



Cause-specific premature death from ambient PM_{2.5} exposure in India: Estimate adjusted for baseline mortality



Sourangsu Chowdhury, Sagnik Dey *

Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

ARTICLE INFO

Article history:

Received 24 November 2015

Received in revised form 2 March 2016

Accepted 6 March 2016

Available online xxxx

Keywords:

Ambient PM_{2.5} exposure

India

Cause-specific premature death

Baseline mortality adjustment

District level statistics

ABSTRACT

In India, more than a billion population is at risk of exposure to ambient fine particulate matter (PM_{2.5}) concentration exceeding World Health Organization air quality guideline, posing a serious threat to health. Cause-specific premature death from ambient PM_{2.5} exposure is poorly known for India. Here we develop a non-linear power law (NLP) function to estimate the relative risk associated with ambient PM_{2.5} exposure using satellite-based PM_{2.5} concentration (2001–2010) that is bias-corrected against coincident direct measurements. We show that estimate of annual premature death in India is lower by 14.7% (19.2%) using NLP (integrated exposure risk function, IER) for assumption of uniform baseline mortality across India (as considered in the global burden of disease study) relative to the estimate obtained by adjusting for state-specific baseline mortality using GDP as a proxy. 486,100 (811,000) annual premature death in India is estimated using NLP (IER) risk functions after baseline mortality adjustment. 54.5% of premature death estimated using NLP risk function is attributed to chronic obstructive pulmonary disease (COPD), 24.0% to ischemic heart disease (IHD), 18.5% to stroke and the remaining 3.0% to lung cancer (LC). 44,900 (5900–173,300) less premature death is expected annually, if India achieves its present annual air quality target of 40 µg m⁻³. Our results identify the worst affected districts in terms of ambient PM_{2.5} exposure and resulting annual premature death and call for initiation of long-term measures through a systematic framework of pollution and health data archive.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Long-term exposure to ambient PM_{2.5} (particulate matter with diameter <2.5 µm) has potential health risks including premature death (Brauer et al., 2012; Pope et al., 2002). In fact, it has been cited as the sixth largest cause of premature death in South Asia in the global burden of disease (GBD 2010) study (Lim et al., 2012). India, home to more than 1.2 billion people, has been recognized as a regional pollution hotspot for persistently high aerosol burden (Dey and Di Girolamo, 2010) and its rapid increase (Dey and Di Girolamo, 2011) in the last decade (2000–2010). In India, respirable particulate matter i.e. PM₁₀ (particulate matter with diameter <10 µm) is routinely monitored at Center Pollution Control Board network sites spread across the country. These data were utilized to estimate relative risk (RR) of all-cause mortality in two major cities - Chennai and Delhi (Health Effects Institute Research Report, 2011). However, PM_{2.5} is more hazardous to human health than PM₁₀ because of the ability of the fine particles to reach the minute airways in the lungs (Pope et al., 2002).

Three major factors contribute to large uncertainty in estimating premature mortality due to ambient PM_{2.5} exposure in India. Firstly, there are very few PM_{2.5} monitoring sites in some of the major cities

and many of them do not have consistent quality-controlled data for a long enough period. Moreover, these sites are located in urban areas, leaving most of India unmonitored. Highly variable aerosol fine mode fraction across space and time (Dey and Di Girolamo, 2010) further suggests the difficulty in inferring PM_{2.5} from PM₁₀ measurements. Secondly, no India-specific cohort study exists to quantify cause-specific RR from ambient PM_{2.5} exposure. Extrapolation of RR for household air pollution exposure to ambient PM_{2.5} exposure using the study conducted recently in India (e.g. Balakrishnan et al., 2013; Smith et al., 2014) does not provide the solution either because of the insufficient data of ambient PM_{2.5} exposure over India as whole. Thirdly, recent estimate of premature death using integrated exposure risk function (IER) considered uniform baseline mortality across India and raw estimate (no calibration against in-situ data from India) of satellite-based PM_{2.5} in the global burden of diseases study (GBD 2013, Murray et al., 2015). Here, we address all these limitations. We utilize our satellite-based PM_{2.5} estimate, validated and bias-corrected against coincident PM_{2.5} measurements from India (Dey et al., 2012) for exposure analysis. We develop a baseline mortality function using GDP as surrogate and adjust estimate of premature death from PM_{2.5} exposure for baseline mortality variation across the country due to socio-economic heterogeneity. Finally, estimate of premature death is apportioned to four major diseases within a defined uncertainty range for ambient PM_{2.5} exposure in the last decade. We also compare our estimate of annual premature

* Corresponding author.

E-mail address: sagnik@cas.iitd.ac.in (S. Dey).

death with the estimate presented in previous studies (such as Apte et al., 2015 and Murray et al., 2015). Statistics are presented at the administrative district level for the first time to provide a better context for the policy makers, unlike any previous study where country specific single estimate was reported as part of global study.

2. Method

2.1. Satellite derived $PM_{2.5}$

Lack of a systematic $PM_{2.5}$ database across India motivated us to utilize satellite aerosol product for this study. Multiangle Imaging SpectroRadiometer (MISR)-retrieved daily columnar aerosol optical depth (AOD) at $0.5^\circ \times 0.5^\circ$ resolution is converted to surface $PM_{2.5}$ using a conversion factor. This spatially varying mean monthly conversion factor is calculated from the ratio of surface $PM_{2.5}$ to AOD simulated by GEOS-Chem model (van Donkelaar et al., 2010) with aerosol vertical distribution constrained by CALIOP measurements. Satellite-retrieved $PM_{2.5}$ is found to be biased low as compared to the coincident direct measurement (Dey et al., 2012). The systematic low bias is corrected following our previous work (Dey et al., 2012) that improves the estimate of $PM_{2.5}$ to within $\pm 8\%$ of the direct coincident measurements. $PM_{2.5}$ statistics for 571 administrative districts (see supplementary information, Table S1 and Fig. S1 for details) are generated by calculating area-weighted average with the help of district boundary shape files overlying the gridded $PM_{2.5}$ data (adjusted for variation in smaller districts as per MISR 17.6 km AOD product within $0.5^\circ \times 0.5^\circ$ grid) in a GIS platform. Mean annual $PM_{2.5}$ (Fig. 1a) varies spatially in a wide range. Kinnaur in Himachal Pradesh is the cleanest district (annual $PM_{2.5} = 3.7 \pm 1 \mu\text{g m}^{-3}$), while Delhi is the dirtiest metropolitan area (annual $PM_{2.5} = 148 \pm 51 \mu\text{g m}^{-3}$). We classify the exposure of the population in all the districts to 7 vulnerability categories (Table 1) based on $PM_{2.5}$ statistics in view of various air quality thresholds. Vulnerability is classified as 'very low' (only 5.2% districts) if $PM_{2.5}$ is less than 5th percentile. Upper limit of $12.5 \mu\text{g m}^{-3}$ for this class is very close to recently adopted US EPA (United States Environmental Protection Agency) standard. 4.7% districts are of 'low' vulnerability category, where $PM_{2.5}$ lies in 5th–10th percentile range with the upper limit of $14.8 \mu\text{g m}^{-3}$ similar to WHO interim target – III. Vulnerability in 16.1% and 25.3% districts are classified as 'moderate' ($PM_{2.5}$ in 10th–25th percentile range with its upper limit of $22.1 \mu\text{g m}^{-3}$ close to EU (European Union) air quality standard), and 'high' ($PM_{2.5}$ in 25th–50th percentile range with its upper limit of $37.3 \mu\text{g m}^{-3}$ close to

Indian standard). 'Very high', 'severe' and 'extreme' vulnerability are categorized for annual $PM_{2.5}$ in the range 50th–75th, 75th–90th and >90th percentile respectively (Fig. 1b). To summarize, more than 45% of the districts housing ~50% of India's population are exposed to $PM_{2.5}$ concentration above the Indian standard ($40 \mu\text{g m}^{-3}$). Only 0.06% of the Indian population is breathing safe air as per the WHO air quality guideline.

2.2. Relative risk and premature death

We choose four diseases - chronic obstructive pulmonary disease (COPD), ischemic heart disease (IHD), stroke and lung cancer (LC), which have direct causal links to ambient $PM_{2.5}$ exposure (Wellenius et al., 2012; Pope et al., 2002; Rich et al., 2010). Premature death ($\Delta M_{i,j}$) due to a particular disease j for district i attributable to $PM_{2.5}$ exposure has been estimated using the traditional epidemiological relation (Anenberg et al., 2010; Silva et al., 2013):

$$\sum_{i,j=1}^N \Delta M_{i,j} = \sum_{i,k,j=1}^N Y_{i,k,j} \times \frac{\sum_{i=1}^N RR_{i,j} - 1}{\sum_{i=1}^N RR_{i,j}} \times \sum_{i=1}^N P_i, \quad (1)$$

where $Y_{i,k,j}$ is the baseline mortality for disease j in a district i within the jurisdiction of a state k . Exposed adult population (P_i) for each district is obtained from Indian Population Census database (population above 25 years of age is considered for exposure assessment to be consistent with GBD study). One critical issue is the extrapolation of relative risk ($RR_{i,j}$) to the observed $PM_{2.5,i}$ concentration for each of the four diseases using the existing epidemiological studies for ambient $PM_{2.5}$ exposure. The conventional approach to calculate RR is either to use concentration-response function derived from cohort studies (Pope et al., 2002) or use some standard model (e.g. Lin 50 model, Cohen et al., 2006), where RR is considered to be uniform at $PM_{2.5} > 50 \mu\text{g m}^{-3}$. All the cohort studies for ambient $PM_{2.5}$ exposure were carried out at developed countries, where ambient $PM_{2.5}$ concentration was much lower than what is usually observed in India. Hence those values are not truly representative of Indian condition. Cohort studies for household air pollution exposure and smoking (Pope et al., 2011; Smith et al., 2014) clearly suggest a higher RR for cardiovascular and lung diseases; which can never be obtained using any standard model for ambient $PM_{2.5}$ exposure. The upper limit of the ambient $PM_{2.5}$ exposure is close to the lower limit of some of the reported household air pollution exposure.

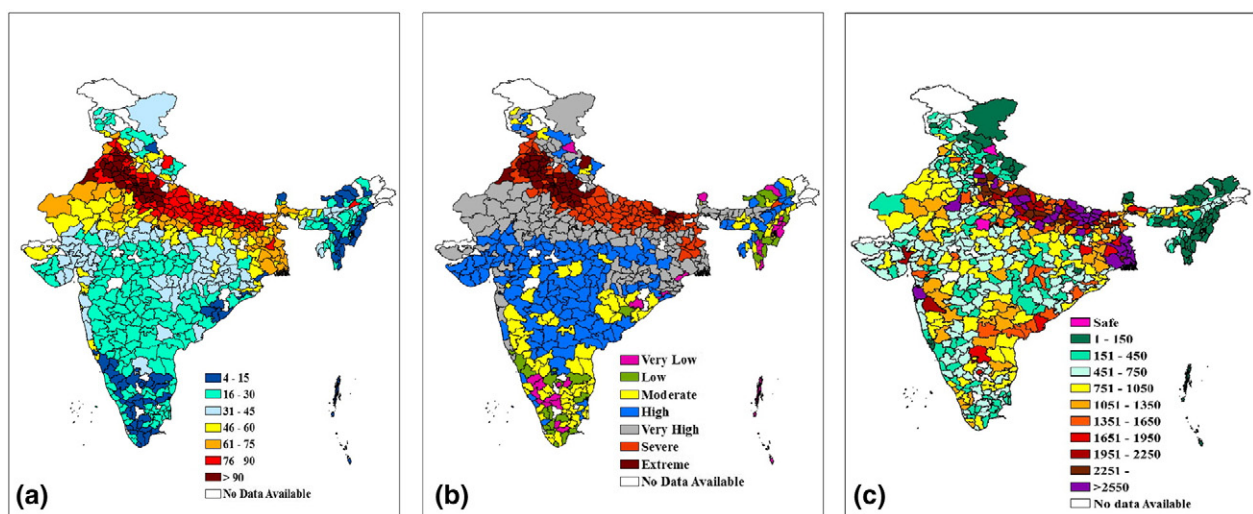


Fig. 1. (a) Mean annual $PM_{2.5}$ concentration (in $\mu\text{g m}^{-3}$) over India for the period 2000–2010. The district-level statistics are generated from the gridded data as described in the text. (b) Classification of districts to 7 vulnerability categories defined in Table 1. (c) Burden of premature death per year adjusted for state-specific baseline mortality in each district obtained by adding up premature deaths due to COPD, IHD, stroke and lung cancer attributed to ambient $PM_{2.5}$ exposure in India.

Table 1

Qualitative classification of vulnerability due to ambient PM_{2.5} exposure in India and the annual premature death estimated using bias-corrected MISR-PM_{2.5} and NLP and IER risk functions, adjusted for state-specific baseline mortality.

Percentile	Annual PM _{2.5} (µg m ⁻³)	Category	# districts (in %)	# exposed population (>25 yrs)	Annual premature death per district using NLP (IER)	Mean baseline mortality per 100,000 population for COPD, IHD, stroke	Comment
<5th	<12.5	Very low	5.2	12,971,441	260 (270)	133, 162, 113	Similar to US-EPA air quality guideline of 12 µg m ⁻³
5th - 10th	12.5–14.8	Low	4.7	15,419,579	390 (500)	132, 160, 112	Upper limit similar to WHO Interim Target - III
10th - 25th	14.8–22.1	Moderate	16.1	64,473,347	500 (730)	133, 161, 113	Lower limit close to EU Air quality threshold
25th - 50th	22.1–37.3	High	25.3	107,079,474	610 (1080)	148, 171, 121	Upper limit close to WHO Interim Target - I and Indian Standard
50th - 75th	37.3–69.8	Very High	23.9	110,089,990	950 (1620)	182, 193, 137	Upper limit is double of WHO Interim Target - I
75th - 90th	69.8–87.3	Severe	15.6	78,465,779	1610 (2420)	213, 228, 164	Upper limit exceeds twice of Indian Standard
>90th	>87.3	Extreme	9.2	42,143,518	1090 (2020)	169, 179, 127	~9 times of WHO air quality standard

Since we only consider PM_{2.5} mass concentration and not the chemical composition to estimate RR, the approach seems to be justified (same philosophy was followed in developing IER function, Burnett et al., 2014). We collect estimates of RR from 141 published cohort studies (see the complete list in supplementary information, Table S2) across the globe spanning exposure from ambient, household air pollution, active and second hand smoking, and partitioned them into four diseases. A non-linear power law (NLP) function is developed for each of these diseases (Fig. 2) to quantify RR at any given PM_{2.5} exposure in the following form:

$$RR_{i,j} = 1 + \alpha_j \times (\Delta PM_{2.5})_i^{\beta_j} \tag{2}$$

Two constants, α and β are estimated to be 0.11 (0.016–0.25) and 0.44 (0.31–0.56), 0.05 (0.02–0.075) and 0.31 (0.26–0.36), 0.06 (0.014–0.1) and 0.3 (0.23–0.37), and 0.07 (0.03–0.11) and 0.75 (0.68–0.82), respectively for COPD, IHD, stroke and LC. $(\Delta PM_{2.5})_i$ is the change in PM_{2.5} concentration of a district i from the counterfactual concentration. The function is derived in a way that the RR becomes equal to 1 below the counterfactual concentration of PM_{2.5} exposure, in which case, difference between ambient and counterfactual PM_{2.5} concentration (i.e. $\Delta PM_{2.5}$) is zero. In other words, no risk exists below the counterfactual concentration, which is considered to be 5.8 µg m⁻³ (lower limit of counterfactual concentration used in estimating premature death in Lim et al., 2012 and Murray et al., 2015).

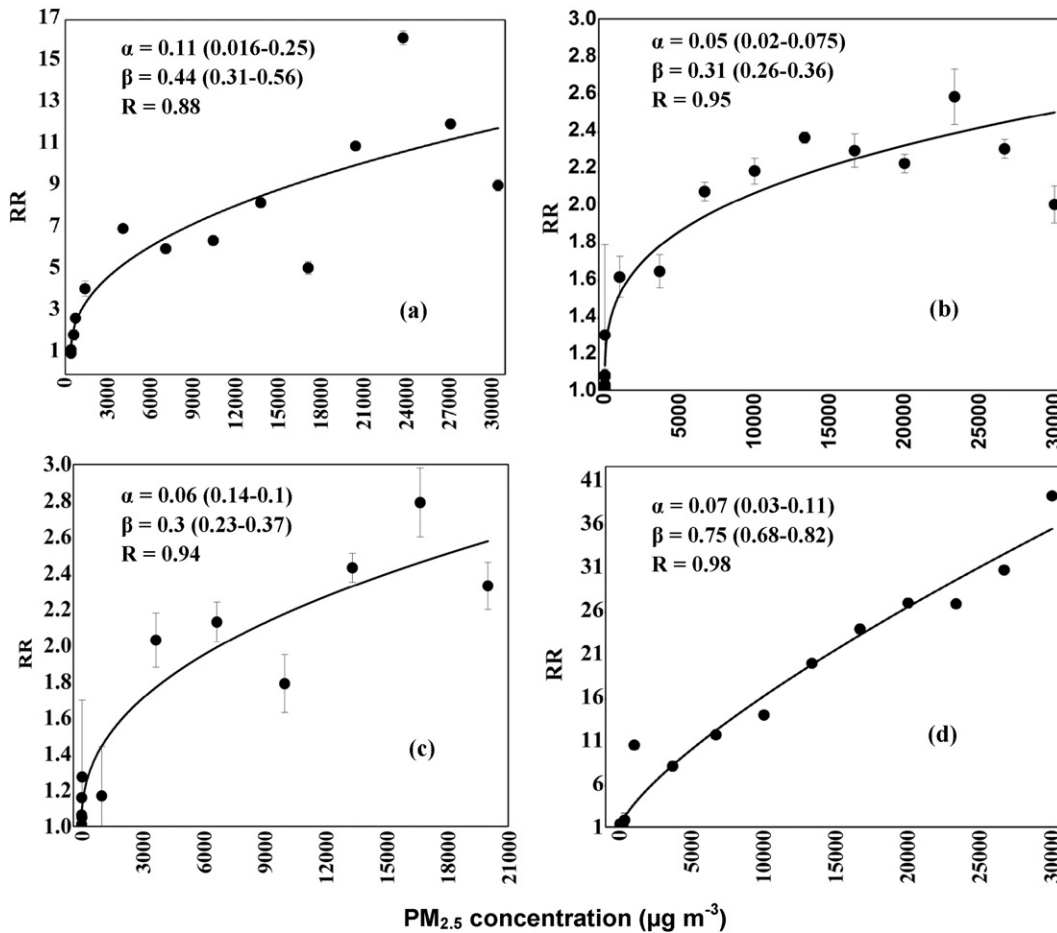


Fig. 2. Power-law fit for RR estimate associated with exposure to PM_{2.5} for (a) COPD, (b) IHD, (c) Stroke and (d) LC. Each point in the figure represents RR determined by an individual cohort study (listed in Table S2) with the uncertainty shown as error bar. Values of α and β in the power law function, $RR = 1 + \alpha \times (\Delta C)^\beta$, are also displayed along with uncertainty range ($\pm\%$) for each of the diseases. The correlation coefficients are statistically significant at 99% CI following t -test.

2.3. Baseline mortality adjustment

Single baseline mortality values (142.1, 165.8, 116.4 and 6.5 per 100,000 population for COPD, IHD, stroke and LC, respectively) for India are available from WHO statistics (<http://www.who.int/whosis/whostat/2011/en/>). In a country like India with varying socio-economic condition, assumption of uniform baseline mortality across India does not make sense. Our approach to examine this issue is based on the following logic. Premature mortality from a disease can be prevented to some extent if adequate healthcare support is affordable. Since access to 'healthcare support' depends on the socio-economic condition of the population, it is related to baseline mortality. We consider per capita gross domestic product (GDP) as a surrogate to the socio-economic condition and obtain GDP at constant price for the year 2010–11 for the Asian countries from World Bank statistics and baseline mortality of the four diseases from WHO statistics 2011. Data are available for all the countries in the world, but they are restricted to Asian countries only to minimize the heterogeneity in ethnicity of the exposed population. Following this logic it would have been even better to restrict the data to South-Asian countries only; but it could not be done due to very less data limiting robust statistics. We assume that the states with higher GDP are expected to spend more in healthcare facilities that would reduce the mortality for COPD, stroke and IHD from all possible sources including ambient PM_{2.5} exposure. We obtain the gross per capita GDP of 32 Indian states and union territories from the Reserve Bank of India statistics (2014) for the year 2010–11. Non-linear functions (statistically significant at 95% CI) are derived

in the form of $GDP = a \times [(1 + y)^b]$ for IHD and $GDP = a \times y^b$ for COPD and stroke (Fig. 3; here, 'a' and 'b' are disease specific constants) from data of GDP and baseline mortality of all the Asian countries. These non-linear models are further applied to find the baseline mortality for COPD, IHD and stroke for each state. LC baseline mortality does not exhibit any relation with GDP; perhaps because even best possible healthcare may not be enough to prevent mortality from LC, which is a fatal and progressive disease. Baseline mortality (mean $\pm 1\sigma$) averaged for all the states (140.5 ± 67 , 164.9 ± 44 and 115.8 ± 34 for COPD, IHD and stroke respectively per 100,000 population) are estimated to be very similar (within <1.2%) to the national average from WHO, 2011 statistics mentioned earlier, thereby validating the non-linear functions. State-specific baseline mortality for COPD, IHD and stroke estimated using these functions is shown in Fig. 4. Baseline mortality (per 100,000 population) for COPD ranges between 48 (for the state of Goa) to 355 (for the state of Bihar), while the corresponding ranges for IHD and stroke are 91 (Goa) to 289 (Bihar) and 60 (Goa) to 211 (Bihar) respectively. The geographical positions of the states are shown in supplementary information Fig. S1. After adjusting for the spatial variation of baseline mortality, the baseline mortality of some states in central-north India like Bihar increases by 126%, 103% and 67% for COPD, IHD and stroke respectively relative to the assumption of uniform baseline mortality values in the previous studies across India. In some of the states in north-western India, baseline mortality decreases (e.g. by 52%, 31% and 36% for COPD, IHD and stroke respectively in Delhi). All the states in South India show a decrease in baseline mortality from the Indian average (WHO, 2011) for all the three diseases.

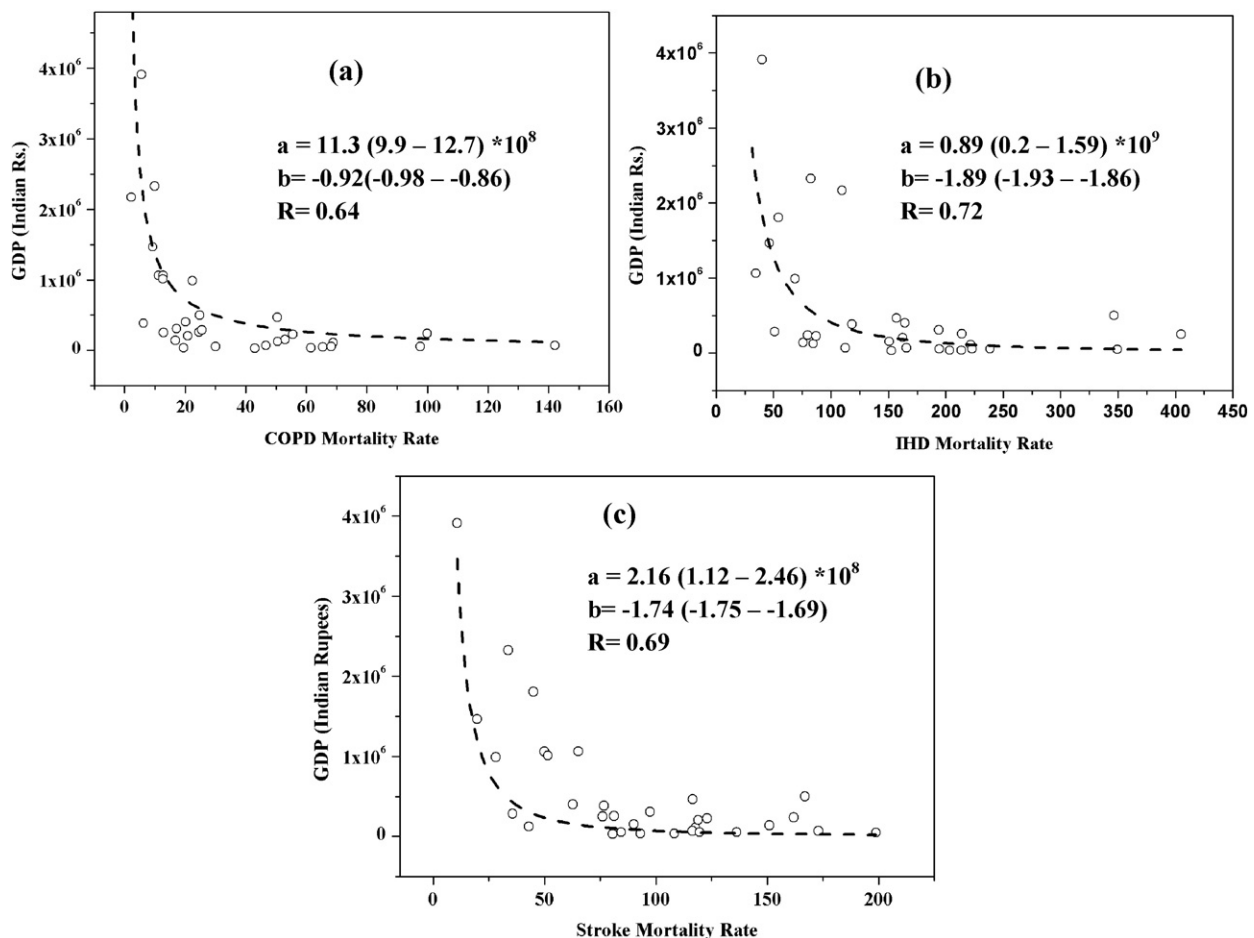


Fig. 3. Non-linear fit for per capita GDP against baseline mortality for (a) COPD, (b) IHD and (c) stroke in the Asian countries. The function for COPD and stroke is $GDP = a \times y^b$ and $GDP = a \times (1 + y)^b$ for IHD, where 'a' and 'b' are disease specific constants.

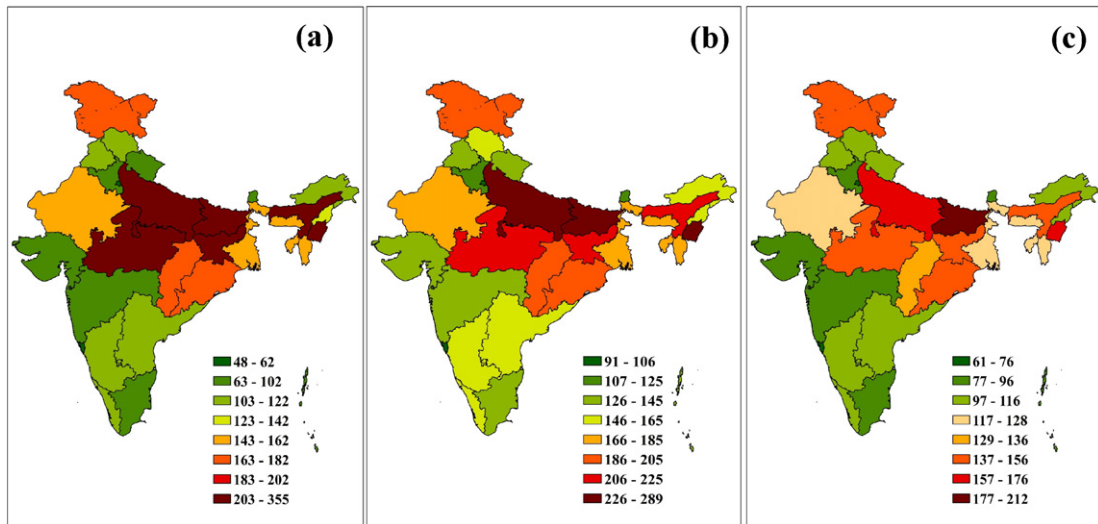


Fig. 4. State-specific baseline mortality for (a) COPD, (b) IHD and (c) stroke. Transition from green shade to red signifies crossing over the all-India average baseline mortality. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4. Estimate of premature death using IER function

Burnett et al. (2014) developed nonlinear integrated exposure-response function that constrains the shape of the concentration response relationship unlike the vintage concentration response relation developed by Pope et al., 2002, by considering RR of premature mortality across a wide range of exposure including ambient and household air pollution and active and passive smoking. The IER framework can be generalized as Eq. (3), and depends on the concentration of $PM_{2.5}$ based on meta-analysis of observed data from cohort studies performed around the world.

$$RR_{i,j} = 1 + \alpha_j \left[1 - \exp\left(-\gamma_j (\Delta PM_{2.5})_i^{\delta_j}\right) \right], \quad (3)$$

where RR_i represents the relative risk of a disease j at the district i for a specific $PM_{2.5}$ concentration. $(\Delta PM_{2.5})_i$ denotes the change in $PM_{2.5}$ concentration from the counterfactual concentration of $5.8 \mu g m^{-3}$. α_j , γ_j and δ_j are parameters specific for each disease. We use the mean values of RR for each disease (age specific values of RR has been used for stroke and IHD) that are provided as a look up table by Apte et al. (2015) to estimate the premature death and compare with our estimates using the NLP function. The previous studies (Lim et al., 2012; Apte et al., 2015) used the satellite based estimates of $PM_{2.5}$ from van Donkelaar et al. (2010) with 25% uncertainty, which were not bias corrected and therefore were biased low over India. (Dey et al., 2012) has bias-corrected and validated satellite derived $PM_{2.5}$ concentration over India with coincident direct measurements to generate improved $PM_{2.5}$ estimate with a lower (8%) uncertainty.

3. Results and discussion

We estimate all-India annual premature death adjusted for state-specific baseline mortality of 271,900 (14,400–795,800) for COPD, 110,700 (37,800–220,000) for IHD, 88,700 (13,200–219,700) for stroke and 14,800 (7700–19,300) for LC by the NLP risk function. Largest (54.5%) contribution comes from COPD and the smallest (3.0%) from LC. Premature death for a district estimated by using the NLP function is high (Fig. 1c), if either of $PM_{2.5}$ concentration, population of the district and the baseline mortality of the district or all three of these factors are large. Average annual premature death is 48% higher for the districts categorized as ‘very high’ (~1600 per district) compared to the districts of ‘extreme’ (~1100 per district) vulnerability because of larger (~1.8

times) exposed population and higher baseline mortality for the former class. We quantify various factors driving the observed spatial heterogeneity of estimated premature death. Attributable population at district i and for disease j is defined as following:

$$\sum_{i,j=1}^N AP_{i,j} = \frac{\sum_{i=1}^N RR_{i,j} - 1}{\sum_{i=1}^N RR_{i,j}} \times \sum_{i=1}^N P_i. \quad (4)$$

We note that, in addition to RR and exposed population, baseline mortality variation (Fig. 4) also plays an important role in spatial distribution of premature death. Spatial variability of attributable population for each of the diseases is shown in supplementary information Fig. S2. Higher population is at risk of dying prematurely due to ambient $PM_{2.5}$ exposure from these diseases in the districts of northern India along the Indo-Gangetic plain (relative to districts of southern India), because both the population and $PM_{2.5}$ concentration in these districts are higher. Some districts in the state of Andhra Pradesh (see supplementary Fig. S1 for location) also has higher attributable population as the exposed population is very high although the $PM_{2.5}$ exposure for those districts is comparatively lower than the districts lying on the Indo-Gangetic plain (supplementary Table S1).

It can be identified that in the districts of Indian states of Uttar Pradesh, Bihar, and West Bengal, all the three factors driving premature death (viz. $PM_{2.5}$ concentration, exposed population and baseline mortality) are high thus resulting in higher number of premature deaths in these states; whereas in the districts of the states of Punjab, Haryana and Delhi in northern India, even though the $PM_{2.5}$ concentration and exposed population are high, low baseline mortality results in relatively lower annual premature death (relative to the districts of Uttar Pradesh, Bihar, and West Bengal). In the districts of South India, even though the exposed population is high, low $PM_{2.5}$ concentration and low baseline mortality restrain the annual premature death due to ambient $PM_{2.5}$ exposure.

District-wise statistics is summarized in supplementary Table S1 along with the range of uncertainty within parentheses. The states of Uttar Pradesh, Bihar, West Bengal, Maharashtra and the Delhi metropolitan area (see supplementary Fig. S1 for their locations) are the most vulnerable and contribute 25%, 15%, 7.6%, 5.4% and 1.7% to total all-India premature death from ambient $PM_{2.5}$ exposure in the previous decade. The analysis (Table 2) further reveals that the annual premature death is estimated to be lower by 14.7% for assuming uniform baseline mortality (for COPD, IHD and stroke) across India compared to the estimate adjusted for state-specific baseline mortality. Fig. 5 depicts the

Table 2

Comparison between estimates of annual premature death from ambient PM_{2.5} exposure calculated using four combinations (denoted as 'cases') of baseline mortality and PM_{2.5} concentration for two different risk functions - NLP and IER. Case 1 refers to spatially varying baseline mortality and bias-corrected PM_{2.5} concentration. Case 2 refers to spatially varying baseline mortality and raw estimate of PM_{2.5}. Case 3 refers to single value baseline mortality for India and bias-corrected PM_{2.5} estimate. Case 4 refers to single value baseline mortality for India and raw estimate of PM_{2.5} and is the most representative of the GBD estimate.

RR function	Case 1	Case 2	Case 3	Case 4
NLP	486,100	414,500	413,300	356,900
IER	811,000	655,300	698,300	580,600

relative change (in %) in premature death for adjustment of baseline mortality using the NLP risk function with respect to the estimates using a single value of baseline mortality (WHO, 2011). The negative (positive) change depicted by the shades of red (green) denotes the districts where use of single baseline mortality for each of the diseases accounts for lower (higher) premature death than the estimate adjusted for varying baseline mortality. The negative change is mostly confined to the economically weak and populous districts in the central and northern parts of the country (Fig. 5). On the other hand, positive change can be noticed by a widely varying range of 10% to more than 60% per year in districts in the Peninsular and western India and north-western states of Punjab, Haryana and Delhi.

Several factors contribute to the uncertainty range provided for the estimated premature death. PM_{2.5} concentration derived from MISR-

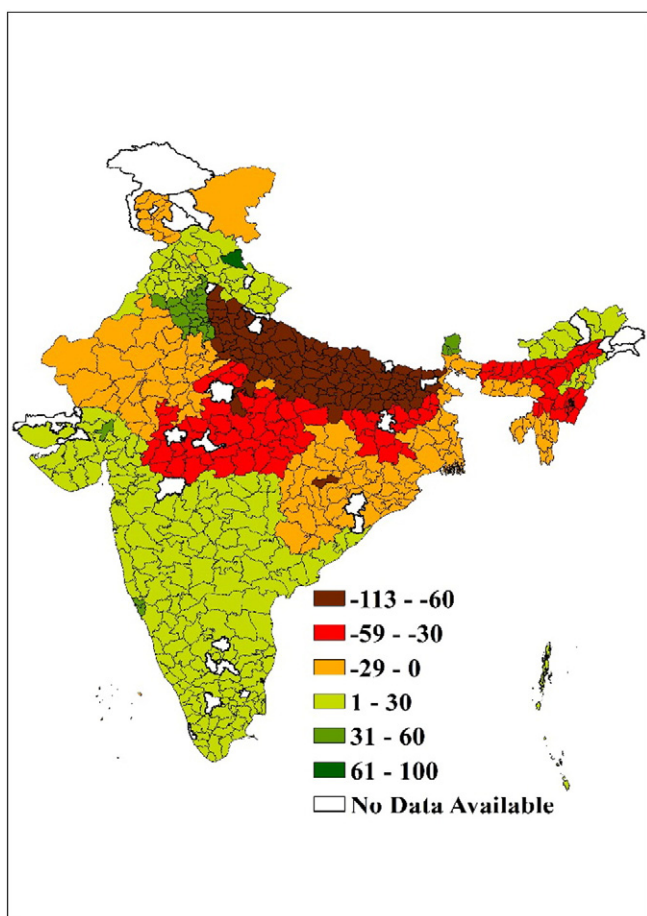


Fig. 5. Percentage change in estimated annual premature death adjusted for state-specific baseline mortality relative to uniform baseline mortality. Shades of reddish (greenish) tinge denote lower (higher) estimate in annual premature death using single value of baseline mortality across India. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

AOD has an overall 8% uncertainty (with respect to coincident direct measurement). Estimated RR and state-specific baseline mortality have uncertainties due to errors in the coefficients ' α ' and ' β ' (Fig. 2) and ' a ' and ' b ' (Fig. 3) of the non-linear functions. Our estimate of 486,100 (73,200–1,254,800) annual premature deaths is lower than some of the recent estimates for India as part of global study (Pope et al., 2002; Cohen et al., 2006; Apte et al., 2015) because of the difference in methodology. The risk function in the first two studies (Cohen et al., 2006; Pope et al., 2002) was developed based on a single cohort study. The last study (Apte et al., 2015) utilized IER (Burnett et al., 2014) developed from a range of exposure studies, but used uniform baseline mortality across India and raw estimate of satellite-PM_{2.5} (without bias correction). It is not possible to directly validate our premature death estimate because of non-availability of cause-specific mortality records in India at district level. However we point out that each individual cohort study used to develop RR function was based on quality-controlled health data across a heterogeneous population. We provide estimate within a defined range of variability. Furthermore, we provide a detailed comparison to understand the sensitivity of the estimates for various factors described above.

We adjust the premature death estimates reported in the recent GBD study (Murray et al., 2015 and Apte et al., 2015) by introducing the spatially varying baseline mortality and employing improved estimates of satellite-PM_{2.5} (due to bias correction). Our estimate shows a higher ambient PM_{2.5} exposure (all-India average of 46.5 $\mu\text{g m}^{-3}$) compared to a much lower value considered in the previous studies (e.g. 28 $\mu\text{g m}^{-3}$ in Apte et al., 2015). The model-derived PM_{2.5} exposure (20.4 $\mu\text{g m}^{-3}$ reported in Anenberg et al., 2010) is even lower. Table 2 compares the annual premature death estimated with two different risk functions (NLP and IER) for four combinations (referred to as 'Cases') as described below. Case 1 considers spatially varying baseline mortality and bias-corrected estimate of PM_{2.5} (as reported in Dey et al., 2012). Case 2 considers spatially varying baseline mortality and raw estimate of PM_{2.5} (without bias correction using coincident data from India). In Case 3, single value of baseline mortality for entire India and bias-corrected estimate of PM_{2.5} are used, while a single value of baseline mortality for entire India and raw estimate of PM_{2.5} are used for Case 4. District-level statistics provided in Fig. 1 and Table S1 are derived for Case 1 using NLP and IER. Estimate by IER for Case 4 (580,600) best approximates annual premature death (587,000) obtained by Apte et al., 2015 and Murray et al., 2015 for India. The small difference (~7000) may be attributed to the fact that those two studies used MODIS-MISR combined PM_{2.5} data, while we use only MISR-PM_{2.5}.

Annual premature death estimate of 587,000 reported for India in GBD study (Murray et al., 2015) using raw estimate of PM_{2.5}, uniform baseline mortality across India and IER function increases to 698,300 (Case 3) when bias-corrected PM_{2.5} is used. This further increases to 811,000 (Case 1), when it is adjusted for state-specific baseline mortality. On the other hand, simply, change of risk function (to NLP) decreases the corresponding estimates to 356,900 (Case 4), 413,300 (Case 3) and 486,100 (Case 1) respectively. Estimate of annual premature death increases by 15.8% (20.3%) when we consider better estimate of PM_{2.5} and by 36.2% (39.7%) when we adjust for baseline mortality also for NLP (IER) risk functions respectively. If raw satellite-PM_{2.5} concentrations are used, but baseline mortality variation is adjusted, annual premature death estimate increases by 16.1% and 12.9% for NLP and IER risk functions respectively. To summarize, the estimate varies widely for various assumptions and chosen risk functions. Logically, estimate using bias-corrected PM_{2.5} and adjusted for varying baseline mortality is more realistic. Annual premature death of 811,000 (spatial variation is shown in supplementary information Fig. S3) estimated for India using bias-corrected PM_{2.5} and adjusted for varying baseline mortality with IER is larger than our estimate (486,100) using NLP risk function for similar combination. It is difficult to assess which risk function is more appropriate for India in absence of any robust health data and beyond scope

of this study. However, we note that the mean estimate using IER lies in the uncertainty range of our estimate using NLP (73,200 to 1,254,800). This emphasizes on developing India-specific risk function in future to settle this argument.

We note that the results should be interpreted keeping in mind the following three key assumptions; (1) RR does not vary significantly with PM_{2.5} composition, (2) disease-specific risk functions have negligible co-morbidity and (3) adding up the annual premature death caused by COPD, IHD, stroke and lung cancer due to exposure to ambient PM_{2.5} to obtain total premature death assumes no overlap of the causes of premature death. Our district level statistics are critically important for policy implementation to mitigate the burning problem of air pollution in the second most populous country in the world. India has recently adopted annual ambient PM_{2.5} concentration of 40 µg m⁻³ as air quality standard. Note that this standard cannot be realistically considered as counterfactual limit for estimating premature mortality in India because cohort studies (e.g. from Burnett et al., 2014) clearly suggest RR > 1 at PM_{2.5} larger than 5.8 µg m⁻³. Consideration of 40 µg m⁻³ as counterfactual results in 35% underestimation of estimated annual premature death, which may be misleading. However given high background ambient PM_{2.5} concentration in India (median annual PM_{2.5} concentration of 37.3 µg m⁻³ is close to the Indian annual standard), attempting a target of low PM_{2.5} like US-EPA or EU standard is not prudent. It is important to initiate India-specific cohort study spanning across a wide range of demography and PM_{2.5} exposure in future to improve RR estimate. This requires a long-term national level effort through which cause-specific mortality data needs to be archived. However, the country cannot afford to wait long for better estimate of RR, because even our lowest possible estimate of annual premature death within the defined range of uncertainty (i.e. 73,200) is large enough to worry about.

If India manages to meet the national air quality standard in all the districts exceeding Indian standard, annual premature death will be reduced by 9.5%. A large fraction of PM_{2.5} in the eastern part of the Indo-Gangetic Basin (districts in east Uttar Pradesh, Bihar, Jharkhand and West Bengal) is transported from the northwest India (districts in Delhi, west Uttar Pradesh and Punjab) in the post-monsoon and winter seasons (Dey and Di Girolamo, 2010). Regional transport of PM_{2.5} should be considered during implementation of local mitigation measures. Household and biomass-burning emissions are two major contributors to PM_{2.5} in the Indo-Gangetic Basin covering these states (Venkataraman et al., 2005). Controlling biomass emission (dominant in Punjab and Haryana) will reduce transported fraction of PM_{2.5} in the downwind districts. Reduction in household emission will reduce household PM_{2.5} exposure (Smith et al., 2014) in addition to betterment of ambient air, especially when a large fraction of ambient PM_{2.5} in India is contributed by household cooking with solid fuels (Chafe et al., 2014). This will further reduce the baseline mortality in the economically weak states; thereby contributing to additional reduction in premature mortality and co-benefit to climate mitigation through reduction in short-lived climate pollutants emitted from solid fuel burning. Majority of the districts categorized as 'very high' and 'severe' in terms of vulnerability (see Fig. 1b) have large rural population, while 'extreme' vulnerable districts are dominated by urban population. Reducing PM_{2.5} concentration in the rural population dominated districts in the states of Bihar, Uttar Pradesh, West Bengal, Orissa, Jharkhand, Rajasthan and Chhattisgarh (see Fig. S1) to achieve Indian air quality standard will lead to 32,200 (4900–130,400) less premature deaths per year, ~2.5 times higher than the corresponding number of 12,700 (1000–42,900) in the urban dominated districts (using NLP risk functions for Case 1). We also note that the districts where annual PM_{2.5} is less than the Indian air quality standard should also be targeted for further reduction in ambient PM_{2.5} exposure to diminish mortality and morbidity burden. India spends only 4% of its GDP in public healthcare, most of which is focused in urban areas. Our results suggest that careful planning is required to address these issues (in terms of GDP share in public health)

through prioritization of the targeted districts for maximum health benefit.

4. Summary and conclusions

In this study, we present district-level statistics of annual premature death (from COPD, stroke, IHD and lung cancer) for ambient PM_{2.5} exposure in India, which is adjusted for state-specific baseline mortality. Statistics presented at district level provides insight into the spatial heterogeneity, which are difficult to extract from global studies. These details and region-specific adjustments for baseline mortality are critical for the administration to prioritize the vulnerable districts in taking an informed decision to address this important problem under the context of varying socio-economic condition across the country. Key conclusions of our study are:

1. Exposure analysis reveals that 50% population living in 45% districts of India is exposed at PM_{2.5} exceeding Indian air quality standard of 40 µg m⁻³. Kinnaur in Himachal Pradesh is identified as the cleanest district (annual PM_{2.5} is 3.7 ± 1 µg m⁻³), while Delhi is the dirtiest metropolitan area (annual PM_{2.5} is 148 ± 51 µg m⁻³).
2. Annually the premature death estimates are 272,000 (14,400–795,800) for COPD, 110,600 (37,800–220,000) for IHD, 88,700 (13,200–219,700) for stroke and 14,800 (7700–19,300) for lung cancer respectively in India using NLP risk function with adjusted baseline mortality and bias-corrected estimates of PM_{2.5}. In terms of relative contribution, 54.5% of the total premature death is attributed to COPD, while the least (3.0%) is to lung cancer.
3. Without baseline mortality adjustment, all-India annual premature death from ambient PM_{2.5} exposure is lower by 14.7%. However, in economically stronger state, the premature death is estimated to be higher (e.g. in Delhi by ~39%).
4. If India manages to achieve the national air quality target of 40 µg m⁻³, 44,900 (5900–173,300) less annual premature death is expected.

Acknowledgements

We acknowledge financial support from the Department of Science and Technology, Govt. of India through a research grant (DST/CCP/(NET-2)/PR-36/2012(G)) under the network program of climate change and human health, currently operational at IIT Delhi (RP2726). We thank Prof. Kirk Smith (UC Berkeley) and Prof. Aaron Cohen (IHME, Washington University) for helpful discussion. We acknowledge the anonymous reviewers for comments that helped improving the earlier version of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.envint.2016.03.004>.

References

- Anenberg, S.C., Horowitz, L.W., Tong, D.Q., West, J.J., 2010. An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling. *Environ. Health Perspect.* 118 (10), 1189–1195. <http://dx.doi.org/10.1289/ehp.0901220>.
- Apte, J.S., Marshall, J.D., Cohen, A.J., Brauer, M., 2015. Addressing global mortality from ambient PM_{2.5}. *Environ. Sci. Technol.* 49, 8057–8066 ([10.1021/acs.est.5b01236](https://doi.org/10.1021/acs.est.5b01236)).
- Balakrishnan, K., Ghosh, S., Ganguli, B., Sambandam, S., Bruce, N., Barnes, D.F., Smith, K.R., 2013. State and national household concentrations of PM_{2.5} from solid cookfuel use: results from measurements and modeling in India for estimation of the global burden of disease. *Environ. Heal. Glob. Access Sci.* 12 (1), 77. <http://dx.doi.org/10.1186/1476-069X-12-77>.
- Brauer, M., Amann, M., Burnett, R.T., Cohen, A., Dentener, F., Ezziati, M., ... Thurston, G.D., 2012. Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. *Environ. Sci. Technol.* 46, 652–660. <http://dx.doi.org/10.1021/es2025752>.

- Burnett, R.T., Arden Pope, C., Ezzati, M., Olives, C., Lim, S.S., Mehta, S., ... Cohen, A., 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ. Health Perspect.* 122 (4), 397–403. <http://dx.doi.org/10.1289/ehp.1307049>.
- Chafe, Z. a, Brauer, M., Klimont, Z., Van Dingenen, R., Mehta, S., Rao, S., Riahi, K., Dentener, F., Smith, K.R., 2014. Household cooking with solid fuels contributes to ambient PM2.5 air pollution and the burden of disease. *Environ. Health Perspect.* 122, 1314–1320. <http://dx.doi.org/10.1289/ehp.1206340>.
- Cohen, A.J., Ross Anderson, H., Ostro, B., Pandey, K.D., Krzyzanowski, M., Künzli, N., ... Smith, K., 2006. The global burden of disease due to outdoor air pollution. *J. Toxicol. Environ. Health A* 68 (October 2013), 1301–1307. <http://dx.doi.org/10.1080/15287390590936166>.
- Dey, S., Di Girolamo, L., 2010. A climatology of aerosol optical and microphysical properties over the Indian subcontinent from 9 years (2000–2008) of Multiangle Imaging Spectroradiometer (MISR) data. *J. Geophys. Res. Atmos.* 115, 1–22. <http://dx.doi.org/10.1029/2009JD013395>.
- Dey, S., Di Girolamo, L., 2011. A decade of change in aerosol properties over the Indian subcontinent. *Geophys. Res. Lett.* 38 (May), 1–5. <http://dx.doi.org/10.1029/2011GL048153>.
- Dey, S., Di Girolamo, L., van Donkelaar, A., Tripathi, S.N., Gupta, T., Mohan, M., 2012. Variability of outdoor fine particulate (PM2.5) concentration in the Indian subcontinent: a remote sensing approach. *Remote Sens. Environ.* 127, 153–161. <http://dx.doi.org/10.1016/j.rse.2012.08.021>.
- Health Effects Institute Research Report, 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. *Health (San Francisco)* 157 Retrieved from <http://pubs.healtheffects.org/getfile.php?u=623>.
- Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., ... Ezzati, M., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380, 2224–2260. [http://dx.doi.org/10.1016/S0140-6736\(12\)61766-8](http://dx.doi.org/10.1016/S0140-6736(12)61766-8).
- Murray, C., et al., 2015. Global, regional and national comparative risk assessment of 76 behavioural, environmental, occupational and metabolic risks or clusters of risks in 188 countries 1990–2013 : a systematic analysis for the GBD Background. *Lancet* 6736, 1–27. [http://dx.doi.org/10.1016/S0140-6736\(15\)00128-2](http://dx.doi.org/10.1016/S0140-6736(15)00128-2).
- Pope, A., Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Thurston, G.D., 2002. Cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Lung Cancer* 287 (9).
- Pope, C.A., Burnett, R.T., Turner, M.C., Cohen, A., Krewski, D., Jerrett, M., ... Thun, M.J., 2011. Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure–response relationships. *Environ. Health Perspect.* 119 (11), 1616–1621. <http://dx.doi.org/10.1289/ehp.1103639>.
- Rich, D.Q., et al., 2010. Triggering of transmural infarctions, but not nontransmural infarctions, by ambient fine particles. *Environ. Health Perspect.* 118, 1229–1234.
- Silva, R. a, West, J.J., Zhang, Y., Anenberg, S.C., Lamarque, J.-F., Shindell, D.T., ... Zeng, G., 2013. Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change. *Environ. Res. Lett.* 8, 034005. <http://dx.doi.org/10.1088/1748-9326/8/3/034005>.
- Smith, K.R., Bruce, N., Balakrishnan, K., Adair-Rohani, H., Balmes, J., Chafe, Z., ... Rehfuess, E., 2014. Millions dead: how do we know and what does it mean? Methods used in the comparative risk assessment of household air pollution. *Annu. Rev. Public Health* 35, 185–206. <http://dx.doi.org/10.1146/annurev-publhealth-032013-182356>.
- Van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., Villeneuve, P.J., 2010. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ. Health Perspect.* 118 (6), 847–855. <http://dx.doi.org/10.1289/ehp.0901623>.
- Venkataraman, C., et al., 2005. Residential biofuels in South Asia: carbonaceous aerosol emissions and climate impacts. *Science* 307, 1454–1456.
- Wellenius, G.A., Burger, Mary R., Coull, Brent A., S., J., Suh, Helen H., Koutrakis, Petros, Schlaug, Gottfried, G., D.R., M., A.M., 2012. Ambient air pollution and the risk of acute ischemic stroke. *Arch. Intern. Med.* 127 (3), 358–366. <http://dx.doi.org/10.1016/j.jsbmb.2011.07.002.Identification>.
- WHO, 2011. *World Health Statistics 2011*. WHO Libr. Cat. Data 1 170 (doi:978 92 4 156419 9).