaches could have been tried and compared. The Committee noted that there may have been residual confounding by weather in the main analyses. Recent insights reveal that more complex distributed lag models may be required to capture residual confounding by temperature effects in many cities. The Committee thought the Delhi team had conducted a limited but informative set of sensitivity analyses to assess the effects on mortality risk estimates. As was observed for Chennai, the results may be sensitive to temperature lag. Risk estimates based on single-monitor pollution data were sensitive to which single monitor was included in the sensitivity analyses, providing further support for averaging pollutant data across multiple monitors to estimate population exposure in the core model.

The Committee noted some unusual features of the statistical models used by the Delhi team. The degrees of freedom that describe the smoothness of the mortality function against temperature, relative humidity, and PM_{10} in sensitivity analyses (for the PM_{10}–mortality curve) were specified per year. However, the complexity of these curves (e.g., the temperature–mortality curves) was unlikely to increase over the duration of the study. Normal practice is to select degrees of freedom as a function of the number of years only for smooth functions of time.

DISCUSSION AND CONCLUSIONS

The broad general consistency of the results of the two Indian studies described here and in other Asian time-series studies of mortality with those in Europe and North America is reassuring. It suggests that the continued use of data from Western cohort studies to estimate the Asian burden of disease attributable to short-term exposure to air pollution is defensible. However, developing Asia currently differs from the United States and Europe with regard to energy use, air quality, and population health, which are also dynamically changing. The Indian studies highlight that regional differences in demographics (in particular, age structure and general health status of the population) may affect health outcomes of interest. Thus, estimates of the risk of mortality associated with air pollution that are based on even the most carefully executed U.S. studies must be used with appropriate caveats.

Given the data limitations faced by the investigator teams, they are to be commended for making the most of limited resources. They have blazed a trail for improved quality epidemiologic studies of air pollution in India. However, considerable uncertainties remain due to data limitations, potential residual confounding, and potential methodologic sensitivity, all of which will need to be revisited in any future epidemiologic studies. As the investigators pointed out, data limitations prevented a number of more in-depth analyses standard in time-series studies, for example, of specific causes of death or — in Chennai — different pollutant models. Such detailed analyses will only be possible once more detailed, consistent air pollution monitoring and health record collection are implemented.

The methodology applied in the PAPA time-series studies can provide a stronger foundation for further research in developing Asia. The lack of data on air quality and mortality, especially cause-specific mortality, remains a major impediment to conducting such studies in many parts of developing Asia. As a result, major population centers in South and Southeast Asia (India, Pakistan, Vietnam, Philippines, Indonesia, and Malaysia) remain understudied — although the two PAPA studies in India are starting to fill in some of those gaps. Expanded, coordinated multicity studies conducted across Asia could provide more definitive answers if they are designed and analyzed consistently with the additional methodologic improvements noted above and given rigorous quality control of air quality and health data.

Part 1

Short-Term Effects of Air Pollution on Mortality: Results from a Time-Series Analysis in Chennai, India

Kalpana Balakrishnan, Bhaswati Ganguli, Santu Ghosh, S. Sankar,
Vijaylakshmi Thanasekaraan, V.N. Rayudu, Harry Caussy

Short-Term Effects of Air Pollution on Mortality: Results from a Time-Series Analysis in Chennai, India

Kalpana Balakrishnan, Bhaswati Ganguli, Santu Ghosh, S. Sankar, Vijaylakshmi Thanasekaraan, V.N. Rayudu, Harry Caussy

Department of Environmental Health Engineering (K.B., S.G., S.S.) and Department of Chest Medicine (V.T.), Sri Ramachandra University, Porur, Chennai; University of Calcutta, Calcutta (B.G.); Advanced Environmental Laboratory Tamil Nadu Pollution Control Board, Chennai (V.N.R.); and World Health Organization South-East Asia Regional Office, New Delhi (H.C.) — India

ABSTRACT

This report describes the results of a time-series analysis of the effect of short-term exposure to particulate matter with an aerodynamic diameter ≤ 10 µm (PM_{10}*) on mortality in metropolitan Chennai, India (formerly Madras). This was one of three sites in India chosen by HEI as part of its Public Health and Air Pollution in Asia (PAPA) initiative. The study involved integration and analysis of retrospective data for the years 2002 through 2004. The data were obtained from relevant government agencies in charge of routine data collection.

Data on meteorologic confounders (including temperature, relative humidity, and dew point) were available on all days of the study period. Data on mortality were also available on all days, but information on cause-of-death (including accidental deaths) could not be reliably ascertained. Hence, only all-cause daily mortality was used as the major outcome for the time-series analyses. Data on PM_{10}, nitrogen dioxide (NO_2), and sulfur dioxide (SO_2) were limited to a much smaller number of days, but spanned the full study period. Data limitations resulting from low sensitivity of gaseous pollutant measurements led to using only PM_{10} in the main analysis. Of the eight operational ambient air quality monitor (AQM) stations in the city, seven met the selection criteria set forth in the common protocol developed for the three PAPA studies in India. In addition, all raw data used in the analysis were subjected to additional quality assurance (QA) and quality control (QC) criteria to ensure the validity of the measurements.

Two salient features of the PM_{10} data set in Chennai were a high percentage of missing readings and a low correlation among daily data recorded by the AQMs. The latter resulted partly because each AQM had a small footprint (approximate area over which the air pollutant measurements recorded in the AQM are considered valid), and partly because of differences in source profiles among the 10 zones within the city. The zones were defined by the Chennai Corporation based on population density. Alternative exposure series were developed to control for these data features. We first developed exposure series based on data from single AQMs and multiple AQMs. Because neither was found to satisfactorily represent population exposures, we subsequently developed an exposure series that disaggregated pollutant data to individual zones within the city boundary. The *zonal series*, despite some uncertainties, was found to best represent population exposures among other available choices. The core model was thus a *zonal model* developed using disaggregated mortality and pollutant data from individual zones.

We used quasi-Poisson generalized additive models (GAMs) with smooth functions of time, temperature, and relative humidity modeled using penalized splines. The degrees of freedom (*df*) for these confounders were selected to maximize the precision with which the relative risk for

This Investigators' Report is one part of Health Effects Institute Research Report 157, which also includes a Critique by the Health Review Committee and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr. Kalpana Balakrishnan, Department of Environmental Health Engineering, Sri Ramachandra University, Porur, Chennai 600116, India; e-mail: *kalpanasrmc@vsnl.com*.

The PAPA Program was initiated by the Health Effects Institute in part to support the Clean Air Initiative for Asian Cities (CAI-Asia), a partnership of the Asian Development Bank and the World Bank to inform regional decisions about improving air quality in Asia. Additional funding was obtained from the William and Flora Hewlett Foundation. The contents of this document 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.

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

PM_{10} was estimated. This is a deviation from the traditional approaches to degrees of freedom selection, which usually aim to optimize overall model fit. Our approach led to the use of 8 *df*/year for time, 6 *df*/year for temperature, and 5 *df*/year for relative humidity.

The core model estimated a 0.44% (95% confidence interval [CI] = 0.17 to 0.71) increase in daily all-cause mortality per 10-µg/m^3 increase in daily average PM_{10} concentrations. Extensive sensitivity analyses compared models constructed using alternative exposure series and contributions of model parameters to the core model with regard to confounder degrees of freedom, alternative lags for exposure and meteorologic confounders, inclusion of outliers, seasonality, inclusion of multiple pollutants, and stratification by sex and age. The sensitivity analyses showed that our estimates were robust to a range of specifications and were also comparable to estimates reported in previous time-series studies: PAPA, the National Morbidity, Mortality, and Air Pollution Study (NMMAPS), Air Pollution and Health: A European Approach (APHEA), and Air Pollution and Health: A European and North American Approach (APHENA).

While the approaches developed in previous studies served as the basis for our model development, the present study has new refinements that have allowed us to address specific data limitations (such as missing measurements and small footprints of air pollution monitors). The methods developed in the study may allow better use of routine data for time-series analysis in a broad range of settings where similar exposure and data-related issues prevail. We hope that the estimates derived in this study, although somewhat tentative, will facilitate local environmental management initiatives and spur future studies.

INTRODUCTION

BACKGROUND

Systematic investigations conducted over several decades provide convincing evidence that outdoor air pollution affects human health in many ways. The World Health Organization (WHO) and the American Thoracic Society have identified a broad range of adverse health outcomes that are specifically related to air pollution exposure (WHO 1987; American Thoracic Society 2000). Moreover, recent toxicologic studies provide evidence of plausible biologic mechanisms in support of such an association. Advances in many disciplines — including clinical physiology, population and molecular epidemiology, toxicology, environmental genomics, and biostatistics — have made it possible to detect fairly subtle health effects with a high degree of resolution. Many recent reports such as those of HEI (2001, 2002), the U.S. Environmental Protection Agency (2004), and WHO (2005) support these observations.

Scientific advances in the field, however, have not been concomitantly accompanied by efforts to reduce the burden of disease attributable to outdoor air pollution in many regions of the world. Outdoor air pollution is estimated to account for approximately 800,000 deaths and 4.6 million lost life-years worldwide (WHO 2002), with the burden being unequally distributed across countries (e.g., nearly two-thirds of the deaths occur in developing countries of Asia). Moreover, the bulk of scientific literature providing evidence for the health impacts of air pollution comes from studies conducted in developed country settings within North America and Western Europe. The differences between developed and developing countries in exposure concentrations, the nature of pollutants, baseline health, and determinants of susceptibility add uncertainties while extrapolating exposure–response relationships across countries. This is further compounded by the limited technical and administrative resources available for air quality management in developing countries. Thus, risk perception and risk communication mechanisms are handicapped by the lack of locally-derived relationships and methodological understanding. This limits transferability across settings and results in a gross lack of initiative for environmental management in the communities that most need them.

HEI initiated the PAPA program, which is based on a needs assessment, to generate scientific evidence within Asia and build local capacity to implement and sustain air quality management initiatives. This report summarizes the results of the PAPA study conducted in Chennai, India. The primary objective of the study was to develop models to estimate the short-term health effects of selected criteria air pollutants using a retrospective time-series design. While previous studies conducted in the United States and Europe (NMMAPS [Samet et al. 2000 a,b], APHEA [Katsouyanni et al. 1996, 2001], and APHENA [Samoli et al. 2008; Katsouyanni and Samet et al. 2009]) and the first wave of PAPA studies (HEI Public Health and Air Pollution in Asia Program 2010) served as the basis for developing these models, several alternative approaches have been developed to address the specific data challenges encountered by the authors in this setting.

RATIONALE FOR TIME-SERIES STUDIES IN INDIA

The HEI International Scientific Oversight Committee (ISOC) study (2004) provides a comprehensive review and meta-analysis of original epidemiologic studies of the health effects of outdoor air pollution conducted in Asia; it is based on 138 papers published between 1980 and 2003. Most of the studies were conducted in the East Asian countries of China, Hong Kong, South Korea, and Japan with a few in India and Southeast Asia. They collectively examined the association of air pollutants with mortality, hospital admissions, respiratory symptoms, pulmonary function, and adverse reproductive outcomes. Most (70%) were either cross-sectional prevalence studies of chronic respiratory symptoms and pulmonary function or were time-series studies of the effects of short-term exposure on daily mortality or hospital admissions. Although the estimated effects of the impact on mortality and morbidity from the Asian studies are similar to those reported by early studies in North America and Europe (Pope and Dockery 1999), they do not reflect the wide diversity of Asian cities. The review emphasizes the need to initiate studies that use a common protocol, are spread across a representative spectrum of exposure settings, and include countries where few studies have been carried out.

India, like other developing countries, can be seen from a time-scale perspective as being somewhere in the middle of the environmental health risk transition. Risks posed by traditional environmental health hazards such as unsafe food and drinking water, inadequate sanitation, infections from animals and vectors, and poor housing compete with risks posed by modern environmental health hazards such as air pollution, chemical exposures, and traffic accidents. In most cities, deteriorating outdoor air quality often has to compete with a range of other risk factors for resource allocation and risk mitigation. Although the available national statistics on environmental quality in India clearly identify the need for addressing outdoor air pollution within environmental management plans, few studies have performed direct assessments of health impacts from local environmental degradation. The critical missing piece is usually the lack of good quality exposure data. Even when available, the data are not easily linked with health data sets on the same population within the same time frame.

Routine collection of environmental data has a relatively short history in India. The State and Central Pollution Control Boards have been vested with a large share of this responsibility along with other national agencies such as the National Environmental Engineering Research Institute. The collection of these data is largely driven by the need to document compliance with India's National Ambient Air Quality Standards (NAAQS; CPCB 2009), without explicit requirements for modeling health impacts on the exposed population. Few detailed exposure assessment, dispersion modeling, or biologic monitoring exercises have been done. Also, environmental epidemiologic studies have mostly focused on gathering cross-sectional prevalence information on environmental health endpoints. The HEI ISOC study (2004) presents a more detailed literature review of all major Indian studies.

One of the earliest assessments of the health impacts of air pollution in India was conducted by the World Bank (1995). In that assessment the economic costs of environmental degradation for several cities in India were estimated. Economic costs associated with health impacts and productivity impacts were calculated for several priority areas that were identified under the National Environmental Action Program of the Indian government. The study used exposure–response information from other countries. Although the study estimated that the economic burden from air pollution may amount to a staggering 4%–6% of the national gross domestic product, it had little impact on local policy. The lack of locally-derived exposure–response functions was pointed out as a major impediment for accurate assessment of health burdens that were attributable to air pollution. A time-series study conducted in Delhi more than a decade ago related total suspended particulate matter (TSP) concentrations to mortality (Cropper et al. 1997). However, much has changed with respect to both air quality and mortality recordings in most Indian cities (e.g., TSP is no longer considered a good marker for health effects and is no longer routinely monitored). This provided the most important justification for conducting a time-series study of the health effects of air pollution in India (as such studies have often served as catalysts for further detailed studies) and led to the initiation of a coordinated set of time-series studies in Chennai, Delhi, and Ludhiana, under the HEI PAPA initiative. A first step in the conduct of these three studies was to develop a common protocol (see Appendix B; available on the Web).

RATIONALE FOR STUDY IN CHENNAI

Chennai is the fourth largest metropolis in India and the capital of the southern coastal state of Tamil Nadu. A brief listing of relevant demographic variables and sources related to air pollution is provided in Table 1. The topography of the city is almost flat with mean ground level rising gradually to about 7 meters above mean sea level. Temperatures range from around 21°C (between December and February) to around 37°C (between March and September). The city's population of nearly 4.3 million people is spread across an area of approximately 174 square kilometers, with an additional 2.6 million people living just outside the city in the Greater Chennai metropolitan area (Census of India 2001).

Urban development in Chennai has been rapid over the last two decades. Much of this has been the result of initiatives supported by the World Bank and the International Development Association under the Madras Urban

Table 1. Demographic Data for Chennai City[a]

	Total	Male	Female
Age group (years)			
0–4	299,356	151,721	147,635
5–44	2,288,225	1,166,698	1,121,527
45–64	686,927	359,005	327,918
≥ 65	225,856	111,185	114,671
Total population	4,343,645	2,219,539	2,124,106
Literacy rate (%)	85.53	90.01	80.44
Population density	24,231/km^2		

[a] Source: Census of India 2001.

Development Projects and the Tamil Nadu Urban Development Projects. Chennai was considered to be one of the cleanest metropolitan cities in terms of air quality in the early 1990s, but the situation has deteriorated significantly since then. Despite a relatively high per capita annual income of about $500 U.S. dollars (USD), as compared with the national average of $330 USD, and a literacy rate of over 85%, the air quality in the city is well below international air quality standards. Several factors may have contributed to this: 1) increasing population density (24,321/km^2) with nearly a third of the population living in slums close to air pollution hot spots; 2) increasing vehicular traffic (1.8 million vehicles) and congestion; 3) a vehicle fleet with high-polluting diesel and gasoline engines with nearly a third of the gasoline engines being two-stroke; and 4) more than 939 hazardous industries (e.g., petrochemical, automotive, and power plants) are located within the Greater Chennai Metropolitan area.

The earliest efforts to address environmental health risks within a sustainable development framework in Chennai were initiated by the Sustainable Chennai project of the United Nations Development Program. The project identified several strategies for urban growth and development that would be compatible with environmental preservation. An exhaustive environmental profile was prepared and several areas of environmental concern within the city were identified, including identifying air pollution as a priority area for environmental management programs. More recently, the same investigators have conducted comparative risk assessments of environmental concerns at selected municipalities within Chennai. These studies have shown that the health and economic risks of PM_{10} exposure may rank higher than risks of exposure to microbial or chemical contaminants in water or to solid waste (Balakrishnan et al. 2003; Balakrishnan and Parikh 2005), and that people residing in industrial zones face nearly double the risk of respiratory symptoms or illness compared with those living in residential zones. These studies have also shown that citywide monthly or yearly average air pollutant concentrations reported in national statistics are poor surrogates of exposure concentrations for the city's population. Several parts of the city, including the northernmost industrial corridors, experience air pollutant concentrations that are much higher than the city average, and temporal variations are poorly reflected in those summary statistics. Although the city has a network of eight routine AQM stations (i.e., AQM stations for routine data collection) that measure all criteria pollutants (including PM_{10}, SO_2, NO_2, and lead), the daily averages are either not consistently available or easily accessible. Finally, the status of health records in the public health departments and private hospitals does not readily lend itself to establishing linkages with the available environmental data. Health impact assessments for air pollutants in Chennai have therefore been limited to a few cross-sectional studies that have established the baseline prevalence of a few health outcomes related to air pollution (including respiratory symptoms and prevalence of upper and lower respiratory tract illnesses). Until now, data limitations precluded the design and implementation of a time-series study.

However, Chennai presented a unique opportunity to initiate a time-series study built on the groundwork laid out by previous projects related to air pollution. Several government representatives expressed the need for such a study during numerous formal and informal meetings related to previous projects. The Tamil Nadu Pollution Control Board (TNPCB) was already involved in the implementation of routine environmental monitoring schemes under the National Environmental Protection Act. Sri Ramachandra University, the primary institution conducting this study, has had a long-standing collaborative association with state agencies such as the TNPCB, the Chennai Metropolitan Development Authority (CMDA), and the Directorate of Public Health. In addition to granting permission to conduct this study, these agencies further extended their cooperation by sharing the raw data along with available quality control information. The team at Sri Ramachandra University was also supported by an ongoing collaborative arrangement with the Department of Statistics at the University of Calcutta. Chennai represents one of nine cities selected by the Ministry of Environment and Forests, Government of India, to create a national environmental health profile; these studies are currently being conducted by the same set of partners.

Thus, major factors in choosing Chennai city as the location for this study include pressing environmental conditions, a good infrastructure for collection of secondary health and environmental data, adequate technical capacities to conduct the proposed project, a favorable administrative climate for data collection and implementation of potential study recommendations, and the availability of a committed and eager multidisciplinary core group of investigators. The investigators trained through this project will be able to transfer their skills and knowledge to others to improve future studies.

STUDY OBJECTIVES

Specific objectives of the study are as follows:

1. describe the distribution of daily averages of criteria air pollutant concentrations across all ambient AQM stations located within the greater Chennai area over a 3-year period from 2002 to 2004;
2. examine the patterns and quality of data from individual monitors or combinations of monitors to establish criteria for inclusion or exclusion in alternative exposure series;
3. create a single best exposure series that maximizes the use of available data and adequately represents population exposure;
4. retrieve and organize data on all-cause mortality over the same period;
5. develop a common protocol (see Appendix B) for a time-series analysis using the exposure and mortality data sets for the three Indian PAPA study sites and identify special requirements for the Chennai data set;
6. develop statistical models to estimate short-term impacts of criteria air pollutants on all-cause mortality in Chennai and conduct sensitivity analyses according to criteria laid out in the common protocol; and
7. contribute to the pool of Indian and Asian studies under the PAPA program for coordinated analysis and meta-analysis.

METHODS

DATA COLLECTION FRAMEWORK

The Chennai metropolitan area is governed by two administrative divisions, the Chennai Corporation and the CMDA. Figure 1 shows the boundaries of areas under the jurisdiction of the CMDA and the Chennai Corporation, and locations of AQM stations, meteorologic stations, major roads, and industrial plants. The Chennai Corporation boundary encloses 10 municipal zones. Approximately 98% of the births and deaths occurring within these 10 zones are recorded by the Chennai Corporation office (according to communications provided by the Vital Statistics Department functioning under the Directorate of Public Health). The CMDA boundary defines the Greater Chennai Metropolitan area.

Most of the major industrial point sources that are likely to contribute to criteria air pollutant exposures for the greater Chennai area population are located outside the Chennai Corporation boundary but within the CMDA boundary. Several ambient AQMs and one of the two meteorologic stations are also located outside the Chennai Corporation boundary but within the CMDA boundary. Most major hospitals are located within the Chennai Corporation boundary.

Ambient air quality and meteorologic data were collected from monitors within the CMDA boundary (as this provided the largest coverage of available data for air pollution exposures and meteorologic covariates); mortality data were provided by the Chennai Corporation (as this provided the single largest death registry for the residents of Chennai). Ideally it would have been useful to collect mortality data from the entire greater Chennai area (i.e., including the CMDA area); however, because deaths outside the Chennai Corporation boundary were recorded by five different municipalities, logistic challenges precluded collection of those data. Figure 2 shows the locations of AQM stations relative to the average population density across the city zones.

This study used retrospective data collected between January 2002 and December 2004 from three principal sources. Air quality data were obtained from the TNPCB, which operates ambient AQM network stations at eight locations within the CMDA boundary. Data on daily all-cause mortality were collected from the Chennai Corporation office, which records all births and deaths within the Chennai Corporation boundary at 10 zonal offices linked to a central registry. Meteorologic data were collected from the India Meteorological Office, which operates two stations within the CMDA boundary.

AIR POLLUTION DATA

The TNPCB collects daily air quality data on weekdays from eight routine AQM stations in the greater Chennai area and operates a continuous AQM station colocated with one of these stations. Air quality monitoring is conducted across all monitoring stations in the city at a preset periodicity that spans the entire year. Air quality is monitored either on designated days or on every third

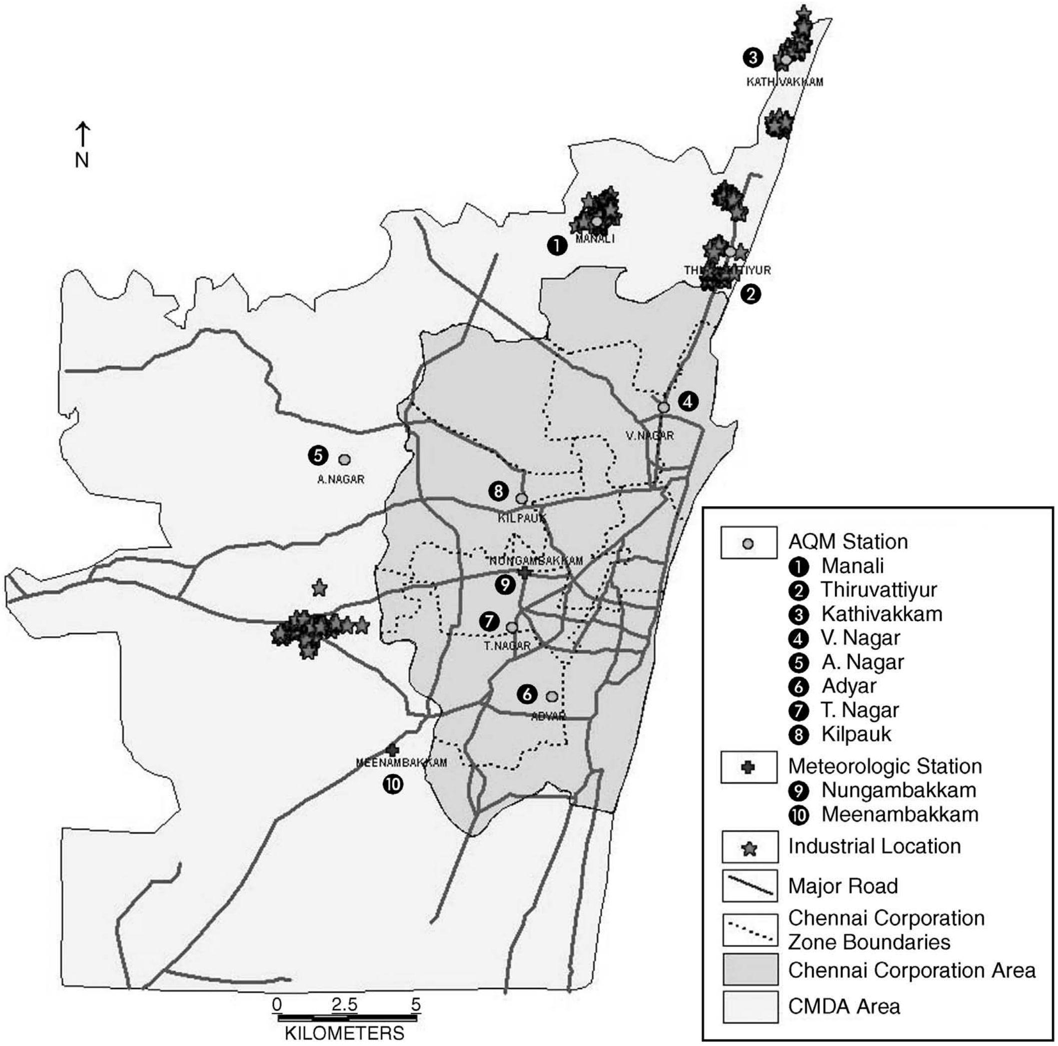


Figure 1. Map of Chennai city showing the CMDA area, the Chennai Corporation area with its zone boundaries, and the locations of AQMs and meteorologic stations. The continuous AQM is colocated with AQM 3. Source: Ms. Tata Consultancy Services for the Department of Environmental Health Engineering, Sri Ramachandra University, Chennai, India.

day, resulting in 100 to 120 observations per year per station. It is not monitored on weekends and public holidays. The data from each station are electronically compiled by the TNPCB.

After obtaining the data, we considered analyzing 24-hour averages of PM_{10}, SO_2, and NO_2. However, the utility of the SO_2 and NO_2 data was limited because low-resolution wet-chemical methods were used to measure concentrations of these gaseous pollutants. Therefore, only the PM_{10} data were used to guide exposure series construction. The core model was thus a single-pollutant model centered on PM_{10}; inclusion of the gaseous pollutant data was restricted to subsequent sensitivity analyses.

AQMS

AQM Classification

AQMs operated by TNPCB are classified as industrial (AQMs 1, 2, and 3), commercial (AQMs 4, 7, and 8) or residential (AQMs 5 and 6) (see Table 2). These classifications are based on land-use criteria set by the Central Pollution Control Board (CPCB), which is the nodal government agency in charge of the National Ambient Air Quality Monitoring Programme in India. Accordingly, the criteria for a residential AQM are: (1) it is located in an area where the population density is greater than 4000 people per square kilometer; (2) no industrial activities are within

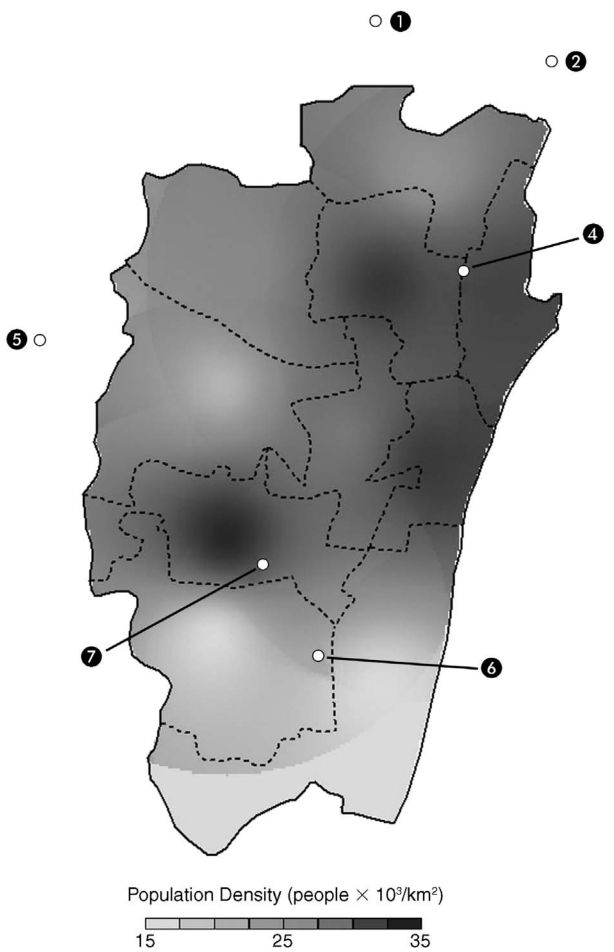


Figure 2. Map of Chennai Corporation zones showing AQM locations relative to population density. Source: Ms. Tata Consultancy Services for the Department of Environmental Health Engineering, Sri Ramachandra University, Chennai, India.

two kilometers of the AQM; and (3) the AQM is located more than 100 meters away from streets with a vehicular density greater than 500 vehicles per day. Commercial AQMs are located in areas near roads (~ 10 m away) with at least 10,000 vehicles per day (as vehicular density increases, the required distance from the road also increases). Industrial AQMs are located in areas within a cluster of air-polluting industries. In addition, city zone classifications (residential, commercial, or industrial) are based on land-use information. The classification of AQMs is not only characterized by varying proximities to sources as described above, but also reflects the general land-use pattern of the zone in which it is sited. For example, AQM 4 is designated a commercial AQM both because of its placement close to highly trafficked roads, but also because it is in a commercial zone characterized by high traffic densities. The spatial scale of the AQMs is estimated by TNPCB to be at the lower end of the range of an urban scale AQM (which ranges from 4 to 100 km). AQMs are usually positioned on the roofs of buildings with two to four floors. Tables 2 and 3 give details of the monitoring stations and the methods as provided by the TNPCB. Aerial views of the AQM stations are in Appendix G (available on the Web).

AQM Selection Criteria

AQM stations included for analyses were first selected on the basis of preset site selection criteria and data quality criteria as outlined in the common protocol (see Appendix B). Accordingly, sites were included if: (1) they were located within the greater metropolitan city boundary as defined by the Chennai Metropolitan Authority (see Figure 1 for general location and Figure 2 for location relative to population density); (2) the sites were large enough to ensure the availability of space for monitoring, were located in flat space and were elevated at least 10 feet above the ground; (3) were not curbside AQMs; and (4) were located at least five meters upwind from building exhausts, vents, or chimneys and at least two meters from walls. Thus, AQM 8 was disqualified for the study because it was a curbside AQM.

Data Quality Criteria

TNPCB provided the complete set of QA and QC information available in their records for cross-verification by investigators. After the AQMs that met site-selection criteria were selected, the raw data on QA and QC parameters were manually abstracted from the field data cards that had been maintained by the individual AQM stations. The electronic data set was compared with the field data cards for selection of valid data points as per protocols recommended by the CPCB (described in Appendices C, D, and E; available on the Web). The averaging period for individual data points in a day ranged from 4 to 8 hours in the TNPCB data set.

Guidelines set forth in the common protocol were used to select valid days. Accordingly, for PM_{10}, days were included in the analysis only if all three 8-hour shifts had been monitored. For NO_2 and SO_2, days were included if at least 75% of the total monitoring duration had been covered. Application of QA and QC criteria resulted in the exclusion of several periods (of two and sometimes up to four months) from AQMs 6 and 7 (reported by TNPCB as being due to frequent shifts in the AQM locations). The remaining five AQMs (i.e., AQM 1–5) had valid data that were available for the full three years. The core (zonal) model developed in the study, by disaggregating exposures on the basis of nearest available AQMs, utilized valid data from all AQMs for all available periods within the three years.

Table 2. Background Information for AQMs

AQM	Global Positioning System Coordinates	Sampling Height (feet)	Site Classification	Monitoring Schedule Days of Week, Number of Days in Year 2002/2003/2004
1 Manali	13°09.951N 80°15.502E	17	Industrial (suburban)	Mon/Thu, 99/99/98
2 Thiruvattiyur	13°09.374N 80°18.122E	19	Industrial (suburban)	Wed/Fri, 96/94/102
3 Kathivakkam	13°13.083N 80°19.218E	12	Industrial (suburban)	Tue/Thu, 99/97/102
4 V. Nagar	13°06.345N 80°16.809E	10	Commercial (city)	Weekdays, 177/93/95
5 A. Nagar	13°05.319N 80°10.515E	35	Residential (city)	Weekdays, 173/104/101
6 Adyar	13°00.744N 80°14.665E	25	Residential (city)	Weekdays, 176/98/97
7 T. Nagar	13°02.074N 80°13.814E	25	Commercial (city)	Weekdays, 180/91/97

Table 3. Air Quality Monitoring Techniques[a]

	Pollutants		
	PM_{10}	SO_2	NO_2
Equipment	RDS	Gaseous sampling unit attached to HVS/RDS	Gaseous sampling unit attached to HVS/RDS
Flow measuring device	Pressure drop across orifice in the hopper	Rotameters/Orifice Nanometers	Rotameters/Orifice Nanometers
Flow rate	0.8–1.3 m^3/min	1 L/min	1 L/min
Sampling period (over 24 hours)	8 hours × 3	4 hours × 6	4 hours × 6
Sampling frequency	Twice per week	Twice per week	Twice per week
Analytical methods[b]	Gravimetric	West & Gaeke	Jacobs & Hochheiser
Minimum detection limit	1 µg/m^3	0.04 µg/m^3	0.03 µg/m^3
Minimum reporting value	10 µg/m^3	6 µg/m^3	3 µg/m^3
Maximum absorption wave length	NA	560 nm	550 nm

[a] RDS indicates respirable dust sampler; HVS indicates high-volume sampler; NA indicates not applicable.
[b] Described in Appendices C, D, and E (available on the Web).

Finally, data were also obtained from a continuous AQM station that was colocated with AQM 3 and operated by TNPCB. Parameters recorded at this station included readings every half hour with 8-hour and 24-hour averages for each of the three pollutants. The continuous AQM station used a β gauge for PM_{10} measurements and chemiluminescence methods for gases, while all the other AQM stations used an indigenous cyclone-based gravimetric method for PM_{10} and wet chemical methods for gases. Data from the continuous AQM station were used to serve as a broad indicator of comparability of methods, because the availability of data from the continuous AQM station was limited to less than a year within the study period. Figure 3 is a schematic diagram showing how the Chennai air pollution database was constructed.

Table 4 lists descriptive statistics for the air pollutant data set for AQMs 1 through 5. Figure 4 provides time-series plots of daily average PM_{10} concentrations at those AQM sites with an overlay of a smooth function of time over the full study period.

24-hour averages for PM_{10} consistently exceeded the current NAAQS value (CPCB 2009) of 100 μg/m^3 at AQM stations located in industrial zones (AQMs 1–3), while

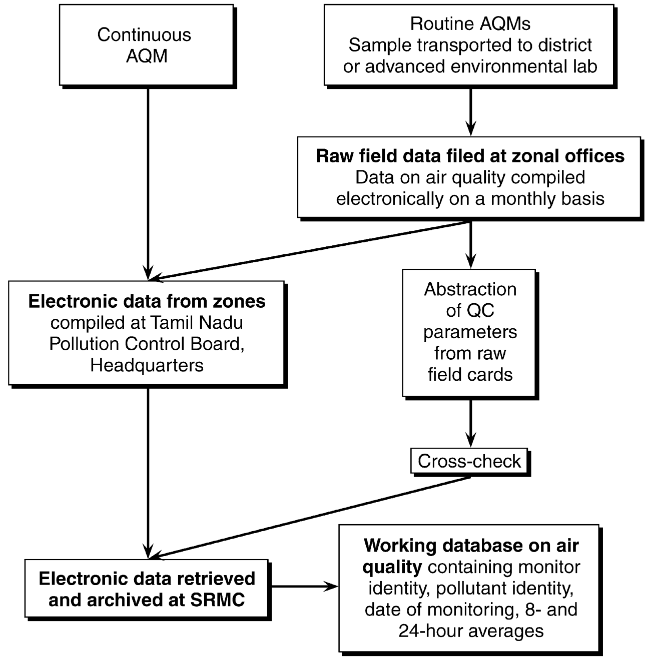


Figure 3. Schematic of air pollution data collection steps. SRMC indicates Sri Ramachandra Medical College.

Table 4. Summary Statistics for Pollutants[a]

Pollutant/ Site	*n*	Mean (μg/m^3)	Median (μg/m^3)	SD	QD	Minimum	Maximum
PM_{10}							
AQM 1	280	108.0	107.55	27.48	18.85	39.6	218
AQM 2	288	95.9	94.3	26.02	17.55	35.9	183
AQM 3	293	90.7	89.9	25.5	18.0	35.0	195.8
AQM 4	360	110.7	101.0	47.23	23.40	32	417.0
AQM 5	376	71.73	67.0	33.42	16.37	15	317.0
NO_2							
AQM 1	276	31.2	28.3	14.35	8.51	6.4	91.9
AQM 2	286	28.9	26.2	12.35	8.73	5.4	69.9
AQM 3	291	30.5	23.8	11.94	7.93	9.6	76.1
AQM 4	361	27.75	24.4	13.66	8.51	9.0	93.0
AQM 5	378	15.08	13.5	6.75	3.61	6.3	69.0
SO_2							
AQM 1	276	24.4	22.3	11.53	6.76	5.1	77.9
AQM 2	286	23.6	21.6	11.48	6.22	4.3	66.1
AQM 3	291	28.4	23.5	16.11	9.55	3.8	102.3
AQM 4	361	5.86	5.0	2.49	1.60	4.0	18.0
AQM 5	378	4.76	4.0	1.94	0.40	3.6	25.6

[a] All pollutant concentrations are 24-hour averages. *n* indicates the number of days that each monitor had available data over the study period; SD indicates standard deviation; QD indicates quartile deviation.

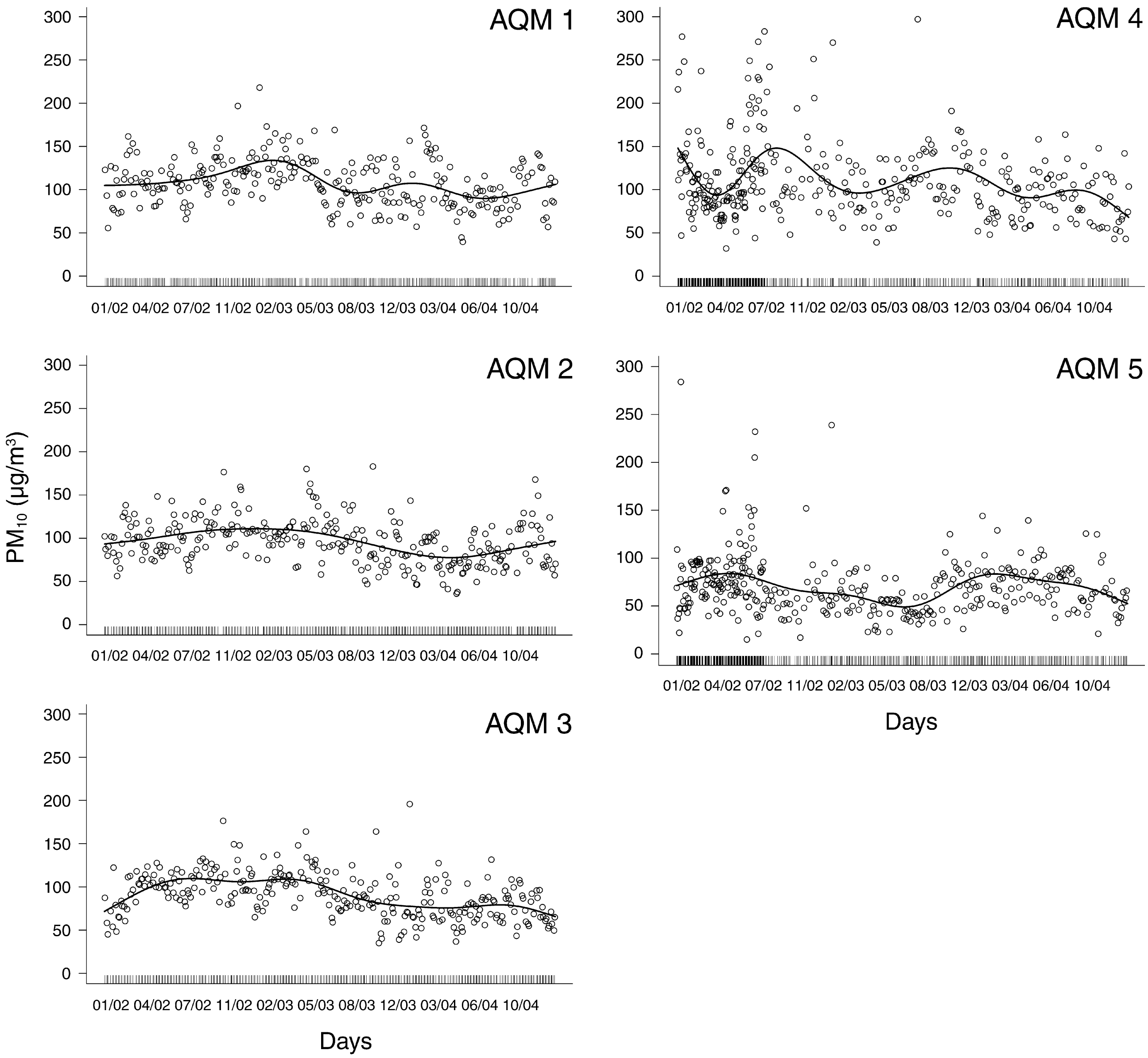


Figure 4. Time-series plots of daily PM_{10} concentrations at AQMs 1–5.

AQM stations in residential and commercial zones (AQMs 4 and 5) generally recorded 24-hour averages near that value. 24-hour averages for SO_2 frequently exceeded the current NAAQS value of 80 µg/m^3 at industrial stations and were below that value at commercial and residential zone AQMs. Nearly 60% of the values were clustered at the lower detection limit of 4 µg/m^3. 24-hour averages for NO_2 frequently exceeded the current NAAQS value of 80 µg/m^3 at industrial stations; residential and commercial stations on average remained below guideline values of 80 µg/m^3, but the daily averages varied greatly. Nearly 30% of the values were clustered at the lower detection limit value of 4 µg/m^3.

MORTALITY DATA

We had access to two data sets on mortality. The primary data set included all reported deaths provided by the Chennai Corporation. A modified data set was provided by the National Cancer Institute in Chennai (NCI-Chennai). The Chennai Corporation is responsible for compiling and maintaining mortality data and vital statistics for the city of Chennai. Birth and death registrars in the individual zonal offices under the Chennai Corporation manually collect the mortality data. This information is compiled electronically at the zonal offices. A computerized wide area network reporting system links information between the zonal offices and the electronic data processing cell at the Chennai Corporation office. We used the electronic database furnished by the Chennai Corporation as the primary database. The NCI-Chennai database had death records of all noninfant deaths for people who had lived within the Chennai Corporation limits. NCI-Chennai collected these data independently from the Chennai Corporation records to create a cancer registry within the city. We linked the two electronic databases provided by the Chennai Corporation and NCI-Chennai to track discrepancies in a few data entry fields. Subsequently, discrepancies were resolved by verification against individual records available at the individual zonal offices. Only the following fields were identified as useful after cross-verification: date-of-death, place-of-death, place-of-residence, age, and sex. Cause-of-death information was retrieved, but the data were not used because the reliability was considered to be low. For example, fewer than 160 deaths out of more than 100,000 deaths recorded during the entire study period were reported as being due to accidental causes, and only 60% had any information on the cause of death. The mortality data used for analyses thus refer to all-cause mortality in all results generated by the study.

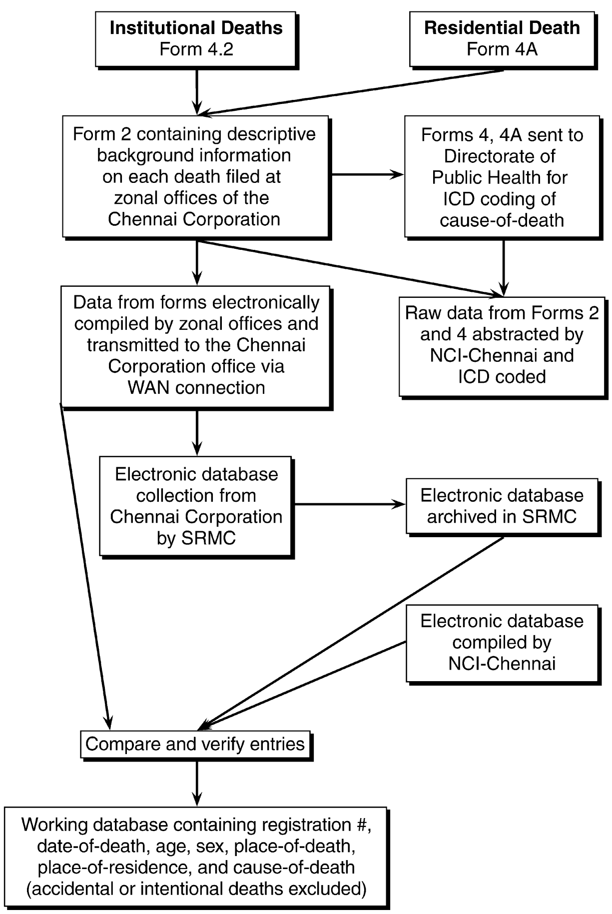


Figure 5. Schematic of mortality data collection steps. ICD indicates International Classification of Diseases; WAN indicates wide area network; SRMC indicates Sri Ramachandra Medical College.

Deaths were registered according to the place at which the death occurred. Deaths of those residing within but dying outside the city limits were therefore not included in the city records. Conversely, deaths of people who lived outside the Chennai Corporation limits but died at a major hospital within the city were included. Therefore, all deaths occurring within the Chennai Corporation boundaries (and recorded at the Chennai Corporation office) were included for analysis.

Daily all-cause mortality ranged from 60 to 229 deaths in a year, with male mortality being consistently higher than female mortality in most age groups. People 65 years or older accounted for the greatest proportion of all deaths. Figure 5 shows the data processing steps for mortality data, and Table 5 shows the summary statistics of the mortality data set; Figure 6 is a time-series plot of daily mortality, and Figure 7 shows time-series plots stratified by sex and age.

Table 5. Summary Statistics for Mortality[a]

	Mean	Median	SD	QD	Min	Max
Female	39.126	39	8.48	5.5	16	96
Male	57.762	57	11.15	7.0	25	133
Age (years)						
0–4	4.895	4	3.31	1.5	0	21
5–44	21.067	20	6.28	3.5	5	109
45–64	26.308	26	6.45	4.0	10	64
≥ 65	43.763	43	9.93	6.0	19	90
Total	96.888	96	16.55	10.0	61	229

[a] Data are summarized for the 1096 days that had available data over the study period. SD indicates standard deviation; QD indicates quartile deviation; Min indicates minimum number; Max indicates maximum number.

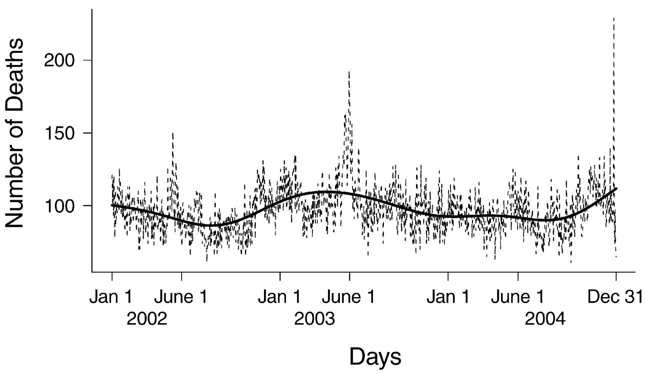


Figure 6. Time-series plot of daily total all-cause mortality in Chennai city.

METEOROLOGIC DATA

Information on temperature, relative humidity, dew point, wind speed, and wind direction was collected from the Regional Meteorological Centre in Chennai. Two meteorologic monitoring stations, one at the Regional Meteorological Centre in Nungambakkam (monitor 9) and the other at the airport in Meenambakkam (monitor 10), recorded two daily measurements for each of the meteorologic parameters. We used the average of the daily morning readings and of the daily evening readings from the two stations. Data on barometric pressure were available but were not abstracted from the records, so they could not be used in subsequent analyses. Field data were entered on data cards

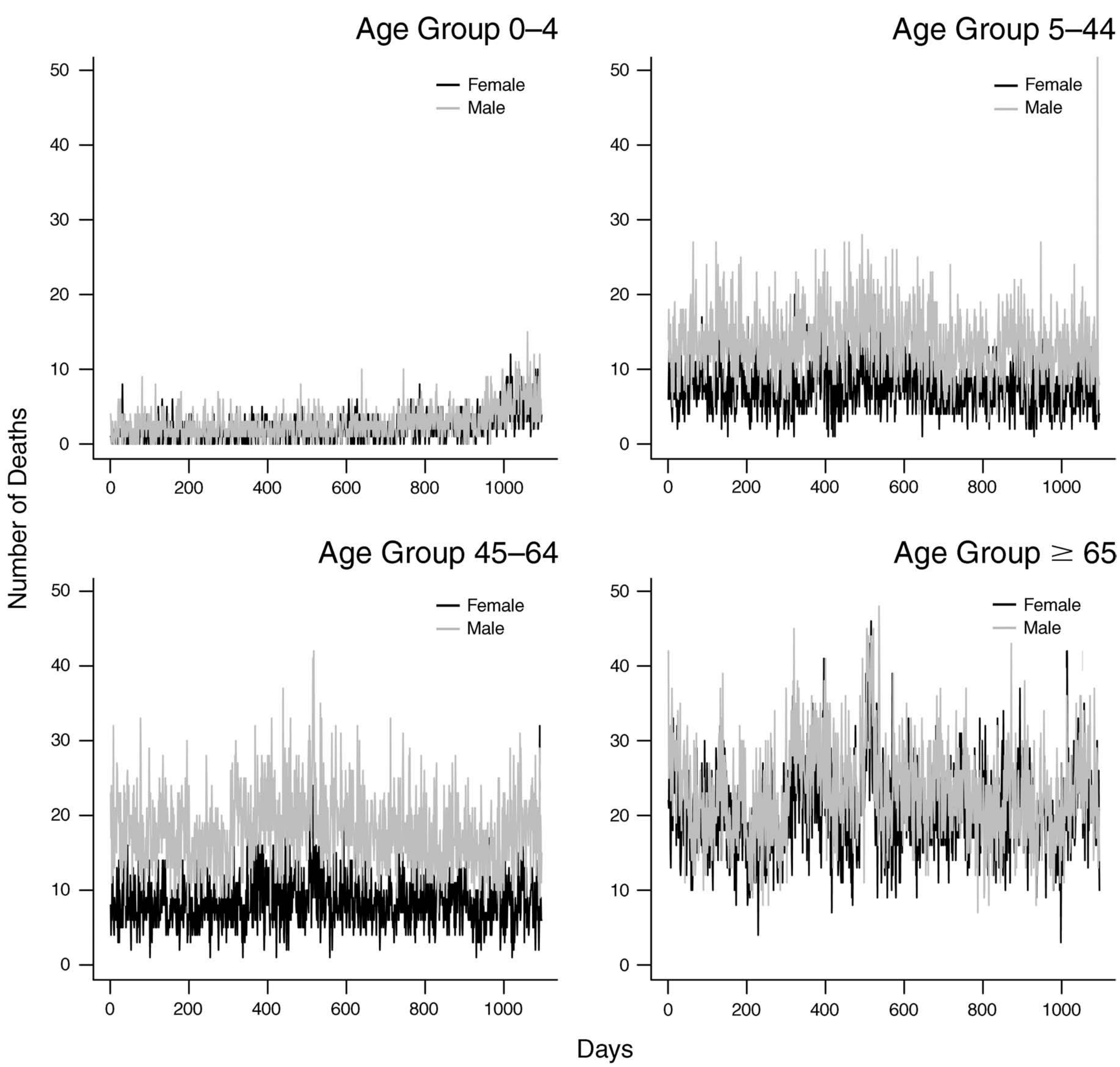


Figure 7. Time-series plots of daily total all-cause mortality in Chennai city stratified by sex and age.

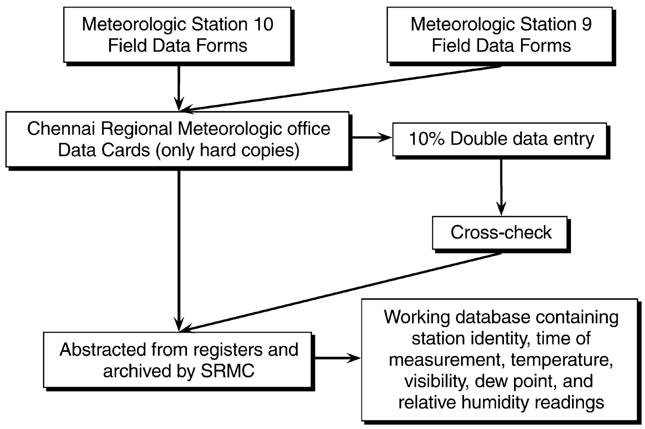


Figure 8. Schematic of meteorologic data collection steps. SRMC indicates Sri Ramachandra Medical College.

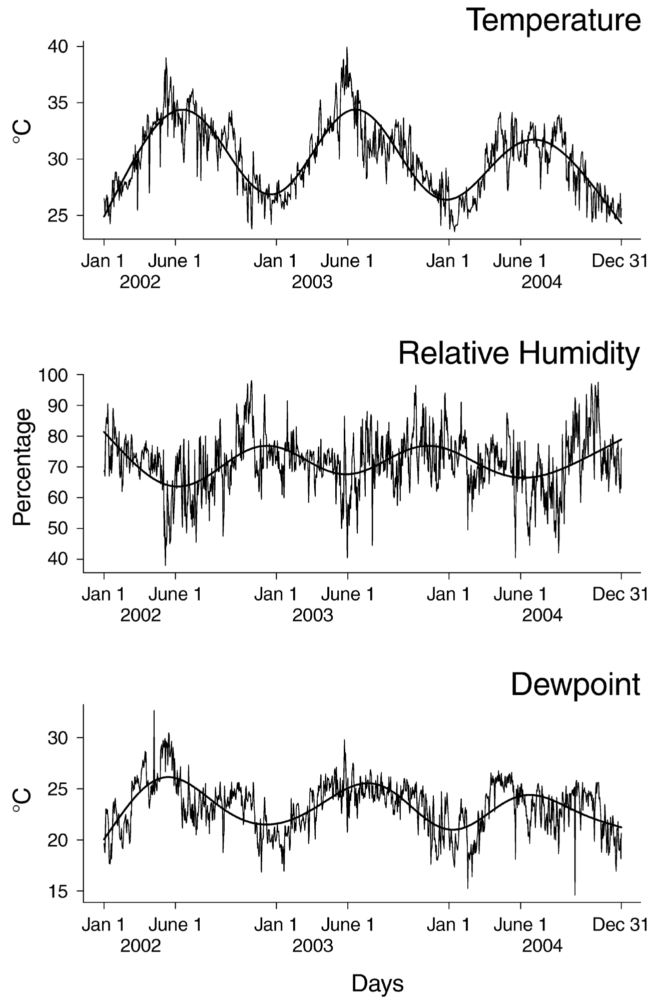


Figure 9. Time-series plots of meteorologic parameters.

and records were maintained as handwritten registers at the Chennai Regional Meteorological office. This information was abstracted manually into an electronic database to serve as the primary meteorologic data set. Ten percent of the data points were entered in duplicate to check for discrepancies. Data processing steps used for meteorologic data are shown in Figure 8.

Table 6 describes the meteorologic data set and Figure 9 provides the time-series plots for key meteorologic covariates over the study period. Annual temperature ranged from 23–48°C, relative humidity from 38%–98% and dew point from 15.0–34.4°C. Wind direction and wind speed data were also retrieved to determine their contributions to spatial gradients in air pollution concentrations. Data on actual visibility were not available because the meteorologic records only provided categorized visibility (with ranges of 1340–2300 m, 2300–4000 m, 4000–7500 m, 7500–12,000 m, and > 12,000 m). Limited variations could be discerned using the categorical recordings. Morning visibility was greater than 12,000 m on nearly 700 days and between 7500 m and 12,000 m on 380 days. Morning visibility for 16 days was between 4000 m and 7500 m. Evening visibility was consistently better, with nearly 1070 days having visibility greater than 12,000 m. Lack of continuous data precluded further use of visibility data in analyses (such as in imputing missing values in air pollutant data series).

Table 6. Summary Statistics for Meteorologic Parameters[a]

Variable	Mean	Median	SD	QD	Min	Max
Temperature (°C)	29.77	30.1	3.06	2.24	23.0	48.0
Relative humidity (%)	75.2	76.0	10.77	6.5	38.0	98.0
Dewpoint (°C)	23.49	23.4	2.85	1.65	15.0	34.4
Wind speed (km/hr)	8.76	8.2	5.82	5.2	0.0	25.0

[a] Data are summarized for the 1096 days that had available data over the study period. SD indicates standard deviation; QD indicates quartile deviation; Min indicates minimum number; Max indicates maximum number.

CONSTRUCTION OF ALTERNATIVE EXPOSURE SERIES AND SELECTION OF A SINGLE BEST EXPOSURE SERIES FOR CORE MODEL DEVELOPMENT

Two salient features of our data, a high percentage of missing readings and a low correlation among daily readings recorded by various AQMs, precluded the use of the daily mean of all AQM readings as a meaningful measure of population exposure. We elaborate on these features in the remainder of this section and describe the rationale for the construction of alternative exposure series and the selection of a *best available* exposure series for core model development. We also justify our approach through comparisons with approaches used in previous studies.

Missing Values

A significant number of AQM readings were missing. As per the protocol of the CPCB, readings were to be taken at three 8-hour intervals over a 24-hour period. Each site was to be monitored at least twice per week, so that 100 to 120 monitoring days per year per AQM would be available. Table 7 shows the percentage of missing observations for each AQM. In addition to missing days, whenever AQMs were operated on designated days, some days of the week were consistently missed. For example, readings for AQM 1 were available for 83% of all Mondays but only 8% of all Tuesdays and 0% of all Wednesdays in the period surveyed. Also, weekend data were virtually nonexistent and hence could not be included in any analysis. Similarly, AQMs did not operate on festivals and public holidays, which added approximately 10–15 missing days per year.

AQM Correlation

Another important feature of the data is the low correlation between most pairs of AQMs. Figures 10 to 13 show the partial correlations between readings at individual AQMs after eliminating the effect of temperature and relative humidity. A problem with computing correlations is that these are based on a limited number of days on which readings at both AQMs of a given pair have been recorded (Table 8). For example, AQMs 1 and 2 have no days in common, so it was not possible to estimate the correlation between them. As an alternative we also plotted the time-smoothed PM_{10} readings to assess whether the series corresponding to each AQM behaved similarly over time (see Figure 4). Only the industrial AQMs (1–3) displayed similar movements. There was substantial heterogeneity among the other AQMs with respect to their average PM_{10} concentration as well as to the extent and nature of variation over time. A Bartlett test for equality of variances of the various AQMs suggested significant variability in AQM readings ($P = 2.2 \times 10^{-16}$).

We examined possible reasons for the low correlations before proceeding to construct the representative exposure series. First, to address measurement errors we compared the limited data set available from the continuous AQM station that used a β gauge for PM_{10} measurements with its colocated routine AQM station (AQM 3). Some residual data quality concerns seem likely because the correlation was modest ($r = 0.76$). However, given our limited abilities to further validate instrumentation selection, we relied on rigorous QA and QC checks to validate the precision with which measurements were made. Because of the cooperation extended by the TNPCB, all raw field data cards were made available to investigators in addition to routine checks made by CPCB. Based on the QA and QC considerations described earlier, compliance to stated procedures was high and there seemed little reason to believe that systematic errors in measurement across days or AQMs could be responsible for low correlations between AQMs. Second, contributions from changes in wind direction, which would affect plumes from local sources, could not be reliably assessed because only wind direction information was available. Although low correlations persisted even after adjustments for wind direction, no information could be gathered on plume height in relation to sampler height. Finally, we met with officials of the TNPCB and the CPCB to ascertain the spatial scale for each AQM to see if small AQM footprints together with differences in source profiles may have been responsible for the lack of correlation and movement over time.

Table 7. Percentage of Missing Observations at AQMs by Day-of-the-Week

	AQM				
	1	2	3	4	5
Sunday	100	100	100	98	99
Monday	17	100	100	58	49
Tuesday	92	100	4	52	53
Wednesday	100	7	100	54	48
Thursday	12	100	8	58	58
Friday	100	10	100	53	54
Saturday	100	100	100	97	98

Table 8. Percentage Overlap of Monitoring Days Across AQMs

	AQM				
	1	2	3	4	5
1	100	0.0	37.0	16.4	34.7
2	0.0	100	0.0	33.1	13.8
3	37.0	0.0	100	17.4	34.7
4	16.4	33.1	17.4	100	6.1
5	34.7	13.8	34.7	6.1	100

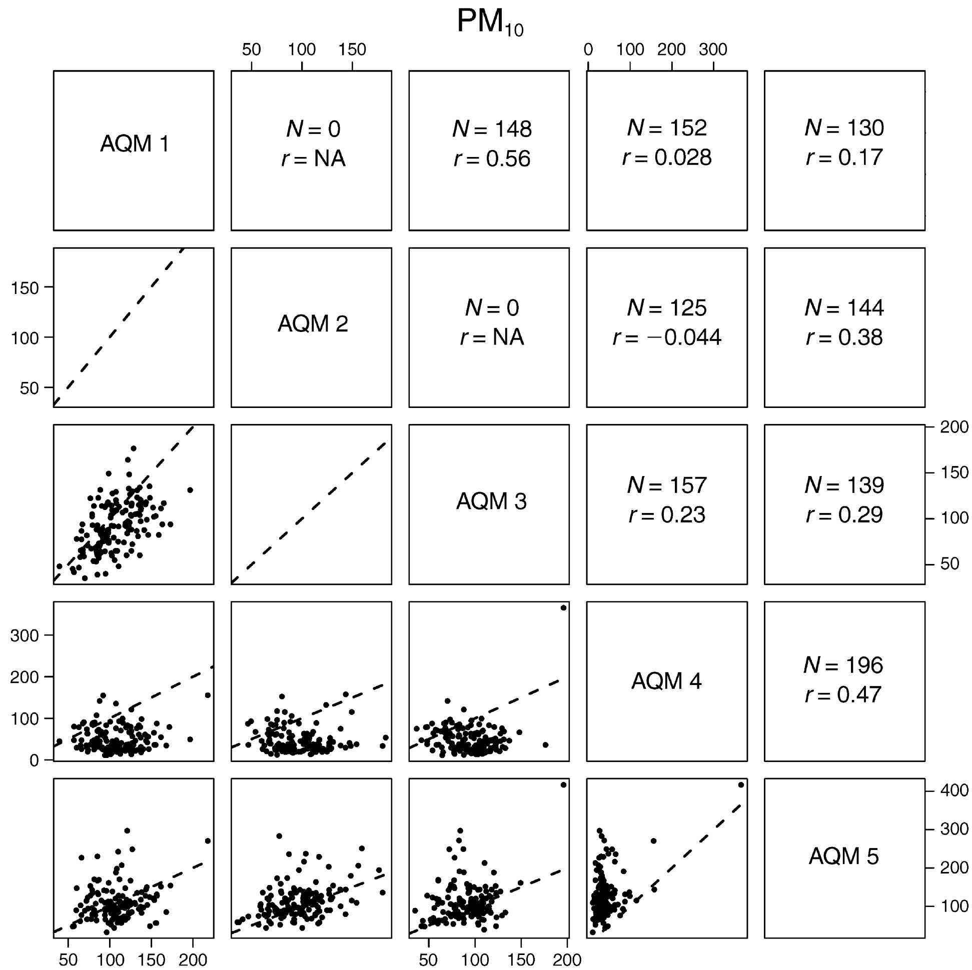


Figure 10. Pairwise scatter plots of PM_{10} (μg/m^3) across monitors. *N* indicates the total number of overlapping observations during the study period; *r* indicates the partial correlation coefficient adjusted for season and meteorologic parameters.

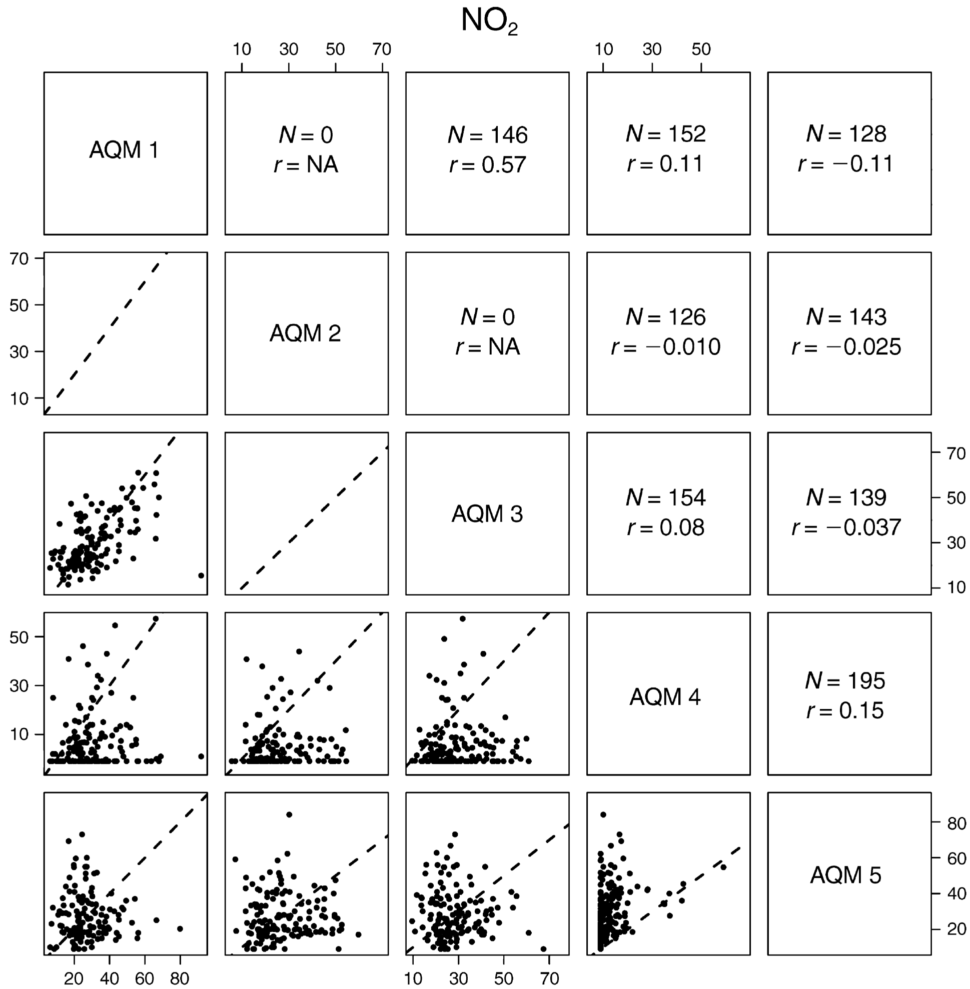


Figure 11. Pairwise scatter plots of NO_2 (μg/m^3) across monitors. *N* indicates the total number of overlapping observations during the study period; *r* indicates the partial correlation coefficient adjusted for season and meteorologic parameters.

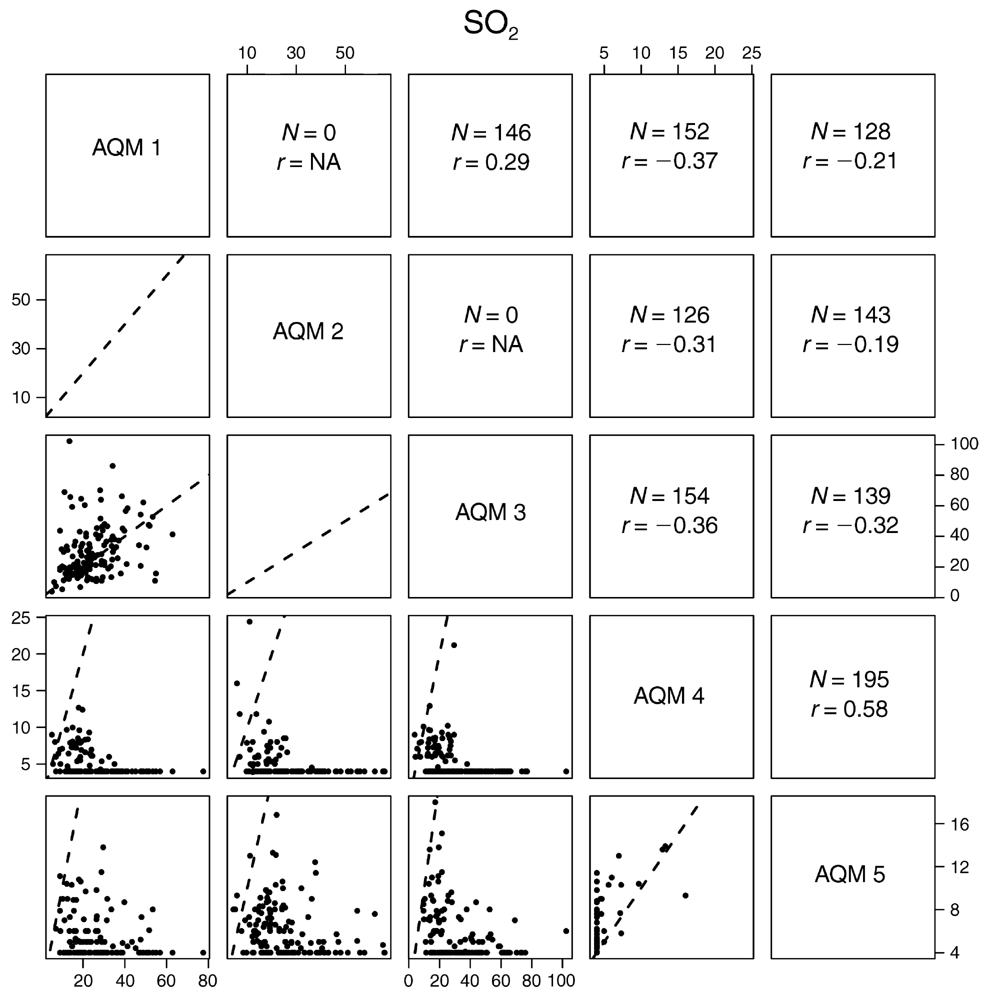
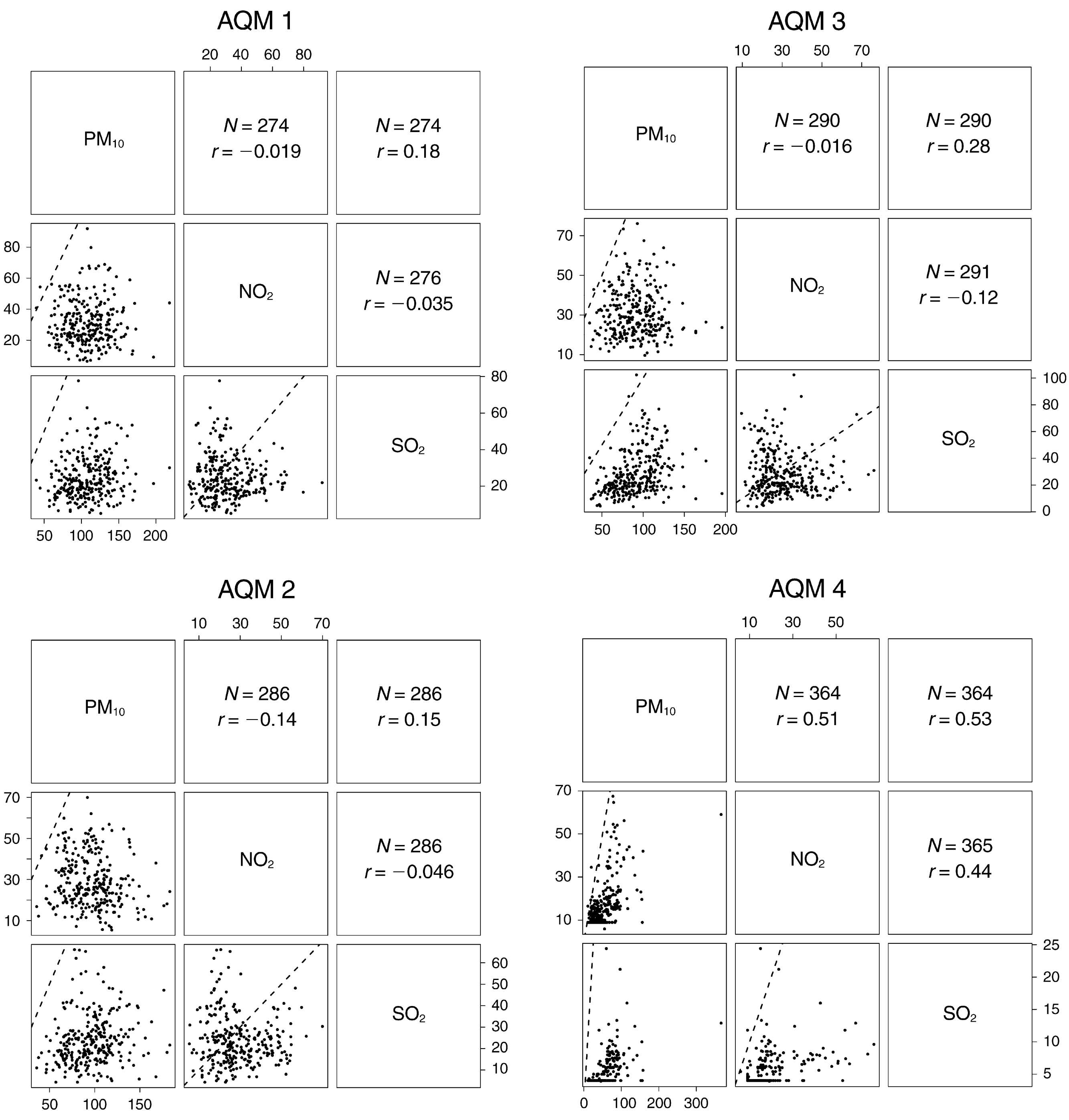


Figure 12. Pairwise scatter plots of SO_2 (μg/m^3) across monitors. *N* indicates the total number of overlapping observations during the study period; *r* indicates the partial correlation coefficient adjusted for season and meteorologic parameters.



(*Figure continues on next page*)

Figure 13. Monitorwise scatter plots of pollutants for AQMs 1–5. *N* indicates the total number of overlapping observations during the study period; *r* indicates the partial correlation coefficient adjusted for season and meteorologic parameters. All units are in μg/m^3.

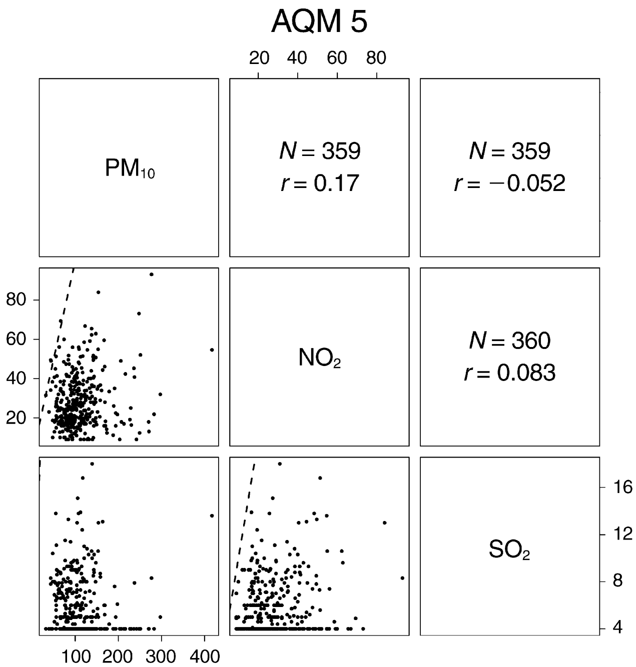


Figure 13 (*Continued*).

Spatial Scale of AQMs

According to information provided by TNPCB, the spatial scale (footprint) for each AQM was estimated to be in the range of 4–10 km. We were unable to find any documentation to support this estimation, but presume that it may be available in official documentation validating the setup of the ambient air quality network in the city. TNPCB emphasized that the small footprint served as the rationale for operating multiple AQMs within the city to capture the spatial variations and gradients in air quality. As described earlier, AQMs were located in zones with assigned land-use classifications of residential, commercial, or industrial. They were also placed at varying distances from pollutant sources to represent relative exposure potentials for populations in that zone from alternative sources. For example, populations in commercial zones are expected to be most affected by traffic-related exposures, populations in industrial zones are expected to be most affected by industry-related exposures, and populations in residential zones are expected to experience only background-level pollutant exposures.

In the absence of detailed emissions inventories for the city, given the differences in types and strengths of sources (such as industrial or vehicular), the pollutant concentrations recorded across AQMs with small footprints (with additional unknown contributions from dispersion-related variables) could thus be expected to be uncorrelated. Most nonindustrial AQMs were located more than 10 km from each other (4–10 km being the estimated footprint), offering a possible explanation for the low correlations. Industrial AQMs displayed good correlation, presumably because of their close proximity to each other (5–10 km) and the similarity in source types and strengths. Because the population density was high across all zones of the city (Figure 2), it became clear that more than one neighborhood-scale AQM would be needed to adequately represent pollutant exposure for the entire city population. None of the AQMs used in this study could be excluded as they were likely to represent specific *typical exposures* for significant proportions of the population. In the next section we describe a review of past approaches to combine AQMs and provide justification for a modified approach to address the unique challenges of the exposure data set described above.

Approaches Used in Previous Studies

While previous studies such as NMMAPS and APHENA have used the daily average of AQM readings to measure ambient PM exposure, the lack of correlation among AQMs in Chennai suggests that the daily average of all available AQM readings is not an appropriate measure of ambient PM concentrations across the city. For example, the average may increase sharply from one day to another simply because a high reading is recorded at one AQM rather than because of any real increase in the average PM concentrations across the city. A regression of total mortality on such a daily average would lead to an underestimate of the true risk ratio because fluctuations in the average concentration would be artifactual, depending on AQM availability, and would not be accompanied by any corresponding change in mortality. We also did not use a centering technique to combine the AQM readings. This technique first centers nonmissing observations x_{ij} corresponding to AQM i on day j by the annual AQM mean $\bar{x}_i$. The centered data are then added to the annual mean of all AQMs $\bar{x}$ to get the centered observations $d_{ij} = x_{ij} - \bar{x}_i + \bar{x}$. The exposure series is the daily average of all available d_{ij}. Effectively, centering shifts the base of each AQM to the same level $\bar{x}$. This is logical if there is not much variation between mean pollutant concentrations of the individual AQMs. The Chennai data show distinct variations among AQM readings in residential, commercial, and industrial zones. Ideally, statistical analysis should be based on mortality data disaggregated by zone. An acceptable alternative could be to model AQM effects separately (either singly or severally) rather than by combining them into a spurious level of concentration that may be quite distinct from what is observed in different parts of the city. An argument in favor of centering would be that because we are primarily interested in a slope coefficient, a change of base should not be of much concern. However, the pooled data have an inflated variability owing to the between-AQM differences. This is not a reflection of the true variability in the exposure concentrations (as given by each of the within-AQM variabilities), but is simply induced by the pooling of the data. The likely consequences would be to get underestimates of the slopes, with smaller and somewhat spurious standard errors of estimates. The heterogeneity in the pooled data would be even more difficult to explain if the within-AQM variability were not the same — as was the case for Chennai.

While we recognize that approaches other than simple averaging and centering (which for reasons described earlier are not applicable to the Chennai data) may have been used in individual studies, we did not see a parallel to the limitations of our exposure data described in the literature. We therefore have had to develop new methods which address the specific features of our exposure data set.

Approaches Used in the Chennai Study

We developed several alternative approaches to address heterogeneity among AQMs. First, we modeled AQM effects separately instead of combining them (in the face of significant missing values) to avoid calculating a spurious level of concentration that may have been quite distinct from what was observed in different parts of the city. We used the simplest approach, which was to create a separate exposure series for each of the five AQMs with the most complete data sets (i.e., AQMs 4 and 5 and the industrial AQMs [1–3]). However, owing to the limited footprints of

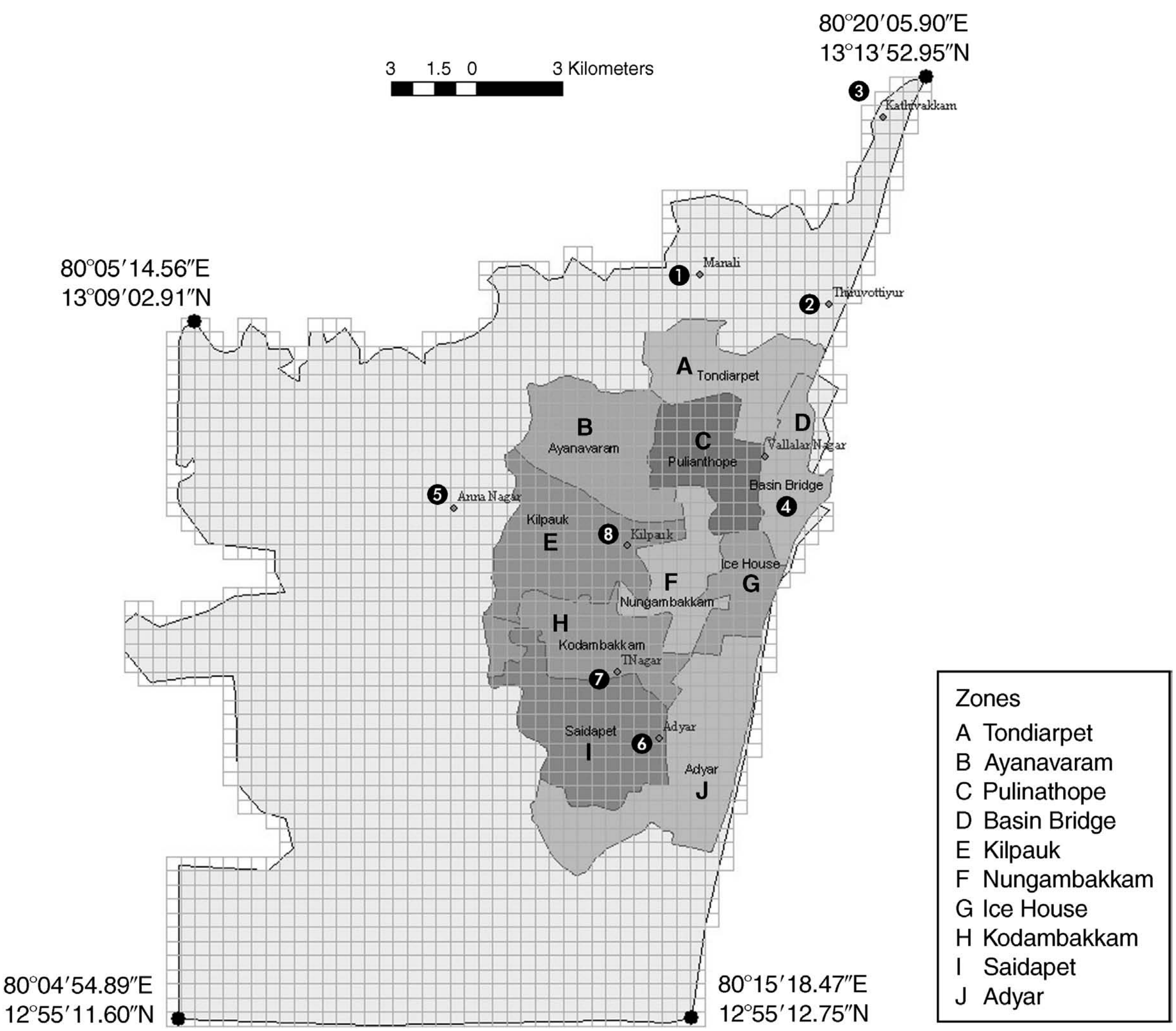


Figure 14. Map of Chennai city superimposed by a grid with 0.5 km^2 squares, showing the 10 zones and 8 AQMs. Source: Prepared by SRMC using base maps for Chennai city provided by Ms. Tata Consultancy Services.

AQMs, none of the individual exposure series could be expected to reflect average population exposures to pollutants across the entire city. Moreover, the differences in the behavior of the individual series over time imply that each had a different correlation with total mortality. We then developed a multiple-AQM model that included a separate exposure series for each AQM in a single regression model. Such a model would allow for a better fit. The low correlations between pairs of AQMs would mean that the slope coefficients corresponding to each AQM would not change appreciably from the corresponding coefficient in the single-AQM model. Although the multiple-AQM exposure series attempted to address the issue of AQM heterogeneity, two of the four nonindustrial AQMs (AQM 6 and 7) had many missing and invalid data periods; therefore, the multiple-AQM model for the most part was a two- or three-AQM model. Considerable potential for exposure misclassification could still be expected in the multiple-AQM model, although misclassification may have been somewhat lower compared with that of single-AQM models.

This reasoning led to the development of a zonal model which disaggregates exposures and mortality at the level of individual zones (administrative units defined by the Chennai Corporation as per Figure 14). The zonal exposure series, in our judgment, represents the best available exposure series for creation of the core (zonal) model, as explained in the next sections.

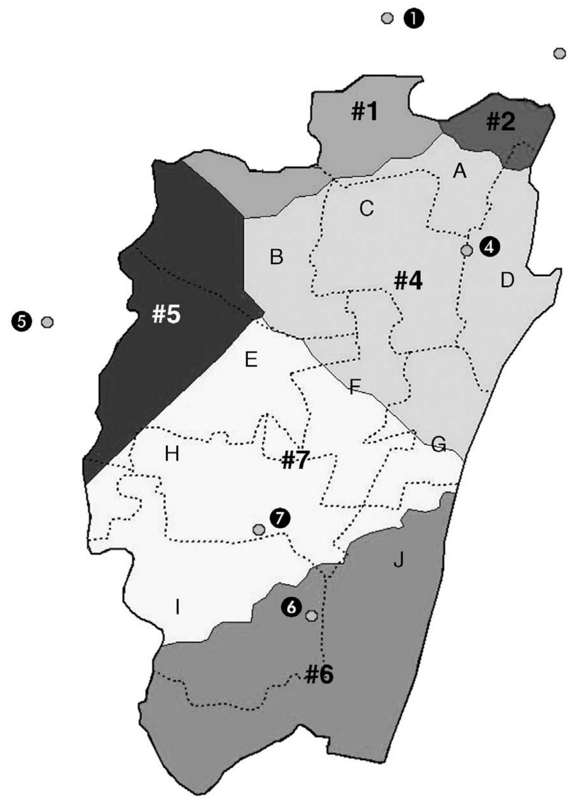


Figure 15. Map of Chennai city showing the areas (#) allocated to AQMs 1, 2, 4, 5, 6, and 7. Source: Prepared by SRMC using base maps for Chennai city provided by Ms. Tata Consultancy Services.

Creation of a Zonal Exposure Series for Core (Zonal) Model Development

Given the problems of poor inter-AQM correlation and the small footprints of AQMs, we explored spatial modeling of the data by assigning deaths to the nearest available AQM. However, a full-blown geostatistical analysis (Diggle and Ribeiro 2007) of the data could not be conducted because residential address information was unavailable for more than 70% of the recorded deaths. The only consistently available information was the zone where a death occurred because mortality was recorded by the individual zonal offices according to the place of death. To fit a regression of zone-specific mortality on average daily exposure within the same zone, we would then either need an AQM in each zone (which was not available) or a derivation of daily zone-specific average exposure through spatial interpolation. Spatial interpolation deals with evaluating workable estimates for exposure data for nonsampled locations within a zone, based on data at sampled locations (i.e., AQM stations). The concentration at a nonsampled site within a zone could be best described by the concentration at the nearest sampled site. However, not all points in a zone have the same nearest AQM. For example, in Figure 14, locations on or near the southern boundary of zone A are closest to AQM 4, whereas locations on the northwest boundary are closest to AQM 1. To address this we superimposed a grid with squares equal to 0.5 square kilometers on the map in Figure 14. To each grid cell we then assigned the exposure recorded at the AQM nearest to the cell's centroid. Using this inverse distance weighting approach, four out of ten city zones were assigned to a single AQM (zones C, D, H, J); four other zones had two *nearest AQMs* available (zones E, F, G, I); and two zones had three *nearest AQMs* available (zones A and B) as shown in Figure 15. Because the maximum distance between a grid cell and its nearest available AQM was of the same order as the estimated footprint of the AQMs (as a result of a network of AQMs being available across the city), this offered the best available approach to minimize exposure misclassification for a grid (whenever a measurement on its assigned AQM was available). Also, while the zone exposure on a given day could still be driven by the availability of the most influential AQMs, because mortality is aggregated and assumed to be uniformly distributed within a zone, all nearest AQMs could be assumed to provide a reasonable exposure assignment for the zone. For these reasons, we believe the zonal series best represents population exposure to pollutants and the core (zonal) model developed in the study is thus based on the zonal exposure series.

In the next sections we describe only the core model in detail to facilitate easy readability and provide details of all other models constructed using alternative exposure series (single- and multiple-AQM models) in Appendix F (available on the Web). We have included details of model comparisons in the sensitivity analyses and discussion sections.

STATISTICAL MODELS

The zonal model that allows disaggregation of exposures and deaths by zone served as the core model for analysis. In addition, models created using alternative exposure series were used in the sensitivity analyses.

CORE (ZONAL) MODEL SPECIFICATIONS

As mentioned earlier, a full blown geostatistical analysis of the data could not be conducted because address information was not consistently available. This means that it was not possible to fit a fully disaggregated regression model that regresses individual deaths on individual exposures. However, as mentioned earlier, because deaths could be disaggregated by zones, we fitted a quasi-Poisson model of the following form:

$$\log(E[mortality_{it}]) = \alpha_{0i} + \beta_1 PM_{i(t-1)} + f_1(t) + f_2(temp_t) + f_3(rh_t), \qquad (1)$$

where $mortality_{it}$ is the all-cause mortality for zone i on day t, α_{0i} is the intercept for zone i and $PM_{i(t-1)}$ is the average PM_{10} concentration for zone i on day $t-1$. The function $f_1(t)$ accounts for any temporal variations in mortality not explained by variations in daily PM_{10} concentrations or in weather. For functions $f_2(temp_t)$ and $f_3(rh_t)$, $temp_t$ and rh_t are the average daily temperature and average daily relative humidity, respectively, on day t. Although $mortality_{it}$ is directly available from the observed data, we needed a reasonable approximation for PM_{it}. Because not all zones have AQMs and not all points within a zone have a unique nearest AQM, we superimposed a grid of 0.5 km^2 squares over all 10 zones within the Chennai Corporation boundary (see Figure 14).

For a fixed zone, i, PM_{it} is measured as the simple average of the readings from all grid cells (located in zone i on day t [i.e., for a fixed zone i on a given day t]),

$$PM_{it} = \frac{\sum_{k \epsilon S_{it}} PM_{itk}}{N_i} \qquad (2)$$

where S_{it} is the set of N_i grid cells with centroids located in zone i for which a valid exposure is available on day t (i.e., for which readings have been recorded at the nearest AQM on the day). PM_{itk} is the exposure assigned to cell k in S_{it}. While an entire zone cannot always be assigned to one unique AQM (see Figure 15), each cell in the grid is sufficiently small for all points within the cell to have one single nearest AQM (see Figure 14). For the sake of uniqueness, we assigned to each grid cell the exposure recorded at the AQM closest to its centroid.

To illustrate how the daily average PM_{10} concentration over a zone is calculated, we will use zone B as an example. As shown in Figure 16, the zone may be approximated by a grid of 214 squares, each with an area of 0.5 km^2. For this set of squares, AQM 1 is the nearest available monitor for 43 squares, AQM 5 is the nearest for 86 squares, AQM 4 is the nearest for 84 squares, and AQM 7 is the nearest for 1 square.

If on a particular day, readings have been recorded at all four of those AQMs and are, for example, 85 µg/m^3 (AQM 1), 74 µg/m^3 (AQM 5), 95 µg/m^3 (AQM 4) and 86 µg/m^3 (AQM 7), the average PM_{10} concentration for zone B on this day would be estimated as:

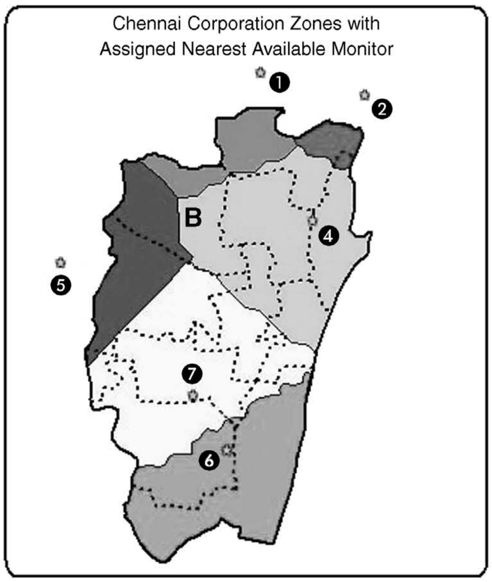
$$\frac{43 \times 85 + 86 \times 74 + 84 \times 95 + 1 \times 86}{43 + 86 + 84 + 1} = 84.6 \ \mu g/m^3.$$

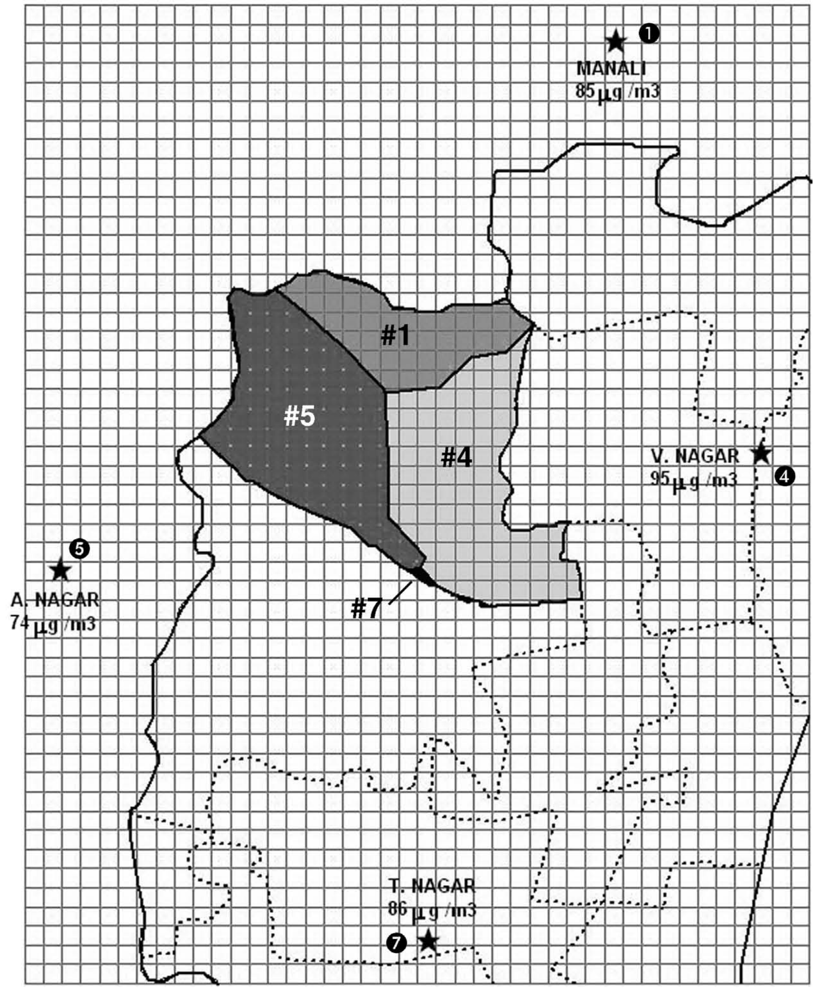
If, however, AQM readings were recorded only at AQMs 1 and 4 with the same readings as in the previous example, the average PM_{10} concentration for zone B would then be estimated as:

$$\frac{43 \times 85 + 84 \times 95}{43 + 84} = 91.6 \ \mu g/m^3.$$

If none of the four AQM readings were recorded on a particular day, the average PM_{10} zone B concentration would be treated as a missing observation, and deaths from zone B on the same day would be automatically excluded from the analyses.

This approach implies that on days when data from all AQMs representing a zone are available, the average zone exposure would be adequately weighted by the proportion of grids close to particular AQMs, and on days when none of the grids in a zone have an available value they would be excluded from the analysis. A limitation of this approach is that on days when data from at least one of the AQMs influencing a zone are missing, there would be some exposure misclassification. On such days, we ignored all cells assigned to AQMs with missing data, and instead used the average overall grid cells for which the reading at the nearest AQM had been recorded. This effectively replaced the grid cells that had missing readings with the average value of grid cells for which readings





AQM	PM_{10}	Number of Squares
1	85	4
4	95	84
5	74	86
7	86	1

Figure 16. An example using zone B to illustrate how average PM_{10} concentration is calculated for a given day. The area of zone B is approximated by a grid of 214 squares (0.5 km^2 each). Source: Prepared by SRMC using base maps for Chennai city provided by Ms. Tata Consultancy Services.

had been recorded, or in terms of AQMs, substituted the values of a weighted average of adjacent AQM readings for a missing value. This leads to some misclassification but we can assume that the degree of misclassification is small if the spatial gradient in exposures is gradual. This was true for Chennai, as can be seen from the map of average grid-wise daily PM_{10} exposures based on available data given in Figure 17. Further, as mentioned previously, because the distribution of deaths is assumed to be uniform across a zone, all nearest AQMs can be assumed to represent likely zonal exposures. Finally, since the zonal model allows individual zones to be included or excluded depending on AQM availability on each day, the data sets for AQMs 6 and 7, despite having many missing observations, could be included for all valid periods (such as most of 2003 and all of 2004). Also, industrial AQMs outside the city boundary were included in case they were the closest available AQM for select zones (or parts thereof).

Because data were missing for multiple days we only considered a lag of one day in modeling the effect of PM_{10} on mortality (i.e., total daily mortality on day t is assumed to be influenced by PM_{10} concentrations on day $[t-1]$). Fitting distributed lag models to our data for a general lag of order $P > 1$ is complicated by the high percentage of missing observations, as such models require that PM_{10} readings are recorded for $P + 1$ consecutive days. By contrast, including a single lag term can be accomplished by a shift of the PM_{10} series. We have considered the effect of shifts of size 2 and 3 in the section on sensitivity analysis.

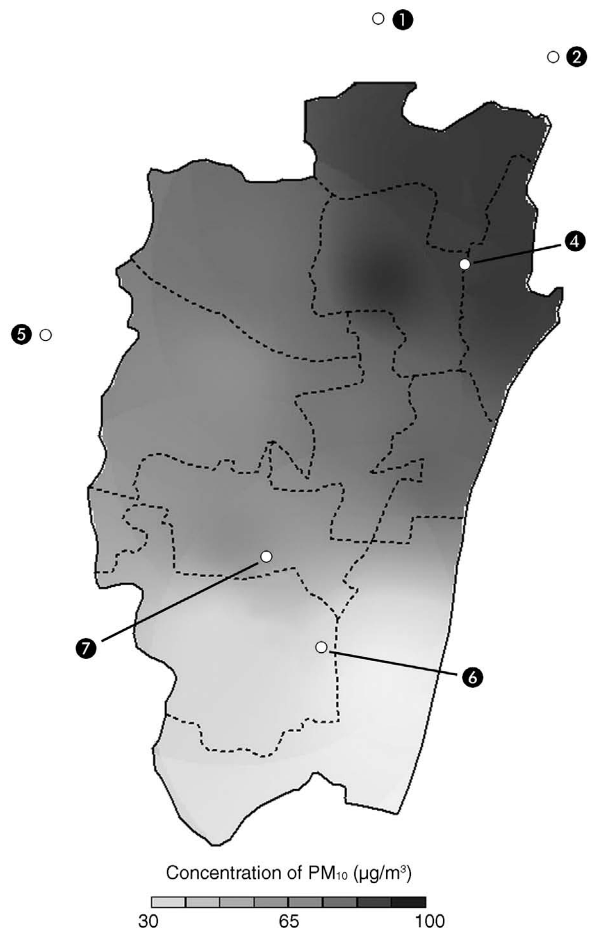


Figure 17. Spatial gradients in PM_{10} concentration with AQM locations (generated using annual averages assigned to each grid from the nearest available AQMs for the study period). Source: Prepared by SRMC using base maps for Chennai city provided by Ms. Tata Consultancy Services.

MODELS USED IN SENSITIVITY ANALYSES

The zonal exposure series used for core model development evolved from an elaborate examination of alternative exposure series created using data from single or multiple AQMs. While the zonal series is likely the best choice among available exposure series, we used models developed using all alternative series in sensitivity analyses. Accordingly, we developed:

- single-AQM models using available exposure data from
  - AQM 4;
  - AQM 5;
  - the average of the industrial AQMs (because industrial AQMs 1–3 displayed similar movements over time and showed good correlations, we combined their readings);
- multiple-AQM models using:
  - available exposure data from both nonindustrial monitors AQM 4 and AQM 5;
  - available data from both nonindustrial monitors AQM 4 and AQM 5 together with the averaged data from the 3 industrial monitors (average of AQMs 1, 2, and 3); and
  - available and imputed data (to fill in missing values) from both nonindustrial monitors AQM 4 and AQM 5 together with the averaged data from the 3 industrial monitors (average of AQMs 1, 2, and 3).

We provide details of the models in Appendix F but discuss comparisons with the core model in the main sections of this report.

SELECTION OF CONFOUNDER DEGREES OF FREEDOM

Given an exposure series $\{x_t\}$, for the core or alternative models (used in sensitivity analyses) we have used quasi-Poisson GAMs of the form:

$$\log(E[mortality_t]) = \alpha_0 + \beta x_t + f_1(t) + f_2(temp_t) + f_3(rh_t), \quad (3)$$

where $mortality_t$ is the total number of all-cause mortality on day t; x_t is PM_{10} concentration on day t; $temp_t$ is the average daily temperature on day t; and rh_t is the average daily relative humidity on day t.

The function $f_1(t)$ accounts for any temporal variations in mortality not explained by variations in daily PM_{10} concentrations or in weather. The effect of known confounders such as temperature and relative humidity is modeled through the functions $f_2(temp_t)$ and $f_3(rh_t)$ where $temp_t$ and rh_t are the average daily temperature and average daily relative humidity, respectively, on day t.

For each of the exposure series choices used, the functions $f_1(t)$, $f_2(temp_t)$, and $f_3(rh_t)$ are assumed to be smooth and are modeled using penalized splines (Hastie and Tibshirani 1990; Ruppert et al. 2003). The degrees of freedom for the smooth terms were selected based on the approach of Dominici and colleagues (2004). Many approaches for selection of degrees of freedom have been discussed in the smoothing literature and most suggest selection based on optimization of a measure of goodness-of-fit such as the generalized cross validation criterion (GCV; Wood 2004) or the Akaike Information Criterion (Ruppert et al. 2003). We used the method specified by Dominici and colleagues (2004) because it recognizes that the objective of time-series studies of air pollution and mortality is not to optimize the predictive power of the model but to maximize the precision with which the coefficient β_1 is estimated. We provide an algorithm for selection of degrees of freedom for each smooth term based on the bootstrap technique after an extensive consideration of the complications of concurvity in such models. Concurvity is the curve equivalent of multicollinearity and arises due to the dependence of both exposure and mortality on the same covariates (i.e., temperature, relative humidity, and time).

To estimate the degrees of freedom for the time component $f_1(\bullet)$ in equation 3, the algorithm involved four steps.

1. Estimate the degrees of freedom, $\hat{d}$ which best predict the exposure, x_t as a function of t in the conventional sense of minimizing the GCV.
2. Fit the model in equation 3 with x_t and t on the right side using degrees of freedom larger than $\hat{d}$ for $f_1(\bullet)$. We used $d^* = 3\hat{d}$ degrees of freedom for $f_1(\bullet)$. This leads to an unbiased estimate of β_1 but it will have a large variance.
3. Perform a bootstrap analysis by generating values of $mortality_t$ from the model in equation 3 (again with only x_t and t on the right side) assuming d^* degrees of freedom for $f_1(\bullet)$.
4. Fit the model using these generated values and assuming 1, 2, . . ., d^* degrees of freedom for $f_1(\bullet)$. Find the bias and variance of each of the estimates and choose that value d which leads to an unbiased estimate with the smallest variance.

To get the degrees of freedom for the smooth terms corresponding to temperature and relative humidity, we repeated the above algorithm, substituting temperature or relative humidity for time. Figure 18 illustrates the application of the above approach and the selection of degrees of freedom by optimizing square bias and variance of the estimates. The zonal model does not permit direct application of the procedure of Hastie and Tibshirani (1990) because that procedure requires a single-exposure series. Degrees of freedom were calculated separately for each confounder for each of the single-series models, and the maximum of these was used for the zonal model. Applying this approach to the zonal series led to the selection of 8 *df* per year for time, 6 *df* per year for temperature, and 5 *df* per year for relative humidity. This choice was subsequently applied in all models used for sensitivity analyses as well.

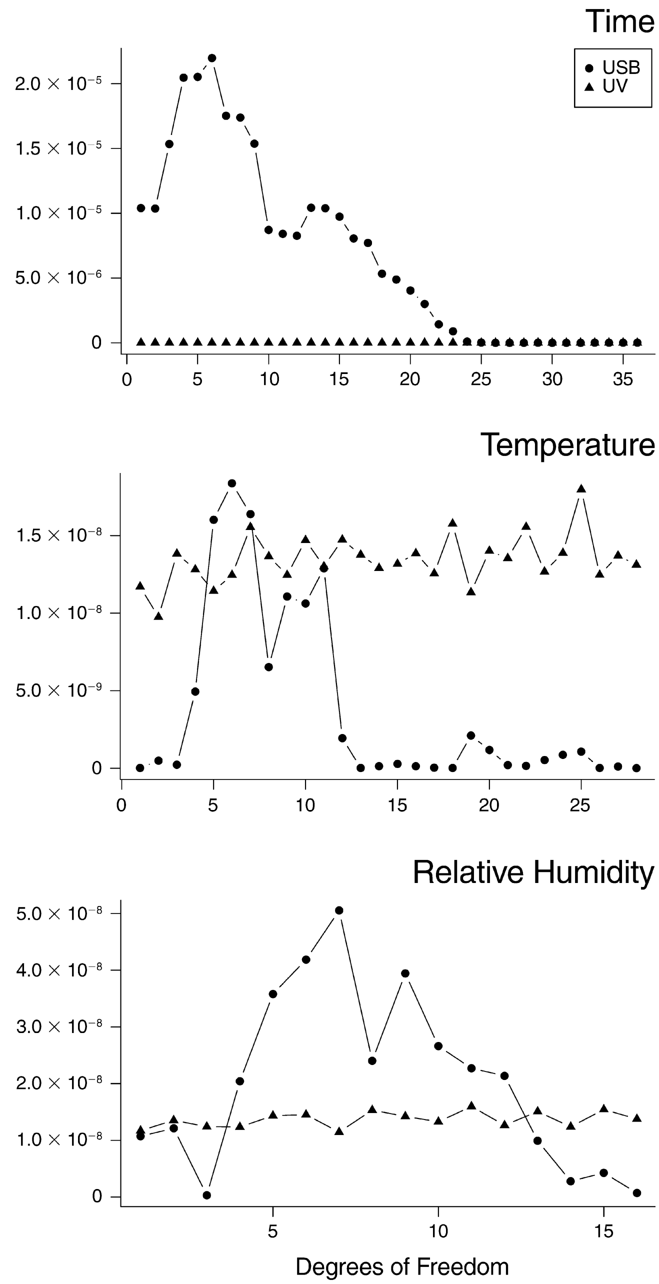


Figure 18. Confounder degrees of freedom selection by optimizing square bias and variance of the estimated RR for PM_{10} following the method of Dominici and colleagues (2004). USB indicates unconditional squared bias; UV indicates unconditional variance.

RESULTS

CORE (ZONAL) MODEL

Detailed Results

Tables 9 and 10 provide details of the core (zonal) model results. The model estimates a log mortality ratio of 0.00044 (95% CI = 0.00017 to 0.00071). This corresponds to a 0.44% (95% CI = 0.17 to 0.71) increase in mortality per 10-μg/m^3 increase in PM_{10} concentration.

The zonal model uses 8, 6, and 5 degrees of freedom per year to adjust for confounding by time, temperature, and relative humidity, respectively. The model explains 50% of the total deviance. The estimated overdispersion parameter for the model is 1.25, which is lower than that for any other model fitted to the data.

Residual Analysis

The studentized deviance residual from the core (zonal) model was plotted against predicted mortality, PM_{10}, time, and temperature (Figure 19). Most residuals are scattered around zero with no specific pattern. Though the points beyond the upper and lower lines indicate the presence of outliers, we have addressed this through outlier analysis in the next section.

SENSITIVITY ANALYSES

Results from all models are presented together with other results of sensitivity analyses in Table 10. The percentage change per 10-μg/m^3 increase in pollutant concentration was estimated by multiplying the β-coefficient from the regression model by 1000.

Results from the Alternative Models Exposure Series

Single-AQM Models The three single-AQM models estimated a change in mortality per 10-μg/m^3 increase in PM_{10} concentration for each of the three exposure series: AQM 4 = 0.43% (95% CI = 0.11 to 0.95); AQM 5 = 0.33% (95% CI = −0.10 to 0.76); and the industrial AQM average = 0.30% (95% CI = −0.11 to 0.71) (Table 10). The three models explain 49.5% (AQM 4), 48.7% (AQM 5), and 43.8% (industrial AQM average) of the total deviance. The estimated overdispersion parameters for the models are 1.40 (AQM 4), 1.54 (AQM 5), and 1.56 (industrial AQM average).

Multiple-AQM Models Using Available Data The multiple-AQM models did not find statistically significant evidence of a differential effect of AQMs ($P = 0.86$). Therefore, we pooled the individual β coefficients using a weighted average of inverse variances to arrive at a common pooled estimate. The two multiple-AQM models estimated a 0.39% (95% CI = 0.17 to 0.62) and 0.30% (95% CI = 0.10 to

Table 9. Output of Core (Zonal) Model[a]

Parameters	Estimate	95% CI	*P* Value
$\hat{\alpha}$	1.7606474	1.705 to 1.815	$< 2 \times 10^{-16}$[b]
$\hat{\alpha}_2$	0.5807508	0.522 to 0.638	$< 2 \times 10^{-16}$[b]
$\hat{\alpha}_3$	0.8815847	0.826 to 0.936	$< 2 \times 10^{-16}$[b]
$\hat{\alpha}_4$	0.3389721	0.282 to 0.396	$< 2 \times 10^{-16}$[b]
$\hat{\alpha}_5$	0.9744544	0.919 to 1.029	$< 2 \times 10^{-16}$[b]
$\hat{\alpha}_6$	0.0195223	−0.042 to 0.081	0.53352
$\hat{\alpha}_7$	0.6474729	0.591 to 0.704	$< 2 \times 10^{-16}$[b]
$\hat{\alpha}_8$	0.4317842	0.374 to 0.489	$< 2 \times 10^{-16}$[b]
$\hat{\alpha}_9$	−0.0060403	−0.056 to 0.068	0.84957
$\hat{\alpha}_{10}$	0.3001306	−0.358 to −0.242	$< 2 \times 10^{-16}$[b]
$\hat{\beta}$	0.0004414	0.00017 to 0.00071	0.00129[c]
f_1 (*time*)		$df = 24$	$< 2 \times 10^{-16}$[b]
f_2 (*temperature*)		$df = 17$	1.44×10^{-7}[b]
f_3 (*relative humidity*)		$df = 14$	0.511

[a] $\hat{\alpha}$ indicates the estimated general intercept; α_i indicates the estimated zone-specific additions to the intercept; $\hat{\beta}$ indicates the estimated coefficient for PM_{10}.
[b] Significant at $P < 0.001$.
[c] Significant at $P < 0.01$.

Table 10. Estimated Effect of PM_{10} from Alternative Models and Sensitivity Analysis for the Core Zonal Model[a]

		10-μg/m^3 PM_{10} Increase		Estimate of Over-dispersion	Deviance Explained (%)	*P* Value
	n	% Change (95% CI)	RR (95% CI)			
Core (Zonal) Model, Lag 1	742	0.44 (0.17 to 0.71)	1.004 (1.002 to 1.007)	1.25	50	0.001
Alternative Exposure Models						
Single AQM						
AQM 4	358	0.43 (0.11 to 0.95)	1.004 (1.001 to 1.008)	1.40	49.5	0.133
AQM 5	375	0.33 (−0.10 to 0.76)	1.003 (0.999 to 1.008)	1.54	48.7	0.007
Average Industrial	712	0.30 (−0.11 to 0.71)	1.003 (0.999 to 1.007)	1.56	43.8	0.153
Multiple AQM[b]						
Without industrial AQMs	512	0.39 (0.17 to 0.62)	1.004 (1.002 to 1.006)	1.49	45.0	0.000
With industrial AQMs	731	0.30 (0.10 to 0.50)	1.003 (1.001 to 1.005)	1.66	45.1	0.002
With industrial AQMs + Imputed	1096	0.40 (0.20 to 0.60)	1.004 (1.002 to 1.006)	1.62	44.5	0.000
Sensitivity Analyses						
Zonal model						
Male	742	0.53 (0.13 to 0.93)	1.005 (1.001 to 1.0093)	1.12	30.0	0.012
Female	742	0.38 (0.03 to 0.73)	1.004 (1.0004 to 1.007)	1.18	39.9	0.025
Age (years)						
0–4	742	0.17 (−0.42 to 1.78)	1.002 (0.986 to 1.018)	1.26	21.6	0.834
5–44	742	0.59 (0.01 to 1.17)	1.006 (1.0001 to 1.012)	1.14	41.9	0.042
45–64	742	0.65 (0.17 to 1.14)	1.007 (1.002 to 1.011)	1.08	25.8	0.007
≥ 65	742	0.31 (−0.05 to 0.70)	1.003 (0.999 to 1.007)	1.14	23.4	0.103
PM_{10} lag 0	743	0.30 (0.04 to 0.57)	1.003 (1.0003 to 1.006)	1.21	50.5	0.021
PM_{10} lag 2	741	−0.0005 (−0.28 to 0.28)	1.000 (0.997 to 1.003)	1.30	48.9	0.997
PM_{10} lag 3	740	−0.01 (−0.30 to 0.27)	1.000 (0.997 to 1.003)	1.31	48.7	0.943
Multiple-pollutant (NO_2)	742	0.55 (0.24 to 0.86)	1.006 (1.002 to 1.009)	1.22	50.0	0.000
Distributed weather lag	742	0.37 (0.10 to 0.65)	1.004 (1.001 to 1.006)	1.21	50.8	0.008
Excluding outlier	742	0.46 (0.19 to 0.73)	1.005 (1.002 to 1.007)	1.21	50.1	0.000
Season						
Mar–Aug	742	0.56 (0.21 to 0.92)	1.006 (1.002 to 1.009)	1.22	50.1	0.002
Sep–Feb	742	0.59 (0.23 to 0.96)	1.006 (1.002 to 1.010)	1.22	50.1	0.002
df (time, temp, RH)						
(25, 20, 15)	742	0.47 (0.20 to 0.74)	1.005 (1.002 to 1.007)	1.21	50.1	0.001
(30, 25, 20)	742	0.45 (0.18 to 0.71)	1.004 (1.002 to 1.007)	1.21	50.4	0.001
(35, 30, 25)	742	0.38 (0.11 to 0.65)	1.004 (1.001 to 1.007)	1.20	50.8	0.007
(40, 35, 30)	742	0.38 (0.10 to 0.66)	1.004 (1.001 to 1.007)	1.20	51.1	0.007

[a] *n* indicates the number of days that had available data over the study period; RH indicates relative humidity. Unless otherwise specified, all RR are for PM_{10} of lag 1.

[b] Multiple-AQM models: *Without industrial AQMs* includes available exposure data from both nonindustrial monitors AQM 4 and AQM 5; *With industrial AQMs* adds averaged data from the 3 industrial monitors (AQMs 1, 2, and 3); *With industrial AQMs + Imputed* adds imputed data (to fill in missing values) from both nonindustrial monitors AQM 4 and AQM 5 together with the averaged imputed data from the 3 industrial monitors (AQMs 1, 2, and 3).

0.50) change in mortality per 10-μg/m^3 increase in PM_{10} concentration, corresponding to exclusion and inclusion of the industrial AQMs, respectively (Table 10). The two models accounted for 45.0% and 45.1% of the deviance, respectively. The estimated overdispersion parameters for the models were 1.49 and 1.66, respectively.

Multiple-AQM Models Using Available and Imputed Data The multiple-AQM model with imputed values for missing days also did not find statistically significant evidence of a differential effect of AQMs (P = 0.56). Accordingly we pooled the individual β coefficients using a weighted average of inverse variances to arrive at a

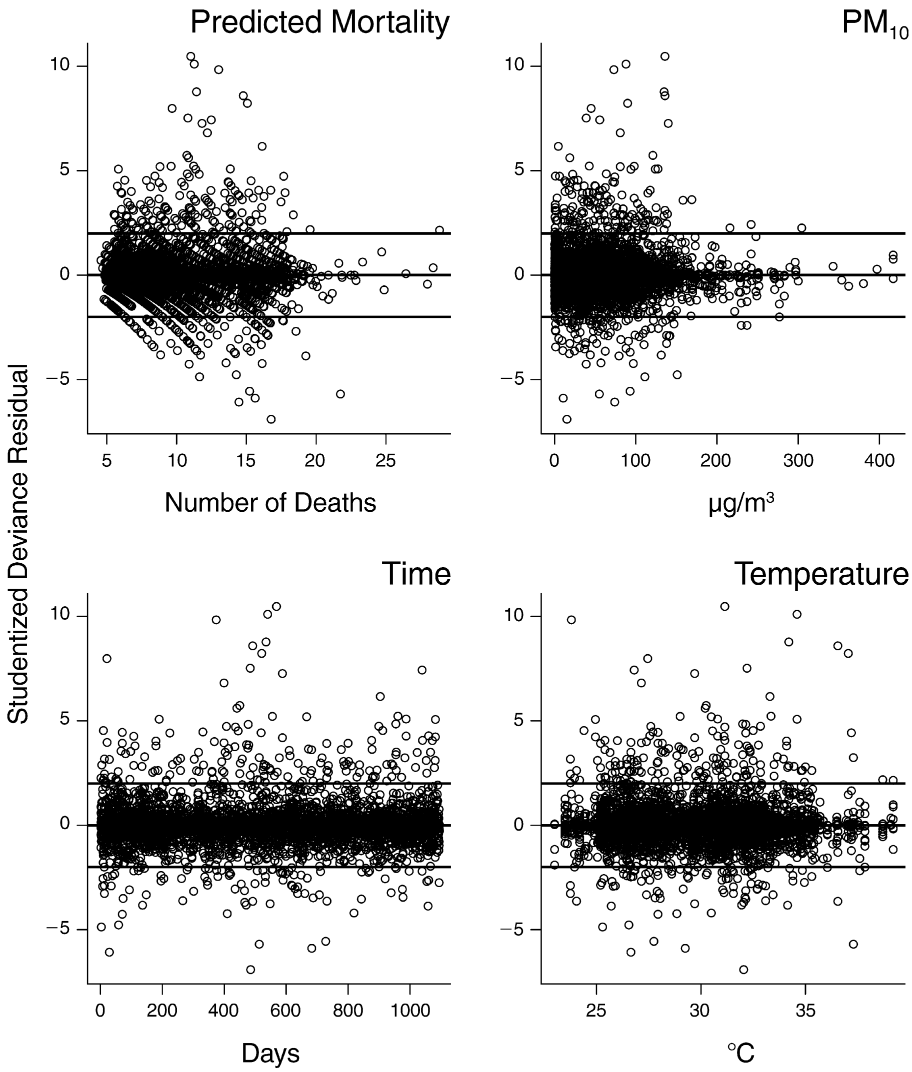


Figure 19. Studentized deviance residual plot against predicted mortality, PM_{10}, time, and temperature for the core model.

common pooled estimate. The model estimated a 0.40% (95% CI = 0.20 to 0.60) increase in mortality per 10-µg/m^3 increase in PM_{10} concentration (Table 10).

Sensitivity Analyses for the Core Model

Confounder Degrees of Freedom To explore the robustness of the relative risk (RR) to choice of confounder degrees of freedom, a detailed sensitivity analysis was carried out for the core model. The results are summarized in Figure 20. The figure shows that with the variation in degrees of freedom from 1 to 45 for time, 1 to 36 for temperature and 1 to 24 for relative humidity, the estimates of RR change slightly from 1.003 to 1.006. This exercise supplemented the optimum choice of degrees of freedom by following the procedure suggested by Dominici and colleagues (2004).

Stratification by Sex and Age All analyses presented in this report are based on total all-cause mortality in Chennai, irrespective of sex or age. Table 10 presents sensitivity of the RR to age- and sex-specific analyses. As before, there was not much variation in the estimated RR, but we observed a general increase in RR from 1.002 (for children between 0 and 4 years) to 1.007 (for adults between 45 and 64 years). Analysis for the group of adults 65 years and older finds an estimated RR of 1.003, but this is likely to be unreliable because the sample size was small. There was

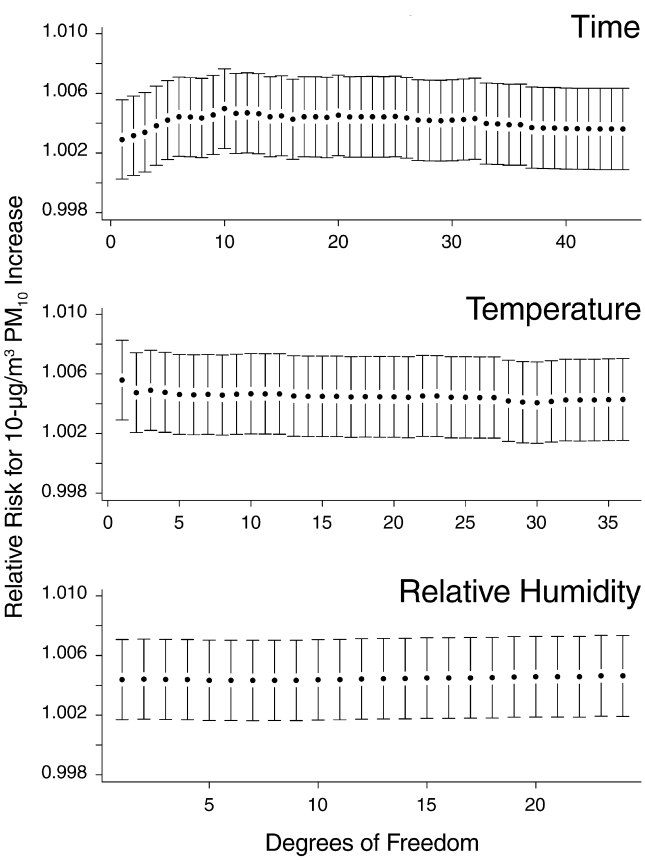


Figure 20. Sensitivity analysis: Changes in the estimated RR for PM_{10} as a result of varying confounder degrees of freedom.

no significant difference in the RR for males (1.005) and females (1.004).

Exposure Lag We have explored the sensitivity of the RR estimate to the lag specification used for PM_{10} concentration. The results are presented in Table 10. The RR changes from 1.004 (95% CI = 1.002 to 1.007) for the core (zonal) model with its 1-day lag to 1.003 (95% CI = 1.0003 to 1.006) for a zero-lag model and is insignificant with 2- and 3-day lags. Another lag specification explored in sensitivity analysis was the lag model assumed for the meteorologic confounders. A distributed lag model with a 7-day lag for temperature and relative humidity was fitted, but no appreciable change was found in the estimates.

Inclusion of Gaseous Pollutants A multiple-pollutant model was fitted to the data with NO_2 as another potential confounder of the effect of PM_{10} concentration. The SO_2 data were not considered reliable enough for inclusion, because many of the values lie below the detection limit. The RR for PM_{10} increased marginally from 1.004 to 1.006 after inclusion of NO_2.

Season-Specific Analysis To examine the effect of season on mortality and to explore whether such seasonal effect confounds PM_{10} concentration, a sensitivity analysis was conducted by incorporating a seasonal effect and a *pollution* × *season* interaction term in the core model. Two seasons were assumed — one from September through February and the other from March through August. The season–exposure interaction was statistically insignificant.

Outliers As is typical of environmental exposures, each of the exposure series is skewed to the right with several outlying observations at the upper end of the exposure scale. Estimates of relative risk typically are highly influenced by such outliers in the sense that inclusion or exclusion of these outliers can have a substantial impact on the estimates. A naive means of assessing the influence of these outliers is to compare the estimates before and after deletion of the top *x*% of exposures. This is not the best approach, for every high exposure is not necessarily an outlier. Moreover, it is not clear how to generalize this approach to the case where the data also contain information on confounders, which in turn may have some outlying values. Formal deletion diagnostics procedures have been developed for several statistical models, and the standard approach is to consider the effect of deletion of each individual in turn on parameter estimates and predicted values. Therefore, we used the formal deletion diagnostic techniques that have been developed for the generalized linear model. While these diagnostics are not directly applicable to the GAM, we could reduce our model to a comparable generalized linear model by using a Poisson link and natural splines with the same degrees of freedom as the corresponding GAM. Figure 21 shows the output obtained from these deletion diagnostics. The figure is a so-called *bubble plot* of studentized DFFIT residuals plotted against fitted values; the size of each bubble is proportional to the corresponding exposure. Vertical reference lines are drawn at two and three times the average fitted value and horizontal reference lines are drawn at −2, 0, and 2 for the studentized residual scale. Influential observations are those which lie outside these reference lines. We have refitted the core zone specific time-series model after dropping the influential values; the results are presented in Table 10. The results show that inclusion or exclusion of outliers had a minor effect on estimated relative risk. The estimate of RR obtained through inclusion of outliers is 1.004 while the estimate is 1.005 when outliers have been excluded from the analysis.

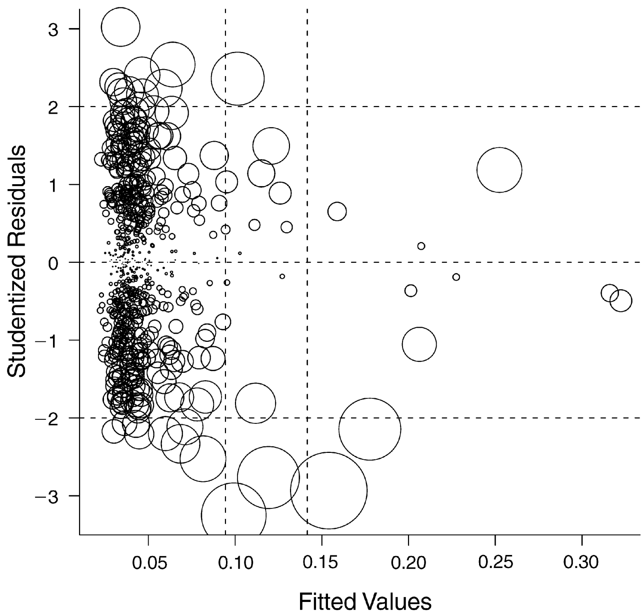


Figure 21. Bubble plot for detection of outliers and influential values for the core (zonal) model.

DOSE–RESPONSE

To test for evidence of nonlinearity we refitted our core model as follows:

$$\log(E[mortality_{it}]) = \alpha_{0i} + f(PM_{i(t-1)}) + f_1(t) + f_2(temp_t) + f_3(rh_t), \quad (4)$$

where $mortality_{it}$ is the mortality for zone i on day t, α_{0i} is the intercept for zone i, $f(PM_{i(t-1)})$ is the smooth function of the average PM_{10} concentration for zone i on day $t-1$. The function $f_1(t)$ accounts for any temporal variations in mortality not explained by variations in daily PM_{10} concentrations or in weather. For functions $f_2(temp_t)$ and $f_3(rh_t)$, $temp_t$ and rh_t are the average daily temperature and average daily relative humidity, respectively, on day t. The estimated dose–response curve is shown in Figure 22A. A curve based on 4 df minimizes the GCV. However, note that the curve is linear over the major range of the data and any nonlinearity is largely driven by a few outlying values at the upper range of exposures. In Figure 22B, we have plotted the results of refitting the dose–response curve after deleting outlying PM_{10} concentrations (concentrations which lie outside the 95% CI for the three-year annual average PM_{10} concentration for the city); the refitted curve is based on 1 df, which corresponds to the minimum GCV. We conclude that there is no significant evidence for nonlinearity in the dose–response model.

Figure 23 shows the results of core and alternative models and the accompanying sensitivity analyses for the core model.

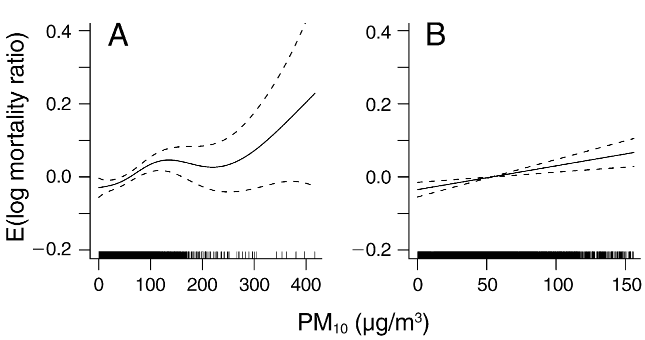


Figure 22. Sensitivity analysis: Dose response curves for PM_{10} using version 1.3-24 of the mgcv package in R and the estimated degrees of freedom for each model. A) The estimated curve from the full data set (4 df); **B)** The estimated curve after deleting outlying exposures (1 df). E(log mortality ratio) indicates expectation of log mortality ratio.

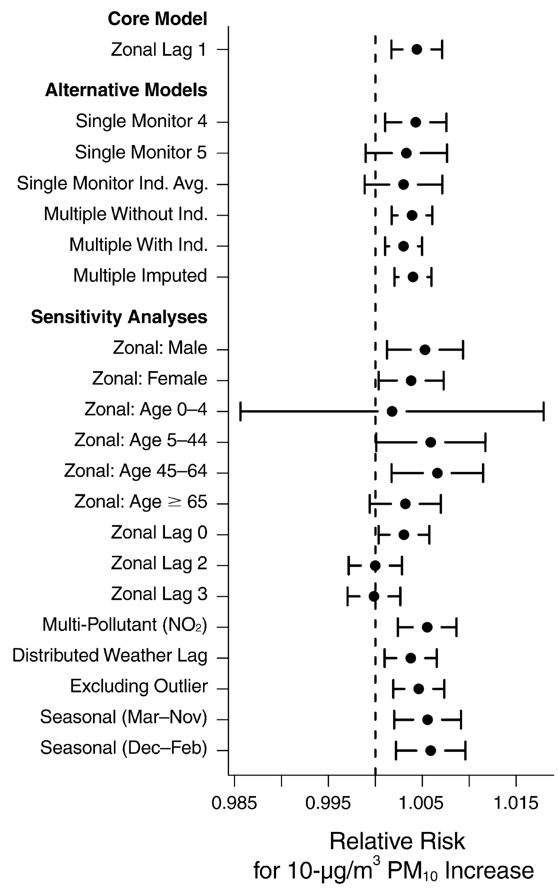


Figure 23. A comparison of the estimated RR's for PM_{10} obtained from the core zonal model, alternative models, and sensitivity analyses. Ind. indicates industrial AQMs; Ind. Avg. indicates average of the industrial AQMs; Multi-pollutant (NO_2) indicates multiple-pollutant model with inclusion of NO_2 as a confounder.

DISCUSSION

The various models may be compared as follows:

1. **Utilizing a high percentage of the available exposure and mortality data:** Data from AQM 4, AQM 5, and the industrial AQMs were recorded for 48%, 50%, and 96% of all days in the study period. By comparison, the multiple-AQM models used pollutant concentration and mortality data for 69% and 98% of all days in the study period, corresponding to exclusion or inclusion of the industrial AQMs, respectively. The zonal model used pollutant concentration and mortality data for 100% of the days in the study period. Weekends and holidays have been ignored when calculating the total number of days in the study period. Thus the zonal model used all available data, but the models that included industrial AQMs also utilized a high percentage of the available data.
2. **Estimated RR due to PM_{10} concentration:** The RR estimates for the single-AQM series corresponding to AQM 4, AQM 5, and the industrial AQM average are 1.004, 1.003, and 1.003 respectively. The multiple-AQM model based on AQM 4 and AQM 5 estimates an RR of 1.004 while the model including the industrial AQM average estimates an RR of 1.003 for AQM 4, AQM 5, and the industrial AQM average. The multiple-AQM model with imputation of missing values estimates an RR of 1.004. None of the multiple AQM models, however, found sufficient evidence to support the hypothesis of a differential RR for each exposure series. This does not imply that we have convincing evidence that the RR's are identical, rather that our model does not have adequate statistical power to detect a difference. In fact, this is corroborated by the RR for AQM 4 being significant in all multiple-AQM models while the RR's for AQM 5 and the industrial AQM average are insignificant. This also explains the difference in the RR estimates for AQM 5 and the industrial AQM average between the single- and multiple-AQM models. The core (zonal) model finds a statistically significant RR estimate of 1.004.
3. **Interpretation of RR estimates:** The RRs from the single-AQM models measure the impact of changes in PM_{10} recorded by AQM 4, AQM 5, and the industrial AQM average on total all-cause mortality in Chennai. The corresponding coefficients in the multiple-AQM models have the same interpretation as those in the single-AQM models. Owing to the low correlation between AQMs, it can be expected that inclusion or exclusion of another AQM will not serve as a confounder in the multiple-AQM model. The RR from the zonal model measures the effect of increase in the average PM_{10} concentrations on zone-specific mortality and provides an aggregated estimate for the whole city. Given the differences in source strengths, types of sources, dispersion potential, and the limited footprints of the AQMs, it seems likely that the RRs from the zonal model best represent the underlying relationship between citywide variation in PM_{10} level and mortality (this is further corroborated by the smallest width of 95% CIs and degree of overdispersion as described later).
4. **Width of 95% CI for the RR due to PM_{10} concentration:** The single-AQM series corresponding to AQM 4, AQM 5, and the industrial AQMs report widths of 0.007, 0.009, and 0.008 respectively for the 95% CI of the RR. The 95% CI width is 0.004 for all multiple AQM models. The core (zonal) model 95% CI width is 0.005.
5. **Degree of overdispersion:** The single-AQM series corresponding to AQM 4, AQM 5, and the industrial AQM average report overdispersion parameters of 1.40, 1.54, and 1.56 respectively. The multiple-AQM models corresponding to exclusion or inclusion of the industrial AQMs have overdispersion parameters of 1.49 and 1.66 respectively. The multiple-AQM model with imputation has an overdispersion parameter of 1.62. The core (zonal) model has an overdispersion parameter of 1.25.
6. **Restrictiveness and plausibility of assumptions:** The various models compare as follows:
   - Single-AQM models: Assumes that changes in a single AQM can accurately represent changes in average pollutant concentrations across the city.
   - Multiple-AQM model based on observed data: There was no such restrictiveness, but the model formulation was not parsimonious and interpretation of the different β estimates can be difficult. Multicollinearity can arise due to the presence of a large number of indicator variables.
   - Multiple-AQM model based on imputed data: Requires the imputation model to have high predictive power. At the same time, the high predictive power of the model leads to a situation of high concurvity (Peng et al. 2006) and can lead to unidentifiable β estimates.
   - Zonal model: Assumes that the population is distributed within a zone in a homogeneous manner. This assumption could be verified with high resolution zonal data on population density that were not available for our study.

Based on all observations provided above, we believe that the zonal model makes the best use of the limited data that were available to the investigators. It utilizes all available data, affords the best possible assignment of exposure and is based on a relatively strong underlying causal relationship reflected by a short 95% CI and low overdispersion. A significant limitation of this approach has been the inability to address: (1) the impact of bias resulting from the high proportion of missing data at included AQMs, the nonuniform availability of data from all AQMs on any given day, and the assumption of a homogenous distribution of population within a zone; (2) the increased noise in the exposure series created by dropping grid squares with missing data from a zone's daily mean; and (3) the mobility of the population.

COMPARISON WITH PREVIOUS STUDIES IN ASIA

Chennai's annual averages of less than 80 µg/m^3 for PM_{10} (Central Pollution Control Board 2006) were among the lowest recorded in Indian cities but were higher than what had been reported from most other Asian studies. On the other hand, concentrations of SO_2 and NO_2 were consistently well below guideline values (WHO 1999, WHO 2005). Many Asian studies report effects estimates for a range of criteria air pollutants including PM_{10}, $PM_{2.5}$, SO_2, NO_2, carbon monoxide, and ozone for a range of health outcomes including all-cause and cause-specific mortality. However, the present study was confined to estimates for associations between PM_{10} and all-cause mortality, mostly because cause-of-death data were not available and methodological issues precluded inclusion of pollutant data for gases. Although the statistical methods used in this study were similar to what has been used in recent time-series studies in Asia, the United States, and Europe, they have gone beyond routine methods. In particular our methods have customized approaches for settings that encounter data challenges similar to what was experienced in this study.

The core model estimates a 0.44% (95% CI = 0.17 to 0.71) increase in mortality per 10-µg/m^3 increase in daily average PM_{10} concentrations. These effects estimates for PM_{10} are very similar to the summary estimates obtained from the Asian studies (summary effects estimates were 0.41% [95% CI = 0.25 to 0.56] and 0.49% [95% CI = 0.23 to 0.76] for fixed and random effects respectively from four studies; HEI 2004). This is also in the range of estimates found in the APHEA study of 29 European cities (Katsouyanni et al. 2001) of a 0.6% increase in risk of mortality and that of a 0.21% increase in risk of mortality reported in revised estimates of the NMMAPS study in the United States (Samet et al. 2000b). This is not surprising given the similarity in exposure and perhaps baseline demographic and health profiles. In Chennai the annual averages for PM_{10} are at the lower end of what is found in most large cities in India, at the upper end of APHEA and NMMAPS cities, and is comparable with those of many Asian cities. Chennai's health and socioeconomic status indicators are among the best for Indian cities and perhaps reflect additional grounds for similarity to other cities where such studies have been conducted (a formal comparative analysis is being performed).

The dose–response relationship we observed for daily average concentrations of PM_{10} is linear between 0 and 150 µg/m^3 (the 95% upper confidence limit for the 3-year average PM_{10} concentration). This is similar to what has been reported in previous literature over the same concentration range. Our study had too few observations above 150 µg/m^3 to allow reliable prediction of the dose–response curve beyond 150 µg/m^3.

Some Asian studies have reported sensitivity analyses that included age and sex stratifications, lag structures, and multiple-pollutant models. For example, from 11 estimates for PM_{10} and daily mortality in 4 cities, all estimates showed an increased risk of mortality for all ages, with the highest estimates for the age groups of ≥ 50 years or ≥ 65 years (HEI ISOC 2004). In our analyses the mean estimated relative risk was lower for females than for males, although this effect was not pronounced. Age did not appear to have any consistent effect on the results, although the risks for people 45 to 64 years and for people 65 years and older generally tended to be the highest and the most significant.

Few Asian studies have reported details of alternative lag structures used, but they found that 1-day lags consistently yielded the most significant estimates. Similarly, few studies report results of multiple-pollutant models. In the present study, inclusion of NO_2 data in a multiple-pollutant model slightly increased the coefficients for PM_{10}, but not significantly. Over 25% of NO_2 values and up to 50% of SO_2 values in all stations were clustered around the lower detection limit of 4 µg/m^3. Single-pollutant models for gases were therefore not feasible for reasons cited earlier. Similarly, inclusion of cumulative averages was not feasible because of the large number of missing observations.

UNCERTAINTIES OF ENVIRONMENTAL MEASUREMENTS

Environmental characterization analyses have not often been described in Asian studies. Lippmann and coworkers (2000) describe three main contributions to environmental measurement characterization errors with regard to site-to-site variability, person-to-AQM variability, and analytic and instrumentation errors.

Site-to-Site Variability

We examined site-to-site temporal variability in great detail because the lack of correlation between AQMs combined with a high percentage of missing days for each AQM posed the biggest challenges in creating representative exposure series that would result in the least amount of exposure misclassification. Interestingly, the CPCB protocol required that AQMs be placed at sites that are likely to differ in pollutant concentrations because of human activity (such as industrial, commercial, and residential). Also, to capture variations across city zones (and possibly to identify local control strategies), the sampling height was set much lower than the height recommended for true background samplers. While they were labeled as urban scale AQMs (with a range of 4–100 km), because they were at the lower end of that range they functioned more as neighborhood scale AQMs. Therefore, the variations found among AQMs are to be expected owing to the design of the monitoring network. The QA and QC information provided by TNPCB excludes the possibility that systematic measurement error across AQMs or days was responsible for the observed lack of correlation. The small footprints of the AQM stations combined with the differences in source strengths are thus likely to be the primary reason for the site-to-site variation.

Finally, because of the large percentage of missing values against the backdrop of considerable spatial heterogeneity, we refrained from averaging across AQMs or pooling available estimates into a single-exposure series. Our model development was concerned with addressing both the site-to-site variation and the missing data with consequent minimization of exposure misclassification.

Person-to-AQM Variation

The limitations of the data set available for the study and the nonavailability of detailed personal exposure data precluded application of formal methods to assess contributions from exposure measurement errors to the zonal model (such as differentially being able to address classical and Berkson errors as demonstrated in Zeger et al. 2000). Preliminary results from previous studies carried out by the same investigators in Chennai city (Balakrishnan and Parikh 2005) indicate that the observed spatial heterogeneity in PM_{10} concentrations recorded by the ambient AQM stations was mirrored in population exposure reconstructions (performed through multiple microenvironmental measurements and time–activity assessments). Results indicated that in general, industrial zone residents have greater exposures to PM_{10} because of considerably higher outdoor air pollution concentrations. The distribution of other determinants (e.g., environmental tobacco smoke and indoor solid fuel use) was not significantly different across zones. The bulk of the personal exposures was accounted for by *commuting* exposures that tended to be the highest in industrial zones. This spatial gradient of outdoor exposures across zones is also borne out in the pollutant concentration maps, such as for PM_{10} in Figure 17, which were generated using the annual averages assigned to each grid from nearest available AQMs for the study period (see Figure 16). The northernmost zones of the city are the closest to major industries (see Figure 1) and have the highest average pollutant concentrations when compared with southern residential areas. While in general it is well recognized that no ambient AQM would adequately capture personal exposures, because of the nature of the monitoring protocols followed in Chennai and other studies in India, the overwhelming site-to-site variability may have worked in favor of a smaller person-to-AQM variation in the zonal models. The 95% CI for the core (zonal) model is also the smallest, suggesting that this method leads to the most precise treatment of the data because it appears to have a lower degree of exposure misclassification.

Analytic and Instrumentation Errors

Lack of hourly data and sensitivity of gas measurement methods did not permit inclusion of gases in single- or multiple-pollutant models with the exception of one site for NO_2 (AQM 4). We were provided access to the complete set of QA and QC information related to sampling by TNPCB. There was no difference in terms of the procedures used among stations, and compliance with the standard operating procedures was nearly 100%. Thus, a high degree of precision in estimates collected from each site is expected. Since most sampling protocols used by TNPCB and CPCB are either historical methods used by the U.S. Environmental Protection Agency or are modified versions of current methods, we attempted to compare the accuracy of the measurements collected by the indigenous instruments. Most studies in Asia have used β-ray absorption or tapered element oscillating microbalance methods for measuring PM_{10} and chemiluminescence methods for measuring gases, while routine AQM stations in Chennai use gravimetric methods and wet chemical methods for PM_{10} and gases respectively.

We also collected data from a continuous AQM station that was colocated with the routine AQM station 3, and had data available for a year within the period used in this study. Pollutant concentrations recorded at the continuous AQM station, which used a β-gauge method for PM_{10}, were generally higher than what was recorded at AQM 3, and were modestly correlated (r = 0.76). That indicates residual concerns about data quality. The continuous AQM station used chemiluminescence methods for gases, but the available data on gaseous concentrations were too sparse to allow comparisons with the AQM 3 data. For administrative reasons, it was not feasible to further examine the contributions from differences in monitoring methods used by CPCB in comparison with other international protocols.

UNCERTAINTIES IN MORTALITY DATA

Mortality data were made available electronically, a rare practice in Indian cities (e.g., data were abstracted manually in Ludhiana, and parts of data were abstracted manually in Delhi). Data entry checks had already been performed by the Chennai Corporation. We further cross-checked 10% of all entries against hard copies. The entries were also subject to the HEI quality assurance audits and no problems were noted with respect to the fields of entry used for analyses in this study (Appendix A). Secondly, the same raw data (i.e., completed death registration forms) had been independently abstracted into another electronic database by NCI-Chennai. We merged the two databases and flagged all discrepancies for verification against the raw data. A total of 45 such entries were identified that required corrections in fields that would be used in analyses. A high degree of precision could thus be expected in the mortality data sets similar to that of the pollutant data sets. Also, because nearly 98% of all deaths occurring within the Chennai municipal boundary were recorded (according to information provided by the Vital Statistics department under the Directorate of Public Health), uncertainties resulting from unreported deaths are likely to be minimal.

Each zonal office of the Chennai Corporation recorded and registered all deaths occurring within its zone. The Chennai Corporation has a record of all deaths occurring from all zones within the Chennai municipal boundaries. Because the zonal offices did not consistently record the place of residence on the death registration forms, it was difficult to stratify mortality by residents versus nonresidents of Chennai. The lack of information on permanent addresses also precluded a full-blown geostatistical analysis. The NCI-Chennai data set is limited to adults who are Chennai residents. Comparisons of the two data sets indicate that nearly 80% of the deaths recorded are from Chennai residents. Given that Chennai is home to many major medical facilities, including a large number of tertiary care facilities, it is unlikely that a large percentage of Chennai resident deaths occur outside the city, although we did not attempt to estimate this.

Further, parts of the greater Chennai area that fall outside the Chennai Corporation limits, especially the northernmost corridors, experience some of the highest exposures, but we lacked the information required to assess whether a large proportion of these deaths occurred within the Chennai municipal boundaries or in large government facilities located immediately outside the boundary. We propose that future studies include deaths from all major municipal offices in the greater Chennai area to comprehensively account for most deaths within the greater Chennai area. The multiple-AQM and zonal models could be further refined with these additional data.

Lacking cause-of-death information was a major limitation of the mortality data set. This information would have allowed not only the exclusion of accidental deaths but also the examination of a range of cardiovascular and respiratory outcomes and their relative impacts on different age groups.

CONCLUSIONS

The study has been an exploratory attempt to apply a set of methods for time-series analyses that would most closely describe the association between concentrations of PM_{10} and all-cause mortality in Chennai city. This study was based on retrospective data collected for the purpose of routine monitoring. Abilities to inform health based investigations are limited considering the design, statistical power, and level of detail available in the data sets. While approaches developed in previous studies served as the basis for model development, specific data limitations required the development of refinements that allowed the use of routinely collected data to estimate the impacts of air pollution on mortality with a reasonable level of precision.

We were able to complete the data processing steps for the pollutant and mortality data as envisaged because of the cooperation of the local agencies, including access to their raw data. Although reliability was established for these particular data sets, such cooperation may not always be forthcoming in other settings. Also, considerable effort was required for tracking relevant information, cleaning up data sets, and data compilation. This may not be routinely feasible. There is thus a need to initiate steps to structure the data collection and recording procedures in ways that would make them more useful for environmental health professionals, both within the government and in academic settings.

Development of a representative exposure series occupied a central role in model development. Specifically, the issues of missing data and small footprints of AQMs are likely to be encountered in many other Indian cities as well as in other developing countries. Until such time when infrastructural investments allow the design of more sophisticated monitoring mechanisms, the methods developed in this study may allow data currently being collected to be used for baseline assessments in situations where similar exposure issues prevail. Informal discussions about the results of the present study, held with stakeholders in the State and Central Pollution Control Boards (i.e., TNPCB and CPCB), indicate that there is interest in using new methods to collect air pollution data in select locations to facilitate data collection for future environmental health studies.

Improved mortality data recording is recognized as being necessary for many public health programs. There are on-going efforts in Chennai to increase the completeness of the data sets, especially with respect to the cause-of-death fields. Although assessments of mortality impacts are useful, morbidity data provide a much higher resolution for a range of early biologic effects of air pollution. At present, morbidity data are poorly recorded in Chennai, and it is virtually impossible to access routinely collected morbidity data in electronic formats. This represents a serious challenge for the conduct of future follow-up studies. The same can be said about meteorologic data. Improved visibility data and electronic data sets that include hourly readings would substantially enhance data usability. With the increasing impetus on climate change issues, several efforts are likely to be initiated in India to better capture many meteorologic parameters. This may represent an opportunity to maximize the efficiency of utilizing routinely collected data in time-series analyses as well as in other environmental epidemiologic studies. However, risk communication channels will need considerable strengthening, before environmental health is explicitly included in the programs of these local agencies.

The effects estimates for PM_{10} are in the range of summary estimates obtained from the Asian, European, and U.S. studies. However, without additional information on source apportionment, emissions, and detailed information on cause-of-death, it is difficult to judge comparability between Chennai and the other cities, both with respect to exposures and outcomes. This argues for the development of such data sets locally, to guide future studies based on retrospective data. Also, prospective studies such as those using case–crossover designs may allow some of these uncertainties to be addressed.

In conclusion, while we have attempted to extract the maximum amount of information from the data sets available to us, future efforts with better ways to capture population exposure to pollutants are needed before we can develop a realistic picture of mortality associated with PM_{10} and other criteria air pollutants. We hope that the study methods will be useful for application in future time-series analyses in other Indian cities, and that this study, together with future studies, will catalyze policy changes and contribute to the improvement of air quality in India.

ACKNOWLEDGMENTS

We thank project consultants Dr. Sugata Sen Roy of the Department of Statistics, University of Calcutta, Calcutta, India, and Dr. Aditya Chatterjee, Department of Statistics, University of Burdwan, Burdwan, India, for providing overall technical oversight for the statistical analyses performed for this project. Without their keen involvement many methodological refinements used in the project, including the framework for the spatially disaggregated zonal model, would not have been feasible. Their committed support and technical contributions are acknowledged with gratitude.

We also thank project assistants Deepa Rajagopalan and Venkatesan Doss for their support in data collection and database maintenance. Their meticulous efforts in abstracting manual data and verifying QA/QC related information enable the project investigators to successfully comply with HEI's stringent data auditing requirements.

REFERENCES

American Thoracic Society. 2000. What constitutes an adverse health effect of air pollution? Am J Respir Crit Care Med 161:665–673.

Balakrishnan K, Sambandam S, Ramaswary P, Pitani RS. 2003. Comparative health risk assessment for air pollution in Chennai. In: Proceedings from EURUHEIS-SASHU Conference on Health and Environment. March 30–31, 2003. Östersund, Sweden.

Balakrishnan K, Parikh J. 2005. Economic valuation of health damage in North Chennai using a Comparative Risk Assessment Framework. Environmental Economics Working Paper Series, EH-2. Indira Gandhi Institute of Development Research, Mumbai, India. Available from *http://coe.mse.ac.in/eercrep/fullrep/ehe/EHE_FR_Kalpana_Balakrishnan.pdf.*

Census of India. 2001. Census of India: Metadata and brief highlights on slum population. Office of the Registrar General & Census Commissioner, New Delhi, India.

Central Pollution Control Board. 2006. National Ambient Air Quality monitoring series: NAAQMS/29/2006-07. Available at *www.cpcb.nic.in.* Last accessed December 14, 2010.

Central Pollution Control Board. 2009. National Ambient Air Quality Standards. Notification B-21906/20/90/PCI-L. Central Pollution Control Board, New Delhi, India. Available from *www.cpcb.nic.in/National_Ambient_Air_Quality_Standards.php.* Accessed 9/20/2010.

Cropper M, Simon NB, Alberini A, Arora S, Sharma PK. 1997. The health benefits of air pollution control in Delhi. Am J Agr Econ 79:1625–1629.

Diggle PJ, Ribeiro PJ. 2007. Model-based Geostatistics. Springer, New York, NY.

Dominici F, McDermott A, Hastie T. 2004. Improved semiparametric time series models of air pollution and mortality. J Am Stat Assoc 468:938–948.

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

Health Effects Institute. 2001. Airborne Particles and Health: HEI Epidemiologic Evidence. HEI Perspectives. Health Effects Institute, Cambridge, MA.

Health Effects Institute. 2002. Understanding the Health Effects of Components of the Particulate Matter Mix: Progress and Next Step. HEI Perspectives. Health Effects Institute, Boston, MA.

HEI International Scientific Oversight Committee. 2004. Health Effects of Outdoor Air Pollution in Developing Countries of Asia: A Literature Review. Special Report 15. Health Effects Institute, Boston, MA.

HEI Public Health and Air Pollution in Asia Program. 2010. Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Katsouyanni K, Samet JM, Anderson HR, Atkinson R, Le Tertre A, Medina S, Samoli E, Touloumi G, Burnett RT, Krewski D, Ramsay T, Dominici F, Peng RD, Schwartz J, Zanobetti A. 2009. Air Pollution and Health: A North American and European Approach (APHENA). Research Report 142. Health Effects Institute, Boston, MA.

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, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 12:521–531.

Lippmann M, Ito K, Nádas A, Burnett RT. 2000. Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations. Research Report 95. Health Effects Institute, Cambridge, MA.

Peng RD, Dominici F, Louis TA. 2006. Model choice in time series studies of air pollution and mortality (with discussion). J Royal Stat Ser A 169(2):179–203.

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

Ruppert D, Wand MP, Carroll RJ. 2003. Semiparametric Regression. Cambridge University Press, Cambridge, UK.

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

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

Samoli E, Peng R, Ramsay T, Pipikou M, Touloumi G, Dominici F, Burnett R, Cohen A, Krewski D, Samet J, Katsouyanni K. 2008. Acute effects of ambient particulate matter on mortality in Europe and North America: Results from the APHENA study. Environ Health Perspect 116: 1480–1486.

U.S. Environmental Protection Agency. 2004. Air Quality Criteria for Particulate matter, Volume II of II. EPA/600/P-99/002bF. Office of Research and Development, Research Triangle Park, NC.

Wood SN. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. J Am Stat Assoc 99:673–686.

World Bank. 1995. Annual World Bank Conference on Development Economics. (Bruno M, Pleskovic B, eds.). World Bank, Washington, DC.

World Health Organization. 1987. Air Quality Guidelines for Europe. European Series No 23. WHO Regional Publications, World Health Regional Office for Europe, Copenhagen, Denmark.

World Health Organization. 1999. Guidelines for Air Quality. WHO, Geneva, Switzerland.

World Health Organization. 2002. The World Health Report 2002: Reducing Risks, Promoting Healthy Life. WHO, Geneva, Switzerland.

World Health Organization. 2005. WHO Air Quality Guidelines Global Update 2005. Report on a Working Group Meeting, Bonn, Germany, 18–20 October. WHO Regional Office for Europe, Copenhagen, Denmark.

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

APPENDIX A. HEI Quality Assurance Statement*

The conduct of this study was subjected to periodic, independent audits by a team from Hoover Consultants. This team consisted of auditors with experience in toxicology, epidemiology and air quality data. The audits included in-process monitoring of study activities for conformance to the study protocols and examination of records and supporting data. The dates of each audit are listed below with the phase of the study examined.

December 5–6, 2005

Records from this study were obtained by the investigators from external groups and the audit did not extend beyond these records to the original data sources. The 5- and 10-month progress reports were included in the audit. This study was conducted in accordance with two protocols: an individual protocol that included the unique features of the Chennai project and a combined protocol for the coordinated time series analysis.

Dr. V.N. Rayudu of the TNPCB provided a tour of the air sampling monitoring sites. The air quality data audit focused on the National Ambient Air Quality Management Network data from October through December 2004. Meteorologic data were available for 2004 at two stations, collected by the study team by manual entry of records at the Chennai office of the India Meteorological Department. Electronic records were not checked against the original records because the office could not be reached due to local flooding. The data for visibility and relative humidity were examined for their potential as an indicator of daily particle concentrations.

Mortality data were obtained electronically from the Corporation of Chennai. The audit compared the data abstracted by the study team with the electronic file. A total of 1923 mortality data points selected from two zones were included in the audit. Variables included: date-of-death, date-of-registration, age, gender, place-of-death, native place, religion, occupation, whether medical attention was available, if the death was medically certified, cause-of-death, antecedent causes, if the death was pregnancy-related, and history of smoking and alcohol use. Recommendations were presented for more standardized data collection and cleaning procedures.

November 19, 2010

A draft titled *unedited version for QA* of the final study report (and associated appendices) was examined for internal consistency and conformance with the study protocols. No comments were noted and no report was provided to HEI.

A written report of the December 2005 audit was provided to the Director of Science of the Health Effects Institute who transmitted these findings to the Principal Investigator. These quality assurance audits demonstrated that the study was conducted by experienced professionals in accordance with both study protocols. The final report appears to be an accurate representation of the study.

B. Kristin Hoover
Hoover Consultants

* The QA Statement was edited for brevity.

APPENDICES AVAILABLE ON THE WEB

Appendices B–G contain supplemental material not included in the printed report. They are available on the HEI Web site *http://pubs.healtheffects.org*.

Appendix B. Common Protocol for Time-Series Studies of Daily Mortality in Indian Cities

Appendix C. Method for Determination of Respirable Suspended Particulate Matter (RSPM) in the Ambient Air (Gravimetric technique with high volume sampling)

Appendix D. Determination of Sulphur Dioxide in Ambient Air (Improved West and Gaeke Method)

Appendix E. Determination of Nitrogen Dioxide in Ambient Air (Jacob and Hochheiser Method)

Appendix F. Single and Multiple Monitor Models

Appendix G. Aerial View of AQM Locations

ABOUT THE AUTHORS

Kalpana Balakrishnan is professor and head of the Department of Environmental Health Engineering at Sri Ramachandra University, Chennai, India. She received her Ph.D. degree in Biophysics and subsequent post-doctoral training in Environmental Health Engineering from the Johns Hopkins University, Baltimore, Maryland, U.S.A. As the Director of the Center for Advanced Research on Environmental Health for the Indian Council of Medical Research, Government of India and the WHO-Collaborating center on Occupational Health at her university, she heads several national and international efforts in environmental and occupational health research specializing in the areas of indoor and outdoor air pollution, in addition to directing graduate programs. A recipient of several awards, she also serves as a member of national and international technical committees concerned with environmental health guidance. Her special interests are in developing sophisticated methods to address unique exposure challenges in developing countries.

Harry Caussy served as an environmental epidemiologist in the Department of Evidence for Information and Policy at the World Health Organization, South-East Asia Region, prior to his retirement in 2007. He holds a Ph.D. degree in Epidemiology from McMaster University, Ontario, Canada. He has been involved in several regional global efforts in the area of infectious diseases, occupational and environmental health, and health policy.

Bhaswati Ganguli is associate professor in the Department of Statistics, Calcutta University, Kolkatta, India. She received her Sc.D. degree in Biostatistics from Harvard University, Cambridge, Massachusetts, U.S.A. She serves as a lead investigator on several extra-mural grants in the areas of environmental epidemiology, exposure modeling, and spatial statistics.

Santu Ghosh is lecturer in the Department of Environmental Health Engineering at Sri Ramachandra University, Chennai, India. He received his Master's degree in Statistics from Calcutta University, Kolkatta, India. He is involved in several multicentric international data analysis efforts in the areas of outdoor and indoor air pollution, water and sanitation, and occupational risk assessment.

S. Sankar is assistant professor in the Department of Environmental Health Engineering at Sri Ramachandra University, Chennai, India. He received his Master's degree in Environmental Toxicology from Madras University, Chennai, India. He is involved in numerous national and international efforts in the areas of exposure assessment for indoor and outdoor air pollution, chemical and physical hazards in the work environment, and health risk assessment.

Vijayalakshmi Thanasekaraan is professor in the Department of Chest Medicine, Sri Ramachandra University, Chennai. She holds an M.D. degree in Chest and General Medicine from the Madras Medical College, Chennai, India. She is a leading clinician and researcher in the areas of tuberculosis and related chest diseases, asthma and allergic diseases, and air-pollution-related chest diseases.

ABBREVIATIONS AND OTHER TERMS

APHEA	Air Pollution and Health: A European Approach
APHENA	Air Pollution and Health: A European and North American Approach
AQM	air quality monitor
CI	confidence interval
CMDA	Chennai Metropolitan Development Authority
CPCB	Central Pollution Control Board
df	degrees of freedom
ER	excess risk
GAM	generalized additive model
GCV	generalized cross validation criterion
ISOC	International Scientific Oversight Committee
NAAQS	National Ambient Air Quality Standards (India)
NCI-Chennai	National Cancer Institute in Chennai
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
NO_2	nitrogen dioxide
PAPA	Public Health and Air Pollution in Asia
PM	particulate matter
$PM_{2.5}$	PM with an aerodynamic diameter $\leq$ 2.5 µm
PM_{10}	PM with an aerodynamic diameter $\leq$ 10 µm
QA	quality assurance
QC	quality control
RR	relative risk
SO_2	sulfur dioxide
SRMC	Sri Ramachandra Medical College
TNPCB	Tamil Nadu Pollution Control Board
TSP	total suspended particulate matter
WHO	World Health Organization

Part 2
Time-Series Study on Air Pollution and Mortality in Delhi

Uma Rajarathnam, Meena Sehgal, Subramanya Nairy,
R.C. Patnayak, Sunli Kumar Chhabra, Kilnani, K.V. Santhosh Ragavan

INVESTIGATORS' REPORT

Time-Series Study on Air Pollution and Mortality in Delhi

Uma Rajarathnam, Meena Sehgal, Subramanya Nairy, R.C. Patnayak, Sunli Kumar Chhabra, Kilnani, K.V. Santhosh Ragavan

The Energy and Resources Institute, New Delhi, India (U.R., M.S., S.N., K.V.S.R.); Municipal Corporation of Delhi, Delhi, India (R.C.P.); Vallabhbhai Patel Chest Institute, University of Delhi, Delhi, India (S.K.C.); and All-India Institute of Medical Sciences, New Delhi, India (K.)

ABSTRACT

INTRODUCTION

Air pollution concentrations in most of the megacities in India exceed the air quality guidelines recommended by the World Health Organization and may adversely affect human health in these cities. Particulate matter (PM*) is the pollutant of concern in many Indian cities, particularly in the capital city of Delhi. In recent years, several actions have been taken to address the growing air pollution problem in Delhi and other Indian cities; however, few studies have been designed to assess the health effects of air pollution in Indian cities. To bridge the gap in scientific knowledge and add evidence to the ongoing studies in other Asian cities, a retrospective time-series study on air pollution and mortality in Delhi was initiated under the HEI Public Health and Air Pollution in Asia (PAPA) program.

APPROACH

The study used retrospective time-series data of air quality and of naturally-occurring deaths recorded in Delhi to identify changes in the daily all-natural-cause mortality rate that could be attributed to changes in air quality. The 3-year study period included the years 2002 through 2004. The methodology involved: (1) collecting data on ambient air quality for major pollutants from all monitoring stations in Delhi; (2) collecting meteorologic data (temperature, humidity, and visibility); (3) collecting daily mortality records from the Registrar of Births and Deaths; (4) statistically analyzing the data using the common protocol for Indian PAPA studies, which included city-specific modifications.

RESULTS AND IMPLICATIONS

The study findings showed that increased concentrations of PM with an aerodynamic diameter ≤ 10 µg/m^3 (PM_{10}) and of nitrogen dioxide (NO_2) were associated with increased all-natural-cause mortality. It was found that every 10-µg/m^3 change in PM_{10} was associated with only a 0.15% increase in total all-natural-cause mortality. When NO_2 alone was considered in the model, daily all-natural-cause mortality increased 0.84% for every 10-µg/m^3 increase in NO_2 concentration.

No significant effect was observed for changes in sulfur dioxide (SO_2) concentrations. The study provides insight into the link between air pollution and mortality in local populations and contributes information to the existing body of knowledge.

INTRODUCTION

The effects of air pollution on health have been established by a number of epidemiologic studies in various parts of the world. The main air pollutants, SO_2, NO_2, and suspended particulate matter (SPM) have been associated with both acute and chronic respiratory morbidity and mortality (Bascom et al. 1996; Ackermann-Liebrich and Rapp 1999; Samet et al. 2000). Epidemiologic studies conducted in North American cities showed that a 10-µg/m^3 increase in PM_{10} was associated with an increase in mortality of

This Investigators' Report is one part of Health Effects Institute Research Report 157, which also includes a Critique by the Health Review Committee and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr. Uma Rajarathnam, Enzen Global Solutions, 90, Madiwala, Hosur Road, Bangalore 560 068, India; e-mail: *uma.r@enzenglobal.com*.

The PAPA Program was initiated by the Health Effects Institute in part to support the Clean Air Initiative for Asian Cities (CAI-Asia), a partnership of the Asian Development Bank and the World Bank to inform regional decisions about improving air quality in Asia. Additional funding was obtained from the William and Flora Hewlett Foundation. The contents of this document 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.

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

about 0.3% to 0.5% (Samet et al. 2000). Studies from Europe support similar observations (Katsouyanni et al. 1997). These results play an important role in establishing PM air quality standards in the European Union and the United States. However, only a few such studies have been done in Asia, and they were concentrated in developed Asian cities. Because of the limitations of epidemiologic studies, many Asian countries have used air quality guidelines recommended by the World Health Organization. These guidelines are largely based on studies conducted in western countries but are extrapolated to developing countries to establish their air quality standards. Asia differs from western countries, both in the nature of air pollution and in population characteristics. Some of the highest concentrations of outdoor air pollution in the world were recorded in Asian cities.

Many cities in India are considered to be among the polluted megacities of the world, which are inhabited by millions of people. In recent years, several actions have been taken, mainly in the capital city of Delhi, to address the growing air pollution problem in Indian cities that have suffered for decades from declining air quality. However, few studies have assessed the health effects of air pollution in Indian cities. To bridge the gap in scientific knowledge and add information to the ongoing studies in other Asian cities, HEI initiated pollution and mortality studies in Chennai, Delhi, and Ludhiana, under the PAPA project. This report presents the methodology and results of this research that explores the relationship between air pollution concentrations and daily all-natural-cause mortality in Delhi, India between 2002 and 2004.

OBJECTIVES

Main objectives of the study were:

- assess the time-series data on air quality parameters and all-natural-cause mortality to study the relationship between air pollution (mainly PM_{10}) and all-natural-cause mortality in Delhi; and
- assess the daily change in all-natural-cause mortality in relation to changes in air quality after controlling for potential confounding variables.

STUDY DESIGN

In this study, retrospective data on air pollution and meteorologic parameters for the city of Delhi were collected and analyzed by following a common protocol with suitable modifications for the city of Delhi (Appendix B; available on the Web). The main components of the proposed study were:

- collecting 24-hour average concentrations of criteria air pollutants for the study period (2002–2004) across the 10 locations being monitored by the Central Pollution Control Board (CPCB), the National Environmental Engineering Research Institute (NEERI), and The Energy and Resources Institute (TERI);
- collecting information about the air quality monitor (AQM) sites to evaluate the site-selection criteria and to determine the suitable sampling locations that could represent the population exposure in the city of Delhi;
- reviewing the monitoring protocols, and the quality assurance (QA) and quality control (QC) procedures for data harmonization;
- collecting meteorologic data;
- reviewing the common protocol in consultation with other Indian investigators to harmonize the methodology for Indian cities;
- collecting mortality records from the Registrar of Births and Deaths, Municipal Corporation of Delhi (MCD) and from the New Delhi Municipal Council (NDMC), then extracting records on natural deaths based on the information available on the primary cause of death;
- statistically analyzing the data to determine if there is a relationship between changes in daily all-natural-cause mortality and daily air pollutant concentrations after controlling for weather parameters.

DESCRIPTION OF STUDY AREA

The study was conducted in Delhi, the national capital of India, located in northern India. It shares borders with the states of Haryana and Uttar Pradesh. Delhi has a total area of 1483 km^2 and is 216 m above sea level. Of this total area, 591.91 km^2 is rural and 891.09 km^2 is urban. According to the 2001 census, Delhi had a population of 13.78 million with a growth rate of 3.81%. The corresponding population density was 9294 persons per km^2, with a ratio of 821 women per 1000 men, and a literacy rate of 81.82%.

Delhi was considered to be a single district for the 1991 population census. In 1996, the Government of National Capital Territory Delhi created 9 districts. So the 2001 population census for Delhi was based on these 9 districts. A demographic profile of Delhi is presented in Table 1.

Three different agencies are responsible for recording information on deaths occurring in Delhi. The MCD is the largest municipal agency among the three, covering 1397.3 km^2. The NDMC covers 42.7 km^2, and the Delhi Cantonment Board covers 43 km^2. Of the total population, the MCD comprises 97%; the NDMC and Delhi Cantonment Board comprise 2% and 1% respectively.

Table 1. 2001 Demographic Profile of Delhi[a]

Districts	Population: Rural	Population: Urban	Population: Total	Females/1000 Males: Rural	Females/1000 Males: Urban	Females/1000 Males: Total
Northwest	263,487	2,583,908	2,847,395	806	822	820
North	46,586	733,202	779,788	812	827	826
Northeast	141,528	1,622,184	1,763,712	850	851	851
East	18,123	1,430,647	1,448,770	809	845	845
West	85,304	2,034,337	2,119,641	763	833	830
Southwest	223,688	1,525,804	1,749,492	822	778	783
South	184,499	2,073,868	2,258,367	776	799	797
Central	—	644,005	644,005	—	843	843
New Delhi	—	171,806	171,806	—	791	791
Total	963,215	12,819,761	13,782,976	806	822	821

[a] Source: Government of National Capital Territory of Delhi 2002.

Delhi has a semiarid climate with large temperature differences between the summer (May–June) and the winter (November–December). The temperature ranges from −0.6°C (30.9°F) during the winter to 47°C (117°F) during the summer. The annual mean temperature is 25°C (77°F). The average annual rainfall is approximately 714 mm (28.1 inches), most of which occurs during the monsoon season (July–August).

The high concentration of air pollution in the city is a major environmental problem. Major sources of air pollution include vehicular emissions, industrial pollution, burning organic material (e.g., wood, crop residues, dung cake), and background sources such as dust that is blown in from the desert during dry summer days. Urbanization and economic growth in the last two decades has resulted in an increased number of vehicles on the roads. Most vehicles in the city run on gasoline and most of these are two-wheeled vehicles with highly polluting 2-stroke engines.

During this study, the PM_{10} concentration exceeded the national ambient air quality standard (NAAQS) for most of the monitored periods. The annual mean PM_{10} concentration reached 170 μg/m^3 in 2003, which is almost three times higher than the NAAQS of 60 μg/m^3 for residential areas (Central Pollution Control Board, 2003).

Available air quality data suggest that PM pollution is of concern in Delhi. However, the characterization and sources of PM pollution have not been studied in detail. CPCB has performed limited monitoring of PM with an aerodynamic diameter ≤ 2.5 μg/m^3 ($PM_{2.5}$) at different locations in Delhi (Figure 1). The results indicate that the $PM_{2.5}/PM_{10}$ ratio varies from 48% during the summer to 59% during the winter (Badwar and Patak 2007). Few PM_{10} samples were analyzed for organic carbon and elemental carbon. In samples from two residential areas (AQMs 5 and 10), the ratios of organic carbon to elemental carbon were greater than 2, indicating that nonvehicular sources such as burning organic material contribute to PM_{10} concentrations. At all AQM sites, lower carbonaceous fractions were observed in the summer months compared with the winter months. This suggests that natural dust from dry desert winds contributes to the high PM concentrations in the summer.

METHODS

DATA COLLECTION

Air Quality Data

Ambient air quality is being monitored for the criteria air pollutants PM_{10}, SO_2, and NO_2 at 10 AQM locations (see Figure 1). CPCB directly operates six AQM sites under the National Air Quality Monitoring Program (NAMP); NEERI operates three AQM sites on behalf of the CPCB as part of the network under the NAMP, and TERI operates one AQM site. Table 2 lists the monitoring agency and reporting organization for each AQM site.

All AQM sites are located in urban areas and typically cover major land-use categories: residential, industrial, and commercial. AQMs 1, 2, 3, 5, 6, 7, and 10 are located in residential areas. AQMs 4, 8, and 9 are located in industrial areas. AQM 1 represents a business and cultural complex with many offices and restaurants. It is located on Lodi road along with residential colonies, schools, government

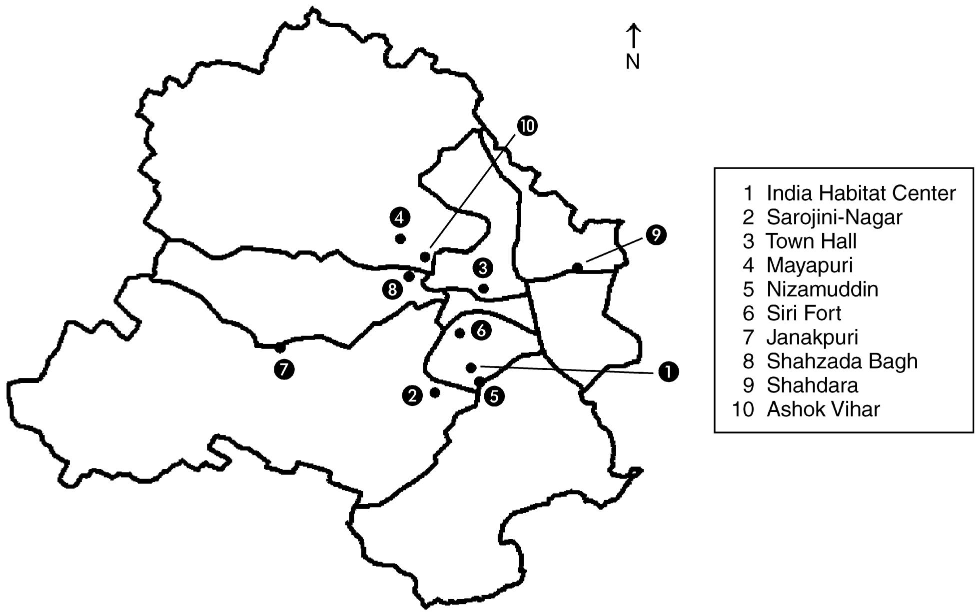


Figure 1. Map of Delhi showing AQM sites. Source: Survey of India.

offices, and commercial sites. In Delhi, mixed land-use patterns are common (i.e., commercial activities and small-scale industry operations prevail in residential areas). Commercial activities are especially predominant in the area covered by AQM 3. The existing AQM stations are not evenly distributed throughout the city. AQMs 1, 2, 5, and 6 are located within 10 km of each other. Population density and economic status also varied among AQM sites. For example, AQM 3 is in a densely populated location. Considering the geographic and demographic variations, the pollution exposure of this large population could not be represented by any single AQM.

At all 10 monitoring sites, the sampling equipment is placed on a flat surface that is elevated 3 to 20 meters above the ground. According to the common protocol (Appendix B; available on the Web), they are supposed to be located 5 meters upwind from the building exhaust and 2 meters from the walls. Institutes and organizations involved in the air quality monitoring program in Delhi and other Indian cities follow common measurement techniques (shown in Table 3) recommended by the Bureau of Indian Standards or the CPCB. Standard operating procedures for measuring PM_{10}, SO_2, and NO_2 are given in Appendices B–D (available on the Web). In all these

Table 2. AQM Sites in Delhi

AQM Site	Location	Type	Monitoring Agency	Reporting Agency
1 India Habitat Center	South Delhi	Residential	TERI	TERI
2 Sarojini-Nagar	South Delhi	Residential	NEERI	CPCB
3 Town Hall	Central North Delhi	Mixed	NEERI	CPCB
4 Mayapuri	West Delhi	Industrial	NEERI	CPCB
5 Nizamuddin	Southeast	Residential	CPCB	CPCB
6 Siri Fort	South	Residential	CPCB	CPCB
7 Janakpuri	Southwest	Residential	CPCB	CPCB
8 Shahzada Bagh	Northwest	Industrial	CPCB	CPCB
9 Shahdara	Northeast	Industrial	CPCB	CPCB
10 Ashok Vihar	North-northwest	Residential	CPCB	CPCB

Table 3. Air Quality Monitoring Techniques[a]

	Pollutants			
	SPM	PM_{10}	SO_2	NO_2
Equipment	HVS	RDS	Gaseous sampling unit attached to HVS/RDS	Gaseous sampling unit attached to HVS/RDS
Flow measuring device	Pressure drop across orifice in the hopper	Pressure drop across orifice in the hopper	Rotameters/orifice nanometers	Rotameters/orifice nanometers
Flow rate	0.8–1.3 m^3/min	0.8–1.3 m^3/min	1 L/min	1 L/min
Sampling period (over 24 hours)	8 hours × 3	8 hours × 3	4 hours × 6	4 hours × 6
Sampling frequency	Twice per week	Twice per week	Twice per week	Twice per week
Analytical method[b]	Gravimetric	Gravimetric	West & Gaeke	Jacobs & Hochheiser
Minimum detection limit	1 μg/m^3	1 μg/m^3	0.04 μg/mL	0.03 μg/mL
Minimum reporting value	10 μg/m^3	10 μg/m^3	6 μg/m^3	3 μg/m^3
Maximum absorption wavelength	NA	NA	560 nm	550 nm

[a] HVS indicates high-volume sampler; RDS indicates respirable dust sampler; NA indicates not applicable.
[b] Described in Appendices C, D, and E (available on the Web).

locations, AQM readings are taken at three 8-hour intervals over a 24-hour period, and the values are averaged and reported as 24-hour averages. Based on the field visits and information collected from the monitoring agencies, the site selection criteria have been evaluated. Details of AQM sites in Delhi are given in Appendix H (available on the Web).

The monitoring frequency varies among AQM stations. At AQM 1, air quality is monitored daily. Monitoring at AQMs 2–10 follows the NAMP criteria of 108 monitoring days per year. However, the days of the week for which monitoring is to be done are not fixed. The investigators noticed that monitoring at these nine locations was performed on weekdays with no particular frequency. Therefore, air quality data on weekends and holidays from 9 of the 10 AQM sites were not available for the study.

Sampling and Measurement Techniques Sampling PM_{10} involved collecting PM_{10} on preweighed glass fiber paper placed in a high-volume air sampler with a cyclone-based separator for separating large particles. The PM_{10} was measured using gravimetric analysis. Wet chemical methods were followed for analyses of the gaseous pollutants (SO_2 and NO_2). These gaseous pollutants were absorbed in suitable media and then analyzed colorimetrically.

Air Quality Data Collection and AQM Site Visits Daily average air quality data for SPM, PM_{10}, SO_2, and NO_2 from the 10 AQM sites were collected from CPCB and TERI in an electronic format. Sample data recording forms and information about data handling procedure were also collected. The CPCB follows QA and QC procedures to validate the data and remove outliers.

QA and QC The CPCB collects data from the network of AQM sites in Delhi and ensures quality through analytical procedures. In 1997, the CPCB established a calibration laboratory with assistance from the German Technical Cooperation Agency. It was used to ensure uniformity in the analytical procedure and the quality of data. The laboratory was also used to train State Board officials in QC and QA techniques.

The program employed a *static injection system*, which carried out a *ring test* for gaseous pollutants such as SO_2 and NO_2. In 1999, four rounds of interlaboratory exercises were undertaken involving different state pollution control boards and pollution control committees, including the CPCB.

CPCB Initiatives to Ensure Data Quality Recognizing the need to improve the quality of data, the CPCB has initiated several measures. These are:

1. periodic inspection of the AQM site and laboratory;
2. organizing a training program on ambient air quality monitoring for field and laboratory staff of various monitoring agencies;
3. analytical QC exercises using a ring test facility given regularly every year at CPCB, in which various monitoring agencies, including TERI and NEERI, have participated;
4. review meetings held at CPCB Zonal Offices to discuss data problems and deficiencies with the monitoring agencies;
5. software training program for data management was conducted for monitoring agencies;
6. days with a monitoring period less than 16 hours are not considered when calculating the daily average; and
7. systematic data entry checks for data validation.

Data Processing The CPCB follows a standard QA procedure for validating air quality data. We performed data entry crosschecking and descriptive analysis to check for data entry errors, completeness, and outliers. We then brought any queries or doubts about the CPCB air quality data to the attention of the CPCB officials for their verification.

Limitations of Air Quality Data Existing AQM locations are all in urban areas; background air pollution data from rural areas were not available.

Air quality data for the nine stations operated under NAMP were available for fewer than 100 days per year. The criteria for the exact days of monitoring in these locations were not fixed. Monitoring was not done on weekends or on holidays.

The wet chemical method used for gaseous pollutant sampling is obsolete. With this method, the absorption efficiency of gases varies with temperature and sampling flow rate.

Meteorologic Data

Meteorologic data of the Safdarjung station in New Delhi were collected from the India Meteorological Department. The meteorologic parameters for collecting these data are:

- daily maximum and minimum temperature (in °C);
- daily total rainfall (in mm);
- daily minimum and maximum relative humidity (in %); and
- visibility — eight observations per day at 3-hour intervals (in meters).

The printed data provided by the India Meteorological Department were entered onto an electronic spreadsheet. An independent research assistant verified the accuracy of the data entries.

Mortality Data

In Delhi three different administrative agencies (i.e., MCD, NDMC, and Delhi Cantonment) are responsible for recording information on deaths occurring in Delhi. About 73% of total deaths are registered by the MCD office, about 25% by the NDMC office, and less than 2% by the Delhi Cantonment. The Delhi Cantonment area is under the jurisdiction of the Indian Army; mortality records from this area are not available for research studies.

As per the Delhi Registration of Births & Deaths Rules, 1999 (effective January 1, 2000), all births, still births, and deaths must be reported within 21 days of occurrence to the registrar or subregistrar of the local area. For deaths occurring in institutions (i.e., deaths in hospitals, health centers, maternity homes, nursing homes, jails, hostels, or similar institutions), the person in charge of the institution is responsible for reporting deaths to the registrar in the respective area. Information from police records is also reported to the NDMC and MCD. If a death occurs at home, the head of the household (the head being the person so recognized by the household) is responsible for reporting the death. If the head is not present at the time of death, responsibility falls first to the nearest relative of the head, and then to the oldest adult present in the house. About 55% of reported deaths are reported by institutions; about 45% are reported by household members.

At the time of this study, the medical certification of cause-of-death in Delhi was confined to the deaths occurring in the hospitals, nursing homes, and other private institutions. For various reasons, cause-of-death was not reported for the domiciliary deaths.

Information on deaths in the NDMC and MCD jurisdictions is recorded on mortality report forms and submitted to the registrar of their respective areas. The form has a legal section containing personal information and a statistical section with information such as age, sex, place-of-death, and cause-of-death. Mortality reports from the NDMC were collected and entered in the VITAL database (created in Visual Fox Pro software), and were maintained by the NDMC. After data entry, records were transferred to an electronic spreadsheet and brought to TERI.

MCD mortality data for the first part of the study period (years 2002 and 2003) were obtained in an electronic format. In 2004, MCD introduced an online mortality registration system, in which all hospitals register mortality data directly to an online database. Data for domiciliary deaths were entered into the online system through the respective MCD local centers. MCD mortality data for the year 2004 were downloaded from their website (*www.mcdonline.gov.in/*). The data have been cross-checked for duplicate entries, missing information, and misclassification. Details of mortality data collection and cleaning up data sets are presented in Appendix I (available on the Web).

Limitations of Mortality Data Births and deaths are registered at the place of the event rather than at the place of residence. Delhi, being a capital city, has a large number of major hospitals that provide medical care to patients from neighboring cities. For example, the population of the NDMC area is only about 2% of the total population in Delhi, but approximately 25% of all deaths recorded in Delhi are from the NDMC. This makes it difficult to measure the true population exposure.

- Despite several initiatives, mortality cases are underreported, especially for deaths of infants, children, and women.
- Medical certification of cause-of-death in Delhi is confined to institutional deaths only.

DATA ANALYSIS

Descriptive Analyses

Air quality data collected from the nine AQM locations operated under the NAMP by CPCB and the AQM 1 data obtained from TERI were further analyzed for:

- summary statistics of pollutants;
- correlations of individual pollutants from various AQM sites and correlations between various pollutants; and
- time-series analysis of pollutants.

Similarly, descriptive and trend analyses were done for weather parameters and mortality data.

Core Model for Mortality

We have followed the common protocol to develop the core model for analysis (Appendix B; available on the Web). Quasi-Poisson regressions were used to develop models to study the effect of air pollutants on all-natural-cause mortality, the health outcome. Models were developed by adopting a generalized additive model (GAM) with penalized spline smoothers in R (from the mgcv package, version 1.3-24).

The air quality data obtained from CPCB and TERI for the 10 AQM sites were considered for developing suitable exposure series. As discussed earlier, one of the major limitations of the air quality data set was the large amount of missing data. Among the 10 AQM stations, only AQM 1 met the data completeness criteria (i.e., data available for more than 350 days per year). Other AQM stations recorded data for fewer than 100 days per year. It was also noted that AQM 1 values were higher compared with those of other AQM stations. This would cause simple averages to be unduly influenced by AQM 1 values. To address these limitations, we used the centering approach described by Wong and colleagues (2001). Outliers for each AQM location were removed. Nonmissing daily concentration data were first centered for each AQM station (i.e., individual daily concentrations X_{ij} were subtracted by an annual station mean X_i for each day j). The centered data from all stations were then combined and added to the annual mean of all stations X to form X'_{ij} as shown in the following formula:

$$X'_{ij} = (X_{ij} - X_i + X) \tag{1}$$

The daily exposure concentration of individual pollutants was computed by taking the mean centered individual pollutant X'_{ij} over all stations.

Mortality data collected from NDMC and MCD were aggregated to get records of all-cause mortality in Delhi. This accounted for about 98% of all-cause mortality that occurred in Delhi — mortality records from the Cantonment area were not available. Due to the limitations associated with cause-of-death coding, analysis was limited to deaths from all natural causes. Ambiguity in non-natural death reporting was considered to be minimal as the majority of such deaths (e.g., from accidents or fire) were recorded by institutions or were registered in police records. The core model accounted for time variation and for the weather parameters temperature (*temp*) and relative humidity (*rh*) as follows:

$$\log(E[mortality]) = \alpha + \beta(pollutant) + s(time) + s(temp) + s(rh) + \gamma \times DOW \tag{2}$$

where pollutant refers to daily average ambient air pollutant level, s denotes the smooth function of the covariate (*time*, *temp*, *rh*), and γ is the vector of the regression coefficient associated with indicator variables for day-of-week (*DOW*).

Both penalized and natural spline models have been examined in this study. The penalized spline model is a flexible, nonparametric approach using cubic splines and a term that penalizes the curvature of the smoothing function (Wood 2000). The natural spline model is a parametric approach that fits piece-wise polynomial functions joined at knots, which typically are evenly placed throughout the distribution of the variable. The function is constrained to be continuous at each knot (Ruppert et al. 2003). The number of knots used determines the overall smoothness of the data.

We conducted preliminary analyses with different degrees of freedom (*df*) (3–15 *df*/year) for the time variable. For temperature and relative humidity, we used 3 *df*/year. The criteria for choosing a core model were: minimum partial autocorrelation function (PACF) of the residuals, low generalized cross validation score, high deviance explained, low-scale estimate, and statistical significance of the estimate. Overdispersion in the core model was further adjusted by reducing the PACF value to less than one using the autoregression function. After stabilizing the core model with autoregression, pollutants (PM_{10}, NO_2, SO_2) were introduced separately in the model to study the association between individual pollutants and daily all-natural-cause mortality. Technical details of the best core model are given in Appendix J (available on the Web). Pollutant effects on lags of 0 to 3 days and the cumulative average of lag 0–1 day were considered in the data analysis. We also considered two-pollutant models, using PM_{10} and one of the gaseous pollutants (either SO_2 or NO_2) and a multi-pollutant model using PM_{10}, SO_2, and NO_2 together.

Sensitivity Analyses

Detailed analyses were conducted to investigate sensitivity of the model with:

1. various degrees of freedom using natural spline smoothers;
2. 3–15 *df*/year for the time variable;
3. different exposure series using pollutant values of individual AQM locations, the average of all AQM stations excluding AQM 1, the average of all 10 AQM stations without centering; and
4. differing lag days to control for the confounding effect of temperature. Five sets of temperature lag days (i.e., lag 1, lag 2, lag 3, the cumulative average of lag 0–7, and the cumulative average of lag 8–14) were used with pollutant exposure for the current day in the sensitivity analyses.

Final results are expressed as the percentage change in daily all-natural-cause mortality per 10-μg/m^3 increase in pollutant concentration. The percentage change per 10-μg/m^3 increase in pollutant concentration was estimated by multiplying the β-coefficient from the regression model by 1000.

For each pollutant concentration in the core model, a concentration–response curve was developed by introducing mortality as a smooth function with 3 *df*/year.

Table 4. Summary Statistics for Air Quality Data from 10 AQM Sites in Delhi[a]

Pollutant/ Site	Min (μg/m^3)	Max (μg/m^3)	Mean (μg/m^3)	Median (μg/m^3)	SD (μg/m^3)	Valid *n*
PM_{10}						
AQM 1	20	1566	266	238	166	1094
AQM 2	33	582	161	136	103	282
AQM 3	9	626	182	156	120	287
AQM 4	31	689	213	185	133	176
AQM 5	29	416	131	123	58	229
AQM 6	20	382	124	113	66	205
AQM 7	29	402	144	132	69	201
AQM 8	35	441	156	140	69	173
AQM 9	18	371	140	129	65	207
AQM 10	16	482	138	122	77	189
Average	30	1015	221	204	116	1096
NO_2						
AQM 1	7	159	58	56	19	1094
AQM 2	10	130	47	44	21	286
AQM 3	16	158	57	53	27	285
AQM 4	11	129	51	47	24	176
AQM 5	21	61	43	43	7	224
AQM 6	9	54	32	31	8	223
AQM 7	12	61	41	42	8	196
AQM 8	17	102	40	38	11	211
AQM 9	15	65	36	35	10	208
AQM 10	10	60	33	32	11	199
Average	7	159	50	48	15	1094
SO_2						
AQM 1	6	30	8	7	3	1094
AQM 2	4	16	7	6	3	287
AQM 3	4	55	11	9	7	287
AQM 4	4	33	12	11	7	176
AQM 5	5	25	12	12	3	224
AQM 6	4	16	9	9	3	224
AQM 7	4	22	12	11	3	196
AQM 8	4	23	9	9	3	210
AQM 9	5	57	12	11	6	208
AQM 10	4	34	8	7	4	200
Average	4	26	9	9	3	1094

[a] Min indicates minimum concentration; Max indicates maximum concentration; SD indicates standard deviation; Valid *n* indicates the number of data points that met the QA/QC criteria.

RESULTS

AIR QUALITY DATA

Table 4 shows the summary statistics of PM_{10}, NO_2 and SO_2 parameters recorded at the AQM sites during 2002 to 2004. Among these, only AQM 1 had data for more than 350 days/year. The other 9 AQM stations had data for fewer than 100 days/year (less than 30%).

As shown in Table 4, mean PM_{10} concentration recorded at various AQM stations ranged from 124 to 266 μg/m^3. Figure 2 shows a box plot of PM_{10} concentration at the 10 AQM locations. PM_{10} concentrations were highest at AQM 1.

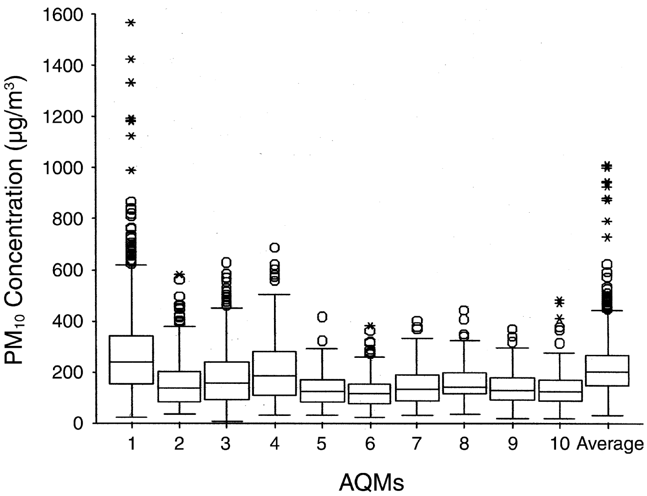


Figure 2. Box plot of PM_{10} concentration at the 10 AQM sites. The bottom and top of the box show the 25th and 75th percentiles. The middle line of the box represents the median. The ends of the I-bar represent the minimum and maximum PM_{10} concentrations. Small circles represent outliers and * represents extreme values.

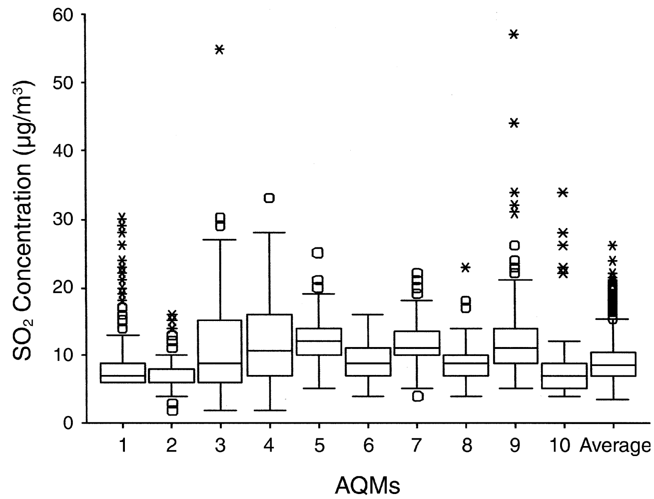


Figure 4. Box plot of SO_2 concentration at the 10 AQM sites. The bottom and top of the box show the 25th and 75th percentiles. The middle line of the box represents the median. The ends of the I-bar represent the minimum and maximum PM_{10} concentrations. Small circles represent outliers and * represents extreme values.

Mean NO_2 concentration recorded at various AQM stations ranged from 32 to 58 µg/m^3. Annual average NO_2 concentrations at AQMs 1 and 3 were close to the NAAQS of 60 µg/m^3 for residential areas. As shown in Figure 3, median NO_2 concentrations were high at AQMs 1 and 3.

The mean SO_2 concentration of various AQM stations ranged from 7 to 12 µg/m^3 (see Table 4). Median SO_2 concentrations at the 10 AQM locations were within 20 µg/m^3 of each other with very little variation among locations (Table 4, Figure 4).

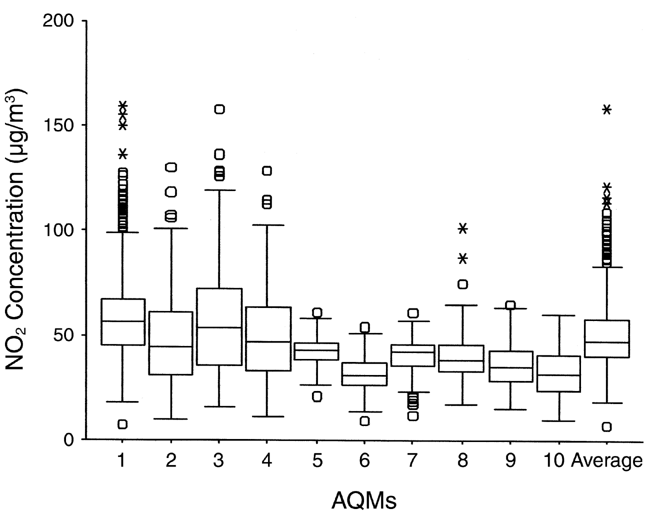


Figure 3. Box plot of NO_2 concentration at the 10 AQM sites. The bottom and top of the box show the 25th and 75th percentiles. The middle line of the box represents the median. The ends of the I-bar represent the minimum and maximum PM_{10} concentrations. Small circles represent outliers and * represents extreme values.

Time-series plots of PM_{10} concentration at AQMs 1–10 (Figure 5) reveal yearly bimodal distribution in PM_{10} concentration; the NAAQS 24-hour average standard for residential areas is 100 µg/m^3. One increase occurred during the summer months of May and June and the other during the winter months of November and December. Dust storms are common in May and June; dust may have contributed to the high PM_{10} concentrations recorded during those months. Many festivals (Dassera, Diwali and Christmas) are celebrated during November and December. Vehicular traffic increases during these festivals. Diwali, one of the main Hindu festivals, is celebrated with large numbers of firecrackers; the pollution concentrations were very high during Diwali; PM_{10} concentrations as high as 1200 µg/m^3 were recorded. During the rainy season (July through September) the PM_{10} concentrations were much lower. Though there was considerable variation in the mean PM_{10} concentration across the AQM stations in Delhi, all sites had reasonably similar trends in daily PM_{10} concentrations.

Time-series plots of NO_2 at the 10 AQM locations are shown in Figure 6. Daily average NO_2 values recorded at AQMs 5–10 were below the NAAQS 24-hour average standard for residential areas (80 µg/m^3). At other locations the daily average NO_2 concentration exceeded the standard for only a few days per year.

Time-series plots of SO_2 concentration at the 10 AQM locations are shown in Figure 7. 24-hour average SO_2 concentration for most of the monitoring days at all AQM

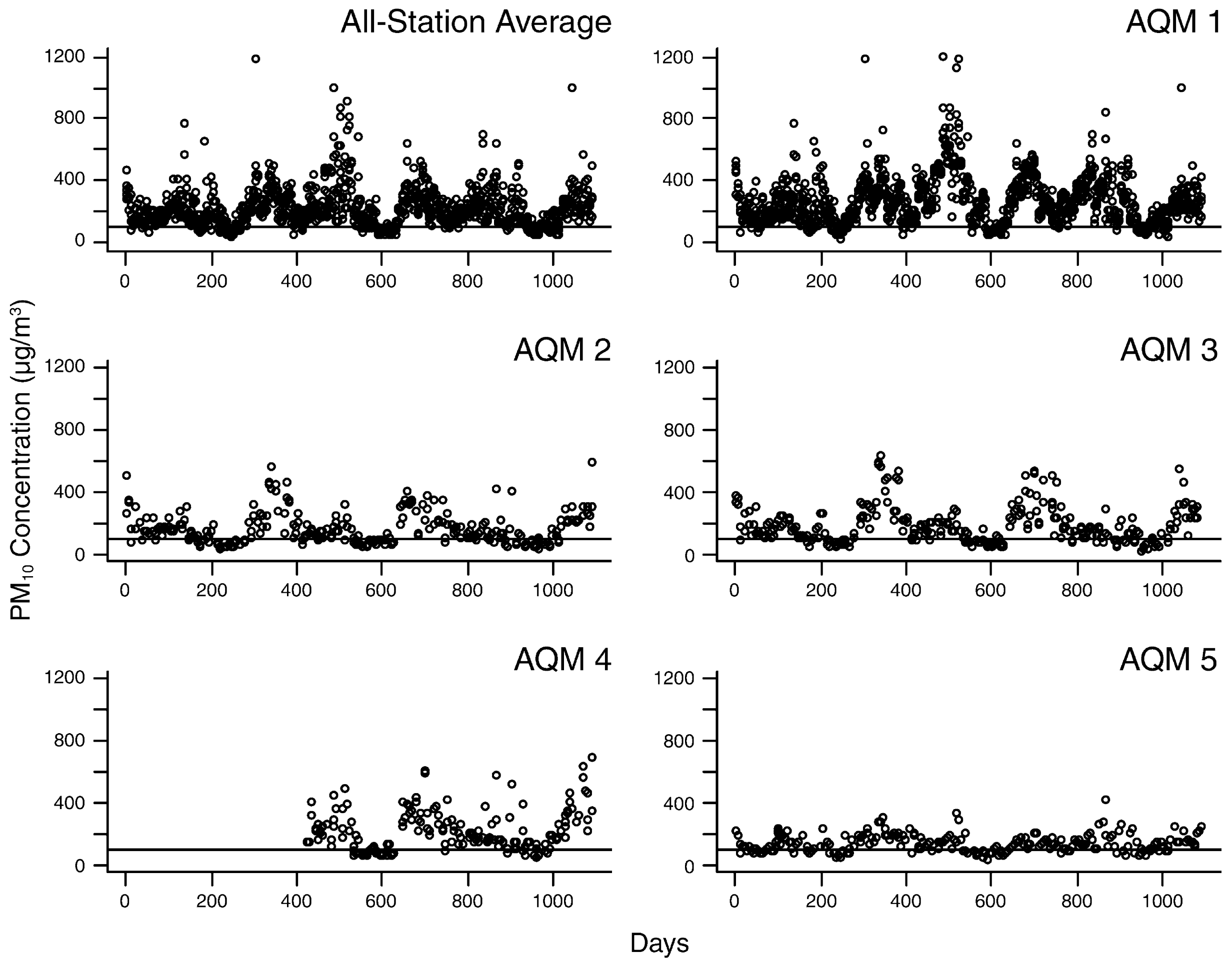


Figure 5. Time-series plots of PM_{10} concentration at the 10 AQM sites and the all-station average. The line at 100 µg/m^3 indicates the NAAQS 24-hour average standard for residential areas.

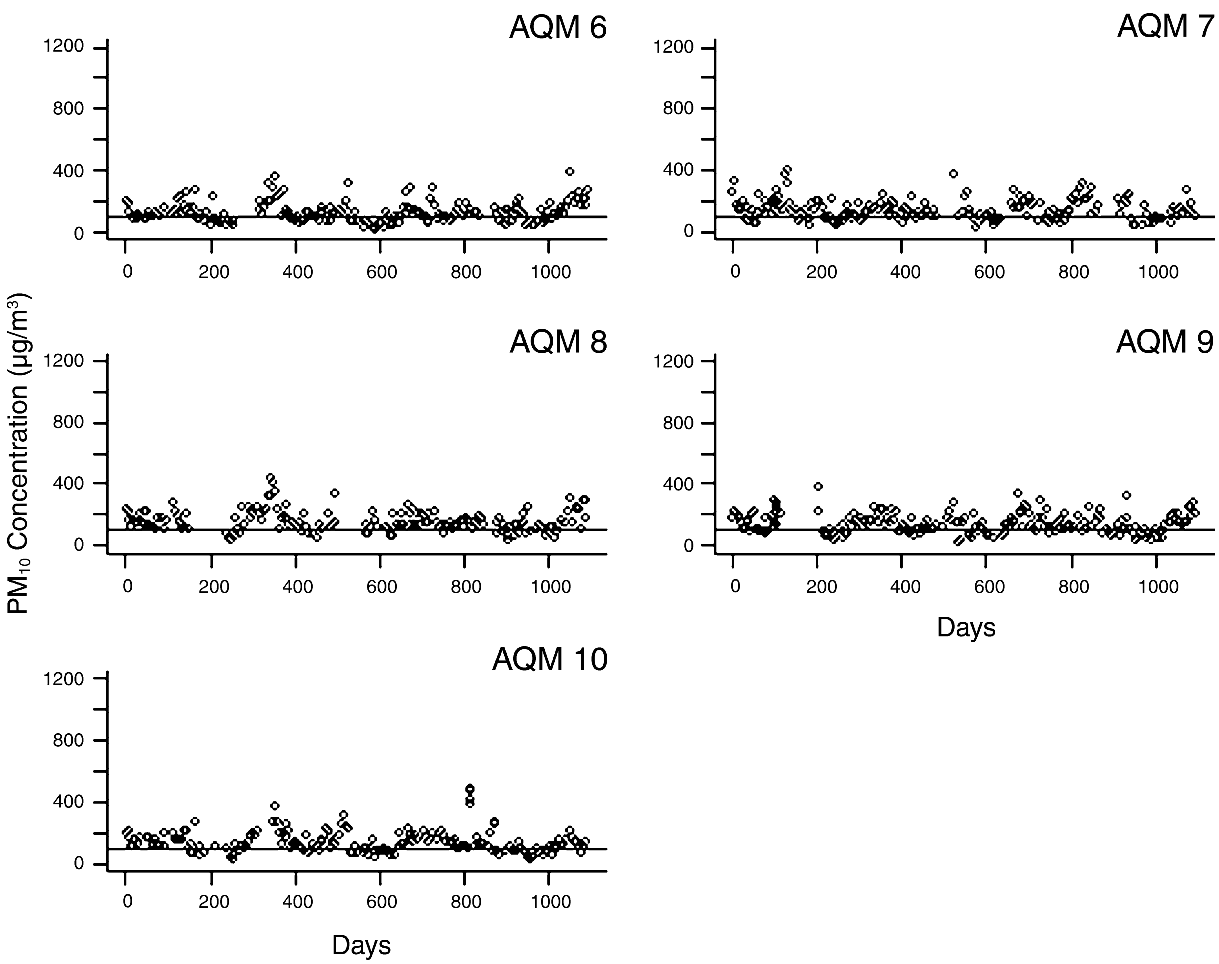


Figure 5 (*Continued*).

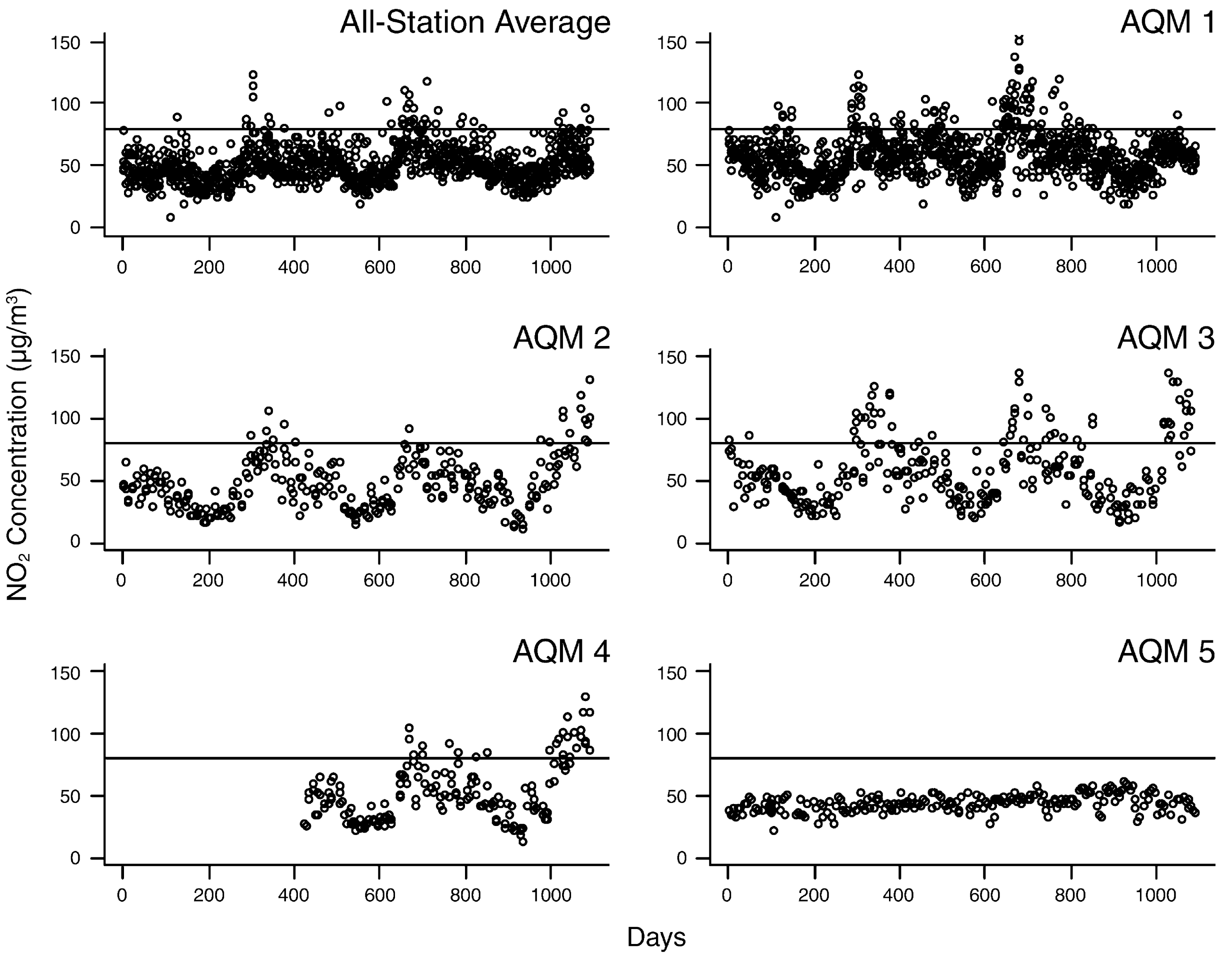


Figure 6. Time-series plots of NO_2 at the 10 AQM sites and the all-station average. The line at 80 µg/m^3 indicates the NAAQS 24-hour average standard for residential areas.

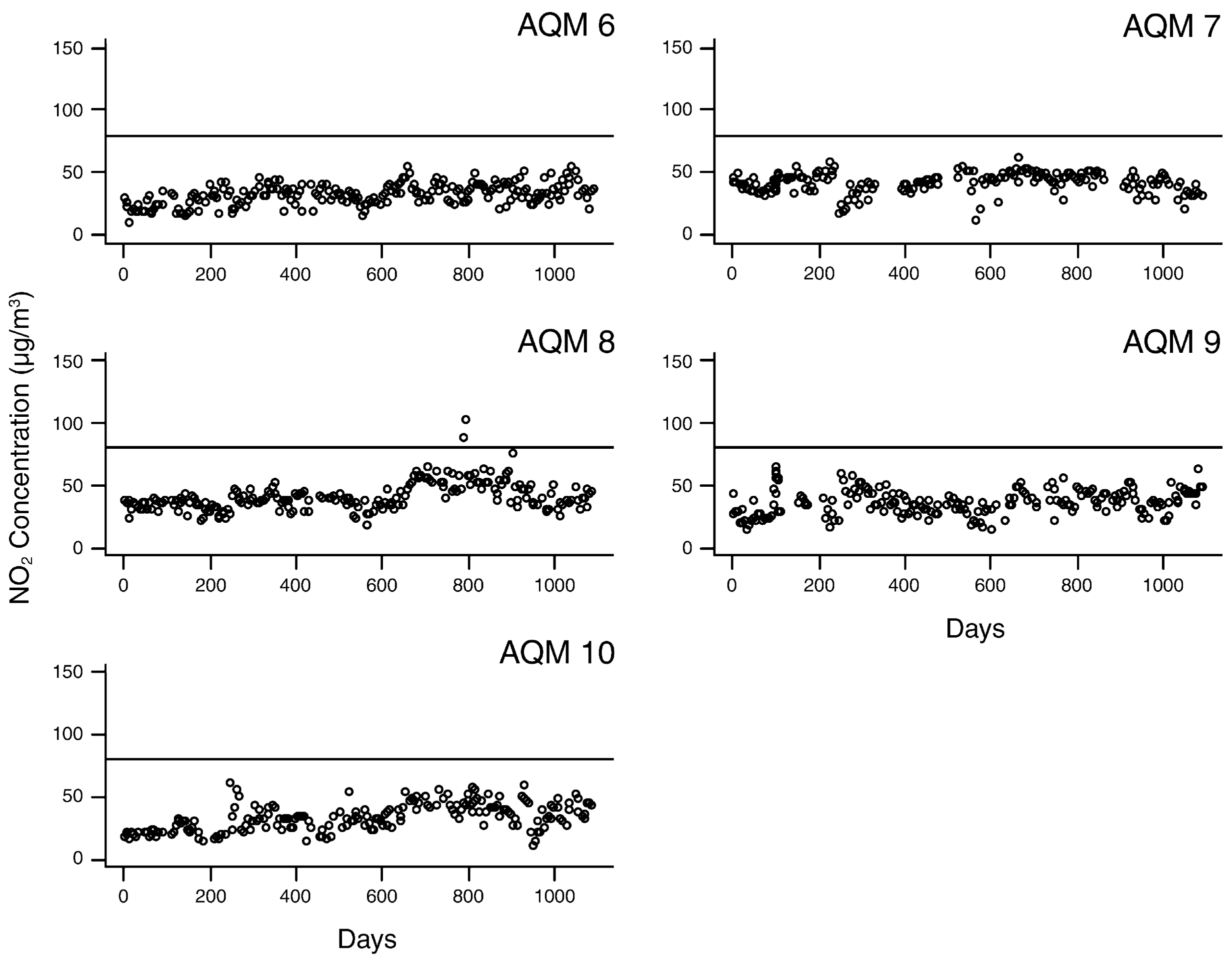


Figure 6 (*Continued*).

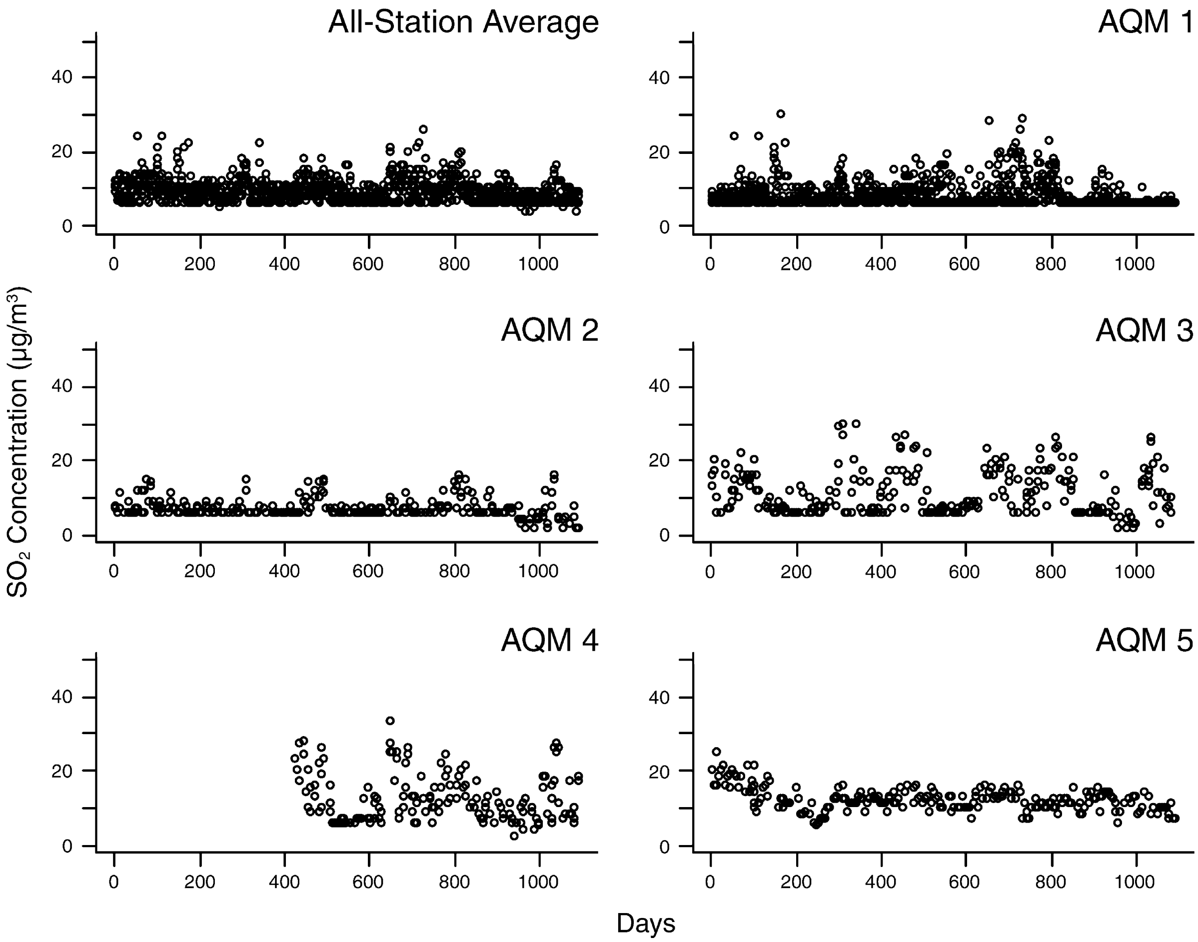


Figure 7. Time-series plots of SO_2 concentration at the 10 AQM sites and the all-station average. Concentrations were well within the NAAQS 24-hour average standard of 80 µg/m^3 for residential areas.

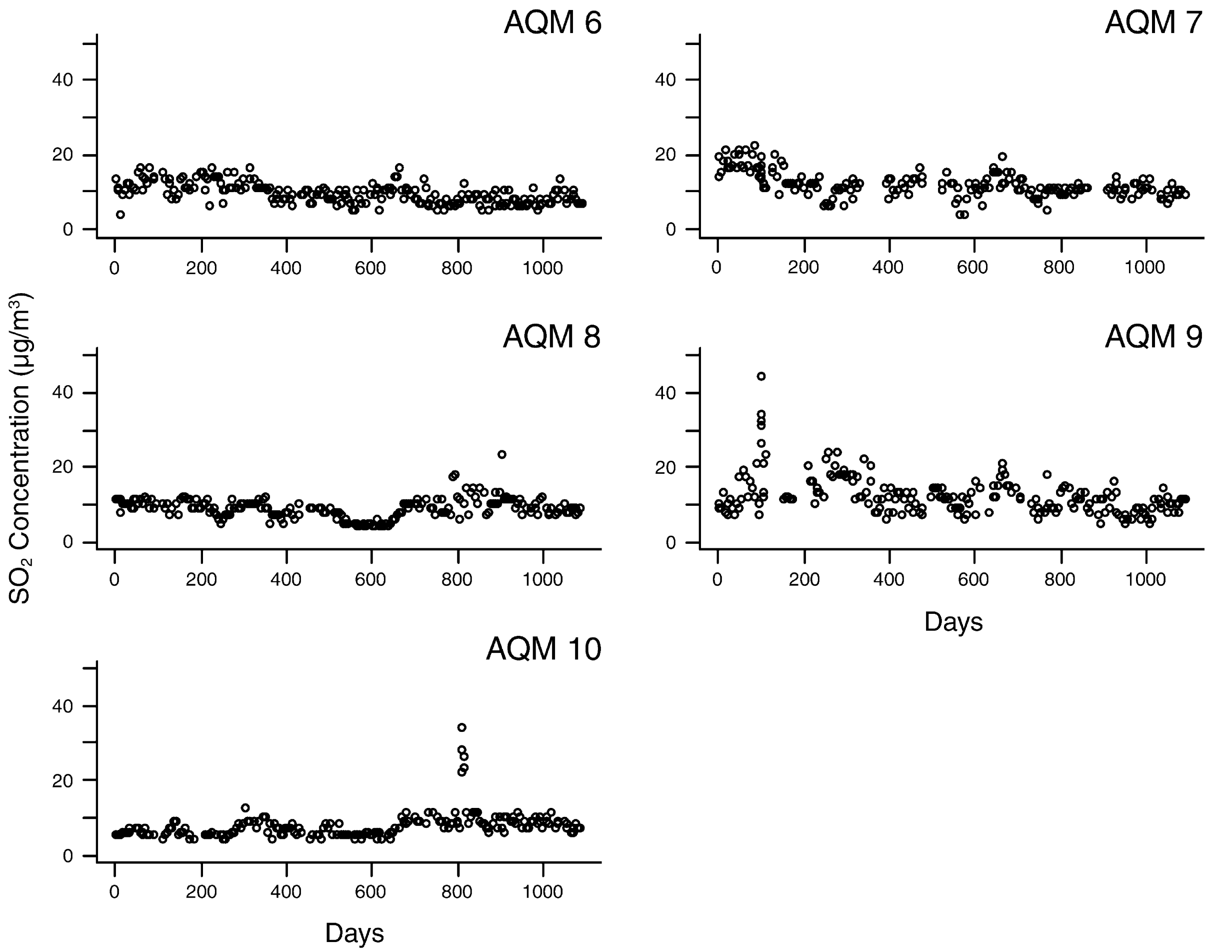


Figure 7 (*Continued*).

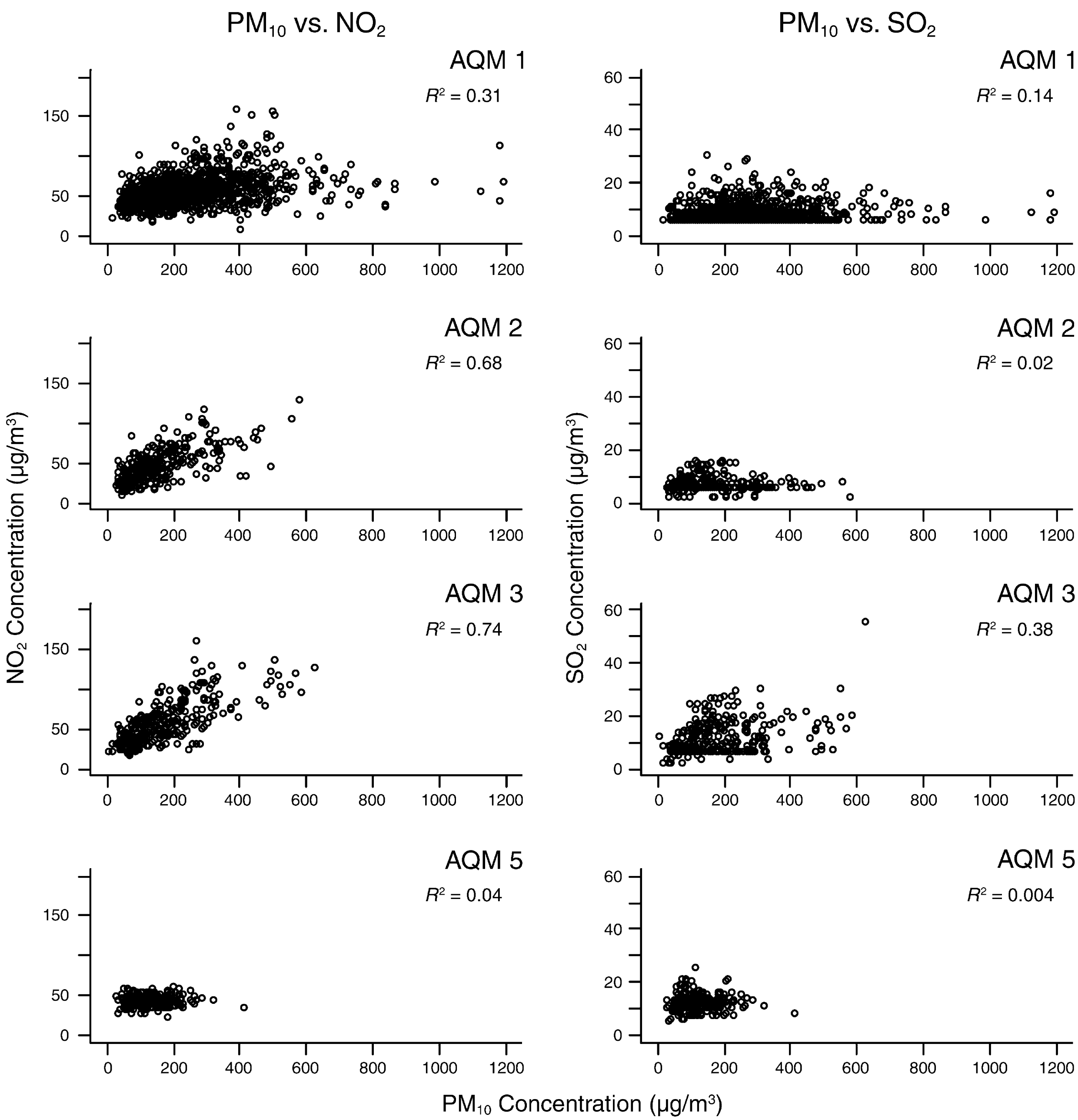
sites was well within the NAAQS 24-hour average standard for residential areas (80 μg/m^3).

Pairwise Pearson correlation coefficients for all available pairs for each of the air pollution parameters for all available days are shown in Table 5. Spatial correlations differed among various AQM locations. PM_{10} correlation varied from moderate (0.3–0.7) to high (> 0.7). PM_{10} was not strongly correlated with either NO_2 or SO_2 at most stations (Figure 8). The all-station average PM_{10} showed moderate to high correlation with individual AQM stations (Table 5). This indicates that sources other than vehicular emissions also contributed to PM_{10} pollution in Delhi.

Table 5. Pairwise Pearson Correlation Coefficients for the 10 AQM Sites by Pollutant[a]

Pollutant/ Site	AQM										
	1	2	3	4	5	6	7	8	9	10	Average
PM_{10}											
AQM 1	1.000										
AQM 2	0.537	1.000									
AQM 3	0.522	0.892	1.000								
AQM 4	0.515	0.867	0.847	1.000							
AQM 5	0.588	0.626	0.585	0.563	1.000						
AQM 6	0.441	0.652	0.605	0.655	—	1.000					
AQM 7	0.478	0.352	0.359	0.408	−0.545	—	1.000				
AQM 8	0.478	0.614	0.573	0.582	0.550	—	—	1.000			
AQM 9	0.513	0.650	0.622	0.627	0.459	—	0.396	—	1.000		
AQM 10	0.390	0.402	0.317	0.144	0.996	0.514	—	—	—	1.000	
Average	0.866	0.857	0.845	0.751	0.690	0.721	0.653	0.734	0.704	0.611	1.000
NO_2											
AQM 1	1.000										
AQM 2	0.518	1.000									
AQM 3	0.596	0.876	1.000								
AQM 4	0.473	0.887	0.839	1.000							
AQM 5	0.087	0.002	0.054	−0.127	1.000						
AQM 6	0.209	0.333	0.302	0.313	—	1.000					
AQM 7	0.168	0.019	0.061	0.187	−0.446	—	1.000				
AQM 8	0.182	0.241	0.316	0.289	0.377	—	—	1.000			
AQM 9	0.163	0.281	0.441	0.568	0.078	—	0.510	1.000	1.000		
AQM 10	0.181	0.166	0.183	0.158	—	0.392	—	—	—	1.000	
Average	0.760	0.878	0.912	0.877	0.306	0.514	0.320	0.523	0.446	0.480	1.000
SO_2											
AQM 1	1.000										
AQM 2	0.217	1.000									
AQM 3	0.228	0.568	1.000								
AQM 4	0.029	0.417	0.668	1.000							
AQM 5	0.013	0.143	0.332	0.007	1.000						
AQM 6	−0.029	−0.080	0.114	0.427	—	1.000					
AQM 7	0.058	0.143	0.195	0.196	−0.128	—	1.000				
AQM 8	0.028	0.017	0.091	−0.001	0.227	—	—	1.000			
AQM 9	0.126	0.106	0.388	0.338	0.097	—	0.188	1.000	1.000		
AQM 10	0.234	0.405	0.408	0.199	—	−0.152	—	—	—	1.000	
Average	0.648	0.629	0.853	0.800	0.477	0.352	0.485	0.510	0.723	0.625	1.000

[a] — indicates not available.



(*Figure continues on next page*)

Figure 8. Correlations between PM_{10} concentration and NO_2 or SO_2 concentrations at AQMs 1–3, 5, 8–10, and the all-station average.

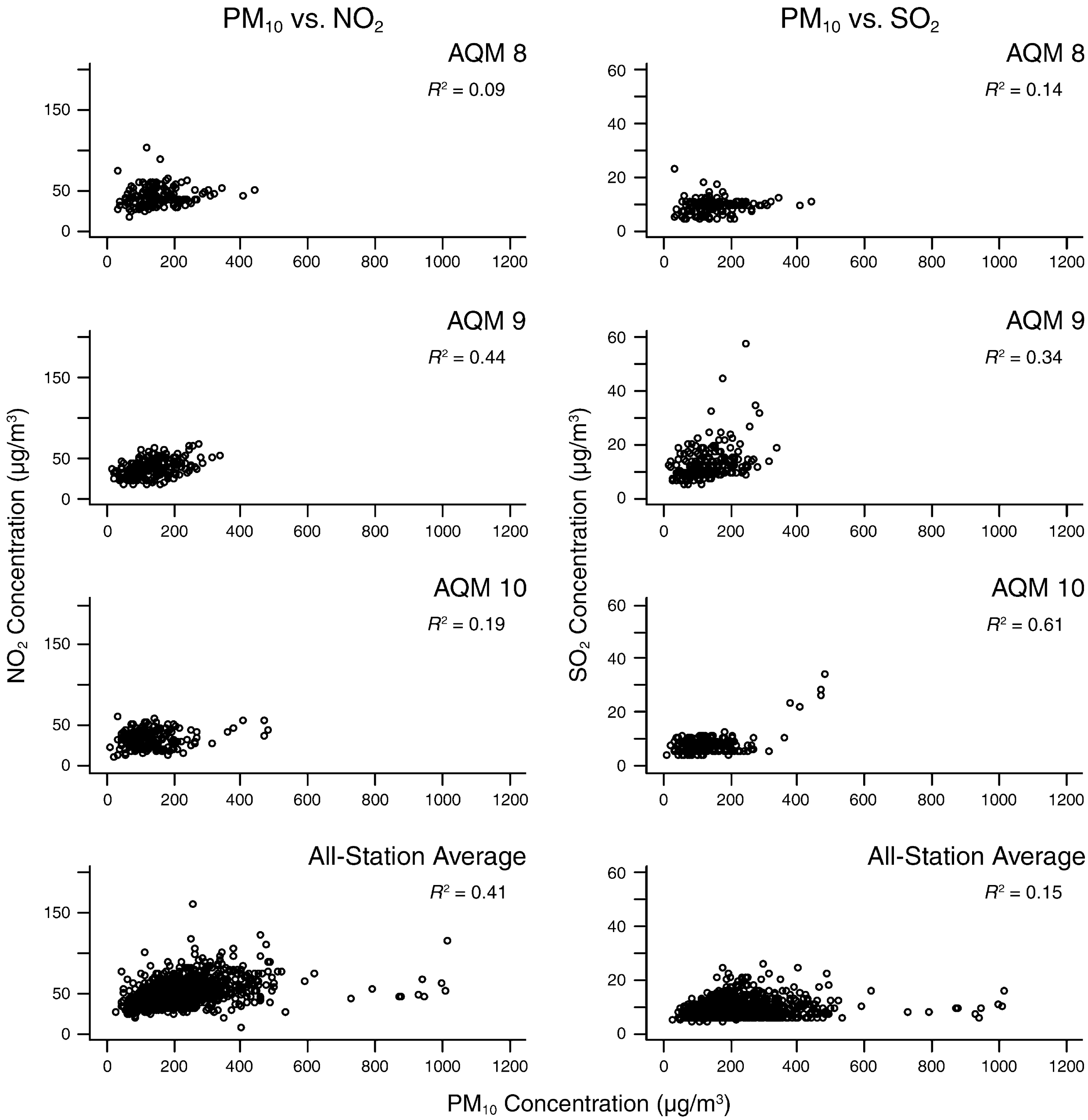


Figure 8 (*Continued*).

Summary statistics of meteorologic parameters are shown in Table 6. Average daily minimum relative humidity ranged from 8% to 98% for 2002, from 6% to 94% for 2003, and from 5% to 93% for 2004. Average daily maximum relative humidity ranged from 24% to 100% for 2002, 27% to 100% for 2003, and 37% to 100% for 2004. Time-series plots of average temperature and average relative humidity are shown in Figure 9. Minimum temperature in Delhi went as low as 3.5°C during the winters. In the summers the maximum temperature recorded was 45°C to 46°C (Table 6).

Average daily visibility is shown in Table 6. Visibility varied throughout the day. In 2002 the visibility recorded at 5:30 am ranged from 10 to 3000 meters, at 5:30 pm it ranged from 500 to 5000 meters, and at 8:30 pm it ranged from 300 to 3000 meters. In 2003 visibility ranged from 10 to 3000 meters at 5:30 am, 400 to 5000 meters at 5:30 pm, and 400 to 3000 meters at 8:30 pm. In 2004 visibility ranged from 20 to 4000 meters at 5:30 am, 200 to 6000 meters at 5:30 pm, and 200 to 4000 meters at 8:30 pm.

Total annual rainfall for years 2002, 2003, 2004 was 566 mm, 1166 mm, and 574 mm, respectively. The number of rainy days for each of those years was 49, 76, and 48 with average rainfall on those days being 11.5 mm, 15.3 mm, and 11.9 mm respectively.

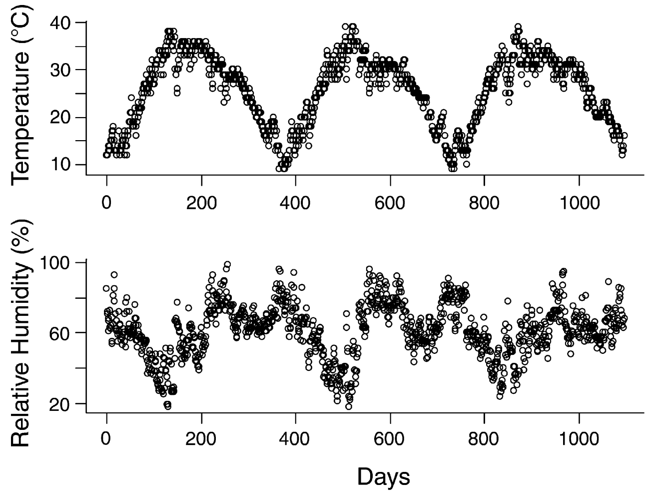


Figure 9. Time-series plots of daily average temperature and daily relative humidity.

Table 6. Summary Statistics of Meteorologic Data[a]

Year/ Daily Values	Daily Average: Visibility (m)	Min RH (%)	Max RH (%)	Min Temp (°C)	Max Temp (°C)
2002					
Mean	2439	39.7	80.5	19.5	32.2
Median	2583	37.0	87.0	21.0	33.4
Min	298	8.0	24.0	4.4	14.8
Max	4000	98.0	100.0	32.0	46.0
2003					
Mean	1931	42.9	81.9	19.0	30.9
Median	2000	39.0	89.0	20.0	32.7
Min	211	6.0	27.0	3.5	11.2
Max	3250	94.0	100.0	35.0	45.6
2004					
Mean	2339	39.8	80.5	19.4	31.9
Median	2500	37.0	84.0	20.8	33.4
Min	178	5.0	37.0	3.7	12.4
Max	5688	93.0	100.0	33.6	44.5

[a] Min indicates average daily minimum; Max indicates average daily maximum; RH indicates relative humidity; Temp indicates temperature.

MORTALITY DATA

Descriptive statistics for the mortality data are given in Table 7. Total daily all-natural-cause mortality varied from 126 to 368 with an average of 222 deaths per day. The age distribution reveals that about 65% of mean total deaths in Delhi occurred before the age of 65. The mean number of deaths for females was lower than that for males (38% vs. 62%). Time-series plots of daily all-natural-cause mortality from the NDMC and the MCD with a total for Delhi are shown in Figure 10.

Table 7. Descriptive Statistics for All-Natural-Cause Mortality Data[a]

	Min	Max	Mean	Median	SD
Total daily mortality	126	368	222	219	33
Female	40	142	85	83	15
Male	77	244	139	137	21
Age (years)					
0–4	5	39	19	18	6
5–44	28	102	64	64	11
45–64	23	125	61	60	12
≥ 65	28	158	76	74	28

[a] Min indicates minimum value; Max indicates maximum value; SD indicates standard deviation.

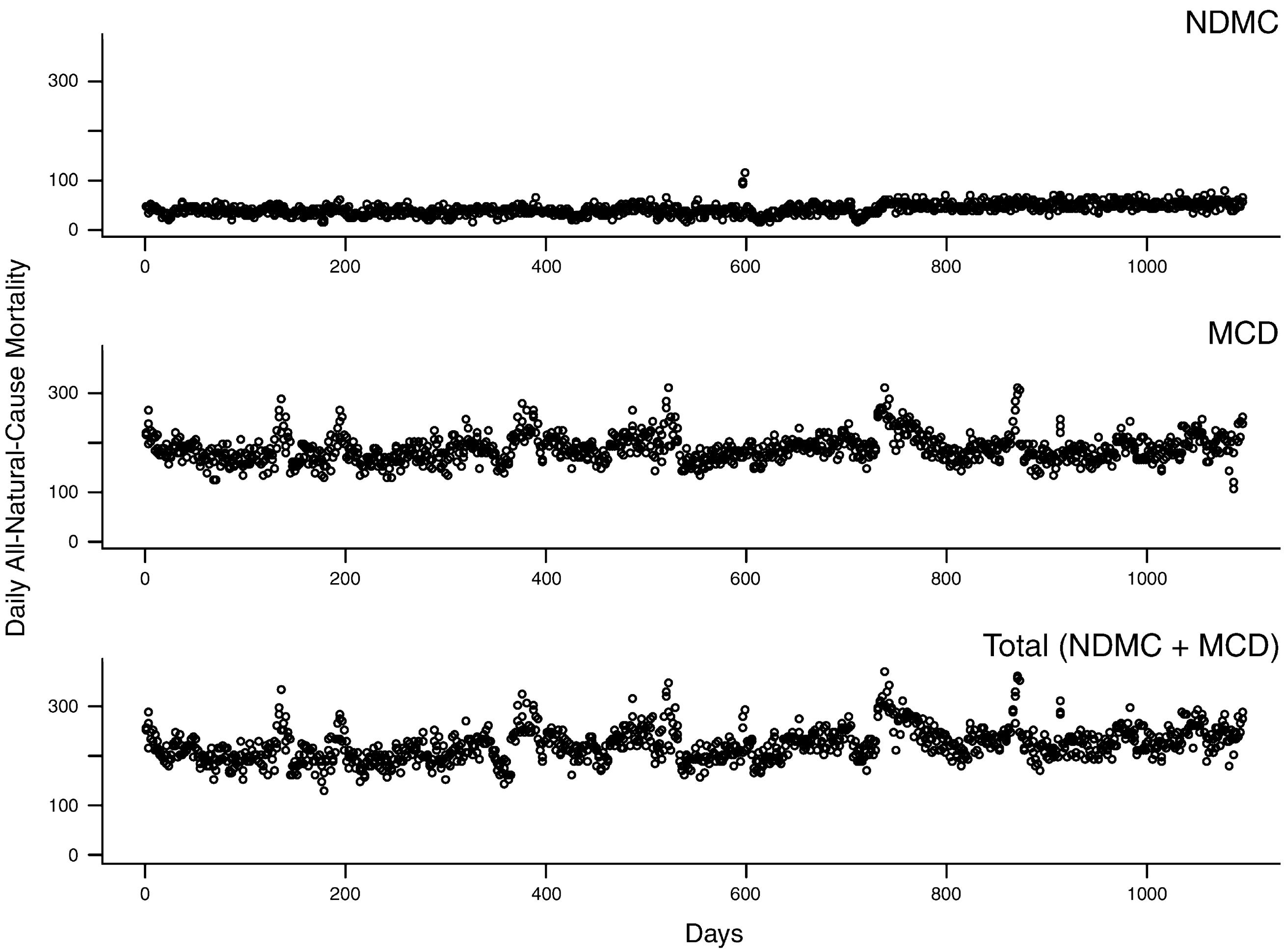


Figure 10. Time-series plots of daily all-natural-cause mortality for each of the two agencies (NDMC and MCD) and for the agencies combined (Total).

MODEL OUTPUT

Figure 11 shows the GAM plot for smoothing of all-natural-cause mortality over the covariates temperature, time, and relative humidity. As shown in Figures 12 and 13, noise in the core model was reduced after autoregression. The results of the GAM analysis using the core model specifications are presented in Table 8. The effect of PM_{10} was found to be significant with a 0.15% increase in daily all-natural-cause mortality for every 10-µg/m^3 increase in PM_{10} concentration. NO_2 also had a positive and significant effect on all-natural-cause mortality with a 0.84% increase in mortality for every 10-µg/m^3 increase in NO_2 concentration. The effect for SO_2 was negative but not significant. The effect of PM_{10} on all-natural-cause mortality increased to 0.17% when high PM_{10} concentration values ($\geq$ 400 µg/m^3) were removed.

The effect of PM_{10} on all-natural-cause mortality decreased for lag days 1, 2, and 3 compared with the core model (lag 0). However the cumulative mean of PM_{10} for lag days 0 and 1 had a greater effect on all-natural-cause mortality (0.17% for every 10-µg/m^3 increase in PM_{10} concentration) than did PM_{10} for lag 1 day (about 0.12% for

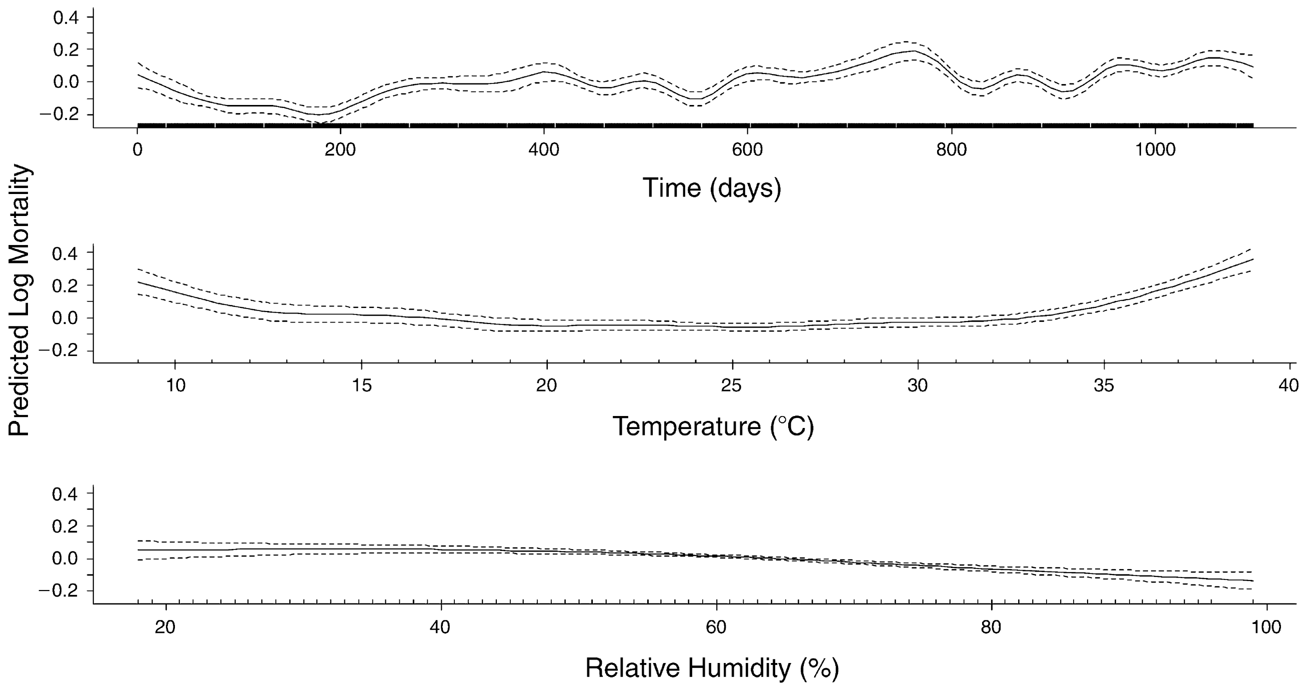


Figure 11. GAM model plot for smoothing of all-natural-cause mortality over time, temperature, and relative humidity.

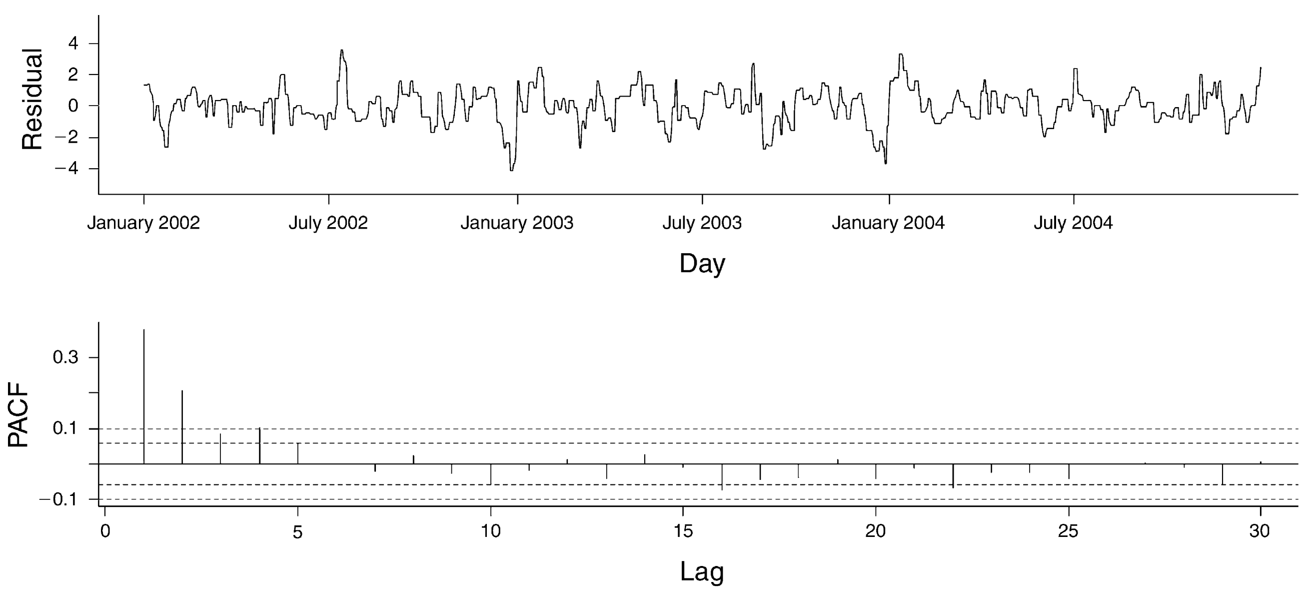


Figure 12. Residual and PACF plot of the core model before autoregression. Note that the y-axes of Figures 12 and 13 differ in scale.

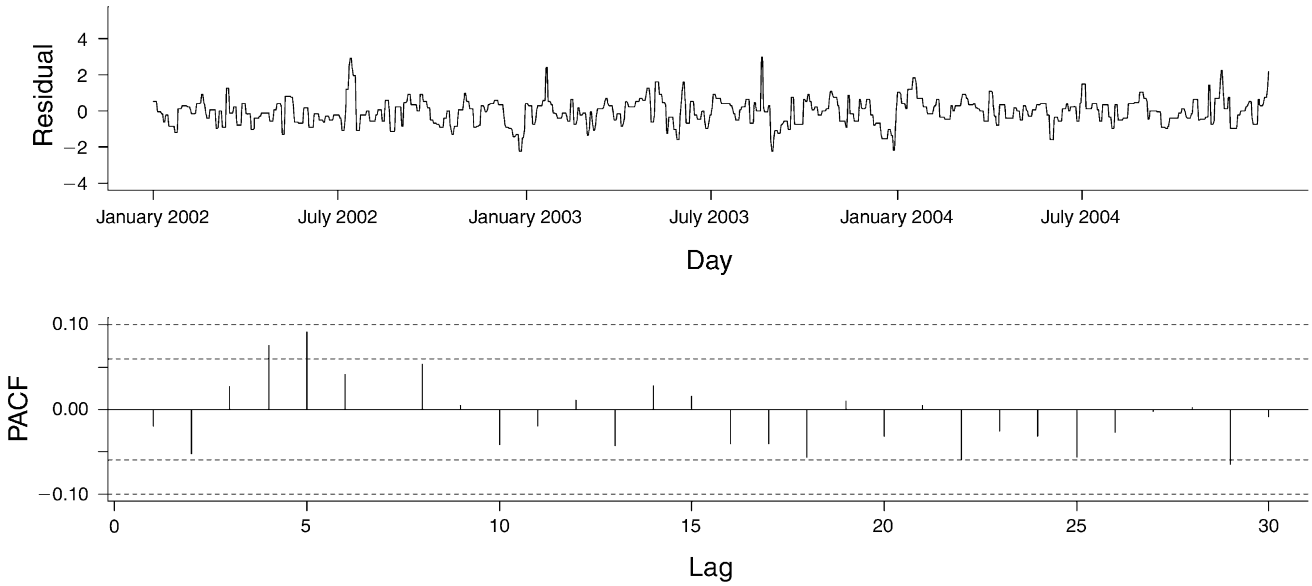


Figure 13. Residual and PACF plot of the core model after autoregression. Note that the y-axes of Figures 12 and 13 differ in scale.

Table 8. Results of GAM Analyses Using Core Specifications

Model	Parameters	β Coefficient	Change in Mortality per 10-μg/m^3 Increase in Pollutant Concentration % (95% CI)	*P* Value
Core model	Intercept	5.392		0.000[a]
Core model + PM_{10}	PM_{10}	0.00015	0.15 (0.07 to 0.23)	0.0007[a]
Core model + PM_{10} < 400 μg/m^3	PM_{10}	0.00017	0.17 (0.09 to 0.25)	0.0030[a]
Core model + PM_{10}, lag 1 day	PM_{10}, lag 1	0.00012	0.12 (0.04 to 0.20)	0.005[a]
Core model + PM_{10}, lag 2 day	PM_{10}, lag 2	0.00012	0.12 (0.04 to 0.20)	0.003[a]
Core model + PM_{10}, lag 3 day	PM_{10}, lag 3	0.00009	0.09 (0.01 to 0.17)	0.029[a]
Core model + PM_{10}, lag 0–1[b]	PM_{10}, lag 0–1[b]	0.00017	0.17 (0.09 to 0.25)	0.000[a]
Core model + NO_2	NO_2	0.00084	0.84 (0.29 to 1.4)	0.00271[a]
Core model + SO_2	SO_2	−0.0014	−1.4 (−3.56 to 0.76)	0.261
Core model + PM_{10} + NO_2	PM_{10}	0.00012	0.12 (0.03 to 0.21)	0.0072[a]
	NO_2	0.00065	0.65 (0.04 to 1.26)	0.088
Core model + PM_{10} + SO_2	PM_{10}	0.00015	0.15 (0.07 to 0.23)	0.000[a]
	SO_2	−0.0017	−1.7 (−4.2 to 0.85)	0.37
Core model + PM_{10} + NO_2 + SO_2	PM_{10}	0.00009	0.091 (0.012 to 0.17)	0.035[a]
	NO_2	0.0008	0.8 (0.21 to 1.4)	0.011[a]
	SO_2	−0.00215	−2.1 (−4.7 to 0.45)	0.094

[a] Significant at $P < 0.05$.

[b] Lag 0–1 is the cumulative mean of lags 0 and 1.

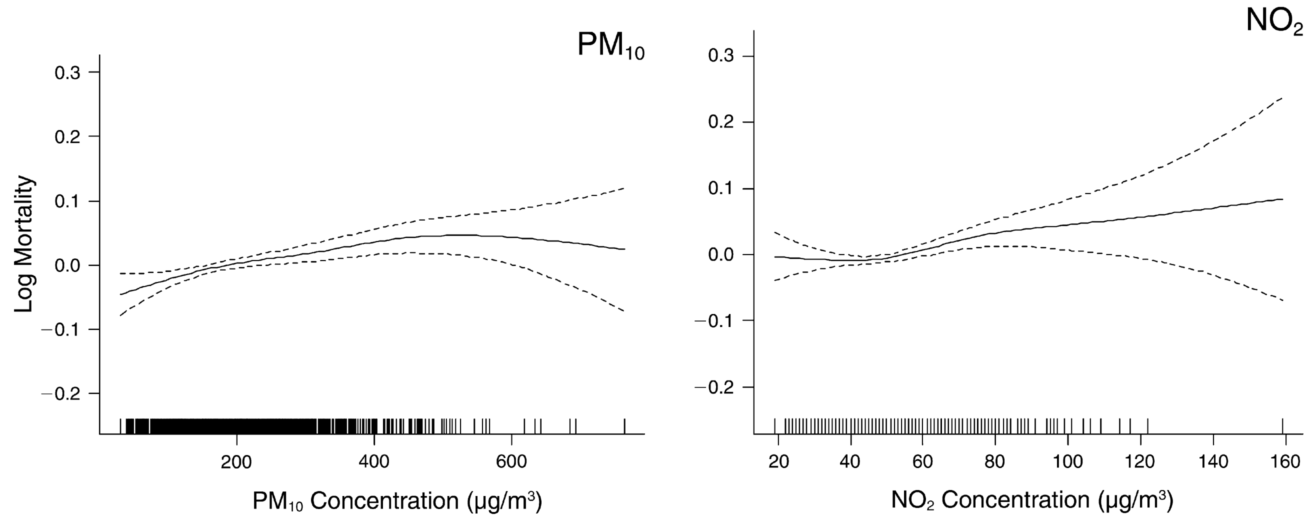


Figure 14. Dose–response curves for PM_{10} and NO_2 concentrations.

every 10-µg/m^3 increase in PM_{10} concentration). When more than one pollutant was considered in the model (PM_{10} + NO_2 or PM_{10} + NO_2 + SO_2), the effect was slightly reduced compared with a model that included PM_{10} only. The dose–response curves for PM_{10} and for NO_2 show curvilinear relationships with the linear component at lower concentrations and a slight flattening (lower slope) at higher concentrations (Figure 14).

Analysis by different age groups (Table 9) reveals that the maximum effect occurred in the 15–44 age group.

SENSITIVITY ANALYSES

Various degrees of freedom (3–15 *df*/year) using natural spline smoothers were tried in GAM analysis. Percentage change in the mean number of daily all-natural-cause mortality with 95% confidence intervals (CI) is shown in Figure 15. The overall effect was low with few degrees of freedom (3–4 *df*/year), and there was a marginal decline with degrees of freedom above nine per year.

Table 9. Summary of Model Results by Sex and Age Groups for Increases in PM_{10}

Population Subgroup	β Coefficient	Change in Daily All-Natural-Cause Mortality per 10-µg/m^3 PM_{10} Increase % (95% CI)	*P* Value
Female	0.00017	0.17 (0.06 to 0.28)	0.0011[a]
Male	0.00009	0.09 (−0.008 to 0.19)	0.226
Age (years)			
0–4	0.000077	0.077 (−0.02 to 0.18)	0.340
5–44	0.00015	0.15 (0.05 to 0.25)	0.0039[a]
45–64	0.00011	0.11 (0.02 to 0.20)	0.036[a]
≥ 65	0.000098	0.098 (−0.06 to 0.26)	0.056

[a] Significant at $P < 0.05$.

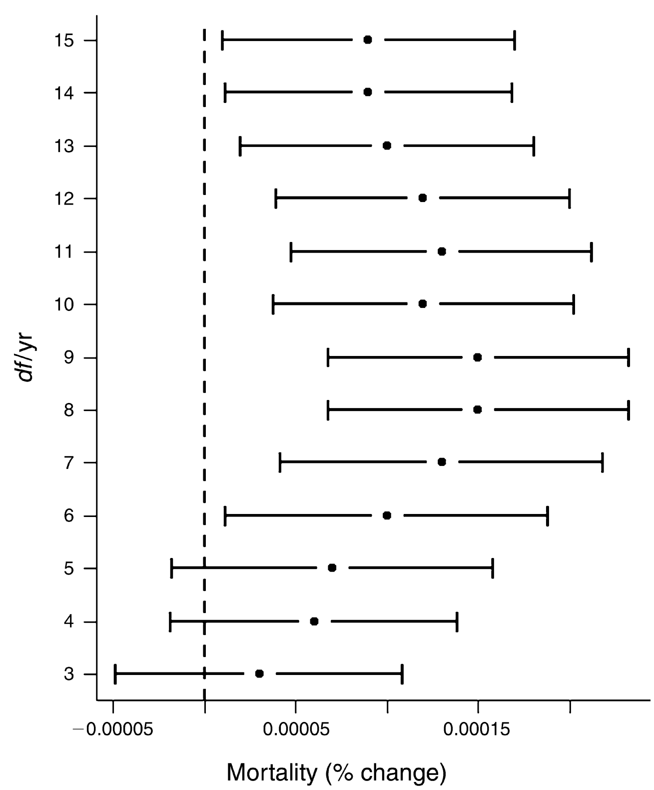


Figure 15. Percentage change in daily all-natural-cause mortality, with 95% CIs, for each 10-µg/m^3 PM_{10} increase using natural spline smoothers and various degrees of freedom.

Table 10. Percentage Change in Daily All-Natural-Cause Mortality per 10-µg/m^3 PM_{10} Increase by AQM Site

Site	Change in Daily All-Natural-Cause Mortality per 10-µg/m^3 PM_{10} Increase (%)	*P* Value
AQM 1	0.14	0.0001[a]
AQM 2	0.22	0.0184[a]
AQM 3	0.38	0.000[a]
AQM 4	0.26	0.0017[a]
AQM 5	0.31	0.0494[a]
AQM 6	−0.12	0.460
AQM 7	−0.29	0.0512
AQM 8	−0.011	0.945
AQM 9	0.47	0.0009[a]
AQM 10	−0.06	0.638
All-station average (10 AQMs)	0.12	0.00004[a]
Average of 9 AQMs (excluding AQM 1)	0.15	0.00006[a]

[a] Significant at $P < 0.05$.

GAM output of PM_{10} values for individual AQM locations are given in Table 10. The percentage change in daily all-natural-cause mortality for each 10-µg/m^3 increase in PM_{10} concentration at individual AQM locations ranged from −0.29 to 0.47. The percentage changes at AQMs 6, 7, 8, and 10 were negative but not significant. None of the individual AQM station values can be considered to be a true representation of population exposure. All could have been influenced by local sources of air pollutants. In addition, AQMs 2–10 have a large number of missing values. Because the average of all AQMs shows moderate to high correlation with individual AQM stations (Table 5), it is considered to be a good representation of population exposure. It shows an increase of 0.12% all-natural-cause mortality for each 10-µg/m^3 increase in PM_{10} concentration. When AQM 1 values (which were consistently higher than values at other stations) were removed and average exposure values were estimated using the nine CPCB AQM values, the overall estimate increased marginally.

Table 11 summarizes the results of the model with one to three day lag periods for temperature. The table reveals that there is not much change in the overall effect when short duration lag effects of temperature (i.e., one to three days) were considered in the model.

Table 11. Percentage Change in Daily All-Natural-Cause Mortality per 10-µg/m^3 PM_{10} Increase for the Effect of Temperature with Various Lagged Days

Lag Period for Temperature	Change in Daily All-Natural-Cause Mortality per 10-µg/m^3 PM_{10} Increase (%)	*P* Value
1 Day	0.14	0.0007[a]
2 Days	0.13	0.0017[a]
3 Days	0.13	0.0014[a]
Cumulative average of lag 1 and 2 days	0.13	0.021[a]
Cumulative average of lag 0–7 days	0.10	0.059
Cumulative average of lag 8–14 days	0.09	0.04[a]

[a] Significant at $P < 0.05$.

DISCUSSION

STUDY RESULTS

Results of the present study reveal that the effect of PM_{10}, while statistically significant, is small when compared with the results of other studies (Table 12). This is even lower than the 0.23% change in daily all-natural-cause mortality with a 10-µg/m^3 SPM increase in Delhi during 1991 to 1994 that was reported by Cropper and colleagues (1997).

There are many plausible explanations for this small effect. One such explanation is the variation in the nature of PM. PM toxicity depends on the size and chemical characteristics of the particles. Though source characterization of PM in Delhi is not studied in detail, reports from a few studies indicate that the contribution from road dust or PM from natural sources, both of which may be less toxic than PM from vehicles or industrial sources, could be as high as 40% (Energy Sector Management Assistance Program 2004). Since late 1990, the government has taken various actions to control air pollution, including relocating industries from Delhi, tightening vehicular emission standards, and shifting to cleaner fuels (e.g., compressed natural gas). These actions have helped to arrest the growing air pollution problem in Delhi. There could also be changes in chemical characteristics of PM that have not been studied. The small effect of PM_{10} could have also been affected by changes in

Table 12. Comparison of Present Study Results with the Results from Other Studies

Location	Change in Daily All-Natural-Cause Mortality per 10-μg/m^3 Particulate Increase (%)	Particulate Measured	Source
Birmingham, Alabama, USA	0.6	PM_{10}	Schwartz 1993
29 European cities	0.41[a]	PM_{10}	Katsouyanni et al. 2001
Bangkok, Thailand	0.8	PM_{10}	Ostro et al. 1999
Delhi, India	0.2	SPM	Cropper et al. 1997
Delhi, India	0.114	PM_{10}	Present study

[a] Fixed effects estimate.

the demographic characteristics of the city. In the present study, about 65% of the deaths in Delhi occurred in people under the age of 65. The Delhi population may be more susceptible to other risk factors such as cardiovascular disease, possibly because of dietary patterns or nutritional status, waterborne diseases, or other factors.

Analysis by different age groups (Table 9) reveals that the maximum effect occurred in the 15–44 age group. Similar findings have been reported in an earlier study in Delhi (Cropper et al. 1997). This implies that in the 15–44 age group, more life years are lost per person due to air pollution than in younger or older age groups.

STUDY LIMITATIONS

The present study is one of very few studies undertaken in India that attempted to measure the short-term effect of air pollution on mortality. However the study has following limitations:

- Regular measurement of air quality was limited to urban areas and to criteria pollutants, namely PM (SPM, PM_{10}), SO_2, and NO_2. Neither the nature nor the chemical composition of the PM pollution was determined.
- Though air quality was monitored at 10 AQM sites, 9 of the sites had large numbers of missing data. This also influenced the overall average pollutant concentrations used in the exposure series.
- Despite the efforts taken through centering techniques to minimize the influence of missing data, exposure measurement errors were likely. Assuming such measurement errors were less differential with respect to the population at risk, the overall estimates were likely to be biased downward.
- The influence of possible geographic confounding factors could not be studied owing to unmeasured spatially-varying factors.
- Because of the limitations associated with overall cause-of-death coding, analyses were limited to all-cause natural deaths. Disease-specific outcomes could not be studied.
- Underreporting the mortality data was also expected, particularly for infants, children, and women.

CONCLUSIONS

Findings from this study in Delhi are broadly in agreement with those of previous studies that found positive associations between daily changes in all-natural-cause mortality and PM_{10} air pollution. The magnitude of the effect was small but statistically significant after controlling for meteorologic parameters and time trends. The relatively smaller effect of PM pollution on all-natural-cause mortality, compared with that in other studies, could be due to chemical characteristics of PM and its associated toxicity. Differences in demographic characteristics and life style might also alter the susceptibility to air pollution. The coefficients determined in the present study may be more suitable for measuring the economic effect of air pollution in the cities of developing countries that have high air pollutant concentrations. The study provides insight into the link between air pollution and all-natural-cause mortality in local populations and contributes information to the existing body of knowledge.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. B. Sengupta (Member Secretary), Dr. Naresh Badwar, and Dr. Makijani of the CPCB for their active cooperation and support in air quality data collection. We thank Dr. Ramesh Kumar, NDMC, for his support in mortality data collection. We gratefully acknowledge the technical support provided by the HEI staff members and the International Scientific Oversight Committee members of HEI. We also acknowledge Dr. Kalpana Balakrishnan, and Mr. Santanu Ghosh of Sri Ramachandra Medical College for their support. We thank Dr. R.K. Pachauri (Director General of TERI) for his encouragement and support. We acknowledge Ms. Beena, Mr. Ritesh Jha, and Mr. Shenoy Mooken for their secretarial assistance.

REFERENCES

Ackermann-Liebrich U, Rapp R. 1999. Epidemiological Effects of Oxides of Nitrogen, Especially NO_2. In: Air Pollution and Health (Holgate ST, Samet JM, Koren HS, Maynard R, eds.). Academic Press, London, UK.

Badwar N, Patak A. 2007. Findings on particulate matter. Presentation in Workshop on Some Recent Advances in the Field of Air Quality Management. September 23, 2007, Indian National Science Academy, New Delhi, India. Unpublished.

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 1. Am J Respir Crit Care Med 153:3–50.

Central Pollution Control Board. 2003. Air Quality of Delhi. *www.cpcb.delhi.nic.in/bulletin/bul.htm*. Accessed 07/2005.

Cropper ML, Simon NB, Alberini A, Sharma PK. 1997. The health effects of air pollution in Delhi, India. PRD working paper 1860 (unpublished). New Ideas in Pollution Regulation, World Bank, Washington, D.C. Available from: *www.worldbank.org/nipr/work_paper/1860/index.htm*. Accessed 04/2006.

Energy Sector Management Assistance Program (ESMAP). 2004. Towards cleaner urban air in South Asia: Tackling transport pollution, understanding sources. Report 281/04. The World Bank, Washington, DC.

Government of National Capital Territory of Delhi. 2002. Economic Survey of Delhi, 2001–2002. Planning Department, Government of National Capital Territory of Delhi, India.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 12:521–531.

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 time series data from the APHEA project. BMJ 314:1658–1663.

Ostro B, Chestnut L, Vichit-Vadakan N, Laixuthai A. 1999. The impact of particulate matter on daily mortality in Bangkok, Thailand. J Air Waste Manage Assoc 49:PM100–PM107.

Ruppert D, Want MP, Carroll RJ. 2003. Semiparametric Regression. Cambridge University Press, Cambridge, UK.

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.

Schwartz J. 1993. Air Pollution and Daily Mortality in Birmingham, Alabama. Am J Epidemiol 137:1136–1147.

Wong CM, Ma S, Hedley AJ and Lam TH. 2001. Effect of air pollution on daily mortality in Hong Kong. Environ Health Perspect 109:335–340.

Wood SN. 2000. Modeling and smoothing parameter estimation with multiple quadratic penalties. J R Stat Soc Series Stat Methodol B 62:413–428.

APPENDIX A. HEI Quality Assurance Statement*

The conduct of this study was subjected to periodic, independent audits by a team from Hoover Consultants. This team consisted of auditors with experience in toxicology, epidemiology and air quality data. The audits included in-process monitoring of study activities for conformance to the study protocols and examination of records and supporting data. The dates of each audit are listed below with the phase of the study examined.

December 8–9, 2005

Records from this study were obtained by the investigators from external groups and the audit did not extend beyond these records to the original data sources. The 5- and 10-month progress reports were included in the audit. This study was conducted in accordance with two protocols: an individual protocol that included the unique features of the New Delhi project and a combined protocol for the coordinated time-series analysis.

Mortality data were collected by the Municipal Corporation of Delhi (MCD) for approximately 97% of the total population of 13.78 million. The New Delhi Municipal Corporation (NDMC) collected 2% and the remaining 1% was collected by the Delhi Cantonment under the control

* The QA Statement was edited for brevity.

of the military (excluded from study). The audit team compared hard copy death information against electronic files. In addition, Dr. Ramesh Kumar, Chief Medical Officer for the Community Health Administration of the New Delhi Municipal Council (NDMC) met with the audit team at TERI to discuss procedures that NDMC follows in the collection of mortality data. Dr. Kumar provided the audit team with a tour of the NDMC office and explained procedures for data obtained from reporting family members, entry into the computer system and archival of hard copies. The audit team was able to audit data sets from both MCD and NDMC and audited the electronic files against hard copy. Data from 270 deaths were audited for registration number, date-of-death, date-of-registration, name, father or husband's name, address-of-deceased, gender, age, place-of-death, living place-of-deceased, district, religion, occupation, whether medical attention was obtained, if the death was certified, cause-of-death, if the death was pregnancy-related, and history of alcohol and tobacco use.

Data from the India Habitat Center observatory (i.e., AQM 1) are generated and processed at TERI and were available to the audit team for review, from hand-written initial record through electronic data-reduction spreadsheet to tabulated 24-hour averages. Data for the other sites were obtained by the investigators as electronic files of summary 24-hour averages. All 24-hour averages at the India Habitat Center during 2003 for PM_{10}, SO_2, and NO_2 were audited against values from the data-reduction spreadsheet, and all spreadsheet values for 16 of these days were audited against the original hand-written initial record.

TERI investigators obtained hard-copy meteorological records for the Safdarjung Airport (VIP airport) from the India Meteorological Department. The data, for daily maximum and minimum temperature and relative humidity, daily total rainfall, and visual range observations every three hours, were entered in electronic spreadsheets that were independently verified by a TERI research assistant. The audit team merged visual range and relative humidity data from the electronic files with the pollutant data, and examined the empirical dependence of visual range on relative humidity and PM_{10} concentrations. The consistency with theoretical expectations of these wholly independent measurements supports the quality of both data sets.

July 28, 2010

A "cleaned, but unedited draft" of the final study report was examined for internal consistency and conformance with the study protocols. Comments were provided to HEI via e-mail.

A written report of the December 2005 audit was provided to the Director of Science of the Health Effects Institute who transmitted these findings to the Principal Investigator. These quality assurance audits demonstrated that the study was conducted by experienced professionals in accordance with both study protocols. The final report appears to be an accurate representation of the study.

B. Kristin Hoover
Hoover Consultants

APPENDICES AVAILABLE ON THE WEB

Appendices B–E and H–J contain supplemental material not included in the printed report. They are available on the HEI Web site *http://pubs.healtheffects.org*.

Appendix B. Common Protocol for Time-Series Studies of Daily Mortality in Indian Cities

Appendix C. Method for Determination of Respirable Suspended Particulate Matter (RSPM) in the Ambient Air (Gravimetric technique with high volume sampling)

Appendix D. Determination of Sulphur Dioxide in Ambient Air (Improved West and Gaeke Method)

Appendix E. Determination of Nitrogen Dioxide in Ambient Air (Jacob and Hochheiser Method)

Appendices F and G are associated with the Chennai portion of this HEI Report.

Appendix H. Details of Air Quality Monitoring Locations in Delhi

Appendix I. Mortality Data Registration System in Delhi

Appendix J. Technical Summary

ABOUT THE AUTHORS

Uma Rajarathnam received her Ph.D. in the field of energy and environment. She has more than 17 years of experience in the area of environmental studies. Her area of research includes air pollution monitoring, indoor air pollution studies, exposure assessment, and health studies. She worked with TERI for 13 years. Currently she is heading the environment practice at Enzen Global Solutions Pvt. She was awarded the Fulbright Indo American Environmental Leadership Program fellowship.

Meena Sehgal M.P.H, M.Sc., has worked extensively on design and management of community and public health programs in developed and developing countries. She has undertaken a variety of research assignments and authored position papers on various community health issues. Ms. Seghal has applied principals of epidemiologic planning and analysis to a wide range of research issues such as the effect of asthma on quality of life, the prevalence of disability arising from chronic conditions, and self-reported alcohol-impaired driving. At TERI she is currently involved in air pollution health effect studies and the effect of climate change on health.

Subramanya Nairy received his Ph.D. in biostatistics. He has more than 13 years of experience in teaching and research. His field of research includes statistical analysis of environmental data, epidemiology, and statistical computing.

R. C. Patnayak M.B.B.S., M.D., is chief medical officer at MCD where his responsibilities include birth and death registration and data collection. He has introduced the online registration system for birth and death data.

Sunil Kumar Chhabra M.B.B.S., M.D., is professor of cardiorespiratory disease at Vallabai Patel Chest Institute in Delhi. He has more than 20 years of experience in teaching and research. His area of research includes air pollution and epidemiology.

Kilnani M.B.B.S., M.D., is professor of medicine at All India Institute of Medical Sciences in Delhi. He has special interest in pulmonary and critical care medicine. He has been working in the field of health effects of outdoor air pollution for several years.

K. V. Santhosh Ragavan has a Master of Science degree in Environmental Health. He was with TERI for over five years. Currently he is a consultant at Enzen Global Solutions.

ABBREVIATIONS AND OTHER TERMS

APHEA	Air Pollution and Health: A European Approach
AQM	air quality monitor
CI	confidence interval
CPCB	Central Pollution Control Board
df	degrees of freedom
ER	excess risk
DOW	day-of-the-week
GAM	generalized additive model
ISOC	International Scientific Oversight Committee
MCD	Municipal Corporation of Delhi
NAAQS	National Ambient Air Quality Standard
NAMP	National Air Quality Monitoring Program
NDMC	New Delhi Municipal Council
NEERI	National Environmental Engineering Research Institute
NMMAPS	National Morbidity, Mortality, and Air Pollution Study
NO_2	nitrogen dioxide
PACF	partial autocorrelation function
PAPA	Public Health and Air Pollution in Asia
PM	particulate matter
$PM_{2.5}$	PM with an aerodynamic diameter ≤ 2.5 µm/m^3
PM_{10}	PM with an aerodynamic diameter ≤ 10 µm/m^3
QA	quality assurance
QC	quality control
RR	relative risk
SO_2	sulfur dioxide
SPM	suspended particulate matter
TERI	The Energy and Resources Institute
WHO	World Health Organization

CRITIQUE
Health Review Committee

Research Report 157, *Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Two Indian Cities*, K. Balakrishnan et al. (Part 1) and U. Rajarathnam et al. (Part 2)

INTRODUCTION

Epidemiologic time-series studies are commonly used to evaluate the short-term effects of ambient air pollution on public health (e.g., HEI 2003; Pope and Dockery 2006). The time-series approach relies upon day-to-day variations to estimate the association between short-term exposure to air pollutant concentrations and mortality or morbidity, such as hospital admissions for respiratory and cardiovascular diseases. Time-series studies have provided useful information about the association between air pollution and health outcomes in single geographic locations, but it is often challenging to compare results across locations because of differences among the methods used to collect and analyze the data. Recent multicity time-series studies have been conducted in cities of Canada, Europe, and the United States using similar analytic approaches (Samet et al. 2000; Katsouyanni et al. 2001, 2009; Dominici et al. 2003). Those studies have considerably strengthened the evidence for the association between short-term increases in concentrations of particulate matter (PM*) and mortality, and have contributed to the scientific basis for setting regulatory standards for air pollutants in the United States and elsewhere.

However, considerable uncertainties remain regarding the extrapolation of this evidence to developing countries, including large populations in Asia. High concentrations of air pollution continue to be observed in the rapidly growing economies of China and India, with many locations exceeding the World Health Organization (WHO) air quality guidelines (Bell et al. 2004; Gurjar et al. 2004; WHO 2006). The composition of the air pollution mixture can differ dramatically between Asia and other geographic regions and vary among cities within Asia due to substantial differences in pollution sources. For example, cities in Asia have been shown to experience unique patterns of traditional and modern sources of exposure (HEI International Scientific Oversight Committee 2004; 2010). Furthermore, public health impacts from air pollution can be magnified in population-dense megacities (Gurjar et al. 2008), particularly in developing countries in Asia where various other demographic factors and disease patterns interact with air pollution to influence population health (Cohen et al. 2004). The distinct and diverse cities in Asia and the relative lack of local studies call for greater attention to these areas, as described in the Preface to this report.

Given the extensive gaps in knowledge on the health effects of air pollution in Asia, HEI funded several coordinated time-series studies of air pollution and health in Asia under the Public Health and Air Pollution in Asia (PAPA) Program to collect information that would be relevant to local populations, with the added goal of supporting capacity building in the region. After a first wave of four studies in China and Thailand was underway (see Preface), HEI's International Scientific Oversight Committee (ISOC) selected three additional investigator teams to conduct time-series studies of air pollution and mortality in Indian cities: Chennai, Delhi, and Ludhiana. The investigators coordinated their approaches by adapting the common protocol of the first wave PAPA studies; differences in data availability and completeness in the Indian cities prompted the investigators to develop city-specific approaches.

Although substantial deficiencies in the available air pollution data in Ludhiana prevented the study from being completed in that city, ISOC commends the Ludhiana team for their efforts in assessing the available health outcomes

Dr. Balakrishnan's 1-year study, "Developing exposure-response functions for air pollutants from time-series analyses — A pilot exercise in Chennai, India," began in July 2005. Total expenditures were $65,700. The draft Investigators' Report from Balakrishnan and colleagues was received for review in September 2007. A revised report was received in June 2008. A second revised report, received in January 2009, was accepted for publication in May 2009.

Dr. Rajarathnam's 1-year study, "Time series study on air pollution and mortality in Delhi," began in July 2005. Total expenditures were $63,029. The draft Investigators' Report from Rajarathnam and colleagues was received for review in September 2007. A revised report was received in June 2008. A second revised report, received in December 2008, was accepted for publication in May 2009.

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

and air pollution data and for their contributions to the workshops and communications that coordinated the study design among the Indian teams. (In a recent publication, the team estimated mortality risks in Ludhiana using visibility as a surrogate for air pollution [Kumar et al. 2010].) The investigators of the studies in Chennai and Delhi aimed to use the most consistent methods possible (given the data constraints) that would allow for a comparison of the relative risks (RR) of mortality for these cities and the exploration of potential sources of variation in the results.

BACKGROUND

Previous assessments of the available time-series studies conducted in Asia have found the results of the Asian literature to be mostly consistent with the wider scientific literature on air pollution and health, and the statistical methods used were comparable to studies conducted in other regions (HEI ISOC 2004; 2010). More than 50 Asian time-series studies on mortality have been published between 1980 and 2007, including three multicity studies in Korea, Japan, and China and Thailand (Lee et al. 2000; Omori et al. 2003; Wong et al. 2008). Nearly all of those Asian studies revealed significant positive associations between air pollution and mortality from respiratory disease, cardiovascular disease, or other causes (HEI ISOC 2010).

However, some differences in the robustness of the association between mortality and specific pollutants (i.e., PM ≤ 10 μm in aerodynamic diameter [PM_{10}] and nitrogen dioxide [NO_2]) were observed with the inclusion of other pollutants between the coordinated first wave PAPA studies in Bangkok, Hong Kong, Shanghai, and Wuhan (Wong et al. 2010b) and two North American and European studies: the National Morbidity, Mortality, and Air Pollution Study (NMMAPS) and Air Pollution and Health: A European Approach (APHEA) (Samet et al. 2000; Katsouyanni et al. 2001; Dominici et al. 2003). Given these remaining uncertainties and the under-representation of Southeast Asia and South Asia in the Asian literature on air pollution and health, additional time-series investigations in those regions are needed to inform the local authorities as well as add to the global body of evidence on air pollution and health.

Information on specific pollution sources, demographics, and other variables in developing countries of Asia indicates that air pollution affects mortality in this region similarly as in other regions, but that there may be undefined distinct air pollution-associated risks and burden of disease (HEI ISOC 2010). Concentrations of PM and other major pollutants frequently exceed WHO air quality guidelines by substantial margins in Delhi, Chennai, and other Indian cities that have one million inhabitants or more. Delhi is one of the world's megacities with a population of 15.9 million (in 2007) and is expected to be the third largest urban agglomeration by 2015, after Tokyo and Mumbai; Chennai has an estimated 7.5 million people (in 2006).

Several studies in India have evaluated the effects of air pollution on respiratory function and disease, cardiac disease, blood lead concentrations, and eye irritation, as summarized in the Public Health and Air Pollution in Asia– Science Access on the Net database (Health Effects Institute 2008). At the time the current studies started, there was only one time-series study that directly estimated the effect of exposure to air pollution on mortality (Cropper et al. 1997), although there were other studies that estimated mortality impacts based on information from other locations (Srivastava and Kumar 2002; Joseph et al. 2003; Parikh and Hadker 2003; Mukhopadhyay and Forssell 2005).

Cropper and colleagues (1997) used 1991–1994 data collected in Delhi; they found positive, significant associations between PM and mortality in the 5- to 64-year age group of Delhi residents who died from nontraumatic causes as well as from specific respiratory and cardiovascular complications: the excess risk (ER) of mortality associated with a 10-μg/m^3 increase in suspended particulate matter was 0.23 ± 0.1. While the effects estimates were small compared with studies in other countries, the investigators inferred that Delhi may have a greater number of years of life lost because of higher mortality rates in younger age groups (15–44 years), based on the age distribution in India at the time. They and other investigators have suggested that reliance on existing concentration–response functions obtained in populations of developed countries may therefore underestimate the impact of air pollution on population health in developing regions (Cropper et al. 1997; Joseph et al. 2003).

In fact, WHO estimates of deaths and years of life lost suggest that developing countries of Asia face the greatest burden from air pollution worldwide (Ezzati et al. 2002; Cohen et al. 2004). In addition to different distributions of mortality rates across age groups, health status may also differ among Asian populations and those in developed countries. Also, differences in the prevalence of cardiovascular, respiratory, and other diseases exist among populations of East Asia, South Asia, and Southeast Asia, depending on the status of countries' transition toward a more western model (HEI ISOC 2010). There are differences in rates of tobacco smoking and indoor burning of solid fuels that may affect the adult risk of tuberculosis and the childhood risk of morbidity and mortality related to bronchiolitis and pneumonia (also known as acute lower respiratory infection). There is a significant contribution of cardiovascular disease to deaths in India; the population has a high prevalence of hypertension, hypercholesterolemia, dyslipidemia, high glycemic diet, abdominal obesity, and smoking (HEI ISOC 2010).

Outdoor air pollution in India can be traced to a mixture of traditional and modern sources. It is mostly restricted to urban areas, where mobile sources are a major contributor of emissions, and to areas with industrial emissions and power generating facilities (which are mostly coal-based with small amounts of natural gas and oil). A recent comprehensive source apportionment of PM ≤ 2.5 µm in aerodynamic diameter ($PM_{2.5}$) mass for several Indian cities found that fossil fuel combustion contributed between 21% and 57% of $PM_{2.5}$ mass in Delhi, Mumbai, Kolkata, and Chandigarh, whereas biomass combustion contributed 7%–20% in those cities (Chowdhury et al. 2007). On the other hand, biomass contributed a much higher percentage (80%) of black carbon emitted in rural areas of India, affecting a large segment of the population (Gupta et al. 2001).

National and local authorities have been addressing these issues by setting national ambient air quality standards (Critique Table 1) and promoting targeted actions to reduce emissions, such as the use of compressed natural gas and replacing two-stroke with cleaner four-stroke engines to power two- and three-wheeled vehicles, and requirements that new vehicles meet increasingly enhanced auto emissions standards (currently equivalent to EURO 4 standards in the major cities). India's Central Pollution Control Board (CPCB) is targeting 17 major cities for pollution control (CPCB 2006). However, economic growth may overwhelm the introduction of controls and worsen emissions trends; for example, the number of vehicles (including a large proportion of two-wheeled vehicles) is expected to rise from 5 million in 2005 to 250 million in 2025 (Asian Development Bank 2006). Whereas emissions from biomass burning in Asia seem to be stabilizing, combustion of fossil fuels such as coal are on the rise (HEI ISOC 2010).

Considerable challenges remain in conducting epidemiologic assessments of air pollution in Asia owing to data availability and quality, among other factors (HEI ISOC 2010). Widespread deficiencies in the available environmental data and the status of health records from public health departments and private hospitals preclude the regular conduct of such studies. Although some progress has been made to monitor ambient air pollution in selected Asian cities, many environmental agencies lack adequate technical and administrative resources and are understaffed. This has resulted in substantial reliance of Asian air quality decisions on health impact assessments from other regions. This is of particular concern for rapidly developing economies, such as India's, that have distinct population and pollution characteristics. Thus, the current studies in Chennai and Delhi provide welcome additional information about air pollution and health in local Indian populations. As research efforts in India continue, additional results will become available; for example, two cohort studies of long-term effects of exposure to indoor and outdoor air pollution were initiated recently (see Balakrishnan et al. 2011).

Critique Table 1. National Ambient Air Quality Standards in India[a]

Pollutant	Time weighted average	Industrial, Residential, Rural, Other	Ecologically Sensitive	WHO Guideline
PM_{10}	Annual	60	60	20
	24-hour	100	100	50
$PM_{2.5}$	Annual	40	40	10
	24-hour	60	60	25
NO_2	Annual	40	30	40
	24-hour	80	80	100[b]
SO_2	Annual	50	20	—
	24-hour	80	80	20
Ozone	8-hour	100	100	100
	1-hour	180	180	—

[a] Data are µg/m^3. (Sources: Central Pollution Control Board 2009; World Health Organization 2006.) — indicates concentration not specified.

[b] The concentration was measured using a 1-hour averaging time.

STUDY AIMS

The investigator teams in Chennai and Delhi conducted time-series studies on the relationship between daily all-cause mortality and daily concentrations of PM_{10} for the period between January 1, 2002, and December 31, 2004. The India studies did not focus on cause-specific causes of death (e.g., deaths related to respiratory or cardiovascular disease) owing to concerns about the completeness and accuracy of cause-of-death coding. To the extent possible, the investigators followed the PAPA Common Protocol for Coordinated Time-Series Studies of Daily Mortality in Asian Cities (Wong et al. 2010b) with some modifications required because of limitations in the mortality and pollutant monitoring data in Chennai and Delhi. The main features of the common protocol used in the Indian studies were:

Mortality data were provided by local registries of births and deaths in each of the two cities and were coded by trained medical professionals using the WHO's International Classification of Diseases, 10th revision, to exclude deaths from non-natural causes, such as accidents, trauma, or suicide.

Pollutant data for NO_2, SO_2, and PM_{10} were provided by the local government agencies in each city and met local quality control and assurance standards. Exposure metrics used for NO_2, SO_2, and PM_{10} were 24-hour average concentrations. Investigators initially followed an independent, standardized procedure with regard to ensuring both the completeness and representativeness of the average daily exposure of the population. Because of data limitations, the investigators subsequently developed alternative exposure models for their specific location: a novel *zonal* approach in Chennai and centering techniques in Delhi. Gaseous pollutants were included only in the Delhi analyses.

A generalized additive modeling approach was used to obtain the ER of daily mortality associated with daily increases in pollutant concentrations. The investigators fitted quasi-Poisson regression models to the data and carried out sensitivity analyses, such as the inclusion of various degrees of freedom (*df*) for model parameters, different temperature lags, and alternative exposure models.

TECHNICAL SUMMARY OF PART 1: CHENNAI

STUDY OBJECTIVES

The primary objective of this study was to develop models for estimating the association between short-term exposure to criteria air pollutants and daily mortality in Chennai, India, using a time-series approach. The study had the following specific objectives:

1. describe the distribution of daily averages of criteria air pollutant concentrations across all ambient air quality monitoring (AQM) stations located within the greater Chennai area over a 3-year period from 2002 to 2004;
2. examine the patterns and quality of data from individual monitors or combinations of monitors to establish criteria for inclusion or exclusion in alternative exposure series;
3. create a single best exposure series that maximizes the use of available data and adequately represents population exposure;
4. retrieve and organize data on all-cause mortality over the same period;
5. develop a common protocol for a time-series analysis using the exposure and mortality data sets for the three Indian PAPA study sites and identify special requirements for the Chennai data set;
6. develop statistical models to estimate short-term impacts of criteria air pollutants on mortality in Chennai and conduct sensitivity analyses according to criteria laid out in the common protocol; and
7. contribute to the pool of Indian and Asian studies under the PAPA program for coordinated analysis and meta-analysis.

METHODS

Data Collection

Air quality data for 2002–2004 were obtained from the eight stations in the Chennai AQM network operated routinely by the Tamil Nadu State Pollution Control Board. Each station monitors PM_{10}, NO_2, and SO_2 concentrations for an average of 100–140 days per year, either on designated days or every third day, excluding weekends and public holidays. Using preset criteria, the investigators excluded one monitor that was located curbside and below the specified height requirement. Of the seven remaining stations, five had valid data available for the full three-year period. Thus the investigators used data from three suburban *industrial* monitors (AQM 1–3) and two urban monitors (*commercial*: AQM 4, and *residential*: AQM 5).

PM_{10} was measured gravimetrically in 8-hour periods for 24 hours, twice per week. Readings were included only if there were valid data for all three 8-hour periods. NO_2 and SO_2 were measured using a wet chemistry method in 4-hour periods for 24 hours, twice per week. Additional air quality data were obtained from a continuous AQM station that was colocated with AQM 3. Data included 30-minute readings and 8- and 24-hour averages for PM_{10} (using a β-gauge), NO_2, and SO_2 (using chemiluminescence). Because data were available from the continuous station for less than one year, they were used only to evaluate the comparability of the measuring techniques. Meteorologic data on temperature, dew point, relative humidity, wind speed and direction, visibility, and rainfall for 2002–2004 were obtained from the Regional Meteorological Center in Chennai.

Mortality data for 2002–2004 were obtained from the Chennai Corporation and the National Cancer Institute in Chennai (NCI-Chennai). The Chennai Corporation maintains an electronic database of deaths registered across 10 zonal offices (administrative units) within the city, which served as the primary data source. The NCI-Chennai database provided additional information for the noninfant deaths. Records were obtained for all deaths that took place within the Chennai Corporation boundary, including residents of Chennai as well as residents from outside the boundary who died at one of the major hospitals, for example. With a population of 6 million, an average 60 to 100 deaths were recorded every day. After cross-verification of the data sets, the investigators retrieved information on age, sex, residence, date-of-death, place-of-death, and specific cause-of-death.

Air quality and mortality data sets were cleaned according to Quality Assurance / Quality Control (QA/QC) procedures set forth in the common protocol for Indian PAPA investigators.

Development of a Zonal Model of Exposure for PM_{10}

After examining the monitoring schedule at each site and heterogeneity between monitors, the investigators found large spatial variations among pollutant concentrations in residential, commercial, and industrial areas. A high percentage of missing values and low correlation among monitor readings for daily PM_{10} averages prompted the investigators to forgo using the daily average of monitor readings or centering techniques routinely used in time-series studies, and to construct an alternative exposure series. Because of the low resolution of the analytic method for gaseous pollutants, the construction of the exposure series was based on PM_{10} data. NO_2 and SO_2 data were used only in sensitivity analyses (described later).

Several alternative exposure series were created. The first three were single series for the entire city: (1) single-monitor models based on observed data only; (2) multiple-monitor models based on observed data only; and (3) multiple-monitor models based on observed data and imputation of missing data. However, because of spatial variability of exposure series within the city, the primary analyses were carried out using a fourth approach, (4) a *core zonal model* which estimated exposure at the level of 10 individual zones. Daily concentrations for each zone were based on averaging concentrations over the 0.5×0.5 km grid squares within zones, with grid concentrations derived as inverse distance weighted means of pollutant concentrations recorded at one or more of the nearest available monitors.

Statistical Analyses

The investigators broadly followed the common protocol (Appendix B; available on the Web) to construct generalized additive quasi-Poisson models with spline smoothing functions. The analyses were restricted to all-cause mortality, allowing for time trends and controlling for potential confounding by same-day temperature and relative humidity. In a development going beyond the common protocol, degrees of freedom were chosen using a procedure suggested by Dominici and colleagues (2004).

Two sets of sensitivity analyses were conducted. First, performance of the core zonal model was compared with the alternative, nonzonal exposure models using data from AQMs 4 and 5 and the average of the three industrial monitors (AQMs 1, 2, and 3). Some analyses were done with inclusion or exclusion of the industrial monitors. Each of these analyses selected degrees of freedom separately, using the Dominici approach described above. Second, the investigators conducted sensitivity analyses of the core zonal model by varying the degrees of freedom for time, temperature and relative humidity; considering different pollution exposure lags (1 to 3 days); assessing stratification by sex, age, or season; considering inclusion of gaseous pollutants; and removing outliers in the exposure series. Finally, the investigators assessed linearity of the concentration–response curves.

TECHNICAL SUMMARY OF PART 2: DELHI

STUDY OBJECTIVES

The primary objective of this study was to develop models for estimating the association between short-term exposure to criteria air pollutants and daily mortality in Delhi, India, using a time-series approach. The study had the following study objectives:

1. collect ambient air quality data for major criteria pollutants (i.e., PM_{10}, SO_2, and NO_2) for all monitored stations over a 3-year period (2002–2004) and analyze the trend in conjunction with daily changes in weather condition;
2. collect death certificates from the Registrar of Births and Deaths, Municipal Corporation of Delhi (MCD) and New Delhi Municipal Council (NDMC), and code the primary cause of death as per the International classification system;
3. statistically analyze the data for: (a) changes in daily deaths (total deaths excluding death due to accidents) associated with changes in daily air quality levels, and (b) changes in age-specific death due to cardiorespiratory disease associated with changes in daily air quality levels.

METHODS

Data Collection

Air quality data for 2002–2004 were obtained from 9 AQM stations operated by the CPCB and one station operated by The Energy and Resources Institute (TERI). Each station monitors PM_{10}, NO_2, and SO_2 concentrations twice per week for an average of fewer than 100 days per year, except for the station operated by TERI, which collects data on a daily basis. The days in the week that are monitored vary and no measurements are made on weekends and public holidays. All stations are located in urban areas: seven are in residential areas (AQMs 1, 2, 5–7, and 10), three are in industrial areas (AQMs 4, 8, and 9), and one is in an area with mixed commercial and residential use (AQM 3). Measurement methods were identical to those in Chennai. All data were corrected for outlying values. Daily averages for temperature, rainfall, relative humidity, and visibility for 2002–2004 were obtained from the Indian Meteorological Department.

Approximately 900,000 deaths per year are registered in Delhi by three different administrative agencies — MCD, NDMC, and Delhi Cantonment (military). Data from the cantonment were not available but comprised less than 2% of the deaths. NDMC data (about 25% of deaths) were not in electronic format and were entered into spreadsheets by trained field staff. Each death certificate was screened and the cause of death was coded according to the International Classification of Diseases. After cross-verification of the data for duplicate entries, missing information, and misclassification, the investigators retrieved information on age, sex, place-of-death, and cause-of-death. Due to limitations with the cause-of-death coding, statistical analyses were limited to deaths from all natural causes.

Development of a Core Model of Exposure

The investigators noted a high percentage of missing values for the CPCB monitors and high readings for daily pollutant averages at AQM 1 (operated by TERI) compared with the other monitors. This led the investigators to construct an alternative exposure series, instead of using the daily average of monitor readings. The investigators used a data centering approach described by Wong and colleagues (2001); they removed outliers at each monitoring location and then *centered* each daily concentration by subtracting the annual mean for that station and adding the annual mean of all stations.

Statistical Analyses

The investigators followed the common protocol (Appendix B; available on the Web), to construct generalized additive models with penalized spline smoothing functions to control for confounding. The analyses were restricted to all-cause mortality, allowing for time trends and controlling for potential confounding by temperature and relative humidity. Preliminary quasi-Poisson regression analyses were done with 3 to 15 *df* for the time variable and 3 *df* for temperature and relative humidity. The investigators selected the core model based on criteria for autocorrelation, the extent to which deviance was explained, overdispersion, and other statistical criteria. The investigators considered single-, two-, and three-pollutant models for PM_{10}, NO_2, and SO_2, and considered the effects of pollutants on mortality lagged at 0 to 3 days, as well as using a 2-day average of 0- and 1-day lags.

Sensitivity analyses were conducted with various degrees of freedom using natural spline smoothing functions and different lags for temperature (1–3 days and cumulative averages of 0–7 days and 8–14 days). Additional analyses were conducted using different exposure series with pollutant values of individual monitoring locations, the average of 9 monitoring stations (excluding AQM 1), and the average of 10 monitoring stations without the centering technique. Finally, the investigators assessed concentration–response curves by introducing mortality as a smooth function with 3 *df* for the pollutant concentration.

SUMMARY OF KEY RESULTS

AIR QUALITY

Chennai

A summary of air quality and meteorologic data for the five Chennai AQM stations is presented in Critique Table 2. The investigators reported that the concentrations of PM_{10} and NO_2 measured at the industrial and commercial AQM stations were often above the annual national standards; the same was true for SO_2 concentrations at industrial (but not commercial) stations. Concentrations at residential monitors were generally at or below the standards for residential areas. The investigators reported that nearly 30% of the NO_2 and 60% of the SO_2 measurements were around the detection limit of 4 μg/m^3. They therefore focused their main statistical analyses on PM_{10}; NO_2 data were used in sensitivity analyses only.

Delhi

A summary of air quality and meteorologic data for the 10 Delhi AQM stations is presented in Critique Table 2. The investigators noted spatial and seasonal variation in pollutant concentrations. Average PM_{10} and NO_2 concentrations at AQM 1 and 3 were higher than at the other stations, driving up the average concentration across all monitors (Part 2 Investigators' Report Figures 2 and 3). The investigators noted that PM_{10} concentrations were high in May and June owing to dust storms, and in November and December owing to holiday fireworks (e.g., a peak PM_{10} level of 1200 μg/m^3 was observed during the festival of Diwali). During the rainy season (July through September) PM concentrations are generally low. The investigators reported that in residential areas, NO_2 concentrations occasionally

Critique Table 2. Average 24-Hour Pollutant Concentrations, Temperature, and Relative Humidity in Chennai and Delhi[a]

Pollutant	City	Mean	Median	SD	Minimum	Maximum
PM_{10} (µg/m^3)	Chennai	95.4	91.9	31.9	25.1	266.2
	Delhi	221	204	116	30	1015
NO_2 (µg/m^3)	Chennai	26.7	23.2	11.8	7.3	80.0
	Delhi	50	48	15	7	159
SO_2 (µg/m^3)	Chennai	17.4	15.3	8.71	4.2	56.0
	Delhi	9	9	3	4	26
Temp (°C)	Chennai	29.8	30.1	3.1	23.0	40.0
	Delhi	25.5	26.9	—	3.9	45.4
RH (%)	Chennai	75.2	76.0	10.8	44.0	98.0
	Delhi	60.9	62.2	—	6.3	100

[a] Based on valid measurements observed at 5 AQM stations in Chennai and 10 AQM stations in Delhi during 2002–2004. Pollutant concentrations are the average of all AQM stations in each location and thus reflect both spatial and temporal variability. SD indicates standard deviation; Temp indicates temperature; RH indicates relative humidity; — indicates data not provided.

exceeded the national standard but that SO_2 concentrations were well within the standard. Correlations among monitoring stations were moderate to high for PM_{10} but low for NO_2 and SO_2 (Part 2 Investigators' Report Table 5). Correlations between PM_{10} and either NO_2 or SO_2 at individual monitors were low as well (Part 2 Investigators' Report Figure 8).

MORTALITY RISK

Chennai

Using the core zonal model, the investigators reported an increase in the RR for nonaccidental, all-cause mortality of 1.004 (95% confidence interval [CI] = 1.002–1.007) per 10-µg/m^3 increase in PM_{10} concentration on the previous day (Part 1 Investigators' Report Table 10). Alternative, nonzonal exposure series using single- or multiple-monitor models did not appreciably change the RR for mortality.

Additional sensitivity analyses of the core zonal model results showed that there was not much variation in RR between males and females, but that the RRs for the age groups 5–44 years and 45–64 years were slightly higher than for the younger or older age groups. The RR for all-cause mortality associated with PM_{10} was slightly lower at lag 0 (same day) than at lag 1 (previous day), and was elevated at 2- or 3-day lags. No change in RR was found with a 7-day distributed lag for temperature and relative humidity compared with the core model (details for the distributed lag model were not specified). Varying the degrees of freedom for time, temperature, and relative humidity, stratifying by season, or excluding outliers in the exposure series did not appreciably change the RRs (Part 1 Investigators' Report Table 10).

Delhi

Using the core model with centered air quality data, the investigators reported an increase in the RR for nonaccidental, all-cause mortality of 1.0015 (95% CI = 1.0007–1.0023) per 10-µg/m^3 increase in PM_{10} concentration on the same day (Part 2 Investigators' Report Table 8). The RR was slightly lower at 1 to 3 day lags and slightly higher at a cumulative lag of 0–1 days and when PM_{10} concentrations exceeding 400 µg/m^3 were excluded from the analysis. The investigators also reported an increase in RR for daily mortality of 1.0084 (95% CI = 1.0029–1.014) associated with a 10-µg/m^3 increase in NO_2 concentration; there was no evidence of an association between SO_2 concentration and mortality. Using two- and three-pollutant models, the RRs associated with PM_{10} and NO_2 concentrations were slightly attenuated.

Additional sensitivity analyses of the core model showed that the RR associated with PM_{10} was lower in males than females; the RR for the age group 5–44 years was higher than for the other age groups (Part 2 Investigators' Report Table 9). There was considerable variability in RRs associated with PM_{10} among individual monitoring stations, although using all ten stations, the one station operated by TERI (which reported consistently higher PM concentrations), or the nine stations other than the station operated by TERI gave similar RRs (Part 2 Investigators' Report Table 10). Using natural spline smoothing functions for the pollutant effect, RRs were lowest for the smallest number of degrees of freedom (df = 3) and also lower for the largest numbers of degrees of freedom (Part 2 Investigators' Report Figure 15). Using different lags for temperature did not appreciably change the RR, except at longer cumulative lags (0–7 or 8–14 days; Part 2 Investigators' Report Table 11).

HEALTH REVIEW COMMITTEE EVALUATION

The two time-series studies of air pollution and daily mortality in Chennai, India, and Delhi, India, have provided useful additional information on air pollution and health outcomes in developing countries. Previously, only one other time-series study of all-cause mortality and air pollution had been conducted in India (Cropper et al. 1997). Results from Chennai and Delhi suggest a generally similar risk of mortality associated with PM exposure compared with the first four PAPA studies, as well as with multicity studies conducted in South Korea, Japan, Europe, and North America (Critique Table 3). The association of mortality with exposure to NO_2 observed in Delhi was also similar to values reported from other studies in Asia (Wong et al. 2010b; HEI ISOC 2010). The associations were fairly consistent in spite of the fact that concentrations of criteria air pollutants were substantially higher than those observed in the United States and Europe (HEI ISOC 2004, 2010; Wong et al. 2010b); for example, PM_{10} concentrations in Chennai, Delhi, and elsewhere in Asia frequently exceed national ambient air quality standards and WHO air quality guidelines.

The investigator teams (including the team of the third study in Ludhiana) invested considerable effort to establish air quality and mortality data sets, verify the data, and develop a common protocol for their statistical analyses. This protocol was adapted from the common protocol that had been developed for the first wave PAPA studies in Bangkok, Hong Kong, Shanghai, and Wuhan (Wong et al. 2010b). All three teams encountered substantial difficulties in terms of data quality, most importantly the fact that there were many missing values at AQM stations and a lack of reliable information on specific causes of deaths. Given substantial limitations and uncertainties in the Ludhiana data, it was not clear that an interpretable result using air quality data would be possible; the study was not completed. In Chennai and Delhi, the investigator teams focused their analyses on all-cause mortality and developed

Critique Table 3. Percent Excess Risk of All-Natural-Cause Mortality Associated With a 10-µg/m^3 Increase in PM Concentration Among Selected Cities in Asia, Europe, and North America[a]

	Pollutant	% Excess Risk of Mortality (95% CI)	Lag	Reference
Single Cities				
Chennai	PM_{10}	0.4 (0.2 to 0.7)	1 day	Balakrishnan et al. (this report)
Delhi	PM_{10}	0.15 (0.07 to 0.23)	0–1 day	Rajarathnam et al. (this report)
Delhi	SPM	0.23 (0.1)[b]	2 days	Cropper et al. 1997
Ludhiana	Visibility[c]	0.7 (0.1 to 1.2)	1 day	Kumar et al. 2010
Bangkok	PM_{10}	1.25 (0.82 to 1.69)	0–1 day	Wong et al. 2010b
Hong Kong	PM_{10}	0.53 (0.26 to 0.81)	0–1 day	Wong et al. 2010b
Shanghai	PM_{10}	0.26 (0.14 to 0.37)	0–1 day	Wong et al. 2010b
Wuhan	PM_{10}	0.43 (0.24 to 0.62)	0–1 day	Wong et al. 2010b
Multiple Cities				
4 Asian cities	PM_{10}	0.55 (0.26 to 0.85)	0–1 day	Wong et al. 2010b
7 South Korean cities	SPM	0.17 (0.08 to 0.26)	0–1 day	Lee et al. 2000
13 Japanese cities	SPM	0.49 (0.38 to 0.60)	0 days	Omori et al. 2003
Meta-analysis of 82 Asian studies	Various	0.27 (0.12 to 0.42)	Various	HEI ISOC 2010
12 Canadian cities	PM_{10}	0.86 (0.32 to 1.4)	1 day	Katsouyanni et al. 2009
32 European cities	PM_{10}	0.33 (0.22 to 0.44)	1 day	Katsouyanni et al. 2009
89 U.S. cities	PM_{10}	0.29 (0.18 to 0.4)	1 day	Katsouyanni et al. 2009

[a] SPM indicates suspended particulate matter.
[b] The value is the standard deviation.
[c] Because of the extent of missing data, the investigators used reduction of visibility by 1000 meters as a surrogate for air pollution.

city-specific exposure series, using a novel zonal approach and a centering technique, respectively. The remaining sections of this critique evaluate the specific approaches used in each study and conclude with a general discussion of common issues.

EVALUATION OF THE CHENNAI STUDY

Balakrishnan and colleagues concluded that their findings for the effect of PM_{10} on daily mortality in Chennai were in the range reported by other time-series studies in Asia, Europe, and North America (Critique Table 3). They also concluded that as a result of data inadequacies (i.e., lack of data on concentrations of $PM_{2.5}$ and gaseous pollutants, specific causes of death) it remained difficult to apply more sophisticated time-series approaches.

In its independent assessment of the study, the HEI Health Review Committee thought that the investigators had applied an innovative approach for exposure estimation in Chennai, assigning nearest ambient monitor values to each grid cell in the 10 zones to address heterogeneity in PM_{10} concentrations among monitors. The investigators were thoughtful in their handling of exposure data and related scientific issues, and conducted a good set of sensitivity analyses. Because the results from the core zonal model and alternative exposure series were fairly consistent, the Committee thought it would be unlikely that a different approach to exposure estimation would have yielded very different risk estimates, adding to confidence in the risk estimates. However, the Committee identified several caveats regarding the zonal approach.

- When data from one or more monitors are missing, the zonal approach appears to ignore missing data. Thus, the final time series may include spurious jumps in a zone's exposure estimate when data from one of two monitors contributing to a zone's concentration are missing for a period.
- The investigators observed poor correlation among measurements at different sites and large differences among monitors in seasonal trends and other variations over time, which raises the likelihood that there may be important spatiotemporal variation, even within zones. The availability of data from only a few monitoring locations makes it virtually impossible to understand and correct for sources of spatio-temporal variation.
- The monitor *footprint* may be different than the 5–10 km radius assumed by the investigators. Monitors located in close proximity to specific sources, such as the industrial monitors located to the north, may have footprints that are much smaller than 5–10 km, whereas others may have larger footprints. Further analysis is needed to justify what size footprint is appropriate. Treating monitors as having similar footprints introduces additional uncertainty in the exposure estimates.
- The zonal approach does not strictly impose a "weighting of monitors in proportion to the population," as stated by the investigators, but rather in proportion to the number of area grids, unless the population is uniformly distributed across grids.

The Committee noted that the five monitors selected for inclusion in the core model were located in the central-western and northern areas of the city, leaving no direct monitoring coverage of the southern neighborhoods in the model. It could have been useful to use a centering approach as an alternative sensitivity analysis, to provide further understanding of the effect of different exposure approximations on the results.

In its evaluation of the mortality data analysis model, the Committee noted that the investigators used a relatively new, ambitious approach to select degrees of freedom for smoothing of time and meteorology covariates. This approach seeks to minimize PM_{10} coefficient variability and bias using a resampling method. Whether this goal is achievable is uncertain and is the topic of ongoing HEI-funded work by Robins and colleagues (2010). The Committee commented that the degrees of freedom selected for temperature and relative humidity seemed high, which may indicate that those variables are strong predictors of pollution concentrations. The Committee noted that it is not advisable to use different degrees of smoothing for individual monitors, in particular because of the high amount of missing data at a given site; instead it would be more appropriate to apply the degrees of freedom from the model based on all monitors to all sensitivity analyses considering exposure metric. However, because the degrees of freedom were relatively similar across the different approaches (see Appendix F; available on the Web), it is unlikely that the amount of temporal smoothing would substantially affect the results. Further, the Committee thought that using the scale of overdispersion as a measure of goodness-of-fit may not be an optimal approach when comparing models fit to aggregated and zonally disaggregated data, because the mean counts, and hence expected random variation, differ greatly. The investigators did not include an indicator for day-of-week in their analyses, which may be an additional source of uncertainty given the day-of-week patterns in missing data.

The inclusion of a fairly large number of sensitivity analyses was a strength of the study, and broadly gave assurance that the primary measure of association of mortality with PM_{10} was robust to several potential sources of bias. In particular, the robustness of the pollution coefficient to degrees of freedom in the smoothing terms was reassuring. However, the investigation of robustness to temperature effects beyond lag 0 was partial. Temperature has been widely observed to have effects on mortality beyond lag 0 (the temperature lag considered in the core model), and the sensitivity analyses only considered an unspecified distributed lag model over 7 days. If the distributed lag model merely replaced the smooth function of lag 0 temperature with the mean temperature over lags 0–6 days, this would be only a partial exploration of residual confounding by temperature.

EVALUATION OF THE DELHI STUDY

Rajarathnam and colleagues concluded that the changes in mortality associated with short-term exposure to PM_{10} concentrations in Delhi were in agreement with other time-series studies of air pollution and mortality in Asian cities and elsewhere (Critique Table 3). They noted that the magnitude of the effect of air pollution on daily mortality was small but statistically significant after controlling for meteorologic parameters and time trends.

In its independent assessment of the study, the HEI Health Review Committee thought that the use of a centering approach was appropriate for estimating exposure in the context of missing data. However, during the initial review stages the Committee expressed concern about the apparent dominance of the only monitoring station that had continuous data and also showed consistently higher PM_{10} concentrations compared with other stations. That station may not have been representative of population exposures because it was perhaps influenced by industrial sources and may have biased the results. The investigators responded to the Committee's feedback by calculating risk estimates based on 9 versus 10 monitors; excluding the monitor with continuous data appeared to slightly increase the mortality risk estimate. The similarity of the risk estimate of this series with the risk estimate of the series using 10 monitors was reassuring.

The Committee noted that there may have been residual confounding by weather in the main analyses. In their sensitivity analyses of the ER associated with PM_{10}, the investigators replaced the core model smooth function based on lag 0 temperature with smooth functions based on alternative single-day lags or means over multiple-day lag intervals. However, recent insights suggest that more complex distributed lag models may be required to capture temperature effects in many cities (Anderson and Bell 2009). The sensitivity analyses with alternative temperature terms considered in Delhi showed moderate impacts on both the risk estimates and their standard errors. For example, using temperature mean over lags 0–7 days, the investigators reported a smaller ER of 0.10 with $P = 0.059$ (Part 2 Investigators' Report Table 11); back-calculating a standard error yields a 95% CI of about 0.0 to 0.20, which is wider than the CI reported for the core model (0.15; 95% CI = 0.07 to 0.23).

The mortality data also had limitations; as noted by the investigators, not all noninstitutional or domiciliary deaths may have been captured. They encountered additional difficulties in terms of missing entries, multiple entries, misclassification in coding of death, and lacking addresses of deceased persons. However, the large number of daily deaths in Delhi may have offset the loss of statistical power due to missing exposure data.

The investigators conducted a limited but informative set of sensitivity analyses to assess the effects on mortality risk estimates, including the effect of varying time lags, varying confounder models, single- and multiple-pollutant models, age groups, and increasing the temperature lags. For example, changing the degrees of freedom for the smooth function of time from 3 to 15 per year — and using natural splines rather than penalized splines — showed that the highest ERs of mortality occurred at 8 *df* (Part 2 Investigators' Report Figure 15). The ER declined with degrees of freedom below and above 8; the percent increase in risk ranged from about 0.05 to 0.10, indicating sensitivity to the stringency of time control. The results may also be sensitive to temperature lag, as has been discussed for Chennai. Risk estimates based on single-monitor pollution data were sensitive to which single monitor was included in the sensitivity analyses, providing further support for averaging pollutant data across multiple monitors to estimate population exposure in the core model.

The Committee noted some unusual features of the statistical models used. In the core model, the initial rather than final "effective" degrees of freedom were those specified by the authors (8 *df*/year for time, 3 *df*/year for temperature and humidity; see Appendix J, available on the Web). Given that the investigators used a penalized rather than a fixed degree of freedom spline (by specifying the default "fx = F" option in the *s* function in the R package mgcv), the effective degrees of freedom may have been less than these.

In addition, the Committee noted that the degrees of freedom that describe the smoothness of the mortality function against temperature, relative humidity, and PM in sensitivity analyses (for the PM-mortality curve) were specified per year. However, the complexity of these curves (e.g., the temperature-mortality curves) was not expected to increase over the duration of the study. Normal practice is to select degrees of freedom as a function of the number of years only for smooth functions of time. The impact of this specification is likely to be small, given that the splines were penalized.

GENERAL DISCUSSION

One of the important observations from these and other Asian studies is that the ER of all-cause mortality that was associated with an increase of 10 μg/m^3 is fairly comparable to the mortality risk observed in Europe and North America, even though average annual pollutant concentrations in Asia are much higher and there are notable differences in demographics, source mixtures, public health status, and other factors.

Mortality Trends in Asia

Mortality counts in India reported in these two PAPA studies reflect regional patterns found across countries in East Asia, Southeast Asia, and South Asia. For example, in India relatively high mortality was observed in people 45 years and younger (Critique Table 4). The percentages of daily deaths of children under the age of 5 were higher in Chennai (4.9%) and Delhi (8.6%) than those observed in the first wave PAPA studies in Bangkok (2.9%; Vichit-Vadakan et al. 2010), Hong Kong (0.7%; Wong et al. 2010a), Shanghai (0.3%; Kan et al. 2010), and Wuhan (1.5%; Qian et al. 2010). The proportion of daily deaths in people 65 years and older was 45% in Chennai and 34% in Delhi, which was comparable to Bangkok (36%) but considerably lower than in Hong Kong (77%), Shanghai (84%), and Wuhan (72%) (Wong et al. 2010b). The proportionally higher mortality in younger age groups in India suggests that even moderate efforts to improve air quality could have substantial benefits in terms of life expectancy and general health of the working population and associated economic benefits.

The proportions of daily mortality counts for females in both Chennai and Delhi were considerably lower than for males. A similarly higher proportion of male deaths (64%) was reported in Bangkok (Vichit-Vadakan et al. 2010) but not in Shanghai (52.5%; Kan et al. 2010) or Wuhan (55.5%; Qian et al. 2010). The investigators proposed two possible explanations: a higher proportion of males in cities related to jobs and other factors, and deaths of females often being underreported. It remains unclear how the mortality rates differ by sex among the age groups. A recent report by WHO (2009) observed that men between the ages of 15 and 60 have much higher risks of dying than women in the same age category in every region of the world, mainly because of injuries (including violence and conflict) and higher levels of heart disease, factors that may also have contributed to the rates observed in Chennai and Delhi.

Because of the different mortality rates among age groups, the investigators stratified the risk estimates by the age groups listed in Critique Table 4. The ER estimates in Delhi were fairly similar for ages 5–44 (ER = 0.15; 95% CI = 0.05 to 0.25), 45–64 (0.11; 0.02 to 0.20), and ≥ 65 (0.10; −0.06 to 0.26). While the estimates in Chennai were generally higher, the patterns across age groups were also fairly similar: ages 5–44 (0.59; 0.01 to 1.2), ages 45–64 (0.65; 0.2 to 1.1), and ages ≥ 65 (0.31; 0.0 to 0.7).

Other studies have focused more specifically on the oldest age groups (i.e., ≥ 65 and ≥ 75). Vichit-Vadakan and colleagues (2010) reported ER estimates associated with a 10-μg/m^3 increase in PM_{10} that were also fairly similar across their age groupings of 5–44 (ER = 0.9; 95% CI = 0.2 to 1.7), 45–64 (1.1; 0.4 to 1.9), and ≥ 65 (1.5; 0.5 to 2.1);

Critique Table 4. Daily All-Natural-Cause Mortality in Chennai and Delhi by Sex and Age Group for the Period 2002–2004[a]

	Chennai				Delhi			
	Mean ± SD	%	Minimum	Maximum	Mean ± SD	%	Minimum	Maximum
Total	96.8 ± 16.6	100	61	229	222 ± 33	100	126	368
Female	39.1 ± 8.5	40.4	16	96	85 ± 15	38.3	40	142
Male	57.7 ± 11.2	59.6	25	133	139 ± 21	62.6	77	244
0–4	4.8 ± 3.3	4.9	0	21	19 ± 6	8.6	5	39
5–44	21.1 ± 6.3	21.8	5	109	64 ± 11	28.8	28	102
45–64	26.3 ± 6.4	27.5	10	64	61 ± 12	27.5	23	125
≥ 65	43.7 ± 9.9	45.1	19	90	76 ± 24	34.2	28	158

[a] SD indicates standard deviation.

however, the highest ER was observed for individuals 75 years and older (2.2; 1.3 to 3.0). Kan and colleagues (2010) reported no associations between exposure to PM_{10} or other pollutants and mortality in Shanghai for the age groups 5–44 (0.04; −0.52 to 0.59) and 45–64 (0.17; −0.11 to 0.45); their ER estimates were highest for the age group ≥ 65, but still consistent with the estimates they obtained in the other groups (0.26; 0.15 to 0.38). Risk estimates observed in other studies for age groups ≥ 65 and ≥ 75 have been consistently higher than for younger age groups (e.g., Katsouyanni et al. 2009; Wong et al. 2010b), which is clearly different from the findings in the two Indian cities and, to some extent, Bangkok and Shanghai. In these other studies the risk estimates in the oldest age groups are mainly associated with deaths due to cardiovascular and cerebrovascular disease (Katsouyanni et al. 2009; Wong et al. 2010b).

Because there were no reliable daily data on specific causes of deaths that could be used in the Indian studies, one can only speculate on the underlying causes of death in the various age groups and how they may have been affected by exposure to air pollution. Country-specific data compiled by WHO (2008) on causes of death for three age groups (0–14, 15–59, and ≥ 60) indicate that the age-specific percentages of deaths from cardiovascular disease are very similar in India, China, and the United States (Critique Table 5), as are deaths from injuries for ages 15–59 and ≥ 60. However, compared with China and the United States, India has higher percentages of deaths from communicable diseases, including respiratory infections for all age groups as well as diarrheal diseases for ages 0–14 and tuberculosis for ages 15–59.

Evaluation of the Methods

As discussed above, the two Indian studies were more limited in scope than the first wave PAPA studies owing to limitations of the mortality and pollutant data sets. A detailed general evaluation of the common protocol of the first wave PAPA studies is provided in the Integrated Discussion accompanying the combined analyses of time series in Bangkok, Hong Kong, Shanghai, and Wuhan (Wong et al. 2010b). Here, we provide an overview of the main discussion points and indicate differences between the first wave PAPA studies and the two Indian studies.

Air Monitoring and Exposure Estimation Features that were considered important for the first wave PAPA studies were: (1) correlation analyses to identify local sources and decide on exclusion of specific monitors; (2) calculating average monitor concentrations to arrive at an average population exposure (instead of relying on a single monitor); and (3) sensitivity analyses using a centering approach or including certain sets of monitors to evaluate sensitivity of the modeling to different exposure estimates. A recommendation was made for future studies to add colocated monitors to provide additional quality control of the air quality measurements.

These observations and recommendations apply equally to the Indian studies. However, an important difference is that the first wave PAPA studies had access to more complete monitoring data than did the studies in Chennai and Delhi. As described above, the large amount of missing data and poor correlation among monitors, especially in

Critique Table 5. Percentage Estimated Deaths by Cause and Age Group in India, China, and the United States[a]

	0–14 years			15–59 years			≥ 60 years		
Cause of Death	China	India	U.S.	China	India	U.S.	China	India	U.S.
Population (× 10^6)	291	373	62	879	661	185	141	82	48
Communicable diseases									
Diarrheal	9.0	17.3	0.1	0.1	0.7	0.1	0.1	0.8	0.2
Malaria	0	0.3	0	0	0.2	0	0	0	0
Respiratory (infections)	8.9	14.9	1.7	1.3	4.7	1.1	2.5	10.5	2.9
Tuberculosis	0.2	0.3	0	3.9	7.1	0	2.0	1.8	0
Total	65.7	86.1	46.1	9.6	29.1	6.8	5.2	18.2	5.3
Noncommunicable diseases									
Cardiovascular	0.7	1.2	3.2	19.6	21.1	23.0	43.3	43.7	40.6
Respiratory (asthma, COPD)	0.4	0.4	1.7	5.2	6.5	3.0	22.2	12.3	8.1
Total	17.4	8.3	36.3	63.2	47.0	69.1	91.0	76.9	92.0
Injuries	16.8	5.6	17.6	27.2	23.9	24.1	3.8	4.9	2.6

[a] Source: WHO 2008.

Chennai, prompted the Indian teams to develop alternative exposure metrics. Interestingly, effect estimates were more sensitive to individual monitors in Delhi than in Chennai. Because of preset selection criteria and data limitations, only 5 monitors (of which 3 were averaged) contributed data to the Chennai analyses as opposed to 10 monitors in Delhi. In Chennai, all monitors were located in central-western and northern areas of the city, without coverage of southern areas in the model. In Delhi, some monitors were located in relatively close proximity to each other, but it remains unclear if their pollutant concentrations showed higher correlations. Other factors that may have affected the exposure estimations are the type of neighborhoods in which monitors were located (e.g., industrial versus residential) and possibly also monitor-specific patterns in missing data.

In both cities, it would have been helpful to have more detailed information about the possible sources affecting individual monitors, including their proximity to major roads. Another difference was that the Indian studies did not have information on ozone, and the information on NO_2 and SO_2 concentrations was limited to the extent that they could not be included in the Chennai study.

Time Series Modeling Using natural cubic splines as a smooth function of time was considered a reasonable choice for the first wave PAPA studies. However, the constraint that the time-smooth had to have 4–6 *df* per year was considered too restrictive because it may not have completely controlled for temporal confounding, and the data-driven approach to choosing degrees of freedom was considered somewhat problematic. A more stringent approach to control confounding by temperature was also advised, in particular considering distributed lag effects (at least in sensitivity analyses). Detailed recommendations for future studies based on these experiences are listed in the Summary and Conclusions section below.

There are some notable differences with the Indian studies, however. The Delhi and Chennai teams used different approaches, as noted above, but both ended with a maximum of 8 *df*/year in their core models (given that the smooth functions were penalized the effective degrees of freedom may have been smaller), and both reported sensitivity analyses with other choices (up to 15 *df*/year in Delhi and 40 *df*/year in Chennai). The Committee was unconvinced that any approach to degrees of freedom selection could guarantee optimal control for confounding; however, the use of 8 *df*/year is close to values used in some important benchmark studies (e.g., NMMAPS, which used 7 *df*). In addition, the sensitivity analyses revealed relatively little impact to the degree of smoothing.

As noted in the evaluation of each study above, the Committee considered the Indian studies' investigation of sensitivity to alternative temperature lags a step in the right direction, but it remains uncertain whether the powerful and sometimes complex distributed lag associations of temperature with mortality have been controlled in these analyses. Other aspects of weather, for example rainfall, might also be important uncontrolled confounders in these cities.

The Committee also queries the way that degrees of freedom were specified for temperature and relative humidity. As noted in the evaluation of the Delhi study, specifying degrees of freedom of the time smooth based on a "per year" index is sensible and standard procedure. However, doing the same for temperature and humidity (as reported for both Chennai and Delhi) is unusual and yields an exceptionally large number of degrees of freedom (e.g., $6 \times 3 = 18$ for temperature in Chennai). However, since the smoothed functions were penalized in the core models, the effective degrees of freedom in the temperature and relative humidity smooth functions are most likely much smaller than the maximum number.

Concentration–Response Curve The assumption of a linear concentration–response curve is common in air pollution time-series studies. Results from the first wave PAPA studies suggest generally linear and increasing patterns across all four cities and pollutants, with some exceptions (e.g., ozone in Shanghai). Sensitivity analyses did not suggest a presence of thresholds or any other clear evidence against a linear concentration–response relationship. However, considering the high concentrations of pollutants, it would be interesting to know whether there may be a *leveling off* of the concentration–response relationship at very high PM concentrations. The results from the first wave PAPA studies were not indicative of such leveling off, but are compatible with it.

This comment also broadly applies to the two Indian studies. The PM_{10} concentration–response curve for Delhi (Part 2 Investigators' Report Figure 14) showed suggestive evidence for leveling off at concentrations above 400 $\mu g/m^3$ and was compatible with leveling off at concentrations above 200 $\mu g/m^3$. The concentration–response curve in Chennai (Part 1 Investigators' Report Figure 22) is either consistent with a linear fit (for the dataset that excludes the highest PM concentrations), or shows a *dip* around 250 $\mu g/m^3$, above which the curve resumes a steady increase. However, the confidence intervals are wide, the data above 200 $\mu g/m^3$ are limited, and the realized degree of smoothing in these two pairs of curves is uncertain. Therefore, the observed functional form in the Chennai curve should not be over-interpreted.

SUMMARY AND CONCLUSIONS

The two time-series studies of air pollution and daily mortality in Chennai, India, and Delhi, India, have provided useful additional information on air pollution and health outcomes in developing countries. Results from Chennai (0.4% increase in risk per 10-μg/m^3 increase in PM_{10} concentration) and Delhi (0.15% increase in risk per 10-μg/m^3 increase in PM_{10} concentration) suggest a generally similar risk of mortality associated with PM_{10} exposure compared with the first four PAPA studies, as well as with multicity studies conducted in South Korea, Japan, Europe, and North America.

The broad general consistency of the results of the two Indian studies described here and in other Asian time-series studies of mortality with those in Europe and North America is reassuring. It suggests that the continued use of data from Western cohort studies to estimate the Asian burden of disease attributable to short-term exposure to air pollution is defensible. However, developing Asia currently differs from the United States and Europe with regard to energy use, air quality, and population health, which are also dynamically changing. The Indian studies highlight that regional differences in demographics (in particular, age structure and general health status of the population) may affect health outcomes of interest. Thus, estimates of the risk of mortality associated with air pollution that are based on even the most carefully executed U.S. studies must be used with appropriate caveats.

Given the data limitations faced by the investigator teams, they are to be commended for making the most of limited resources. They have blazed a trail for improved quality epidemiological studies of air pollution in India. However, considerable uncertainties remain due to data limitations, potential residual confounding, and potential methodological sensitivity, all of which will need to be revisited in any future epidemiological studies. As the investigators pointed out, data limitations prevented a number of more in-depth analyses standard in time-series studies, for example, of specific causes of death or — in Chennai — different pollutant models. Such detailed analyses will only be possible once more detailed, consistent air pollution monitoring and health record collection are implemented.

Based on the experience from the first wave PAPA studies and these two Indian studies, the Committee has made the following recommendations for future studies:

- The analysis strategy should avoid reliance on identification of an *optimal* confounder model, since no such strategy can guarantee against residual confounding. Instead, the protocol should specify an a priori primary analysis, supplement this with a comprehensive set of analyses of sensitivity to model construction, and ensure inclusion in models of known determinants of fluctuations in mortality.
- Extensive analyses of sensitivity to confounder control should be undertaken. These are often overlooked for *second order* investigations, such as putative effect modification, concentration–response modeling, or multipollutant models.
- Weather is often a powerful determinant of mortality at lags extending well beyond zero and is associated with pollution. As such, it is a strong potential confounder and needs careful modeling, including consideration of distributed lag models, in main and sensitivity analyses.
- Assessment of concentration–response relationships, as was done in these studies, should be included as part of the sensitivity analyses.
- The use of zonal exposure series in cities that exhibit poorly correlated concentration series among different monitors is worth further investigation. Potentially important data features and methodological limitations remain to be explored, however, in particular because poor correlation among available monitors suggests high spatiotemporal variability, which may be considerable even within zones.
- Investigators should strive for complete transparency in their reporting. One step in this direction is to clearly present the statistical models that were fitted to the data, for example by including annotated R code in an appendix, as was done in part by these investigators.

The methodology applied in the PAPA time-series studies can provide a stronger foundation for further research in developing Asia. The PAPA studies have added to a growing number of time-series studies across Asia (HEI ISOC 2010). These studies have been consistent in showing increases in daily mortality associated with short-term exposure to measured air pollutants, but have been conducted largely in China and South Korea. The lack of data on air quality and mortality, especially cause-specific mortality, remain major impediments to conducting such studies in many parts of developing Asia. As a result, major population centers in South and Southeast Asia (India, Pakistan, Vietnam, Philippines, Indonesia, and Malaysia) remain understudied — although the two PAPA studies in India and two additional HEI-funded studies in Vietnam are starting to fill in some of those gaps. Expanded, coordinated multicity studies conducted across Asia could provide more definitive answers if they are designed and analyzed consistently with the additional methodologic improvements noted above and given rigorous quality control of air quality and health data.

ACKNOWLEDGMENTS

The Health Review Committee thanks the ad hoc reviewers for their help in evaluating the scientific merit of the Investigators' Reports. The Committee is also grateful to Sumi Mehta for her oversight of the studies, Amara Ezeamama for her oversight of the review process, and Annemoon van Erp, Marian Berkowitz, and Tiffany North for their assistance in preparing the studies' Critique; to Carol Moyer for the science editing of the Reports and the Critique; and to Fred Howe, Suzanne Gabriel, Bernard Jacobson, and Flannery Carey McDermott for their roles in publishing these Research Reports.

REFERENCES

Anderson BG, Bell ML. 2009. Weather-related mortality: How heat, cold, and heat waves affect mortality in the United States. Epidemiol 20:205–213.

Asian Development Bank. 2006. Energy Efficiency and Climate Change Considerations for On-Road Transport in Asia. Asian Development Bank, Manila, Philippines. Available from *www.cleanairnet.org/caiasia/1412/articles-70656_finalreport.pdf*. Accessed 10/1/2010.

Balakrishnan K, Dhaliwal RS, Shah B. 2011. Integrated urban–rural frameworks for air pollution and health-related research in India: The way forward. Environ Health Perspect 119:a12–a13.

Bell ML, Samet JM, Dominici F. 2004. Time-series studies of particulate matter. Ann Rev Pub Health 25:247–280.

Central Pollution Control Board. 2006. Air Quality Trends and Action Plan for Control of Air Pollution from Seventeen Cities. National Ambient Air Quality Monitoring Series NAAQMS/29/2006-07. New Delhi, India. Available from *www.indiaenvironmentportal.org.in/files/airquality17cities.pdf*. Accessed 9/21/2010.

Central Pollution Control Board. 2009. National Ambient Air Quality Standards. Notification B-21906/20/90/PCI-L. Central Pollution Control Board, New Delhi, India. Available from *www.cpcb.nic.in/National_Ambient_Air_Quality_Standards.php*. Accessed 9/20/2010.

Chowdhury Z, Zheng M, Schauer JJ, Sheesley RJ, Salmon LG, Cass GR, Russell AG. 2007. Speciation of ambient fine organic carbon particles and source apportionment of $PM_{2.5}$ in Indian cities. J Geophys Res 112:D15303.

Cohen AJ, Anderson HR, Ostro B, Pandey KD, Krzyzanowski M, Künzli N, Gutschmidt K, Pope CA III, Romieu I, Samet JM, Smith KR. 2004. Urban air pollution. In: Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors (Ezzati M, Lopez AD, Rodgers A, Murray CJL eds.). World Health Organization, Geneva, Switzerland.

Cropper M, Simon NB, Alberini A, Arora S, Sharma PK. 1997. The health benefits of air pollution control in Delhi. Am J Agr Econ 79:1625–1629.

Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. 2003. Revised Analyses of the National Morbidity, Mortality, and Air Pollution Study. Part II: Mortality Among Residents of 90 Cities. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston, MA.

Dominici F, McDermott A, Hastie T. 2004. Improved semiparametric time series models of air pollution and mortality. J Am Stat Assoc 468:938–948.

Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJ; Comparative Risk Assessment Collaborating Group. 2002. Selected major risk factors and global and regional burden of disease. Lancet 360:1347–1360.

Gupta PK, Prasad VK, Kant Y, Sharma C, Ghosh AB, Sharma MC, Sarkar AK, Jain SL, Tripathi OP, Sharma RC, Badarinath KVS, Mitra AP. 2001. Study of trace gases and aerosol emissions due to biomass burning at shifting cultivation sites in East Godavari District (Andhra Pradesh) during INDOEX IFP-99. Curr Sci Ind 80:186–196.

Gurjar BR, Butler TM, Lawrence MG, Lelieveld J. 2008. Evaluation of emissions and air quality in megacities. Atmos Environ 42:1593–1606.

Gurjar BR, van Aardenne JA, Lelieveld J, Mohan M. 2004. Emission estimates and trends (1990–2000) for megacity Delhi and implications. Atmos Environ 38:5663–5681.

Health Effects Institute. 2003. Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report. Health Effects Institute, Boston, MA.

Health Effects Institute. 2008. Public Health and Air Pollution in Asia – Science Access on the Net. Table 10, Research in India. Available at *www.healtheffects.org/Asia/Table10_India.pdf*. Accessed 10/4/2010.

HEI International Scientific Oversight Committee. 2004. Health Effects of Outdoor Air Pollution in Developing Countries of Asia: A Literature Review. Special Report 15. Health Effects Institute, Boston, MA.

HEI International Scientific Oversight Committee. 2010. Outdoor Air Pollution and Health in the Developing Countries of Asia: A Comprehensive Review. Special Report 18. Health Effects Institute, Boston, MA.

Joseph A, Ad S, Srivastava A. 2003. PM(10) and its impacts on health—A case study in Mumbai. Int J Environ Health Res 13:207–214.

Kan H, Chen B, Zhao N, London SJ, Song G, Chen G, Zhang Y, Jiang L. 2010. Part 1. A time-series study of ambient air pollution and daily mortality in Shanghai, China. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Katsouyanni K, Samet JM, Anderson HR, Atkinson R, Le Tertre A, Medina S, Samoli E, Touloumi G, Burnett RT, Krewski D, Ramsay T, Dominici F, Peng RD, Schwartz J, Zanobetti A. 2009. Air pollution and health: A European and North American Approach (APHENA). Research Report 142. Health Effects Institute, Boston, MA.

Katsouyanni K, Touloumi G, Samoli E, Gryparis A, Le Tertre A, Monopolis Y, Rossi G, Zmirou D, Ballester F, Boumghar A, Anderson HR, Wojtyniak B, Paldy A, Braunstein R, Pekkanen J, Schindler C, Schwartz J. 2001. Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. Epidemiology 12:521–531.

Kumar R, Sharma SK, Thakur JS, Lakshmi PVM, Sharma MK, Singh T. 2010. Association of air pollution and mortality in the Ludhiana City of India: A time-series study. Indian J Public Health 54:98–103.

Lee JT, Kim H, Hong YC, Kwon HJ, Schwartz J, Christiani DC. 2000. Air pollution and daily mortality in seven major cities of Korea, 1991–1997. Environ Res 84:247–254.

Mukhopadhyay K, Forssell O. 2005. An empirical investigation of air pollution from fossil fuel combustion and its impact on health in India during 1973–1974 to 1996–1997. Ecol Econ 55:235–250.

Omori T, Fujimoto G, Yoshimura I, Nitta H, Ono M. 2003. Effects of particulate matter on daily mortality in 13 Japanese cities. J Epidemiol 13:314–322.

Parikh JK, Hadker N. 2003. Economic impacts of urban air pollution: Valuation for Mumbai, India. Int J Environ Pollut 19:498–515.

Pope CA III, Dockery DW. 2006. Health effects of fine particulate air pollution: Lines that connect. J Air Waste Manage Assoc 56:709–742.

Qian Z, He Q, Lin H-M, Kong L, Zhou D, Liang S, Zhu Z, Liao D, Liu W, Bentley CM, Dan J, Wang B, Yang N, Xu S, Gong J, Wei H, Sun H, Qin Z. 2010. Part 2. Association of daily mortality with ambient air pollution, and effect modification by extremely high temperature in Wuhan, China. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Robins JM, Zhang P, Ayyagari R, Logan R, Ling L, Tchetgen Tchetgen E, van der Vaart A. 2010. A new approach to semiparametric regression in time-series studies of the health effects of pollutants: Results and comparison with previous approaches. HEI 2010 Annual Conference Presentation. Health Effects Institute, Boston, MA.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, 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.

Srivastava A, Kumar R. 2002. Economic valuation of health impacts of air pollution in Mumbai. Environ Monit Assess 75:135–143.

Vichit-Vadakan N, Vajanapoom N, Ostro B. 2010. Part 3. Estimating the effects of air pollution on mortality in Bangkok, Thailand. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Wong C-M, Ma S, Hedley AJ, Lam TH. 2001. Effect of air pollution on daily mortality in Hong Kong. Environ Health Perspect 109:335–340.

Wong C-M, Thach TQ, Chau PYK, Chan EKP, Chung RY-N, Ou C-Q, Yang L, Peiris JSM, Thomas GN, Lam T-H, Wong T-W, Hedley AJ. 2010a. Part 4. Interaction between air pollution and respiratory viruses: Time-series study of daily mortality and hospital admissions in Hong Kong. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Wong C-M on behalf of the PAPA teams: Bangkok, Hong Kong, Shanghai, and Wuhan. 2010b. Part 5. Public Health and Air Pollution in Asia (PAPA): A combined analysis of four studies of air pollution and mortality. In: Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities. Research Report 154. Health Effects Institute, Boston, MA.

Wong C-M, Vichit-Vadakan N, Kan H, Qian Z. 2008. Public Health and Air Pollution in Asia (PAPA): A multicity study of short-term effects of air pollution on mortality. Environ Health Perspect 116:1195–1202.

World Health Organization. 2006. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. WHO/SDE/PHE/OEH/06.02. WHO Press, Geneva, Switzerland.

World Health Organization. 2008. The Global Burden of Disease: 2004 Update. Death and DALY estimates for 2004 by cause for WHO Member States. Available at *www.who.int/healthinfo/global_burden_disease/estimates_country/en/index.html*. Accessed 10/6/2010.

World Health Organization. 2009. Global Health Risks. Mortality and Burden of Disease Attributable to Selected Major Risks. WHO Press, Geneva, Switzerland. Available at *www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf*. Accessed 10/6/2010.

RELATED HEI PUBLICATIONS

Number	Title	Principal Investigator	Date*
Research Reports			
154	Public Health and Air Pollution in Asia (PAPA): Coordinated Studies of Short-Term Exposure to Air Pollution and Daily Mortality in Four Cities		2010
	Part 1. A Time-Series Study of Ambient Air Pollution and Daily Mortality in Shanghai, China	H. Kan	
	Part 2. Association of Daily Mortality with Ambient Air Pollution, and Effect Modification by Extremely High Temperature in Wuhan, China	Z. Qian	
	Part 3. Estimating the Effects of Air Pollution on Mortality in Bangkok, Thailand	N. Vichit-Vadakan	
	Part 4. Interaction Between Air Pollution and Respiratory Viruses: Time-Series Study of Daily Mortality and Hospital Admissions in Hong Kong	C.-M. Wong	
	Part 5. Public Health and Air Pollution in Asia (PAPA): A Combined Analysis of Four Studies of Air Pollution and Mortality	C.-M. Wong	
142	Air Pollution and Health: A European and North American Approach (APHENA)	K. Katsouyanni and J.M. Samet	2009
140	Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality	D. Krewski	2009
139	Effects of Long-Term Exposure to Traffic-Related Air Pollution on Respiratory and Cardiovascular Mortality in the Netherlands: The NLCS-AIR Study	B. Brunekreef	2009
137	The Influence of Improved Air Quality on Mortality Risks in Erfurt, Germany	A. Peters	2009
123	Time-Series Analysis of Air Pollution and Mortality: A Statistical Review	F. Dominici	2004
98	Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany	H.-E. Wichmann	2000
	Part I. Role of Particle Number and Particle Mass		
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	M.S. Goldberg	2000
94	The National Morbidity, Mortality, and Air Pollution Study		
	Part II. Morbidity and Mortality from Air Pollution in the United States	J.M. Samet	2000
	Part III. Concentration–Response Curves and Thresholds for the 20 Largest US Cities	M.J. Daniels	2004
	Part IV. Hierarchical Bivariate Time-Series Models — A Combined Analysis of PM_{10} Effects on Hospitalization and Mortality	F. Dominici	2005

Continued

** Reports published since 1998.*
Copies of the reports can be obtained from the Health Effects Institute, and many are available at www.healtheffects.org.

RELATED HEI PUBLICATIONS

Number	Title	Principal Investigator	Date*
Special Reports			
18	Outdoor Air Pollution and Health in the Developing Countries of Asia: A Comprehensive Review		2010
17	Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects		2010
15	Health Effects of Outdoor Air Pollution in Developing Countries of Asia: A Literature Review		2004
	Revised Analyses of Time-Series Studies of Air Pollution and Health		2003
	Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality: A Special Report of the Institute's Particle Epidemiology Reanalysis Project		2000
HEI Communications			
13	Public Health and Air Pollution in Asia (PAPA): Key Results from Bangkok, Hong Kong, Shanghai, and Wuhan		2008
Web			
	Public Health and Air Pollution in Asia: Science Access on the Net (PAPA-SAN). Database available at *www.healtheffects.org/Asia/papasan-home.htm*		2008

** Reports published since 1998.*
Copies of the reports can be obtained from the Health Effects Institute, and many are available at www.healtheffects.org.

HEI BOARD, COMMITTEES, and STAFF

Board of Directors

Richard F. Celeste, Chair *President, Colorado College*

Sherwood Boehlert *Of Counsel, Accord Group; Former Chair, U.S. House of Representatives Science Committee*

Enriqueta Bond *President Emeritus, Burroughs Wellcome Fund*

Purnell W. Choppin *President Emeritus, Howard Hughes Medical Institute*

Michael T. Clegg *Professor of Biological Sciences, University of California–Irvine*

Jared L. Cohon *President, Carnegie Mellon University*

Stephen Corman *President, Corman Enterprises*

Gowher Rizvi *Vice Provost of International Programs, University of Virginia*

Linda Rosenstock *Dean, School of Public Health, University of California–Los Angeles*

Henry Schacht *Managing Director, Warburg Pincus; Former Chairman and Chief Executive Officer, Lucent Technologies*

Warren M. Washington *Senior Scientist, National Center for Atmospheric Research; Former Chair, National Science Board*

Archibald Cox, Founding Chair *1980–2001*

Donald Kennedy, Vice Chair Emeritus *Editor-in-Chief Emeritus,* Science; *President Emeritus and Bing Professor of Biological Sciences, Stanford University*

Health Research Committee

David L. Eaton, Chair *Associate Vice Provost for Research and Director, Center for Ecogenetics and Environmental Health, School of Public Health, University of Washington–Seattle*

David T. Allen *Gertz Regents Professor in Chemical Engineering; Director, Center for Energy and Environmental Resources, University of Texas–Austin*

David E. Foster *Phil and Jean Myers Professor, Department of Mechanical Engineering, Engine Research Center, University of Wisconsin–Madison*

Uwe Heinrich *Professor, Medical School Hannover; Executive Director, Fraunhofer Institute for Toxicology and Experimental Medicine, Hanover, Germany*

Grace LeMasters *Professor of Epidemiology and Environmental Health, University of Cincinnati College of Medicine*

Sylvia Richardson *Professor of Biostatistics, Department of Epidemiology and Public Health, Imperial College School of Medicine, London, United Kingdom*

Richard L. Smith *Director, Statistical and Applied Mathematical Sciences Institute, University of North Carolina–Chapel Hill*

James A. Swenberg *Kenan Distinguished Professor of Environmental Sciences, Department of Environmental Sciences and Engineering, University of North Carolina–Chapel Hill*

HEI BOARD, COMMITTEES, and STAFF

International Scientific Oversight Committee

Frank Speizer, Chair *Professor of Environmental Science, Department of Environmental Health, Harvard Medical School*

H. Ross Anderson *Professor of Epidemiology and Public Health, Division of Community Health Sciences, St. George's, University of London, and Medical Research Council–Health Protection Agency Centre for Environment and Health, London, United Kingdom*

Michael Brauer *Director, School of Environmental Health, University of British Columbia, Canada*

Bingheng Chen *Professor, School of Public Health, Fudan University, Shanghai, China*

Kenneth L. Demerjian *Director and Professor, Atmospheric Sciences Research Center and Department of Earth and Atmospheric Science, University at Albany, State University of New York*

Jiming Hao *Professor of Environmental Science, Tsinghua University; Dean, Institute of Environmental Science and Engineering, Beijing, China*

Anthony J. Hedley *Honorary Professor, School of Public Health, University of Hong Kong, Hong Kong*

Jitendra N. Pande *Senior Consultant in Medicine and Chest Diseases, Sitaram Bhartia Institute of Science and Research, New Delhi, India*

C. Arden Pope III *Mary Lou Fulton Professor of Economics, Brigham Young University*

Kirk R. Smith *Professor of Global Environmental Health and Chair, Graduate Group in Environmental Health Studies, School of Public Health, University of California–Berkeley*

Mark J. Utell *Professor of Medicine and Environmental Medicine, University of Rochester Medical Center*

Paul Wise *Richard E. Behrman Professor of Child Health and Society; Professor of Pediatrics, Stanford University Medical School*

Health Review Committee

Homer A. Boushey, Chair *Professor of Medicine, Department of Medicine, University of California–San Francisco*

Ben Armstrong *Reader in Epidemiological Statistics, Public and Environmental Health Research Unit, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, United Kingdom*

Michael Brauer *Professor, School of Environmental Health, University of British Columbia, Canada*

Bert Brunekreef *Professor of Environmental Epidemiology, Institute of Risk Assessment Sciences, University of Utrecht, the Netherlands*

Mark W. Frampton *Professor of Medicine and Environmental Medicine, University of Rochester Medical Center*

Stephanie London *Senior Investigator, Epidemiology Branch, National Institute of Environmental Health Sciences*

Armistead Russell *Georgia Power Distinguished Professor of Environmental Engineering, School of Civil and Environmental Engineering, Georgia Institute of Technology*

Lianne Sheppard *Professor of Biostatistics, School of Public Health, University of Washington–Seattle*

HEI BOARD, COMMITTEES, and STAFF

Officers and Staff

Daniel S. Greenbaum *President*
Robert M. O'Keefe *Vice President*
Rashid Shaikh *Director of Science*
Barbara Gale *Director of Publications*
Jacqueline C. Rutledge *Director of Finance and Administration*
Helen I. Dooley *Corporate Secretary*

Kate Adams *Staff Scientist*
Aaron J. Cohen *Principal Scientist*
Maria G. Costantini *Principal Scientist*
Philip J. DeMarco *Compliance Manager*
Suzanne Gabriel *Editorial Assistant*
Hope Green *Editorial Assistant (part time)*
L. Virgi Hepner *Senior Science Editor*
Anny Luu *Administrative Assistant*
Francine Marmenout *Senior Executive Assistant*
Sumi Mehta *Senior Scientist*
Nicholas Moustakas *Policy Associate*
Hilary Selby Polk *Senior Science Editor*
Sarah Rakow *Science Administrative Assistant*
Robert A. Shavers *Operations Manager*
Geoffrey H. Sunshine *Senior Scientist*
Annemoon M.M. van Erp *Senior Scientist*
Katherine Walker *Senior Scientist*
Morgan Younkin *Research Assistant*

HEALTH
EFFECTS
INSTITUTE

101 Federal Street, Suite 500
Boston, MA 02110, USA
+1-617-488-2300
www.healtheffects.org

RESEARCH
REPORT

Number 157
March 2011