

# Child Survival and Early Lifetime Exposures to Ambient Fine Particulate Matter in India: A Retrospective Cohort Study

Jiawen Liao,<sup>1,2</sup> Yang Liu,<sup>1</sup> Kyle Steenland,<sup>1</sup> Ajay Pillarisetti,<sup>1,3</sup> Lisa M. Thompson,<sup>4</sup> Sagnik Dey,<sup>5,6</sup> Kalpana Balakrishnan,<sup>7</sup> and Thomas Clasen<sup>1</sup>

<sup>1</sup>Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

<sup>2</sup>Department of Population and Public Health Sciences, Keck School of Medicine of the University of Southern California, Los Angeles, California, USA

<sup>3</sup>Environmental Health Sciences, School of Public Health, University of California, Berkeley, California, USA

<sup>4</sup>Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, Georgia, USA

<sup>5</sup>Centre for Atmospheric Sciences, Indian Institute of Technology Delhi (IIT Delhi), New Delhi, India

<sup>6</sup>Centre of Excellence for Research on Clean Air, IIT Delhi, New Delhi, India

<sup>7</sup>Department of Environmental Health Engineering, ICMR Center for Advanced Research on Air Quality, Climate and Health, Sri Ramachandra Institute for Higher Education and Research (Deemed University), Chennai, Tamil Nadu, India

**BACKGROUND:** Ambient fine particulate matter [PM  $\leq 2.5$   $\mu\text{m}$  in aerodynamic diameter (PM<sub>2.5</sub>)] is a major health risk for children, particularly in South Asia, which currently experiences the highest PM<sub>2.5</sub> levels globally. Nevertheless, there is comparatively little epidemiological evidence from this region to quantify the effects of PM<sub>2.5</sub> on child survival.

**OBJECTIVES:** We estimated the association between PM<sub>2.5</sub> exposure and child survival in India.

**METHODS:** We constructed a large, retrospective, and nationally representative cohort of children <5 years of age, born between 2009–2016, from the publicly available, cross-sectional 2015–2016 Demographic Health Surveys in India. *In utero* and post-delivery lifetime average ambient PM<sub>2.5</sub> exposures were estimated with data from satellite remote sensing, meteorology, and land use information (model  $R^2 = 0.82$ ). We used Cox proportional hazards regression to estimate the association between both average *in utero* and post-delivery lifetime PM<sub>2.5</sub> and all-cause child mortality, controlling for individual- and household-level covariates, seasonality, location, and meteorology.

**RESULTS:** Over 7,447,724 child-months of follow-up, there were 11,559 deaths at <5 years of age reported by the children's mothers. The mean concentrations of 9-month *in utero* and post-delivery lifetime average ambient PM<sub>2.5</sub> exposure were 71.1  $\mu\text{g}/\text{m}^3$  (range: 20.9–153.5  $\mu\text{g}/\text{m}^3$ ) and 73.7  $\mu\text{g}/\text{m}^3$  (range: 14.0–247.3  $\mu\text{g}/\text{m}^3$ ), respectively. Estimated child mortality adjusted hazard ratios were 1.023 [95% confidence interval (CI): 1.008, 1.038] and 1.013 (95% CI: 1.001, 1.026) per 10- $\mu\text{g}/\text{m}^3$  increase of *in utero* and post-delivery lifetime PM<sub>2.5</sub>, with both exposures in the model.

**DISCUSSION:** This study adds to the growing body of evidence about the adverse health effects of PM<sub>2.5</sub> by demonstrating the association between exposure, both *in utero* and post-delivery, on child survival at the national level in India. Strategies to reduce ambient air pollution levels, including steps to minimize *in utero* and early life exposures, are urgently needed in India and other countries where exposures are above recommended guideline values. <https://doi.org/10.1289/EHP8910>

## Introduction

Fine particulate matter [airborne particles with an aerodynamic diameter of  $\leq 2.5$   $\mu\text{m}$  (PM<sub>2.5</sub>)] has been shown to be associated with adverse health effects in children, including preterm birth (Balakrishnan et al. 2018; Huynh et al. 2006; Li et al. 2018; Stieb et al. 2012), low birth weight (LBW) (Bell et al. 2007; Pedersen et al. 2013), low height-for-age z-score (Spears et al. 2019), respiratory infections (Gurley et al. 2013; Smith et al. 2011), and mortality (Heft-Neal et al. 2018; Son et al. 2017; Woodruff et al. 2006). In India, ambient PM<sub>2.5</sub> levels are among the highest in the world, with an estimated 91.7  $\mu\text{g}/\text{m}^3$  population-weighted annual average (India State-Level Disease Burden Initiative Air Pollution Collaborators 2021) in 2019. The Global Burden of Disease (GBD) Project identified ambient PM<sub>2.5</sub> as the second largest risk

factor for disease burden in India (GBD 2019 Diseases and Injuries Collaborators 2020) and estimated that >980,000 premature deaths every year are associated with ambient PM<sub>2.5</sub> air pollution (India State-Level Disease Burden Initiative Air Pollution Collaborators 2021), >49,000 of which are premature deaths in children at <5 years of age (India State-Level Disease Burden Initiative Air Pollution Collaborators 2019, 2021). To date, however, most evidence on child mortality due to exposure to ambient air pollution during pregnancy and early childhood was from regions with lower ambient PM<sub>2.5</sub> levels (Glinianaia et al. 2004; Hans et al. 2011; Son et al. 2017; Woodruff et al. 1997, 2006). Although prospective pregnant mother–child cohort studies have recently been initiated in India to examine the association between air pollution exposure and indicators of child health, such as birth-weight (Balakrishnan et al. 2015, 2018; Clasen et al. 2020), to our knowledge, estimates of child mortality have not been reported. Furthermore, owing to the low density of ambient air pollution monitoring stations in India and other low- and middle-income countries (LMICs) (Brauer et al. 2019; Landrigan et al. 2018), current studies of air pollution and infant mortality in India and LMICs are mainly based on annual average ambient PM<sub>2.5</sub> prediction models (Shaddick et al. 2018; van Donkelaar et al. 2016). Such exposure estimates may be prone to error or misclassification, especially for outcomes such as child mortality, which has been associated with monthly or daily changes in air pollution levels (Hans et al. 2011; Son et al. 2017). The recent expansion in the number of ground PM<sub>2.5</sub> monitoring stations in India and the development of spatial PM<sub>2.5</sub> prediction models enable more spatiotemporally resolved estimation of ambient PM<sub>2.5</sub> exposure throughout India.

The primary objective of our study was to evaluate the association between child mortality and ambient PM<sub>2.5</sub> exposure during *in utero* and post-delivery lifetime periods among a nationally

---

Address correspondence to Jiawen Liao, Department of Population and Public Health Sciences, Keck School of Medicine of the USC, SSB #209-17, 2001 N. Soto Street Health Sciences Campus, Los Angeles, CA 90032, USA. Email: [liaojiaw@usc.edu](mailto:liaojiaw@usc.edu), or Yang Liu, Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, 1518 Clifton Rd. NE, Atlanta, GA 30322 USA. Email: [yang.liu@emory.edu](mailto:yang.liu@emory.edu)

Supplemental Material is available online (<https://doi.org/10.1289/EHP8910>). All authors declare they have no actual or potential competing financial interest.

Received 1 January 2021; Revised 17 November 2021; Accepted 20 December 2021; Published 26 January 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 [ehponline@niehs.nih.gov](mailto:ehponline@niehs.nih.gov). Our staff will work with you to assess and meet your accessibility needs within 3 working days.

representative, retrospective cohort based on maternal-reported data in the Demographic and Health Surveys [DHS, also known as the National Family Health Survey 2015–2016 (NFHS-4)] in India. To estimate time-dependent ambient PM<sub>2.5</sub> exposure for the cohort, we developed a monthly ambient PM<sub>2.5</sub> model with 0.1° resolution between 2009–2018 over the entire Indian subcontinent. Our ground station–validated model was developed using machine learning, with multiple inputs from satellite remote sensing, meteorological data, and land use information.

## Methods

### *Study Design and Population*

We conducted a retrospective cohort study of a nationally representative sample of children 0–59 months of age born throughout India (as of 2019, 29 states and 5 Union Territories, excluding the Andaman and Nicobar Islands and Lakshadweep Union Territories), whose mothers participated in the 2015–2016 DHS/NFHS-4 between January 2015 and November 2016. The design of the DHS/NFHS-4 has been described elsewhere (IIPS and ICF 2017). Briefly, DHS/NFHS-4 is cross-sectional survey using a two-stage stratified sampling framework to collect information from a nationally representative sample of households with women 15–49 years of age and all of their children. Each respondent answered survey questions separately for all children no matter whether the child was still alive at the time of interview. We included children born to women no more than 5 y preceding the survey. Three tiers of all-cause mortality—including neonatal mortality (death at <1 month of age), infant mortality (death at <12 months of age), and child mortality (death at <60 months of age)—were outcome variables in our study. Child birth year and month, child death year and month (if the child died before interview), and potential confounders were reported by mothers during the DHS/NFHS-4 interview during the 2015–2016 period. The child-, maternal-, and household-level covariates were also reported by mothers at the end of the follow-up period. We calculated follow-up time in months, starting from child birth until child death as the failure event and the month of the DHS/NFHS-4 survey or passing the at-risk age as the censoring event.

Because this study analyzed only publicly available data without personal identifiers, the institutional review board (IRB) of Emory University determined this study not to be human subject research and did not require IRB review. Data analyses were conducted between March and September 2019.

### *Ambient PM<sub>2.5</sub> Exposure and Meteorological Variables*

In this study, we developed a random forest (RF) model to predict monthly ambient PM<sub>2.5</sub> concentrations at a 0.1° spatial resolution (~ 11 km at the equator) over the entire Indian subcontinent, from 2009 to 2018 (10 y). This approach has been shown to have reliable historical prediction capabilities, to allow a variety of input variables, and to have good agreement with ground measurements from previous studies for modeling ambient PM<sub>2.5</sub> concentrations in East Asia and Latin America (Vu et al. 2019; Xiao et al. 2018). Other advantages of RFs are that these models can work with correlated predictors, can provide variable importance vectors to compare the influence of variables, and are not likely to overfit (Xiao et al. 2018).

First, we collected daily ambient PM<sub>2.5</sub> concentration between 2015 and 2018 from air monitoring stations across India from the Central Pollution Control Board (<https://app.epcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing>) and hourly data at the U.S. Embassy and Consulates in India from AirNow (<https://airnow.gov/>).

We excluded daily PM<sub>2.5</sub> concentrations that were outside of 3 standard deviations of the mean for the log-transformed data (0.05%; exclusions were values outside of 4.2–845 µg/m<sup>3</sup>) because we believed these measurements may not be plausible. For hourly PM<sub>2.5</sub> concentration data from U.S. embassies or consulates, we removed data that lacked valid quality checks and daily average concentrations with at <18 hourly measurements (2.2%).

Daily average PM<sub>2.5</sub> measurements from stations within the same 0.1° × 0.1° grid were averaged, resulting in 56,834 grid-day PM<sub>2.5</sub> measurements (Table S1). We averaged gridded daily PM<sub>2.5</sub> measurements into gridded monthly PM<sub>2.5</sub> averages if more than 15 d had valid daily measurements for a given grid cell in a given month. We ended up with 1,446 grid-months of PM<sub>2.5</sub> measurements during 2017–2018 from 134 ground air pollution monitoring stations for model development and 465 grid-months of PM<sub>2.5</sub> measurements during 2015–2016 from 57 stations for model hindcasting evaluation.

We used multiple satellite-retrieved aerosol, nitrogen dioxide (NO<sub>2</sub>), fire spot data, and vegetation index products; global meteorological and aerosol reanalysis models; land use information; and population density data for the period between 2009 and 2018 as inputs for our model. Briefly, we used Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 level-2 aerosol optical depth (AOD) products (MOD04\_L2 and MYD04\_L2) at a 10-km resolution collected by the Aqua and Terra satellites, from the Distributed Active Archive Center (<https://ladsweb.modaps.eosdis.nasa.gov/>) (Levy and Hsu 2015). We used Deep Blue (DB) and Deep Blue/Dark Target combined parameters (Combined) from MODIS retrievals and assigned and averaged the centroid of each retrieval to create 0.1° × 0.1° grid cells and calculate monthly average AOD value for DB and Combined algorithms. We then conducted gap-filling of the missing AOD values for the period similar to previous study (Xiao et al. 2018). Table S2 summarizes the AOD data missingness patterns, and Figure S1 shows the Combined AOD parameter and DB AOD parameter spatial distribution after gap-filling. The aerosol absorbing index in visible and ultraviolet light and tropospheric NO<sub>2</sub> density data were obtained from the Ozone Monitoring Instrument, downloaded from the Goddard Earth Sciences Data and Information Service Center (<https://mirador.gsfc.nasa.gov/>). We also collected active fire data obtained from the Information for Resource Management System (<https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms>) and the Normalized Difference Vegetation Index from the MODIS Vegetation Indices in the Distributed Active Archive Center (<https://ladsweb.modaps.eosdis.nasa.gov/>).

Daily meteorological data and aerosol diagnostics were obtained from the Goddard Earth Observing System Data Assimilation System 5/Modern-Era Retrospective analysis for Research and Applications (MERRA-2; <https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/>) (Bosilovich et al. 2015). Elevation data was collected from the National Aeronautics and Space Administration's Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (version 2). Highway, primary road, and other road density parameters were obtained from the Global Roads Inventory Project (Meijer et al. 2018). Yearly varying population numbers for 2009–2018 were obtained from the LandScan Global Population Database. All of the above variables were assigned into 0.1° × 0.1° grid cells, and their monthly averages were used as predictors in the PM<sub>2.5</sub> model. Details of these variables are listed in Table S3.

We trained the RF model on gridded monthly mean PM<sub>2.5</sub> concentrations from the 2017–2018 period ( $N = 1,446$ ) and evaluated model performance using standard, spatial, and temporal 10-fold cross-validation (CV). Spatial CV refers to using data from 90% of grids for developing the model and then testing the model on the

remaining 10% of grids; temporal CV relies on using 90% of monthly data to develop the model and then testing the model based on the remaining 10% of monthly data. We also evaluated the hindcasting performance of the model using monthly data from the 2015–2016 period ( $N = 456$ ). Root-mean-squared error (RMSE) and  $R^2$  values from 10-fold cross-validation and from hindcasting were reported to evaluate the historical prediction performance of the model. We then estimated population-weighted annual mean  $PM_{2.5}$  concentrations over India during 2009–2018, evaluated the change of  $PM_{2.5}$  levels over those 10 y, and compared our estimates with other published databases (India State-Level Disease Burden Initiative Air Pollution Collaborators 2019; Shaddick et al. 2018; van Donkelaar et al. 2016). Predicted ambient monthly  $PM_{2.5}$  levels were matched to each child with geo-coordinates at the cluster level [a primary sampling unit (PSU) or a segment of a PSU, of around 100–150 households] provided by DHS/NFHS-4. For each child, we assigned the 9-month average  $PM_{2.5}$  concentrations prior to birth as the *in utero*  $PM_{2.5}$  exposure. We treated post-delivery average  $PM_{2.5}$  exposures as time dependent and defined them as the average  $PM_{2.5}$  concentration from the calendar month of child birth through each month the child was at risk, until the month of death or censoring.

Meteorological variables including monthly air temperatures at 2 m and monthly rainfall precipitation were included as well, and they were also matched with the geo-coordinates of households (in the same way as performed for  $PM_{2.5}$  exposure). Monthly air temperatures at 2 m values were extracted from MERRA-2 data (Bosilovich et al. 2015), and monthly rainfall precipitation in mm/month were extracted from Climate Hazards Group InfraRed Precipitation with Station data (Funk et al. 2015). R (version 3.5; R Development Core Team) with the *randomForestSRC* (version 4.6) package was used for  $PM_{2.5}$  modeling, prediction, and assignment of  $PM_{2.5}$  and meteorological variables for each child.

### Statistical Analysis

We used Cox proportional hazards regression models with follow-up in calendar months as the time variable and child death as the outcome. We included 9-month *in utero* and time-dependent, post-delivery average ambient  $PM_{2.5}$  exposure to estimate adjusted hazard ratios (HRs) relating child mortality risk with  $10\text{-}\mu\text{g}/\text{m}^3$  increases of  $PM_{2.5}$  exposure. In this model, the estimated effects of  $PM_{2.5}$  exposure during *in utero* and post-delivery periods were adjusted for each other. We also used single-exposure models with  $PM_{2.5}$  exposures from either the *in utero* or post-delivery periods for comparison. We used a set of confounders from the previous literature and a directed acyclic graph (Figure S2) to identify covariates to be included in the Cox regression model (Heft-Neal et al. 2018; Spears et al. 2019; Subramanian et al. 2009). The prespecified time-independent covariates included child sex, birth month and year, birth order, and location of birth (institutional birth or not); multibirths; and mother's age at child birth, height, marital status, education (whether above secondary level), maternal smoking (yes/no), secondhand smoke exposure in the home (yes/no), wealth index, urban or rural location of households, geographical region/zonal council (Northern, Central, North Eastern, Eastern, Western, and Southern, based on <https://www.mha.gov.in/zonal-council>), primary cooking fuel, and toilet facilities. The household wealth index was calculated and provided by DHS/NFHS-4 using principle component analysis of household ownership of assets (Staveteig and Mallick 2014). We dichotomized toilet facilities as whether they were safely managed based on the World Health Organization (WHO) and United Nations Children's Fund Joint Monitoring Program (Croft et al. 2018; WHO/UNICEF JMP for Water Supply, Sanitation and Hygiene 2018), and dichotomized cooking fuel as clean (electricity, liquefied petroleum gas, natural

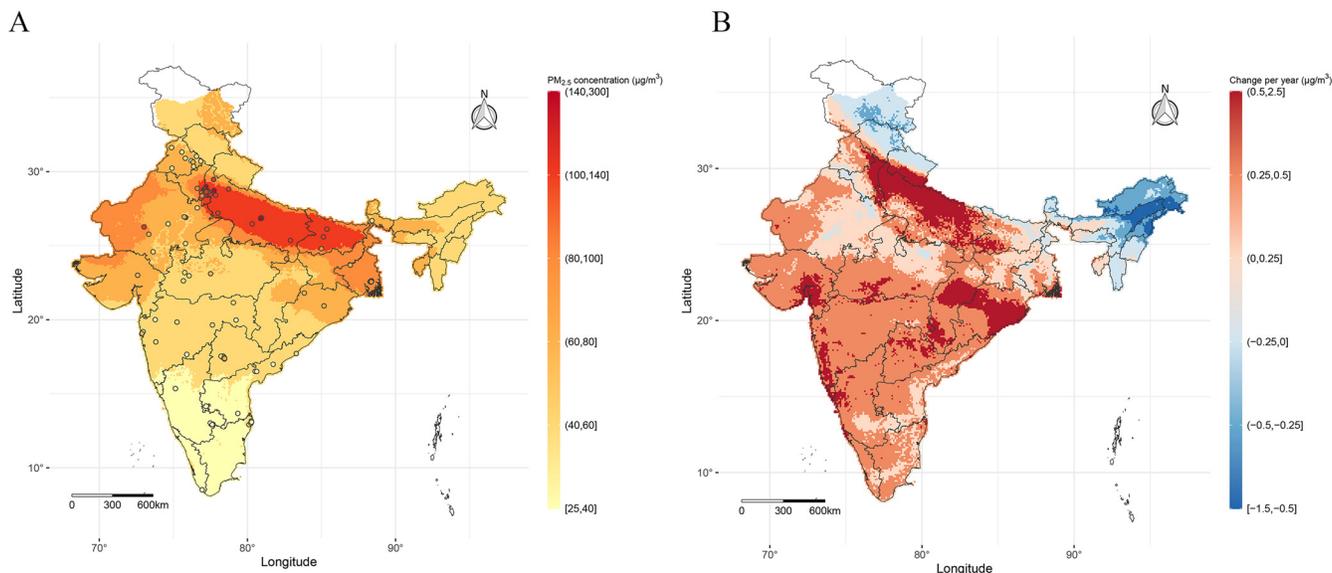
gas, and biogas) or dirty (kerosene, coal, wood, straw, agriculture crop residue, and dung). Time-dependent covariates included monthly rainfall precipitation and monthly temperature. Child birth year and month, sex, and the geographical region of households were modeled as strata. Unique baseline hazard functions were given for all 927 strata of combinations of these strata. Robust variance estimation for coefficients was used. For tied outcomes, Breslow's method was used. Table S4 provides the details on the variables used in the Cox regression models. Although we did not include children born with LBW because it has been found to be a potential intermediate variable between air pollution and childhood death (Hernández-Díaz et al. 2006), we conducted sensitivity analysis adjusting for LBW in our study population, as described below. We also tested the association between LBW and *in utero*  $PM_{2.5}$  exposure, and between LBW and child mortality. We stratified the Cox regression by child birth month and year, child sex, and geographical regions (927 strata) to account for spatiotemporal and sex differences in child survival. In addition, we included the sample probability weight from DHS/NFHS-4 in the Cox regression (Equation 1).

$$h_g(t, Z, X(t)) = h_{0g}(t) \exp[\beta^T Z + \delta^T X(t)]. \quad (1)$$

$h_g(t)$  is the hazard function (probability of death) at month  $t$ , and  $g$  refers to the (1–927) strata of child birth month, sex, and geographical location.  $Z$  is the vector containing time-independent variables, including *in utero*  $PM_{2.5}$  exposure, and  $X(t)$  are the time-dependent variables including post-delivery  $PM_{2.5}$  exposure. The association between  $PM_{2.5}$  and child mortality was reported as the HR per  $10\text{-}\mu\text{g}/\text{m}^3$  increase of  $PM_{2.5}$ , and we also calculated the HR per interquartile range (IQR) increase of  $PM_{2.5}$  ( $40.6\text{ }\mu\text{g}/\text{m}^3$ ) to compare other studies. To evaluate whether effects of  $PM_{2.5}$  differed across covariates, we conducted subgroup stratified analysis of the two-exposure models. The proportional hazards assumption for *in utero* and post-delivery average  $PM_{2.5}$  exposure was tested on the basis of Schoenfeld residuals on time, and we found it did not violate the proportional hazards assumption ( $p > 0.05$ ). The correlation matrix of the predictor estimation from the Cox regression model was calculated to check multicollinearity, and the Akaike information criterion (AIC) was calculated to compare single- and two-exposure models. We conducted subgroup analysis by conducting separate models by stratifying variables including sex, LBW, delivery mode, whether multibirth, maternal age, maternal height, education, maternal smoking, secondhand smoking, household wealth index, cook fuel type, toilet type, urbanity, marital status, and geographical location, and we tested the hypothesis that there were no interaction effects between *in utero* and post-delivery  $PM_{2.5}$  effects with each of the stratifying variables.

To examine the robustness of our results, we conducted additional sensitivity analyses. We added LBW back to the model (modify 1); conducted analysis using stratified variables as time-independent covariates (modify 2); adding additional cluster-level covariates, including the percentage of household with improved toilets and clean cooking fuels (modify 3); performing unweighted analysis without sample weight (modify 4); using generalized estimating equations to account for the additional correlation within household (modify 5) and cluster level (modify 6); replacing post-delivery lifetime  $PM_{2.5}$  exposure with monthly exposure (modify 7); and performing crude analysis without adjusting covariates (modify 8). In addition, we also conducted an analysis treating *in utero* and post-delivery  $PM_{2.5}$  as categorical variables using quartiles of exposure in the model and tested the linear trend of the categorized exposure.

Last, we calculated the population attributable fraction (PAF) as the burden of child mortality from ambient  $PM_{2.5}$  exposure.



**Figure 1.** Spatial patterns of annual  $PM_{2.5}$  concentrations. (A) Model-predicted annual  $PM_{2.5}$  concentrations with measurements from 133 ground-based monitoring stations in subcontinental India in 2018. (B) Annualized change rate of  $PM_{2.5}$  concentrations between 2009–2018. Circles represent ground station with an annual mean measurement of  $PM_{2.5}$ . Note:  $PM_{2.5}$ , particulate matter  $\leq 2.5$   $\mu m$  in aerodynamic diameter (fine particulate matter).

We estimated the PAF from the modeled *in utero*  $PM_{2.5}$  exposure distribution by every 10th percentile, and the risk of mortality (derived from the single-exposure model only) at a given exposure level compared baseline exposure levels. Stata (version 14; StataCorp) and R (version 3.6; R Development Core Team) were used in the statistical analyses.

## Results

The RF  $PM_{2.5}$  prediction model characterized monthly  $PM_{2.5}$  levels well, with CV  $R^2 = 0.82$  at the monthly level (RMSE of  $25 \mu g/m^3$ ) in 2017–2018. The spatial and temporal CV  $R^2$  (RMSE) of the RF model was 0.77 ( $26 \mu g/m^3$ ) and 0.77 ( $25 \mu g/m^3$ ) respectively, indicating a stable and good fit of monthly ambient  $PM_{2.5}$  measurements from ground stations. Compared with historical measurements in 2015–2016, our model provided good predictions of monthly  $PM_{2.5}$  with  $R^2$  (RMSE) of 0.82 ( $32 \mu g/m^3$ ). Figure S4 shows the performance of the CV and hindcasting of the RF model. Figure S3 shows changes in model performance across different numbers of decision trees and the variable importance matrix for the model. The historical prediction shows that our model underestimated  $PM_{2.5}$  levels during periods with extreme levels, when monthly  $PM_{2.5}$  levels were  $>350 \mu g/m^3$ . Regardless, our model still had relatively good hindcasting capacity to predict historical monthly  $PM_{2.5}$  levels. The annual within-sample RMSE from the model CV in 2017–2018 was  $10.9 \mu g/m^3$ , and the annual RMSE from historical hindcasting in 2015–2016 was  $21.7 \mu g/m^3$ , with better performance compared with previous ambient air pollution models used in the GBD study in the South Asia Region (Shaddick et al. 2018), which had within-sample RMSE of  $17.6 \mu g/m^3$ .

Figure 1 shows the predicted annual  $PM_{2.5}$  in India in 2018 and the annualized change of  $PM_{2.5}$  over a 10-y period. The predicted annual population-weighted mean  $PM_{2.5}$  in India in 2018 was  $71.7 \mu g/m^3$ . The highest annual  $PM_{2.5}$  concentrations ( $>120 \mu g/m^3$ ) were in the Indo-Gangetic Plain Region, covering the states of Haryana, Uttar Pradesh, and Bihar and the National Capital Territory of Delhi. These values are three times  $>40 \mu g/m^3$ , the recommended limit set by the National Ambient Air Quality Standards of India (Gautam 2009). We observed

increased ambient  $PM_{2.5}$  concentrations over a 10-y period in most areas, with the strongest increase of ambient  $PM_{2.5}$  concentrations at  $\sim 1 \mu g/m^3$  per year in the Indo-Gangetic Plain Region and along the west coast of the Western Region, whereas the Northeastern Region and the north part of the Northern Region showed decreases in ambient  $PM_{2.5}$  levels. Details of predicted  $PM_{2.5}$  in India (Figures S5 and S6), major metropolitan areas (Delhi, Mumbai, Kolkata, Chennai, and Bangalore) (Table S5), and population-weighted  $PM_{2.5}$  exposures from our model vs. other global models are listed in Table S6. The predicted ambient  $PM_{2.5}$  air pollution can be downloaded at <https://doi.org/10.15139/S3/RFILYH>.

The present study included data from 259,627 live-birth children born 5 y preceding the survey from 175,865 women in the DHS/NFHS-4 data. We excluded 6,839 (2.6%) children in the analysis due to a) living in the Andaman and Nicobar Islands and Lakshadweep Union Territories, which are islands more than 200 km from the India subcontinent ( $n = 2,282$ , 0.9%); b) not living with their mothers ( $n = 1,912$ , 0.7%); or c) missing variables ( $n = 2,645$ , 1.1%). A total of 252,788 children born between January 2010 and November 2016 were included from 27,853 clusters (Table 1). In the follow-up period of 7,447,724 child-months, 11,559 deaths were reported in children  $<5$  years of age by their mothers. Among these deaths, 7,520 deaths occurred before the first month of life, and 10,862 deaths occurred during the first year of life (including 7,520 neonatal deaths). For the entire cohort, the mean (SD) of 9-month *in utero* and post-delivery lifetime average  $PM_{2.5}$  exposure until failure or censoring was 71.1 (28.2) and 73.3 (29.8)  $\mu g/m^3$ , respectively. The characteristics of the cohort are shown in Table 1, stratified by 9-month *in utero* average  $PM_{2.5}$  exposure, at levels  $<49.7 \mu g/m^3$ , 49.7 to  $<62.4 \mu g/m^3$ , 62.4 to  $<90.3 \mu g/m^3$ , and  $\geq 90.3 \mu g/m^3$ . In general, compared with children in the lowest quartile of exposure were more likely to reside in rural areas (80.1% vs. 70.8%), were born to mothers with shorter stature (47.4% vs. 34.2% of mothers were  $<149.9$  cm in height), lower educational levels (44.3% vs. 19.4% were below the secondary education level), and lived in households with a lower wealth index (35.6% vs. 15.6% were in the lowest quartile of the wealth index). Children residing in central (Chhattisgarh,

**Table 1.** Unweighted characteristics of children <5 years of age by four quartiles of 9-month *in utero* PM<sub>2.5</sub> exposure who were born between 2010–2016 in the Demographic and Health Surveys [DHS, also known as the National Family Health Survey 2015–2016 (NFHS-4)] in India.

Characteristics	Average 9-month <i>in utero</i> PM <sub>2.5</sub> exposure quartile (µg/m <sup>3</sup> ) (Q)				
	Entire cohort	Q1 (20.9–<49.7)	Q2 (49.7–<62.4)	Q3 (62.4–<90.3)	Q4 (90.3–153.5)
Average 9-month <i>in utero</i> PM <sub>2.5</sub> exposure (µg/m <sup>3</sup> )					
Children (n)	252,788	63,197	63,197	63,197	63,197
Total follow-up in month (n)	7,447,724	1,860,527	1,891,613	1,873,198	1,822,386
Child death at <5 years of age (n)	11,559	2,187	2,688	3,001	3,683
All-cause child mortality rate (n/1,000 child-years)	18.62	14.11	17.05	19.22	24.25
Average 9-month <i>in utero</i> PM <sub>2.5</sub> {µg/m <sup>3</sup> [mean (SD)]}	71.1 (28.2)	41.4 (6.0)	55.4 (3.6)	75.0 (8.1)	112.6 (13.9)
Average post-delivery lifetime PM <sub>2.5</sub> exposure until death or censoring {µg/m <sup>3</sup> [mean (SD)]}	73.7 (29.8)	46.6 (10.6)	57.1 (11.5)	83.6 (25.6)	107.6 (20.1)
Child covariates <sup>a</sup>					
Sex (%)					
Female	47.9	48.4	47.8	47.6	47.3
Male	52.1	51.6	52.1	52.4	52.6
LBW (%) <sup>b</sup>	13.1	14.2	13.4	13.9	11.2
Birth order (%)					
First	36.9	40.8	38.0	37.1	31.5
Second	30.7	34.2	31.1	30.0	27.5
Third	16.1	14.2	15.7	16.4	18.1
≥Fourth	16.3	10.8	15.2	16.5	22.9
Multiple birth (%)	1.65	1.55	1.53	1.81	1.71
Birth at institution (%)	75.3	81.8	75.7	74.8	68.7
Maternal covariates <sup>a</sup>					
Age at child birth [y (%)]					
<19	12.2	12.6	11.3	12.9	11.8
20–24	43.4	43.4	42.2	44.5	43.4
25–29	28.4	28.6	29.2	27.8	28.3
≥30	16.0	15.4	17.5	14.8	16.5
Mothers height [cm (%)]					
<149.9	38.3	34.2	34.0	37.5	47.4
150–154.9	33.8	35.3	34.6	33.3	32.2
155–159.9	19.7	21.5	21.9	20.4	15.0
>160	8.2	9.0	9.5	8.8	5.4
Marital status (%)					
Not married	1.4	1.9	1.8	1.1	0.9
Married	98.6	98.1	98.2	98.9	99.1
Education (%)					
Below secondary level	31.5	19.4	27.3	34.8	44.3
Above secondary level	68.5	81.6	72.7	66.2	55.7
Current smoker (%)	10.0	13.4	14.0	7.5	5.2
Secondhand smoke exposure (%)	56.6	51.8	60.5	56.5	58.3
Household covariates <sup>a</sup>					
Cook with clean fuel (%)	27.8	38.5	28.4	23.1	21.1
Access to improved toilet (%)	41.5	48.0	44.7	39.3	33.9
Urban/Rural (%)					
Urban area	23.6	29.2	23.3	22.1	19.9
Rural area	76.4	70.8	76.7	77.9	80.1
Zone area (%)					
Northern	16.6	3.5	25.7	29.4	7.9
North Eastern	13.8	20.9	23.9	9.8	0.5
Central	31.8	18.6	31.2	27.2	50.3
Eastern	21.1	5.0	9.6	28.5	41.3
Western	7.1	14.5	8.7	5.1	0
Southern	9.6	37.5	0.9	0	0
Wealth quartiles (%) <sup>c</sup>					
First (lowest)	25	15.6	21.7	27.1	35.6
Second	25	23.4	26.3	24.5	25.9
Third	25	31.3	26.4	22.4	19.8
Forth	25	29.7	25.6	26.0	18.7
Meteorological covariates <sup>d</sup>					
Temperature {°C [mean (SD)]}	27.1 (6.4)	26.7 (4.9)	24.7 (7.8)	27.6 (6.0)	29.4 (5.5)
Total rainfall {cm [mean (SD)]}	103 (144)	115 (144)	98 (134)	90 (145)	110 (151)

Note: LBW, low birth weight; PM<sub>2.5</sub>, particulate matter ≤2.5 µm in aerodynamic diameter (fine particulate matter); SD, standard deviation.

<sup>a</sup>Child, maternal and household-level covariates collected at the end of the follow-up during the DHS/NFHS-4 interview between 2015–2016.

<sup>b</sup>Assuming children with missing LBW values are not LBW infants.

<sup>c</sup>The household wealth index developed by the DHS/NFHS-4 was calculated using principle component analysis of household ownership of assets.

<sup>d</sup>Meteorological covariates were averaged post-delivery monthly estimates.

Madhya Pradesh, Uttarakhand, and Uttar Pradesh) and eastern (Bihar, Jharkhand, Odisha, and West Bengal) India states had higher estimated ambient PM<sub>2.5</sub> levels compared with other states. Most children included in the analysis were born

between 2012 and 2014, with the least number of children born in 2016 (Table 2). The correlation between average *in utero* and post-delivery lifetime PM<sub>2.5</sub> exposure for the same child was 0.74.

**Table 2.** Number of children and clusters and their estimated annual ambient *in utero* and post-delivery lifetime PM<sub>2.5</sub> exposure by child birth year for children in the Demographic and Health Survey [DHS, also known as the National Family Health Survey 2015–2016 (NFHS-4)] in India.

	2010	2011	2012	2013	2014	2015	2016	All years
Children/cluster (n)								
Children	15,206	40,727	51,960	50,769	51,668	33,951	8,507	252,788
Cluster <sup>a</sup>	8,445	19,766	22,548	22,353	22,419	17,217	5,590	27,835
<i>In utero</i> PM <sub>2.5</sub> exposure (µg/m <sup>3</sup> ) <sup>b</sup>								
Mean (SD)	72.5 (31.7)	68.3 (27.0)	72.3 (28.4)	72.1 (29.1)	71.2 (28.0)	70.8 (26.3)	72.0 (27.8)	71.1 (28.2)
Min	21.2	20.9	22.6	22.4	23.1	24.7	23.7	20.9
Q1	45.7	47.9	51.1	49.9	49.7	50.6	50.2	49.7
Median	61.3	60.1	63.6	63.2	62.4	62.5	62.0	62.4
Q3	96.2	86.9	91.8	91.1	90.4	88.2	95.6	90.3
Max	145.1	142.1	153.5	153.0	144.1	144.2	141.4	153.5
Post-delivery lifetime PM <sub>2.5</sub> exposure (µg/m <sup>3</sup> ) <sup>c</sup>								
Mean (SD)	73.7 (30.7)	73.0 (28.9)	72.7 (28.3)	73.1 (28.8)	76.0 (31.7)	75.2 (31.3)	67.9 (28.5)	73.7 (29.8)
Min	15.7	14.4	14.0	14.7	15.6	15.1	17.4	14.0
Q1	48.2	50.5	50.5	50.5	51.0	52.6	48.6	50.6
Median	61.5	63.0	62.8	63.3	65.1	67.2	61.6	64.0
Q3	106.4	101.5	98.8	98.5	101.3	90.9	81.6	98.3
Max	211.2	238.3	247.3	234.6	241.8	232.2	228.6	247.3

Note: Max, maximum; min, minimum; PM<sub>2.5</sub>, particulate matter ≤2.5 µm in aerodynamic diameter (fine particulate matter); Q, quartile; SD, standard deviation.

<sup>a</sup>Cluster is a primary sampling unit (PSU) or a segment of PSU, of around 100–150 households, in the two-stage clustering sampling design of the DHS/NFHS-4.

<sup>b</sup>9-month average of ambient PM<sub>2.5</sub> exposure prior to child birth month.

<sup>c</sup>Average of post-delivery lifetime ambient PM<sub>2.5</sub> exposure from month of child birth until the month of death, or month of DHS/NFHS-4 interview of the child's mother (censoring).

In the two-exposure Cox regression models, both *in utero* and post-delivery lifetime average PM<sub>2.5</sub> exposures were significantly ( $p < 0.05$ ) associated with increases in child mortality, infant mortality, and neonatal mortality, after controlling for covariates (Table 3). A 10-µg/m<sup>3</sup> increase of *in utero* PM<sub>2.5</sub> exposure was associated with an increase in neonatal mortality [HR = 1.018; 95% confidence interval (CI): 1.001, 1.035], infant mortality (HR = 1.021; 95% CI: 1.006, 1.037), and child mortality (HR = 1.023; 95% CI: 1.008, 1.038). A 10-µg/m<sup>3</sup> increase of post-delivery average PM<sub>2.5</sub> exposure was associated with an increase in neonatal mortality (HR = 1.017; 95% CI: 1.003, 1.030), infant mortality (HR = 1.015; 95% CI: 1.003, 1.027), and child mortality (HR = 1.013; 95% CI: 1.001, 1.026). In the single-exposure Cox regression models, with either *in utero* PM<sub>2.5</sub> exposure or post-delivery average PM<sub>2.5</sub> exposure, we found similar but relatively stronger PM<sub>2.5</sub>-mortality HRs compared with HRs in the two-exposure models (Table 3). In the parameter correlation estimates in the two-exposure models, we found moderate-to-high correlation between exposure *in utero* and post-delivery PM<sub>2.5</sub> (Table S7). However, the two-exposure models had lower AICs compared with the single-exposure models (Table S8). Table S9 lists the HRs and 95% CIs of covariates in the two-exposure models of child mortality, and Table S10 lists the HRs and 95% CIs of mortality per IQR increase of PM<sub>2.5</sub> exposure in the single- and two-exposure models.

We found little evidence of effect modification by most covariates on PM<sub>2.5</sub> exposure on mortality in subgroup analysis. Figure 2 shows the estimated *in utero* and post-delivery lifetime average ambient PM<sub>2.5</sub> effects on child mortality across covariates, and the results were similar for infant mortality and neonatal mortality (Figures S7 and S8). We observed some heterogeneity of post-delivery PM<sub>2.5</sub>-mortality HRs across different maternal height, maternal age, and LBW, indicating some interactions between these covariates and post-delivery PM<sub>2.5</sub> exposure. Numerical values for these results are shown in Table S11. For children with LBW, we observed *in utero* PM<sub>2.5</sub>-mortality HRs of <1. In addition, we found that LBW was associated with both *in utero* PM<sub>2.5</sub> exposure and child mortality, as shown in additional analyses (Tables S12 and S13). Our results were robust to different variable selections (modify 1–3 in Figure S9, Table S14), unweighted analysis (modify 4 in Figure S9, Table S14), and Cox regressions with generalized estimating equations to account for the correlation of outcomes within households and clusters (modifies 5 and 6 in Figure S9, Table S14), inclusion of time-dependent post-delivery monthly PM<sub>2.5</sub> exposure to replace post-delivery lifetime PM<sub>2.5</sub> exposure (modify 7 in Figure S9, Table S14), and crude models with single or two exposures (modify 8 in Figure S9, Table S14). The results for the categorized exposure models are listed in Table S15, showing a consistent linear trend for both categorized *in utero* and post-delivery PM<sub>2.5</sub> exposure on mortality.

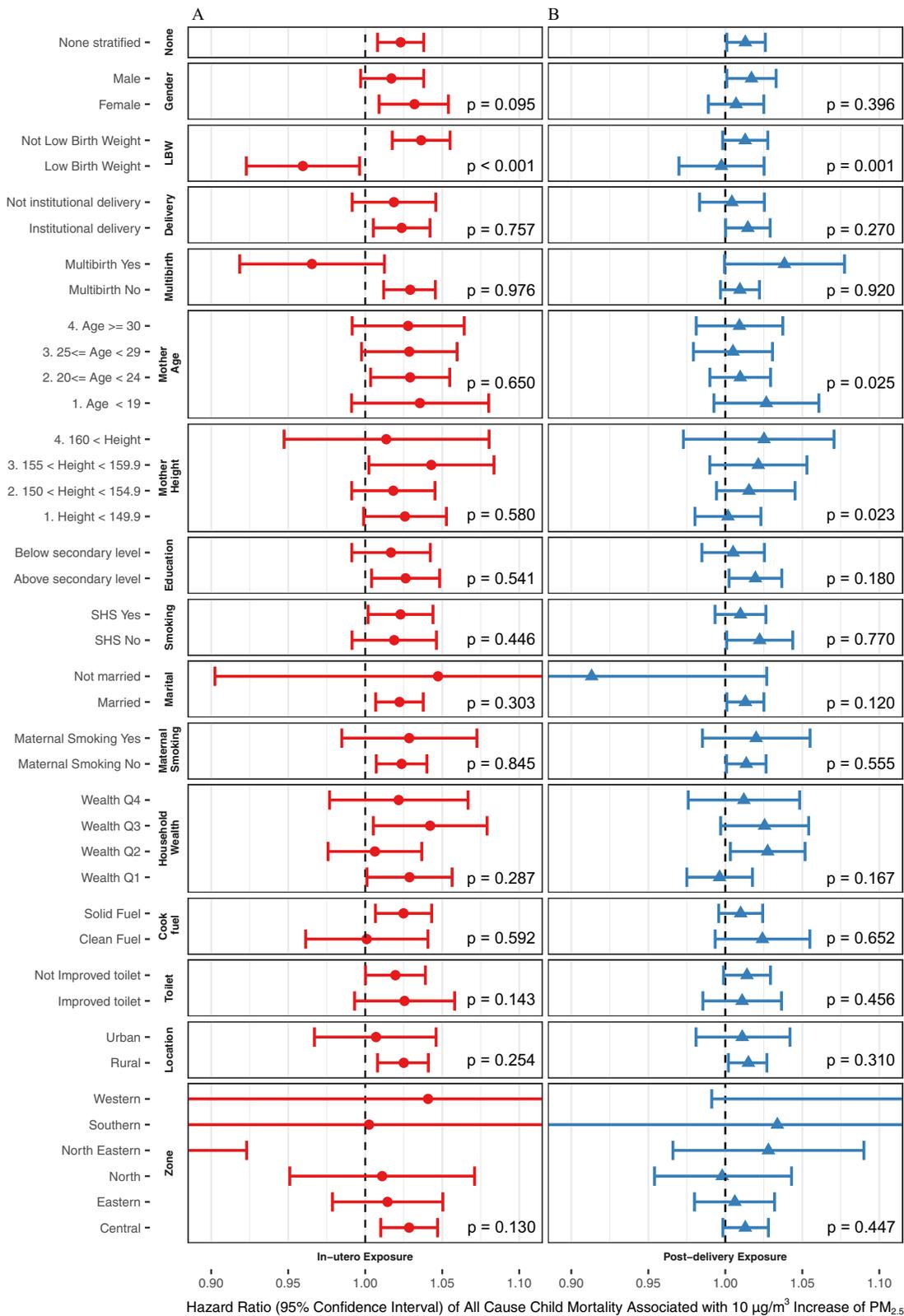
**Table 3.** Hazard ratios (95% confidence intervals) of neonatal (<1 month), infant (<12 months), and child (<60 months) mortality per 10-µg/m<sup>3</sup> increase of ambient PM<sub>2.5</sub> exposure for children <5 years of age in the Demographic and Health Surveys [DHS, also known as the National Family Health Survey 2015–2016 (NFHS-4)] in India.

Models		All-cause neonatal mortality (<1 month of age)	All-cause infant mortality (<12 months of age)	All-cause child mortality (<60 month of age)
Deaths (n)		7,520	10,862	11,559
Two-exposure models <sup>a</sup>	<i>In utero</i> PM <sub>2.5</sub>	1.018 (1.001, 1.035)	1.021 (1.006, 1.037)	1.023 (1.008, 1.038)
	Post-delivery average PM <sub>2.5</sub>	1.017 (1.003, 1.030)	1.015 (1.003, 1.027)	1.013 (1.001, 1.026)
Single-exposure models <sup>b</sup>	<i>In utero</i> PM <sub>2.5</sub>	1.032 (1.019, 1.045)	1.033 (1.021, 1.044)	1.033 (1.022, 1.044)
	Post-delivery average PM <sub>2.5</sub>	1.025 (1.015, 1.035)	1.025 (1.016, 1.034)	1.025 (1.016, 1.033)

Note: Adjusted hazard ratios and 95% confidence intervals for all-cause neonatal mortality, infant mortality and child mortality are shown for 10-µg/m<sup>3</sup> increase of ambient PM<sub>2.5</sub> during 9-month *in utero* period before child birth and post-delivery periods. Models stratified on child sex, birth month and year, geographical zone (927 strata), adjusted for birth order, multibirth, birth location, mother's age, height, marital, education, maternal smoking, household wealth, secondhand smoking, cooking fuel, improved toilet, urban or rural location of households, monthly temperature, monthly precipitation. PM<sub>2.5</sub>, particulate matter ≤2.5 µm in aerodynamic diameter (fine particulate matter).

<sup>a</sup>Two-exposure models include both ambient PM<sub>2.5</sub> exposure during 9-month *in utero* period before child birth and post-delivery lifetime average until death, or censoring.

<sup>b</sup>Single-exposure models include either one of PM<sub>2.5</sub> exposure during 9-month *in utero* period before child birth or post-delivery average until death, or censoring.



**Figure 2.** Hazard ratios (HRs) and 95% confidence intervals (CIs) of child mortality per 10- $\mu\text{g}/\text{m}^3$  increase of ambient  $\text{PM}_{2.5}$  exposure, according to subgroup analysis for the two-exposure models for children <5 years of age in the Demographic and Health Surveys [DHS, also known as the National Family Health Survey 2015–2016 (NFHS-4)] in India. Adjusted HRs and 95% CIs for all-cause child mortality are shown for 10- $\mu\text{g}/\text{m}^3$  increases in ambient  $\text{PM}_{2.5}$  exposure during both (A) *in utero* and (B) post-delivery lifetime periods, stratified on subgroups by individual and household-level covariates. Some error bars and points are off the limits outside the image. All numerical values can be found in Table S8. All model specifications are the same as for the main analysis except for the subgroup analysis of low birth weight (LBW), which is not included in the main analysis. Red circles (A) represent effects of *in utero*  $\text{PM}_{2.5}$  exposure, and blue triangles (B) represent the effects of post-delivery lifetime average  $\text{PM}_{2.5}$  exposure. *p*-Values for testing differences among strata for each stratifying variable are listed. Note:  $\text{PM}_{2.5}$ , particulate matter  $\leq 2.5 \mu\text{m}$  in aerodynamic diameter (fine particulate matter); Q, quartile; SHS, secondhand smoking.

As for the child mortality burden from air pollution, we estimated that 17.6% (95% CI: 12.3%, 22.4%) of mortality could be reduced if exposures were reduced to 10  $\mu\text{g}/\text{m}^3$  and that 9.3% (95% CI: 6.5%, 11.9%) of mortality could be reduced if exposures were reduced to the National Ambient Air Quality Standard of India of 40  $\mu\text{g}/\text{m}^3$  (Table S16).

## Discussion

In a large, nationally representative retrospective cohort of children <5 years of age in India, we found a consistent association between exposure to ambient  $\text{PM}_{2.5}$  during *in utero* and post-delivery periods and elevated child mortality. The estimated associations were stronger in single-exposure models compared with two-exposure models, perhaps due to the strong correlation between *in utero* and post-delivery exposures. Similar results were observed for infant and neonatal mortality. Even though *in utero* and post-delivery lifetime  $\text{PM}_{2.5}$  exposure showed some multicollinearity in the two-exposure models, two-exposure models had smaller AICs compared with single-exposure models, indicating the proper fit of the two-exposure models. To our knowledge, this is the first cohort study conducted in India to estimate the effects of ambient  $\text{PM}_{2.5}$  exposure on child survival in a nationally representative cohort of children.

To estimate monthly  $\text{PM}_{2.5}$  concentrations, we developed a model based on multiple inputs from satellite observations, meteorological data sets, and land use information from between 2009 and 2018. We identified disproportionately high levels of  $\text{PM}_{2.5}$  in the Indo-Gangetic Plain Region, where the population density is high. In addition, we found increasing levels of ambient  $\text{PM}_{2.5}$  over the past 10 y, emphasizing an urgent need to control ambient air pollution in India. Children included in our study had high exposure levels to ambient  $\text{PM}_{2.5}$  at nearly twice the National Ambient Air Quality Standard of India (40  $\mu\text{g}/\text{m}^3$ ) (Gautam 2009) and more than seven times the WHO Air Quality Guidelines (10  $\mu\text{g}/\text{m}^3$ ) (WHO 2005). Given the high levels of ambient  $\text{PM}_{2.5}$  throughout India, reducing ambient air pollution levels could provide a relatively large reduction in child mortality and substantial health benefits.

The child mortality-HR estimates per 10  $\mu\text{g}/\text{m}^3$  of  $\text{PM}_{2.5}$  exposure in our study were in the range of 1–3%, lower than previous estimates in LMICs. In sub-Saharan Africa, Heft-Neal et al. (2018) found a 9% increase in risk of infant mortality associated with a 10- $\mu\text{g}/\text{m}^3$  increase in annual  $\text{PM}_{2.5}$  concentration, and Loomis et al. (1999) found a 7% increase of infant mortality associated with a 10- $\mu\text{g}/\text{m}^3$  increase of  $\text{PM}_{2.5}$  concentration during the 3–5 d before death in Mexico City. In the present study, given that the  $\text{PM}_{2.5}$  exposure range was three times higher than those in previous studies, the estimated increases of mortality per IQR (one IQR = 40.6  $\mu\text{g}/\text{m}^3$ ) increase of  $\text{PM}_{2.5}$  were in the range of 5–12%, more similar to earlier findings (Heft-Neal et al. 2018; Loomis et al. 1999). One study, which assessed early life ambient  $\text{PM}_{2.5}$  exposure and child mortality in 43 LMICs, did not find a significant association between air pollution and child mortality (Goyal et al. 2019). This could be due to their use of annual  $\text{PM}_{2.5}$  exposure as a proxy for exposure at <1 year of age, which thus introduced measurement error and which may have biased results toward the null (Zeger et al. 2000). Another study conducted in Beijing using a time-series design also did not find a significant association between current-month  $\text{PM}_{2.5}$  exposure and infant mortality (Wang et al. 2019).

The present study has several strengths. First, we used machine learning to develop a high-quality model of ambient  $\text{PM}_{2.5}$  concentrations that had monthly temporal resolution and high spatial resolution that could reflect seasonal differences in ambient air

pollution. These estimates of ambient  $\text{PM}_{2.5}$  levels could be potentially applied in future health impact assessment and epidemiological analysis in India. Second, we used a large, retrospective, and nationally representative cohort of children <5 years of age in India based on the well-documented and -conducted DHS/NHFS-4 survey. Third, we applied Cox regression with time-dependent air pollution, temperature, and rainfall precipitation on child mortality using the calendar month as the time variable. This approach allowed us to examine the longitudinal association between air pollution and child mortality and compare post-delivery  $\text{PM}_{2.5}$  exposure for children with the same follow-up time. This approach could largely reduce exposure misclassification compared with previous studies (Goyal et al. 2019; Heft-Neal et al. 2018), which were based on yearly  $\text{PM}_{2.5}$  averages to assess a binary outcome of child death at the end of 1 year of age. Our analysis controlled for potential bias from seasonality and different weather patterns, in addition to using individual-level, household-level, and cluster-level covariates. Notably, we did not include LBW as a covariate in the model due to the potential selection bias of including LBW as a covariate, and for children with LBW, we found protective effects of *in utero*  $\text{PM}_{2.5}$  exposure. This is similar to the protective effects of maternal smoking on infant mortality for children with LBW observed in the United States as a result of selection bias (Hernández-Díaz et al. 2006). Finally, the  $\text{PM}_{2.5}$ –mortality associations were consistently positive in single- and two-exposure models, and these associations were robust in additional sensitivity analyses.

Our study also has some limitations. First, no national  $\text{PM}_{2.5}$  monitoring data were available during the study period of 2009–2014, so we lacked ground-level  $\text{PM}_{2.5}$  measurement data before 2015 with which to validate our  $\text{PM}_{2.5}$  prediction model. The monitoring stations were more heavily concentrated in urban areas, which could have led to differential errors across the country as well. Similar to previous  $\text{PM}_{2.5}$  prediction models (Xiao et al. 2018), our model tended to underestimate some very high  $\text{PM}_{2.5}$  values, when monthly  $\text{PM}_{2.5}$  concentrations were >350  $\mu\text{g}/\text{m}^3$  (Figure S3). Second, we used only total ambient  $\text{PM}_{2.5}$  mass concentration at the cluster level as our exposure for participants at the household level, without estimation of different  $\text{PM}_{2.5}$  components, and did not include the impact of other ambient air pollutants, such as PM with an aerodynamic diameter of  $\geq 10 \mu\text{m}$ ,  $\text{NO}_2$ , and ozone, or direct household air pollution exposures from cooking, which are significant in many parts of India. Third, our retrospective cohort was based on a cross-sectional survey of the DHS/NHFS-4. We focused only on all-cause mortality reported in the survey, whereas it is likely that ambient  $\text{PM}_{2.5}$  is associated with specific cause of death. Child birth year, birth month, vital information, and covariates were collected at the end of the study and recalled by mothers and may have differed from values at the time the children were born. In addition, we did not have the children's gestational age at birth and assigned *in utero*  $\text{PM}_{2.5}$  exposure levels as a 9-month average prior to the birth, which may not be the case for preterm infants who have a shorter *in utero* exposure period. Given this data collection limitation, our study could have suffered from recall bias and misclassification. However, we believe most of the maternal and household-level covariates will be nondifferential misclassification of outcomes and could most likely bias our results toward the null. Last, as with all observational studies, our study may have potential unmeasured confounders that were inadequately controlled for in this analysis. Our results indicate a need for future prospective cohort studies (Balakrishnan et al. 2015), including intervention trials of reducing air pollution levels (Clasen et al. 2020), to investigate air pollution effects on child health.

In conclusion, this study found increased mortality risk associated with ambient  $\text{PM}_{2.5}$  during the 9-month *in utero* and post-delivery periods for children <5 years of age. This study, based

on a nationally representative retrospective cohort in India, provides evidence suggesting that *in utero* and post-delivery PM<sub>2.5</sub> exposure contribute to child mortality in developing countries. Given the high levels of ambient PM<sub>2.5</sub> throughout India, expanding air pollution monitoring stations (both in rural and urban areas), adding more epidemiological research, and making a substantial effort to reduce ambient air pollution and early life exposures are needed.

## Acknowledgments

J.L., Y.L., and T.C. designed this study; J.L. conducted the analysis with contribution from Y.L., S.D., and K.S. J.L. wrote the first draft of the paper with contributions from Y.L., K.S., L.M.T., A.P., K.B., and T.C. All authors provided critical input to paper, interpreted the data, revision of the manuscript, and contributed significant intellectual content.

Funding was provided by the National Aeronautics and Space Administration (NASA): The work of Y.L. was supported by the NASA Applied Sciences Program (grants NNX16AQ28Q and 80NSSC19K0191) and by the Multi-Angle Imager for Aerosols Science Team at the NASA Jet Propulsion Laboratory, California Institute of Technology (contract 1588347). The work of S.D. was supported by the Indian Institute of Technology Delhi Institute Chair Fellowship. A.P. was partially supported by the National Institutes of Health/National Institute of Environmental Health Sciences HERCULES Exposome Research Center (grant P30ES019776). The authors took a neutral position with respect to territorial claims in published maps.

The data set on predicted ambient monthly air PM<sub>2.5</sub> concentration can be accessed online at <https://doi.org/10.15139/S3/RFILYH>.

## References

Balakrishnan K, Ghosh S, Thangavel G, Sambandam S, Mukhopadhyay K, Puttaswamy N, et al. 2018. Exposures to fine particulate matter (PM<sub>2.5</sub>) and birthweight in a rural-urban, mother-child cohort in Tamil Nadu, India. *Environ Res* 161:524–531, PMID: 29227900, <https://doi.org/10.1016/j.envres.2017.11.050>.

Balakrishnan K, Sambandam S, Ramaswamy P, Ghosh S, Venkatesan V, Thangavel G, et al. 2015. Establishing integrated rural-urban cohorts to assess air pollution-related health effects in pregnant women, children and adults in Southern India: an overview of objectives, design and methods in the Tamil Nadu Air Pollution and Health Effects (TAPHE) study. *BMJ Open* 5(6):e008090, PMID: 26063570, <https://doi.org/10.1136/bmjopen-2015-008090>.

Bell ML, Ebisu K, Belanger K. 2007. Ambient air pollution and low birth weight in Connecticut and Massachusetts. *Environ Health Perspect* 115(7):1118–1124, PMID: 17637932, <https://doi.org/10.1289/ehp.9759>.

Bosilovich MG, Akella S, Coy L, Cullather R, Draper C, Gelaro R, et al. 2015. *MERRA-2: Initial Evaluation of the Climate. Technical Report Series On Global Modeling And Data Assimilation, version 43*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Brauer M, Guttikunda SK, K A N, Dey S, Tripathi SN, Weagle C, et al. 2019. Examination of monitoring approaches for ambient air pollution: a case study for India. *Atmos Environ* 216:116940, <https://doi.org/10.1016/j.atmosenv.2019.116940>.

Clasen T, Checkley W, Peel JL, Balakrishnan K, McCracken JP, Rosa G, et al. 2020. Design and rationale of the HAPIN study: a multicountry randomized controlled trial to assess the effect of liquefied petroleum gas stove and continuous fuel distribution. *Environ Health Perspect* 128(4):047008, PMID: 32347766, <https://doi.org/10.1289/EHP6407>.

Croft TN, Marshall AMJ, Allen CK. 2018. *Guide to DHS Statistics. DHS-7: The Demographic and Health Surveys Program*. Rockville, MD: U.S. Agency for International Development. [https://www.dhsprogram.com/pubs/pdf/DHSG1/Guide\\_to\\_DHS\\_Statistics\\_DHS-7.pdf](https://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_DHS-7.pdf) [accessed 6 January 2022].

Funk C, Peterson P, Landsfeld M, Pederos D, Verdin J, Shukla S, et al. 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci Data* 2:150066, PMID: 26646728, <https://doi.org/10.1038/sdata.2015.66>.

Gautam SP. 2009. National Ambient Air Quality Standards—Central Pollution Control Board Notification. New Delhi, India. [https://cpcb.nic.in/uploads/National\\_Ambient\\_Air\\_Quality\\_Standards.pdf](https://cpcb.nic.in/uploads/National_Ambient_Air_Quality_Standards.pdf) [accessed 10 January 2022].

GBD 2019 Diseases and Injuries Collaborators. 2020. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396(10258):1204–1222, PMID: 33069326, [https://doi.org/10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9).

Glinianaia SV, Rankin J, Bell R, Pless-Mulloli T, Howel D. 2004. Does particulate air pollution contribute to infant death? A systematic review. *Environ Health Perspect* 112(14):1365–1370, PMID: 15471726, <https://doi.org/10.1289/ehp.6857>.

Goyal N, Karra M, Canning D. 2019. Early-life exposure to ambient fine particulate air pollution and infant mortality: pooled evidence from 43 low- and middle-income countries. *Int J Epidemiol* 48(4):1125–1141, PMID: 31074784, <https://doi.org/10.1093/ije/dyz090>.

Gurley ES, Homaira N, Salje H, Ram PK, Haque R, Petri W, et al. 2013. Indoor exposure to particulate matter and the incidence of acute lower respiratory infections among children: a birth cohort study in urban Bangladesh. *Indoor Air* 23(5):379–386, PMID: 23906055, <https://doi.org/10.1111/ina.12038>.

Hans S, Mwailili SM, Faes C, Fierens F, Nemery B, Nawrot TS. 2011. Does air pollution trigger infant mortality in Western Europe? A case-crossover study. *Environ Health Perspect* 119(7):1017–1022, PMID: 21269926, <https://doi.org/10.1289/ehp.1002913>.

Heft-Neal S, Burney J, Bendavid E, Burke M. 2018. Robust relationship between air quality and infant mortality in Africa. *Nature* 559(7713):254–258, PMID: 29950722, <https://doi.org/10.1038/s41586-018-0263-3>.

Hernández-Díaz S, Schisterman EF, Hernán MA. 2006. The birth weight “paradox” uncovered? *Am J Epidemiol* 164(11):1115–1120, PMID: 16931543, <https://doi.org/10.1093/aje/kwj275>.

Huynh M, Woodruff TJ, Parker JD, Schoendorf KC. 2006. Relationships between air pollution and preterm birth in California. *Paediatr Perinat Epidemiol* 20(6):454–461, PMID: 17052280, <https://doi.org/10.1111/j.1365-3016.2006.00759.x>.

IIPS (International Institute for Population Sciences), ICF (ICF International). 2017. *National Family Health Survey (NFHS-4), 2015–16*. Mumbai, India: IIPS.

India State-Level Disease Burden Initiative Air Pollution Collaborators. 2019. The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: the Global Burden of Disease Study 2017. *Lancet Planet Health* 3(1):E26–E39, PMID: 30528905, [https://doi.org/10.1016/S2542-5196\(18\)30261-4](https://doi.org/10.1016/S2542-5196(18)30261-4).

India State-Level Disease Burden Initiative Air Pollution Collaborators. 2021. Health and economic impact of air pollution in the states of India: the Global Burden of Disease Study 2019. *Lancet Planet Health* 5(1):E25–E38, PMID: 33357500, [https://doi.org/10.1016/S2542-5196\(20\)30298-9](https://doi.org/10.1016/S2542-5196(20)30298-9).

Landrigan PJ, Fuller R, Acosta NJR, Adeyi O, Arnold R, Basu NN, et al. 2018. The Lancet Commission on pollution and health. *Lancet* 391(10119):462–512, PMID: 29056410, [https://doi.org/10.1016/S0140-6736\(17\)32345-0](https://doi.org/10.1016/S0140-6736(17)32345-0).

Levy R, Hsu C. 2015. MODIS Atmosphere L2 Aerosol Product. NASA MODIS Adaptive Processing System. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center. [https://doi.org/10.5067/MODIS/MOD04\\_L2\\_006](https://doi.org/10.5067/MODIS/MOD04_L2_006).

Li Q, Wang YY, Guo Y, Zhou H, Wang X, Wang Q, et al. 2018. Effect of airborne particulate matter of 2.5 μm or less on preterm birth: a national birth cohort study in China. *Environ Int* 121(pt 2):1128–1136, PMID: 30352698, <https://doi.org/10.1016/j.envint.2018.10.025>.

Loomis D, Castillejos M, Gold DR, McDonnell W, Borja-Aburto VH. 1999. Air pollution and infant mortality in Mexico City. *Epidemiology* 10(2):118–123, PMID: 10069245, <https://doi.org/10.1097/00001648-199903000-00006>.

Meijer JR, Huijbregts MAJ, Schotten KCGJ, Schipper AM. 2018. Global patterns of current and future road infrastructure. *Environ Res Lett* 13(6):064006, <https://doi.org/10.1088/1748-9326/aabd42>.

Pedersen M, Giorgis-Allemand L, Bernard C, Aguilera I, Andersen AMN, Ballester F, et al. 2013. Ambient air pollution and low birthweight: a European cohort study (ESCAPE). *Lancet Respir Med* 1(9):695–704, PMID: 24429273, [https://doi.org/10.1016/S2213-2600\(13\)70192-9](https://doi.org/10.1016/S2213-2600(13)70192-9).

Shaddick G, Thomas ML, Amini H, Broday D, Cohen A, Frostad J, et al. 2018. Data integration for the assessment of population exposure to ambient air pollution for Global Burden of Disease assessment. *Environ Sci Technol* 52(16):9069–9078, PMID: 29957991, <https://doi.org/10.1021/acs.est.8b02864>.

Smith KR, McCracken JP, Weber MW, Hubbard A, Jenny A, Thompson LM, et al. 2011. Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial. *Lancet* 378(9804):1717–1726, PMID: 22078686, [https://doi.org/10.1016/S0140-6736\(11\)60921-5](https://doi.org/10.1016/S0140-6736(11)60921-5).

Son JY, Lee HJ, Koutrakis P, Bell ML. 2017. Pregnancy and lifetime exposure to fine particulate matter and infant mortality in Massachusetts, 2001–2007. *Am J Epidemiol* 186(11):1268–1276, PMID: 29121205, <https://doi.org/10.1093/aje/kwx015>.

Spears D, Dey S, Chowdhury S, Scovronick N, Vyas S, Apte J. 2019. The association of early-life exposure to ambient PM<sub>2.5</sub> and later-childhood height-for-age in India: an observational study. *Environ Health* 18(1):62, PMID: 31288809, <https://doi.org/10.1186/s12940-019-0501-7>.

Staveteig S, Mallick L. 2014. *Intertemporal Comparisons of Poverty and Wealth with DHS Data: A Harmonized Asset Index Approach*. DHS Methodological

- Reports 15. Rockville, MD: ICF International. <https://www.dhsprogram.com/pubs/pdf/MR15/MR15.pdf> [accessed 6 January 2022].
- Stieb DM, Chen L, Eshoul M, Judek S. 2012. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. *Environ Res* 117:100–111, PMID: 22726801, <https://doi.org/10.1016/j.envres.2012.05.007>.
- Subramanian SV, Ackerson LK, Smith GD, John NA. 2009. Association of maternal height with child mortality, anthropometric failure, and anemia in India. *JAMA* 301(16):1691–1701, PMID: 19383960, <https://doi.org/10.1001/jama.2009.548>.
- van Donkelaar A, Martin RV, Brauer M, Hsu NC, Kahn RA, Levy RC, et al. 2016. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environ Sci Technol* 50(7):3762–3772, PMID: 26953851, <https://doi.org/10.1021/acs.est.5b05833>.
- Vu BN, Sánchez O, Bi J, Xiao Q, Hansel NN, Checkley W, et al. 2019. Developing an advanced PM<sub>2.5</sub> exposure model in Lima, Peru. *Remote Sens (Basel)* 11(6):641, PMID: 31372305, <https://doi.org/10.3390/rs11060641>.
- Wang J, Cao H, Sun D, Qi Z, Guo C, Peng W, et al. 2019. Associations between ambient air pollution and mortality from all causes, pneumonia, and congenital heart diseases among children aged under 5 years in Beijing, China: a population-based time series study. *Environ Res* 176:108531, PMID: 31226628, <https://doi.org/10.1016/j.envres.2019.108531>.
- WHO (World Health Organization). 2005. *Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide. Global Update 2005: Summary of Risk Assessment*. WHO/SDE/PHE/OEH/06.02. Copenhagen, Denmark: WHO.
- WHO/UNICEF JMP for Water Supply, Sanitation and Hygiene (WHO/United Nations International Children's Fund Joint Monitoring Programme for Water Supply, Sanitation and Hygiene). 2018. *JMP Methodology 2017 Update & SDG Baselines*. <https://washdata.org/report/jmp-methodology-2017-update> [accessed 6 January 2022].
- Woodruff TJ, Grillo J, Schoendorf KC. 1997. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. *Environ Health Perspect* 105(6):608–612, PMID: 9288495, <https://doi.org/10.1289/ehp.97105608>.
- Woodruff TJ, Parker JD, Schoendorf KC. 2006. Fine particulate matter (PM<sub>2.5</sub>) air pollution and selected causes of postneonatal infant mortality in California. *Environ Health Perspect* 114(5):786–790, PMID: 16675438, <https://doi.org/10.1289/ehp.8484>.
- Xiao Q, Chang HH, Geng G, Liu Y. 2018. An ensemble machine-learning model to predict historical PM<sub>2.5</sub> concentrations in China from satellite data. *Environ Sci Technol* 52(22):13260–13269, PMID: 30354085, <https://doi.org/10.1021/acs.est.8b02917>.
- Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, et al. 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect* 108(5):419–426, PMID: 10811568, <https://doi.org/10.1289/ehp.00108419>.