Research

Mortality Associated with Ambient PM_{2.5} Exposure in India: Results from the Million Death Study

Patrick E. Brown,¹[®] Yurie Izawa,¹ Kalpana Balakrishnan,²[®] Sze Hang Fu,¹ Joy Chakma,³ Geetha Menon,³[®] Rajesh Dikshit,⁴ R.S. Dhaliwal,³[®] Peter S. Rodriguez,¹[®] Guowen Huang,¹ Rehana Begum,¹ Howard Hu,⁵ George D'Souza,^{6*} Randeep Guleria,^{7*} and Prabhat Jha^{1*}

¹Centre for Global Health Research (CGHR), St Michael's Hospital and Dalla Lana School of Public Health, The University of Toronto, Toronto, Ontario, Canada

²Department of Environmental Health Engineering, Sri Ramachandra Institute of Higher Education and Research, Chennai, India

³The Indian Council of Medical Research, New Delhi, India

⁴Centre for Cancer Epidemiology, Tata Memorial Centre, Mumbai, India

⁵Department of Preventive Medicine, Keck School of Medicine of University of Southern California, Los Angeles, USA

⁶St. John's Medical College, St. John's Research Institute, Bangalore, India

⁷All India Institutes of Medical Sciences, New Delhi, India

BACKGROUND: Studies on the extent to which long-term exposure to ambient particulate matter (PM) with aerodynamic diameter $\leq 2.5 \ \mu m \ (PM_{2.5})$ contributes to adult mortality in India are few, despite over 99% of Indians being exposed to levels that the World Health Organization (WHO) considers unsafe.

OBJECTIVE: We conducted a retrospective cohort study within the Million Death Study (MDS) to provide the first-ever quantification of national mortality from exposure to PM_{2.5} in India from 1999 to 2014.

METHODS: We calculated relative risks (RRs) by linking a total of ten 3-y intervals of satellite-based estimated $PM_{2.5}$ exposure to deaths 3 to 5 y later in over 7,400 small villages or urban blocks covering a total population of 6.8 million. We applied using a model-based geostatistical model, adjusted for individual age, sex, and year of death; smoking prevalence, rural/urban residency, area-level female illiteracy, languages, and spatial clustering and unit-level variation.

RESULTS: $PM_{2.5}$ exposure levels increased from 1999 to 2014, particularly in central and eastern India. Among 212,573 deaths at ages 15–69 y, after spatial adjustment, we found a significant RR of 1.09 [95% credible interval (CI): 1.04, 1.14] for stroke deaths per 10-µg/m³ increase in $PM_{2.5}$ exposure, but no significant excess for deaths from chronic respiratory disease and ischemic heart disease (IHD), all nonaccidental causes, and total mortal-ity (after excluding stroke). Spatial adjustment attenuated the RRs for chronic respiratory disease and IHD but raised those for stroke. The RRs were consistent in various sensitivity analyses with spatial adjustment, including stratifying by levels of solid fuel exposure, by sex, and by age group, addition of climatic variables, and in supplementary case–control analyses using injury deaths as controls.

DISCUSSION: Direct epidemiological measurements, despite inherent limitations, yielded associations between mortality and long-term $PM_{2.5}$ inconsistent with those reported in earlier models used by the WHO to derive estimates of $PM_{2.5}$ mortality in India. The modest RRs in our study are consistent with near or null mortality effects. They suggest suitable caution in estimating deaths from $PM_{2.5}$ exposure based on MDS results and even more caution in extrapolating model-based associations of risk derived mostly from high-income countries to India. https://doi.org/10.1289/EHP9538

Introduction

Epidemiological studies of the health consequences of exposure to particulate matter (PM) with aerodynamic diameter $\leq 2.5 \ \mu m$ (PM_{2.5}) have been conducted mostly in high-income countries with lower levels of exposure than those in India. In 2014, average Indian exposure to PM_{2.5} was 47 $\mu g/m^3$ based on satellite-derived estimates (Brauer et al. 2016), with over 99% of India's population exposed to levels above the ceiling of 10 $\mu g/m^3$ recommended as "safe" by the World Health Organization (WHO) (WHO 2006). In India, household solid cooking fuels contribute to 20%–50% of ambient PM_{2.5} exposure, a much higher proportion than in other countries (Chowdhury et al. 2019). With such high levels of exposure, even modest excess mortality caused by PM_{2.5} would result

Supplemental Material is available online (https://doi.org/10.1289/EHP9538).

in a substantial number of deaths. Indian hazards likely differ substantially from those studied in high-income countries given distinct disease patterns, including high age-standardized death rates for ischemic heart disease (IHD) and stroke (Ke et al. 2018).

In the absence of nationally representative evidence on the extent to which PM_{2.5} exposure increases cause-specific mortality rates in India, the WHO and others have relied on modeled associations that apply various complex transformations to hazards observed in studies mainly in high-income countries (Burnett et al. 2018). The WHO thus estimates that 1.1 million Indians of all ages died from exposure to ambient air pollution in 2016, to which $PM_{2.5}$ is the major contributor (WHO 2018). In these model-based death totals, about three-quarters were from respiratory and vascular disease in adults. Here, we conducted a retrospective cohort study within the Million Death Study (MDS) in India to quantify the association of cause-specific adult mortality with long-term exposure to PM2.5. We focused on quantification of the risks specific to the Indian population and considered misclassification of exposures and diseases. Given that many previous assessments have not appropriately considered clustering of exposure and disease, we pay specific attention to showing the variation in estimated risks with and without spatial adjustment.

Methods

Satellite-Based PM_{2.5} Estimates

We obtained satellite-derived modeled estimates of ambient PM_{2.5} data from 1998 to 2014 from van Donkelaar et al. (2015). Briefly,

^{*}Joint Senior authors.

Address correspondence to Patrick Brown, Centre for Global Health Research, St. Michael's Hospital and University of Toronto, ON M5B 2C5 Canada. Telephone: (416) 864-6042. Email: CanadaPatrick.Brown@Utoronto.ca

The authors declare they have no actual or potential competing financial interests.

Received 22 April 2021; Revised 11 July 2022; Accepted 2 August 2022; Published 14 September 2022.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

as described by Boys et al. (2014), van Donkelaar et al. (2015) combined data from three types of satellite instruments together with the GEOS-Chem chemical transport model (version 9-01-03; http://geos-chem.org) to estimate global surface PM2.5 concentrations at a resolution of 0.1 degrees (approximately 11 km at the equator) for the period 1998-2012. The satellite-based data do not distinguish between sources of PM2.5, which differ substantially across countries (e.g., emissions from motor vehicles, industrial facilities, household solid fuel use) and within India [e.g., in northern India during seasons of crop burning, and in rural areas with high levels of use of household solid fuels; Balakrishnan et al. (2013)]. van Donkelaar et al. (2015) derived 3-y medians of the annual satellite-based PM2.5 grids to reduce noise in the annual satellite-derived values. In addition, van Donkelaar et al. (2015) evaluated their 3-y medians of PM2.5 satellite-derived estimates using data from established ground-level monitoring networks in Canada, the United States, and Europe and collected 210 ground-based observations from the literature for areas outside these regions. Correlations between satellite-based estimates and ground-level data were stronger with more prolonged measurements. For example, for ground data collected for a year or less, the correlation coefficient, R, was 0.67, whereas R was 0.79 for ground-level data collected over 10 y. Correlations varied also by region: for Canada and the United States, the R was 0.76; for Europe, the R was 0.73; and for the rest of the world the *R* was 0.81.

With respect to India, van Donkelaar et al. (2015) evaluated their PM_{2.5} estimates using 55 different locations (representing >250 points in time) distributed over 24 major cities (Figure S1). Although some high population density areas were not represented, all Indian mega-cities (population >4 million) except Surat were represented. Using New Delhi as an example, the authors reported a local underestimation in annual mean PM_{2.5} in comparison with ground observations that seemed to be amplified during the winter season (Figure S2C). Similarly, the authors found that the mean (population–weighted) PM_{2.5} concentration for South Asia was 34.6 μ g/m³ [standard deviation (SD) = 15.8] for the period 2001–2010 and that concentrations increased by 2.9%/year [95% confidence interval (CI): 2.2%, 3.6%] for this region in the period 1998–2012.

We further compared alternative satellite $PM_{2.5}$ data from the public domain to ground monitoring data sources from various locations and times (Supplemental Material, "Comparison of $PM_{2.5}$ Data Sources"). Table S1 lists these satellite-derived $PM_{2.5}$ data sources together with a few ground-level sources, and Figure S3 displays the various satellite-derived $PM_{2.5}$ data for the year 2010. We found that the data set by van Donkelaar et al. (2015) has the highest correlation (0.88; p < 0.001) with January ground-level data (which is the peak month for $PM_{2.5}$ exposure) and the second highest correlation with the ground-level annual averages (Figure S2A,B).

The MDS

Full details of the methods, strengths, and limitations of the MDS have been published earlier (Aleksandrowicz et al. 2014; Dikshit et al. 2012; Menon et al. 2019; Million Death Study Collaborators 2017). Briefly, in collaboration with the Registrar General of India, the MDS monitored approximately 14 million people in 2.4 million nationally representative households in India from 1998 through 2013 in the Sample Registration System (SRS) (RGI and Census Commissioner 2011a). The sample frames comprised 6,671 (2001–2003) and 7,597 (2004–2013) sampling units, typically villages or small census urban blocks, drawn from the preceding 1991 and 2001 censuses, respectively. Nationally representative demographic

statistics can be derived from each sampling frame. Each sampling unit undergoes population enumeration at the beginning of each 10-y period. We used the 2004-2013 SRS sample frame to assign PM2.5 estimates and deaths because we had population denominators of the sampling units. Each sampling unit comprised $\sim 150-300$ households. The SRS follows the population within these sampling units prospectively for a 10-y period for any deaths. Every 6 months, ~900 full-time, trained, nonmedical surveyors update the population composition in these households, capture the date of death, and investigate each death using a well-validated verbal autopsy instrument [based on the WHO (2012) instrument, including a detailed two-page structured list of symptoms and a half-page local language narrative of the circumstances, symptoms, and treatments, if any, prior to death] (Aleksandrowicz 2014). Each field report of a death is converted to an electronic form and randomly assigned to 2 of 404 specially trained physicians. The two physicians independently assigned a cause of death based on the nonmedical surveyors' field report of the death, using the WHO's International Classification of Diseases-10th revision (ICD-10) (WHO 2004), according to the cause-of-death assignment manual specifically designed for the MDS (RGI, SRS Collaborators, and CGHR 2011; Menon et al. 2019). Where the two diagnoses did not agree, they underwent anonymous reconciliation by the same two physicians, and a third senior physician anonymously adjudicated persisting differences (which occurred $\sim 18\%$ of the time; Table S2). The percentage of deaths requiring adjudication differed little by sex, urban-rural setting, geographic region, or PM_{2.5} exposure level but was higher among people older than 69 y of age (Table S2). Various quality control procedures resulted in generally low levels of illdefined deaths before age 70 y, which is a robust measure of the quality of the field collection and central physician assignment of causes (Aleksandrowicz et al. 2014; Gomes et al. 2017). Ethics approval for the MDS was obtained from St. John's Research Institute and St. Michael's Hospital, Toronto, Canada. All SRS participants provided oral consent at the beginning of each sampling frame. Households were free to withdraw, but <1% did so (Jha et al. 2006).

Deaths

We included the entire population who resided in the 2004–2013 sampling units. We focused on deaths from respiratory disease, IHD, stroke (Ke et al. 2018), and all nonaccidental deaths (ICD-10 codes shown in Table 1 footnote) at ages 15-69 y for the main analysis (Table S3). Indian life expectancy in 2018 was 69 y, and we used this age range to consider the younger age distribution of vascular death in India than in high-income countries (Ke et al. 2018). Moreover, the verbal autopsy instrument has lower levels of misclassification before age 70 y than above (Jha et al. 2019). To minimize misclassification of causes, the main analyses was of cause-specific deaths when both physicians initially agreed on the diagnosis, a smaller number than based on adjudication. In India, cardiovascular and respiratory diseases accounted for $\sim 30\%$ of total adult deaths in the period 2010–2013 (RGI and CGHR 2015). Lung cancer deaths are, as documented earlier, uncommon in India (Dikshit et al. 2012; Jha et al. 2008) and were too few to be quantified reliably.

Geocoding and Ambient PM_{2.5} Estimates Linkage

We geocoded 7,416 MDS sampling units using the postal codes in urban areas and village names in rural locations. The SRS records sampling areas as either rural or urban relying on established Census of India's definitions, based chiefly on population

Variable	Stratum	Chronic respiratory	IHD	Stroke	Nonaccidental	All-cause	Population
Death count/population count		13,185	29,834	13,318	177,686	212,573	678,889,985
Age [y (median; 25th, 75th		62 (56, 66)	57 (47, 63)	60(53, 65)	56 (44, 64)	54(39, 63)	27 (12, 42)
percentile)]							
PM _{2.5} [μg/m ³ ; median (25th, 75th percentile)]	I	27.1 (18.7, 44.5)	23.5 (16.3, 39.8)	24.4(16.8, 35.2)	24.4 (16.9, 38.6)	24.3 (16.8, 38.3)	25.5(17.3, 41.4)
Percentage of household using		87.1% (67.2%, 94.0%)	72.3% (38.7%, 88.2%)	83.8% (57.0%, 93.4%)	81.5% (54.1%, 93.1%)	81.5% (54.7%, 93.0%)	81.8% (50.7%, 93.4%)
solid cooking fuel at subdistrict level ^a fmedian (25th: 75th							
percentile)							
Percentage of respondent smoked		11.9% (6.1%, 20.0%)	7.3% (3.2%, 13.3%)	9.8% (4.8%, 17.0%)	9.3% (4.4%, 16.9%)	$9.3\% \ (4.4\%, 16.7\%)$	10.0% (4.5%, 18.2%)
at sampling unit level ^{b} [median							
(25th, 75th percentile)]							
Percentage of female illiteracy at		57.4% ($44.6%$, $69.8%$)	46.9% (35.3%, 59.7%)	49.5%(37.7%, 61.5%)	51.2% (38.5%, 64.3%)	51.3% $(38.8%, 64.3%)$	51.0% (38.2%, 64.8%)
subdistrict level [median (25th,							
75th percentile)]							
Sex [count (%)]	Female	5,267 (39.9%)	8,805 (29.5%)	5,388(40.5%)	71,193 (40.1%)	82,002 (38.6%)	331,433,770 (48.8%)
	Male	7,918(60.1%)	21,029 (70.5%)	7,930(59.5%)	106,493 $(59.9%)$	130,571 (61.4%)	347,456,201 (51.2%)
Residence [count (%)]	Rural	10,854 (82.3%)	20,754 (69.6%)	9,823 (73.8%)	133,599 (75.2%)	161,092 (75.8%)	499,531,510 (73.6%)
	Urban	2,331 (17.7%)	9,080 (30.4%)	3,495(26.2%)	44,087 (24.8%)	51,481 (24.2%)	179,358,480 (26.4%)
Indian region [count (%)]	Northern	8,291 (62.9%)	13,408 (44.9%)	8,382 (62.9%)	101,485 (57.1%)	119,843 (56.4%)	419,655,310 (61.8%)
1	Southern	4,894(37.1%)	16,426 (55.1%)	4,936(37.1%)	76,201 (42.9%)	92,730 (43.6%)	259,234,670 (38.2%)
Note: Deaths were due to asthma and chr G83), nonaccidental causes (ICD-10: A00 ized — no data PM5 e marticulate marte	onic obstructive)-R99), and all- r with aerodyna	e pulmonary disease (chronic 1 cause (ICD-10: A00-R99, U00 unic diameter equal to or small	respiratory) (ICD-10: J40-J47 0-Y99). Covariate informatior ller than 2.5 um	r, 163, 193), ischemic heart dis n included in main analyses fo	ease (IHD) (ICD-10: 120-125, I r ages 15–69 y is presented, an	44, 146, 170, R55, R96), stroke d death counts used on stratifie	(TCD-10: 160-169, G45-G46, G81- d sensitivity analyses are summar-

ized. —, no data; PM₂₅, particulate matter with aerodynamic diameter equal to or smaller than 2.5 µm. "The percentages of households using solid fuels for cooking (coal, firewood, dung, and crop residue) at the subdistrict-level were obtained from the Census of India 2001 and 2011; the percentages were linked to each death based on the resi-"The percentages of households using solid fuels for cooking (coal, firewood, dung, and crop residue) at the subdistrict-level were obtained from the Census of India 2001 and 2011; the percentages were linked to each death based on the resi-dence locations, with deaths from 2004–5 and 2006–13 used percentages from Census 2001 and 2011, respectively. ^bSmoking percentage (prevalence of either bidis or cigarettes) were used as a covariate in Poisson models.

Table 1. Descriptive statistics of covariates for cause-specific (n = 56, 337), nonaccidental (n = 177, 686), and all-cause (n = 212, 573) deaths, as well as statistics for the general population (n = 678, 889, 985) from

size and population density (RGI and Census Commissioner 2011b). Urban sampling units were geocoded to postal code because postal code information was commonly available for the death records, and rural sampling units to village locations because rural sampling units are located within villages. Only 0.2% of urban sampling units and 4.0% of rural sampling units could not be geocoded due to unavailability of postal code or village information. Figure S4 displays the locations of the geocoded units. We overlaid the geocoded sampling units on the grid cells of the satellite-derived PM_{2.5} estimates. We performed exploratory analyses on lag years of PM_{2.5} by linking the 3-y median of ambient PM_{2.5} estimates to deaths that occurred 2-4 y, 3-5 y, and 4-6 y later based on year of death (e.g., for a 3- to 5-y lag, 3-y medians of 1999–2001 for $PM_{2.5}$ were linked to deaths from year 2004; see Figure S5). We found that using lag 2-4 y would cause PM_{2.5} exposures to be protective on respiratory disease and IHD, whereas using lag 4-6 y the effect of PM_{2.5} exposures on stroke was smallest in comparison with lag 2-4 y and lag 3-5 y (Figure S6). Thus, to be consistent for all three disease outcomes and avoiding an implausible protective effect of $PM_{2.5}$, we chose to use lag 3–5 y. Using the 3-5 y lag, the effect of PM2.5 exposures had stronger association with respiratory and IHD mortality, and intermediate association with stroke mortality (Figure S6). We assigned the 3-y median PM_{2.5} values from the period of 1999–2010 (i.e., a total of ten 3-y intervals) and linked 3-y median of ambient PM2.5 estimates to deaths 3-5 y later in each village or urban block. We were able to link 99.8% of deaths to a satellite-based PM2.5 estimates.

Statistical Methods

We modeled each individual's mortality risk as a function of individual-level risk factors (age and sex), sampling unit-level risk factors (urban/rural, smoking prevalence), a time trend, and spatially attributable risk factors ($PM_{2.5}$, female illiteracy, dominant language). Further, we allowed for random variation at the unit-level and a smoothly varying spatial random effect. We used model-based generalized linear geostatistical model (Diggle and Ribeiro 2007), currently the state-of-art method for modeling point-referenced spatial data in a hierarchical modeling framework (Gelfand and Banerjee 2017). The approach is broadly similar to the methods used by Bhatt et al. (2015) for estimating malaria prevalence in Africa.

In our model, an individual in age-sex group *j* living in sampling unit *i* at time *t* has a mortality risk $\lambda_{it}\theta$, where θ is the baseline rate for group j in the Indian population, and λ_{it} is the relative risk (RR). Because the probability of death in any given year is small, we can approximate the binary-valued living/dead response variable with the commonly used Poisson model for the unit-level annual mortality counts (Lawson 2018). The RRs by age and sex group are estimated by fitting a simple Poisson regression model with age, sex, and time as explanatory variables; with a large data set the uncertainty in the estimated θ_i is negligible. By treating these age-sex "reference rates" as fixed, the observed death counts and age-sex adjusted expected counts for each unit and year are sufficient statistics for making inference on the unit-level and spatial risk factors (see Li et al. 2012). The resulting response variable Y_{it} is the total death count at sampling unit *i* and time *t*, which is Poisson distributed with mean $\lambda_{it} P_{it}^{T} \theta$, where P_{it} is the vector of population counts by age-sex group j. Our model ignored multiple adult deaths in one household because these are rare (on average, 0.2% within 1 study year) in the MDS. The model is fit separately for each outcome (respiratory disease, IHD, and stroke).

Our models have the following formula for the RR λ_{it} for individuals living in sampling unit *i* at time *t*:

$$\begin{split} Y_{it} \sim \text{Poisson}(\lambda_{it}E_{it})\\ \log(\lambda_{it}) = X_{it}\beta + W(s_i,t)\alpha + g[PM_{2.5}(s_i,t);\gamma] + U(s_i) + Z_i\\ Z_i \sim N(0,\tau^2)\\ U(s_i) \sim N(0,\sigma^2)\\ cor[U(s+h),U(s)] = Matern(|h|/\varphi;1). \end{split}$$

Here E_{it} is the expected number of cases (see Supplemental Material, "Statistical Methods"), s_i denotes the spatial location of sampling unit *i*, X_{it} are covariates at the unit-level, and spatially referenced covariates are denoted (s_i, t) . The satellite-derived PM_{2.5} value $PM_{2.5}$ (s_i, t) has a nonlinear effect on risk, specified by the function $g(\cdot)$ which can be thought of as a "wiggly" line. More specifically, $g(\cdot)$ is a second-order random walk with variance parameter γ , the smaller the value of γ the straighter $g(\cdot)$ will be. To be compatible with the most recent global exposure of mortality model (GEMM) (Burnett et al. 2018), we assumed that PM_{2.5} exposure in the range of $0-2\,\mu g/m^3$ (which is rare in India) had no effect on mortality (values below 2 were discretized to the value of 2). The X_{it} includes an intercept, a linear time trend, urban/rural residency, and smoking prevalence (bidi or cigarette) in the sampling unit (i.e., percentage of smokers among the MDS survey respondents). The spatially referenced covariates (s_i, t) include subdistrict-level (each subdistrict area contains approximately 1-2 MDS sampling units) percentage of female illiteracy (continuous; as a proxy for poverty) and dominant language groups (as a proxy for cultural and dietary differences, which may also be clustered geographically); both were derived from the 2011 Indian census (Figure S7). We did not add subdistrict-level solid fuel use from the 2011 Indian census as a covariate because it contributes to ambient PM_{2.5}.

To account for potential spatial variation, the model included smoothly varying spatial random effects (i.e., spatial autocorrelation or clustering) $U(s_i)$. Similarly, we included independent unitlevel random effects Z_i to account for sampling unit-level variation. Unlike the spatial effect $U(s_i)$, two villages in close proximity can have very different values of Z_i . Thus, the Z_i can be thought of as accounting for short-scale spatial variation or village-level risk factors not included in the model as covariates, whereas $U(s_i)$ can be interpreted as representing any unknown or unobserved spatially varying risk factors. The spatial random effect $U(s_i)$ has the correlation between the RRs at different sampling unit locations specified by a Matérn spatial correlation function with a shape parameter fixed at 1.0. The spatial range parameter ϕ controls how quickly correlation decays with distance, with large values of ϕ leading to smooth surfaces. The variance parameter σ^2 controls the importance that this spatial effect has in determining mortality risk, with σ^2 near zero corresponding to a small effect (see Moraga 2019).

The regression coefficients β , α were given uninformative normal priors. The prior distributions for τ , σ , and $1/\phi$ have exponential priors following the "penalized complexity" framework of Simpson et al. (2017). These priors discourage a spatial effect [wanting $U(s_i)$ flat and close to zero] unless the data indicate a clear preference for a spatial model. When the posterior estimate for σ is far away from zero and its credible interval is relatively narrow, there is evidence that spatial adjustment is necessary as informed by the data.

We applied two additional variations on our model for the main analysis. First, we treated PM_{2.5} exposure as a linear effect, estimating RRs for a $10 \,\mu g/m^3$ increase [with $g(\cdot)$ being linear with a slope parameter]. Second, we ran linear and nonlinear models that did not include adjustment for spatial clustering, that is, we omitted the $U(s_i)$ term in our model (or equivalently set $\sigma = 0$). We compared our results of nonlinear effects of PM_{2.5} to the GEMM (Burnett et al. 2018) and the Prospective Urban Rural

Epidemiology (PURE) study, which examined nonfatal vascular events and deaths in about 750 communities in 21 countries (Hystad et al. 2020). The cohort studies included in the GEMM adjusted for individual characteristics and contextual variables, but most studies did not include community-level random effects or spatial random effects (Burnett et al. 2018). The PURE study adjusted for individual- and household-level characteristics, geographic covariates, as well as community-level random effects. We used R packages geostatsp (Brown 2015) and R-INLA (Rue et al. 2009) for statistical analyses. We report the median RR or excess risks (defined as RR minus 1) and 95% credible intervals (CI) from the posterior distributions of the Bayesian estimates. Further details of the model fitting and the prior distributions used can be found in Supplemental Material, "Statistical Methods."

We performed sensitivity analyses focusing on the linear effect of PM_{2.5} with spatial adjustment (by the smoothly varying spatial random effects) to explore whether the RRs of PM2.5 exposure differed across subgroups using stratified models. Interactions between PM2.5 and other potential risk factors were explored by fitting stratified models. We stratified by age (15-44, 45-69, and 70 y and above), by sex, in areas with high (87.3%-100.0%), medium (44.6%-87.3%), and low percentage (0.4%-44.6%) of households using solid cooking fuels (including firewood, crop residue, dung, and coal) at the subdistrict-level. We obtain solid cooking fuel use percentages at the subdistrict-level from the Census of India 2001 and 2011; the percentages were linked to each death based on the residence locations, with deaths from 2004-2005 and 2006-2013 used percentages from 2001 and 2011 censuses, respectively. We used tertiles to classify each deceased individual into these high, medium, and low solid fuel use levels. Stratification also included rural sampling units (urban sampling units were too few to produce estimate reliably) and northern vs. southern region of India (based on zonal councils https://www. mha.gov.in/zonal-council). Additional sensitivity analyses excluded language group from the covariates and included climatic variables. For climatic variables, we included temperature and relative humidity as covariates in the models. We obtained monthly gridded data of 2.5-degree resolution for surface temperature and relative humidity (Kalnay et al. 1996). We calculated their annual monthly averages and obtained 3-y medians of these annual averages. Furthermore, we explored the possible confounding effect of secondhand smoking exposure by focusing on deceased females who lived with or without male smokers as the SRS respondent in the household. Because very few females in India smoke, women are the population most likely exposed only to secondhand smoke (Jha et al. 2006).

Finally, we calculated odds ratio (OR) for chronic respiratory, IHD, and stroke deaths as cases and injury deaths as controls (Table S3). We used injury deaths as controls because in exploratory analysis using 2001-2003 MDS deaths, we found that death rates for injury deaths did not vary substantially by different levels of PM2.5 exposures (Figure S8). We fitted a logistic regression model with a spatial random effect to the case or control status of each subject, adjusting for individual sex, year, urban/rural residency, age (continuous years), current smoking (none, only cigarettes, only bidi, both as was done by Jha et al. 2008 and other studies using the MDS data) interacted with sex, and subdistrict percentage of female illiteracy and dominant language groups derived from the 2011 Indian census. As in the original cohort model, the logistic model included terms to consider potential spatial and sampling unit-level variation. The benefit of this "case-control" approach is to act as a check on possible misclassification of PM2,5 exposures, which should not affect injury deaths and cases differently.

Results

Over 99% of India's population lived throughout the study period in settings where satellite-based $PM_{2.5}$ exceeded $10 \,\mu g/m^3$. The

exposure to $PM_{2.5}$ rose in India (3-y median of $25 \,\mu g/m^3$ in 1998–2000 to $40 \,\mu g/m^3$ in 2012–2014), with notable hot spots in the north of India, south of the Himalayan mountain range, and West Bengal (Figure 1). The level of $PM_{2.5}$ varies by regions, with central and eastern regions having the highest increase in $PM_{2.5}$ level and the southern region with the least increase (Figure S9).

At ages 15 y or older, we excluded 1.2% of 96,359 respiratory and vascular disease deaths because of missing covariates (Table S4). The excluded population were slightly older, located in sampling units with higher 3-y median of ambient PM_{2.5} satellitebased estimates lagged 3-5 y and had other slight differences from those included (Table 1; Table S4). The main comparison was of 27,553 deaths from chronic respiratory disease, 42,950 from IHD and 24,940 from stroke, or a total of 95,443 deaths, including 56,337 at ages 15-69 y (Table S3; Table S5-7), drawn from a total population of 6.8 million (Table 1). The remaining deaths occurred above age 70 y (Table S8). Table 1 provides summary statistics of the covariates by disease condition. At ages 15–69 y, the median ages at death from chronic respiratory disease, IHD, stroke, all-cause, and nonaccidental causes were older than the median age of the Indian population, and hence all analyses adjusted for age. Deaths from chronic respiratory disease were more rural than in the overall population, had the highest mean PM_{2.5} levels of exposure, more smokers in the sampling unit, higher female illiteracy, and a higher percentage of solid fuel use in the area than in the overall population. IHD deaths occurred more in urban residents than the overall population, and thus these deaths reported lower PM_{2.5} levels of exposure, fewer smokers in the sampling unit, lower solid fuel use, and lower female illiteracy. Stroke deaths showed levels of PM_{2.5} exposure, smoking prevalence at sampling unit level, urban residence, and female illiteracy at subdistrict-level that were generally similar to those of the overall population. The geographic distributions of the three main conditions were distinctive as evident from the distributions by language group regions (Table S5; Figure S7A). More deaths occurred in the northern region, with the exception of IHD, which occurred more in the southern region (Table 1).

After adjusting for death year, urban/rural residency, smoking prevalence in sampling units, subdistrict-level female illiteracy, and dominant language groups, the RRs at ages 15-69 y for both sexes for every $10-\mu g/m^3$ increase in 3-y median of ambient PM_{2.5} lagged 3-5 y varied with adjustment for spatial clustering (Figure 2). The RR for chronic respiratory disease deaths, which occurs at much higher rates in north India (RGI and CGHR 2015), were substantially lower after spatial adjustment, from 1.05 (95% CI: 1.03, 1.08) to 1.01 (95% CI: 0.96, 1.05), representing a nonsignificant 1% excess risk. Further stratification for rural areas, and additional climatic covariates led to only minor changes in the RR. Similarly, the RR for IHD became nonsignificant after spatial adjustment, from 1.02 (95% CI: 1.01, 1.04) to 1.00 (95% CI: 0.97, 1.03) and varying little in rural areas or with additional climatic covariates (Figure 2). By contrast, the RR for stroke was higher with spatial adjustment [1.09 (95% CI: 1.04, 1.14)] than without spatial adjustment [0.91 (95% CI: 0.89, 0.93)], representing a statistically significant excess risk of 9% (Figure 2). For each of the above three conditions, the OR for death compared to injury controls in the case-control analyses showed similar results to the cohort analysis (Figure 2). The association of $PM_{2.5}$ exposure lagged 3–5 y with all nonaccidental deaths varied only slightly without and with adjustment for spatial clustering [RR 0.99 (95% CI: 0.98, 1.00) vs. 1.01 (95% CI: 1.00, 1.03), respectively]. For all-cause mortality, the spatially adjusted RR was 1.02 (95% CI: 1.01, 1.03), and remained statistically significant in the stratified analyses, except



Figure 1. Annual median satellite-based $PM_{2.5}$ values in micrograms per cubic meter from 2012 to 2014 and trends nationally and for selected cities in India from 1998 to 2014 (table). Three-year median of $PM_{2.5}$ levels (in micrograms per cubic meter) for selected cities at different time periods are displayed in the figure table. $PM_{2.5}$, particulate matter with aerodynamic diameter equal to or smaller than 2.5 µm. State administrative boundaries came from the 2011 Indian census. We used a simplified version of the state boundaries, which was created in house. Map was created using R software (version 4.1.0; R Development Core Team) with the "sp," "raster," "mapmisc," and "RColorBrewer" packages. Note: "The proportion of population exposed to ambient $PM_{2.5}$ (annual median over the years 2012–2014) was weighted using the Gridded Population of the World (version 4) for the year 2010 (https://sedac. ciesin.columbia.edu/data/set/gpw-v4-population-count-adjusted-to-2015-unwpp-country-totals-rev11). $\dagger PM_{2.5}$ for all of India are population-weighted values from van Donkelaar et al. (2015).

in the southern region (Figure 2). The 2% excess risk in all-cause mortality was mostly driven by stroke mortality, as excluding stroke from all-cause mortality risks yielded a nonsignificant RR of 1.01 (95% CI: 0.99, 1.02). The RRs in the southern region were consistently lower than in the northern region, but with wider credible confidence levels (Figure 2). The RRs excluding language as a covariate were similar to the main results (Figure S10). Table S9 provides the effects estimates for all covariates and parameters of random effects for models with spatial adjustment. The variance parameters for unit-level (τ) and spatial random effects (σ) have narrow CIs and are far away from zero, showing that both random effects are necessary to be included in the models.

Table 2 presents results stratified by sex, age group, and different levels of solid fuel use. Results for men and women were broadly similar, as were the results for ages 70 y or older. Stratification by high, medium, or low levels of solid fuel use at the local area also yielded broadly similar results, although the RR for chronic respiratory and stroke deaths in men showed higher RR than in women. Sensitivity analysis showed that the



Figure 2. Linear RRs per 10 microgram per cubic meter increase in $PM_{2.5}$ for chronic respiratory disease (n = 13,185), IHD (n = 29,834), stroke (n = 13,318), nonaccidental causes (n = 177,686), and all-cause mortality (n = 212,573) at ages 15–69 y, both sexes combined, from India in 2004–2013, for every 10 micrograms per cubic meter increase in $PM_{2.5}$ with and without spatial adjustment and in various sensitivity analyses. Linear effects for RRs were estimated within the generalized linear model framework, using the Poisson regression model. Models adjusted for death year, urban/rural residency, smoking prevalence in the sampling unit, subdistrict-level percentage of female illiteracy, and dominant language groups (from 2011 Indian census) and included smoothly varying spatial random effects and independent unit-level random effects. Box sizes are weighted by the sample sizes. Horizontal lines represent the 95% CI. Note: CI, credible interval; IHD, ischemic heart disease; OR, odds ratio; PM_{2.5}, particulate matter with aerodynamic diameter equal to or smaller than 2.5 µm; RR, relative risk. *We fitted a logistic model within the generalized linear regression, adjusting for individual sex, death year, urban/rural residency, age (continuous years), smoking (none, only cigarettes, only bidi, both) interacted with sex, and subdistrict percentage of female illiteracy and language groups. The logistic model used injury deaths as controls, which were shown to have zero association with exposure to $PM_{2.5}$ in a preliminary analysis based on MDS deaths occurring from 2001 to 2003 and $PM_{2.5}$ exposure levels in a special survey done in February 1998 (Figure S8). The reported estimates from the logistic model are OR.

RRs for female deaths were broadly similar when stratified by households with or without male smokers (Figure S11).

The estimated nonlinear models showed that RRs, without adjustment for spatial clustering, at ages 15–69 y in both sexes were higher with increasing levels of $PM_{2.5}$ for chronic respiratory and IHD deaths, but these RRs were lower notably with increasing $PM_{2.5}$ after adjustment for spatial clustering (Figure 3). By contrast, stroke deaths showed lower RR without adjustment but after adjusting for spatial clustering, these RRs were higher with increasing $PM_{2.5}$. Similarly, nonaccidental and all-cause deaths showed higher RR after adjusting for spatial clustering, with increasing $PM_{2.5}$ associated with higher RR. For all three conditions, nonaccidental, and all-cause deaths, RRs at higher levels of exposure had wide credible intervals.

The exposure–response relationships for chronic respiratory disease and IHD deaths in the MDS were much flatter than the modeled relationship of hazard ratios from the GEMM but were more comparable for stroke deaths (Figure 4). The exposure–response relationships of $PM_{2.5}$ with mortality in the MDS were similar to those of incident IHD or stroke in the comparable analyses used by the PURE study (Hystad et al. 2020; Figure 4). GEMM showed the steepest relationship at the youngest adult ages, whereas the MDS showed no such differences by age (Table 2). Even the 95% upper CI of the estimates showing RRs of 1.05, 1.03 and 1.14, respectively, for chronic respiratory disease, IHD, or stroke (Figure 2), were considerably lower than the mean GEMM modeled estimates (Burnett et al. 2018).

Discussion

To our knowledge, our large retrospective cohort study provides the first nationally representative epidemiological evidence of the association between cause-specific mortality and long-term PM_{2.5} exposure based on lagged satellite-derived PM2.5 exposure. We find a statistically significant excess risk of 9% (range 4%-14%) for stroke deaths at ages 15–69 y per $10-\mu g/m^3$ increase in exposure, but no significant excess risks for chronic respiratory disease, IHD deaths, or all nonaccidental deaths. The overall mortality excess risk was 2% (range 1%-3%) but was nonsignificant once stroke was excluded. The RRs included adjustment for both spatial and unitlevel variation in risks unexplained by covariates included in the models and were consistent in various sensitivity analyses including stratifying by age, sex, levels of solid fuel use, in rural areas, and focusing only on the northern regions where PM2.5 exposure is most widespread. Results on case-control analyses using injury controls showed nearly identical results once spatial clustering was considered.

The MDS association of PM2.5 exposure with IHD deaths is consistent with a zero excess risk. Respiratory disease deaths also showed close to a zero excess risk, although we noted higher risks among men than women, perhaps reflecting the residual effects of smoking history not captured in the community measures of smoking prevalence (given that few women smoke in India). The story for stroke is more nuanced and even our observed 9% excess risk may not be real. Stroke mortality is lowest in several northern states adjacent to the Himalayas: Punjab, Haryana, Rajasthan, Uttar Pradesh, and Bihar (Ke et al. 2018), states which coincidentally have the highest levels of satellite-measured ambient PM2.5 estimates (Figure 1). In the analysis without spatial adjustment, therefore, PM_{2.5} appears to be protective against stroke, an artifact that reduces with spatial adjustment. The model with spatial adjustment looked for a consistent relationship between more regional variation in PM2.5 and stroke mortality. Once we considered a smooth spatial effect to explain the large region of low stroke mortality in the north, we identified a modest but significant excess risk.

The RRs for adult respiratory disease mortality are modest in comparison with published model-based estimates. Chronic respiratory disease is more prevalent in the northern states, Rajasthan, Uttar Pradesh, and Maharashtra (Salvi et al. 2018). For IHD, higher prevalence is found in Punjab in the north and in Andhra Pradesh and Tamil Nadu in the south (Menon et al. 2019). Thus, deaths from chronic respiratory diseases (and to a lesser extent IHD) cluster in the same areas as does PM_{2.5} exposure, even after taking into account differences in subdistrict-level illiteracy rate or sampling unit–level smoking prevalence. The PURE study, which included community-level random effects in their models, also showed a lower RR for nonvascular deaths. Our estimates for stroke are comparable to the PURE results (Hystad et al. 2020).

The exposure variable is 3-y median of ambient and satellitederived PM2.5 estimates, which is widely used in environmental epidemiology due to its availability. At a minimum, our results show that the large RRs reported for satellite-derived PM2.5 elsewhere are not reproducible in India. However, the MDS results are not directly comparable to the modeling results from GEMM or the Global Burden of Disease (GBD) for several reasons. First, although our IHD RRs are comparable to the RRs in the PURE study and in some of the larger studies that were inputs to the GEMM/GBD meta-analyses, the modeled GEMM/GBD results comprise a metaanalysis of 41 cohort studies, with various additional transformations to create exposure-hazard relationships. Unfortunately, most of the input studies did not consider spatial clustering (Burnett et al. 2018). In contrast, the MDS estimates did not need to rely on any prior assumptions about exposure-response relationships or transformations of PM_{2.5} estimates, yielding markedly divergent exposure-response curves for IHD and chronic respiratory disease but more similar curves for stroke. Second, the range of 3-y median of ambient PM2.5 estimates studied in the MDS is far larger (and captures much higher ambient PM2.5 levels) than the relatively narrow range in the GEMM input studies. Last, the MDS RRs for the key conditions are smaller in comparison with GEMM HRs-0% to 9% excess mortality risks with a $10 - \mu g/m^3$ increase in ambient PM_{2.5} exposure. Even for stroke, we did not observe the near 2-fold hazards that GEMM reported for ambient PM_{2.5} exposure in highly polluted areas (like Delhi and Varanasi), which are close to the excess mortality documented for stroke from bidi or cigarette smoking in India (Jha et al. 2008). Modest risks from ecological exposures are difficult to quantify reliably. Hence, suitable caution is warranted in interpreting the MDS evidence, and even more caution is warranted in extrapolating GEMM/GBD's current model-based associations to India.

Our findings, if true, suggest a need for a substantial downward revision of the WHO estimates (which effectively use the GBD or GEMM results), which suggest that ambient PM_{2.5} accounts for 0.29 million, 0.39 million, and 0.14 million deaths, respectively, from chronic respiratory diseases, IHD, and stroke at ages 25 y and above in India (WHO 2018). These model-based estimates find that IHD contribute more than one-quarter of the total mortality due to exposure to ambient air pollution in India and worldwide (WHO 2018). However, studies linking ambient PM_{2.5} exposure to IHD have been mostly done in high-income countries where the levels of PM2.5 concentrations are lower and the range of exposure narrower (Burnett et al. 2018; Kaufman et al. 2016). We find a negligible association for IHD with $PM_{2.5}$ exposure in India. This finding might reflect differences in the susceptibility to PM2.5 in the Indian population and/or a difference in the composition and resulting toxicity of PM2.5 exposure in India. On the exposure side, we cannot exclude possibly missing more refined PM2.5 exposure that could increase vascular risk, such as ultrafine particles that may impact vascular end points (Schulz et al. 2005) and that may differ across countries



Figure 3. Nonparametric RR per unit (micrograms per cubic meter) increase in $PM_{2.5}$ exposure for chronic respiratory disease (n = 13,185), IHD (n = 29,834), stroke (n = 13,318), nonaccidental causes (n = 177,686), and all-cause mortality (n = 212,573) at ages 15–69 y, both sexes combined, from India in 2004–2013, with and without spatial adjustment. Nonparametric effects for RRs were estimated using the Poisson regression model. Models adjusted for death year, urban/rural residency, smoking prevalence in the sampling unit, subdistrict-level percentage of female illiteracy, and dominant language groups (from 2011 Indian census) and included smoothly varying spatial random effects and independent unit-level random effects. Solid curves are the median RR estimates; dotted curves are 95% CI (graph is limited to RRs between 0.0 and 2.5). Vertical black arrows at *x*-axis indicate national population-weighted average exposure of 47 µg/m³ ambient PM_{2.5} in 2014. Note: CI, credible interval; IHD, ischemic heart disease; RR, relative risk.



Figure 4. Comparison of exposure-response relationships of PM2.5 risk for chronic respiratory disease (top panel, n = 13,185), IHD (middle panel, n = 29,834), and stroke (lowest panel, n = 13,318) deaths in GEMM, PURE, and MDS from India, 2004-2013. Estimates are at various ages in GEMM (based on hazard ratios) and incidence of IHD and strokes in PURE at ages 35-70 y (based on hazard ratios) and in the MDS at ages 15-69 y (based on RRs). GEMM composed of a meta-analysis of 41 cohort studies with various transformations to create exposure-hazard relationships. Estimates for COPD from GEMM is not age specific. MDS estimates were adjusted for death year, urban/rural residency, smoking prevalence in the sampling unit, subdistrict-level percentage of female illiteracy, and dominant language groups (from 2011 Indian census) and included smoothly varying spatial random effects and independent unit-level random effects. PURE estimates were adjusted for age, sex, baseline year, smoking status, physical activity, PURE diet index score, waist-to-hip ratio, INTERHEART risk score, use of solid fuels for cooking, education level, household wealth index, occupational class, baseline cardiovascular disease and chronic conditions, cardiovascular disease medication use, hypertension, geographical covariates (urban or rural location, baseline country gross domestic product per capita, Night Light Development Index score, and a national or regional Health Care Access and Quality Index), and community random effect (Hystad et al. 2020). The comparison is to PURE model 3, which the authors of the study indicate is their preferred model. For GEMM's estimates for IHD and stroke, younger ages have higher hazard ratios (or steeper exposure-response curves) than older ages. Shaded color band represent regions of 95% CI for MDS estimates. Note: CI, credible interval; COPD, chronic obstructive pulmonary disease; GEMM, global exposure of mortality model; IHD, ischemic heart disease, MDS, Million Death Study; PURE, Prospective Urban Rural Epidemiology.



Figure 4. (Continued.)

(Burney and Amaral 2019) but are not well captured in satellitebased estimates. Notably, the PURE study found a 3% significant excess risk for vascular disease deaths with every $10-\mu g/m^3$ increase in PM2.5 exposure, but this excess risk disappeared once adjustment was made for dust and salt in satellite-based ambient PM_{2.5} estimates (Hystad et al. 2020). Thus, the effect of PM_{2.5} exposure is highly sensitivity to the type of exposure data and spatial adjustment. More speculatively, low levels of PM_{2.5} exposure, which are uncommon in India, might play a disproportionate role in increasing IHD (Papadogeorgou et al. 2019). Moreover, most IHD deaths in India involve conventional risk factors such as smoking (Jha et al. 2008), elevated blood pressure and blood lipids, and, unusually, low body mass index (BMI) (Gajalakshmi et al. 2018; Ke et al. 2021). Although our estimates of IHD mortality from ambient PM_{2.5} adjusted for smoking prevalence in sampling units, it is possible that smoking type and other risk factors measured at the individual level, such as BMI, diet, alcohol drinking, education level, and medical history, might have confounding effect on the mortality-exposure associations, as might access to effective treatments (Ke et al. 2018). In Chinese populations adjacent to northeast India, the relationship of ambient PM2.5 with stroke is largely with hemorrhagic strokes (Chen et al. 2018). Because verbal autopsy cannot distinguish between hemorrhagic and occlusive strokes, it is possible that the notably high stroke mortality in northeastern India noted earlier (Ke et al. 2018) might represent a different relationship than that found in China.

Several limitations exist in the MDS study, including measurement error in the ambient PM2.5 estimates, the potential omission of additional confounders (i.e., solid fuel use), and possible misclassification of the mortality outcomes. The key consideration is whether these various forms of model misspecification have attenuated our estimated RRs toward one, meaning we were unable to detect modest but perhaps clinically relevant excess risks. Overall, we do not believe we have systematically underestimated the RRs. First, adjustment for spatial clustering is appropriate for dealing with residual correlation in spatial data (Dormann et al. 2007). Cohort analyses for the United States have generally included random effects to represent spatial patterns in the data and found that hazards were significantly impacted by their inclusion (Krewski et al. 2009). Second, the uncertainties that exist in our PM2.5 exposure assessment method could cause either overestimation or underestimation of RRs (van

Table 2. Stratification an IHD $(n = 42, 841)$, stroke	nalyses of RRs per micrograms [$(n = 24, 899)$, nonaccidental cau	per cubic meter lises $(n = 315, 45)$	increase in $PM_{2.5}$ eror 7), and all causes (<i>n</i>	xposure, by = 358,662	y age group and solid () in India, 2004–201	d fuel use 3.	level for male and f	emale deat	hs from chronic res	piratory di	sease $(n = 27, 504)$,
		Chro	onic respiratory		DHI		Stroke	Nonac	cidental causes		All-cause
Group	Strata	Cases	RR (95% CI)	Cases	RR (95% CI)	Cases	RR (95% CI)	Cases	RR (95% CI)	Cases	RR (95% CI)
Male											
A 200 20000 (11)	15 40	521	1 00 /0 /0 1 10/	6699		1 400	1 07 /0 07 1 10/	000 20		100 22	

		Curc	onic respiratory		UHI		Suroke	INONAC	cidental causes		All-cause
Group	Strata	Cases	RR (95% CI)	Cases	RR (95% CI)	Cases	RR (95% CI)	Cases	RR (95% CI)	Cases	RR (95% CI)
Male											
Age group (y)	15-49	921	1.08(0.98, 1.19)	6,628	0.97(0.93, 1.02)	1,498	1.07 (0.97, 1.18)	37,390	0.99 (0.96, 1.02)	55,224	0.99 (0.97, 1.02)
	50-69	6,997	1.03 (0.98, 1.09)	14,401	1.02 (0.98, 1.05)	6,432	1.05 (0.99, 1.12)	69,103	1.00 (0.98, 1.02)	75,347	1.01 (0.99, 1.03)
	70+	8,288	0.97 (0.92, 1.01)	8,016	1.03 (0.98, 1.08)	5,953	1.05 (0.99, 1.11)	70,835	1.02 (1.00, 1.04)	75,029	1.02 (1.00, 1.04)
Solid fuel use level ^a	Low	1,149	1.08 (0.98, 1.18)	5,938	1.00 (0.96, 1.05)	1,559	0.98 (0.89, 1.07)	23,466	1.00 (0.97, 1.03)	28,092	1.00 (0.98, 1.03)
	(0.4% - 44.6%, median = 16.7%)										
	Medium	2,897	1.02 (0.95, 1.10)	9,656	1.02 (0.97, 1.07)	3,012	1.08 (1.00, 1.18)	42,232	1.03 (1.00, 1.06)	52,706	1.03 (1.01, 1.06)
	(44.6% - 87.3%, median = 72.3%)										
	High	3,871	1.07 (1.00, 1.14)	5,430	0.99 (0.94, 1.05)	3,358	1.07 (0.98, 1.17)	40,780	0.98 (0.96, 1.01)	49,755	0.99 (0.96, 1.01)
	(87.3%-100.0%, median = 94.7%)		~		~		~		~		~
Female											
Age group (y)	15-49	673	0.94(0.86, 1.03)	2,255	1.03 (0.98, 1.07)	938	0.91 (0.83, 1.01)	24,438	1.01 (0.98, 1.04)	31,928	1.01 (0.98, 1.04)
	50-69	4,594	1.01 (0.95, 1.07)	6,550	1.03 (0.98, 1.08)	4,450	1.04 (0.97, 1.12)	46,755	1.02 (1.00, 1.04)	50,074	1.02 (1.00, 1.04)
	70+	6,031	1.04 (0.99, 1.10)	4,991	1.06 (1.00, 1.13)	5,628	1.09 (1.03, 1.17)	66,936	1.04 (1.02, 1.06)	71,060	1.04 (1.02, 1.06)
Solid fuel use level ^a	Low	767	0.98 (0.89, 1.07)	2,432	1.01 (0.96, 1.05)	987	0.93 (0.84, 1.03)	14,160	1.03 (0.99, 1.06)	16,025	1.02 (0.99, 1.05)
	(0.4%-44.6%, median = 16.7%)										
	Medium	1,882	0.98 (0.90, 1.07)	3,860	1.04 (0.98, 1.11)	1,987	0.99(0.90, 1.09)	27,061	1.02 (0.99, 1.05)	31,453	1.02 (0.99, 1.05)
	(44.6% - 87.3%, median = 72.3%)										
	High	2,617	1.01 (0.94, 1.09)	2,512	0.99 (0.92, 1.06)	2,413	1.00 (0.92, 1.10)	29,963	1.01 (0.98, 1.04)	34,513	1.01 (0.98, 1.04)
	(87.3% - 100.0%, median = 94.7%)										
		001 001		0,000						:	

Note: ICD-10 codes for chronic respiratory are 140-147, 163, 193; for IHD are 120-125, 144, 146, 170, R55, R96; for stroke are 160-169, G45-G46, G81-G83; for nonaccidental causes are A00-R99; and for all-cause mortality are A00-R99, U00-Y99, Estimates were derived from a model-based generalized linear geostatistical model with a Poisson distribution. All models were adjusted for death year, urban/ural residency, sampling unit-level smoking prevalence, subdistric-level from a model-based generalized linear geostatistical model with a Poisson distribution. All models were adjusted for death year, urban/ural residency, sampling unit-level smoking prevalence, subdistric-level from a model-based generalized linear geostatistical model with a Poisson distribution. All models were adjusted for death year, urban/ural residency, sampling unit-level smoking prevalence, subdistric-level from a model-based generalized linear geostatistical model with a Poisson distribution. All models were adjusted for death year, urban/ural residency, sampling unit-level smoking prevalence, subdistric-level from at disease; R, rald vorinsant language groups and included smoothly varying spatial random effects and independent unit-level random effects. Credible intervals (CI) of 95% were reported for the RR. CI, credible interval; IHD, ischemic for the at disease; R, rald vorinsant is a solid cooking fuels were used to classify each deceased individual into high, medium, and low solid fuel use categories. Data range and median values for each category are provided in the parenthese.

130(9) September 2022

Smeden et al. 2020). Any substantial daily variation in ambient $PM_{2.5}$ exposure is not likely captured in our use of 3-y medians (Figure S12) and the coarse spatial resolution of satellite-derived $PM_{2.5}$ data (~11 km) would also lead to smoothing out local peaks and troughs, possibly biasing the mortality–exposure association in either direction (Brenner et al. 1993, van Smeden et al. 2020). Because errors in measurements by satellites may also differ by region, this would further complicate the bias direction for the overall estimates. As with any analysis with ecological variables, there is potential for the combination of confounding and measurement error to cause bias in the results presented here, although the bias will not necessarily cause effects to be underestimated.

Third, separating household solid fuel use from other sources of ambient pollution that contribute to satellite-based PM_{2.5} estimates is difficult (Chowdhury al. 2019). However, the percentage of the population using solid fuel was broadly similar between those who died and the overall population (Table 1), and disease-specific RRs were similar in areas of high and low solid fuel use (Table 2) and in the model assessing confounding effect of secondhand smoking exposure among females (Figure S11). Women have higher levels of exposure to solid fuel use than men (Choudhuri and Desai 2020), so the broad consistency of the RRs for PM_{2.5} exposure by sex also suggest that our RRs are not likely to be greatly confounded by solid fuel use. Fourth, there are inherent uncertainties of verbal autopsy assignment of cause of death, even though the distinction between stroke and IHD deaths is reasonably clear with more misclassification between IHD and respiratory disease deaths (Jha et al. 2019). We addressed this uncertainty by relying on dual physicians' initial agreement on the deceased person's cause of death, which was independent of PM_{2.5} level (Table S2). Nevertheless, there remain residual uncertainties in the cause-of-death assignment even using these strict criteria. Use of this strict criteria may also lead to capturing fewer number of cause-specific deaths. Furthermore, such misclassification of causes is greater at ages 70 y or more, but we observed similar RRs in this age group and in the younger age group for the three causes. Our main results focus on ages 15–69 y, where the number of ill-defined causes of death were low in comparison with older ages, to ensure higher accuracy in the reported estimates. Moreover, the use of all nonaccidental deaths yielded similar results to those for IHD and respiratory disease.

Fifth, our analyses have some limitations which are countered by the fact that the sample was nationally representative and drawn from over 7,400 discrete areas and that our models considered both unit-level and spatially varying unmeasured risk factors. We had fewer individual-level covariates than did the PURE study, but our results were very similar. The PURE study also showed little variation in RRs by adding more individual covariates (Hystad et al. 2020). Use of community-level variables potentially introduces residual confounding in the models. We minimized the risk of residual confounding by adjusting for deprivation (via female illiteracy), smoking prevalence in sampling units, various sociocultural risk factors by including language, and, most important, by the adjustment for unit-level variation and spatial clustering. Moreover, we saw nearly identical results from casecontrol analyses, where any bias affecting assessment of the ambient PM2.5 exposure in the cases should similarly affect assessment of the exposure in control deaths. Our model does not produce hazard ratios typically used in prospective studies. However, our model is equivalent to a survival model, assuming (reasonably) that the hazard function is piecewise constant changing every 5 y of age. Because we are modeling a rare event, with nearly all the death dates being right censored, a survival model with a spatial random effect would be challenging to fit to a data set of this size. Last, we do not report on childhood deaths, including childhood pneumonia, which is the leading cause of childhood mortality in India (MDS Collaborators 2017) and has been linked to $PM_{2.5}$ exposure (WHO 2018). However, the overall age-standardized death rates from pneumonia at ages 0–4 y and 5–14 y have fallen substantially in India from 2000 to 2015, including in the states where $PM_{2.5}$ exposure is most common (MDS Collaborators 2017; Fadel et al. 2019).

Given these uncertainties in our estimates, we did not derive the population attributable risks, because we believe it is premature to do so until direct epidemiological studies can further quantify the relationship between ambient PM2.5 exposure and mortality and establish causality (which model-based estimates presuppose). Naturally, errors in measuring PM2.5 apply also to GEMM/WHO estimates, which also rely on these satellite-based estimates. Indeed, our study emphasizes the need for additional reliable nationwide mortality studies that incorporate ambient PM_{2.5}, smoking, and other relevant risk factors in countries with relatively high PM2.5 exposure. Better ground-based measurement to capture the spatiotemporal variations of PM2.5 exposures (such as daily exposure near each residence) is warranted (Kaufman et al. 2016). Our population model, ideally with individual-level smoking type and other confounders and including examining possible interactions between sex, age group, and solid fuel use, could help to resolve some of the uncertainties in documenting mortality from ambient $PM_{2.5}$ exposure in China (Yin et al. 2017) and elsewhere.

However, action to reduce $PM_{2.5}$ exposure is justified regardless of these uncertainties. There is substantial earlier evidence and biological plausibility that chronic respiratory morbidity and mortality in children from acute respiratory infection can be reasonably attributable to $PM_{2.5}$ exposure (Gehring et al. 2013). Limiting air pollutant output from traffic, industrial, and household sources is a relevant strategy in India to reduce overall $PM_{2.5}$ exposure (Karagulian et al. 2015).

Acknowledgments

P.J. conceived the study. R.G., G.D.S., P.B., and P.J. were the study principal investigators. The MDS Collaborators implemented the data collection procedures. P.B., Y.I., P.R., S.H.F., and P.J. conducted the analysis. P.J., P.B., G.D.S., and S.H.F. wrote the first draft, and all authors interpreted the data, reviewed results, and contributed to the report. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. P.J. is the guarantor. The authors thank P. Hystad and S. Yusuf for providing data from the PURE study and H. Gelband for useful comments.

This study was supported by the Indian Council of Medical Research, Canadian Institutes of Health Research Foundation grant (FDN 154277), the University of Toronto, the National Institutes of Health grants (1R01TW007939).

This paper expresses the views of the authors and working group and does not necessarily represent the views of the Government of India. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing the report.

Data availability: Please refer to https://censusindia.gov.in/ census.website/data/SRSSTAT for public reports. For MDS data access procedures, please contact the Office of the Registrar General, RK Puram, New Delhi, India. The satellitederived ambient PM_{2.5} data from van Donkelaar et al. 2015 can be obtained free of charge from https://sites.wustl.edu/ acag/datasets/surface-pm2-5/. The 2011 Indian census data (including data on female illiteracy, language groups, and religion) can be accessed from https://data.gov.in. The Gridded Population of the World (version 4) can be downloaded from https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-adjusted-to-2015-unwpp-country-totals-rev11.

Code availability: The geostatsp and R-INLA packages for the statistical software R were used for all analyses. A sample R script with synthetic data is available upon request from the first author.

References

- Aleksandrowicz L, Malhotra V, Dikshit R, Gupta PC, Kumar R, Sheth J, et al. 2014. Performance criteria for verbal autopsy-based systems to estimate national causes of death: development and application to the indian million death study. BMC Med 12(1):21, PMID: 24495287, https://doi.org/10.1186/1741-7015-12-21.
- Balakrishnan K, Ghosh S, Ganguli B, Sambandam S, Bruce N, Barnes DF, et al. 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 Health 12(1):77, PMID: 24020494, https://doi.org/10. 1186/1476-069X-12-77.
- Bhatt S, Weiss D, Cameron E, Bisanzio D, Mappin B, Dalrymple U, et al. 2015. The effect of malaria control on plasmodium falciparum in africa between 2000 and 2015. Nature 526 526(7572):207–211, PMID: 26375008, https://doi.org/10.1038/ nature15535.
- Boys BL, Martin RV, Van Donkelaar A, MacDonell RJ, Hsu NC, Cooper MJ, et al. 2014. Fifteen-year global time series of satellite-derived fine particulate matter. Environ Sci Technol 48(19):11109–11118, PMID: 25184953, https://doi.org/10. 1021/es502113p.
- Brauer M, Freedman G, Frostad J, Van Donkelaar A, Martin RV, Dentener F, et al. 2016. Ambient air pollution exposure estimation for the global burden of disease 2013. Environ Sci Technol 50(1):79–88, PMID: 26595236, https://doi.org/10. 1021/acs.est.5b03709.
- Brenner H, Savitz DA, Gefeller O. 1993. The effects of joint misclassification of exposure and disease on epidemiologic measures of association exposure and disease on epidemiologic measures of association. J Clin Epidemiol 46(10):1195– 1202, PMID: 8410104, https://doi.org/10.1016/0895-4356(93)90119-L.
- Brown PE. 2015. Model-based geostatistics the easy way. J Stat Soft 63(12):1–24, https://doi.org/10.18637/jss.v063.i12.
- Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA, et al. 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc Natl Acad Sci USA 115(38):9592–9597, PMID: 30181279, https://doi.org/10.1073/pnas.1803222115.
- Burney P, Amaral AF. 2019. Air pollution and chronic airway disease: is the evidence always clear? Lancet 394(10215):2198–2200, PMID: 31761449, https://doi.org/10. 1016/S0140-6736(19)32537-1.
- Chen Z, Iona A, Parish S, Chen Y, Guo Y, Bragg F, et al. 2018. China Kadoorie Biobank orative group. 2018. Adiposity and risk of ischaemic and haemorrhagic stroke in chinese men and women: a prospective study of 0.5 million adults. Lancet Glob Health 6(6):e630–e640, PMID: 29773119, https://doi.org/10.1016/ S2214-109X(18)30216-X.
- Choudhuri P, Desai S. 2020. Gender inequalities and household fuel choice in India. J Clean Prod 265:121487, PMID: 32831484, https://doi.org/10.1016/j.jclepro.2020.121487.
- Chowdhury S, Dey S, Guttikunda S, Pillarisetti A, Smith KR, Di Girolamo L. 2019. Indian annual ambient air quality standard is achievable by completely mitigating emissions from household sources. Proc Natl Acad Sci USA 116(22):10711– 10716, PMID: 30988190, https://doi.org/10.1073/pnas.1900888116.
- Diggle PJ, Ribeiro PJ. 2007. *Model-Based Geostatistics*. New York: Springer-Verlag.
- Dikshit RP, Gupta PC, Ramasundarahettige CF, Gajalakshmi V, Aleksandrowicz L, Badwe R, et al. 2012. Cancer mortality in India: a nationally representative survey. Lancet 379(9828):1807–1816, PMID: 22460346, https://doi.org/10.1016/S0140-6736(12)60358-4.
- Dormann CF, McPherson JM, Araújo MB, Bivand R, Bolliger J, Carl G, et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30(5):609–628, https://doi.org/10.1111/j.2007. 0906-7590.05171.x.
- Fadel SA, Boschi-Pinto C, Yu S, Reynales-Shigematsu LM, Menon GR, Newcombe L, et al. 2019. Trends in cause-specific mortality among children aged 5–14 years from 2005 to 2016 in India, China, Brazil, and Mexico: an analysis of nationally representative mortality studies. Lancet 393(10176):1119–1127, PMID: 30876707, https://doi.org/10.1016/S0140-6736(19)30220-X.
- Gajalakshmi V, Lacey B, Kanimozhi V, Sherliker P, Peto R, Lewington S. 2018. Body-mass index, blood pressure, and cause-specific mortality in India: a prospective cohort study of 500,810 adults. Lancet Glob Health 6(7):e787–e794, PMID: 29903379, https://doi.org/10.1016/S2214-109X(18)30267-5.
- Gehring U, Gruzieva O, Agius RM, Beelen R, Custovic A, Cyrys J, et al. 2013. Air pollution exposure and lung function in children: the ESCAPE project. Environ

Health Perspect 121(11–12):1357–1364, PMID: 24076757, https://doi.org/10.1289/ehp.1306770.

- Gelfand AE, Banerjee S. 2017. Bayesian modeling and analysis of geostatistical data. Annu Rev Stat Appl 4:245–266, PMID: 29392155, https://doi.org/10.1146/ annurev-statistics-060116-054155.
- Gomes M, Begum R, Sati P, Dikshit R, Gupta PC, Kumar R, et al. 2017. Nationwide mortality studies to quantify causes of death: relevant lessons from India's Million Death Study. Health Aff (Millwood) 36(11):1887–1895, PMID: 29137507, https://doi.org/10.1377/hlthaff.2017.0635.
- Hystad P, Larkin A, Rangarajan S, AlHabib KF, Avezum Á, Calik KBT, et al. 2020. Associations of outdoor fine particulate air pollution and cardiovascular disease in 157 436 individuals from 21 high-income, middle-income, and lowincome countries (PURE): a prospective cohort study. Lancet Planet Health 4(6):e235–e245, PMID: 32559440, https://doi.org/10.1016/S2542-5196(20)30103-0.
- Jha P, Gajalakshmi V, Gupta PC, Kumar R, Mony P, Dhingra N, et al. 2006. Prospective study of one million deaths in India: rationale, design, and validation results. PLoS Med 3(2):e18, PMID: 16354108, https://doi.org/10.1371/journal. pmed.0030018.
- Jha P, Jacob B, Gajalakshmi V, Gupta PC, Dhingra N, Kumar R, et al. 2008. A nationally representative case–control study of smoking and death in India. N Engl J Med 358(11):1137–1147, PMID: 18272886, https://doi.org/10.1056/NEJMsa0707719.
- Jha P, Kumar D, Dikshit R, Budukh A, Begum R, Sati P, et al. 2019. Automated versus physician assignment of cause of death for verbal autopsies: randomized trial of 9374 deaths in 117 villages in India. BMC Med 17(1):1–11, PMID: 31242925, https://doi.org/10.1186/s12916-019-1353-2.
- Jiang H, Brown PE, Rue H, Shimakura S. 2014. Geostatistical survival models for environmental risk assessment with large retrospective cohorts. J Royal Stat Soc Series A 177(3):679–695, https://doi.org/10.1111/rssa.12041.
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, et al. 1996. The NCEP/NCAR 40-year reanalysis project. Bull Amer Meteor Soc 77(3):437–470, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.C0;2.
- Karagulian F, Belis CA, Dora CFC, Prüss-Ustün AM, Bonjour S, Adair-Rohani H, et al. 2015. Contributions to cities' ambient particulate matter (PM): a systematic review of local source contributions at global level. Atmos Environ 120:475–483, https://doi.org/10.1016/j.atmosenv.2015.08.087.
- Kaufman JD, Adar SD, Barr RG, Budoff M, Burke GL, Curl CL, et al. 2016. Association between air pollution and coronary artery calcification within six metropolitan areas in the USA (the multi-ethnic study of atherosclerosis and air pollution): a longitudinal cohort study. Lancet 388(10045):696–704, PMID: 27233746, https://doi.org/10.1016/S0140-6736(16)00378-0.
- Ke C, Gupta R, Shah BR, Stukel TA, Xavier D, Jha P. 2021. Association of hypertension and diabetes with ischemic heart disease and stroke mortality in India: the Million Death Study. Glob Heart 16(1):69, PMID: 34692394, https://doi.org/10.5334/gh.1048.
- Ke C, Gupta R, Xavier D, Prabhakaran D, Mathur P, Kalkonde YV, et al. 2018. Divergent trends in ischaemic heart disease and stroke mortality in India from 2000 to 2015: a nationally representative mortality study. Lancet Glob Health 6(8):e914–e923, PMID: 30012272, https://doi.org/10.1016/s2214-109x(18)30242-0.
- Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, et al. 2009. Extended follow-up and spatial analysis of the American cancer society study linking particulate air pollution and mortality. Boston, MA: Health Effects Institute.
- Lawson AB. 2018. Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. Boca Raton, FL: Chapman and Hall/CRC.
- Li Y, Brown P, Rue H, Al-Maini M, Fortin P. 2012. Spatial modelling of lupus incidence over 40 years with changes in census areas. J R Stat Soc Ser C Appl Stat 61(1):99–115, https://doi.org/10.1111/j.1467-9876.2011.01004.x.
- Lindgren F, Rue H. 2008. On the second-order random walk model for irregular locations. Scandinavian J Stat 35(4):691–700, https://doi.org/10.1111/j.1467-9469. 2008.00610.x.
- MDS (Million Deaths Study) Collaborators. 2017. Changes in cause-specific neonatal and 1–59-month child mortality in India from 2000 to 2015: a nationally representative survey. Lancet 390:1972–1980, PMID: 28939096, https://doi.org/10. 1016/S0140-6736(17)32162-1.
- Menon GR, Singh L, Sharma P, Yadav P, Sharma S, Kalaskar S, et al. 2019. National burden estimates of healthy life lost in India, 2017: an analysis using direct mortality data and indirect disability data. Lancet Glob Health 7(12):e1675–e1684, PMID: 31708148, https://doi.org/10.1016/S2214-109X(19) 30451-6.
- Moraga P. 2019. *Geospatial Health Data: Modeling and Visualization with r-INLA and Shiny*. Boca Raton, FL: Chapman Hall/CRC.
- Papadogeorgou G, Kioumourtzoglou M-A, Braun D, Zanobetti A. 2019. Low levels of air pollution and health: effect estimates, methodological challenges, and future directions. Curr Environ Health Rep 6(3):105–115, PMID: 31090042, https://doi.org/10.1007/s40572-019-00235-7.
- RGI (Registrar General of India), Census Commissioner. 2011a. Rural Urban Distribution of Population (Provisional Population Totals). https://censusindia.gov. in/nada/index.php/catalog/1430 [accessed 25 August 2022].

- RGI, Census Commissioner. 2011b. Vital Statistics Sample Registration. https:// censusindia.gov.in/census.website/data/SRSSTAT [accessed 25 August 2022].
- RGI, CGHR (Centre for Global Health Research). 2015. Causes of Death Statistics 2010–2013. New Delhi, India: Registrar General of India.
- RGI, Sampling Registration System (SRS) Collaborators, Centre for Global Health Research (CGHR). 2011. *Health Care Professionals Manual for Assigning Causes of Death from Verbal Autopsy*. New Delhi, India: Registrar General of India.
- Rue H, Martino S, Chopin N. 2009. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. J R Stat Soc Series B Stat Methodol 71(2):319–392, https://doi.org/10.1111/j.1467-9868.2008.00700.x.
- Salvi S, Kumar GA, Dhaliwal R, Paulson K, Agrawal A, Koul PA, et al. 2018. The burden of chronic respiratory diseases and their heterogeneity across the states of India: the global burden of disease study 1990–2016. Lancet Glob Health 6(12):e1363–e1374, PMID: 30219316, https://doi.org/10.1016/S2214-109X (18)30409-1.
- Schulz H, Harder V, Ibald-Mulli A, Khandoga A, Koenig W, Krombach F, et al. 2005. Cardiovascular effects of fine and ultrafine particles. J Aerosol Med 18(1):1– 22, PMID: 15741770, https://doi.org/10.1089/jam.2005.18.1.
- Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. 2017. Penalising model component complexity: a principled, practical approach to constructing priors. Statist Sci 32(1):1–28, https://doi.org/10.1214/16-STS576.
- van Donkelaar A, Martin RV, Brauer M, Boys BL. 2015. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate

matter. Environ Health Perspect 123(2):135–143, PMID: 25343779, https://doi.org/ 10.1289/ehp.1408646.

- van Smeden M, Lash TL, Groenwold RH. 2020. Reflection on modern methods: five myths about measurement error in epidemiological research. Int J Epidemiol 49(1):338–347, PMID: 31821469, https://doi.org/10.1093/ije/dyz251.
- WHO (World Health Organization). 2004. The International Statistical Classification of Diseases and Health Related Problems, 10th revision. Geneva, Switzerland: World Health Organization. https://apps.who.int/iris/handle/10665/42980 [accessed 20 February 2020].
- WHO. 2006. WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide: Global Update 2005: Summary of Risk Assessment. Geneva, Switzerland: World Health Organization. https://apps.who.int/iris/ handle/10665/69477 [accessed 20 February 2020].
- WHO. 2012. Verbal Autopsy Standards: The 2012 WHO Verbal Autopsy Instrument. https://cdn.who.int/media/docs/default-source/mca-documents/rmncah/verbalautopsy-standards-the-2012-who-verbal-autopsy-instrument—release-candidate-1.pdf?sfvrsn=68f9b5a4_1 [accessed 20 February 2020].
- WHO. 2018. Ambient (AAP) air pollution attributable burden of disease, 2016 (May 2018). Geneva, Switzerland: World Health Organization. https://www.who.int/airpollution/ data/aap_bod_may2018_v0.xlsx?ua=1 [accessed 20 February 2020].
- Yin P, Brauer M, Cohen A, Burnett RT, Liu J, Liu Y, et al. 2017. Long-term fine particulate matter exposure and nonaccidental and cause-specific mortality in a large national cohort of chinese men. Environ Health Perspect 125(11):117002, PMID: 29116930, https://doi.org/10.1289/EHP1673.