Chapter 6

Discussion and Conclusions

6.1 Discussion

With increasing rates of urbanization, population growth and development, the impacts of worsening air quality on health is becoming a serious concern in cities of all rapidly developing countries, including India. Health impact and burden of disease assessments in the past relied heavily on results from exposure-response studies conducted in North America and Europe. In the recent years there has been an increasing base of research studies from Asia that have addressed uncertainties introduced by higher pollutant concentrations, different pollutant mixes and underlying disease conditions commonly prevalent in the region. The present study adds important new information to this growing pool of regional evidence. In particular, it fills important gaps created by the paucity of information from India. A detailed discussion of study methods, data quality and modeling approaches used in the study (in relation to what has been used in previous studies) is provided below. The chapter concludes with implications of study results for policy concerning air quality and health in India as well as future research.

6.1.1 Exposure Data

This study has collected and summarized information on PM$_{10}$ concentrations, from multiple monitors operated by TNPCB, over a 7 year period in Chennai city. The city average concentrations of 24-hr PM$_{10}$ ranged from 69.8 to 100.2 $\mu g/m^3$ for the years 2002-2008. The increasing trend in PM$_{10}$ concentrations in commercial areas and residential areas of Chennai (based on monitors at Anna Nagar and Vallalar Nagar shown in Figure 5.1)
indicate contributions from the rapidly expanding vehicular fleets, while the relatively stable concentrations in industrial areas (based on monitors at Kathivakkam, Thiruvottiyur and Manali shown in Figure 5.1) are presumably indicative of the stringent Governmental regulations and control of industrial emissions. Chennai’s annual average of around 80 \(\mu g/m^3\) for PM\(_{10}\) are amongst the lowest recorded in Indian cities but are comparable (and in some cases higher) than what has been reported from other Asian studies (HEI 2010).

The study used data on air pollution concentrations and meteorological parameters reported from routine monitoring networks similar to other multicity studies in North America, Europe and Asia (Katsouyanni & Samet, 2009; C. M. Wong et al., 2010b; J. T. Lee et al., 2000; Omori et al., 2003; HEI, 2010). Since these networks are subject to continuous and uniform standards of quality assurance, biases introduced by variations in protocols across monitoring sites are substantively reduced. Methods developed using such datasets also become more readily applicable in other cities within the country.

A detailed examination of spatio-temporal variation of pollutant concentrations (i.e. PM\(_{10}\) in this study) across monitors allowed an exploration of the representativeness and adequacy of data from individual monitors for estimating population exposures. Many previous time series studies relied on use of background monitors and used datasets spanning over several years. The data from monitors thus were fairly well correlated and missing-ness could be minimized by pooling data across available monitors in these studies (C. M. Wong et al., 2001). In contrast, the Chennai dataset was characterized by a lack of correlation between monitors and when combined with a high percentage of missing days for each monitor, it created a major challenge for creating a representative exposure series. Before developing statistical methods to address this challenge, biases from measurement errors had to be ruled out. For this details were gathered on the monitor siting protocols from CPCB. The NAMP protocol followed throughout the country requires the placement of monitors at sites (such as industrial, commercial and residential sites) that are likely to differ in concentrations based on human activity. Also, the height of the sampler is set much lower than the recommended height for true background samplers in order to capture variations across city zones and possibly identify local control strategies. The QA and QC information provided by TNPCB excludes the possibility that systematic measurement error across monitors or days was responsible for the observed lack of correlation. The small footprints of the monitoring stations combined with the

---


---
differences in source strengths are thus likely to be the primary reason for the site-to-site variation.

Lack of hourly data and sensitivity of gas measurement methods used by TNPCB did not permit inclusion of data on gases in single- or multiple-pollutant models as has been done in recent multi-city studies (C. M. Wong et al., 2010b) in Asia.

6.1.2 Health Data

Health data too was collected from the databases of the Chennai Corporation and medical records of two large Governmental hospitals. Classification of causes of death using ICD-9 or ICD-10 found in these records is similar to what has been reported in previous Asian and North American studies. Relying on broad classifications such as due to cardiovascular or respiratory causes has also been shown to quite reliable, when compared using sub-classifications such as stroke or COPD in other time-series studies (HEI, 2010).

6.1.3 Modeling Methods

This study explored several approaches to improve outputs of models for estimating the exposure-response relationship between PM10 and mortality or morbidity through time series analyses. First, it developed a core model that used relatively simple methods of exposure estimations (through use of daily city level averages of available data) while addressing nonlinear confounding by weather parameters and uncertainties introduced by model parameters through extensive sensitivity analyses, similar to approaches taken in recent multi-city Asia studies in Shanghai, Bangkok, Wuhan and Hong Kong (C. M. Wong et al., 2008)

Next, it explored alternative ways to construct exposures using alternative combinations of monitors both to optimize abilities to utilize a high percentage of the available exposure data with mortality data as well as address spatial heterogeneity. Single monitor models based on data from single monitors (at Anna Nagar and Vallalar Nagar) and the industrial monitors (as an average of available industrial monitors) utilised ∼48, 50%, and 96% of all days in the study period, respectively. By comparison, the multiple-monitor 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 monitors, respectively. The zonal model used pollutant concentration and mortality data for 100% of

---

the days in the study period (Note: weekends and holidays could not be included in the total number of days in the study period for any of the models). These models utilized methods built on earlier datasets on Chennai (Balakrishnan et al., 2011) but were able use data over an extended period of time to provide more refined and robust effects estimates for all-cause mortality.

Finally, the development of an autoregressive Poisson model to address serial correlation in outcomes in the face of missing exposure data together with simulation studies to validate the model has significantly enhanced the ability to address model uncertainties. Ad-hoc methods, such as centering (C. M. Wong et al., 2001), have been used by researchers in other Asia studies to tackle small amounts of missing information but they essentially require monitor readings throughout the study locale to be highly correlated, an assumption which does not hold for the Chennai dataset (as evidenced by the dataset used in this study and previous studies (Balakrishnan et al., 2011)). An additional problem which needed to be addressed for this dataset is a non-linear time trend in the log-mortality rate and additional non-linear confounding by meteorological parameters. Such confounding has been addressed by using smoothing and regression splines by Hastie and Tibshirani (1990); Ruppert et al. (2003). This study used a seamless likelihood based approach for fitting Poisson auto-regressive models using random effects to model missing co-variates as well as to account for simultaneous non-linear confounding. The Monte Carlo Expectation Maximization (MCEM) algorithm as discussed by Dempster et al. (1977) was used for parameter estimation while the Metropolis-Hastings algorithm (Chib & Greenberg, 1995) and Adaptive Rejection Sampling (ARS) (Gilk et al., 1995) have been used to implement the E-step. Simulation results indicate that the efficiency of estimation depends essentially on the amount of missing data and to some extent on the between monitor correlation. The availability of multiple sets of exposure series allowed an exploration of the uncertainties under alternative conditions of missing-ness and monitor correlation. Under conditions of low correlation observed in the Chennai dataset, the simulation studies demonstrate the utility of the zonal model to minimize exposure uncertainties. There has been virtually no precedence for the application of such methods in time-series analyses concerning air pollution and health effects. These methods are thus likely to be useful in numerous other cities where similar data challenges and/or exposure heterogeneities prevail.

6.1.4 Comparison of effects estimates from this study with results from previous studies

The study presents effects estimates for excess risks of all-natural cause mortality, mortality from cardio-vascular and respiratory causes and hospital admissions from cardio-vascular and respiratory illness in relation to increases in PM$_{10}$ exposure using a 7 year dataset for Chennai. Previous studies in India have focused on one or a few such endpoints using datasets covering much shorter periods with a limited range of sensitivity analyses (Cropper et al., 1997; Pande et al., 2002). The present study thus provided estimates for a much broader range of outcomes using extended datasets together with results from an extensive range of sensitivity analyses.

The core model using city average PM$_{10}$ exposures estimated a 0.67% (95% CI 0.31-1.04%) increase or RR of 1.007 (95% CI 1.003-1.010) for daily all-natural cause mortality, a 0.94% (95% CI 0.41-1.46%) increase or RR of 1.009 (95% CI 1.004-1.015) for daily mortality due to cardiovascular and respiratory causes per 10$\mu g/m^3$ increase in PM$_{10}$ exposure (expressed as a 4-day cumulative average) and a 1.39% (95% CI 0.3-2.37%) increase or an RR of 1.014 (95% CI 1.003-1.024) for hospital admission due to cardiovascular and respiratory illness per 10$\mu g/m^3$ increase in PM$_{10}$ exposure (expressed as a 3-day cumulative average) respectively. The best performing model, the 'zonal model' estimated a 0.56% increase in all-natural cause mortality per 10$\mu g/m^3$ increase in PM$_{10}$ exposure. Application of the autoregressive Poisson regression model estimated a leads to an estimated 0.32% (95% CI 0.10-0.55 %) increase in all natural cause mortality per 10$\mu g/m^3$ increase in PM$_{10}$.

The earliest time-series study available from India (Cropper et al., 1997) used total suspended particulate matter (TSP) data from the early 1990s in Delhi, a period during which neither air quality nor mortality-related information was being collected efficiently. Nevertheless, the study (Cropper et al., 1997) documented a significant positive association consistent with the results from the newer India studies. The earlier co-ordinated time-series India studies used a dataset covering the period 2002-2004 and estimated a 0.44%(with 95 % CI being 0.17 %-0.70 %) increase in Chennai (Balakrishnan et al., 2011) and a 0.19% increase in Delhi (Rajarathnam et al., 2011) for all natural cause mortality. The nearly threefold difference between effects estimates is comparable to intercity differentials observed in other co-ordinated time-series studies within countries or regions.

HEI, 2010. The study in Ludhiana was unable to use PM$_{10}$ data in the analysis directly (owing to significant missing air pollution data), but when researchers used measured visibility as a proxy for air quality, they reported a 2.4% increase in all-cause mortality for every 1-km decrease invisibility at midday (R. Kumar et al., 2010). However, the relationship between visibility and air pollutant concentration was not specified.

The excess risk estimates obtained across models in the study are very similar to the summary estimates obtained from the meta-analyses of all Asian studies, namely, 0.27% (95% CI 0.120.42; HEI 2010), and to those from coordinated PAPA studies in four Asian cities, i.e. 0.6% (95%CI 0.30.9; C. M. Wong et al., 2008; HEI, 2010) per 10 $\mu$g/m$^3$ increase in pollutant concentration over a mean PM$_{10}$ concentration range of 51.6141.8 $\mu$g/m$^3$ (similar to the 7 year annual average in Chennai that ranged between 69 to 100 $\mu$g/m$^3$). This risk estimate also falls in the range of estimates found in the European meta-analysis of 29 studies(Katsouyanni et al., 2001), namely, a 0.6% increase in risk of mortality and in the range of revised estimates of the NMMAPS study in the USA (Samet et al. 2000b), i.e. , a 0.21% increase in the risk of mortality, and over a much lower concentration range (mostly < 50$\mu$g/m$^3$). Excess risk estimates were observed to be higher for the age groups $\geq 65$ and $\geq 75$ years than for younger age groups similar to what has been consistently reported in studies in other regions(e.g. Katsouyanni et al. 2009; Wong et al. 2001)

The dose response relationship observed for daily average concentrations of PM$_{10}$ is linear in the range under 150$\mu$g/m$^3$ (the 95% upper confidence limit for the 7 year average PM$_{10}$ concentration). This is similar to what has been reported in previous literature over the same concentration range. This study had too few observations beyond this range to allow reliable prediction of the dose response curve beyond 150 $\mu$g/m$^3$. Similar linear dose-response relationships have been reported in recent Asia studies(Wong et al. 2010)

Finally, similar to studies in Asia and North America, the associations with daily hospital admissions due to cardiovascular and respiratory illness was stronger than what was obtained for all-natural cause mortality. This study showed a 1.4% increase with 95% CI (0.3-2.4)% in hospital admissions due to cardiovascular and respiratory illness due to cardiovascular and respiratory illness per 10$\mu$g/m$^3$ increase in PM$_{10}$ exposure.

A comparison of estimates from other previous time series studies related to air pollution and health studies conducted in Asia, are provided in Figure 6.1

Figure 6.1: A comparison of RR estimates with 95% confidence interval of Health outcome due to PM$_{10}$ exposure reported by time series studies in Asia.

### 6.1.5 Limitations of study

Despite the overall similarity of the excess risk estimates between the India studies and those reported across other Asia studies, certain features merit additional consideration. A recent report by WHO (2008) describes different rates of mortality among age groups, with a higher mortality among men between the ages of 15 and 60 years, and underlying causes of death. Country-specific data compiled by WHO (2008) on causes of death for three age groups (0-14, 15-59 and ≥60 years) indicate that the age-specific percentages of deaths from cardiovascular disease are very similar in India, China and the USA, as are deaths from injuries for ages 15-59 and ≥60 years. However, compared with China and the USA, India has higher percentages of deaths from communicable diseases, including respiratory infections for all age groups as well as for diarrheal diseases for age group 0-14 years and for tuberculosis for age group 15-59 years. Although information on specific causes of death was ascertained, it was not possible to gain a full understanding of the underlying reasons for the observed differences and the differential role of air pollution exposure across age groups. The results point to the need for careful consideration of...
such data in future studies.

The dose-response relationship observed for daily average concentrations of PM$_{10}$ in Chennai appears to be linear over a broad range of daily average levels of PM$_{10}$ of up to approximately 100 µg/m$^3$. This result is similar to findings reported recently for other Asian cities (HEI 2010a). Although there appears to be some non-linearity at higher levels, it is possible that the plot over the entire range is plausible; however, our study had too few observations to reliably estimate the shape of the dose-response curve at higher levels. Therefore, the shape of the curve observed in this study has to be considered to be tentative and await future analyses with a multi-city data set for validation.

The data on gaseous pollutants could not be analysed as the methods of sampling and analysis did not provide the needed resolution for time-series analyses. PM$_{2.5}$ data too was not available as routine monitoring networks did not include measurement of PM$_{2.5}$ during the study period. Future studies would need to consider additional pollutants in single or multi-pollutant models to strengthen the evidence obtained from using PM$_{10}$ data.

Finally, while the data on mortality could be analysed using multiple modeling approaches, the morbidity data collected was limited and could only be explored in the core model. Further examination would be needed through datasets that are collected across a range of health care facilities that cover the city’s population and provide a more representative and a more complete morbidity picture.

6.2 Conclusions

The study has been an exploratory attempt to apply a set of methods for time-series analyses that would reliably describe the association between concentrations of PM$_{10}$, all-cause and cause specific mortality and morbidity in Chennai city. 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.

The data processing steps for the pollutant and mortality data could be completed 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 monitors 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 and morbidity 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 and cause of hospital admissions. Improved visibility data and electronic data sets that include hourly readings would also 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 epidemiological 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 additional 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 casecrossover designs may allow some of these uncertainties to be addressed.

In conclusion, this study has attempted to extract the maximum level of information from the datasets available, future efforts that capture population exposure in better ways would be needed before a realistic picture of mortality and morbidity impacts associated with PM$_{10}$ and other criteria air pollutants can be fully developed. Issues such as missing measurements and small monitor footprints (which result in available air quality data being inadequate to readily represent population exposure) are likely to be encountered in many other Indian cities as well as perhaps in other developing countries. Until investments in infrastructure allow the design of more sophisticated monitoring mechanisms, the methods developed in this study may therefore allow data currently being collected only to be used for baseline assessments in situations where similar data challenges prevail. This analysis supports the inference that relative effects are likely to be similar across regions of the world that vary markedly in levels of PM$_{10}$, but that insufficient attention to the nature of local datasets and the application of models that do not adequately address exposure misclassification prevalent in these settings may provide inconsistent results.

It is hoped that this study methods will be useful for application in time series analysis in other Indian cities in the future and this study together with future studies will catalyze policy changes and contribute to the improvement of air quality and public health in India.

---


---