**Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Jiawen Liao Date

Advanced Exposure Assessment of Air Pollution and its Effects on Maternal and Child Health in Low-income Settings

By

Jiawen Liao
Doctor of Philosophy

Environmental Health Sciences

Thomas F. Clasen, Ph.D., J.D.
Advisor

Lisa M. Thompson, Ph.D.
Advisor

Yang Liu, Ph.D.
Committee Member

Kyle Steenland, Ph.D.
Committee Member

John P. McCracken, Ph.D.
Committee Member

Jeremy Sarnat, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

# Advanced Exposure Assessment of Air Pollution and its Effects on Maternal and Child Health in Low-income Settings

By

Jiawen Liao
M.S., University of California, Berkeley

Advisor: Thomas F. Clasen, Ph.D., J.D.
Advisor: Lisa M. Thompson, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Environmental Health Sciences
2020

Abstract

Advanced Exposure Assessment of Air Pollution and its Effects on Maternal and Child Health in Low-income Settings

By Jiawen Liao

In this dissertation, the author explored advancing exposure assessment methods of air pollution and analyzed the maternal and child health effects of air pollution exposure from solid fuel combustion and ambient environment in low-income settings in Guatemala and India. Household air pollution from solid fuel combustion is a leading health risk factor for disease in low-income settings. Using data from the Household Air Pollution Intervention Network (HAPIN) trial in Guatemala, the author showed the feasibility of indirectly assessing household air pollution exposure in children by personal use of Bluetooth signal emitters for microenvironment classification. Based on a prospective cohort of pregnant women from HAPIN trial in Guatemala, the author then investigated the effects of gestational blood pressure effects of fine particulate matter ($PM_{2.5}$) exposure. In additional to household air pollution, there are growing risks of ambient air pollution in low- and middle-income countries such as China and India, due to rapid industrialization and urbanization. The author developed a monthly ambient $PM_{2.5}$ exposure prediction model using a machine-learning algorithm, based on satellite remote sensing, meteorology and land use information, cross-validated with ground-based $PM_{2.5}$ monitoring stations over 10-year period in India. Ultimately, the author examined the adverse effects of ambient $PM_{2.5}$ exposure on child all-cause mortality under age of five in India, linking predicted ambient air pollution and a retrospective birth cohort.

# Advanced Exposure Assessment of Air Pollution and its Effects on Maternal and Child Health in Low-income Settings

By

Jiawen Liao
M.S., University of California, Berkeley

Advisor: Thomas F. Clasen, Ph.D., J.D.
Advisor: Lisa M. Thompson, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Environmental Health Sciences
2020

# Acknowledgments

Many people were involved in making my academic career here at Emory a success. I feel very privileged to be guided and advised by my six committee members, who provided unmatchable academic and moral supports all along the way: Dr. Thomas Clasen and Dr. Lisa Thompson, my committee chairs, for their enormous mentorships, guidance and friendship throughout all my time at Emory and in Guatemala; Dr. Kyle Steenland, for his extensive guidance and expertise in environmental epidemiology and statistical analysis throughout my dissertation that equips me with strong epidemiological and statistical skills; Dr. Yang Liu, for his diligent mentorship in exposure assessment and air pollution modelling; Dr. John McCracken, for his strong and supportive guidance and leadership throughout my involvement in Guatemala HAPIN project field work and data analysis; Dr. Jeremy Sarnat, for his thoughtful advice on exposure assessment, that challenges me to become an exposure scientist.

I also need to express my gratitude towards collaborators from HAPIN projects, Anaité Diaz, Oscar de Leon, Erick Molliendo and Eduardo Canuz and many others from Universidad del Valley de Guatemala (UVG), for the great and fruitful summer I spent in 2016 in Guatemala. I also thank other collogues and experts in HAPIN projects, from who I expanded my knowledge in environmental epidemiology and exposure assessment in household air pollution: Dr. Jennifer Peel, Dr. William Checkley, Dr. Joshua Rosenthal and Dr. Michael Johnson, Dr. Ricardo Piedrahita and many others. Their dedication to HAPIN project truly inspired me to keep learning.

In addition, I feel greatly indebted to be advised and supported by numerous faculty members at Rollins School of Public Health: Dr. Paige Tolbert, the department chair, Dr. Carmen Marsit, Dr. Ajay Pillarisetti, Dr. Barry Ryan, Dr. Dana Barr, Dr. Stefanie Ebelt, Dr. Qiang Zhang, Dr. Noah Scovronick, Dr. Mattew Freeman, Dr. Donghai Liang, Dr. Eri Saikawa, Dr. Lance Waller and Dr. Howard Chang for their professional teaching and encouragement in all means along my academic pursuit.

Thank you to group members and collogues from Clasen Research Group and Liu Research Group: Dr. Miles Kirby, Amy Lovorn, Dr. Sheela Sinharoy, Dr. Valerie Bauza, Wenlu Ye, Frederick Goddard, Jianzhao Bi, Lin Wang and many others, for their diligent collaboration, heartful help and truly friendship. I also want to thank my classmates at Environmental Health Sciences PhD project, for all the time we spent together at this program and all unforgettable memories we shared together at Emory.

In addition, in this unprecedented time COVID-19 pandemic and public health crisis, I want to express my sincere appreciation for health care, public health and essential business workers who make my completion of my dissertation possible. They unwaveringly encouraged me to continue my academic and professional career in protecting public health.

Lastly, I want to thanks my parents Xiaojing Yang and Tao Liao for their endless love.

# Table of Contents

# Chapter 1 **Introduction**

## 1.1 Motivation

Air pollution, both household and ambient, is among the major health risk factors globally, responsible for over 4.5 million premature death in 2017 (Roth et al., 2018). Even though all countries, rich or poor, are adversely affected by air pollution, the health burden is disproportionally large among low- and middle-income countries (LMICs) (Health Effects Institute, 2019). There are several reasons for this. First, in LMICs, a large proportion of the population rely on solid fuel combustion for cooking and heating, generating high levels of household air pollution (HAP) with $PM_{2.5}$ and carbon monoxide (CO) concentrations that are orders of magnitude greater compared to those using cleaner fuels (Shupler et al., 2018). Second, populations in LMICs are also exposed to high level of ambient air pollution (AAP), the air pollutant in the outdoor air environment. Based on global ambient $PM_{2.5}$ exposure model, nearly all populations exceeding the $PM_{2.5}$ exposure annual mean of 10 $\mu g/m^3$, a World Health Organization (WHO) annual guideline live in LMICs (Shaddick et al., 2018; World Health Organization, 2005b). The higher population average air pollution exposure in LMICs results in the larger life expectancy loss (Health Effects Institute, 2019). Populations living in LMICs experience a double burden of air pollution, both from HAP and AAP.

In LMICs, pregnant women and young children are particularly vulnerable to air pollution exposures. Women often cook meals for their households and thus have higher HAP exposures (Dasgupta et al., 2006). Biological susceptibility is increased in gestating fetuses and young children due to their physiologic immaturity and larger inhalation of air pollutant per body weight compared to adults. Children's airways are narrower compared to adults, and pulmonary function is immature until just before adulthood (Makri & Stilianakis, 2008). Maternal and child health effects of ambient and household air pollution have been widely studied; multiple adverse maternal and child health outcomes include adverse pregnancy outcomes (Amegah et al., 2014; Sun et al., 2015),

low birth weight (D. P. Pope et al., 2010a; Stieb et al., 2012), acute respiratory infections (ARI) (Dherani et al., 2008; Gurley et al., 2013; Romieu et al., 2002; Smith et al., 2011b) and child mortality (Glinianaia et al., 2004; Son et al., 2017; WHO Regional Office for Europe, 2013).

Despite the high absolute health burden of air pollution in populous LMICs like India and China, most of the epidemiological evidence of maternal and child health effects of air pollution comes from high-income regions, mainly in North America and Europe (Burnett et al., 2014, 2018). To inform the health risks of air pollution in LMICs and resource-limited settings, several research priorities have been identified, including: 1) improve characterization and mapping of air pollution exposure to elucidate health effects, especially in LMICs (Clark et al., 2013; Landrigan et al., 2018); 2) design informative observational studies that can collect sufficient personal air pollutant exposure data; and 3) design randomized controlled trials of contextualized and realistic interventions that take into account broader macroeconomic, infrastructure and environmental factors (Majid Ezzati & Baumgartner, 2017). Development and validation of air pollution measurement technologies and approaches to assess personal exposure to air pollutants, both AAP and HAP, plays a crucial role in designing effective air pollution interventions and informing policies.

Recently, global collaborative efforts have been taken to improve air pollution exposure assessment, especially with the refinement of estimates in the Global Burden of Disease (GBD) project in 2012 (Lim et al., 2012). The GBD 2012 project developed complex modelling techniques for HAP and AAP exposure globally, largely improving exposure data coverage for LMICs (Lim et al., 2012; Stanaway et al., 2018). While this was a substantial improvement over previous models, due to the lack of personal air pollution exposure assessment in household air pollution studies (Clark et al., 2013; Mortimer et al., 2017; Quansah et al., 2017) and the scarcity of ambient air pollution monitoring stations in low-resources settings (Brauer et al., 2019), accurate assignment of air pollution exposure in these regions is still lacking.

There are large research gaps in the understanding of air pollution health effects in lower-income settings. On the one hand, there is limited air pollution exposure data; on the other hand, the dearth of the high-quality health data available for epidemiological investigation limits an air pollution exposure-response analysis of health outcomes (Balakrishnan et al., 2018). A few cohort studies have been conducted to investigate health effects of air pollution exposure in low-income settings, such as in Western Highlands of Guatemala (Heinzerling et al., 2016; McCracken et al., 2007; Northcross et al., 2010; Smith et al., 2011a; Thompson et al., 2011), and in Tamil Nadu of India (Balakrishnan, Ghosh, et al., 2018a). However, knowledge gaps still exist in LMICs, especially for information on air pollution exposure-response relationships, which is mostly derived from studies in developed regions, even though LMICs have the highest exposures to air pollution (Burnett et al., 2014, 2018).

Motivated by the need to clarify the health effects of household air pollution, the Household Air Pollution Intervention Network (HAPIN) trial was designed to investigate potential health benefits of liquefied petroleum gas (LPG) cookstoves use in settings relying mainly on solid biomass fuels for cooking. The HAPIN trial was designed as a randomized controlled trial of an LPG cookstove and fuel intervention in 3,200 households with pregnant women from four low-income rural settings in India, Guatemala, Peru and Rwanda. The trial started recruiting in mid-2018, and follow-up will continue until children reach 1 year of age. The HAPIN trial aims to determine the effects of a randomized LPG cookstove and fuel intervention on health effects including birth weight, child severe pneumonia, child stunting and adult cardiovascular outcomes, including women's blood pressure. The trial will also investigate exposure-response relationships between household air pollution and these and other health outcomes. This is a valuable opportunity not only in air pollution and health research in LMICs but also an important project that can fill knowledge gaps in sustainable development goals (SDGs) regarding Good Health and Well-being (SDG 3), Affordable and Clean Energy (SDG 7), Climate Action (SDG 13) (Rosenthal et al., 2018).

In furtherance of the goals of the HAPIN trial, this dissertation aims to advance the knowledge in air pollution exposure assessment and provide evidence on the exposure-response relationship of air pollution exposure health effects, especially for children and pregnant women in resource-limited settings. This research will provide the basis of air pollution exposure assessment for newborn infants in the HAPIN trial and contribute to the epidemiological evidence on the gestational health effects of air pollution exposure.

While the HAPIN trial primarily focuses on the health effects of HAP, there is an increasing concern about AAP in LMICs. The number of ambient air pollutant monitoring stations in LMICs has increased in recent decades, particularly in China and India where air pollution levels are the highest in the world (Brauer et al., 2019). As of early 2020, there are approximately 1,800 air pollution monitoring stations in China and 430 monitoring stations in India (Ministry of Environment & Forests of Government of India, 2018; OpenAQ, 2020a, 2020b). The rapid development of ground monitoring stations in India over the past 4 years created an opportunity to validate a prediction model for temporally and spatially resolved high-quality ambient air pollution concentrations for use in epidemiological studies of adverse health outcomes among children exposed to ambient air pollution. Thus, the research presented here covers both HAP and AAP.

## 1.2 Dissertation Aims

As described above, the overall goal of the research presented in this dissertation is to advance the understanding of air pollution exposures and add evidence on the health effects of air pollution, particularly for low-resource settings where air pollution levels are highest. Chapter 2 of this dissertation reviews the current knowledge regarding household air pollution exposure assessment and epidemiological evidence of air pollution effects on maternal and child health in low-income settings, focusing on household air pollution effects on gestational blood pressure and ambient air pollution health effects on child health. Chapter 3 – 5 contains the research output of this dissertation research. Lastly, chapter 6 summarizes the main findings of this dissertation

research, reflections of my dissertation research during 4 years at Emory University, and future air pollution research directions.

Chapter 3 and 4 are based on HAPIN trial data from the Guatemala site, using data collected in the pilot phase (chapter 3) and main phase of trial (chapter 4). In the study described in chapter 3, I used data on personal direct and indirect $PM_{2.5}$ measurements of exposure in mothers and children from 20 households using biomass measured four times, twice before and twice after an LPG pilot intervention. In this study, I compared the indirect $PM_{2.5}$ exposure assessment method based on Bluetooth emitter beacon system with the direct method. I showed that this indirect measurement is a feasible approach to estimate $PM_{2.5}$ in larger HAPIN main trial, when direct measurement is not feasible, such as on children under 2 year of age.

Chapter 4 presents a study on the effects of $PM_{2.5}$ air pollution exposure on gestational blood pressure among pregnant women enrolled in the HAPIN trial in Guatemala. This study is based on data from 800 households in Guatemala cooking with biomass at baseline and at follow-up visits after households were randomized to the LPG intervention during the HAPIN main trial. Using exposure-response analysis, I investigated the effects of 24-hour personal $PM_{2.5}$ exposure effects on gestational blood pressure levels among HAPIN participants in Guatemala in a longitudinal follow-up with three observations before child birth. It shows the exposure to higher level of $PM_{2.5}$ air pollution from cooking with solid fuel is associated with increase of systolic blood pressure, but not diastolic blood pressure, adjusting for covariates.

Chapter 5 of this dissertation moves beyond household air pollution from solid fuel combustion and assesses the effect on child survival of ambient air pollution concentration predicted by a machine-learning model in a retrospective birth cohort in India. This Chapter is based on data from multiple publicly available sources, including both ambient air pollution data source as well as health outcomes. I collected the ground-based $PM_{2.5}$ monitoring station data from Central Pollution Control Board (CPCB) of Indian government, satellite remote sensing data, meteorology data from National Aeronautics and Space Administration (NASA) and land use

information, and built a machine-learning model based on random forest algorithm to predict monthly ambient $PM_{2.5}$ concentrations in India over 10 years from 2009 – 2018. Linking estimated ambient $PM_{2.5}$ exposure with a retrospective cohort constructed using most recent Demographic and Health Survey (DHS), in this study, I showed that ambient $PM_{2.5}$ air pollution during both the gestational period and post-delivery during early childhood were associated with child mortality. Nearly 18% premature mortality under year of 5 can be averted in India if ambient air pollution level were reduced to 10 $\mu g/m^3$, the ambient air quality guideline of World Health Organization.

Overall, this dissertation research will contribute new information to the air pollution research community and clean household energy implementation community. It is hoped that this research will provide new directions for air pollution exposure methods as well as enrich the knowledge base of air pollution exposure in resource-limited settings.

# Chapter 2 **Background**

Air pollution can be categorized mainly as ambient air pollution (AAP) and household air pollution (HAP) (Stanaway et al., 2018). AAP, coming from residential, transportation and industrial sources, is a complex air pollutant mixture than HAP. AAP includes carbon monoxide (CO), PM, nitrogen dioxide ($NO_2$), sulfur dioxide ($SO_2$), ozone ($O_3$) and lead (Pb), which are primary pollutants included in National Ambient Air Quality Standard (NAAQS)(US EPA, 2014). HAP is primarily generated from incomplete combustion of solid biomass fuels such wood, crop residue, animal dung and coal for cooking, heating and lighting (Clark et al., 2013; Naeher et al., 2007). Pollutants from HAP include CO, PM, polycyclic aromatic hydrocarbons (PAH), black carbon and other toxic chemicals. Due to infeasibilities of analyzing many of the HAP constituents in rural settings of developing countries, the primary focus of HAP is measuring PM and CO. Even though other constituents of solid fuel incomplete combustion have adverse effects, fine particulate matter— particulate matter that have a diameter of less than 2.5 micrometers ($PM_{2.5}$)— is mostly used to assess the exposure to HAP by current available instrument. Because $PM_{2.5}$ is a mixture of multiple incomplete combustion components, and is moderately correlation with other pollutants(McCracken et al., 2013), it is usually regarded as a surrogate for the effects of HAP (Clark et al., 2013; Gordon et al., 2014).

## 2.1 Air Pollution and Maternal and Child Health

The adverse health effects of air pollution, both from household sources and outdoor ambient environments, have been well documented since early 2000s(Brook et al., 2010; World Health Organization, 2005b, 2014). Particulate matter air pollution is positively associated with an increase in acute lower respiratory infection and childhood pneumonia, the leading cause of the child mortality in LMICs(Smith et al., 2014; Stanaway et al., 2018). The effects of air pollution on maternal and child health is substantial, responsible for around 7% of mortality under-5 globally (IHME, 2020), and associated with multiple adverse birth outcomes, such as low birth weight and preterm birth (Stieb et al., 2012). The most recent GBD 2018 project shows that AAP and

HAP are associated with over 370,000 premature death globally in under 5 of age, and particulate matter air pollution combined from AAP and HAP ranks as forth leading risk factor of mortality for children under 5 of age (IHME, 2020; Stanaway et al., 2018). In the following sections, I review the current knowledge base and gaps in evidence on of air pollution effects on maternal and child health. I will focus on the current knowledge base on HAP exposure and high blood pressure health outcomes during pregnancy, as well as children's health related to air pollution exposure.

### 2.1.1 Household Air Pollution and Gestational Blood Pressure

Both AAP and HAP exposures have been associated with cardiovascular morbidity and mortality worldwide (Gakidou et al., 2017; Yu et al., 2018). Elevated blood pressure (BP), a risk factor for cardiovascular disease and associated with negative outcomes during pregnancy, has also been shown to be associated with air pollution in the general population (Liang et al., 2014) and in pregnant women (Pedersen et al., 2014).

The biological mechanism of elevated BP after air pollution exposure has been widely investigated in animal and human studies, and has been reviewed by Brook et al (Brook et al., 2010). Generally, after inhalation, $PM_{2.5}$ can penetrate deeply into alveoli region of the lung. It can then 1) stimulate systemic pro-inflammatory responses, 2) alter systemic autonomic nerve system activity and induce imbalance and 3) dissolve parts of fine particulate matter constituents into the circulatory system. All of the above three pathways could lead to short-term and long-term increases in BP (Brook et al., 2010). **Figure 2-1** illustrates the pathways of $PM_{2.5}$ exposure and biological response after exposure.

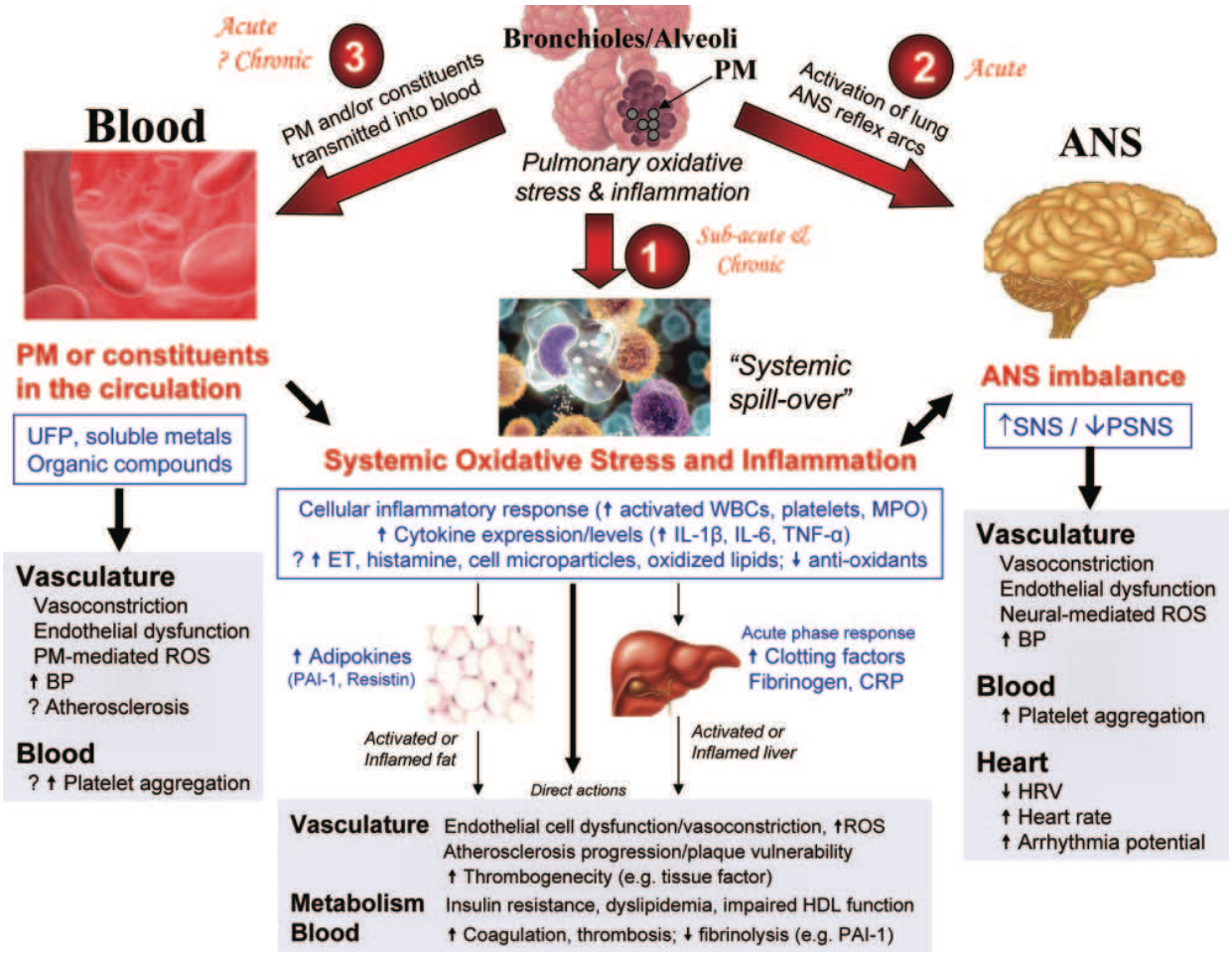



**Figure 2-1.** Biological pathways linking PM exposure with Biological Response, adopted from Brook et al. reference (Brook et al., 2010)

As noted in figure 2, three pathways can link air pollutant exposure and vascular vessel response. Through pathway 1, the BP response to long term air pollution exposure is a sign of vascular injury as cardiovascular system damage caused by oxidative stress, endothelial dysfunction and systemic inflammation, induced long term air pollutant exposure. Another type of BP change (pathway 2) is the reflection of cardiovascular activation induced vasoconstriction through the imbalance of autonomic nerve system in lung caused by exposure of air pollutant. This biological response does not necessarily imply negative biological effect. However, this mechanism may cause complications for population with pre-existing medical conditions, such as heart failure

(Floras, 2009). The third pathway (pathway 3) of BP change is caused by vessel constriction and oxidative stress induced by dissolved constituents of PM in the blood.

A recent systematic review and meta-analysis showed that the use of solid fuel for cooking is associated with increased risks of hypertension (Li et al., 2020). The epidemiological studies on the association between HAP exposure and BP have been mostly focused on non-pregnant women (Alexander et al., 2015; Baumgartner et al., 2011; Clark et al., 2011; Dutta et al., 2011; Fatmi et al., 2019; McCracken et al., 2007; Neupane et al., 2015; Peña et al., 2015; Young et al., 2019); only few focused on this relationship in pregnant women (Agrawal & Yamamoto, 2015; Alexander et al., 2017; Quinn et al., 2016; Thompson et al., 2011) (**Table 2-1**). For non-pregnant populations, epidemiological studies indicated that HAP could induce both short-term and long-term perturbations in BP, after adjusting for potential confounders of this association. However, in pregnant women, there are only a five studies that have investigated effects of HAP exposure on BP level or maternal hypertension during pregnancy (Agrawal & Yamamoto, 2015; Alexander et al., 2017; Quinn et al., 2016; Thompson et al., 2011; Wylie et al., 2015), and the results are not as consistent as those in non-pregnant population. Some studies showed a positive association between air pollution and gestational blood pressure (Alexander et al., 2017; Quinn et al., 2016, 2017) or hypertensive disorder such as preeclampsia (Agrawal & Yamamoto, 2015), others did not find any association (Thompson et al., 2011), and one study found an inverse association (Wylie et al., 2015). This latter finding is consistent with some studies that showed protective effects of smoking on gestational preeclampsia (Glinianaia et al., 2004). Some posited that this association may be from vascular protective effects of CO through lowering sFly1 and sEng production in endothelial and placental cells, thus potentially lowering blood pressure levels (Karumanchi & Levine, 2010). Others also explained that this association is an example of selection bias caused by omitting important confounders associated with both exposure and outcomes and controlling for gestational age at delivery (Luque-Fernandez et al., 2016).

**Table 2-1**. Epidemiological studies of high blood pressure and household air pollution among non-pregnant and pregnant population.

| Study | Population | Study design | Outcome | Exposure | Main Finding, Point estimate (95% CI) | Covariates controlled |
|---|---|---|---|---|---|---|
| Non-pregnant population | | | | | | |
| Guatemala (McCracken et al., 2007) | N = 120, age > 38 years | Randomized control trial (RCT) | Repeated BP measurements in improved biomass stove group vs. open fire group | 24-hour $PM_{2.5}$ monitoring before BP measurements | SBP: 3.1 (0.8, 5.3)<br>DBP: 1.9 (0.4, 3.9) mmHg lower in intervention group | Age, BMI, daily average apparent temperature, rainy season, day of week, time of day, use of a sauna bath, household electricity, asset index, smoking, and SHS. |
| India (Dutta et al., 2011) | N = 480 women, aged 22-41 | Cross-sectional | Three measurements of SBP and DBP with an interval of 24h | Biomass group vs. LPG group<br>Three consecutive days of 8-hour kitchen $PM_{2.5}$ and $PM_{10}$ monitoring at kitchen | Hypertension prevalence 29.5% in biomass group vs. 11% in LPG group<br>OR of hypertension 1.35 (1.14, 1.95) and 1.41 (1.22, 2.08) associated with $PM_{10}$ and $PM_{2.5}$ | None<br>Education, family income, kitchen location |
| Nicaragua (Clark et al., 2011) | N = 124 households cooking with open-fire, with women aged 35 (16) | Cross-sectional | Morning (8am – 12pm) SBP and DBP measurements | 48-hour kitchen $PM_{2.5}$, kitchen CO and personal CO | No statistical significant association being found between BP and air pollution levels<br>SBP: 0.45 (-2.61, 3.52)<br>DBP: -0.12 (-2.15, 1.9) mmHg increase per IQR increase of $PM_{2.5}$ | Age, BMI, SHS, education |
| China (Baumgartner et al., 2011) | N = 280 women aged >24 in rural households using biomass fuel | Cohort study | Repeated BP measurements before and after $PM_{2.5}$ monitoring at participants' home, in summer and winter | 24-hour personal $PM_{2.5}$ in winter and summer | For mixed effect model:<br>SBP: 2.2 (0.8, 3.7)<br>DBP: 0.5 (-0.4, 1.3) mmHg increase per 1 unit increase of log personal $PM_{2.5}$, stronger effects on older women (age > 50) | Age, waist circumference, physical activity, SES, salt intake, day of week, time of day, temperature<br>CME model can account for day-to-day variability in $PM_{2.5}$ exposure better |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | Only statistically significant association being found in older women:<br>SBP: 2.1 (1.9, 2.9)<br>DBP: 2.3 (0.0, 4.7) mmHg increase per 1 unit increase of log personal $PM_{2.5}$ | |
| Peru (Peña et al., 2015) | N = 1004 individual aged 55.3, (47% male, 53% female) | Cross-sectional (baseline of a population cohort study) | BP measured by field workers | Self-report daily biomass fuel use for cooking or heating for >6 months | SBP: 7.0 (4.4, 9.6)<br>DBP: 5.9 (4.2, 7.6) mmHg higher in participants reporting daily use of biomass vs. non-daily use of biomass, effects stronger in men than women | Age, sex, BMI, height, wealth index, education, depressive symptoms, smoking, alcohol abuse, and physical activity |
| Nepal (Neupane et al., 2015) | N = 519 cooks cooking with biogas and firewood age >25 | Cross-sectional with propensity score matching | BP measured at homes | Stove type and fuel type obtained by interview and observation | SBP: -9.84 (-20, 0.8)<br>DBP: -6.49 (-12, -1) mmHg lower in biomass vs. wood-using cooks aged >50. | Biogas and wood-using cooks matched by age, BMI and SES |
| Bolivia (Alexander et al., 2015) | N =28 women participated in cookstove intervention | Cohort study | Repeated BP measurements before and after intervention | 24-hour kitchen level $PM_{2.5}$ concentration before (23 – 750 ug/$m^3$) and after intervention (8 – 170 ug/$m^3$) | SBP: Pre-intervention: 114.5 (13) vs. post-intervention: 109 (10) mmHg.<br>DBP: Pre-intervention: 71.5 (6.3) vs. post-intervention: 70.1 (7.8) mmHg.<br>SBP and DBP correlate with $PM_{2.5}$ concentrations | None |
| India (Norris et al., 2016) | N = 45 women aged 25-66 cooking with biomass | Cohort study | Repeated ambulatory BP measurement during cooking in summer and winter | Black carbon personal exposure during cooking, 24-hour kitchen black carbon measurement | SBP: 6 (-1.7, 2.2) mmHg increase per IQR increase of BC during cooking.<br>DBP: -0.9 (-1.7, -0.1) mmHg decrease per IQR increase of BC during cooking. | None |
| Honduras (Young et al., 2019) | N = 147 women using biomass stove | Cross-sectional | systolic and diastolic blood pressure in sitting position | $PM_{2.5}$ and black carbon (BC) 24-hour average kitchen and personal | SBP: 0.8 (-2.2, 3.8) mmHg increase per natural log-transformed personal $PM_{2.5}$ increase | Age, beds per person, BMI, physical activity |

| | | | | | | |
|---|---|---|---|---|---|---|
| Pakistan (Fatmi et al., 2019) | N = 850 women above 40 years of age | Cross-sectional | Systolic and diastolic blood pressure measurement | Whether currently use biomass for cooking | OR for hypertension (SBP >= 140 mmHg or DBP >= 90 mmHg) is 1.0 (0.8, 1.4) for biomass users | Age and other covariates with p-value less than 0.1 |
| Pregnant women | | | | | | |
| Nigeria (Alexander et al., 2017) | N = 322 pregnant women aged 14-44 participating ethanol cookstove intervention | RCT, longitudinal design | Repeated BP measurements at 20, 26, 30, 34 and 38 week gestational age. | Control (biomass) vs. intervention (ethanol) group | The change in DBP over time was significantly different between intervention and control subjects. Not significant for SBP. | Individual as random intercept, visit time |
| Ghana (Quinn et al., 2016) | N = 817 pregnant women cooking with biomass | Cross-section (baseline of an RCT) | Baseline BP measurements at clinic at 16 (4) week of gestational age | 72-hour CO personal measurement at baseline | SBP: 0.39 (-0.12, 0.9) mmHg increase is associated with 1 ppm increase of CO<br>DBP: 0.43 (-0.01, 0.86) mmHg increase is associated with 1 ppm increase of CO | Age, gestational age, BMI, location of community to main road |
| Guatemala (Thompson et al., 2011) | N = 266 pregnant women cooking with biomass | RCT, cross-sectional design | Baseline BP measurement at clinics | Control (open-fire) vs. intervention (improved biomass stove) group | SBP: 109.4 (10.2) mmHg in control group vs. 107.7(8.2) mmHg in intervention group<br>DBP: 68.3(8.2) mmHg in control group vs. 66.8(7.8) mmHg in intervention group<br>Not statistically significant | None |
| India (Agrawal & Yamamoto, 2015) | N = 39657 pregnant women aged 15-49 | Cross-sectional from NFHS-3 survey | Self-reported experience of preeclampsia or eclampsia during pregnancy | Self-reported type of fuel used, grouped as: high- and medium-exposure vs. low-exposure group | OR = 2.21 (1.26, 3.87) of preeclampsia/eclampsia symptoms for high-and medium-exposure group vs. low-exposure group. | SES, maternal factor, health- and lifestyle factors |
| India (Wylie, 2015) | N = 1369 pregnant women | Cross-sectional in central India | Measured BP at delivery | Self-reported booking fuel, | Adjusted effect size: SBP: -2.0 mmHg (-3.77 - 0.31), DBP: -1,96 (-3.6, -0.30) | History of hypertension, presence of windows, use of smokeless tobacco |

### 2.1.2 Ambient Air Pollution Effects on Child Mortality

Currently, there is no systematic review assessing the effects of gestational and early child life exposure to air pollution on child mortality in LMICs. However, regulatory agencies such as World Health Organizations (WHO) reviewed air pollution effects and recommended including infant mortality in risk assessment of air pollution (WHO Regional Office for Europe, 2013). To better understand the knowledge base of AAP and child mortality, we searched PubMed for relevant articles published in English using the search terms 'outdoor air pollution' OR 'ambient air pollution' OR 'ambient particulate matter' AND 'child mortality' OR 'infant mortality' OR 'child survival' on Oct 30, 2019. This search identified 25 articles. Most of these studies of AAP and child mortality focused on higher income countries. Overall, they suggest short-term (a few days) and long-term (one month or longer) exposure to ambient particulate matter pollution are associated with elevated risk of child mortality. Among these 25 articles, three articles were identified on ambient particulate matter pollution in LMICs (Goyal et al., 2019; Heft-Neal et al., 2018; Wang et al., 2019). There is some evidence that the strength of association differs by locations across LMICs, and the studies in LMICs are less consistent than studies in high income regions. In Sub-Saharan Africa, Heft-Neal and colleagues based their study on nearly 1 million birth records from over 20 years and found a 9% increase in risk of infant mortality associated with a 10 $\mu g/m^3$ increase in annual $PM_{2.5}$ concentration (Heft-Neal et al., 2018). The second study, based on 1236 child deaths in Beijing, is a time-series design but did not find a significant association between current month $PM_{2.5}$ exposure and infant mortality (Wang et al., 2019). A third study, which assessed the early life ambient $PM_{2.5}$ exposure effect on child mortality in 43 LMICs (with half a million child births), found inconclusive association between AAP and neonatal morality as well, with odds ratio of neonatal death 1.08 (0.95, 1.23) per 10 $\mu g/m^3$ increase of annual $PM_{2.5}$ levels (Goyal et al., 2019). The inconsistencies of findings between AAP exposure and child mortality in LMICs could be due to large errors associated with air pollution exposure measurement in each study, which may have biased results towards the null. Therefore, more research is needed to confirm and clarify the association between AAP health effects of child mortality, using the most appropriate methodologies for outcome, exposure assessment, and control of confounders.

## 2.2 Household Air Pollution Exposure Assessment

Assessing exposure to $PM_{2.5}$ is usually done by sampling personal or microenvironmental airborne concentrations for at least 24 or 48 hours, which is used to represent HAP exposure levels. (Balakrishnan, et al., 2015; Dionisio, Howie, et al., 2012; Fitzgerald et al., 2012; Helen et al., 2015; McCracken et al., 2013; Mukhopadhyay et al., 2012; Naeher et al., 2000; Ni et al., 2016; B. J. Wylie et al., 2017). By placing the monitor directly on the subject, personal exposures can be accurately assessed regardless of the individual's movements within or outside the home. Besides personal exposure, other exposure assessment methods have been used in studies to examine the impact of HAP on health, including fuel/stove type, semi-quantitative measures of fuel/stove usage, area/kitchen pollutant concentration and internal biomarker of exposure/health effects (Clark et al., 2013). As summarized in **Figure 2-2,** with the increased of accuracy from using regional fuel use to using biomarkers as assessment of HAP exposure, the cost of the assessment increases as well.

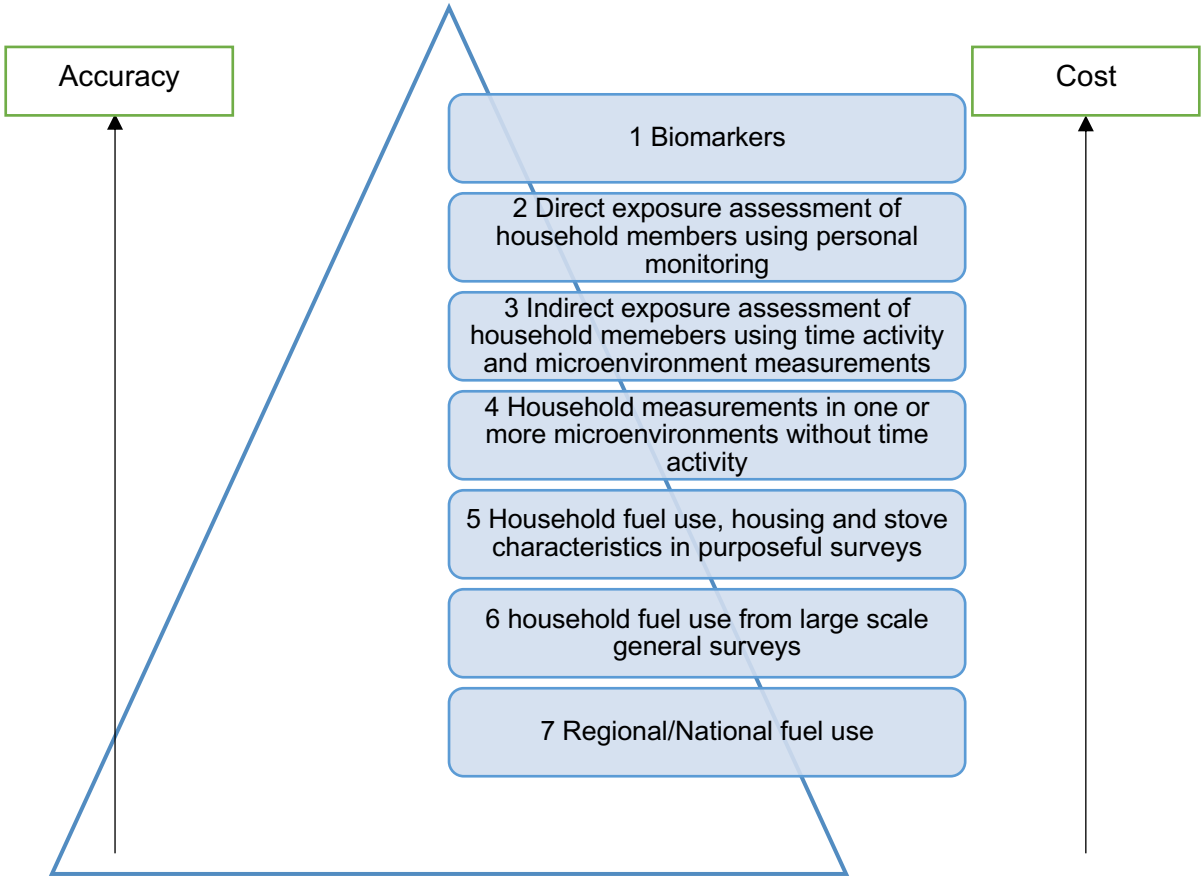


**Figure 2-2**. Schematic for HAP Exposure Assessment Pyramid, adapted from World Health Organization, from reference (World Health Organization, 2005a)

Directly measuring personal $PM_{2.5}$ exposure is challenging, especially among children under 2 years old. Even the latest devices for measuring PM are too large and heavy to be worn by young children who are the most vulnerable to the health risks associated with HAP. Current HAP monitoring devices usually consist of pump, filters and PM size selector such as impactor or cyclone and batteries, and with a size and weight similar to a smart cell phone (Liao et al., 2019; Volckens et al., 2017). These devices are designed to be worn by adults for 24- 48 hours in rural and low-income settings with high exposure level to PM from cookstove. During the time of sampling, these devices are constantly sampling air and collecting PM on a filter used to calculate gravimetric time weighted average concentrations. The size and weight of these device are therefore, too large to be worn by children under 2 year of age.

To estimate child exposure to HAP, many studies sample CO (with small lightweight monitors more easily worn by infants) as a proxy of $PM_{2.5}$ (Dionisio, Howie, Dominici, Fornace, Spengler, Donkor, et al., 2012; Smith et al., 2011a). However, a previous systematic review has shown that CO is not a consistently valid surrogate measurement for $PM_{2.5}$ exposure(Carter et al., 2017), although this finding may have been reinforced by use of poor devices and different sampling strategies. Other studies use less accurate methods of assessing personal exposure by using kitchen area (microenvironmental) $PM_{2.5}$ concentrations (Dionisio, Howie, Dominici, Fornace, Spengler, Adegbola, et al., 2012). However, this method is not very accurate since it does not measure person-time spent outside the kitchen and in other environments (World Health Organization, 2005a). To improve the accuracy of $PM_{2.5}$ exposure measurement for children, there is an increasing need for improved $PM_{2.5}$ monitoring methods that are more likely to capture pollutants to which children are actually exposed as they move about within the home.

Recently, an innovative and low-cost exposure assessment method has been developed using Bluetooth low energy (BLE) beacons to assess the location of a child, coupled with real-time area $PM_{2.5}$ concentration to

estimate personal $PM_{2.5}$ exposure (Piedrahita, 2017). This method will be used in the ongoing HAPIN trial to estimate child HAP exposure for exposure-response analyses of pneumonia, stunting and other child health outcomes. Therefore, validation studies of the indirect exposure assessment method in the field are needed. It is also necessary to determine to what extent the indirect exposure assessment and microenvironmental HAP measurement could represent personal exposure levels. In order to facilitate and guide HAP monitoring and evaluation of future clean stove intervention programs, there is an urgent need to compare different exposure assessment methods to balance both precision and cost. In Chapter 3, I will present a pilot validation study of indirect exposure assessment using BLE beacons to locate participants and assign microenvironmental $PM_{2.5}$ levels to calculate 24-hour average $PM_{2.5}$ exposure.

## 2.3 Ambient Air Pollution Exposure Assessment

The adverse health effects of long-term AAP exposure have been well documented (Brook et al., 2010; Burnett et al., 2018; World Health Organization, 2005b). Most of the epidemiological studies on health effects of $PM_{2.5}$ are based on concentrations monitored by ground-based stations in urban areas. Notably, the Harvard Six City Study and American Cancer Society studies applied centrally located air pollution monitoring stations in communities to represent the exposure to AAP for cohort participants in each city during baseline and follow-up period between 1970s to 1990s (Dockery et al., 1993; C. A. Pope et al., 2002). This exposure assessment approach applied the same AAP exposure level to all participants from the same city/community, thus introducing Berkson's error. Measurement errors of air pollution monitoring stations will also introduce classical errors. In the epidemiological analysis of AAP exposure-outcome analysis, these two types of errors of exposure assessment will lead to errors of effects of air pollution on the outcomes. As shown by Zeger et al., Berkson's error in exposure assessment will not affect the point estimate of the exposure-outcome association, but will introduce larger standard errors around the estimates, while classical errors will introduce attenuation bias and bias the effects towards the null (Zeger S L et al., 2000). In the cohort study measuring long-term air

pollution health effects, such as effects of monthly or annual air pollution exposure on population mortality, the attenuation bias caused by classical errors could be reduced due to longer averaging periods using station monitoring data, thus making the association estimates of air pollution health effects less biased. Another source of AAP exposure error is the lack of indoor sources of air pollution when using centrally-located monitoring stations to infer air pollution exposure. A lack of indoor air pollution source data in the epidemiological analysis of the health effects of AAP exposure is likely to bias the effects toward the null, making the association between air pollution and outcomes more conservative.

With the increased number of ground monitoring stations and advanced air pollution modelling, AAP exposure can be estimated with higher spatial and temporal resolution, such as at the Zip-code level (Di et al., 2017; Pope C. Arden et al., 2019) or 10 km (Yin Peng et al., 2017) or even 1 km spatial grid (Carey et al., 2013). The increased spatial and temporal resolution of AAP will decrease the Berkson's error and increase the precision of the estimate. Recently, more evidence has accumulated and supported air pollution health effects, both for long-term and short-term air pollution exposure health effects (Burnett et al., 2018; Liu et al., 2019). In addition, with the development of satellite remote sensing and chemical transport modelling, temporally and spatially resolved estimations of AAP at global level have been developed for $PM_{2.5}$, $NO_2$ and $O_3$ globally (Achakulwisut et al., 2019; Crouse et al., 2015; Shaddick et al., 2018), and used for Global Burden of Disease (GBD) estimation. These global ambient air pollutant models provide exposure estimation for many areas without air pollutant monitoring stations, and can provide insight for air pollution health effects.

With all these efforts, there are still limitations in assessing AAP exposure in LMICs, especially in developing Asia where air pollution is the highest in the world. The air pollution monitoring stations in China and India are sparse compared to those in developed countries in North America and Europe (Brauer et al., 2019), and the global AAP models have relatively large root mean squared errors (RMSE) in South Asia and East Asia areas (RMSE = 17.6 $\mu g/m^3$ ) (Shaddick et al., 2018). Additionally, validation studies from ground-based monitoring stations show some global models have under-estimated the true AAP levels in South Asia and East

Asia region where monitored $PM_{2.5}$ concentrations are higher than modelled (van Donkelaar et al., 2016; Xiao et al., 2018). With the increase of ground-based monitoring stations in Asia and the development of AAP modeling, such as the machine-learning algorithms, there is a potential to develop more accurate and less biased AAP models in the India region. In chapter 5 of this thesis, I report on an exposure modelling of AAP in India with sub-annual temporal resolution to understand AAP levels in India.

# Chapter 3 **Indirect Personal Household Air Pollution Exposure Assessment Using Bluetooth Low Energy Beacon Systems**

## 3.1 Introduction

Approximately 3 billion people rely on solid fuel for cooking and heating globally due to lack of access to cleaner fuels (Bonjour et al., 2013). According to Global Burden of Disease (GBD), household air pollution (HAP) generated from cooking and heating with biomass stoves is associated with over 1.6 million premature deaths every year, mainly in low- and middle-income countries (LMICs) (Stanaway et al., 2018). Young children and pregnant women are especially at risk from harmful exposure to HAP, since they spend the majority of time indoors. HAP is associated with childhood acute lower respiratory infections (ALRI) (Gurley et al., 2013; Smith et al., 2011a) and low birth weight (Amegah et al., 2014; D. P. Pope et al., 2010b), both of which are the leading causes of death among children under 5 year old in LMICs (Naghavi et al., 2017; Roth et al., 2018). However, HAP mitigation through cleaner cooking interventions, such as improved biomass stoves, has resulted in inconsistent results (Mortimer et al., 2017; Thakur et al., 2018), and many stove intervention studies have failed to sufficiently reduce HAP exposures.

Accurately assessing exposure to HAP – and thus the effectiveness of interventions to mitigate exposure – is challenging. While personal exposure monitors can be used on adults and older children to directly measure $PM_{2.5}$ both gravimetrically and nephelometrically (continuously), even the newer and more compact devices such as the Enhanced Children's MicroPEM (ECM) are too heavy and large to be worn by children under 12 months for periods of 24 hours (Devakumar et al., 2014). When directly estimating personal exposure to $PM_{2.5}$ is not feasible, some studies measure personal exposure to carbon monoxide (CO) with small, lightweight monitors easily worn by infants as a proxy of $PM_{2.5}$ and HAP exposure (Dionisio, Howie, Dominici, Fornace, Spengler, Donkor, et al., 2012; Smith et al., 2011a). However, a systematic review of 61 studies from 27 countries has shown that CO is not always a consistently valid surrogate measurement for $PM_{2.5}$ exposure(Carter et al., 2017). Furthermore, the $PM_{2.5}$ to CO relationship may not be transportable across

different study settings due to heterogeneous stove and fuel types, combustion conditions, and differences in other energy and housing-related factors. A second approach is to rely on kitchen area $PM_{2.5}$ concentration as a proxy for child exposure (Dionisio, Howie, Dominici, Fornace, Spengler, Adegbola, et al., 2012). However, this method does not incorporate exposures during time spent away from the kitchen (World Health Organization, 2005a). Another approach is to conduct an indirect or microenvironmental exposure assessment, which combines conventional pollutant area measurement approaches in various home microenvironments with a time-activity diary or an objective measure of the location of participants in microenvironments (Balakrishnan et al., 2002; M Ezzati et al., 2000; Sidhu et al., 2017; Zuk et al., 2007). However, many of these studies assessed time-location patterns or microenvironmental locations using questionnaires or self-reported diaries, which are prone to recall bias and may not be accurate.

To improve the accuracy of $PM_{2.5}$ exposure measurement, especially in children for whom it may be unfeasible to conduct direct measurements, there is need for more precise, objective and less intrusive indirect $PM_{2.5}$ monitoring methods. Recently, a Bluetooth® Low Energy (BLE) Beacon proximity sensing system, which consists of signal loggers (sensor) and coin-sized signal emitters, was developed to assess the location of children during monitoring (Clark et al., 2013; World Health Organization, 2005a). The application and accuracy of this BLE Beacon system in indirect $PM_{2.5}$ exposure assessment has not been evaluated in field HAP studies. Here, we report on formative research to evaluate an indirect $PM_{2.5}$ exposure assessment method using the BLE Beacon system with participants including women and children enrolled in the Household Air Pollution Intervention Network (HAPIN) trial in rural Guatemala.

## 3.2 Method

### 3.2.1 Purpose and design

This study was conducted as one part of the formative research phase of the HAPIN trial in one of its intervention research centers in Jalapa, Guatemala (Clasen et al., n.d.). This study was designed as a small LPG cookstove intervention, including a 2-month baseline period followed up with a 2-month LPG fuel and

cookstove intervention period. During the 4-month study period, we conducted monthly visits to each household. This study was approved by the institutional review boards of the Universidad del Valle de Guatemala (146-08-2016/11-2016) and Emory University (00089799). The trial is registered at ClinicalTrials.gov (Identifier NCT02944682).

### 3.2.2 Study sites and populations

This study took place between November 2017 and April 2018, in Xalapán area of the Jalapa Department in rural Guatemala, 150 km east of Guatemala City. At an average elevation of 1500 meters, Xalapán has a tropical wet climate with an average temperature of 20 °C. This pilot study was conducted during the dry season with less than 50 mm rainfall per month. We recruited 20 households (1) that relied on woodstoves or open fires for cooking, (2) where a non-smoking woman over the age of 35 years identified as the primary cook, and (3) who had a child aged <1 year. The selection criteria of households is based on the need for testing standardized operating procedures for the main HAPIN trial. Written informed consent was obtained from all participants.

### 3.2.3 $PM_{2.5}$ Measurements

For each household, we conducted four HAP assessments, two before and two after the LPG fuel intervention, for a total of 80 assessments. At each assessment, we measured 24-hour microenvironmental area concentrations (in kitchens, sleeping area, and outdoor patios) and personal $PM_{2.5}$ exposures using the Enhanced Children's MicroPEM (ECM, RTI International, Durham, NC USA), the same device selected for exposure monitoring in the larger HAPIN main trial (Burrowes, n.d.). In kitchen and sleeping area microenvironments, ECMs and personal locating Beacon loggers (more details in section 2.4.1) were placed 1.5 meters above the floor, usually hanging on the wall, 1 meter away from the edge of the combustion source and at least 1 meter away from windows or doors. In the outdoor patio microenvironment, ECMs and Beacon loggers were placed in a secure area 1 to 2 meters above the ground, usually installed under the outside edge of roof, at least 3 meters away from the kitchen and other rooms. Instruments installed in one microenvironment were not visible from the other microenvironment.

ECMs were programmed to sample $PM_{2.5}$ continuously using a nephelometer at a logging rate of 30 seconds and also collected gravimetric $PM_{2.5}$ samples on a 15 mm Teflon filter (PT15-AN-PF02, MTL LLC., Minneapolis, MN USA) at a flow rate of 0.3 L/min. Gravimetric $PM_{2.5}$ measurements made with the ECM have a limit of detection of 5 $\mu g/m^3$ for 24-hour sampling periods. All Teflon filters were pre- and post-weighed in a temperature- and humidity-controlled laboratory at the University of Georgia with temperatures between 20 - 24°C and relative humidity between 30% - 40%. Filters were stored in a -20 °C freezer after sampling in a laboratory at Universidad del Valle de Guatemala, and were transported in double zip-lock bags in coolers with blue ice to the weighing laboratory. We collected 51 duplicate ECM samples (24-hour side-by-side ECM measurements) and 34 field blank filters. In **Figure 3-1S** (appendix)**,** we showed that duplicate ECM samples had good agreement ($R^2 = 0.90$). For all 34 field blanks, net weight changes were less than 5 μg, with a mean of 0.7 (Standard Deviation, SD: 2) μg.

We calibrated all nephelometric continuous $PM_{2.5}$ concentrations with the run-specific 24-hour filter-based $PM_{2.5}$ measurement. First, we calculated a calibration factor for each ECM deployment as the ratio between the 24-hour filter based gravimetric $PM_{2.5}$ concentration and the corresponding 24-hour average nephelometric $PM_{2.5}$ concentration. Then, we multiplied each continuous nephelometric measurement by the calibration factor for each corresponding run to get the gravimetrically-adjusted nephelometric measurements. Finally, we averaged gravimetrically-adjusted nephelometric measurements into 5-minute intervals to reduce variability of the original 30-second measurements. We used the gravimetrically-adjusted continuous nephelometric $PM_{2.5}$ concentrations to reconstruct $PM_{2.5}$ exposures in this study.

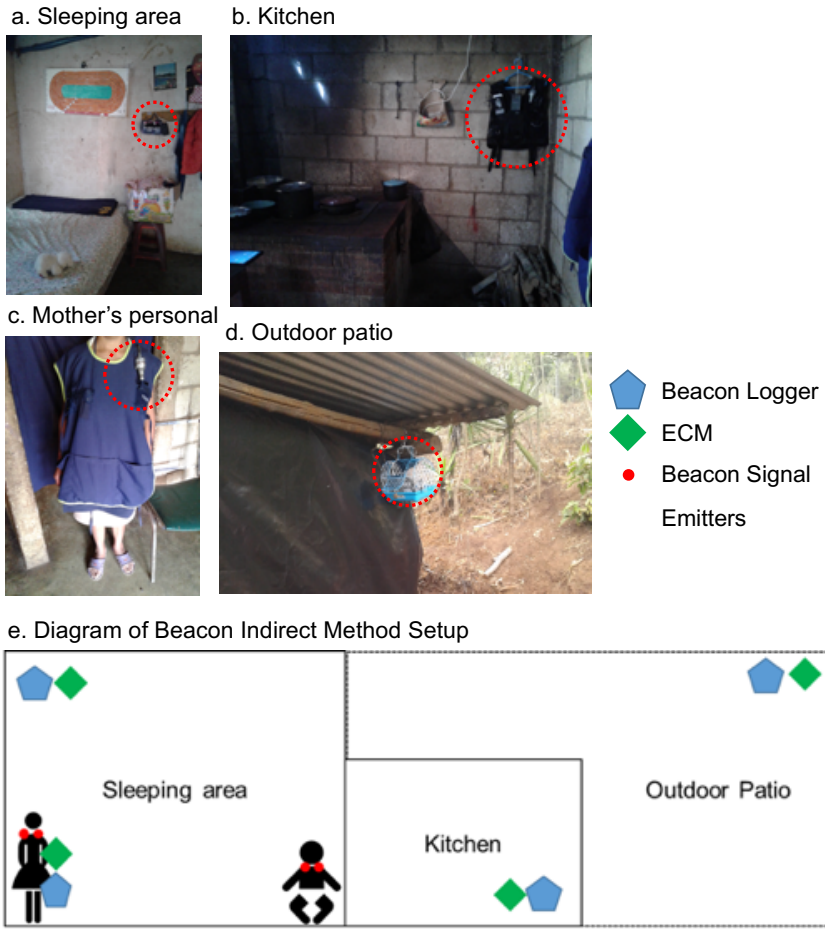



**Figure 3-1**. (a-d): setup of BLE Beacon systems in the sleeping area (a), the kitchen (b), on the patio (d) and on a female participant (c). The dotted red circle in each panel highlights the sampling equipment and Beacon loggers. Panel (e) is a schematic of the Beacon-based indirect exposure assessment method

### 3.2.4 Microenvironment Indirect $PM_{2.5}$ Exposure Measurement Methods

In each microenvironment, we concurrently deployed a fixed-position Beacon logger (Berkeley Air Monitoring Group, California, USA) with the ECM. Beacon loggers receive Bluetooth signals emitted from coin-sized BLE Beacon signal emitters and record the Media Access Control (MAC) address and the Bluetooth Received Signal Strength Indicator (RSSI). Participants wore two emitters each (depicted in **Figure 3-1e**). The RSSI is proportional to the distance between the emitter and the logger and can be used to determine the participants' microenvironmental locations. Signals from the Beacon signal emitter were logged every 20 seconds. We

classified participants' location in five minute intervals as the microenvironment in which the Beacon logger recorded the strongest average RSSI from the two Beacon signal emitters worn by participants.

At the beginning of each deployment, we carried out a 6 to 15 minute long walk-through procedure to assess the accuracy of the Beacon system's location prediction. During the walk-through procedure, field workers wore sampling vests containing all BLE Beacon signal emitters and walked through each microenvironment for 5 minutes, where Beacon loggers have been installed. The start and end times in each microenvironment were recorded and regarded as the 'gold-standard' of microenvironment location classification during the walk-through procedure. We defined the accurate prediction rate of microenvironmental location during the walk-through as the percentage of time when field workers are classified in the same microenvironment as recorded manually. In **Figure 3-2S** (appendix)**,** we show that the correct microenvironmental classification rate increases over time. During initial deployments of the system, due to suboptimal placement of sensors and system failures of loggers, the correct prediction rate of microenvironment was 40% - 50%. At the end of this pilot study, the system was able to classify the microenvironment correctly at an average rate over 85% during walk-throughs. **Equation 3-1** defines the indirect exposure (IE) estimate. IE is the time-weighted average of $PM_{2.5}$ concentrations in microenvironments where participants spend time as classified by the Beacon systems.

$$IE = \frac{\sum_t \sum_j \left(C_{t,j} L_{t,j} \Delta T\right)}{\sum_t \sum_j \left(L_{t,j} \Delta T\right)} = \sum_j IE_j$$

Equation 3-1

$IE$ refers to the total time-weighted average indirect exposure assessment, $IE_j$ refers to the contribution of $PM_{2.5}$ exposure in each microenvironment $j$ to the total time-weighted average indirect exposure. $C_{t,j}$ is the gravimetrically-corrected nephelometric $PM_{2.5}$ concentrations logged by an ECM at time *t* in microenvironment *j*. $L_{t,j}$ is the indicator of the participant's location by the BLE Beacon systems at time *t*, in microenvironment *j*. Specifically, $L_{t,j} = 1$ if the participant is classified in microenvironment *j* at time *t*, otherwise $L_{t,j} = 0$.

Notably, if none of Beacon loggers received Bluetooth signals from Beacon emitters, we classify participants as outside of households, and will not have indirect $PM_{2.5}$ measurements during that period of time. $\Delta T$ refers to the sampling interval, in this case 5 minutes. In **Figure 3-3S** (appendix)**,** we show an example of a time-series plot of RSSI and microenvironmental location classification for one measurement. In **Figure 3-4S** (appendix)**,** we show a time-series plot of indirect exposure and direct personal exposure from the same participant during the same measurement period.

#### 3.2.4.1 Indirect $PM_{2.5}$ exposure for Women

In each household, the primary women cook wore two BLE Beacon signal emitters on their sampling vests along with ECMs to measure their direct personal exposure. Beacon loggers were placed together with ECMs in three microenvironments: kitchen, sleeping area, and outdoor patio. Women's indirect exposure is estimated using gravimetrically-corrected nephelometric $PM_{2.5}$ concentrations from the three fixed microenvironments, when women are classified in the given microenvironment by BLE Beacon systems (Figure 1, a, b, d). Sixty-two (77%) of 80 indirect exposure assessments were valid for women; 18 (23%) measurements were removed due to system failures or suboptimal placement of Beacon loggers (19%) and system failures of ECMs (4%).

#### 3.2.4.2 Indirect $PM_{2.5}$ exposure for Children

In each household, we deployed two Beacon signal emitters on the clothing of each child under 1 year of age and assessed their microenvironmental locations longitudinally. Children's indirect $PM_{2.5}$ exposure is estimated using the gravimetrically corrected nephelometric $PM_{2.5}$ concentrations from the three fixed microenvironment locations (kitchen, sleeping area and outdoor patio) and the women's personal direct microenvironment, when children were classified in the given microenvironment by BLE Beacon systems (Figure 1, a-d). Particularly, we added one microenvironment of women for assessing children's indirect exposure. Direct personal $PM_{2.5}$ measurements from ECM monitors worn by women were used when children were classified in the women's microenvironment. Sixty-one (76%) of 80 indirect exposure assessments were valid for children; 18 (24%)

measurements were removed due to system failures or suboptimal placement of Beacon loggers (19%) or system failures of ECMs (5%).

#### 3.2.5 Statistical Methods

Descriptive statistics, including the arithmetic mean, standard deviation (SD), median and interquartile range (IQR) for 24-hour $PM_{2.5}$ concentrations from area and women's direct (personal) exposure samples were calculated. We reported both mean (SD) and median (IQR) statistics because 24-hour $PM_{2.5}$ concentrations and exposures are not normally distributed (right-skewed). Second, descriptive statistics (mean and SD) for women's and children's time spent in each microenvironment predicted by the Beacon system were calculated. We estimated women and children's indirect $PM_{2.5}$ exposure and calculated descriptive statistics and estimated the mean contribution to indirect $PM_{2.5}$ exposure from each microenvironment. To evaluate the performance of the Beacon-derived indirect exposure methods, we compared women's direct (personal) exposure measurements with indirect measurements and calculated Spearman correlation coefficients. We created Bland-Altman plots to evaluate agreement between direct personal exposure, indirect exposure, and kitchen measurements. We calculated the root mean squared error (RMSE) of indirect exposure estimates and kitchen area $PM_{2.5}$ concentrations compared to direct personal exposure measurements, respectively. Bias was calculated separately as the mean difference of direct personal and indirect measures and the mean difference of the direct personal and kitchen paired $PM_{2.5}$ concentrations, respectively. Data analysis was conducted in R (version 3.5.0, the R foundation).

## 3.3 Results

Among twenty household in this study, most (n = 17, 85%) had a fully enclosed kitchen with a roof and four walls. The walls of households were made of bricks and roofs were made of wood or corrugated metal. The average size of an enclosed kitchen was 14.2 $m^2$, with an average of height of 2.5 m. The kitchens were potentially well ventilated in the households, with an average of 11 windows or apertures. **Table 3-1** shows 24-hour area $PM_{2.5}$ concentrations during the pre-LPG baseline measurements and the post-LPG follow-up period.

We observed high 24-hour area $PM_{2.5}$ concentrations during baseline measures compared to the follow-up period. We found 94%, 79%, and 62% reductions in 24-hour $PM_{2.5}$ levels in the kitchen, sleeping area, and outdoor patio area microenvironments.

**Table 3-1**. Area 24-hour $PM_{2.5}$ Concentration, mean (SD), median (IQR) , unit: $\mu g/m^3$

| | Baseline | | Follow-up | |
|---|---|---|---|---|
| | Mean (SD) | Median (IQR) | Mean (SD) | Median (IQR) |
| Kitchen | 397 (301) | 308 (227) | 21 (14) | 17 (22) |
| Sleeping area | 113 (172) | 34 (101) | 23 (13) | 40 (37) |
| Outdoor patio | 58 (78) | 34 (32) | 22 (18) | 20 (24) |

**Figure 3-2** shows the average estimated hours (over a 24-hour period) that women and children spent in each microenvironment, as well as time outside of the household in the pre- and post-LPG intervention periods. Women spent 12.8 hours in the sleeping area, 6.2 hours in kitchen and 3.5 hours in the outdoor patio. Children spent 11.3 hours with their mothers, 8.2 hours in bedroom and 2 hours in the outdoor patio microenvironment. Women and children spent 0.9 hour outside of the monitored household microenvironments on average. We found that the LPG intervention was not associated with women's time in any of the three microenvironments and was only statistically significantly associated with children's time in the sleeping area (Two-sided t-test $p = 0.01$).

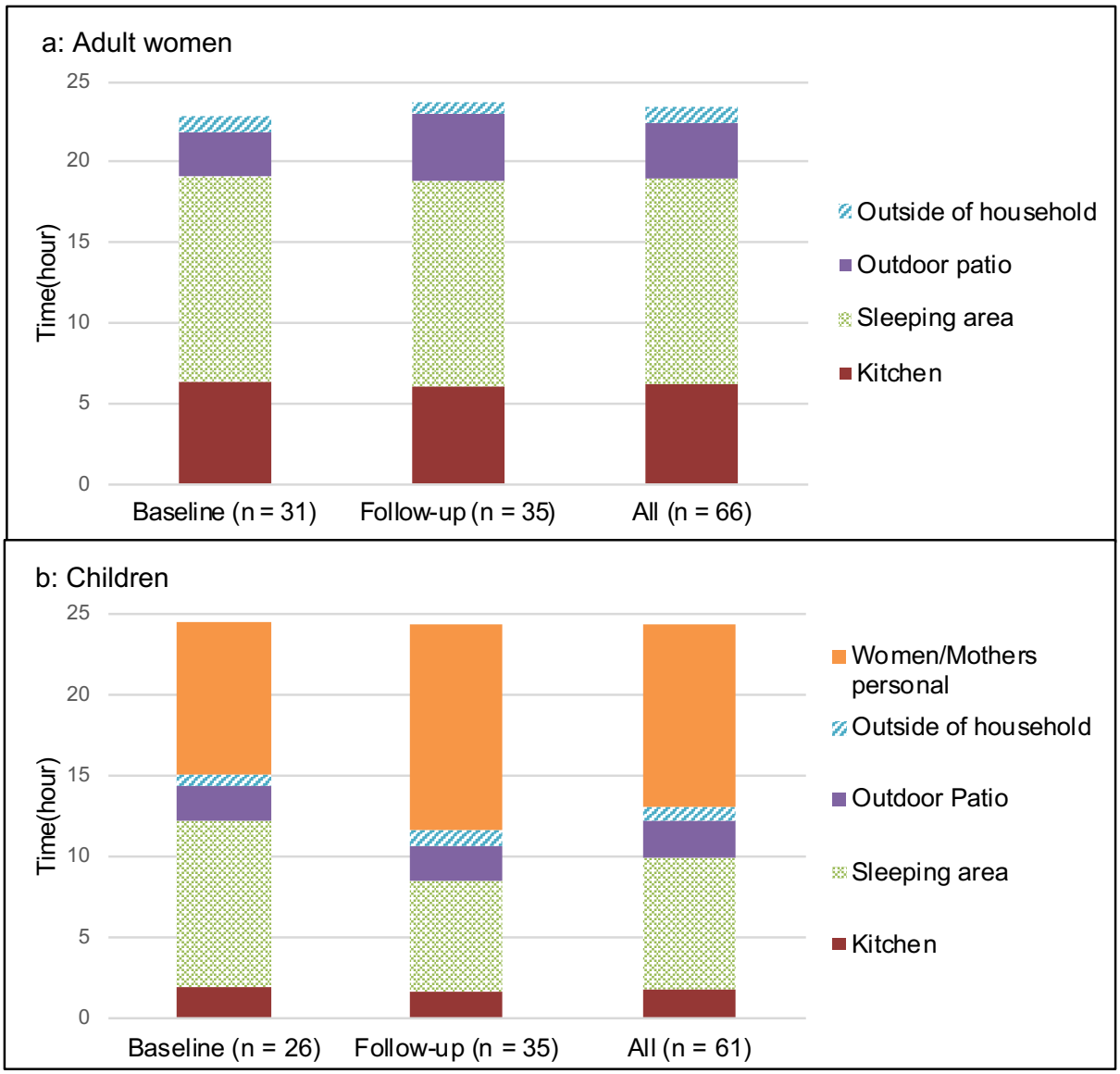



**Figure 3-2**. Daily average time (hour) spent in each microenvironment for women (a) and children (b).

Women participants reported high compliance of wearing sampling vest. The average time not wearing sampling equipment aside from sleeping and bathing was 1.1 hour. **Table 3-2** lists the mean and median 24-hour women's direct exposures and indirect $PM_{2.5}$ exposure reconstructions in pre- and post-LPG periods, along with Spearman correlation coefficients between the indirect and direct measurements. The means of direct and indirect $PM_{2.5}$ exposure are 189 (SD: 138) μg/m$^3$, and 258 (SD: 194) μg/m$^3$, respectively, both of which are well above World Health Organization (WHO) Interim Target 1 guideline of 35 μg/m$^3$. We found a

75% and 91% reduction in direct and indirect 24-hour mean $PM_{2.5}$ exposures after LPG intervention, respectively.

**Table 3-2**. Direct and indirect $PM_{2.5}$ exposure for women, unit: $\mu g/m^3$

| | Baseline n =27 | Follow-up n = 35 | Overall n = 62 |
|---|---|---|---|
| Direct personal $PM_{2.5}$ exposure mean(SD), median (IQR) | 189 (138), 119 (164) | 47 (29), 42 (31) | 109 (116), 66 (79) |
| Indirect $PM_{2.5}$ exposure mean(SD), median (IQR) | 258 (194), 188 (214) | 23 (13), 21 (21) | 125 (172), 39 (135) |
| Spearman correlation coefficient between women' direct and indirect $PM_{2.5}$ measure | 0.63 | 0.66 | 0.81 |

SD: standard deviation, IQR: interquartile range

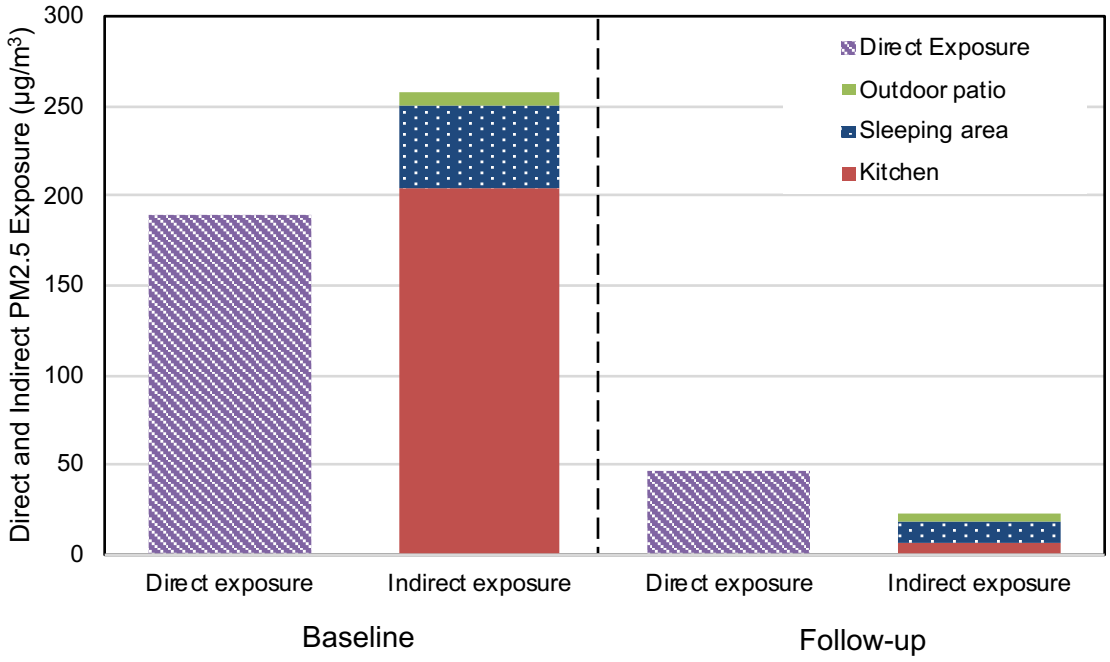



**Figure 3-3**. Direct and indirect $PM_{2.5}$ exposures for women and the contribution of indirect exposure from each microenvironment (kitchen, sleeping area and outdoor patio)

Indirect measures of $PM_{2.5}$ are highly correlated with direct personal measures for women, with a Spearman correlation of 0.81 (**Figure 3-5S** appendix). **Figure 3-3** shows the mean of women's direct $PM_{2.5}$ exposure and indirect $PM_{2.5}$ exposure and the contribution of each microenvironmental $PM_{2.5}$ measurement to the indirect

$PM_{2.5}$ exposure estimates. In the baseline period, indirect exposure estimates were higher than the direct exposure measurements, and $PM_{2.5}$ exposures from the kitchen microenvironment contributed most strongly to the average indirect exposure. In the post-LPG period, direct exposures were higher than indirect exposures and the sleeping area contributed most of indirect exposure for women.

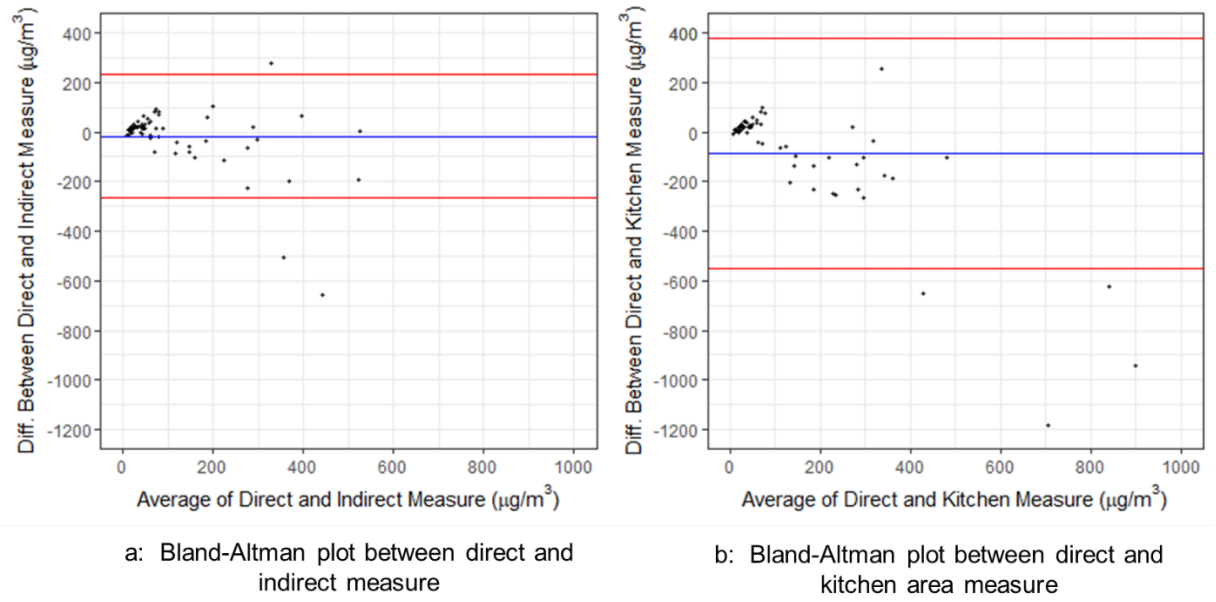



**Figure 3-4**. Bland-Altman Plot of Women's 24-hour Direct and Indirect (a) and Women's Direct and Kitchen Area $PM_{2.5}$ Measure (b)

**Figure 3-4** shows the Bland-Altman plot of 24-hour direct versus indirect $PM_{2.5}$ measurements (left panel) and direct versus kitchen $PM_{2.5}$ measurement (right panel) for women. The x-axis of the plot is the average of two measurements, and the y-axis is the 24-hour direct measurement minus indirect measurement (left panel) or 24-hour direct measurement minus kitchen measurement (right panel), respectively. The blue line is the mean of the measurement differences (y-axis value) and two red lines are 95% confidence interval of the measurement differences. The left panel (a) of Figure 3 shows a smaller difference between two measurements and dots are less deviated from the blue centerline, compared to the right panel (b). Indirect measurements have less bias and have better agreement with direct personal measurement when compared with kitchen measurements (**Figure 3-4**). **Table 3-3** shows the root mean squared error (RMSE) and bias of direct-indirect $PM_{2.5}$ exposure

pairs and direct-kitchen $PM_{2.5}$ concentration pairs by LPG intervention period. When compared to women's direct $PM_{2.5}$ exposure, the RMSE of the women's indirect $PM_{2.5}$ exposure was 128 μg/m³ and the RMSE of kitchen $PM_{2.5}$ concentration was 250 μg/m³. The average bias between direct-indirect $PM_{2.5}$ exposure was -17 μg/m³ (indicating overestimation of the indirect method), and average bias between direct-kitchen $PM_{2.5}$ was -89 μg/m³. Most of the error and bias come from the pre-LPG intervention baseline phase, as indirect exposure and kitchen area measurement overestimated direct personal $PM_{2.5}$ exposure levels (**Table 3-3**).

**Table 3-3**. RMSE and bias between direct-indirect and direct-kitchen paired $PM_{2.5}$, unit: μg/m³

| | | Baseline n =27 | Follow-up n = 35 | Overall n = 62 |
|---|---|---|---|---|
| RMSE | Direct-indirect | 189 | 34 | 128 |
| | Direct-kitchen | 377 | 35 | 250 |
| Bias | Direct-indirect | -70 | 24 | -17 |
| | Direct-kitchen | -230 | 26 | -89 |

RMSE: root mean squared error

**Table 3-4**. Children's Indirect $PM_{2.5}$ Exposure

| Indirect $PM_{2.5}$ exposure estimate | Baseline n = 26 | Follow-up n = 35 | Overall n = 61 |
|---|---|---|---|
| Mean (SD) | 175 (125) | 39 (26) | 97 (107) |
| Median (IQR) | 141 (160) | 35 (30) | 51 (90) |

Children showed high compliance of wearing Beacons. Women participants reported their children had not worn Beacons only for an average of 0.2 hour among 80 measurements, aside from sleeping and bathing. **Table 3-4** lists the mean (SD) and median (IQR) of indirect $PM_{2.5}$ exposures for children by intervention period. We found that children's indirect $PM_{2.5}$ exposure was reduced by 77%, from a mean of 175 (SD, 123) μg/m³ to 39 (SD, 26) μg/m³ after LPG intervention. In **Figure 3-5,** we show the mean of children's indirect $PM_{2.5}$ exposure and the contribution of each microenvironment to indirect $PM_{2.5}$ exposure. In the pre-LPG period, the women/mothers' personal 'microenvironment' contributed most strongly, followed by $PM_{2.5}$ in the kitchen

microenvironment. In the post-LPG period, women/mother's personal microenvironment contributed most to the indirect $PM_{2.5}$ exposure.

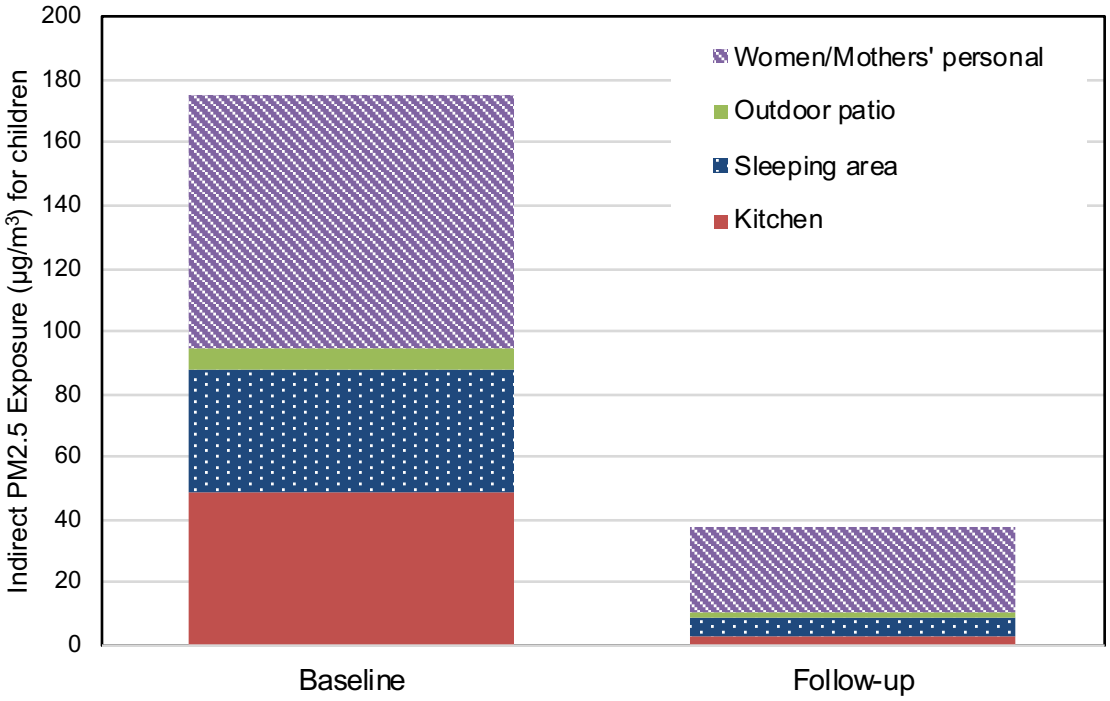



**Figure 3-5**. Indirect $PM_{2.5}$ exposure for children and contribution of indirect exposure from microenvironment locations (kitchen, sleeping area, outdoor patio and women/mothers' personal direct microenvironment)

## 3.4 Discussion

In this study, we demonstrated the feasibility of objectively monitoring the location of participants including adult women and children in their homes using a BLE Beacon proximity sensing system. This system, when combined with ECM $PM_{2.5}$ monitors placed in microenvironments throughout the home, enabled reconstructions of personal exposures that were highly correlated with direct measurements of $PM_{2.5}$ exposure. The same system enabled accurate prediction of the location of children under 1 year of age and enabled reconstructions of their exposure to $PM_{2.5}$ over 24 hour periods.

To our knowledge, this is the first study evaluating indirect exposure to $PM_{2.5}$ using personal locating technology and microenvironment $PM_{2.5}$ monitors in HAP field studies. Previous studies mainly applied time-activity questionnaires or diaries as self-reported records of microenvironmental location (Balakrishnan et al., 2002; Devakumar et al., 2014; M Ezzati et al., 2000; Sidhu et al., 2017; Zuk et al., 2007). A few studies have applied an objective personal locator for time-location assessment in similar settings in Guatemala; those studies relied on an ultrasound emitter and detector to provide a binary presence or absence in a specific microenvironment (Allen-Piccolo et al., 2009; Ruiz-Mercado et al., 2010). Most of the previous studies using indirect exposure approaches did not validate the accuracy of the time-location patterns reported by participants. We conducted walk-through tests by comparing records from field workers (our gold standard) with locations determined by the Beacon logger, and found Beacon systems could accurately predict location 89% of the time on average. This finding of high microenvironment predicting accuracy of the Beacon system is consistent with a previous study that utilized ultrasound personal locator devices (Allen-Piccolo et al., 2009).

For adult women, the Beacon system indicated that they spent half of their time in sleeping area (12.8 hours per day), followed by kitchen (6.2 hours per day) and outdoor patio (3.5 hours per day) microenvironments, and 0.9 hour out of any of these microenvironments. These findings are similar to studies conducted in India (Balakrishnan et al., 2002; Sidhu et al., 2017), Kenya (M Ezzati et al., 2000) and Mexico (Zuk et al., 2007), all of which found women cooks spend around 12 hours per day in the living room or sleeping room, followed by 4-7 hours per day in the kitchen. Notably, we found that time-activity patterns did not seem to change between pre- and post-LPG periods for women. This is consistent with the findings of Zuk et al., who did not find a change in time-activity patterns from an improved biomass stove intervention in rural Mexico (Zuk et al., 2007). For children under 1 year old, we found that they spent most of the time with mothers or in the sleeping area. Notably, we classified children into women/mothers' microenvironment if they were close together, even if they are in the kitchen or sleeping areas. Our findings are consistent with findings from older children in Nepal (Devakumar et al., 2014) and Kenya (M Ezzati et al., 2000), where children spent 12.2 hours per day and 44% of their time in the living room or sleeping area, respectively. Interestingly, we found that the LPG fuel

intervention increased the time children spent with their mothers. However, since our study did not collect self-reported time-activity diaries from participants and due to a relatively limited number of samples, more studies are needed to confirm the effect of LPG interventions on time-activity patterns. In addition, we found that women were not in any of the measured microenvironments for, on average, 1 hour per day. During these periods, no indirect measurement of exposure to $PM_{2.5}$ was captured. This is possibly due to some participants leaving their households during the day to visit friends or relatives, or to go shopping, and also due to a few participants who went to another home to sleep at night. We still included these households in our evaluation of indirect exposure assessment of the women, because we believe these indirect exposure measurements, even lacking a few hours of data, are still useful for predicting daily exposure levels. Sensitivity analysis excluding measurements with more than 4 hours outside of households (n = 3) shown that the time spent and indirect exposure changed less than 10% compared to original results.

Our study illustrates that indirect $PM_{2.5}$ exposure estimates derived from the Beacon system showed a stronger correlation with direct measurements of $PM_{2.5}$ personal exposure ($\rho$=0.81), than did correlations between kitchen microenvironment $PM_{2.5}$ levels and direct personal measurements of $PM_{2.5}$ exposure levels ($\rho$=0.68). As shown in the Bland-Altman plot (**Figure 3-4**), indirect exposure measurements tended to have less bias and agree better with direct personal exposure than kitchen area $PM_{2.5}$ measurements. Therefore, the Beacon indirect exposure method described here better estimates exposures than does simply using area measurements as a proxy for exposure, a common, but perhaps inaccurate, method used to estimate $PM_{2.5}$ exposures for infants(Balakrishnan et al., 2004; Balakrishnan, Ghosh, et al., 2018b). Our findings confirm other recent data from HAPIN formative research indicating that the LPG intervention can reduce $PM_{2.5}$ levels close to the WHO target of 35 $\mu g/m^3$ (Johnson et al., 2018). Prior estimates of an LPG intervention effect were around 70 $\mu g/m^3$ (Steenland et al., 2018). Despite the fact that we provided a 3-month supply of free LPG gas cylinders, it is likely that some continued used of biomass fuel (stove-stacking) and air pollution from neighboring households increased $PM_{2.5}$ exposure above what we would have observed with only gas fuel use.

The new indirect microenvironment exposure approach in our study has a number of advantages over typical indirect exposure assessment. First, we applied a Beacon proximity sensor system, an objective personal locating system to assess microenvironmental locations of participants. This approach can reduce error and recall bias from self-reported time-activity data. Second, we used gravimetrically-corrected continuous microenvironmental $PM_{2.5}$ concentrations to reconstruct indirect exposures. Compared to other similar studies using time-activity patterns or microenvironmental approaches (Devakumar et al., 2014; Zuk et al., 2007), our study has the advantage to capture temporal variation and peaks of $PM_{2.5}$ for indirect exposure.

We also found that indirect exposure estimation from the Beacon system has some limitations and biases. We expect two types of bias would emerge from indirect exposure assessment compared to direct personal exposure. One type of bias is that the indirect method is not able to capture all of the microenvironments participants move through and could mischaracterize locations of participants. Another type of bias emerges when area $PM_{2.5}$ measures differed from true personal direct $PM_{2.5}$ measures, which reflects differences between area ECM $PM_{2.5}$ monitors and personal monitors when participants' locations are known. **Figure 3-4** illustrated heteroscedasticity using indirect exposure to predict indirect exposure, indicating error of indirect exposure increases as $PM_{2.5}$ level increases. **Table 3-4** showed that in the pre-LPG baseline period, indirect exposure overestimated direct exposure but in the post-LPG follow-up period, indirect exposure tended to underestimate direct exposure. The overestimation at baseline may be due to differences between personal monitors and area monitors in households cooking with biomass stoves/open fires, with area monitors being closer to the open fire. The underestimation of indirect exposures in post-LPG follow-up periods may be due to the existence of other sources of air pollution, which is captured by the personal monitor but not necessarily by area monitors, and may have a greater relative importance when kitchen measurements have been sharply lowered. We show in the supplementary materials that compared to personal direct exposure measurement (gold standard), the Beacon indirect method will likely over-estimate personal exposure levels in biomass households and likely

under-estimate personal exposure levels in LPG intervention households, which is in fact what we have observed.

It is also worth noting that we have relatively high failure rates for the Beacon system (19%, 15 measurements out of 80 measurements), mainly due to incorrect set up of 10 (13%) Beacon loggers leading to failures of Beacon logger systems, and 5 (6%) Beacon logger misplacement in outdoor patio areas, leading to the obstruction of Beacon signals. However, we found these failures occurred mainly in the beginning phase of this study and could be largely prevented if additional training of field workers was conducted to ensure proper set up of Beacon loggers. Despite these limitations, our study still showed that the combination of the Beacon system and ECM monitors is a precise and feasible indirect method to assess exposure to $PM_{2.5}$ in low-and-middle income settings for children, especially when direct personal exposure measurement is not practical.

## 3.5 Conclusion

In conclusion, we provided assessment of an indirect, sensor-enabled exposure measurement technique in households using woodstoves at baseline and an LPG cookstove at follow-up. This information adds evidence that indirect exposure assessment using the Beacon system as a microenvironmental location monitor provides an acceptable estimate of personal exposures in both low and high $PM_{2.5}$ exposure settings. We found that indirect exposure methods have higher correlation with direct personal exposure measurements and less bias than do kitchen measurements. In settings where conducting personal direct exposure assessment is not practical, such as in children under 1 year old, the Beacon indirect exposure method is an alternative to better estimate personal exposure to $PM_{2.5}$. The results of this study can inform exposure assessments for future household air pollution studies.

Appendix for Chapter 3

## Appendix Figures and Tables

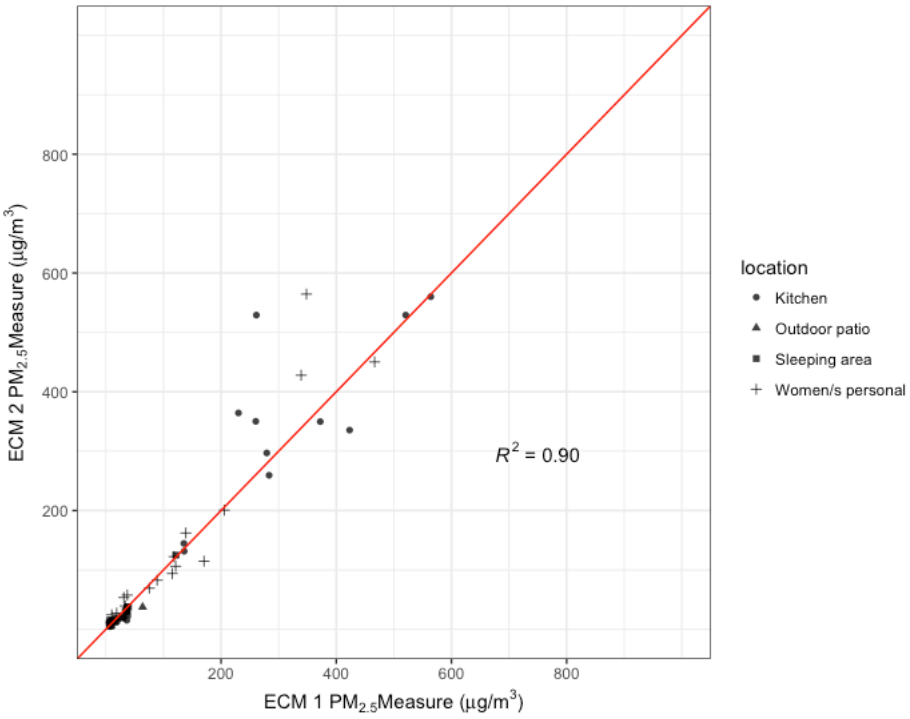



**Figure 3-1S.** Scatter Plot of Duplicate ECM Measure (N = 51 ECM pairs, sleeping area (SAP): 14, mother personal (PEO): 16, kitchen (KAP1): 19, outdoor patio (HOP): 2)

**Figure 3-2S** shows microenvironmental locations prediction correct rate by the Beacon system during the walk-through test, during which we have recorded "true" microenvironment locations of field workers who wear all Beacon signal-emitters. This included microenvironmental locations predicted by malfunctioning Beacon systems. The microenvironment location prediction correct rate increased from 50% to 80% by the end of the study. This is due to troubleshooting of failures of Beacon loggers and the improvement of Beacon logger placement.

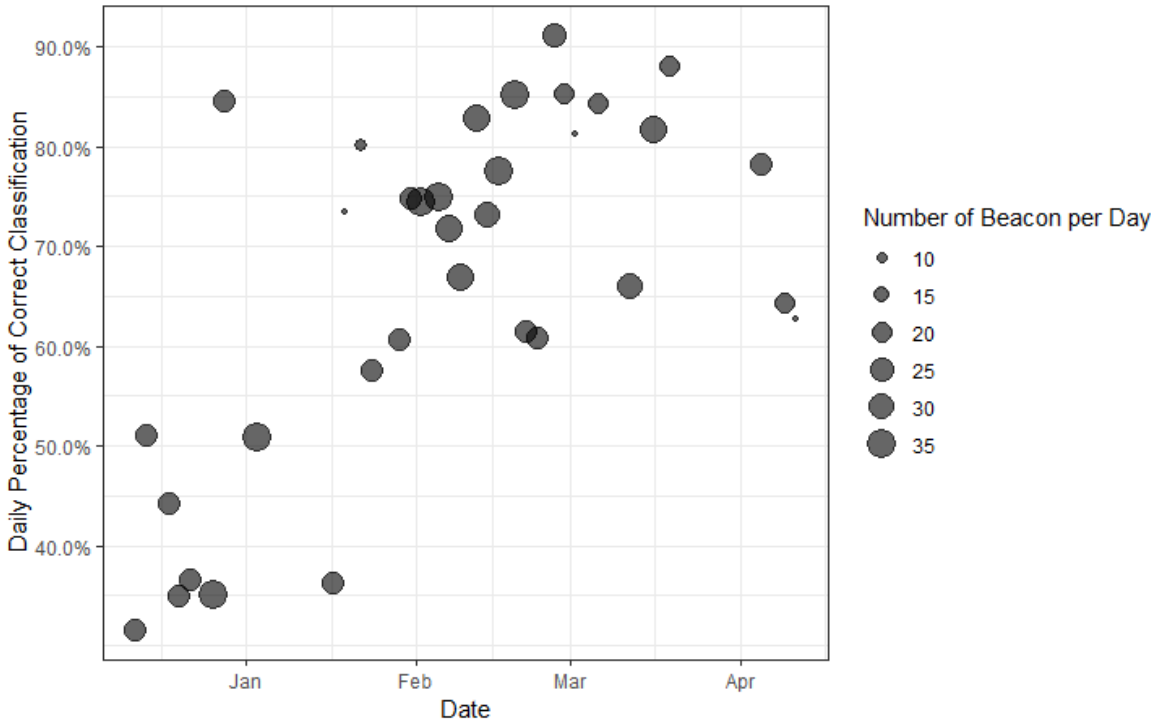



**Figure 3-2S**. Beacon System Microenvironment Location Classification Rate over Time During Walk-through Test. The x axis is calendar date; the y axis is the average percent of correct microenvironmental classifications per day. The size of the dot is the number of Beacons utilized to generate the datapoint,

**Figure 3-3S** shows a time-series plot of Received Bluetooth Signal Strength Indicator (RSSI) and mothers' microenvironment locations from one deployment of the system with receivers in the kitchen, on the outdoor patio and in the sleeping room. Participants are assigned to microenvironments where the Beacon logger received the strongest signal, that is, the largest RSSI over a 5 minute interval. The larger the RSSI is, the closer participants are to the Beacon logger, and thus the participant (who wears emitting devices) is classified in the microenvironment locations with the largest RSSI over the five minute averaging period.

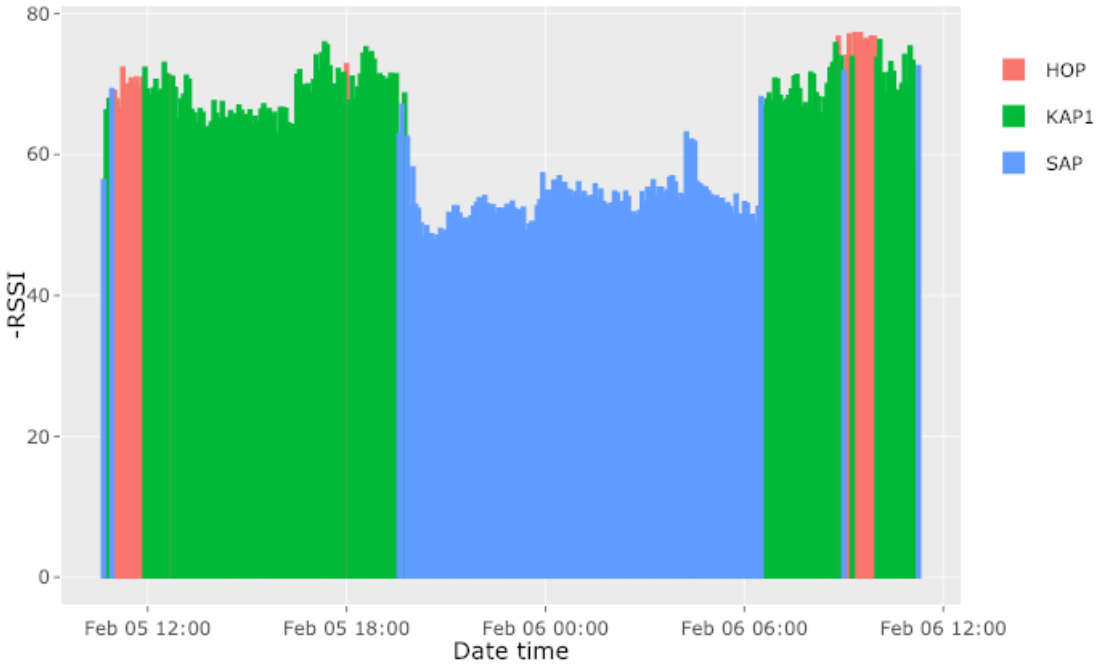



**Figure 3-3S**. Time-series plot of RSSI and microenvironment location classification (HOP: Outdoor patio, KAP1: kitchen, SAP: sleeping area).

**Figure 3-4S** shows the time-series plot of direct and indirect exposures from the same measurement as **Figure 3S**. The bar on the bottom of the plot shows the microenvironment locations (kitchen, sleeping area and outdoor patio) as classified by BLE Beacon proximity system.

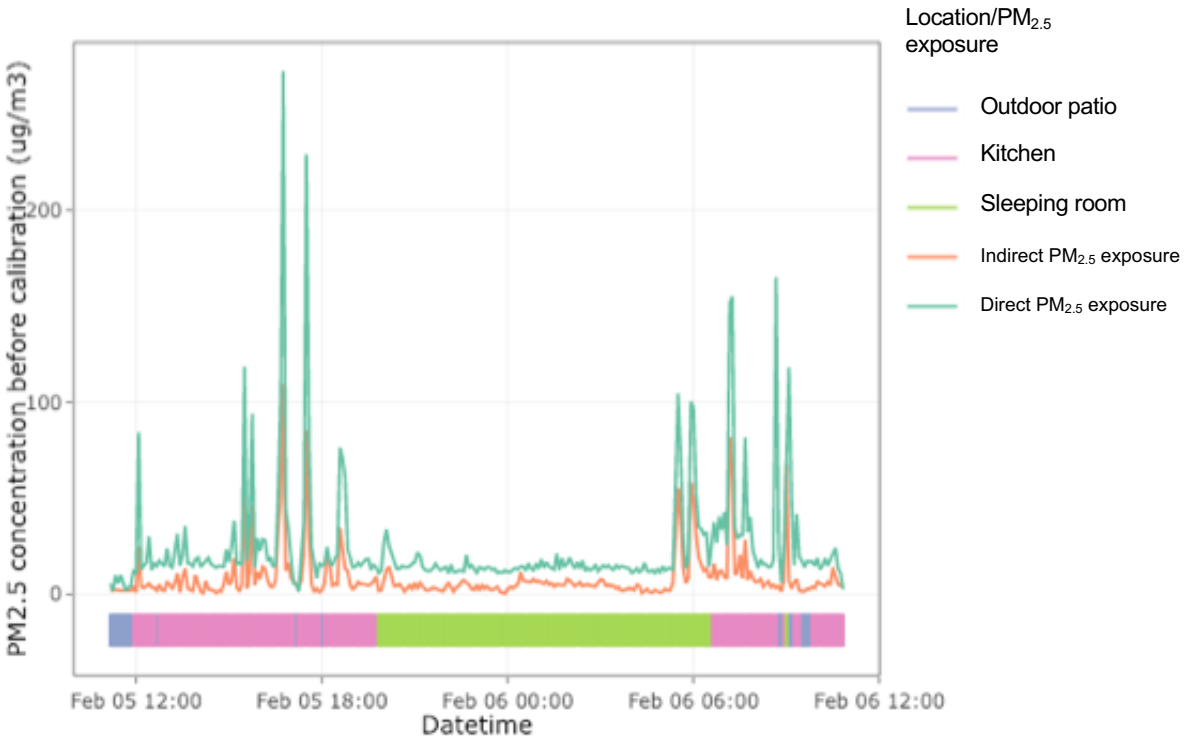



**Figure 3-4S**. Time-series plot of direct and indirect exposure as well as microenvironment classification

**Figure 3-5S** shows the scatter plot of women's direct and indirect $PM_{2.5}$ measure after gravimetric calibration, by each phase. Two measurements in the baselines in the lower right corner of the plot are the most influential in lowering the R-square.

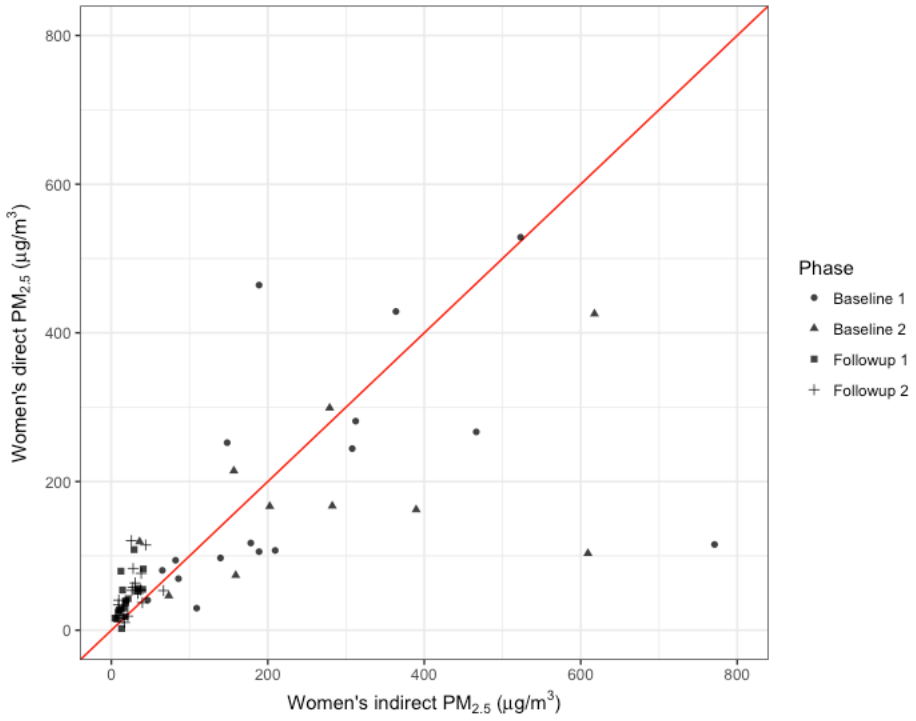



**Figure 3-5S**. Scatter Plot of 24-hour Direct and Indirect $PM_{2.5}$ Measure for Women (BL1/BL2: pre-LPG intervention baseline measurements; FU1/FU2: post-LPG intervention follow-up measurements, solid line is the 1:1 line)

**Bias Estimate for Beacon Indirect Assessment Methods**

True personal exposure metrics integrate exposure from personal $PM_{2.5}$ levels in all microenvironments as well as unobserved microenvironments, such as time spent outside of households. Beacon indirect exposure metrics, however, only accounts for microenvironment exposure levels when participants are classified in the microenvironment locations. Below two equations expressed true personal $PM_{2.5}$ exposure (equ-3-1S) and measured Beacon indirect $PM_{2.5}$ exposure (equ-3-2S).

$$PM_{direct} = \alpha_{micro} PM_{micro} + \alpha_{out} PM_{out}$$

Equ-3-1S

$$PM_{indirect} = PM_{micro}(1 + \tau \times Fuel)$$

Equ-3-2S

$PM_{direct}$: direct personal $PM_{2.5}$ exposure from a personal $PM_{2.5}$ monitor

$\alpha_{micro}$: proportion of time spent in microenvironment (kitchen, bedroom, outdoor patio etc.)

$PM_{micro}$: personal $PM_{2.5}$ levels in microenvironment

$\alpha_{out}$: proportion of time spent outside of the microenvironment (outside of the household)

$PM_{out}$: personal $PM_{2.5}$ levels outside of the microenvironment (outside of the household)

$\tau$: difference between personal $PM_{2.5}$ concentrations and microenvironment $PM_{2.5}$ concentrations. $\tau > 0$ indicating beacon microenvironmental method in biomass household will overestimate personal $PM_{2.5}$.

$Fuel$: Fuel type of the household, if household is using biomass, Fuel = 1, if using LPG, Fuel = 0.

$\delta$: the ratio between $PM_{out}$ and $PM_{micro}$, $\delta = PM_{out}/PM_{micro}$.

$PM_{indirect} = PM_{direct}(1 + \tau \times Fuel)$ (equ-S2) is based on the fact that beacon microenvironmental indirect method can only measure area $PM_{2.5}$ levels and will overestimate true personal $PM_{2.5}$ for biomass household ($\tau$ only exists in biomass households) due to large emissions in microenvironment especially kitchen. In addition, in our study in Guatemala, beacon indirect method only take the microenvironmental $PM_{2.5}$ as the 24-hour average level, not accounting for the time spent outside of the household.

Since the proportion of time spent in microenvironment and outside of the household sum up as 1, i.e. $\alpha_{micro} + \alpha_{out} = 1$, so we can substitute $\alpha_{out}$ and $PM_{out}$ in equ-1, $PM_{direct} = PM_{micro}(\alpha_{micro}(1 - \delta) + \delta)$.

So that

$$PM_{indirect} = \frac{(1 + \tau \times Fuel)}{\alpha_{micro}(1 - \delta) + \delta} PM_{direct}$$

Equ-3-3S

For biomass household, since indoor fuel is biomass, and microenvironments are mostly indoor microenvironment, so $PM_{micro} > PM_{out}$ and then $\delta < 1$, also notice $\alpha_{micro} < 1$. Then, equ-3-3S can be re-written as:

$$PM_{indirect} = \frac{(1 + \tau)}{\alpha_{micro}(1 - \delta) + \delta} PM_{direct} > \frac{(1 + \tau)}{(1 - \delta) + \delta} PM_{direct} > PM_{direct}$$

Equ-3-4S

If $\delta \geq 1$, the relationship $PM_{indirect}$ and $PM_{direct}$ cannot be determined exactly. However, we believe $\delta \geq 1$ is not likely to happen in biomass households, due to $PM_{micro} > PM_{out}$ is usually true in biomass household.

For LPG household, we believe $PM_{micro} \leq PM_{out}$ is usually observed and indoor microenvironment $PM_{2.5}$ levels will be low if there is high compliance of LPG intervention. We can then find $\delta \geq 1$. and equ-3-3S can be re-written as:

$$PM_{indirect} = \frac{1}{\alpha_{micro}(1-\delta)+\delta} PM_{direct} \leq \frac{(1)}{(1-\delta)+\delta} PM_{personal} = PM_{direct}$$

Equ-3-5S

Therefore, beacon indirect method will likely over-estimate personal exposure levels in biomass households and likely under-estimate personal exposure levels in LPG intervention households. Beacon indirect methods will give larger exposure contrast between intervention and control compared to true personal exposure contrast and will bias the exposure-response relationship towards the null.

# Chapter 4 **Associations of Fine Particulate Matter ($PM_{2.5}$) Exposure with Gestational Blood Pressure among Rural Guatemala Women: A Prospective Cohort Study**

## 4.1 Introduction

More than 3 billion people rely on solid fuels for cooking and heating globally (Bonjour et al., 2013). Household air pollution (HAP), including fine particulate matter ($PM_{2.5}$) generated from cooking with solid fuel, is the leading health risk factors in low- and middle-income countries (LMICs) (Gakidou et al., 2017). In Guatemala, HAP is the largest environmental risk factor, responsible for over 5% of total mortality in 2017 and 17% of these estimated deaths are from cardiovascular disease (GBD Compare | IHME Viz Hub—Risk Factors in Guatemala, n.d.).

Elevated blood pressure (BP), a risk factor for cardiovascular disease, has also been shown to be associated with HAP from solid fuel combustion in multiple LMICs, including Guatemala (McCracken et al., 2007), Honduras (Young et al., 2019) and China (Baumgartner et al., 2011). A recent systematic review examining HAP and high BP and hypertension concluded that the use of solid fuel was associated with increased risk of hypertension (Li et al., 2020). The biological mechanisms by which air pollution exposure increases BP have been well demonstrated, through an imbalance of lung autonomic nerve system, systematic oxidative stress and inflammatory, as well as endothelial, dysfunction (Brook et al., 2010, 2017; Giorgini et al., 2016; McCracken et al., 2012).

During pregnancy, elevated BP has important implications for gestational hypertensive disorders such as preeclampsia, a leading cause of maternal, fetal and neonatal mortality (Mol et al., 2016). Because women of reproductive age are usually the primary cooks in the household in rural low-income settings, and studies have

found that women who cook with biomass have an order of magnitude higher exposure to $PM_{2.5}$ than those cooking with clean fuel (Hystad et al., 2019), they are therefore especially vulnerable to HAP exposure.

Unlike consistent results of HAP effects on BP for non-pregnant women, previous studies examining HAP and BP levels show mixed results for pregnant women. Some studies have reported a positive association of personal exposure to CO on systolic blood pressure (SBP) and diastolic blood pressure (DBP) (Quinn et al., 2016, 2017) among pregnant women in Ghana; and another study in Nigeria found an ethanol cookstove intervention can reduce DBP during pregnancy (Alexander et al., 2017). In contrast, a cross-sectional study in India reported that use of biomass cooking fuel is associated with lower BP among pregnant women compared to gas users (B. J. Wylie et al., 2017). This large heterogeneity of effects of HAP on gestational blood pressure is similar to a previous meta-analysis and systematic review of ambient air pollution effects on gestational blood pressure (Pedersen et al., 2014), as well as seen in studies between smoking and gestational pre-eclampsia (Karumanchi S. Ananth & Levine Richard J., 2010). In addition, gestational BP changes naturally during pregnancy (Macdonald-Wallis Corrie et al., 2012), which decreased in first and second trimester and then increased in the third trimester. Longitudinal studies that follow pregnant women could capture the changes of BP and HAP exposure over time during gestation, and overcome potential bias introduced in cross-sectional studies which examined association between HAP exposure and BP at delivery at once (Luque-Fernandez et al., 2016; Blair J. Wylie et al., 2015). To date, however, few longitudinal studies have been conducted to examine this relationship (Alexander et al., 2017).

To address these gaps, we designed this prospective cohort study to assess the longitudinal association between personal 24-hour $PM_{2.5}$ exposure and SBP and DBP levels during pregnancy in an intensive cookstove emission setting. The study uses BP measurements during pregnancy from 800 women prospective cohort enrolled in Household Air Pollution Intervention Network (HAPIN) trial in Guatemala. This study was approved by the institutional review boards of the Universidad del Valle de Guatemala (146-08-2016/11-2016) and Emory University (00089799). The trial is registered at ClinicalTrials.gov (Identifier NCT02944682).

## 4.2 Method

### 4.2.1 Study site and population

This study was based on data from pregnant women participants enrolled in HAPIN trail, a randomized controlled intervention trial among 3200 pregnant women in India, Guatemala, Peru and Rwanda to assess the effects of a liquefied petroleum gas (LPG) cook stove and fuel intervention in settings traditionally relying on solid biomass fuel (wood, charcoal, dung and agricultural residue). In this study, we followed women enrolled from the Guatemala site (N=809) in Xalapán area of Jalapa Department in rural Guatemala, 100 km east of Guatemala City. Xalapán area is a mountainous region in a tropical wet climate zone, with rainy season between May – October, and dry season November – April. The average temperature is 20 °C, and the average elevation is 1500 meters (871m – 2677 m) in the study site. Details of the HAPIN trial have been published elsewhere [ref].

Briefly, in order to be eligible to participate in the study, women were required to be between 19 and 35 of age, non-smokers, between 9 and 20 weeks gestational (confirmed by ultrasound and last menstrual period) and cooking with on biomass. Potentially eligible women were first identified at local prenatal clinics by trained Guatemalan nurses employed by the HAPIN project. If women met eligibility criteria, nurses went to the participant's home to consent the household to participate in the study, usually the same day but up to 2 weeks after the ultrasound visit, and collected baseline clinical and demographic baseline information and scheduled a follow-up visit with the exposure team who would collect data on stove use at baseline and other household characteristics. During that visit, we measured blood pressure and $PM_{2.5}$ air pollution personal exposures over past 24-hour. Health outcomes and personal $PM_{2.5}$ exposure was assessed again at two follow-up visits at participants' homes, after randomized and LPG stove installation in the intervention arm. Due to the fact that HAPIN is ongoing, and treatment arm remains blinded with respect to exposure and health outcomes, this study focused solely on analyzing observed exposures and corresponding responses and, thus, did not include intervention status.

### 4.2.2 Outcome and Covariates

Following informed consent to participate in the study, gestational blood pressure was assessed on enrollment (baseline), and two follow up visits at approximately 24-28 gestational weeks (follow-up 1) and 32 – 36 gestational weeks (follow-up 2), on the next day of the air pollution exposure visit described below. Blood pressure was measured in triplicate using an automatic digital blood pressure machine (Model HEM-907XL, OMRON, Durham, NC, USA) based on a standardized operational protocol. Field workers ensured that the pregnant women participants had not smoked, had alcohol or caffeinated drinks or cooked using biomass in the 30-minute period prior to the blood pressure measurement. Participants rested for 5 minutes in a sitting position in a chair that supported their back and arms with legs uncrossed. Blood pressure was measured in triplicate, 2 minutes between each measurement on the right arm. The second and third blood pressure readings were averaged to represent blood pressure levels at that visit. If a participant was found to have a systolic blood pressure >140 mmHg and/or a diastolic blood pressure >90 mmHg, she was referred to the nearest health center or hospital to receive appropriate treatment. Baseline and follow-up covariates were collected in the case reporting forms (CRFs) on tablets using REDCap software (Vanderbilt University, Nashville, TN, USA). We collected data on participants' demographic information, height, weight, medical history, lifestyle behaviors, anthropometry and self-reported air pollution exposure as potential covariates. We calculated participants household assets index based on a simplified asset indices based on 13 household assets ownership (refrigerator, wash machine, microwave, computer, car, toilet, drinking water source, cooking fuel, roof material, wall material and floor material) at baseline visit using EquityTool (Chakraborty et al., 2016; EquityTool, 2020), which shows agreement with household assets index from Demographic and Health Survey (DHS) in Guatemala. We also calculated participants body mass index (BMI) at each visit time. For participants who reported taking anti-hypertension medicine, we excluded these data points due to we believe these will distort the association between exposure and outcome.

### 4.2.3 $PM_{2.5}$ Exposure

HAP exposure assessment was conducted according to the HAPIN trial protocol. For each household visit, we measured 24-hour personal $PM_{2.5}$ exposures using Children's Enhanced MicroPEM (ECM, RTI International, USA), to represent longer term personal exposure to $PM_{2.5}$. ECMs samplers were programmed to sample $PM_{2.5}$ with an air flow of 0.3 L/min rate and collection on Teflon filters with 15 mm diameter for gravimetric analysis (PT15-AN-PF02, MTL LLC., Minneapolis, MN, USA). All Teflon filters were pre- and post-weighed in a temperature- and humidity-controlled laboratory at the University of Georgia (UGA). Filters were transported in cool conditions inside air-sealed bags between the Guatemala field site and UGA. We used gravimetric $PM_{2.5}$ concentration calculated by mass deposition on filters divided by sampled air volume as the exposure variable. As shown in a previous study, the gravimetric $PM_{2.5}$ measurements made by ECM have a limit of detection around 5 $\mu g/m^3$ (Liao et al., 2019).

### 4.2.4 Statistical Analysis

Firstly, we assessed the overall crude association between blood pressure level between $PM_{2.5}$ exposure (both quartiles and continuous) using simply linear regression. Then, to take into the longitudinal design of this study, we conducted a mixed effects model to assess the longitudinal exposure-response relationship between personal $PM_{2.5}$ exposure (both quartiles and continuous) and gestational BP levels. The mixed effects model included 24-hour personal gravimetric $PM_{2.5}$ exposure (both quartiles and continuous) as the exposure variable, and time-vary and time-invariant covariates (Equation 4-1 and Equation 4-2). Covariates considered in the model are potential risk factors of gestational hypertension and blood pressure, including gestational age, age at baseline, body weight index (BMI), self-reported physical activity, alcohol consumption, caffeinated drink consumption, parity, previous adverse birth event, day of week (weekday or weekend) and time of the day (morning/afternoon) of BP measurement (Hermida et al., 2000; Mol et al., 2016; Shen et al., 2017). Among these covariates, age at baseline, self-reported physical activity, alcohol drinks consumption, caffeinated drinks consumption, previous adverse birth event, education level, and day of week and time of the day of BP

measurements are not associated with both SBP and DBP, and did not significantly affect model parameters. Therefore, we did not include those predictors in the main analysis. For gestational age, since gestational blood pressure level shows a nonlinear relationship with gestational age (Hermida et al., 2000), we included both linear and quadratic term of gestational age in the model, similar to previous study (Alexander et al., 2017).

$$BP_{ij} = \beta_0 + \boldsymbol{\beta_1^T X_{ij}} + \boldsymbol{\beta_2^T Q\_PM_{ij}} + \boldsymbol{\beta_3^T Z_i} + \beta_4 GA_{ij} + \beta_5 GA_{ij}^2 + \delta_i + \varepsilon_{ij}$$

**Equation 4-1**

$$BP_{ij} = \beta_0 + \boldsymbol{\beta_1^T X_{ij}} + \beta_2 ln(PM_{ij}) + \boldsymbol{\beta_3^T Z_i} + \beta_4 GA_{ij} + \beta_5 GA_{ij}^2 + \delta_i + \varepsilon_{ij}$$

**Equation 4-2**

$BP_{ij}$ is the BP (either SBP or DBP) levels for participant i at observation j. $\beta_0$ is the population intercept. $\boldsymbol{X_{ij}}$ is the vector containing time varying covariates except for personal $PM_{2.5}$ exposure and gestational age, in the main analysis, this term only includes BMI at each visit. $\boldsymbol{Q\_PM_{ij}}$ (in equation 4-1) denotes each quartile of $PM_{2.5}$ exposure variables for participants in $i$ in visit $j$. $\boldsymbol{PM_{ij}}$ (in equation 4-2) denotes 24-hour personal $PM_{2.5}$ exposure concentration for participants in $i$ in visit $j$. $\boldsymbol{Z_i}$ is the vector containing individual level time invariant variables at baseline, in the main analysis, this term includes parity and maternal age. $GA_{ij}$ is the gestational age, both for pregnant women i at visit j. $\delta_i$ is the individual random intercept and $\varepsilon_{ij}$ is the model residual, both of which are assumed to be normally distributed.

In order to assess the robustness of the model and avoid residual confounding, we also conducted sensitivity analysis. Firstly, we reported results of model with additional covariates as potential confounders which are removed in the main analysis. These additional covariates include age at baseline, self-reported physical activity, alcohol drinks consumption, caffeinated drinks consumption, previous adverse birth event, education levels, and day of week and time of the day of BP measurements. Secondly, we conducted generalized mixed effects

additive model, adding spline smoothing to continuous variables, including maternal age at baseline, gestational age, BMI and time of measurement (minutes after 12AM in the morning). The details of sensitivity analysis were provided in appendix. The data processing and data statistical analysis was conducted in R software (version 3.6.0), with lme4, gamm4 packages for mixed effects regression and additive models.

## 4.3 Results

At baseline, 809 participants completed assessments. Nine participants subsequently left the study due to voluntary withdrawal (n=4), moving away from study area (n=4) and termination of the pregnancy (n = 1), yielding a total of 800 participants who were randomized and eligible for following up. During follow-up, a total of 57 participants exited the study before their children were born (end of follow-up), due to voluntary withdraw (n = 13), withdrawal of participant by study team (n = 1), moving out of the study area (n = 10) and termination of pregnancy during follow-up (n = 33). An additional 22 participants exited the study before the first follow-up visit, and 13 participants exited before the second follow-up visit. A total of 786 participants completed the follow-up 1, and 708 completed follow-up 2 (**Figure 4-1**).

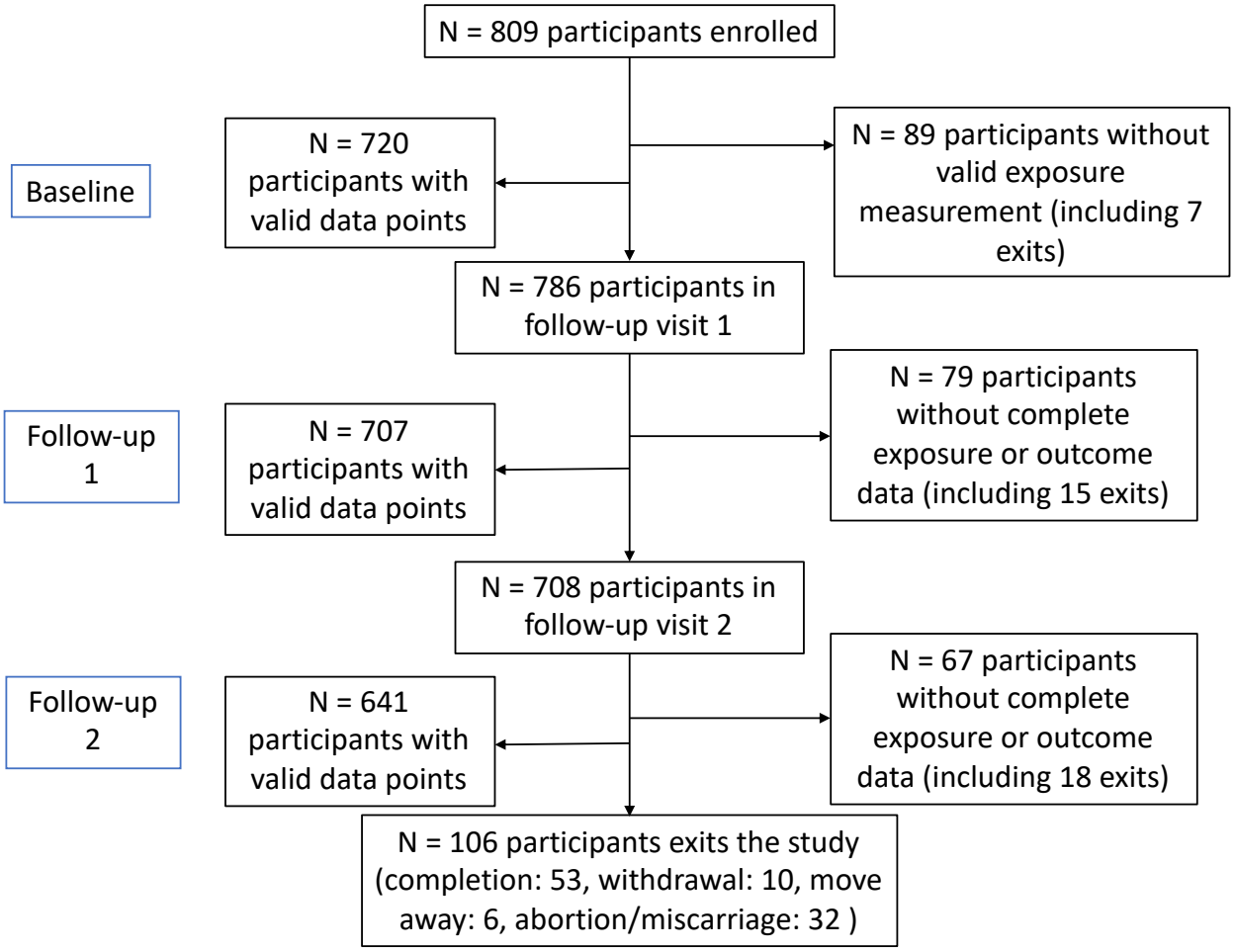



**Figure 4-1**. Flow Diagram of Study Participants at Baseline, Follow-up 1 and Follow-up 2 visits

A total of 2179 gravimetric personal $PM_{2.5}$ measurements were obtained after removing samples (n =258, 10.6%) that failed to pass quality checks or have not been post-weighed, consisting of 720 samples at baseline, 708 at follow-up 1 and 641 at follow-up 2. This sample number contains the 110 side-by-side duplicated samples, which were averaged to represent the personal $PM_{2.5}$ exposure at that visit. **Figure 4-1S** in the appendix shows the details of processing of gravimetric samples. Of 110 duplicated ECM sample pairs, the personal $PM_{2.5}$ exposure levels correlate very well with each other (Spearman R = 0.97). In addition, we collected 96 field blank samples, with a median filter weight change of +2.5 μg (interquartile range, IQR: +1 – +4.25 μg), showing the small change of the filter mass change during the sampling.

After accounting for missing data and data points failing quality check, approximately 10% of samples that are removed and potentially 4% of the data points are still in following up. The current analysis includes measurements from 720 participants in the baseline, 707 participants in follow-up 1 and 641 participants in follow-up 2 visits. The participants we included in the analysis (n = 720) shown similar characteristics in demographic as all participants enrolled in the study at baseline (**Table 4-1**).

**Table 4-1**. Baseline Characteristics of pregnant women and household in HAPIN Trial Guatemala Site

| Variables | Characteristics<br>All participants (N = 809) | Participants included in the analysis at baseline (n=720) |
|---|---|---|
| Maternal Age (years) mean (sd, min – max) | 24.7 (4.4, 18 – 34.9) | 24.8 (4.4, 18 – 34.9) |
| Gestational age (weeks) mean (sd, min – max) | 15.1 (3, 9.4 – 21.6) | 14.9 (3, 9.4 – 21.6) |
| Maternal Education (n,%) | | |
| No formal education | 68 (8.4%) | 60 (8.6%) |
| Primary school incomplete | 314 (39.1%) | 273 (39.2%) |
| Primary school complete | 227 (28.3%) | 199 (28.6%) |
| Secondary school incomplete | 85 (10.8%) | 82 (10.1%) |
| Secondary school complete | 94 (11.9%) | 70 (11.8%) |
| Vocational school | 10 (1.2%) | 9 (1.3%) |
| College | 2 (0.2%) | 2 (0.3%) |
| First pregnancy [n(%)] | 209 (26%) | 170 (24%) |
| BMI ($kg/m^2$) mean (sd, min - max) N = 804 | 23.79 (3.32, 16.42 - 44.24) | 24.44 (18.03, 16.42 – 44.24) |
| Drink alcohol [n(%)] | 6 (1%) | 4 (1%) |
| Employment status [n(%)] | | |
| Employed | 21 (3%) | 20 (3%) |
| Unemployed, work in own house | 787 (97%) | 675 (97%) |
| Known history of hypertension (including pre-existing hypertension and hypertension) [n(%)] | 16 (1.9%) | 12 (1.7%) |
| Known history of adverse pregnancy outcomes (including miscarriages and stillbirth) [n(%)] | 142 (18%) | 127 (18%) |
| Known history of diabetes [n(%)] | 2 (0.2%) | 1 (0.1%) |
| Morning time of the BP measurement | 646 (80.9%) | 572 (82.3%) |
| Day (weekend) of BP measurement | 238 (30%) | 215 (31%) |
| Household characteristics | | |

| | | |
|---|---|---|
| Asset index[1], n(%), N = 808 | | |
| Quintile 1 (lowest) | 616 (76.2%) | 535 (77.1%) |
| Quintile 2 | 130 (16.1%) | 104 (15.0%) |
| Quintile 3 | 54 (6.7%) | 48 (6.9%) |
| Quintile 4 | 8 (1%) | 7 (1%) |
| Quintile 5 (highest) | 0 (0%) | 0 (0%) |
| Kitchen type, [n(%)] | | |
| Not enclosed kitchen | 90 (11.3%) | 83 (11.9%) |
| Fully enclosed kitchen (with roof and 4 walls) | 710 (88.7%) | 612 (88%) |
| Fully enclosed kitchen measurements | | |
| Kitchen area ($m^2$), mean (sd, min - max) | 14.4 (7.5, 3.5 - 129) | 14.4 (7.7, 3.5 – 128) |
| Kitchen height (m), mean (sd, min – max) | 2.35 (0.28,1.65 – 3.33) | 2.35 (0.29, 1.65 – 3.33) |
| Number of stoves (N = 798) | | |
| 1 stove | 271 (33%) | 236 (34%) |
| More than 1 stoves | 527 (67%) | 457 (66%) |
| Primary stove type [n(%)] (N = 798) | | |
| Open fire stove with chimney | 640 (80.2%) | 553 (80.1%) |
| Biomass stove with chimney | 158 (19.8%) | 137 (19.9%) |
| Primary fuel type [n(%)] (N = 796) | | |
| Grass/shrubs | 1 (0.1%) | 0 |
| Agriculture waste | 1 (0.1%) | 1 (0.1%) |
| Wood | 792 (99.5%) | 688 (99.6%) |
| Electricity | 2 (0.2%) | 2 (0.3%) |
| Number of days the primary stove is used, [n(%)] | | |
| Everyday | 757 (94.6%) | 659 (95%) |
| Over 3 days per week – everyday | 43 (5.4%) | 36 (5%) |
| Number of hours per day the primary stove is used, mean (SD, min – max) | 5.9 (2.6, 1 – 18) | 5.9 (2.6, 3 – 18) |

[1]: Based on 13 household assets including refrigerator, wash machine, microwave, computer, car, toilet type, water source, cooking fuel, roof materials, wall materials, floor materials

Blood pressure measurements and gestational age at three visits are shown in **Table 4-2**. In the baseline and first follow-up visit, the mean gestational age was 15.2 (range: 9.4 – 21.6) and 25.5 (range: 21.4 – 31.3) weeks, respectively. Blood pressure levels among study participants are generally low at the baseline and first follow-up visits, with mean systolic blood pressure 103.5 (standard deviation: 8.5) mmHg and 103.7 (standard deviation: 8.3) mmHg and diastolic blood pressure 59.0 (standard deviation: 7.3) mmHg and 59.2 (standard deviation: 7.2) mmHg, respectively. However, in follow-up 2 visit when the mean gestational age is 33.4 (range: 31.1 – 37.4)

weeks, there is an increase of blood pressure levels among participants, with the mean systolic blood pressure and diastolic blood pressure is 106.8 (standard deviation: 8.5) and 62.6 mmHg, respectively. Only four participants were classified as hypertensive, as defined as SBP over 140 mmHg or DBP over 90 mmHg, one in follow-up 1 visit and three in the follow-up 2 visit. All participants who were classified to have hypertension were referred to hospitals, and the participant who shown hypertension at follow up visit 1 existed the study due to miscarriage. Four participants reported taking anti-hypertensive medicine during visit, one in follow-up visit 1 and four in follow up visit 2 (1 participant reported both in follow up 1 and follow up 2). We excluded these 5 data points when anti-hypertension medicine was taken.

**Table 4-2**. Systolic and Diastolic Blood Pressure Measurement (mmHg) at Each Visit for Women

| Visit | n | Gestational week (mean, min – max) | Mean | SD | Hypertension[1], n(%) |
|---|---|---|---|---|---|
| Baseline | | | | | |
| SBP | 715 | 15.1 (9.4 – 21.9) | 103.5 | 8.5 | 0 (0%) |
| DBP | 715 | | 59.1 | 7.4 | |
| Follow-up 1 | | | | | |
| SBP | 703 | 25.3 (21.4 – 31.3) | 103.8 | 8.3 | 2 (0.3%) |
| DBP | 703 | | 59.2 | 7.4 | |
| Follow-up 2 | | | | | |
| SBP | 635 | 33.3 (30.7 – 37.2) | 107 | 8.6 | 3 (0.5%) [2] |
| DBP | 635 | | 63 | 8.0 | |

[1] hypertension was defined as systolic blood pressure (SBP) equal or more than 140 mmHg or diastolic blood pressure equal or more than 90 mmHg.

[2] Three participants developed hypertension at follow-up visit 2 are different from the participants who developed hypertension at follow-up visit 1.

Nearly all (98%) of participants reported they wore the ECM all the time except for sleeping and bathing. The mean sampling time was 24.2 hours (22.2 – 26.4 hours). The personal 24-hour $PM_{2.5}$ exposure levels at baseline showed that exposure was relatively high with mean (standard deviation, SD) of 142.2 (102.7) $\mu g/m^3$. (**Table 4-3**). At follow up visits, when half of the study participants had received the stove intervention, the overall personal $PM_{2.5}$ exposure was substantially lower, with a mean (SD) of 82.0 (95.5) $\mu g/m^3$ and 79.7 (83.3) $\mu g/m^3$ in follow-up 1 and follow-up 2 visits, respectively (**Figure 4-2**). Among all the participants, 525 of them have

three valid observations in the longitudinal follow-ups, 225 of them have two observations and 43 of participants have one data point during the whole period.

**Table 4-3**. Personal 24-hour $PM_{2.5}$ Exposure at Each Visit

| Visit | Personal 24-hour $PM_{2.5}$ exposures ($\mu g/m^3$) |
|---|---|
| Baseline (n = 720) | |
| Mean (SD), range | 142.2 (102.7), 2.3 – 619.5 |
| Median (IQR) | 115.6 (67.6 – 186.6) |
| Follow-up 1 (n = 708) | |
| Mean (SD), range | 82.0 (95.5), 3.3 – 788.4 |
| Median (IQR) | 45.2 (24.4 – 102.3) |
| Follow-up 2 (n = 641) | |
| Mean (SD), range | 79.7 (83.3), 1.7 – 506.7 |
| Median (IQR) | 46.9 (26.4 – 99.9) |

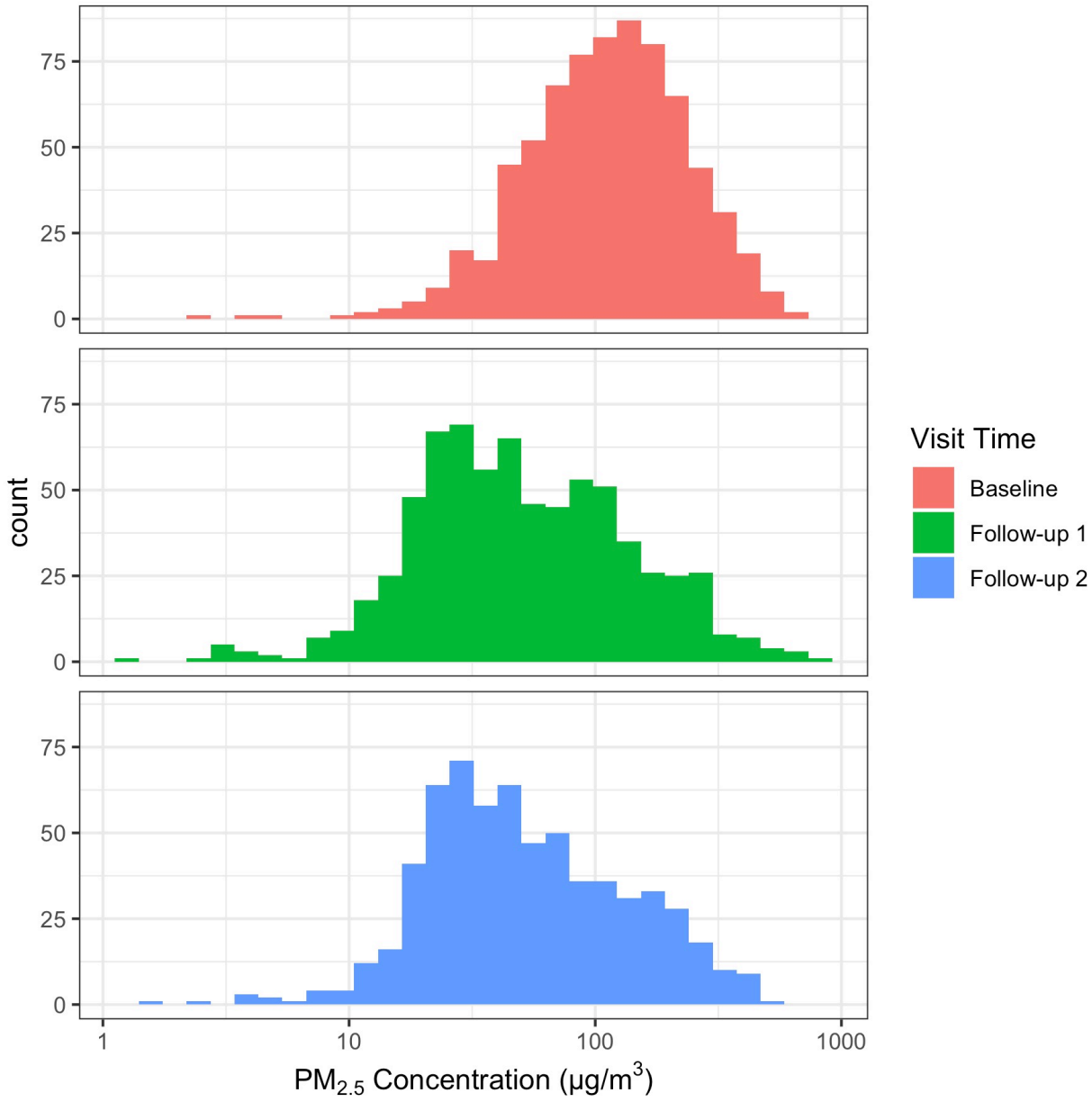



**Figure 4-2**. Distribution of 24-hour Personal $PM_{2.5}$ Exposure. The x-axis is in log-scale.

The scatter plot between personal $PM_{2.5}$ exposure and blood pressure levels are shown in **Figure 4-3**. The figure shows that $PM_{2.5}$ exposure follow the log-normal distribution, and shows reductions in the post-intervention follow-up phases compared to baseline phase. The overall quartile of $PM_{2.5}$ exposure is calculated as: Q1: $PM_{2.5} < 32$ µg/m³, Q2: 32 µg/m³ $< PM_{2.5} < 68$ µg/m³, Q3: 68 µg/m³ $< PM_{2.5} < 137$ µg/m³, and Q4: 137 µg/m³ $< PM_{2.5}$.

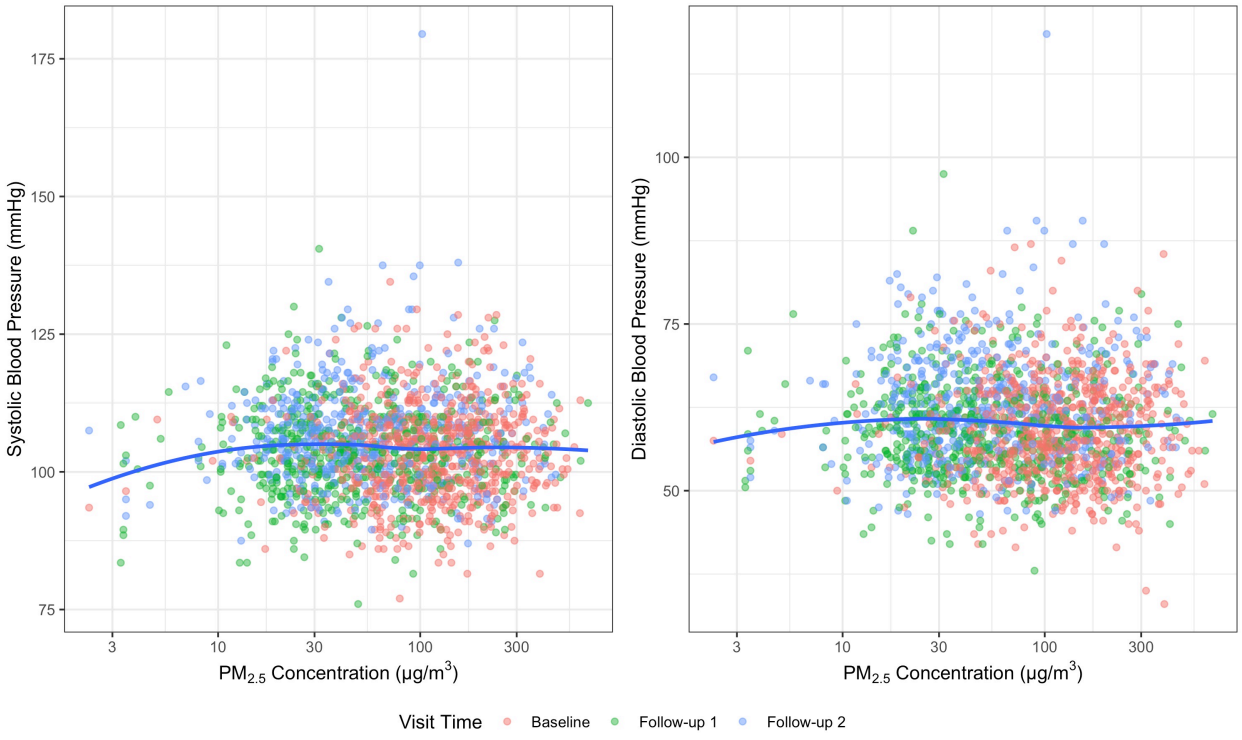



**Figure 4-3**. Scatter Plot of Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP) Between Personal 24-hour $PM_{2.5}$ Exposure with Smoothing Line

The overall crude model shown that $PM_{2.5}$ exposure levels (both continuous and quartile) did not show strong association with systolic blood pressure (SBP) and diastolic blood pressure (DBP) levels (appendix **Table 4-1S**). However, after we controlled potential risk factors of gestational hypertension, including gestational age, BMI and parity, and considering the longitudinal design of the study using mixed effects model, we found statistically significantly association between $PM_{2.5}$ exposure and SBP but not DBP. In **Table 4-4**, we show the model parameters for mixed effects models with random intercept at individual level, between personal $PM_{2.5}$ exposure concentration and blood pressure (SBP and DBP). In the mixed effects model with $PM_{2.5}$ exposure quartiles, the second, third and fourth quartile of personal $PM_{2.5}$ exposure concentration shown a positive association between SBP but not DBP compared to lowest quartile of personal $PM_{2.5}$ exposure, adjusting for a random intercept for each participant and selected covariate (gestational age, BMI and parity). In the second, third and fourth $PM_{2.5}$ exposure quartiles, the SBP is 0.653 (95% CI: -0.147, 1.453), 0. 312 (95% CI: -0.560, 1.184) and 1.065 (95% CI: 0.142, 1.988) mmHg higher than in the first $PM_{2.5}$ exposure quartile, respectively.

These results manifested some evidence of non-monotonic response between $PM_{2.5}$ exposure and SBP, with highest SBP levels in fourth and second quartiles of $PM_{2.5}$ exposure as compared to first quartile.

**Table 4-4**. Model Parameter for Main Analysis using Mixed Effects Model [1]

| | Quartile $PM_{2.5}$ Model [2] | Continuous $PM_{2.5}$ Model [3] |
|---|---|---|
| Systolic Blood Pressure (SBP) | | |
| Q1 ($PM_{2.5} < 32$ μg/m$^3$) | Reference = 0 | NA |
| Q2 (32 μg/m$^3$ <$PM_{2.5}$< 68 μg/m$^3$) | 0.653 (-0.147, 1.453) | NA |
| Q3 (68 μg/m$^3$ <$PM_{2.5}$< 137 μg/m$^3$) | 0.312 (-0.160, 0.784) | NA |
| Q4 (137 μg/m$^3$ <$PM_{2.5}$) | 1.065 (0.1418, 1.988) | NA |
| Continuous $PM_{2.5}$ Natural-logged | NA | 0.342 (0.003, 0.681) |
| Gestational age (week) | -0.691 (-0.922, -0.460) | -0.690 (-0.921, -0.459) |
| Gestational age quadratic term (week$^2$) | 0.018 (0.014, 0.022) | 0.019 (0.015, 0.023) |
| BMI (kg/m$^2$) | 0.037 (0.01, 0.068) | 0.038 (0.007, 0.069) |
| Had birth before | 1.273 (0.091, 2.455) | 1.272 (0.092, 2.452) |
| Diastolic Blood Pressure (DBP) | | |
| Q1 ($PM_{2.5} < 32$ μg/m$^3$) | Reference = 0 | NA |
| Q2 (32 μg/m$^3$ <$PM_{2.5}$< 68 μg/m$^3$) | 0.021 (-0.737, 0.780) | NA |
| Q3 (68 μg/m$^3$ <$PM_{2.5}$< 137 μg/m$^3$) | -0.511 (-1.336, 0.314) | NA |
| Q4 (137 μg/m$^3$ <$PM_{2.5}$) | -0.058 (-0.926, 0.810) | NA |
| Continuous $PM_{2.5}$ Natural-logged | NA | -0.072 (-2.155, 2.011) |
| Gestational age (week) | -0.994 (-1.215, -0.773) | -0.991 (-1.212, -0.769) |
| Gestational age quadratic term (week$^2$) | 0.025 (0.021, 0.0289) | 0.025 (0.021, 0.029) |
| BMI (kg/m$^2$) | 0.018 (-0.009, 0.045) | 0.019 (-0.010, 0.048) |
| Had birth before | 1.017 (-0.125, 2.160) | 1.023 (0.909, 1.137) |

[1] The results are shown as point estimate (95% Confidence Interval (CI)) per 1 mmHg increase of blood pressure levels

[2] Adjusted for gestational age in weeks including linear and quadratic terms, BMI and parity (whether had birth before), exposure variable is quartile of $PM_{2.5}$ exposure concentration.

[3] Adjusted for gestational age in weeks including linear and quadratic terms, BMI and parity (whether had birth before), exposure variable is natural log-transformed continuous $PM_{2.5}$ exposure concentration.

Similarly, in the mixed effects model with continuous $PM_{2.5}$ exposure variable, SBP but DBP is significantly associates with $PM_{2.5}$ exposure. We show that a 1 unit increase of log transformed 24-hour personal $PM_{2.5}$ exposure was associated with a 0.342 (95% confidence interval (CI): 0.003, 0.681) mmHg change of SBP and -0.072 (95% CI: (-2.155, 2.011) mmHg change of DBP, respectively. After additional adjusting for a random intercept for each participant and selected covariate (gestational age, BMI and parity) . The effects of covariates

in mixed effects models are also listed in the **Table 4-4**, and the direction of effects is the same as expected. In the mixed effects model, the variance of random intercept for each participant is larger than the model residual, with intra-class correlation more than 0.55 for models of SBP and DBP as outcomes. This indicates the relatively large between person variability in blood pressure level, explained by random intercept at participant level.

The sensitivity analysis results are shown in appendix. **Table 4-2S** in appendix shows the model parameters of mixed effects models between continuous $PM_{2.5}$ exposure concentration and blood pressure levels with more confounder adjustment, adding age at baseline, self-reported physical activity, alcohol drinks consumption, caffeinated drinks consumption, previous adverse birth event, education levels, and day of week and time of the day of BP measurements. **Table 4-3S** reported additional sensitivity analysis results of $PM_{2.5}$ exposure on blood pressure levels. These analyses include adding regression splines to all continuous variables except for personal $PM_{2.5}$ exposure, and model excluding participants taking hypertension medication. Sensitivity analysis were consistent in in direction and strength of association between $PM_{2.5}$ exposure and SBP level, as reported in main analysis.

## 4.4 Discussion

In this study, we showed that an increase in 24-hour personal $PM_{2.5}$ exposure in a prospective cohort of 800 pregnant women enrolled in HAPIN trial was associated with an increase in their SBP levels during pregnancy, in a longitudinal data analysis. In the overall crude analysis, no strong evidence showing the positive association between $PM_{2.5}$ exposure and both SBP and DBP have been found. But in the longitudinal analysis, after controlling for gestational age, BMI, and parity, we found a positive and statistically significant association between $PM_{2.5}$ exposure concentration and SBP, which tends to show a non-monotonic relationship. The model shows that an increase of 1 unit of log-transformed $PM_{2.5}$ exposure was associated with a 0.342 (95% confidence interval (CI): 0.003, 0.681) mmHg increase of SBP; there was no clear association between $PM_{2.5}$ exposure and DBP. Our model results implied that with the 50% reduction of 24-hour personal $PM_{2.5}$ exposure

is associated with 0.237 (95% CI: (0.002, 0.474) mmHg reduction of SBP. In addition, the mixed effect model with continuous $PM_{2.5}$ exposure with adjusted variables show that the random intercept for each participant explained relatively large variability, compared to model residual. The intra-class correlation is 0.61 and 0.55 for SBP and DBP outcomes.

To our knowledge, this is the first study examining the longitudinal exposure-response association between $PM_{2.5}$ as HAP and BP during pregnancy. Comparing with previous study, the results from this study are consistent with findings from cross sectional studies among pregnant women (Alexander et al., 2017; Quinn et al., 2016, 2017; Thompson et al., 2011) and non-pregnant women (Li et al., 2020; McCracken et al., 2007), even though previous studies have shown heterogeneous findings. Two previous studies assessed the exposure-response association between CO exposure and blood pressure levels among pregnant women (Quinn et al., 2016, 2017). Quinn and collogues in a cross-sectional design found that a 1 ppm increase of 72-hour CO exposure is associated with 0.43 (95% CI: 0.01, 0.86) mmHg increase of DBP among pregnant women cooking with biomass in Ghana in the baseline of an LPG trial. Among the longitudinal follow up of the subset of these 15 pregnant women, they also found that peak CO exposure (defined as hourly concentration >4.1 ppm) was associated with 4.3 (95% CI: 1.1 – 7.4) mmHg increase in ambulatory hourly SBP and 4.5 mmHg (95% CI: 1.9 – 7.2 mmHg) increase in DBP, who cook either with biomass or LPG. Other studies assessed the clean cookstove intervention effects on the blood pressure. One study found in Guatemala, the improved biomass chimney-stove can reduce 1-2 mmHg SBP and DBP level among pregnant women compared to those using open-fire stoves, but these differences are not statistically significant(Thompson et al., 2011). Another ethanol cookstove intervention study in Nigeria found that change in DBP but not SBP over time was significantly different between pregnant women using ethanol compared to those using kerosene or firewood, in a longitudinal follow up (Alexander et al., 2017). Another studies in India, however, found that pregnant women using biomass for cooking has 2 (95%CI: 0.31, 3.77) mmHg lower SBP and 1.96 (95% CI: 0.30, 3.60 ) mmHg DBP at delivery compared to pregnant women using LPG for cooking, in a cross-sectional studies. Our study shows in the longitudinal follow-up, personal 24-hour $PM_{2.5}$ exposure has a stronger association with SBP than

DBP, as we found that 50% increase of $PM_{2.5}$ exposure is associated with 0.237 (95% CI: (0.002, 0.474)) mmHg of SBP, but no clearly association between $PM_{2.5}$ exposure and DBP has been found. Due to the different study design and exposure assessment methods of previous studies, the findings from our paper imply that HAP exposure will influence SBP and will have an adverse effect on gestational hypertension.

As SBP is a better predictor of adverse pregnancy outcomes including adverse events related to central nervous system, cardiorespiratory, hepatic, renal, or hematological morbidity or maternal mortality than DBP did (Mol et al., 2016), the results from our study implied $PM_{2.5}$ exposure may have larger clinical implications than previously reported. The results from our study contribute to the evidence suggesting an increased health risk from HAP. These include hypertensive disorders and other complications during pregnancy. Though the prevalence of the hypertension remains low in our study population (less than 1% have been classified to have hypertension), our study implied that reducing personal $PM_{2.5}$ exposure level could reduce the risk of such risks.

A strength of the study is the repeated measurements of 24-hour $PM_{2.5}$ individual-level exposure and BP three times during pregnancy, which can capture longitudinal changes of exposure and outcomes over time. This study design can reduce the potential selection bias introduced in cross-sectional or case-control studies assessing blood pressure level at delivery that implicitly controlled the gestational age at delivery. At the same time, there are some limitation in our study. Firstly, our study is based on a cohort from randomized controlled trial, participants were randomized to receive free LPG cookstove and fuel or remain using biomass stove after baseline measurements. The LPG cookstove and fuel intervention may introduce a back-door association between $PM_{2.5}$ exposure and blood pressure outcome through dietary changes. If those who receive free fuel in the intervention arm can purchase more food, a weight gain in the intervention group can change through dietary change. We are collecting data on diet and it may affect blood pressure (Mol et al., 2016), and we are be able to control for change in future analyses. Secondly, the results of this study lack consistencies in association between $PM_{2.5}$ exposure and DBP. We did not observe clear association between personal $PM_{2.5}$ exposure and DBP in our study as compared to SBP. Given in inconsistencies of our studies and previous literatures, more

studies are in need to further clarify the association between HAP exposure and gestational blood pressure. Despite these limitations, this investigation shed light on the potential health benefit by reducing HAP levels and switching to clean cooking fuel.

## 4.5 Conclusion

In conclusion, we found that exposure to household air pollution is associated with increases in systolic blood pressure, but not diastolic blood pressure for pregnant women during pregnancy. Consistent with previous studies, our study showed that household air pollution from solid fuel combustion was associated with increases in blood pressure, a risk factor for adverse pregnancy outcome later in the pregnancy. Interventions to reduce HAP level such as clean cooking stove and fuel such as LPG can provide positive impact on blood pressure during pregnancy, and could also provide wider health benefit.

Appendix for Chapter 4

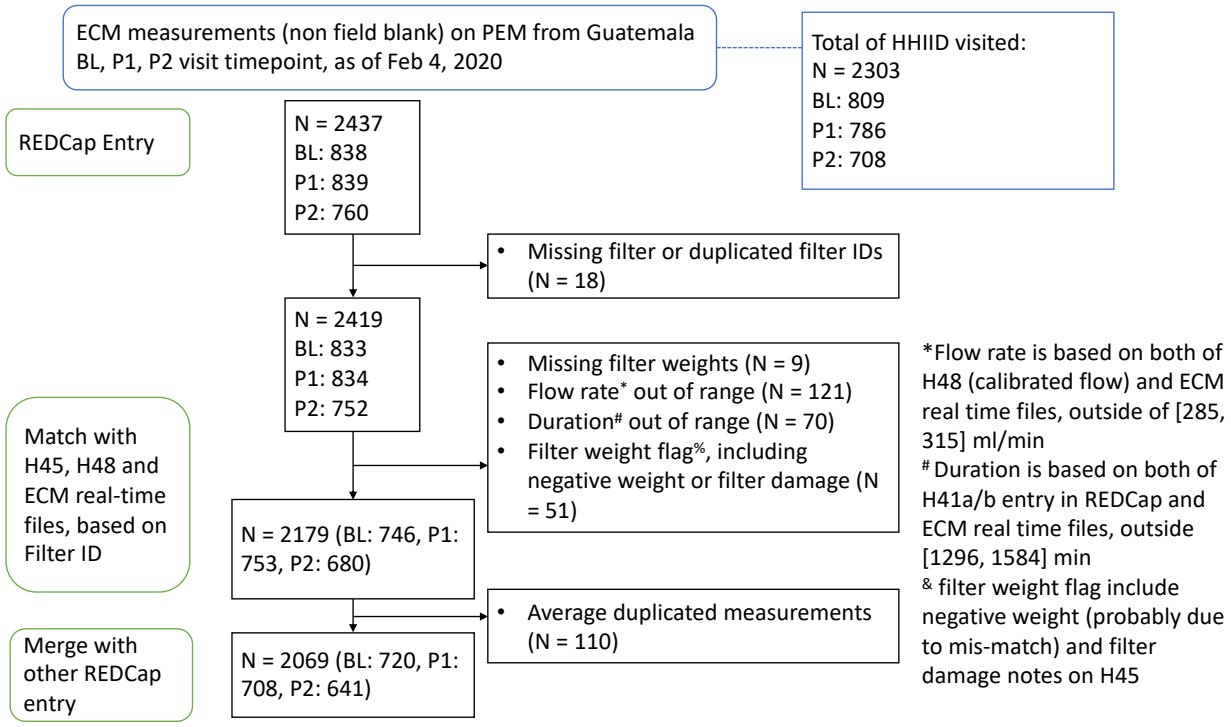



**Figure 4-1S**. Flow Diagram of Personal $PM_{2.5}$ Exposure Samples

**Table 4-1S**. Crude Overall Association between $PM_{2.5}$ exposure and Blood Pressure[1]

| | Crude model [2] |
|---|---|
| Systolic Blood Pressure (SBP) | |
| Q1 ($PM_{2.5}$ < 32 μg/m$^3$) | Reference = 0 |
| Q2 (32 μg/m$^3$ <$PM_{2.5}$ < 68 μg/m$^3$) | 0.088 (-0.965, 1.141) |
| Q3 (68 μg/m$^3$ <$PM_{2.5}$ < 137 μg/m$^3$) | -0.370 (-1.423, 0.683) |
| Q4 (137 μg/m$^3$ <$PM_{2.5}$) | 0.071 (-0.980, 1.122) |
| Continuous $PM_{2.5}$ model | 0.004 (-0.004, 0.003) |
| Diastolic Blood Pressure (DBP) | |
| Q1 ($PM_{2.5}$ < 32 μg/m$^3$) | Reference = 0 |
| Q2 (32 μg/m$^3$ <$PM_{2.5}$ < 68 μg/m$^3$) | -0.510 (-1.463, 0.443) |
| Q3 (68 μg/m$^3$ <$PM_{2.5}$ < 137 μg/m$^3$) | -1.078 (-2.030, -0.125) |
| Q4 (137 μg/m$^3$ <$PM_{2.5}$) | -0.940 (-1.893, 0.0126) |
| Continuous $PM_{2.5}$ model | -0.0025 (-0.006, 0.001) |

[1] The results are shown as point estimate (95% CI) per 1 mmHg increase of blood pressure levels
[2] Only included personal $PM_{2.5}$ exposure quartiles, or natural log-transformed personal $PM_{2.5}$ exposure

**Table 4-2S**. Model Parameters of Mixed Effects Model with Continuous $PM_{2.5}$ Exposure Results and More Covariates Adjustment[1]

| | Model: Systolic blood pressure | Model: Diastolic blood pressure |
|---|---|---|
| Intercept | 128.3 (120.7, 135.9) | 77.75 (70.7, 84.8) |
| Natural log-transformed $PM_{2.5}$ exposure concentration | 0.365 (0.016, 0.714) | -0.086 (-0.41, 0.23) |
| Gestational age in weeks-linear term (week) | -0.69 (-0.925, -0.455) | -1.00 (-1.22, -0.78) |
| Gestational age in weeks-quadratic term ($week^2$) | 0.018 (0.013, 0.023) | 0.025 (0.021, 0.030) |
| Whether conducted vigorous activity over past week | -0.34 (-1.104, 0.424) | 0.20 (-0.58, 0.98) |
| Whether drank alcohol drinks over past 30 minutes | -5.24 (-13.02, 2.54) | 14.57 (3.12, 26.02) |
| Whether drank caffeinate drinks over past 30 minutes | 1.14 (-0.39, 2.67) | -1.59 (-3.16, -0.02) |
| Age in year (year) | 0.004 (-0.125, 0.133) | 0.14 (0.02, 0.26) |
| BMI ($kg/m^2$) | 0.039 (0.007, 0.070) | 0.015 (-0.014, 0.044) |
| Whether had adverse birth event | -0.54 (-1.93, 0.85) | 0.22 (-1.03, 1.47) |
| Whether took anti-hypertensive medicine | 21.97 (16.13, 27.81) | 14.70 (9.09, 20.36) |
| Whether was first pregnancy | 1.45 (0.058, 2.84) | 1.34 (0.09, 2.59) |
| Day (weekend/weekday) of measurement | -0.03 (-0.66, 0.60) | -0.083 (-0.71, 0.54) |
| Time (morning/afternoon) of the measurement | 0.30 (-0.41, 1.01) | -0.201 (-0.92, 0.52) |
| Education level (whether above high school) | -1.26 (-2.85, 0.33) | -0.136 (1.55, 1.27) |

[1] Shown as point estimate (95% CI) per 1 mmHg of increase blood pressure level

**Table 4-4S**. Analysis Results for Mixed Effects Models with Regression Splines and Models Including Participants taking Hypertension Medication[1]

| | Model with regression spline[2] | Models including participants taking hypertension medicine[3] |
|---|---|---|
| Systolic Blood Pressure (SBP) | 0.404 (0.057, 0.751) | 0.368 (0.021, 0.715) |
| Diastolic Blood Pressure (DBP) | -0.032 (-0.357, 0.293) | -0.0590 (-0.380, 0.262) |

[1] Shown as point estimate (95% CI) per 1 mmHg of increase blood pressure level

[2] Models have same layout as models reported in Table 4-2S, but adding a smoothing spline to gestational age, maternal age and BMI, continuous time of BP measurement, also controlled for whether has adverse birth history and whether take anti-hypertensive medicine, whether is the first pregnancy, education level (above high school or not) and day (weekend/weekday) of measurement. Associations are shown as per 1 unit increase of natural log-transformed $PM_{2.5}$ exposure

[3] The model has the same layout as main model (equation 4-2), except including 5 participant-visit data point where participants reported taking hypertension medication.

# Chapter 5 **Gestational and Childhood Exposures to Ambient Fine Particulate matter and Child Survival in India: A Retrospective Cohort Study**

## 5.1 Introduction

Fine particulate matter ($PM_{2.5}$ or particles with an aerodynamic diameter less than 2.5 μm) has been shown to be strongly associated with adverse health effects in children, including preterm birth(Balakrishnan, Ghosh, et al., 2018a; Huynh et al., 2006; Q. Li et al., 2018; Stieb et al., 2012), low birth weight(Bell Michelle L. et al., 2007), low height-for-age z-score(Spears et al., 2019), respiratory infections(Gurley et al., 2013; Smith et al., 2011b) and mortality(Heft-Neal et al., 2018; Son et al., 2017; Woodruff Tracey J. et al., 2006). In India, ambient $PM_{2.5}$ air pollution exposures are among the highest in the world, with an estimated 89.9 $\mu g/m^3$ population weighted annual average(Balakrishnan, Dey, et al., 2018; Chowdhury & Dey, 2016) in 2017. The Global Burden of Disease (GBD) Project identified ambient $PM_{2.5}$ air pollution as the second largest risk factor for disease burden in India, accounting for over 670 000 premature deaths every year, more than 49 000 of which are premature deaths in children under 5 years of age(Balakrishnan, Dey, et al., 2018; Stanaway et al., 2018). To date, however, most evidence on child mortality due to exposure to ambient air pollution during pregnancy and early childhood is from developed regions(Glinianaia Svetlana V. et al., 2004; Scheers Hans et al., 2011; Son et al., 2017; Woodruff T J et al., 1997; Woodruff Tracey J. et al., 2006). Few studies have been conducted in low- and middle-income countries (LMICs)(Goyal et al., 2019; Heft-Neal et al., 2018). While 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 birthweight (Balakrishnan, Ghosh, et al., 2018a; Balakrishnan, Sambandam, Ramaswamy, et al., 2015), impacts on child mortality have thus far not been reported. Furthermore, due to the low density of ambient air pollution monitoring stations in India and other LMICs(Brauer et al., 2019; Landrigan et al., 2018), current studies of air pollution and infant mortality in LMICs are mainly based on annual average ambient $PM_{2.5}$ prediction models(Shaddick et al., 2018; van Donkelaar et al., 2016). Such exposure estimates may be prone to errors or misclassifications, especially for child mortality

which has been associated with monthly or daily changes in air pollution levels(Scheers Hans et al., 2011; Son et al., 2017). The expansion of ground $PM_{2.5}$ monitoring stations during recent years in India, and the development of spatial $PM_{2.5}$ prediction models, now makes it possible to estimate a more spatiotemporally resolved mapping of ambient $PM_{2.5}$ air pollution exposure throughout India.

The primary objective of our study was to evaluate the association between child mortality and ambient $PM_{2.5}$ air pollution exposure during in-utero and post-delivery lifetime periods among a nationally representative, retrospective cohort based on maternal-reported data in the Demographic and Health Survey (DHS, also known as National Family Health Survey (NFHS) in India). To estimate time-dependent ambient $PM_{2.5}$ air pollution exposure for the cohort, we developed a monthly ambient $PM_{2.5}$ 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.

## 5.2 Methods

### 5.2.1 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 (29 states and 5 Union Territories, excluding Andaman and Nicobar Islands and Lakshadweep Union Territories), whose mothers participated in the 2015 – 2016 DHS/NFHS in India between January 2015 – November 2016. The design of the DHS/NFHS has been described elsewhere(IIPS/India & ICF, 2017). Briefly, DHS/NFHS 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 their children. When multiple children were born to the woman respondent, the mother answered separately for each child. We only included children born to women no more than 5 years preceding the survey. Three tiers of all-cause mortality – including neonatal mortality (death occurred less than the first month of age), infant mortality (death occurred less than the 12 months of age), and child mortality (death occurred less than

60 months of age) – were outcome variables in our study. Child birth month, child death month (if the child died), and other potential confounders were reported by mothers. In-utero and post-delivery $PM_{2.5}$ air pollution, temperature, and rainfall data were linked to each child based on geo-coordinates of households at the level of a cluster (a primary sampling unit (PSU) or a segment of PSU, of around 100 – 150 households) provided by DHS. 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 survey or passing the at-risk age as the censoring event. We included prespecified variables that are potential confounders from DHS/ NFHS data as suggested by previous studies(Heft-Neal et al., 2018; Spears et al., 2019; Subramanian et al., 2009). These include child's sex, birth month and year, birth order, location of birth (institutional birth or not), whether multiple birth, mother's age at child birth, mother's height, marital status, education (whether above secondary level), whether maternal smoking, wealth index, whether household members smoke (second-hand smoke exposure), whether households were urban or rural, geographical region/zonal council (Northern, Central, North Eastern, Eastern, Western and Southern), primary cooking fuel, and toilet facilities. We dichotomized toilet facilities as whether they are safely managed based on WHO/UNICEF Joint Monitoring Program (Croft et al., 2018; World Health Organization & United Nations Children's Fund (UNICEF), 2017), and dichotomized cooking fuel as clean (electricity, liquefied petroleum gas (LPG), natural gas and biogas) or solid biomass and others (kerosene, coal, wood, straw, agriculture crop residue, dung). Monthly air temperatures at 2-meters (T2M) were extracted from MERRA-2 (Modern-Era Retrospective analysis for Research and Applications) data(Koster, 2015), and monthly rainfall precipitation in mm/month were extracted from CHIRPS (Climate Hazards Group InfraRed Precipitation with Station) data(Funk et al., 2015), and were matched with the geo-coordinates of households (in the same way as was performed for $PM_{2.5}$ air pollution exposure). We did not adjust for low birth weight (LBW), as we found this variable was likely to be an intermediate variable between air pollution and childhood death, as air pollution predicted LBW, and LBW predicted child death(Hernández-Díaz et al., 2006).

As this study only analyzed 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 2019 – September 2019.

### 5.2.2 Ambient $PM_{2.5}$ Air Pollution Exposure

In this study, we developed a random forest (RF) model to predict monthly ambient $PM_{2.5}$ concentrations at 0.1° spatial resolution (approximately 11km at equator) over the entire Indian subcontinent, from 2009 to 2018 (10 years). This approach has been shown to have reliable historical prediction (hindcast) capabilities in previous studies and has good agreement with ground measurements in East Asia and Latin America(Vu et al., 2019; Qingyang Xiao et al., 2018). Briefly, we used multiple satellite retrieved aerosol, $NO_2$, and fire spot data products; global meteorological and aerosol reanalysis model; land use information; and population density as inputs for the model [Appendix part 1]. We trained the RF model on gridded monthly mean $PM_{2.5}$ concentrations from 2017 – 2018 (N = 1 446) collected by 134 ground air pollution monitoring stations with the above input predictors. We evaluated model performance by comparing predicted monthly $PM_{2.5}$ with observations, using out-of-sample 10-fold cross validation. We also evaluated hindcast performance of the model using monthly measurements in 2015 – 2016 (N = 456) from 57 stations. 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 10 years, and compared our estimates with other published databases(Balakrishnan, Dey, et al., 2018; Shaddick et al., 2018; van Donkelaar et al., 2016) [Appendix part 2]. Predicted ambient monthly $PM_{2.5}$ levels were matched to each child with geo-coordinate in the cluster level. 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 define it as the average of $PM_{2.5}$ concentrations from the calendar month of childbirth through a given month the child is at risk, until the month of death or censoring.

### 5.2.3 Statistical Analysis

We conducted Cox regression models with follow-up time 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 $PM_{2.5}$ air pollution exposure to estimate adjusted hazard ratios (HRs) relating child mortality risk with 10-$\mu g/m^3$ increases of $PM_{2.5}$ exposure. In this model, the effects of $PM_{2.5}$ exposure during in-utero and post-delivery periods were adjusted for each other. We also conducted single-exposure models with $PM_{2.5}$ exposures from either the in-utero or post-delivery periods for comparison. Prespecified time-independent predictors and time-dependent monthly temperature and rainfall precipitation were included in the Cox regression model [appendix part 3]. We stratified the analysis by childbirth month and year, child's gender, and geographical regions (927 strata) to account for spatiotemporal differences in child survival. In addition, we included the sample weight from DHS/NFHS survey in the Cox regression. (**Equation 5-1**)

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

**Equation 5-1**

$h_g(t)$ 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; $\boldsymbol{Z}$ is the vector containing time-independent variables, and $\boldsymbol{X(t)}$ are the time-dependent variables. To evaluate whether effects of $PM_{2.5}$ differed across covariates, we conducted subgroup stratified analysis of the two-exposure model. Lastly, to examine the robustness of our results, we conducted additional sensitivity analysis for different model setups and accounted for potential correlation of outcomes within household and cluster [appendix part 3]. We also tested the proportional hazards assumption in our main model, and found it was not violated for in-utero nor post-delivery $PM_{2.5}$ exposure. Stata (version 14) and R (version 3.6.0) were used in statistical analysis.

## 5.3 Results

The $PM_{2.5}$ prediction model characterized monthly $PM_{2.5}$ levels well with cross-validation (CV) R-square 0.82 at monthly level and with annual CV root mean squared error (RMSE) of 10.9 $\mu g/m^3$ in 2017 – 2018. Compared to historical measurements in 2015 – 2016, our model provides good predictions of monthly $PM_{2.5}$ with R-square of 0.82 [appendix part 4]. **Figure 5-1** shows the predicted annual $PM_{2.5}$ in India in 2018 and the annualized change of $PM_{2.5}$ over 10 years. 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 (over 120 $\mu g/m^3$) were in the Indo-Gangetic Plain (IGP), covering the states of Haryana, Uttar Pradesh, and Bihar and the National Capital Territory of Delhi. These values are three times above 40 $\mu g/m^3$, the recommended limit set by National Ambient Air Quality Standards of India(Gautam, 2019). Details of predicted $PM_{2.5}$ in India over 10 years are in the [appendix part 4, **Table 5-4S, Table 5-5S, Table 5-6S Figure 5-2S**]. We observed increased ambient $PM_{2.5}$ over 10 years in most areas, with the strongest increase of ambient $PM_{2.5}$ concentrations at around 1 $\mu g/m^3$ per year in the IGP region and along the west coast of Western region; while the Northeastern region and north part of the Northern region show decrease in ambient $PM_{2.5}$ levels.

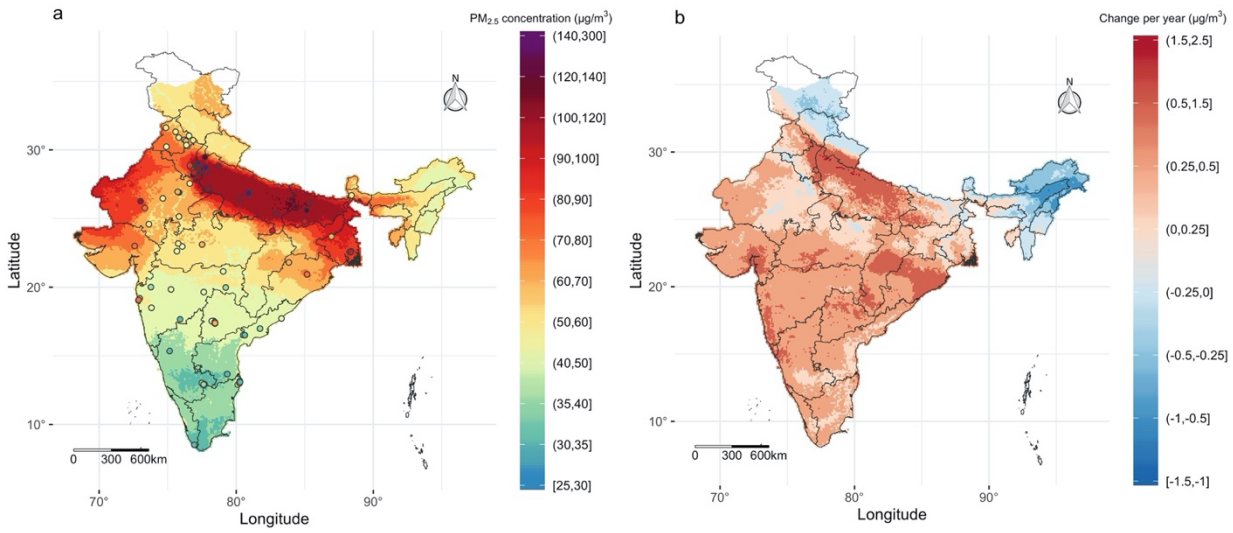



**Figure 5-1. Spatial Patterns of Annual $PM_{2.5}$ Concentration**. a) Model-predicted annual $PM_{2.5}$ concentrations with measurements from 133 ground-based monitoring stations in subcontinental India in 2018. b) Annualized change rate of $PM_{2.5}$ concentrations between 2009 – 2018.

This study included data from 259 627 live-birth children born 5 years preceding the survey from 699 686 women in DHS data. We excluded 6 839 (2.6%) children living in Andaman and Nicobar Islands and Lakshadweep Union Territories, for not living with their mothers or because of missing variables [appendix part 5]. A total of 252 788 children born January 2010 – November 2016 were included. In the follow-up time of 7 447 724 child-months, 11 559 deaths were reported in children under five years of age by their mothers. Among these deaths, 7 520 deaths occurred before the first month of life; 10 862 deaths occurred during the first year of life. 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 5-1**, stratified by 9-month in-utero average $PM_{2.5}$ air pollution exposure, at levels less than 49.7 $\mu g/m^3$, 49.7 to less than 62.4 $\mu g/m^3$, 62.4 to less than 90.3 $\mu g/m^3$, and equal or more than 90.3 $\mu g/m^3$. In general, children with highest quartile of ambient $PM_{2.5}$ air pollution levels tended to reside in rural areas, were born to mothers with shorter stature and lower educational, and lived in households with a lower wealth index. Central and eastern Indian states have higher estimated ambient $PM_{2.5}$ levels and also have lower social-economic status compared to states in other geographical zones. Most children were born between 2013 – 2014, with the least number of children born in 2016 (**Table 5-2**). The correlation between average in-utero and post-delivery lifetime exposure for the same child was 0.74.

**Table 5-1. Unweighted Characteristics of Children Less than 5 Years of Age by four quartiles of 9-month In-utero $PM_{2.5}$ Air Pollution Exposure**

| | Entire Cohort | Average 9-month in-utero $PM_{2.5}$ Exposure Quartile | | | |
|---|---|---|---|---|---|
| Average 9-month in-utero $PM_{2.5}$ exposure, $\mu g/m^3$ | | Quartile 1<br>20.9 – 49.7 | Quartile 2<br>49.7 – 62.4 | Quartile 3<br>62.4 – 90.3 | Quartile 4<br>90.3 – 153.5 |
| Children (n) | 252,788 | 63,197 | 63,197 | 63,197 | 63,197 |
| Total follow-up in month | 7,447,724 | 1,860,527 | 1,891,613 | 1,873,198 | 1,822,386 |
| # of under-5 child death | 11,559 | 2,187 | 2,688 | 3,001 | 3,683 |
| All-cause child mortality rate (per 1000 child-years) | 18.62 | 14.11 | 17.05 | 19.22 | 24.25 |
| Average 9-month in-utero $PM_{2.5}$ exposure range, $\mu g/m^3$ | 20.9 – 153.5 | 20.9 – 49.7 | 49.7 – 62.4 | 62.4 – 90.3 | 90.3 – 153.5 |

| | | | | | |
|---|---|---|---|---|---|
| Average post-delivery lifetime $PM_{2.5}$ exposure until death or censoring, mean(SD), $\mu g/m^3$ | 73.7 (29.8) | 46.6 (10.6) | 57.1 (11.5) | 83.6 (25.6) | 107.6 (20.1) |
| Child covariates† | | | | | |
| Sex | | | | | |
| Male, % | 52.1 | 51.6 | 51.8 | 52.4 | 52.3 |
| Low birth weight, %§ | 13.1 | 13.8 | 13.7 | 14.1 | 10.98 |
| Birth order, % | | | | | |
| First | 36.9 | 41.0 | 37.5 | 38.2 | 30.7 |
| Second | 30.7 | 34.3 | 30.8 | 30.4 | 27.3 |
| Third | 16.1 | 14.1 | 16.0 | 15.8 | 18.4 |
| Forth and above | 16.3 | 10.6 | 15.7 | 15.6 | 23.6 |
| Multiple birth, % | 1.65 | 1.59 | 1.55 | 1.78 | 1.67 |
| Birth at institution, % | 75.3 | 81.6 | 76.3 | 74.8 | 68.5 |
| Second-hand smoke exposure, % | 56.6 | 51.8 | 61.1 | 55.1 | 58.7 |
| Maternal covariates† | | | | | |
| Age at child birth, % | | | | | |
| <19 | 12.2 | 12.5 | 11.3 | 13.8 | 11.2 |
| 20 – 24 | 43.4 | 43.1 | 42.2 | 45.0 | 43.2 |
| 25 - 29 | 28.4 | 28.7 | 29.2 | 27.0 | 28.8 |
| >=30 | 16.0 | 15.7 | 17.3 | 14.2 | 16.9 |
| Mothers height in cm, % | | | | | |
| < 149.9 | 37.7 | 33.3 | 34.2 | 36.9 | 47.9 |
| 150 – 154.9 | 33.7 | 35.2 | 34.4 | 33.3 | 31.8 |
| 155 – 159.9 | 19.6 | 21.6 | 21.6 | 20.5 | 14.6 |
| >160 | 8.7 | 9.9 | 9.8 | 9.3 | 5.7 |
| Marital status, % | | | | | |
| Not married | 1.4 | 1.8 | 2.0 | 1.1 | 0.9 |
| Married | 98.6 | 98.2 | 98.0 | 98.9 | 99.1 |
| Education, % | | | | | |
| Below secondary level | 31.5 | 19.0 | 27.9 | 33.5 | 45.5 |
| Above secondary | 68.5 | 81.0 | 72.1 | 66.5 | 54.5 |
| Currently smoke % | 10.0 | 13.6 | 14.7 | 6.9 | 4.9 |
| Household covariates† | | | | | |
| Cook with clean fuel, % | 29.7 | 41.8 | 28.7 | 24.4 | 23.7 |
| Access to improved toilet, % | 41.5 | 48.9 | 43.7 | 39.4 | 33.8 |
| Urban area, % | 23.6 | 29.6 | 23.1 | 22.0 | 20.0 |
| Zone area, % | | | | | |
| Northern | 16.6 | 4.0 | 25.0 | 60.6 | 6.8 |
| North Eastern | 13.7 | 22.1 | 23.4 | 9.4 | 0 |
| Central | 31.8 | 15.3 | 35.6 | 23.2 | 53.1 |
| Eastern | 21.1 | 4.1 | 8.1 | 32.0 | 40.1 |
| Western | 7.2 | 16.3 | 7.6 | 4.8 | 0 |
| Southern | 9.6 | 38.3 | 0.3 | 0 | 0 |
| Wealth quartiles, % | | | | | |
| First (lowest) | 25 | 15.0 | 22.3 | 26.7 | 36.0 |
| Second | 25 | 23.3 | 26.5 | 24.6 | 25.6 |
| Third | 25 | 31.9 | 26.0 | 22.5 | 19.6 |
| Forth | 25 | 29.8 | 25.2 | 26.2 | 18.9 |
| Meteorological covariates‡ | | | | | |
| Temperature in °C, mean(SD) | 27.1 (6.4) | 26.7 (4.9) | 24.7 (7.8) | 27.6 (6.0) | 39.4 (5.5) |

| | | | | | |
|---|---|---|---|---|---|
| Monthly Total Rainfall in cm, mean(SD) | 103 (114) | 115 (144) | 98 (134) | 90 (145) | 110 (151) |

§Assume children with missing LBW values are not LBW infants.

†Child, maternal and household level covariates collected at the end of the follow-up, during the Demographic and Health Survey/National Family and Health Survey (DHS/NFHS) interview between 2015 – 2016.

‡Meteorological covariates were post-delivery lifetime average estimates.

In the two-exposure Cox regression model, both in-utero and post-delivery lifetime average $PM_{2.5}$ air pollution exposures are significantly associated with increases in child mortality, infant mortality and neonatal mortality, after controlling for covariates (**Table 5-3, Figure 5-2**). A 10 $\mu g/m^3$ increase of in-utero $PM_{2.5}$ exposure is associated with an increase in neonatal mortality (HR: 1.018, 95% CI: 1.001, 1.036), infant mortality (HR: 1.021, 95% CI: 1.006, 1.037) and child mortality (HR: 1.023, 95% CI: 1.000, 1.038). A 10 $\mu g/m^3$ increase of post-delivery average $PM_{2.5}$ exposure is 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 single-exposure Cox regression models, with either in-utero $PM_{2.5}$ exposure or post-delivery average $PM_{2.5}$ exposure, we found stronger $PM_{2.5}$-mortality HRs compared to HRs in the two-exposure model. This could be explained by the correlation between the two exposure estimates. Single exposure models have similar AICs as two-exposure models [Appendix **Table 5-8S**].

**Table 5-3. Hazard Ratio of Mortality per 10 $\mu g/m^3$ Increase of Ambient $PM_{2.5}$ Air Pollution Exposure**§

| | | All-cause neonatal mortality | All-cause infant mortality | All-cause child mortality |
|---|---|---|---|---|
| No. of death | | 7,520 | 10,862 | 11,559 |
| Single-exposure models† | In-utero $PM_{2.5}$ | 1.032 (1.019, 1.045) | 1.033(1.021, 1.044) | 1.033 (1.022, 1.044) |
| | Post-delivery average $PM_{2.5}$ | 1.025 (1.015, 1.035) | 1.025 (1.016, 1.034) | 1.025 (1.016, 1.033) |
| | In-utero $PM_{2.5}$ | 1.017 (1.003, 1.030) | 1.015 (1.003, 1.027) | 1.013 (1.001, 1.026) |

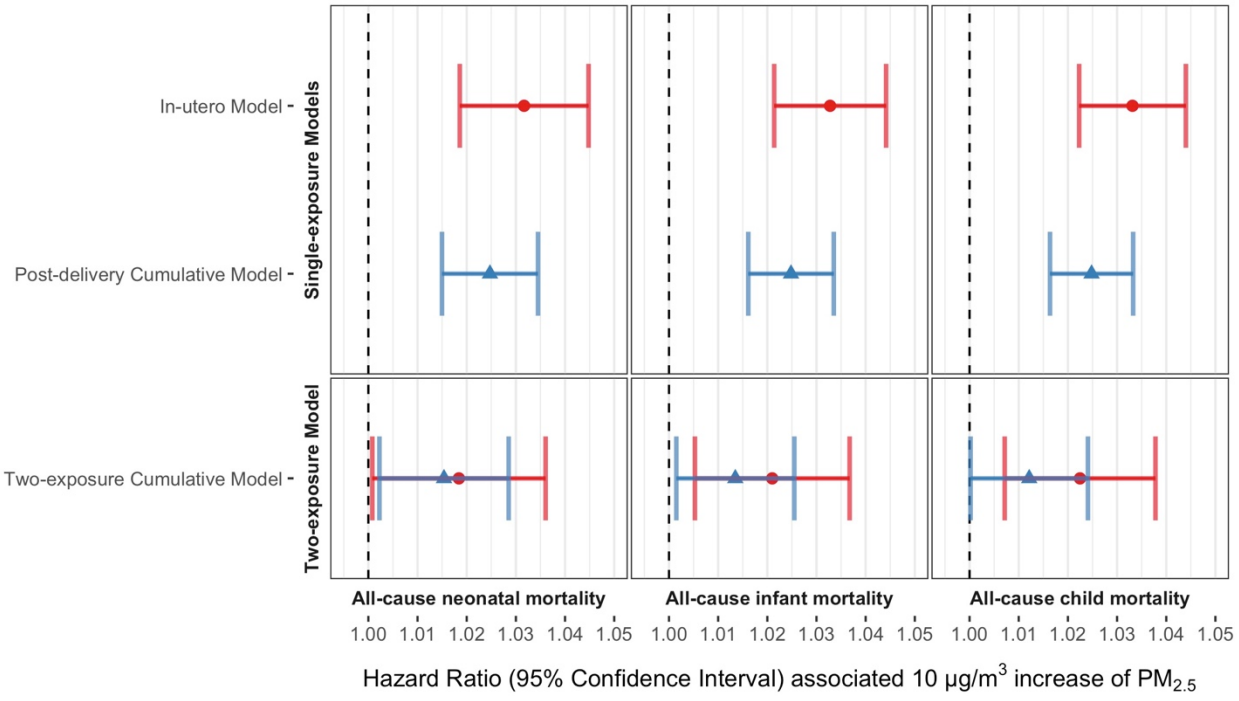
| | | | | |
|---|---|---|---|---|
| Two-exposure model‡ | Post-delivery average $PM_{2.5}$ | 1.017 (1.003, 1.030) | 1.012 (1.001, 1.023) | 1.011 (1.000,1.022) |

§ Adjusted hazard ratios and 95% confidence intervals for all-cause neonatal mortality, infant mortality and child mortality are shown for 10 μg/m³ increase of ambient $PM_{2.5}$ air pollution during 9-month in-utero period before child birth and post-delivery periods. Models stratified on child sex, birth month and year, geographical zone, adjusted for birth order, multi-birth, birth location, mother's age, height, marital, education, maternal smoking, household wealth, second hand smoking, cooking fuel, improved toilet, urban or rural location of households, monthly temperature, monthly precipitation.

†Single-exposure model include either one of $PM_{2.5}$ exposure during 9-month in-utero period before child birth or post-delivery average until death, or censoring.

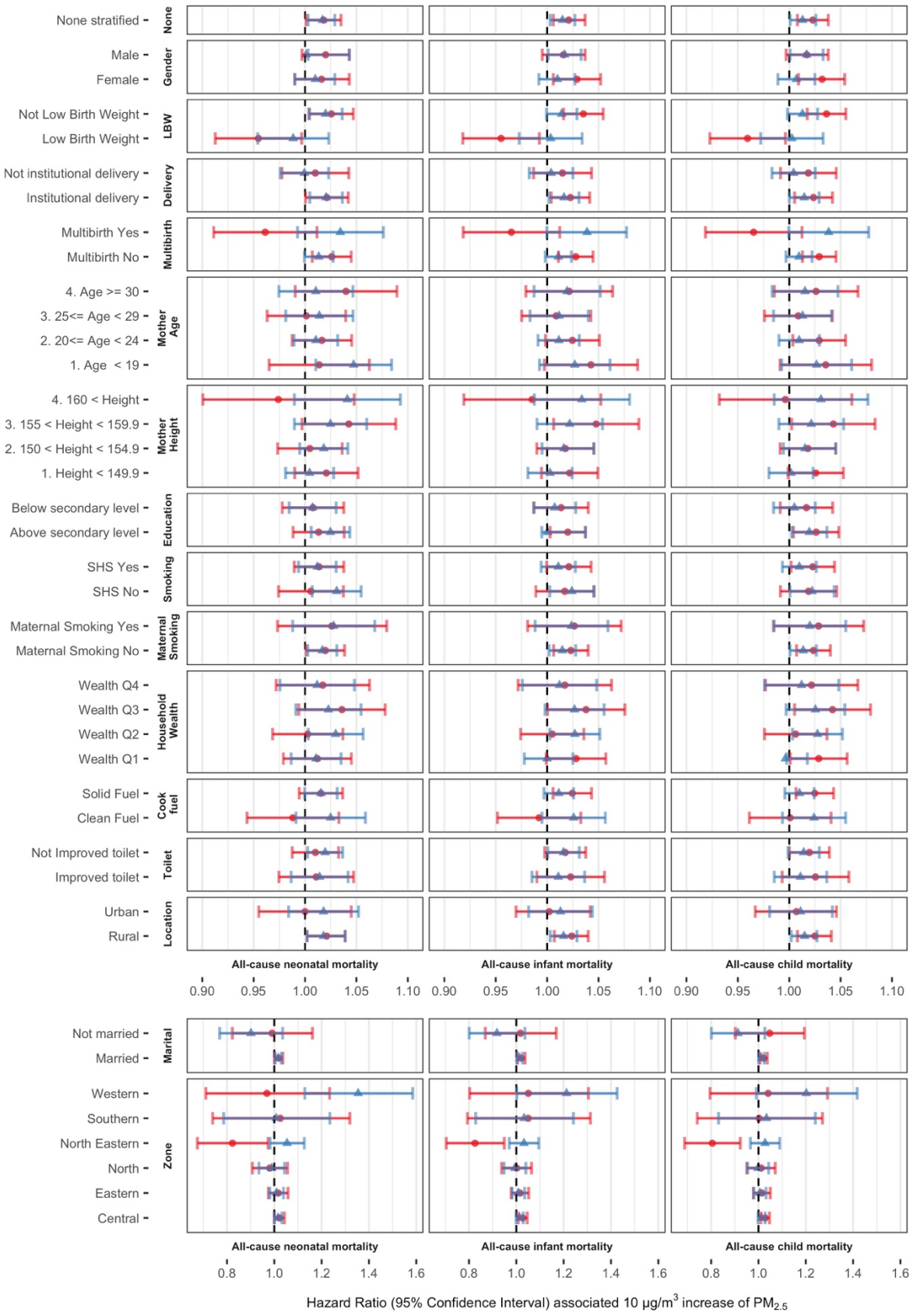
‡ Two-exposure model includes both ambient $PM_{2.5}$ exposure during 9-month in-utero period before child birth and post-delivery lifetime average until death, or censoring.

**Figure 5-2. Hazard Ratio of Mortality per 10 μg/m³ Increase of Ambient $PM_{2.5}$ Air Pollution Exposure.** Adjusted hazard ratios and 95% confidence intervals for all-cause neonatal mortality, infant mortality, and child mortality are shown for 10 μg/m³ increases in ambient $PM_{2.5}$ air pollution. All models have the same specification, except for $PM_{2.5}$ exposure. Single-exposure models only included either $PM_{2.5}$ exposure during the in-utero $PM_{2.5}$ period or during the post-delivery lifetime period. Two-exposure model include both $PM_{2.5}$ exposure during the in-utero and post-delivery lifetime periods. Red circles (●) represent effects of in-utero $PM_{2.5}$ exposure, and blue triangles (▲) represent effects of post-delivery lifetime average $PM_{2.5}$ exposure.



We found little effect modification by covariates with the exception of multi-birth and geographical locations (**Figure 5-3**). The effects of in-utero $PM_{2.5}$ on mortality tend to be lower for multi-birth children as compared to singleton children. This could be due to random error from small sample size of multi-birth children (n = 4 172, 1.7%). For low birth weight (LBW) children, we observed in-utero $PM_{2.5}$-mortality HRs are less than one. This is similar to the protective effect of maternal smoking on infant mortality for LBW children observed in the United States as a result of selection bias(Hernández-Díaz et al., 2006), as LBW could mediate in-utero $PM_{2.5}$ exposure and child mortality [appendix **Table 5-9S, Table 5-10S**]. Additionally, we observed some heterogeneity of $PM_{2.5}$-mortality HRs across different geographical locations, such as in North Eastern states the in-utero $PM_{2.5}$-mortality HRs less than one. Since there is no air pollution measurement in that region, additional validation of our predicted $PM_{2.5}$ air pollution in North Eastern states and investigation of its health effects is needed.

**Figure 5-3. Hazard Ratios of Mortality per 10 μg/m³ Increase of Ambient $PM_{2.5}$ air pollution Exposure, According to Subgroup Analysis for Two-exposure Model.** Adjusted hazard ratios and 95% confidence intervals for all-cause neonatal mortality, infant mortality and child mortality are shown for 10 μg/m³ increases in ambient $PM_{2.5}$ air pollution exposure during both in-utero and post-delivery lifetime periods, stratified on subgroups by individual and household level covariates. All model specifications are the same for main analysis except for subgroup analysis of low birth weight (LBW), which is not included in the main analysis. Red circles (●) represent effects of in-utero $PM_{2.5}$ exposure, and blue triangle (▲) represent effects of post-delivery lifetime average $PM_{2.5}$ exposure.


None stratified
None
Male
Female
Gender
Not Low Birth Weight
Low Birth Weight
LBW
Not institutional delivery
Institutional delivery
Delivery
Multibirth Yes
Multibirth No
Multibirth
4. Age >= 30
3. 25<= Age < 29
2. 20<= Age < 24
1. Age < 19
Mother Age
4. 160 < Height
3. 155 < Height < 159.9
2. 150 < Height < 154.9
1. Height < 149.9
Mother Height
Below secondary level
Above secondary level
Education
SHS Yes
SHS No
Smoking
Maternal Smoking Yes
Maternal Smoking No
Maternal Smoking
Wealth Q4
Wealth Q3
Wealth Q2
Wealth Q1
Household Wealth
Solid Fuel
Clean Fuel
Cook fuel
Not Improved toilet
Improved toilet
Toilet
Urban
Rural
Location
All-cause neonatal mortality
All-cause infant mortality
All-cause child mortality
0.90 0.95 1.00 1.05 1.10
Not married
Married
Marital
Western
Southern
North Eastern
North
Eastern
Central
Zone
All-cause neonatal mortality
All-cause infant mortality
All-cause child mortality
0.8 1.0 1.2 1.4 1.6
Hazard Ratio (95% Confidence Interval) associated 10 µg/m3 increase of PM2.5

Our results are robust to different variable selections, unweighted analysis, and Cox regressions with generalized estimating equation to account for correlation of outcomes within households and clusters, and inclusion time-dependent post-delivery monthly $PM_{2.5}$ exposure to replace post-delivery lifetime $PM_{2.5}$ exposure [Appendix **Figure 5-6S**].

## 5.4 Discussion

In a large, nationally representative retrospective cohort of children under 5 years of age in India, we found a consistent elevated risk for child mortality associated with exposure to ambient $PM_{2.5}$ air pollution during both in-utero and post-delivery periods. An increase of 10 $\mu g/m^3$ of $PM_{2.5}$ exposure during the 9-month in-utero period and post-delivery period were associated with 2.3% (95% CI: 0.8% - 3.8%) and 1.3% (95% CI: 0.1% - 2.6%) increases of all-cause child mortality, respectively, in our two-exposure model. The effects of both these correlated exposures increased in single-exposure models. Similar results have been observed for infant and neonatal mortality. To our knowledge, this is the first cohort study conducted in India to assess the effects of ambient $PM_{2.5}$ air pollution on child survival, in a national representative population.

To estimate monthly $PM_{2.5}$ air pollution concentrations, we developed a model based on multiple inputs from satellite observations, meteorological datasets, and land use information between 2009 – 2018. We identified disproportionally high levels of $PM_{2.5}$ in Indo-Gangetic Plain region, where population density is high. Additionally, we found increasing levels of ambient air pollution over the past 10 years, emphasizing an urgent need to control ambient air pollution in India. Children under 5 years included in our study have high exposure levels to ambient $PM_{2.5}$ air pollution, with in-utero, post-delivery lifetime means (SD) of $PM_{2.5}$ exposure as 71.1 (28.2) $\mu g/m^3$ and 73.7 (29.8) $\mu g/m^3$, respectively. This is nearly twice the National Ambient Air Quality Standards of India (40 $\mu g/m^3$) (Gautam, 2019) and more than seven times the World Health Organization Air Quality Guidelines (10 $\mu g/m^3$) (World Health Organization, 2005b). The difference between in-utero $PM_{2.5}$ levels in highly exposed children ($95^{th}$ percentile, 130.0 $\mu g/m^3$) and the least exposed children ($5^{th}$ percentile, 35.8 $\mu g/m^3$) was 94.2 $\mu g/m^3$. Based on mortality risk estimated from our Cox regression model with of in-utero

$PM_{2.5}$ exposure mean level (71.1 μg/m$^3$) and standard population attributable fraction calculations, we estimate that 18.0% (95% CI: 12.6% - 23.1%) of mortality could be reduced if their exposure were reduced to 10 μg/m$^3$. Given the high levels of ambient $PM_{2.5}$ air pollution throughout India, reduction of ambient air pollution could provide a relatively large reduction in child mortality and substantial health benefits.

The child mortality-HRs estimates per 10 μg/m$^3$ of $PM_{2.5}$ exposure in our study are in the range of 1-3%, lower than previous estimates in LMICs. In Sub-Saharan Africa, Heft-Neal and colleagues found a 9% increase in risk of infant mortality associated with a 10 μg/m$^3$ increase in annual $PM_{2.5}$ concentration(Heft-Neal et al., 2018); and Loomis and colleagues found a 7% increase of infant mortality associated with 10 μg/m$^3$ increase of $PM_{2.5}$ concentration during the 3-5 days before death in Mexico city(Loomis et al., 1999). In our study, however, the $PM_{2.5}$ exposure range is three times higher than those in previous studies, and the HRs per interquartile range (IQR) change of $PM_{2.5}$ are in the range of 5% - 12%, more similar to earlier findings. One study which assessed the early life ambient $PM_{2.5}$ exposure effect on child mortality in 43 LMICs did not find a significant association between air pollution and child mortality(Goyal et al., 2019). This could be because this study used annual $PM_{2.5}$ exposure as a proxy for exposure at less than one year, and thus introduced measurement error which may have biased results towards the null(Zeger S L et al., 2000). Another study conducted in Beijing using a time-series design also did not find a significant association between current-month $PM_{2.5}$ exposure and infant mortality(Wang et al., 2019).

Our study has several strengths. First, we use machine-learning to develop a high-quality model of ground station measured ambient $PM_{2.5}$ air pollution concentrations that has monthly temporal resolution and high spatial resolution that can reflect seasonal differences in ambient air pollution. These estimates of ambient $PM_{2.5}$ air pollution level can be potentially applied in future health impact assessment and epidemiological analysis in India (SI files). Second, we used a large, retrospective and nationally representative cohort of children under 5 in India based on a well-documented and well-conducted DHS/NHFS survey. Third, we applied Cox regression with time-dependent air pollution, temperature and rainfall precipitation on child mortality using

calendar month as time variable. This approach allows us to examine the longitudinal association between air pollution and child mortality, and compare post-delivery $PM_{2.5}$ exposure for children with same follow-up time. This approach will largely reduce exposure misclassification compared to previous studies(Goyal et al., 2019, p.; Heft-Neal et al., 2018), which were based on yearly $PM_{2.5}$ average to assess a binary outcome of child death at the end of one year of age. Additionally, our analysis controlled for individual-level, household-level, and cluster level risk factors of child mortality. Finally, the positive $PM_{2.5}$-mortality HRs are consistent in single-exposure and two-exposure models and are robust in additional sensitivity analysis. The estimated $PM_{2.5}$-mortality HRs were not highly sensitive in sub-cohorts with stratified covariates, except for multi-birth and Northeastern region.

Our study also has some limitations. First, no national $PM_{2.5}$ air pollution monitoring data were available during the study period between 2009 – 2014, so we lacked ground level $PM_{2.5}$ measurement data before 2015 with which to validate our $PM_{2.5}$ prediction model. Similar to previous $PM_{2.5}$ prediction models (Qingyang Xiao et al., 2018), our model tended to underestimate some high $PM_{2.5}$ values [Appendix **Figure 5-3S**]. Secondly, we only used total ambient $PM_{2.5}$ mass concentration as our exposure without estimation of different $PM_{2.5}$ components, and did not include the impact of other ambient air pollutants 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. Child 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 children were born. Fourth, we did not have children's gestational age at birth and assigned in-utero $PM_{2.5}$ exposure levels as 9-month average prior to the birth, which may not be the case for preterm infants who have a shorter in-utero exposure period. Due to this data collection limitation, our study could suffer from recall bias. However, we believe this will would lead to non-differential misclassification of outcomes and covariates, which could bias our results towards the null. Lastly, as with all observational studies, our study may have potential unmeasured confounders inadequately controlled in this analysis. This calls for more future prospective cohort studies(Balakrishnan,

Sambandam, Ramaswamy, et al., 2015), including intervention trials, to investigate air pollution effects on child health.

## 5.5 Conclusion

In conclusion, this study found increased mortality risk associated with ambient $PM_{2.5}$ air pollution during 9-month in-utero and post-delivery periods for children under 5 years of age. This study, based on a nationally representative retrospective cohort in India, substantially expands evidence that in-utero and post-delivery air pollution exposure contributes to child mortality in developing countries. Given the high levels of the ambient $PM_{2.5}$ air pollution throughout India, expanding air pollution monitoring stations, adding more epidemiological research, and making a substantial effort to reduce ambient air pollution and early life exposures are all needed.

Appendix for Chapter 5

## 1. Data inputs for Ambient $PM_{2.5}$ Prediction Model

### 1.1 Ground $PM_{2.5}$ Measurement

We collected daily ambient $PM_{2.5}$ concentrations over 2015 – 2018 from air monitoring stations across India and averaged them into monthly ambient $PM_{2.5}$ concentrations over this period. The daily ambient $PM_{2.5}$ measurement data were downloaded from Central Pollution Control Board (CPCB, https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing). We also collected hourly $PM_{2.5}$ measurements from five US Embassy and Consulates to India from AirNow.gov (https://airnow.gov/) during the same period. We excluded the daily average $PM_{2.5}$ concentrations outside of 3 standard deviations of the mean for the log-transformed daily $PM_{2.5}$ concentrations (0.05% of total data, outside of range 4.2 ug/$m^3$ - 845 ug/$m^3$). We believe these measurements may be implausible and may be due to measurement or other errors. For hourly $PM_{2.5}$ concentration data from US consulates or embassies, we removed data that lacked valid quality checks. Daily concentrations with less than 18 hourly measurements were excluded (2.2%).

**Table 5-1S** shows the number of stations and days of measurements in each year from 2015 – 2018. Daily average $PM_{2.5}$ measurements from stations within the same grid were averaged, resulting in 56,834 grid-day $PM_{2.5}$ measurements. We averaged gridded daily $PM_{2.5}$ measurements into gridded monthly $PM_{2.5}$ averages if more than 15 days have daily measurements for a given grid cell in a given month. We ended up with 1446 grid-months of $PM_{2.5}$ measurement during 2017 – 2018 for model development and 465 grid-months of $PM_{2.5}$ measurement during 2015 – 2016 for hindcasting.

1.2 Satellite Data

We downloaded the Moderate Resolution Imaging Spectroradiometer (MODIS) Collection 6 level-2 aerosol optical depth (AOD) products (MOD04_L2 and MYD04_L2) at 10 km resolution collected by Aqua and Terra satellites, from Distributed Active Archive Center (https://ladsweb.modaps.eosdis.nasa.gov/) during years 2009 – 2018(MODIS Atmosphere Science Team, 2015; MODIS Science Team, 2014). We used Deep Blue (DB) and Deep Blue-Dark Target combined parameters (Combined) from MODIS retrial and assigned and averaged the centroid of each retrial to created $0.1° \times 0.1°$ grid cells and calculated monthly average AOD value for DB and Combined algorithms. MODIS DB algorithm can provide high quality retrievals over bright land areas such as urban regions, where most air pollution monitors are located, while the DT algorithm works better over vegetation covered land. DB-DT combined AOD parameter only used high quality retrievals from DB and DT algorithms, and excluded retrievals when cloud coverage was high during the rainy season. Similar to AOD missing patterns in East Asia(Q. Xiao et al., 2016), we found that 56% – 72% of the monthly AOD data is missing in India between 2009 – 2018.

**Table 5-2S** summarizes the AOD data missing patterns in 2017. Similar to a previous study(Qingyang Xiao et al., 2018), we imputed missing AOD patterns at the monthly level, using random forest algorithm. **Figure S1** shows combined AOD parameter and DB AOD parameter spatial distribution after gap-filling in July 2017.

Active Fire Data were obtained from the Fire Information for Resource Management System (FIRMS, https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms), using Collection 6 NRT Hotspot /

Active Fire Detections product (MCD14DL)(LANCE FIRMS, 2016), from 2009 - 2018. This active fire product is provided from MODIS sensor aboard the Aqua and Terra satellites. We processed the active fire points and assign each active fire event to each grid cell using a 15km buffer and aggregated the number of active fire events in each 0.1° grid cell each month.

Normalized Difference Vegetation Index (NDVI) data were obtained from MODIS Vegetation Indices product (MOD13C1) with 16-day temporal resolution and 0.05-degree spatial resolution from Distributed Active Archive Center (https://ladsweb.modaps.eosdis.nasa.gov/), for 2009 - 2018. This NDVI product incorporates a data fill strategy, based on historic data records, producing a continuous NDVI estimate(K. Didan, 2015, p. 1). We conducted oversampling and calculated the monthly average NDVI value for each 0.1° grid cell, based on original 0.05-degree 16-day data. The aerosol absorbing index (AAI) in visible light and UV light and tropospheric $NO_2$ density data were obtained from Ozone Monitoring Instrument (OMI), downloaded from Goddard Earth Sciences Data and Information Service Center (https://mirador.gsfc.nasa.gov/). We used parameters *AerosolIndexUV* and *AerosolIndexVIS* from OMI Aerosol Extinction Optical Depth and Aerosol types level 2 data (OMAERO) and *UVAerosolIndex* from OMI Near-UV Aerosol Absorption and Extinction Optical Depth and Single Scatter Albedo level 2 data (OMAERUV) to represent aerosol absorbing index (AAI). We extracted parameter *ColumnAmountNO2Trop* from the OMI NO2 level 2 data (OMNO2) to represent tropospheric $NO_2$ density. We excluded retrievals with the cross-track anomaly flag as nonzero, due to a row anomaly starting from 2007(Krotkov et al., 2019; Stein-Zweers & Veefkind, 2019). Then we conducted oversampling and calculated the monthly average AAI and tropospheric $NO_2$ density for each 0.1° grid cell.

1.3 Meteorological Reanalysis and Aerosol Simulation Data

Daily meteorological data and aerosol diagnostics were obtained from Goddard Earth Observing System Data Assimilation System (GEOS 5/MERRA-2, https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/ ), at 0.5° latitude × 0.625° longitude resolution, from 2009 - 2018(Koster, 2015). Air temperature at 2-meter (T2M), eastward wind speed at 10-meter (U10M), northward wind speed at 10-meter (V10M), specific humidity at 2-meter (QV2M), surface pressure (PS), and total liquid water precipitation (TQL), total ice precipitation (TQI) and total water vapor precipitation (TQV) were extracted from 2d assimilated state at hourly intervals (inst1_2d_asm_Nx). Planetary boundary height (PBLH) data were extracted from 2d time-averaged surface flux diagnostics at hourly intervals (tavg1_2d_flx_Nx). Aerosol diagnostics data including organic carbon surface mass concentration (OCSMASS), $SO_2$ surface mass concentration (SO2SMASS), sulfate surface mass concentration (SO4SMASS), black carbon surface mass concentration (BCSMASS), sea salt surface mass concentration – $PM_{2.5}$ (SSSMASS25), dust surface mass concentration – $PM_{2.5}$ (DUSMASS25), dust extinction aerosol optical thickness (AOT) at 550 nm – $PM_{2.5}$ (DUEXTT25) were extracted from 2d time-averaged primary aerosol diagnostics at hourly intervals (tavg1_2d_aer_Nx). We conducted inverse distance weighting to calculate meteorological and aerosol diagnostic data to each 0.1° grid cell and calculated the monthly mean for each grid from 2009 - 2018.

1.4 Land Use and Population Data

We downloaded elevation data from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) version 2 at 1" spatial resolution (approximately 30 meters),

and averaged to 0.1° grid cell(NASA's Land Processes Distributed Active Archive Center (LP DAAC), 2009). Highway, primary road, and other road density were obtained from Global Roads Inventory Project (GRIP)(Meijer et al., 2018) at 5' spatial resolution (approximately 9 km) from GLOBIO (https://www.globio.info/download-grip-dataset) and were averaged and assigned to 0.1° grid cell. Yearly varying population 2009 – 2017 were obtained from LandScan™ Global Population Database at 30" resolution (approximately 1km) and averaged to 0.1° grid cell. The population in 2018 in each grid cell was predicted using linear regression based on population from 2009 to 2017. **Table 5-3S** listed details of variables used in $PM_{2.5}$ prediction model used in this study.

**2. $PM_{2.5}$ Random Forest Model Development and Evaluation**

We chose random forest algorithm (RF), a machine-learning algorithm based on bootstrap aggregating decision trees, to build the prediction model on monthly $PM_{2.5}$ concentrations over study domain. Initially developed by Breiman(Breiman, 2001), the RF algorithm has several advantages allowing both continuous and categorical input variables and being robust to outliers. Another advantage of random forest models is that they can provide variable importance vectors to compare the importance of variables in building the model; random forest algorithms are not likely to overfit.

We trained RF model on monthly averaged $PM_{2.5}$ concentration from 2017 – 2018 for each grid cell using ground based $PM_{2.5}$ measurements and evaluated the model performance using 10-fold cross validation (CV). The number of decision trees chosen is 280, based on minimizing the error rate of prediction. To evaluate temporal and spatial performance of the model, we also conducted 10-fold spatial and temporal 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 relays on using 90% of monthly data to develop the model and then testing the model based on the remaining 10% of monthly data. We further evaluated model prediction and performance by hindcasting on monthly data from 2015 – 2016 (N = 465), before the time period from which the model was trained. Root-mean-squared error (RMSE) and R-squared ($R^2$) from out-of-sample 10-fold cross-validation and from hindcasting were reported to evaluate the historical prediction performance of the model. Lastly, we predicted monthly ambient $PM_{2.5}$ concentrations over India for 10 years (2009 – 2018). We compared our modeled annual $PM_{2.5}$ mean concentrations and model performance with published databases:

global estimates of fine particulate matter by van Donkelaar and collegues(van Donkelaar et al., 2016) and Data integration Model for Air Quality (DIMAQ)(Shaddick et al., 2018) and DIMAQ-2 model(Balakrishnan, Dey, et al., 2018), which are the models used in Global Burden of Disease (GBD) 2016 and GBD 2017 project for ambient air pollution, respectively. R (version 3.5) with the randomForestSRC (version 4.6) package was used for $PM_{2.5}$ modelling, prediction, and assignment for each child.

## 3. Cox Regression and Additional Analysis

### 3.1 Variable included in Cox Regression model

For the two-exposure Cox regression model in the main analysis, 9-month in-utero $PM_{2.5}$, and post-delivery lifetime average $PM_{2.5}$ exposure all satisfied proportional hazard assumptions based on test of nonzero slope in a generalized linear regression of the scaled Schoenfeld residuals on time(StataCorp, 2013). The prespecified time-independent covariates include child's sex, birth month and year, birth order, location of birth (institutional birth or not), whether multiple birth, mother's age at child birth, mother's height, marital status, education (whether above secondary level), maternal smoking, wealth index, whether household member smoke (second-hand smoke exposure), urban or rural location of households, geographical region/zonal council (Northern, Central, North Eastern, Eastern, Western and Southern), primary cooking fuel, and toilet facilities. Time-dependent covariates include monthly rainfall precipitation and monthly temperature. Child birth year and month, sex, and geographical location of households were modelled as strata and unique baseline hazard functions were given for all 927 strata of combinations of these variables. The robust variance estimation for coefficients were used. For the tied outcomes, Breslow's method was used. Children born as low birth weight (LBW, birth weight less than 2 500 g) might mediate in-utero $PM_{2.5}$ and mortality and previous literature suggested not including LBW in the analysis of parental exposure and child mortality(Hernández-Díaz et al., 2006). Therefore, we exclude LBW covariates in the main analysis and conducted sensitivity analysis (explained in part 3.3 and 3.4), and further analyzed whether in-utero $PM_{2.5}$ exposure is associated with LBW using mixed effects logistic regression.

Single-exposure and two-exposure model comparison

We conducted sensitivity analysis and compared $PM_{2.5}$-mortality HRs and Akaike Information Criterion (AIC) for the two-exposure model and single-exposure models using in-utero $PM_{2.5}$ and post-delivery lifetime $PM_{2.5}$ only. All two-exposure model and single-exposure models included covariates specified in the main analysis.

3.3 Additional sensitivity analysis

To check the robustness of our results, we conducted a series of sensitivity analysis on the two-exposure model and the single-exposure models with in-utero or post-delivery life $PM_{2.5}$ exposure. First, to check whether inclusion of LBW will modify the model results, we added LBW to the single-exposure and two-exposure models (modify 1 model). Second, we conducted analysis using the stratified variables (child birth month and year, sex and geographical location) as time-independent predictors in the single-exposure and two-exposure models (modify 2 model). Third, we included additional cluster-level covariates, including the percentage of households with improved toilets, and the percentage of households using clean fuels (modify 3 model). Fourth, we conducted unweighted analysis of two-exposure models without using sample weight from DHS (modify 4 model). Next, due to 76 923 (30%) of children being from the same households and the potential correlation of mortality in these households (n = 175 865) or clusters (n = 27 835) level, we conducted Cox regression analysis with generalized estimating equations to account for the correlation at the household (modify 5 model) and cluster levels (modify 6 model), by specifying vce(*cluster*) in stcox command in Stata. Lastly, we replaced post-delivery lifetime $PM_{2.5}$ exposure with monthly $PM_{2.5}$ exposure, defined as monthly mean $PM_{2.5}$ exposure each month after child birth until child death or censoring, in the two-exposure model (modify 7 model).

### 3.4 Mediation analysis of In-utero $PM_{2.5}$ exposure, LBW and child mortality

In order to check whether LBW mediates associations between in-utero $PM_{2.5}$ exposure and child mortality, we conducted additional modeling assessing the relationship between in-utero $PM_{2.5}$ exposure and LBW and the relationship between LBW and child mortality. First, we conducted a mixed effect logistic regression between in-utero $PM_{2.5}$ exposure and LBW, with cluster-level random intercepts and controlled with related covariates. Then, we conducted Cox regression with LBW, post-delivery single-exposure $PM_{2.5}$ exposure, and other covariates to assess whether LBW can predict child mortality.

**4. Ambient $PM_{2.5}$ Prediction Model Performance and Predicted $PM_{2.5}$ in India Over 10 Years.**

4.1 Performance of $PM_{2.5}$ Prediction Model

**Figure 5-2S** shows the number of decision trees used and variable importance matrix in RF model to fit the ambient $PM_{2.5}$ prediction model. We chose 280 as the number of decision trees in order to minimize the error rate of prediction. The variables with highest prediction importance are simulated aerosol and reanalyzed meteorological variables from the global reanalysis model. Direct satellite retrievals of aerosol optical depth have medium importance among all variables. These includes aerosol absorption index (AAI) from OMI satellite, deep-blue (DB), deep-blue, and dark-target combined (Combined) AOD parameters from Terra and Aqua satellites. **Table 5-4S** shows the RF model CV performance and historical prediction/hindcast performance.

A total of 1 446 data points of monthly $PM_{2.5}$ was included in the RF model from over 134 ground based $PM_{2.5}$ monitoring stations in 2017 - 2018, with overall measured mean $PM_{2.5}$ concentration 76.5 $\mu g/m^3$. The RF model shows 10-fold CV $R^2$ (root-mean-squared error, RMSE) of 0.82 (25 $\mu g/m^3$), with predicted mean of 77.3 $\mu g/m^3$. The spatial and temporal CV $R^2$ (RMSE) of RF model is 0.77 (26 $\mu g/m^3$) and 0.77 (25 $\mu g/m^3$) respectively, indicating stable and good fit of monthly $PM_{2.5}$ air pollution measurements from ground stations. The historical hindcasting based on monthly measurements (N = 465) from 2015 – 2016 shows a $R^2$ (RMSE) of 0.82 (32 $\mu g/m^3$), with predicted mean of 81.8 $\mu g/m^3$ compared to observed mean of 91.2 $\mu g/m^3$. The historical prediction shows our model underestimates $PM_{2.5}$ levels during periods with extreme levels, when monthly $PM_{2.5}$ is over 350 $\mu g/m^3$ (for instance, a few high ambient air pollution measurements in Northern India in January and November of 2016). Regardless, our model still has relatively good hindcasting capacity to predict

historical monthly $PM_{2.5}$ levels. The annual within-sample RMSE from model CV in 2017 - 2018 is 10.9 μg/m$^3$, and annual RMSE from historical hindcasting in 2015 - 2016 is 21.7 μg/m$^3$, with better performance compared to previous ambient air pollution models used in Global Burden of Disease study in the South Asia region(Shaddick et al., 2018), which has within-sample RMSE of 17.6 μg/m$^3$. **Table 5-5S** shows the modelled and measured annual mean $PM_{2.5}$ concentration in Delhi, Mumbai, Kolkata, Chennai and Bangalore in 2015 - 2018, confirming our model has relatively good predictions over large cities in India.

4.2 Predicted Ambient $PM_{2.5}$ Levels

**Figure 5-4S** shows predicted annual mean ambient $PM_{2.5}$ concentrations over 10 years from 2009 to 2018. The map on the upper left shows the predicted annual mean $PM_{2.5}$ concentrations in 2018. The remaining maps show the difference of predicted annual mean $PM_{2.5}$ concentrations between each year from 2009 – 2017 and the predicted concentration in 2018. Annual concentrations show clear spatial patterns, with highest annual $PM_{2.5}$ concentrations in the Indo-Gangetic Plain (IGP) covering Haryana, Uttar Pradesh, Bihar states and National Capital Territory of Delhi of over 120 μg/m$^3$, three times above 40 μg/m$^3$, the recommended limit set by National Ambient Air Quality Standards (NAAQS) of India(Gautam, 2019). On the contrary, southern Indian states have the lowest annual mean $PM_{2.5}$ concentrations, with most of the areas around annual mean $PM_{2.5}$ of 35 μg/m$^3$, the interim target-1 (IT-1) level of World Health Organization and below the Indian NAAQS(Gautam, 2019). The monthly average $PM_{2.5}$ concentration over 10 years shows that $PM_{2.5}$ levels have strong seasonal variations (**Figure 5-5S**), with peak $PM_{2.5}$ between November and March, and lows during the month July and August.

To estimate the change of annual ambient $PM_{2.5}$ air pollution over time, we conducted a linear regression of annual predicted $PM_{2.5}$ over 10 years for each grid. We found that most states shows an increase of annual $PM_{2.5}$ concentrations at the rate of 0.5 $\mu g/m^3$ per year. Particularly, the IGP and west coast of Western region, the ambient $PM_{2.5}$ concentrations increased at around 1 $\mu g/m^3$ per year; while Northeastern region and north part of Northern region shown decrease of ambient $PM_{2.5}$ levels during 2009 -2018 (**Figure 5-1**).

**Table 5-6S** displays population weighted ambient $PM_{2.5}$ concentrations over 2009 – 2018 from our predictions, ground station measurements, and other model predictions of population-weighted ambient $PM_{2.5}$ concentrations. The population weighted-annual ambient $PM_{2.5}$ concentration slightly increased from 2009 – 2015 and remains relative stable from 2016 – 2018. Our model prediction levels are constantly higher than van Donkelaar et al's model prediction(van Donkelaar et al., 2016) but similar to DIMAQ models(Balakrishnan, Dey, et al., 2018; Shaddick et al., 2018)

**5. Demographic and Health Survey (DHS) Data Missing Patterns**

All 259 627 live-birth children born to DHS respondents (women of 15 – 49 age) no more than 5 years preceding the survey were included. We excluded 2 282 (0.9%) children born in the Union Territories of Andaman and Nicobar Islands and Lakshadweep, both of which are more than 200 km away from the subcontinent and excluded 1 117 (0.5%) children without geo-coordinates. In addition, we excluded 1 912 (0.7%) children who were not living in the same households as their mothers during the survey interview, and 1 528 (0.6%) children with missing values of birthplace or mother's height measurements. A total 252 788 children born January 2010 – November 2016 from 27 853 clusters were included in the analysis.

**Supplementary Tables**

**Table 5-1S**. Number of $PM_{2.5}$ measurement stations, the number of station-days of data, and the number of measurement days by year.

| Year | Number of stations | Total station-days | Measurement days for each station | | | |
|---|---|---|---|---|---|---|
| | | | Min | Mean | Median | Max |
| 2015 | 13 | 2721 | 1 | 239 | 254 | 292 |
| 2016 | 56 | 13075 | 1 | 294 | 321 | 366 |
| 2017 | 86 | 17253 | 1 | 274 | 325 | 365 |
| 2018 | 133 | 40655 | 25 | 324 | 347 | 365 |

**Table 5-2S**. Missing Patterns of AOD retrieval from Aqua and Terra satellite.

| Month | Number of days in each month with valid AOD retrieval | | | | | |
|---|---|---|---|---|---|---|
| | Satellite Terra | | Satellite Aqua | | Two Satellite combined | |
| | DB algorithm | DB-DT combined | DB algorithm | DB-DT combined | DB algorithm | DB-DT combined |
| 1 | 10.4 | 8.5 | 13.7 | 11.2 | 17.7 | 13.3 |
| 2 | 4.4 | 4.0 | 13.3 | 11.4 | 14.2 | 11.8 |
| 3 | 12.3 | 11.5 | 11.7 | 10.5 | 17.4 | 14.0 |
| 4 | 11.7 | 10.5 | 11.8 | 9.9 | 16.9 | 13.0 |
| 5 | 11.6 | 9.0 | 10.3 | 6.8 | 16.2 | 9.5 |
| 6 | 5.8 | 3.1 | 4.6 | 2.0 | 8.1 | 2.9 |
| 7 | 1.3 | 0.5 | 1.4 | 0.4 | 2.3 | 0.5 |
| 8 | 0.6 | 0.2 | 1.8 | 0.6 | 2.1 | 0.6 |
| 9 | 6.1 | 3.3 | 5.0 | 2.0 | 8.6 | 3.0 |
| 10 | 11.5 | 8.7 | 10.5 | 7.2 | 15.9 | 10.1 |
| 11 | 13.8 | 10.7 | 13.0 | 9.5 | 18.9 | 12.9 |
| 12 | 13.8 | 10.7 | 13.6 | 10.5 | 19.3 | 13.4 |

DB: deep blue AOD retrieval algorithm, DB-DT combined: deep blue – dark target combined AOD retrieval algorithm

**Table 5-3S**. Variables in Random Forest (RF) Model in Predicting Monthly Ambient $PM_{2.5}$ Levels

| Variable name | Source | Original Product Name | Variable Meaning |
|---|---|---|---|
| MODIS.DB.AOD | Terra and Aqua satellites | MOD04_L2 and MYD04_L2 | Deep-blue (DB) AOD estimation Retrieval |
| MODIS.Combined.AOD | | | Deep-blue and Dark-target combined AOD estimation retrieval |
| MERRA.PS | MERRA-2/GEOS 5 Meteorological Reanalysis and Aerosol Similation Model | PS | Surface pressure |
| MERRA.T2M | | T2M | 2-meter air temperature |
| MERRA.U10M | | U10M | 2-meter eastward wind |
| MERRA.V10M | | V10M | 2-meter northward wind |
| MERRA.QV2M | | QV2M | 2-meter specific humidity |
| MERRA.TQL+TQI | | Derived variable | Total ice and liquid precipitation: *TQI* and *TQL* |
| MERRA.OCSMASS | | OCSMASS | Organic Carbon Surface Mass Concentration |
| MERRA.SO2SMASS | | SO2SMASS | SO2 Surface Mass Concentration |
| MERRA.BCSMASS | | SO4SMASS | SO4 Surface Mass Concentration |
| MERRA.SSSMASS25 | | BCSMASS | Black Carbon Surface Mass Concentration |
| MERRA.DUSMASS25 | | SSSMASS25 | Sea Salt Surface Mass Concentration - PM 2.5 |
| MERRA.DUEXTT25 | | DUSMASS25 | Dust Surface Mass Concentration - PM 2.5 |
| MERRA.PBLH | | DUEXTT25 | Dust Extinction AOT [550 nm] - PM 2.5 |
| OMI.AAI | Aura ozone monitoring instrument (OMI) satellite | OMAERO and OMAERUV | Aerosol absorption index, averaged from *AerosolIndexUV*, *AerosolIndexVIS* and *UVAerosolIndex* variables |
| OMI.NO2 | | OMNO2 | Tropospheric $NO_2$ concentration, based on *ColumnAmountNO2Trop* variable |

| | | | |
|---|---|---|---|
| MODIS.Fire | Terra and Aqua satellites | MCD14DL | Active Fire Spot Detection |
| MODIS.NDVI | Terra satellite | MOD13C1 | Normalized Difference Vegetation Index (NDVI) |
| LS.POPULATION | LandScan population model | LandScan | Global population estimation model |
| GDEM.Elevation | Global Digital Elevation Model | ASTER-GDEM | Global elevation estimation model |
| GRIP.HW | Global Road Inventory Project | GRIP Highway | Highway road density |
| GRIP.MR | | GRIP Major road | Major road density |
| GRIP.OR | | Derived variable | Primary road, secondary road and territory road density |
| GRIP.AllRoad | | Derived variable | All types of roads (highway, major, primary, secondary and territory) density |

**Table 5-4S**. Ambient $PM_{2.5}$ Air Pollution Prediction Model Fitting, 10-Fold Cross-Validation, and Historical Hindcasting at Monthly Level

| | Number of observations | $R^2$ (RMSE ($\mu g/m^3$)) |
|---|---|---|
| Standard 10-fold CV§ | 1446 | 0.82 (25) |
| Spatial 10-fold CV§ | 1446 | 0.77 (27) |
| Temporal 10-fold CV§ | 1446 | 0.77 (26) |
| Historical Prediction/Hindcast ¶ | 465 | 0.81 (32) |

RMSE: rooted mean square error; §based on 2017 – 2018 data; ¶based on external 2015 – 2016 data

**Table 5-5S**. Modelled and Ground Station Measured Annual $PM_{2.5}$ Concentration ($\mu g/m^3$) in Delhi, Mumbai, Kolkata, Chennai and Bangalore 2015 – 2018

| | 2015 | | 2016 | | 2017 | | 2018 | |
|---|---|---|---|---|---|---|---|---|
| City/Region | Modelled PM | Ground station measurements | Modelled PM | Ground station measurements | Modelled PM | Ground station measurements | Modelled PM | Ground station measurements |
| Delhi | 111.8 | 119.4 | 116.1 | 133.2 | 111.5 | 132.05 | 114.6 | 113.6 |
| Mumbai | 51.4 | 50.0 | 52.4 | 54.9 | 53.0 | 65.01 | 53.3 | 58 |
| Kolkata | 89.2 | 68.5 | 88.7 | 84.4 | 81.9 | 76.7 | 82.9 | 95.1 |
| Chennai | 37.1 | 25.8 | 40.9 | 55.3 | 39.5 | 46.9 | 40.6 | 48.4 |
| Bangalore | 33.3 | NA | 35.6 | 48.3 | 34.9 | 34.7 | 35.4 | 34.6 |

NA: not available

**Table 5-6S**. Predicted Population Weighted Annual Average $PM_{2.5}$ in India and Comparison of Other Model

| | | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Our model | Population weighted $PM_{2.5}$ ($\mu g/m^3$) | 70.5 | 67.7 | 70.6 | 72.8 | 69.7 | 72.9 | 70.3 | 73.5 | 72.3 | 71.7 |
| Ground station measurements | $PM_{2.5}$ measurement ($\mu g/m^3$) | NA | NA | NA | NA | NA | NA | 86.5 | 92.3 | 83.4 | 73.1 |
| van Donkelaar et al., 2016 (van Donkelaar et al., 2016) | Population weighted $PM_{2.5}$ ($\mu g/m^3$) | 52.1 | 50.2 | 49.5 | 48.5 | 53.2 | 54.8 | 57.0 | 61.3 | NA | NA |
| DIMAQ/DIMAQ-2(Balakrishnan, Dey, et al., 2018; Shaddick et al., 2018) | Population weighted $PM_{2.5}$ ($\mu g/m^3$) | NA | 64.6 | 63.9 | 66.1 | 69.0 | 71.6 | 75.6 | 75.8 | 89.9 | NA |

DIMAQ/DIMAQ-2: Data integration Model for Air Quality (version 1 and version 2), which was used in GBD2016 and GBD2017 project; NA: not available

**Table 5-7S**. Variables in Cox Regression Model in Predicting Child Mortality

| Variable | Source | Type |
|---|---|---|
| Child's birth month and year | DHS | Stratified variable (categorical) |
| Child's sex | DHS | Stratified variable (binary) |
| Household geographical region | DHS | Stratified variable (categorical) |
| Birth order | DHS | Categorical |
| Institutional birth | DHS | Binary |
| Multiple birth | DHS | Binary |
| Mother's age at childbirth (quartile) | DHS | Categorical |
| Mother's height (quartile) | DHS | Categorical |
| Maternal smoking | DHS | Binary |
| Maternal education | DHS | Binary |
| Other household member smoke | DHS | Binary |
| Household wealth index | DHS | Continuous |
| Household cooking fuel | DHS | Binary |
| Household toilet | DHS | Binary |
| Urbanity of household | DHS | Binary |
| Monthly temperature | MERRA-2 | Continuous |
| Monthly rainfall precipitation | CHIRPS | Continuous |

MERRA-2: Modern-Era Retrospective analysis for Research and Applications; CHIRPS: Climate Hazards Group InfraRed Precipitation with Station

**Table 5-8S**. Akaike Information Criterion (AIC) for single-exposure and two-exposure models§

| | Single-exposure | | Two-exposure |
|---|---|---|---|
| | Model 1 | Model 2 | Model 3 |
| Child mortality | 1,733,004 | 1,733,059 | 1,732,961 |
| Infant Mortality | 1,562,318 | 1,562,346 | 1,562,265 |
| Neonatal Mortality | 1,156,491 | 1,156,481 | 1,156,433 |

Model 1: in-utero model, with only 9-month in-utero $PM_{2.5}$ exposure

Model 2: post-delivery lifetime average model, with only post-delivery lifetime average $PM_{2.5}$ exposure

Model 3: Two-exposure model, with both in-utero and post-delivery lifetime average $PM_{2.5}$ exposure

§All Cox regression models controlled for individual level, household level and cluster level covariates, stratified by child birth year and month, age and geographical zones, and correspond to the models in main analysis.

**Table 5-9S**. Odds Ratio of Low Birth Weight (LBW) per 10 μg/$m^3$ increase of In-utero $PM_{2.5}$ exposure §

| | Odds Radio (95% CI) |
|---|---|
| Crude model | 1.006 (0.999, 1.013) |
| Adjusted model | 1.013 (1.004, 1.022) |

§ based on mixed effect logistic model with cluster level (N = 27 436) random intercept, crude model only included in-utero $PM_{2.5}$ exposure, adjusted model included child birth month and year, sex, birth order, multi-birth, birth location, mother's age at child birth, height, marital status, education level, maternal smoking, household's wealth index, urban or rural location, toilet and cooking fuel.

**Table 5-10S**. Adjusted Hazard Ratios associated with LBW on Child Mortality in Single-exposure Model with In-utero $PM_{2.5}$ exposure§

| | Hazard Ratio (95% CI) |
|---|---|
| Child Mortality | 1.419 (1.326, 1.519) |
| Infant Mortality | 1.442 (1.343, 1.548) |
| Neonatal Mortality | 1.406 (1.296, 1.527) |

§ Controlled for in-utero $PM_{2.5}$, child birth month and year, sex, geographical location, birth order, multi-birth, birth location, mother's age at child birth, height, marital status, education level, maternal smoking, household's wealth index, urban or rural location, toilet and cooking fuel, the same model as modify 1 model in additional sensitivity analysis.

**Supplementary Figures**

**Figure 5-1S**. Gap-filling for Deep-blue (DB) and Deep-blue and Dark-Target combined (Combined) Aerosol Optical Depth Retrieval Algorithm in July 2017

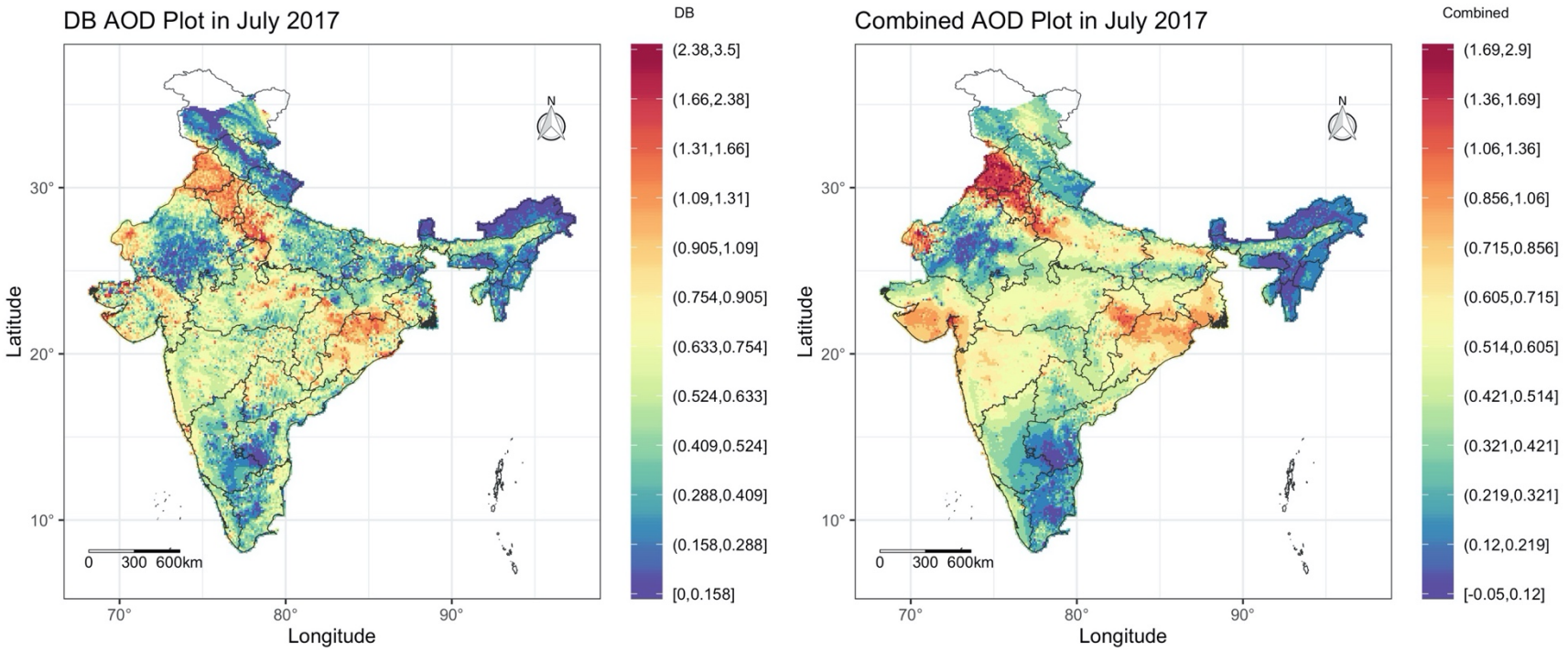


**Figure 5-2S**. Variable importance plot of Ambient $PM_{2.5}$ Random Forest Model Fitting

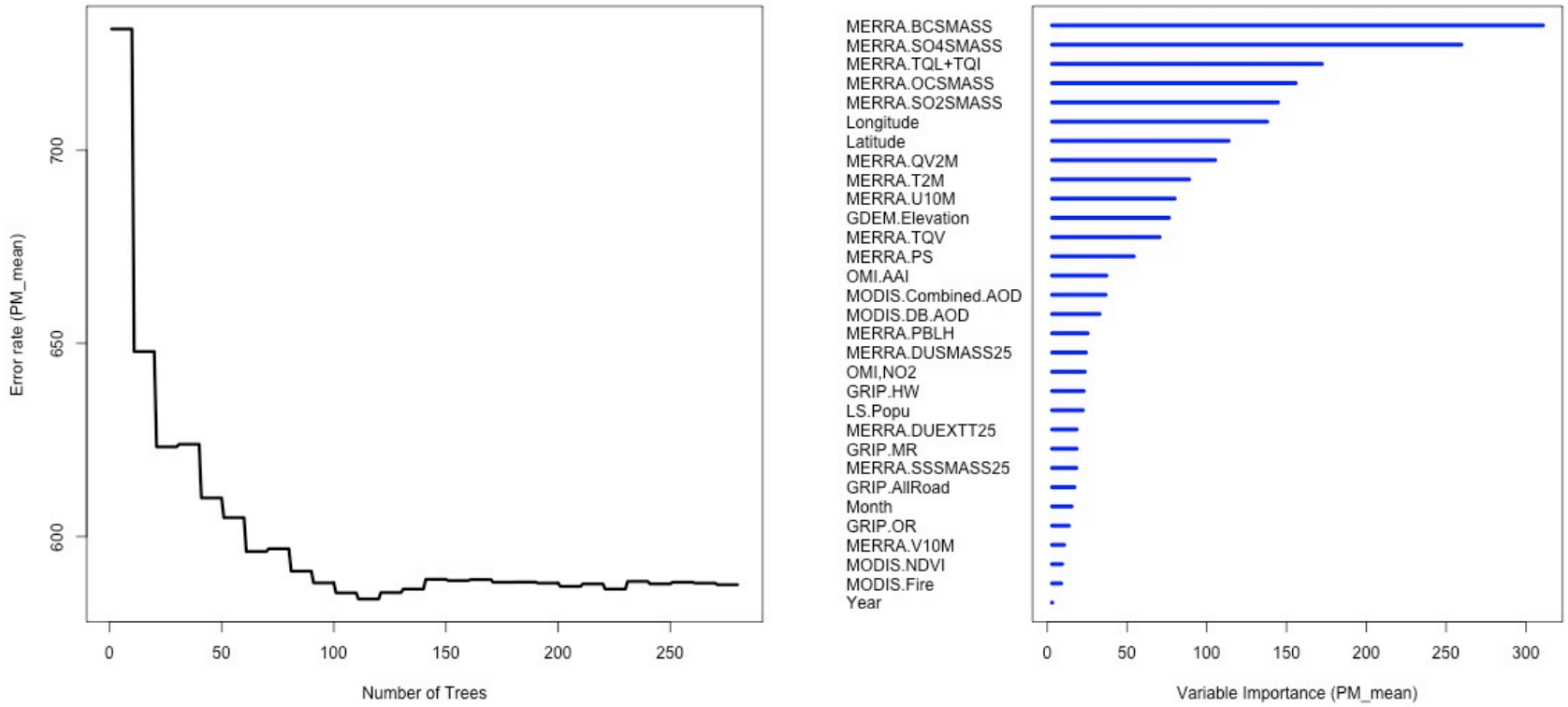


**Figure 5-3S**. Cross-validation and Hindcasting Performance of the Random Forest Model on Monthly $PM_{2.5}$ Measurements

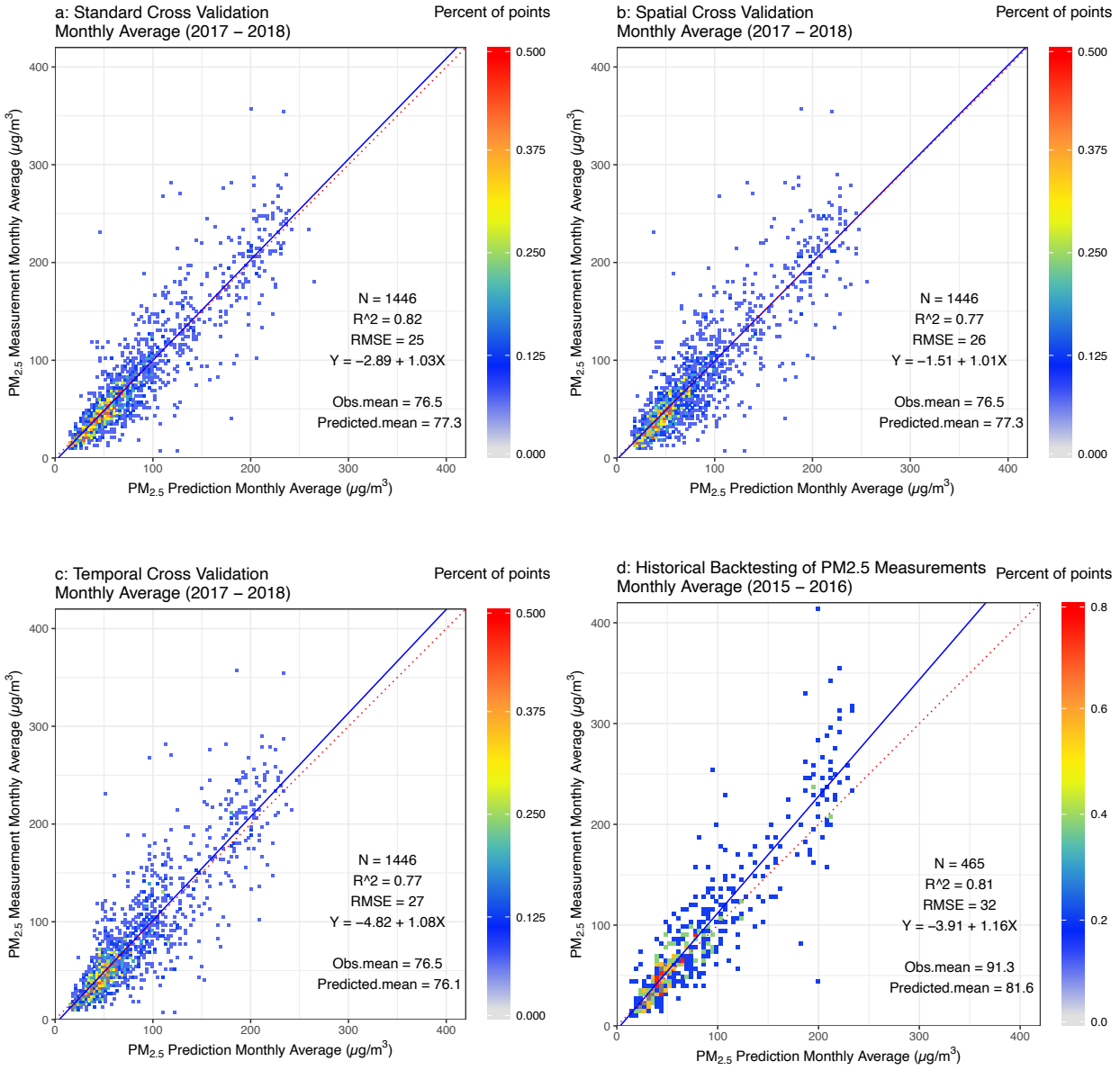


**Figure 5-4S.** Predicted Annual $PM_{2.5}$ Concentrations from 2009 – 2018. The upper left map shows the predicted annual $PM_{2.5}$ Concentrations in 2018, and the rest of maps shows the differences of annual $PM_{2.5}$ Concentrations between 2018 and each year in 2009 – 2017.

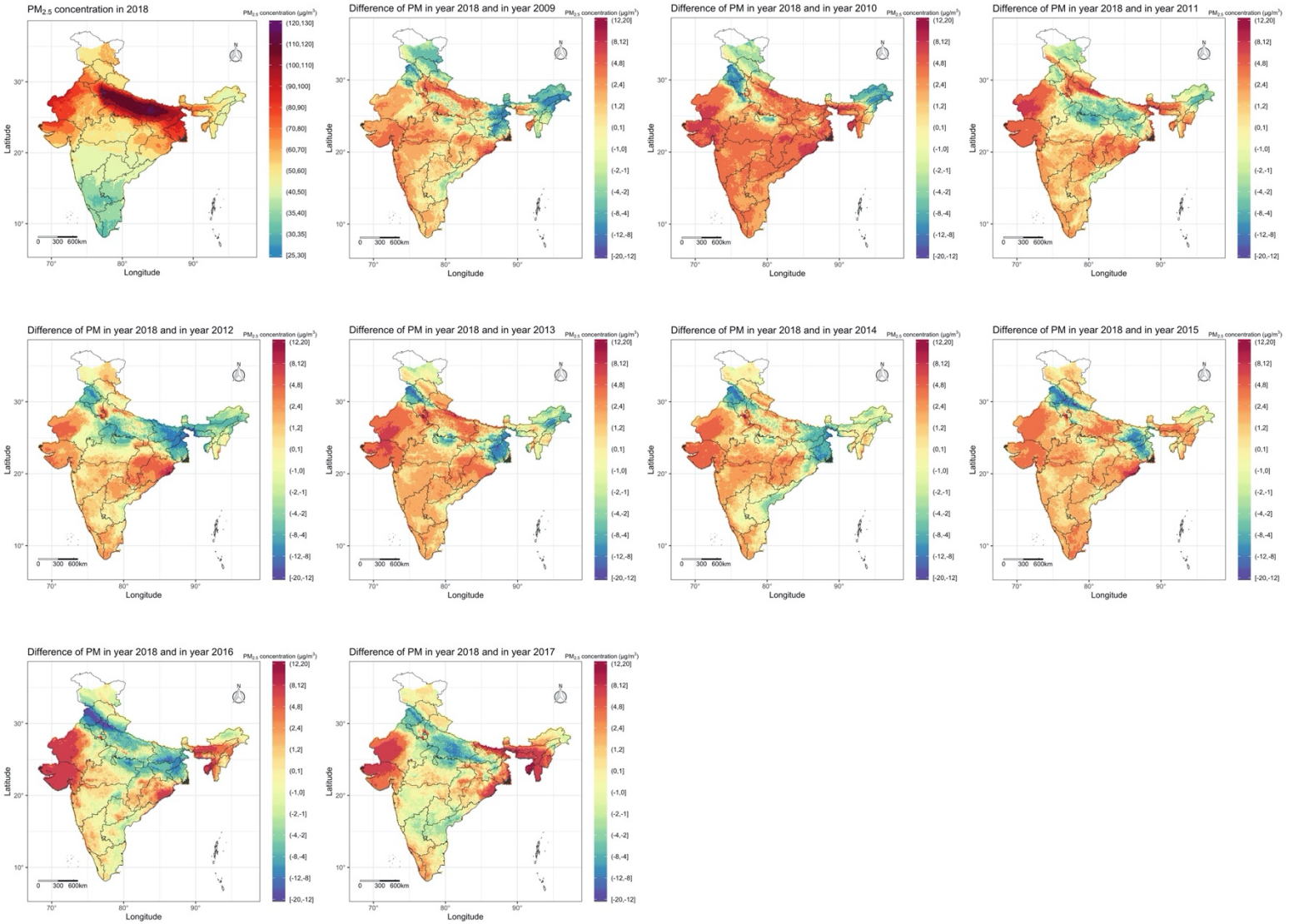


**Figure 5-5S**. Predicted Monthly $PM_{2.5}$ Concentrations, averaged across 10 years 2009 – 2018.

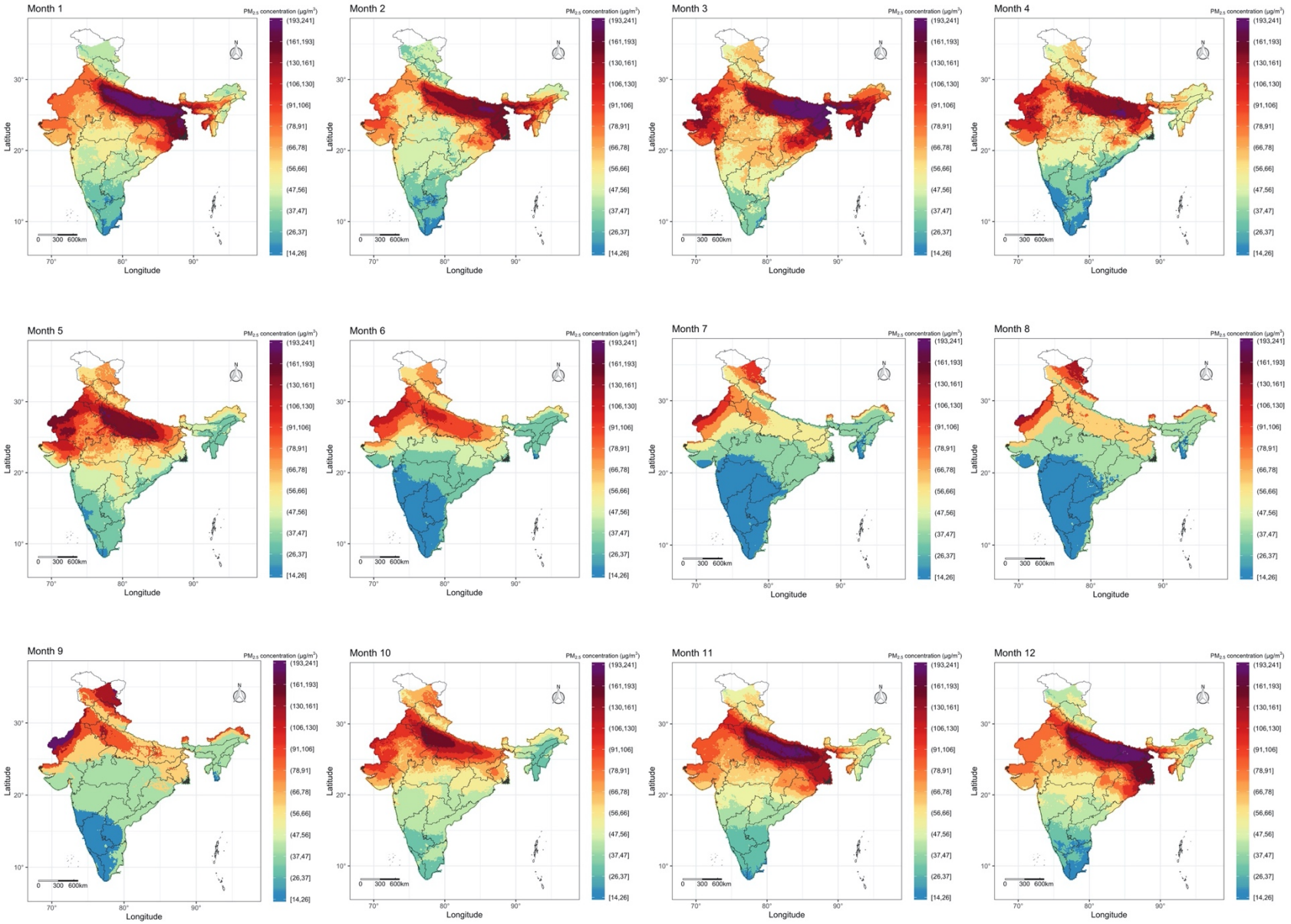
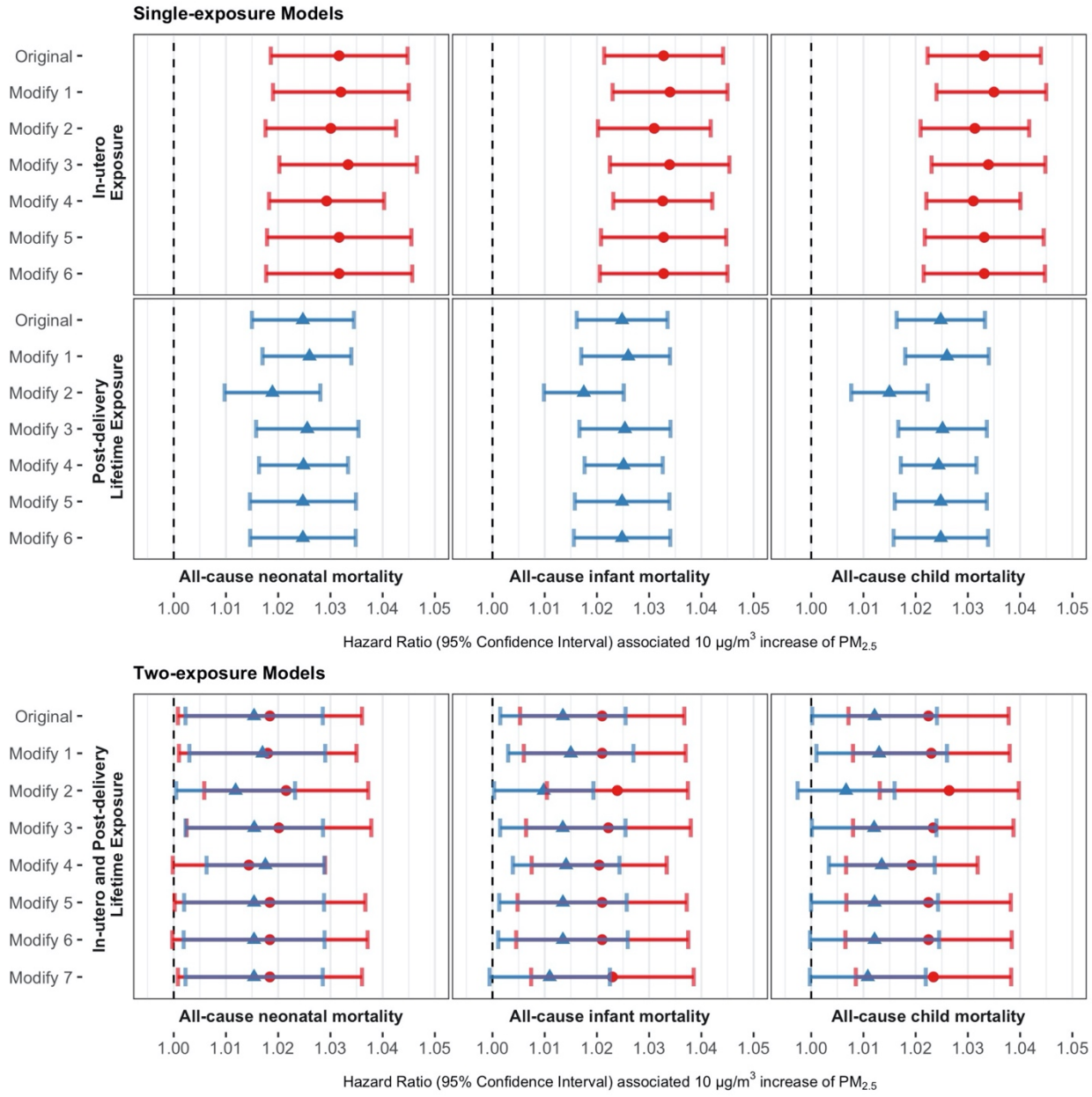

Month 1
Month 2
Month 3
Month 4
Month 5
Month 6
Month 7
Month 8
Month 9
Month 10
Month 11
Month 12
PM2.5 concentration (µg/m3)
(193,241]
(161,193]
(130,161]
(106,130]
(91,106]
(78,91]
(66,78]
(56,66]
(47,56]
(37,47]
(26,37]
[14,26]
Latitude
Longitude
0 300 600km

**Figure 5-6S**. Hazard Ratios (HRs) of Mortality and 95% Confidence Interval per 10 μg/m$^3$ increase of $PM_{2.5}$ Exposure, According to Additional Sensitivity Analysis. Red circle represents effects of in-utero $PM_{2.5}$, and blue triangle represents effects of post-delivery cumulative $PM_{2.5}$.




Original model: represents model in the main analysis

Modify 1: add low birth weight (LBW) as predictor

Modify 2: use stratified variables (child birth month, year, sex and geographical area) as time-independent predictor in the model

Modify 3: add cluster level covariates including percentage of households with improved toilet and clean cooking fuel as predictor

Modify 4: un-weighted analysis

Modify 5: GEE to control correlation within household

Modify 6: GEE to control correlation within cluster

Modify 7: Replace post-delivery lifetime average $PM_{2.5}$ with post-delivery monthly $PM_{2.5}$, only in two-exposure model

# Chapter 6 **Summary, Implications and Further Research**

## 6.1 Summary

The overall objective of the research described in this dissertation was to address knowledge gaps in air pollution exposure assessment and provide evidence on the exposure-response relationship of air pollution, especially for children and pregnant women in resource-limited settings. The research is intended to support air pollution exposure assessment for newborn infants in HAPIN trial and contribute to the epidemiological evidence on the gestational health effects of air pollution exposure. We sought to accomplish this objective by conducting three studies, each of which involved different methods and sources of data.

In Chapter 3, we provided an assessment of an indirect, sensor-enabled exposure measurement technique in households using woodstoves at baseline and an LPG cookstoves at follow-up. This chapter demonstrated that indirect exposure assessment using the Beacon system as a microenvironmental location monitor provides an acceptable estimate of personal exposures in both low and high $PM_{2.5}$ exposure settings. We found that indirect exposure methods have a higher correlation with direct personal exposure measurements and less bias than do kitchen measurements. In settings where conducting personal direct exposure assessment is not practical, such as in children under 1 year old, the Beacon indirect exposure method is an alternative to better estimate personal exposure to $PM_{2.5}$. The results of this study can inform exposure assessments for future household air pollution studies.

In Chapter 4, we applied a mixed effects model to examine the effect of personal $PM_{2.5}$ exposure on gestational blood pressure using a prospective cohort of pregnant women enrolled in the HAPIN intervention trial. We found that exposure to household air pollution is associated with an increase in systolic blood pressure, but not diastolic blood pressure during pregnancy. Consistent with previous studies, our study shows that household air pollution from solid fuel combustion is associated with increases in blood pressure, a risk factor for adverse pregnancy outcomes later in the pregnancy. Interventions to reduce HAP, such as provision of a clean cooking

stoves and LPG fuel, can positively impact blood pressure during pregnancy, with potential benefits on maternal health.

In Chapter 5, we developed a spatial-temporal resolved mapping of ambient $PM_{2.5}$ concentration over India from 2009 – 2018. We demonstrated that the adverse child health effects were associated with high levels of ambient $PM_{2.5}$ exposure, both in utero and post-delivery period, based on a retrospective birth cohort reconstructed from Demographic and Health Survey (DHS). We showed that up to 18% (95% CI: 12.6% - 23.1%) of child mortality can be averted by reducing exposure levels to WHO ambient $PM_{2.5}$ guidelines — a massive disease burden in a country like India due both to the large population and the egregious levels of air pollution.

By advancing exposure assessment of both household air pollution and ambient air pollution through state-of-art technology and data analysis, this thesis demonstrated the feasibility of these improved exposure assessment methods in both intervention studies of a clean cooking stove and observational studies of ambient air pollution. This thesis also provides evidence supporting the adverse maternal and child health effects of air pollution from cookstoves and outdoor ambient environments in low-income settings.

## 6.2 Implications

This research has important implications in exposure science and environmental epidemiological research around air pollution and associated adverse health effects among pregnant women and children in LMICs, where many households rely on solid fuel for cooking and heating and experience high ambient air pollution due to industrialization. The assessment and validation of the Beacon system demonstrate that it offers a promising alternative in assessing the personal exposure to air pollution among very young children who are considered as one of the most vulnerable yet challenging group to monitor. Current personal air monitoring does not suit children for two reasons: 1) the size of monitors which are too large for children to wear and 2) the air pump in the air monitor that is bothersome to operate on the youngest children. Results from the study

will inform instrument selection in future studies. More importantly, overcoming the difficulties in assessing child exposure will encourage more studies to look into this unique age group and add more environmental epidemiological evidence to this important but not well-studied population, especially the effects of early life exposure to air pollution and its' on later health outcomes, such as the development of the respiratory system, nervous system, and metabolic system.

The development of the ambient air pollution model over India for the past 10 years addressed the lack of the ground-based air pollution monitoring stations in the South Asia subcontinent. It provides a basis for future epidemiological studies on the health effects of long-term air pollution exposure in India, where the ambient air pollution remains the highest globally. This advance in air pollution exposure assessment also has important implications for estimating the health effects of air pollution, especially in low-income settings. The study of the gestational blood pressure effects of household air pollution and the study of child mortality effects of ambient air pollution contribute to our knowledge of increased maternal and child health risks from exposure to air pollution – from both around the household and the ambient environment.

Overall, the results of this dissertation research underscore the importance of air pollution intervention strategies at the household level, such as provision of clean cookstoves and fuel, as well the urgency to address ambient air pollution levels in low-income settings with high air pollution exposures.

## 6.3 Further Research

In this dissertation research, we used cutting-edge and innovative methods to characterize individual exposure to air pollution and associated health effects. Accurately assessing personal exposure is the key to environmental health research. In the area of air pollution, the development of monitoring technologies, and advances in artificial intelligence and the focus on human centered design promise a new wave of air pollution evaluation techniques. This promises to continue development of temporal and geographical resolutions and increased application of low-cost sensors and open data sources.

Despite the progress in these areas, research gaps remain in characterizing life-course exposure, as well as understanding how different environmental exposure sources interact with each other over a lifetime. These aspects of exposure are likely to play significant roles in human health, interacting with genetic, metabolic and environmental differences. There is a need to further study the effectiveness of air pollution interventions, identifying critical exposure windows such as during pre-conception, gestation and early-life periods, and examining potential effect differences based on gender and socioeconomic status. Future research should also focus on leveraging the rich dataset generated from high-throughput 'omics' approaches, including exposome, metabolomics, epi-genetics and genomics data, to understand genetic and metabolic pathways linking air pollution exposure during critical periods and the development of disease outcomes and disorders.

## 6.4 Reflections

With the completion of my dissertation research, I want to summarize what I have learned and what could have been done differently. For my research first aim in Chapter 3, I took advantage of an innovative technology and method to assess the air pollution of young children, and trained myself with strong analytical and coding skills. However, the heavy data processing burden and complicated process may prevent the large-scale uptake of this method. The full application of Beacon system in the HAPIN main trial needs more attention and inputs for analysis. More efforts are needed on streamline the data processing, analysis procedures and data integration. Application of new technologies in HAP research and applications should also consider cultural appropriation, acceptance and accessibility in low-resources settings. Future applications of technology to combat household air pollution in rural settings of developing countries should seriously consider these requirements.

In the second research aim of this dissertation, I contributed the epidemiological evidence in understanding maternal effects of HAP in low-income settings, especially among pregnant women where knowledge is currently scarce. To tackle this research question in-depth more, since gestational BP is determined dynamically

by many different factors, designing survey questions to account for potential confounders such as diet and physical activity need to be considered in future research. In addition, my first and second research aims are based on the HAPIN trial in Guatemala, and I contributed to a less extent in data collection, field training and management. This dissertation and my doctoral work have a significant contribution to the science of household air pollution and reducing its health burdens. However, in order to achieve the optimal public health outcomes, better management should always be conducted; therefore I need to train myself more in my academic and professional career in public health project management.

As for my last aim, I learned to use publicly available data sources to improve both ambient air pollution exposure assessment, as well as child health risks of air pollution in India. Both of these dissertation outputs serve as early information regarding health risks in applying ambient air pollution research in global health context in developing Asia. Using this open data source to conduct a retrospective cohort requires a number of assumptions and tends to be more susceptible to confounding and information bias than first-hand prospective cohort study. I am therefore looking to apply more advanced statistical and epidemiological models in prospective cohorts or randomized controlled trial studies in my future studies.

Overall, this dissertation examines advanced exposure assessment methods in HAP and AAP that are to strengthen the base of knowledge to understand the effects of air pollution in low-income settings, especially on maternal and child health. The application of the exposure assessment techniques in the HAPIN trial and future studies will be advanced through this dissertation work. Future epidemiological analysis of air pollution in developing Asia and other rural settings might be inspired by the methodology and findings from this dissertation.

Ultimately, achieving global health equity in the disease burden of air pollution, both from household and ambient environment, especially for vulnerable population like pregnant women and children in low-income settings, requires greater efforts to reduce air pollution levels are urgently needed. This requires rigorous

evidence to inform policy. The findings of this dissertation should assist air pollution exposure assessment, epidemiology research and implementation, and the identification of critical time periods for vulnerable population.

## Reference

Achakulwisut, P., Brauer, M., Hystad, P., & Anenberg, S. C. (2019). Global, national, and urban burdens of paediatric asthma incidence attributable to ambient NO2 pollution: Estimates from global datasets. *The Lancet. Planetary Health*, *3*(4), e166–e178. https://doi.org/10.1016/S2542-5196(19)30046-4

Agrawal, S., & Yamamoto, S. (2015). Effect of Indoor air pollution from biomass and solid fuel combustion on symptoms of preeclampsia/eclampsia in Indian women. *Indoor Air*, *25*(3), 341–352. https://doi.org/10.1111/ina.12144

Alexander, D., Larson, T., Bolton, S., & Vedal, S. (2015). Systolic blood pressure changes in indigenous Bolivian women associated with an improved cookstove intervention. *Air Quality, Atmosphere & Health*, *8*(1), 47–53. https://doi.org/10.1007/s11869-014-0267-6

Alexander, D., Northcross, A., Wilson, N., Dutta, A., Pandya, R., Ibigbami, T., Adu, D., Olamijulo, J., Morhason-Bello, O., Karrison, T., Ojengbede, O., & Olopade, C. O. (2017). Randomized Controlled Ethanol Cookstove Intervention and Blood Pressure in Pregnant Nigerian Women. *American Journal of Respiratory and Critical Care Medicine*, *195*(12), 1629–1639. https://doi.org/10.1164/rccm.201606-1177OC

Allen-Piccolo, G., Rogers, J. V., Edwards, R., Clark, M. C., Allen, T. T., Ruiz-Mercado, I., Shields, K. N., Canuz, E., & Smith, K. R. (2009). An ultrasound personal locator for time-activity assessment. *International Journal of Occupational and Environmental Health*, *15*(2), 122–132. https://doi.org/10.1179/oeh.2009.15.2.122

Amegah, A. K., Quansah, R., & Jaakkola, J. J. K. (2014). Household Air Pollution from Solid Fuel Use and Risk of Adverse Pregnancy Outcomes: A Systematic Review and Meta-Analysis of the Empirical Evidence. *PLOS ONE*, *9*(12), e113920. https://doi.org/10.1371/journal.pone.0113920

Balakrishnan, K., Dey, S., Gupta, T., Dhaliwal, R. S., Brauer, M., Cohen, A. J., Stanaway, J. D., Beig, G., Joshi, T. K., Aggarwal, A. N., Sabde, Y., Sadhu, H., Frostad, J., Causey, K., Godwin, W., Shukla, D. K.,

Kumar, G. A., Varghese, C. M., Muraleedharan, P., … Dandona, L. (2018). The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: The Global Burden of Disease Study 2017. *The Lancet Planetary Health*, *0*(0). https://doi.org/10.1016/S2542-5196(18)30261-4

Balakrishnan, K., Ghosh, S., Thangavel, G., Sambandam, S., Mukhopadhyay, K., Puttaswamy, N., Sadasivam, A., Ramaswamy, P., Johnson, P., Kuppuswamy, R., Natesan, D., Maheshwari, U., Natarajan, A., Rajendran, G., Ramasami, R., Madhav, S., Manivannan, S., Nargunanadan, S., Natarajan, S., … Thanasekaraan, V. (2018a). Exposures to fine particulate matter (PM2.5) and birthweight in a rural-urban, mother-child cohort in Tamil Nadu, India. *Environmental Research*, *161*, 524–531. https://doi.org/10.1016/j.envres.2017.11.050

Balakrishnan, K., Ghosh, S., Thangavel, G., Sambandam, S., Mukhopadhyay, K., Puttaswamy, N., Sadasivam, A., Ramaswamy, P., Johnson, P., Kuppuswamy, R., Natesan, D., Maheshwari, U., Natarajan, A., Rajendran, G., Ramasami, R., Madhav, S., Manivannan, S., Nargunanadan, S., Natarajan, S., … Thanasekaraan, V. (2018b). Exposures to fine particulate matter (PM2.5) and birthweight in a rural-urban, mother-child cohort in Tamil Nadu, India. *Environmental Research*, *161*, 524–531. https://doi.org/10.1016/j.envres.2017.11.050

Balakrishnan, K., Sambandam, S., Ghosh, S., Mukhopadhyay, K., Vaswani, M., Arora, N. K., Jack, D., Pillariseti, A., Bates, M. N., & Smith, K. R. (2015). Household Air Pollution Exposures of Pregnant Women Receiving Advanced Combustion Cookstoves in India: Implications for Intervention. *Annals of Global Health*, *81*(3), 375–385. https://doi.org/10.1016/j.aogh.2015.08.009

Balakrishnan, K., Sambandam, S., Ramaswamy, P., Ghosh, S., Venkatesan, V., Thangavel, G., Mukhopadhyay, K., Johnson, P., Paul, S., Puttaswamy, N., Dhaliwal, R. S., Shukla, D. K., & Team, S.-C. (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. https://doi.org/10.1136/bmjopen-2015-008090

Balakrishnan, K., Sambandam, S., Ramaswamy, P., Mehta, S., & Smith, K. R. (2004). Exposure assessment for respirable particulates associated with household fuel use in rural districts of Andhra Pradesh, India. *Journal of Exposure Science and Environmental Epidemiology*, *14*(S1), S14–S25. https://doi.org/10.1038/sj.jea.7500354

Balakrishnan, K., Sankar, S., Parikh, J., Padmavathi, R., Srividya, K., Venugopal, V., Prasad, S., & Pandey, V. L. (2002). Daily average exposures to respirable particulate matter from combustion of biomass fuels in rural households of southern India. *Environmental Health Perspectives*, *110*(11), 1069–1075.

Baumgartner, J., Schauer, J. J., Ezzati, M., Lu, L., Cheng, C., Patz, J. A., & Bautista, L. E. (2011). Indoor Air Pollution and Blood Pressure in Adult Women Living in Rural China. *Environmental Health Perspectives*, *119*(10), 1390–1395. https://doi.org/10.1289/ehp.1003371

Bell Michelle L., Ebisu Keita, & Belanger Kathleen. (2007). Ambient Air Pollution and Low Birth Weight in Connecticut and Massachusetts. *Environmental Health Perspectives*, *115*(7), 1118–1124. https://doi.org/10.1289/ehp.9759

Bonjour, S., Adair-Rohani, H., Wolf, J., Bruce, N. G., Mehta, S., Prüss-Ustün, A., Lahiff, M., Rehfuess, E. A., Mishra, V., & Smith, K. R. (2013). Solid Fuel Use for Household Cooking: Country and Regional Estimates for 1980–2010. *Environmental Health Perspectives*, *121*(7), 784–790. https://doi.org/10.1289/ehp.1205987

Brauer, M., Guttikunda, S. K., K a, N., Dey, S., Tripathi, S. N., Weagle, C., & Martin, R. V. (2019). Examination of monitoring approaches for ambient air pollution: A case study for India. *Atmospheric Environment*, *216*, 116940. https://doi.org/10.1016/j.atmosenv.2019.116940

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brook, R. D., Newby, D. E., & Rajagopalan, S. (2017). Air Pollution and Cardiometabolic Disease: An Update and Call for Clinical Trials. *American Journal of Hypertension*, *31*(1), 1–10. https://doi.org/10.1093/ajh/hpx109

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Whitsel, L., & Kaufman, J. D. (2010). Particulate Matter Air Pollution and Cardiovascular Disease. *Circulation*, *121*(21), 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S. S., Kan, H., Walker, K. D., Thurston, G. D., Hayes, R. B., … Spadaro, J. V. (2018). Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences*, *115*(38), 9592–9597. https://doi.org/10.1073/pnas.1803222115

Burnett, R., Pope, C. A., III, Ezzati, M., Olives, C., Lim, S. S., Mehta, S., Shin, H. H., Singh, G., Hubbell, B., Brauer, M., Anderson, H. R., Smith, K. R., Balmes, J. R., Bruce, N. G., Kan, H., Laden, F., Prüss-Ustün, A., Turner, M. C., Gapstur, S. M., … Cohen, A. (2014). An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure. *Environmental Health Perspectives*. https://doi.org/10.1289/ehp.1307049

Burrowes, V. (n.d.). *Validation of next-generation portable pollution monitors to measure exposure to PM2.5 from household air pollution in Puno, Peru [unpublished]*.

Carey, I. M., Atkinson, R. W., Kent, A. J., van Staa, T., Cook, D. G., & Anderson, H. R. (2013). Mortality associations with long-term exposure to outdoor air pollution in a national English cohort. *American Journal of Respiratory and Critical Care Medicine*, *187*(11), 1226–1233. https://doi.org/10.1164/rccm.201210-1758OC

Carter, E., Norris, C., Dionisio, K. L., Balakrishnan, K., Checkley, W., Clark, M. L., Ghosh, S., Jack, D. W., Kinney, P. L., Marshall, J. D., Naeher, L. P., Peel, J. L., Sambandam, S., Schauer, J. J., Smith, K. R., Wylie, B. J., & Baumgartner, J. (2017). Assessing Exposure to Household Air Pollution: A Systematic Review and Pooled Analysis of Carbon Monoxide as a Surrogate Measure of Particulate Matter. *Environmental Health Perspectives*, *125*(7), 076002. https://doi.org/10.1289/EHP767

Chakraborty, N. M., Fry, K., Behl, R., & Longfield, K. (2016). Simplified Asset Indices to Measure Wealth and Equity in Health Programs: A Reliability and Validity Analysis Using Survey Data From 16 Countries. *Global Health, Science and Practice*, *4*(1), 141–154. https://doi.org/10.9745/GHSP-D-15-00384

Chowdhury, S., & Dey, S. (2016). Cause-specific premature death from ambient PM2.5 exposure in India: Estimate adjusted for baseline mortality. *Environment International*, *91*, 283–290. https://doi.org/10.1016/j.envint.2016.03.004

Clark, M. L., Bazemore, H., Reynolds, S. J., Heiderscheidt, J. M., Conway, S., Bachand, A. M., Volckens, J., & Peel, J. L. (2011). A baseline evaluation of traditional cook stove smoke exposures and indicators of cardiovascular and respiratory health among Nicaraguan women. *International Journal of Occupational and Environmental Health*, *17*(2), 113–121. https://doi.org/10.1179/107735211799030942

Clark, M. L., Peel, J. L., Balakrishnan, K., Breysse, P. N., Chillrud, S. N., Naeher, L. P., Rodes, C. E., Vette, A. F., & Balbus, J. M. (2013). Health and household air pollution from solid fuel use: The need for improved exposure assessment. *Environmental Health Perspectives*, *121*(10), 1120–1128. https://doi.org/10.1289/ehp.1206429

Clasen, T. F., Peel, J., & Checkley, W. (n.d.). *Household Air Pollution Intervention Network (HAPIN) Trial Protocol [unpublished]*.

Croft, T. N., Aileen, M. J. M., & Courtney, K. A. (2018). *Guide to DHS Statistics*. ICF.

Crouse, D. L., Peters, P. A., Hystad, P., Brook, J. R., van Donkelaar, A., Martin, R. V., Villeneuve, P. J., Jerrett, M., Goldberg, M. S., Pope, C. A., Brauer, M., Brook, R. D., Robichaud, A., Menard, R., & Burnett, R. T. (2015). Ambient PM2.5, $O_3$, and $NO_2$ Exposures and Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environmental Health Perspectives*, *123*(11), 1180–1186. https://doi.org/10.1289/ehp.1409276

Dasgupta, S., Huq, M., Khaliquzzaman, M., Pandey, K., & Wheeler, D. (2006). Who suffers from indoor air pollution? Evidence from Bangladesh. *Health Policy and Planning*, *21*(6), 444–458. https://doi.org/10.1093/heapol/czl027

Devakumar, D., Semple, S., Osrin, D., Yadav, S. K., Kurmi, O. P., Saville, N. M., Shrestha, B., Manandhar, D. S., Costello, A., & Ayres, J. G. (2014). Biomass fuel use and the exposure of children to particulate air pollution in southern Nepal. *Environment International*, *66*, 79–87. https://doi.org/10.1016/j.envint.2014.01.011

Dherani, M., Pope, D., Mascarenhas, M., Smith, K. R., Weber, M., & Bruce, N. (2008). Indoor air pollution from unprocessed solid fuel use and pneumonia risk in children aged under five years: A systematic review and meta-analysis. *Bulletin of the World Health Organization*, *86*(5), 390-398C. https://doi.org/10.2471/blt.07.044529

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., & Schwartz, J. D. (2017). Air Pollution and Mortality in the Medicare Population. *New England Journal of Medicine*, *376*(26), 2513–2522. https://doi.org/10.1056/NEJMoa1702747

Dionisio, K. L., Howie, S. R. C., Dominici, F., Fornace, K. M., Spengler, J. D., Adegbola, R. A., & Ezzati, M. (2012). Household Concentrations and Exposure of Children to Particulate Matter from Biomass Fuels in The Gambia. *Environmental Science & Technology*, *46*(6), 3519–3527. https://doi.org/10.1021/es203047e

Dionisio, K. L., Howie, S. R. C., Dominici, F., Fornace, K. M., Spengler, J. D., Donkor, S., Chimah, O., Oluwalana, C., Ideh, R. C., Ebruke, B., Adegbola, R. A., & Ezzati, M. (2012). The exposure of infants and children to carbon monoxide from biomass fuels in The Gambia: A measurement and modeling study. *Journal of Exposure Science & Environmental Epidemiology*, *22*(2), 173–181. https://doi.org/10.1038/jes.2011.47

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., & Speizer, F. E. (1993). An Association between Air Pollution and Mortality in Six U.S. Cities. *New England Journal of Medicine*, *329*(24), 1753–1759. https://doi.org/10.1056/NEJM199312093292401

Dutta, A., Mukherjee, B., Das, D., Banerjee, A., & Ray, M. R. (2011). Hypertension with elevated levels of oxidized low-density lipoprotein and anticardiolipin antibody in the circulation of premenopausal

Indian women chronically exposed to biomass smoke during cooking. *Indoor Air*, *21*(2), 165–176. https://doi.org/10.1111/j.1600-0668.2010.00694.x

EquityTool. (2020, March). Guatemala—Equitytool. *Equity Tool*. https://www.equitytool.org/guatemala/

Ezzati, M, Saleh, H., & Kammen, D. M. (2000). The contributions of emissions and spatial microenvironments to exposure to indoor air pollution from biomass combustion in Kenya. *Environmental Health Perspectives*, *108*(9), 833–839.

Ezzati, Majid, & Baumgartner, J. C. (2017). Household energy and health: Where next for research and practice? *Lancet (London, England)*, *389*(10065), 130–132. https://doi.org/10.1016/S0140-6736(16)32506-5

Fatmi, Z., Ntani, G., & Coggon, D. (2019). Coronary heart disease, hypertension and use of biomass fuel among women: Comparative cross-sectional study. *BMJ Open*, *9*(8), e030881. https://doi.org/10.1136/bmjopen-2019-030881

Fitzgerald, C., Aguilar-Villalobos, M., Eppler, A. R., Dorner, S. C., Rathbun, S. L., & Naeher, L. P. (2012). Testing the effectiveness of two improved cookstove interventions in the Santiago de Chuco Province of Peru. *Science of The Total Environment*, *420*, 54–64. https://doi.org/10.1016/j.scitotenv.2011.10.059

Floras, J. S. (2009). Sympathetic Nervous System Activation in Human Heart Failure: Clinical Implications of an Updated Model. *Journal of the American College of Cardiology*, *54*(5), 375–385. https://doi.org/10.1016/j.jacc.2009.03.061

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, *2*, 150066. https://doi.org/10.1038/sdata.2015.66

Gakidou, E., Afshin, A., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abdulle, A. M., Abera, S. F., Aboyans, V., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., Abyu, G. Y., Adedeji, I. A., Adetokunboh, O., Afarideh, M., Agrawal, A., Agrawal, S., Ahmadieh, H., … Murray, C. J. L.

(2017). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*(10100), 1345–1422. https://doi.org/10.1016/S0140-6736(17)32366-8

Gautam, S. P. (2019). *National Ambient Air Quality Standards* (B-29016/20/90/PCI-L). Central Pollution Control Board of India.

*GBD Compare | IHME Viz Hub—Risk Factors in Guatemala*. (n.d.). Retrieved January 11, 2020, from http://ihmeuw.org/50c3

Giorgini, P., Di Giosia, P., Grassi, D., Rubenfire, M., Brook, R. D., & Ferri, C. (2016). Air Pollution Exposure and Blood Pressure: An Updated Review of the Literature. *Current Pharmaceutical Design*, *22*(1), 28–51. https://doi.org/10.2174/1381612822666151109111712

Glinianaia Svetlana V., Rankin Judith, Bell Ruth, Pless-Mulloli Tanja, & Howel Denise. (2004). Does Particulate Air Pollution Contribute to Infant Death? A Systematic Review. *Environmental Health Perspectives*, *112*(14), 1365–1370. https://doi.org/10.1289/ehp.6857

Gordon, S. B., Bruce, N. G., Grigg, J., Hibberd, P. L., Kurmi, O. P., Lam, K. H., Mortimer, K., Asante, K. P., Balakrishnan, K., Balmes, J., Bar-Zeev, N., Bates, M. N., Breysse, P. N., Buist, S., Chen, Z., Havens, D., Jack, D., Jindal, S., Kan, H., … Martin, W. J. (2014). Respiratory risks from household air pollution in low and middle income countries. *The Lancet Respiratory Medicine*, *2*(10), 823–860. https://doi.org/10.1016/S2213-2600(14)70168-7

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. *International Journal of Epidemiology*, *48*(4), 1125–1141. https://doi.org/10.1093/ije/dyz090

Gurley, E. S., Homaira, N., Salje, H., Ram, P. K., Haque, R., Petri, W., Bresee, J., Moss, W. J., Breysse, P., Luby, S. P., & Azziz-Baumgartner, E. (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. https://doi.org/10.1111/ina.12038

Health Effects Institute. (2019). *State of Global Air 2019* [Special Report]. Health Effects Institute.

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. https://doi.org/10.1038/s41586-018-0263-3

Heinzerling, A. P., Guarnieri, M. J., Mann, J. K., Diaz, J. V., Thompson, L. M., Diaz, A., Bruce, N. G., Smith, K. R., & Balmes, J. R. (2016). Lung function in woodsmoke-exposed Guatemalan children following a chimney stove intervention. *Thorax*, *71*(5), 421–428. https://doi.org/10.1136/thoraxjnl-2015-207783

Helen, G. S., Aguilar-Villalobos, M., Adetona, O., Cassidy, B., Bayer, C. W., Hendry, R., Hall, D. B., & Naeher, L. P. (2015). Exposure of Pregnant Women to Cookstove-Related Household Air Pollution in Urban and Periurban Trujillo, Peru. *Archives of Environmental & Occupational Health*, *70*(1), 10–18. https://doi.org/10.1080/19338244.2013.807761

Hermida, R. C., Ayala, D. E., Mojón, A., Fernández, J. R., Alonso, I., Silva, I., Ucieda, R., & Iglesias, M. (2000). Blood pressure patterns in normal pregnancy, gestational hypertension, and preeclampsia. *Hypertension (Dallas, Tex.: 1979)*, *36*(2), 149–158.

Hernández-Díaz, S., Schisterman, E. F., & Hernán, M. A. (2006). The birth weight "paradox" uncovered? *American Journal of Epidemiology*, *164*(11), 1115–1120. https://doi.org/10.1093/aje/kwj275

Huynh, M., Woodruff, T. J., Parker, J. D., & Schoendorf, K. C. (2006). Relationships between air pollution and preterm birth in California. *Paediatric and Perinatal Epidemiology*, *20*(6), 454–461. https://doi.org/10.1111/j.1365-3016.2006.00759.x

Hystad, P., Duong, M., Brauer, M., Larkin, A., Arku, R., Kurmi, O. P., Fan, W. Q., Avezum, A., Azam, I., Chifamba, J., Dans, A., du Plessis, J. L., Gupta, R., Kumar, R., Lanas, F., Liu, Z., Lu, Y., Lopez-Jaramillo, P., Mony, P., … [on behalf of Prospective Urban and Rural Epidemiological (PURE) Study investigators]. (2019). Health Effects of Household Solid Fuel Use: Findings from 11 Countries within the Prospective Urban and Rural Epidemiology Study. *Environmental Health Perspectives*, *127*(5), 57003. https://doi.org/10.1289/EHP3915

IHME. (2020, February 1). *GBD Compare—Air Pollution and Child Mortality*. http://ihmeuw.org/511f

IIPS/India, I. I. for P. S.-, & ICF. (2017). *India National Family Health Survey NFHS-4 2015-16*. https://dhsprogram.com/publications/publication-FR339-DHS-Final-Reports.cfm

Johnson, M., Piedrahita, R., Garland, C., Pillarisetti, A., Sambandam, S., Gurusamy, T., Mukhopadhyay, K., Balakrishnan, K., León, O. F. de, Mollinedo, E., McCracken, J. P., Liao, J., Burrowes, V. J., Kephart, J., Rio, M. F. D., Majorin, F., Rosa, G., Kirby, M. A., Steenland, K., … Clasen, T. (2018). Exposures to PM2.5 Associated with LPG Stove and Fuel Interventions in Four Countries: Pilot Results from the HAPIN Trial. *ISEE Conference Abstracts*. https://ehp.niehs.nih.gov/doi/10.1289/isesisee.2018.O02.03.31

K. Didan. (2015). *MOD13C1 MODIS/Terra Vegetation Indices 16-Day L3 Global 0.05Deg CMG V006*. University of Arizona, Alfredo Huete - University of Technology Sydney and MODAPS SIPS - NASA. https://lpdaac.usgs.gov/products/mod13c1v006/

Karumanchi, S. A., & Levine, R. J. (2010). How does smoking reduce the risk of preeclampsia? *Hypertension*, *55*(5), 1100–1101. https://doi.org/10.1161/HYPERTENSIONAHA.109.148973

Karumanchi S. Ananth, & Levine Richard J. (2010). How Does Smoking Reduce the Risk of Preeclampsia? *Hypertension*, *55*(5), 1100–1101. https://doi.org/10.1161/HYPERTENSIONAHA.109.148973

Koster, R. D. (2015). *MERRA-2: Initial Evaluation of the Climate* (NASA/TM–2015-104606/Vol. 43; Technical Report Series on Global Modeling and Data Assimilation, Volume 43). Goddard Space Flight Center. https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf

Krotkov, N., Lamsal, L., Marchenko, S., Celarier, E., & Veefkind, P. (2019). *OMI/Aura Nitrogen Dioxide (NO2) Total and Tropospheric Column 1-orbit L2 Swath 13x24 km V003*. Goddard Earth Sciences Data and Information Services Center (GES DISC). https://disc.gsfc.nasa.gov/datasets/OMNO2_V003/summary?keywords=omno2

LANCE FIRMS. (2016). *NASA Near Real-Time and MCD14DL MODIS Active Fire Detections* [Data set]. NASA ESDIS. https://doi.org/10.5067/FIRMS/MODIS/MCD14DL.NRT.006

Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N. (Nil), Baldé, A. B., Bertollini, R., Bose-O'Reilly, S., Boufford, J. I., Breysse, P. N., Chiles, T., Mahidol, C., Coll-Seck, A. M., Cropper,

M. L., Fobil, J., Fuster, V., Greenstone, M., Haines, A., … Zhong, M. (2018). The Lancet Commission on pollution and health. *The Lancet*, *391*(10119), 462–512. https://doi.org/10.1016/S0140-6736(17)32345-0

Li, L., Yang, A., He, X., Liu, J., Ma, Y., Niu, J., & Luo, B. (2020). Indoor air pollution from solid fuels and hypertension: A systematic review and meta-analysis. *Environmental Pollution*, *259*, 113914. https://doi.org/10.1016/j.envpol.2020.113914

Li, Q., Wang, Y., Guo, Y., Zhou, H., Wang, X., Wang, Q., Shen, H., Zhang, Y., Yan, D., Zhang, Y., Zhang, H., Li, S., Chen, G., Lin, L., Zhao, J., He, Y., Yang, Y., Xu, J., Wang, Y., … Ma, X. (2018). Effect of airborne particulate matter of 2.5 μm or less on preterm birth: A national birth cohort study in China. *Environment International*, *121*, 1128–1136. https://doi.org/10.1016/j.envint.2018.10.025

Liang, R., Zhang, B., Zhao, X., Ruan, Y., Lian, H., & Fan, Z. (2014). Effect of exposure to PM2.5 on blood pressure: A systematic review and meta-analysis. *Journal of Hypertension*, *32*(11), 2130–2140; discussion 2141. https://doi.org/10.1097/HJH.0000000000000342

Liao, J., McCracken, J. P., Piedrahita, R., Thompson, L., Mollinedo, E., Canuz, E., De Léon, O., Díaz-Artiga, A., Johnson, M., Clark, M., Pillarisetti, A., Kearns, K., Naeher, L., Steenland, K., Checkley, W., Peel, J., Clasen, T. F., Aravindalochanan, V., Balakrishnan, K., … HAPIN investigators. (2019). The use of bluetooth low energy Beacon systems to estimate indirect personal exposure to household air pollution. *Journal of Exposure Science & Environmental Epidemiology*. https://doi.org/10.1038/s41370-019-0172-z

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M. A., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., … Ezzati, M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260. https://doi.org/10.1016/S0140-6736(12)61766-8

Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A. M., Guo, Y., Tong, S., Coelho, M. S. Z. S., Saldiva, P. H. N., Lavigne, E., Matus, P., Valdes Ortega, N., Osorio Garcia, S., Pascal, M., Stafoggia, M., Scortichini, M., Hashizume, M., Honda, Y., Hurtado-Díaz, M., Cruz, J., … Kan, H. (2019). Ambient Particulate Air Pollution and Daily Mortality in 652 Cities. *The New England Journal of Medicine*, *381*(8), 705–715. https://doi.org/10.1056/NEJMoa1817364

Loomis, D., Castillejos, M., Gold, D. R., McDonnell, W., & Borja-Aburto, V. H. (1999). Air Pollution and Infant Mortality in Mexico City. *Epidemiology*, *10*(2), 118–123.

Luque-Fernandez, M. A., Zoega, H., Valdimarsdottir, U., & Williams, M. A. (2016). Deconstructing the smoking-preeclampsia paradox through a counterfactual framework. *European Journal of Epidemiology*, *31*(6), 613–623. https://doi.org/10.1007/s10654-016-0139-5

Macdonald-Wallis Corrie, Lawlor Debbie A., Fraser Abigail, May Margaret, Nelson Scott M., & Tilling Kate. (2012). Blood Pressure Change in Normotensive, Gestational Hypertensive, Preeclamptic, and Essential Hypertensive Pregnancies. *Hypertension*, *59*(6), 1241–1248. https://doi.org/10.1161/HYPERTENSIONAHA.111.187039

Makri, A., & Stilianakis, N. I. (2008). Vulnerability to air pollution health effects. *International Journal of Hygiene and Environmental Health*, *211*(3), 326–336. https://doi.org/10.1016/j.ijheh.2007.06.005

McCracken, J. P., Schwartz, J., Diaz, A., Bruce, N., & Smith, K. R. (2013). Longitudinal Relationship between Personal CO and Personal PM2.5 among Women Cooking with Woodfired Cookstoves in Guatemala. *PLOS ONE*, *8*(2), e55670. https://doi.org/10.1371/journal.pone.0055670

McCracken, J. P., Smith, K. R., Díaz, A., Mittleman, M. A., & Schwartz, J. (2007). Chimney Stove Intervention to Reduce Long-term Wood Smoke Exposure Lowers Blood Pressure among Guatemalan Women. *Environmental Health Perspectives*, *115*(7), 996–1001. https://doi.org/10.1289/ehp.9888

McCracken, J. P., Wellenius, G. A., Bloomfield, G. S., Brook, R. D., Tolunay, H. E., Dockery, D. W., Rabadan-Diehl, C., Checkley, W., & Rajagopalan, S. (2012). Household Air Pollution from Solid Fuel

Use: Evidence for Links to CVD. *Global Heart*, *7*(3), 223–234. https://doi.org/10.1016/j.gheart.2012.06.010

Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J., & Schipper, A. M. (2018). Global patterns of current and future road infrastructure. *Environmental Research Letters*, *13*(6), 064006. https://doi.org/10.1088/1748-9326/aabd42

Ministry of Environment & Forests of Government of India. (2018, April 14). *Central Pollution Control Board (CPCB)*. http://www.cpcb.gov.in/CAAQM/frmUserAvgReportCriteria.aspx

MODIS Atmosphere Science Team. (2015). *MOD04_L2 MODIS/Terra Aerosol 5-Min L2 Swath 10km*. NASA Level 1 and Atmosphere Archive and Distribution System. http://modaps.nascom.nasa.gov/services/about/products/c6/MOD04_L2.html

MODIS Science Team. (2014). *MYD04_L2 MODIS/Aqua Aerosol 5-Min L2 Swath 10km* [Data set]. NASA Level 1 and Atmosphere Archive and Distribution System. https://doi.org/10.5067/modis/myd04_l2.006

Mol, B. W. J., Roberts, C. T., Thangaratinam, S., Magee, L. A., de Groot, C. J. M., & Hofmeyr, G. J. (2016). Pre-eclampsia. *Lancet (London, England)*, *387*(10022), 999–1011. https://doi.org/10.1016/S0140-6736(15)00070-7

Mortimer, K., Ndamala, C. B., Naunje, A. W., Malava, J., Katundu, C., Weston, W., Havens, D., Pope, D., Bruce, N. G., Nyirenda, M., Wang, D., Crampin, A., Grigg, J., Balmes, J., & Gordon, S. B. (2017). A cleaner burning biomass-fuelled cookstove intervention to prevent pneumonia in children under 5 years old in rural Malawi (the Cooking and Pneumonia Study): A cluster randomised controlled trial. *The Lancet*, *389*(10065), 167–175. https://doi.org/10.1016/S0140-6736(16)32507-7

Mukhopadhyay, R., Sambandam, S., Pillarisetti, A., Jack, D., Mukhopadhyay, K., Balakrishnan, K., Vaswani, M., Bates, M. N., Kinney, P., Arora, N., & Smith, K. (2012). Cooking practices, air quality, and the acceptability of advanced cookstoves in Haryana, India: An exploratory study to inform large-scale interventions. *Global Health Action*, *5*(1), 19016. https://doi.org/10.3402/gha.v5i0.19016

Naeher, Brauer, M., Lipsett, M., Zelikoff, J. T., Simpson, C. D., Koenig, J. Q., & Smith, K. R. (2007). Woodsmoke health effects: A review. *Inhalation Toxicology*, *19*(1), 67–106. https://doi.org/10.1080/08958370600985875

Naeher, L. P., Leaderer, B. P., & Smith, K. R. (2000). Particulate matter and carbon monoxide in highland Guatemala: Indoor and outdoor levels from traditional and improved wood stoves and gas stoves. *Indoor Air*, *10*(3), 200–205.

Naghavi, M., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., Aboyans, V., Adetokunboh, O., Afshin, A., Agrawal, A., Ahmadi, A., Ahmed, M. B., Aichour, A. N., Aichour, M. T. E., Aichour, I., Aiyar, S., Alahdab, F., Al-Aly, Z., Alam, K., … Murray, C. J. L. (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*(10100), 1151–1210. https://doi.org/10.1016/S0140-6736(17)32152-9

NASA's Land Processes Distributed Active Archive Center (LP DAAC). (2009). *ASTER Global Digital Elevation Model Version 2* [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/aster/astgtm.002

Neupane, M., Basnyat, B., Fischer, R., Froeschl, G., Wolbers, M., & Rehfuess, E. A. (2015). Sustained use of biogas fuel and blood pressure among women in rural Nepal. *Environmental Research*, *136*, 343–351. https://doi.org/10.1016/j.envres.2014.10.031

Ni, K., Carter, E., Schauer, J. J., Ezzati, M., Zhang, Y., Niu, H., Lai, A. M., Shan, M., Wang, Y., Yang, X., & Baumgartner, J. (2016). Seasonal variation in outdoor, indoor, and personal air pollution exposures of women using wood stoves in the Tibetan Plateau: Baseline assessment for an energy intervention study. *Environment International*, *94*, 449–457. https://doi.org/10.1016/j.envint.2016.05.029

Norris, C., Goldberg, M. S., Marshall, J. D., Valois, M.-F., Pradeep, T., Narayanswamy, M., Jain, G., Sethuraman, K., & Baumgartner, J. (2016). A panel study of the acute effects of personal exposure to household air pollution on ambulatory blood pressure in rural Indian women. *Environmental Research*, *147*, 331–342. https://doi.org/10.1016/j.envres.2016.02.024

Northcross, A., Chowdhury, Z., McCracken, J., Canuz, E., & Smith, K. R. (2010). Estimating personal PM2.5 exposures using CO measurements in Guatemalan households cooking with wood fuel. *Journal of Environmental Monitoring: JEM*, *12*(4), 873–878. https://doi.org/10.1039/b916068j

OpenAQ. (2020a, January). *OpenAQ - China*. OpenAQ - China. https://openaq.org/#/countries/CN?_k=c0qotf

OpenAQ. (2020b, January). *OpenAQ - India*. OpenAQ - India. https://openaq.org/#/countries/IN?_k=zbn1uv

Pedersen, M., Stayner, L., Slama, R., Sørensen, M., Figueras, F., Nieuwenhuijsen, M. J., Raaschou-Nielsen, O., & Dadvand, P. (2014). Ambient air pollution and pregnancy-induced hypertensive disorders: A systematic review and meta-analysis. *Hypertension (Dallas, Tex.: 1979)*, *64*(3), 494–500. https://doi.org/10.1161/HYPERTENSIONAHA.114.03545

Peña, M. B., Romero, K. M., Velazquez, E. J., Davila-Roman, V. G., Gilman, R. H., Wise, R. A., Miranda, J. J., & Checkley, W. (2015). Relationship Between Daily Exposure to Biomass Fuel Smoke and Blood Pressure in High-Altitude PeruNovelty and Significance. *Hypertension*, *65*(5), 1134–1140. https://doi.org/10.1161/HYPERTENSIONAHA.114.04840

Piedrahita, R. A. (2017). *On the Assessment of Air Pollution and Behavior Within a Cookstove Intervention Study in Northern Ghana and Development of Improved Measurement Techniques* [Ph.D., University of Colorado at Boulder]. https://search-proquest-com.proxy.library.emory.edu/docview/1904508224/abstract/BC813B643102450CPQ/1

Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., & Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*, *287*(9), 1132–1141. https://doi.org/10.1001/jama.287.9.1132

Pope C. Arden, Lefler Jacob S., Ezzati Majid, Higbee Joshua D., Marshall Julian D., Kim Sun-Young, Bechle Matthew, Gilliat Kurtis S., Vernon Spencer E., Robinson Allen L., & Burnett Richard T. (2019). Mortality Risk and Fine Particulate Air Pollution in a Large, Representative Cohort of U.S. Adults. *Environmental Health Perspectives*, *127*(7), 077007. https://doi.org/10.1289/EHP4438

Pope, D. P., Mishra, V., Thompson, L., Siddiqui, A. R., Rehfuess, E. A., Weber, M., & Bruce, N. G. (2010a). Risk of Low Birth Weight and Stillbirth Associated With Indoor Air Pollution From Solid Fuel Use in Developing Countries. *Epidemiologic Reviews*, *32*(1), 70–81. https://doi.org/10.1093/epirev/mxq005

Pope, D. P., Mishra, V., Thompson, L., Siddiqui, A. R., Rehfuess, E. A., Weber, M., & Bruce, N. G. (2010b). Risk of Low Birth Weight and Stillbirth Associated With Indoor Air Pollution From Solid Fuel Use in Developing Countries. *Epidemiologic Reviews*, *32*(1), 70–81. https://doi.org/10.1093/epirev/mxq005

Quansah, R., Semple, S., Ochieng, C. A., Juvekar, S., Armah, F. A., Luginaah, I., & Emina, J. (2017). Effectiveness of interventions to reduce household air pollution and/or improve health in homes using solid fuel in low-and-middle income countries: A systematic review and meta-analysis. *Environment International*, *103*, 73–90. https://doi.org/10.1016/j.envint.2017.03.010

Quinn, A. K., Ae-Ngibise, K. A., Jack, D. W., Boamah, E. A., Enuameh, Y., Mujtaba, M. N., Chillrud, S. N., Wylie, B. J., Owusu-Agyei, S., Kinney, P. L., & Asante, K. P. (2016). Association of Carbon Monoxide exposure with blood pressure among pregnant women in rural Ghana: Evidence from GRAPHS. *International Journal of Hygiene and Environmental Health*, *219*(2), 176–183. https://doi.org/10.1016/j.ijheh.2015.10.004

Quinn, A. K., Ae-Ngibise, K. A., Kinney, P. L., Kaali, S., Wylie, B. J., Boamah, E., Shimbo, D., Agyei, O., Chillrud, S. N., Mujtaba, M., Schwartz, J. E., Abdalla, M., Owusu-Agyei, S., Jack, D. W., & Asante, K. P. (2017). Ambulatory monitoring demonstrates an acute association between cookstove-related carbon monoxide and blood pressure in a Ghanaian cohort. *Environmental Health*, *16*(1), 76. https://doi.org/10.1186/s12940-017-0282-9

Romieu, I., Samet, J. M., Smith, K. R., & Bruce, N. (2002). Outdoor air pollution and acute respiratory infections among children in developing countries. *Journal of Occupational and Environmental Medicine*, *44*(7), 640–649. https://doi.org/10.1097/00043764-200207000-00010

Rosenthal, J., Quinn, A., Grieshop, A. P., Pillarisetti, A., & Glass, R. I. (2018). Clean cooking and the SDGs: Integrated analytical approaches to guide energy interventions for health and environment goals. *Energy for Sustainable Development*, *42*, 152–159. https://doi.org/10.1016/j.esd.2017.11.003

Roth, G. A., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, H. T., Abebe, M., Abebe, Z., Abejie, A. N., Abera, S. F., Abil, O. Z., Abraha, H. N., … Murray, C. J. L. (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7

Ruiz-Mercado, I., Lam, N., Canuz, E., Acevedi, R., & Smith, K. R. (2010). *Modeling Variability in Kitchen Time-activity of Adult Guatemalan Women Cooking with Biomass*. 20th Annual ISES Meeting, Seoul, South Korea. https://www.ce.berkeley.edu/sites/default/files/news/124/ISEE-ISEA%202010_Ilse%20Ruiz%20Mercado-final%20formated.pdf

Scheers Hans, Mwalili Samuel M., Faes Christel, Fierens Frans, Nemery Benoit, & Nawrot Tim S. (2011). Does Air Pollution Trigger Infant Mortality in Western Europe? A Case-Crossover Study. *Environmental Health Perspectives*, *119*(7), 1017–1022. https://doi.org/10.1289/ehp.1002913

Shaddick, G., Thomas, M. L., Amini, H., Broday, D., Cohen, A., Frostad, J., Green, A., Gumy, S., Liu, Y., Martin, R. V., Pruss-Ustun, A., Simpson, D., van Donkelaar, A., & Brauer, M. (2018). Data Integration for the Assessment of Population Exposure to Ambient Air Pollution for Global Burden of Disease Assessment. *Environmental Science & Technology*, *52*(16), 9069–9078. https://doi.org/10.1021/acs.est.8b02864

Shen, M., Smith, G. N., Rodger, M., White, R. R., Walker, M. C., & Wen, S. W. (2017). Comparison of risk factors and outcomes of gestational hypertension and pre-eclampsia. *PLoS ONE*, *12*(4). https://doi.org/10.1371/journal.pone.0175914

Shupler, M., Godwin, W., Frostad, J., Gustafson, P., Arku, R. E., & Brauer, M. (2018). Global estimation of exposure to fine particulate matter (PM2.5) from household air pollution. *Environment International*, *120*, 354–363. https://doi.org/10.1016/j.envint.2018.08.026

Sidhu, M. K., Ravindra, K., Mor, S., & John, S. (2017). Household air pollution from various types of rural kitchens and its exposure assessment. *Science of The Total Environment*, *586*, 419–429. https://doi.org/10.1016/j.scitotenv.2017.01.051

Smith, K. R., Bruce, N., Balakrishnan, K., Adair-Rohani, H., Balmes, J., Chafe, Z., Dherani, M., Hosgood, H. D., Mehta, S., Pope, D., & Rehfuess, E. (2014). Millions Dead: How Do We Know and What Does It Mean? Methods Used in the Comparative Risk Assessment of Household Air Pollution. *Annual Review of Public Health*, *35*(1), 185–206. https://doi.org/10.1146/annurev-publhealth-032013-182356

Smith, K. R., McCracken, J. P., Weber, M. W., Hubbard, A., Jenny, A., Thompson, L. M., Balmes, J., Diaz, A., Arana, B., & Bruce, N. (2011b). Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): A randomised controlled trial. *The Lancet*, *378*(9804), 1717–1726. https://doi.org/10.1016/S0140-6736(11)60921-5

Smith, K. R., McCracken, J. P., Weber, M. W., Hubbard, A., Jenny, A., Thompson, L. M., Balmes, J., Diaz, A., Arana, B., & Bruce, N. (2011a). Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): A randomised controlled trial. *The Lancet*, *378*(9804), 1717–1726. https://doi.org/10.1016/S0140-6736(11)60921-5

Son, J.-Y., Lee, H. J., Koutrakis, P., & Bell, M. L. (2017). Pregnancy and Lifetime Exposure to Fine Particulate Matter and Infant Mortality in Massachusetts, 2001–2007. *American Journal of Epidemiology*, *186*(11), 1268–1276. 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 PM2.5 and later-childhood height-for-age in India: An observational study. *Environmental Health*, *18*(1), 62. https://doi.org/10.1186/s12940-019-0501-7

Stanaway, J. D., Afshin, A., Gakidou, E., Lim, S. S., Abate, D., Abate, K. H., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, M., Abebe, Z., Abera, S. F., Abil, O. Z., Abraha, H. N., … Murray, C. J. L. (2018). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: A systematic analysis

for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1923–1994. https://doi.org/10.1016/S0140-6736(18)32225-6

StataCorp. (2013). *Stata: Release 13. Statistical Software*. StataCorp LP.

Steenland, K., Pillarisetti, A., Kirby, M., Peel, J., Clark, M., Checkley, W., Chang, H. H., & Clasen, T. (2018). Modeling the potential health benefits of lower household air pollution after a hypothetical liquified petroleum gas (LPG) cookstove intervention. *Environment International*, *111*, 71–79. https://doi.org/10.1016/j.envint.2017.11.018

Stein-Zweers, D., & Veefkind, P. (2019). *OMI/Aura Multi-wavelength Aerosol Optical Depth and Single Scattering Albedo 1-orbit L2 Swath 13x24 km V003*. NASA Goddard Space Flight Center, Goddard Earth Sciences Data and Information Services Center (GES DISC). https://disc.gsfc.nasa.gov/datasets/OMAERO_V003/summary?keywords=omaero

Stieb, D. M., Chen, L., Eshoul, M., & Judek, S. (2012). Ambient air pollution, birth weight and preterm birth: A systematic review and meta-analysis. *Environmental Research*, *117*, 100–111. https://doi.org/10.1016/j.envres.2012.05.007

Subramanian, S. V., Ackerson, L. K., Smith, G. D., & John, N. A. (2009). Association of Maternal Height With Child Mortality, Anthropometric Failure, and Anemia in India. *JAMA*, *301*(16), 1691–1701. https://doi.org/10.1001/jama.2009.548

Sun, X., Luo, X., Zhao, C., Chung Ng, R. W., Lim, C. E. D., Zhang, B., & Liu, T. (2015). The association between fine particulate matter exposure during pregnancy and preterm birth: A meta-analysis. *BMC Pregnancy and Childbirth*, *15*. https://doi.org/10.1186/s12884-015-0738-2

Thakur, M., Nuyts, P. A. W., Boudewijns, E. A., Kim, J. F., Faber, T., Babu, G. R., Schayck, O. C. P. van, & Been, J. V. (2018). Impact of improved cookstoves on women's and child health in low and middle income countries: A systematic review and meta-analysis. *Thorax*, *73*(11), 1026–1040. https://doi.org/10.1136/thoraxjnl-2017-210952

Thompson, L. M., Bruce, N., Eskenazi, B., Diaz, A., Pope, D., & Smith, K. R. (2011). Impact of Reduced Maternal Exposures to Wood Smoke from an Introduced Chimney Stove on Newborn Birth Weight

in Rural Guatemala. *Environmental Health Perspectives*, *119*(10), 1489–1494. https://doi.org/10.1289/ehp.1002928

US EPA, O. (2014, April 10). *NAAQS Table* [Policies and Guidance]. US EPA. https://www.epa.gov/criteria-air-pollutants/naaqs-table

van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin, A., Sayer, A. M., & Winker, D. M. (2016). Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environmental Science & Technology*, *50*(7), 3762–3772. https://doi.org/10.1021/acs.est.5b05833

Volckens, J., Quinn, C., Leith, D., Mehaffy, J., Henry, C. S., & Miller-Lionberg, D. (2017). Development and evaluation of an ultrasonic personal aerosol sampler. *Indoor Air*, *27*(2), 409–416. https://doi.org/10.1111/ina.12318

Vu, B. N., Sánchez, O., Bi, J., Xiao, Q., Hansel, N. N., Checkley, W., Gonzales, G. F., Steenland, K., & Liu, Y. (2019). Developing an Advanced PM2.5 Exposure Model in Lima, Peru. *Remote Sensing*, *11*(6), 641. https://doi.org/10.3390/rs11060641

Wang, J., Cao, H., Sun, D., Qi, Z., Guo, C., Peng, W., Sun, Y., Xie, Y., Liu, X., Li, B., Luo, Y., Pan, Y., Li, Y., & Zhang, L. (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. *Environmental Research*, *176*, 108531. https://doi.org/10.1016/j.envres.2019.108531

WHO Regional Office for Europe. (2013). *Review of evidence on health aspects of air pollution – REVIHAAP Project: Technical Report*. WHO Regional Office for Europe. http://www.ncbi.nlm.nih.gov/books/NBK361805/

Woodruff T J, Grillo J, & Schoendorf K C. (1997). The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. *Environmental Health Perspectives*, *105*(6), 608–612. https://doi.org/10.1289/ehp.97105608

Woodruff Tracey J., Parker Jennifer D., & Schoendorf Kenneth C. (2006). Fine Particulate Matter (PM2.5) Air Pollution and Selected Causes of Postneonatal Infant Mortality in California. *Environmental Health Perspectives*, *114*(5), 786–790. https://doi.org/10.1289/ehp.8484

World Health Organization. (2005a). *Indoor air pollution and household energy monitoring: Workshop resources* (No. 9241593768). World Health Organization. http://www.who.int/iris/handle/10665/43371

World Health Organization. (2005b). *WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide Global Update 2005 Summary of Risk Assessment*. World Health Organization.

World Health Organization. (2014). *WHO Guidelines for Indoor Air Quality: Household Fuel Combustion*.

World Health Organization, & United Nations Children's Fund (UNICEF). (2017). *JMP Methodology 2017 Update & SDG Baselines*. https://www.ircwash.org/sites/default/files/jmp-2017-update-methodology.pdf

Wylie, B. J., Kishashu, Y., Matechi, E., Zhou, Z., Coull, B., Abioye, A. I., Dionisio, K. L., Mugusi, F., Premji, Z., Fawzi, W., Hauser, R., & Ezzati, M. (2017). Maternal exposure to carbon monoxide and fine particulate matter during pregnancy in an urban Tanzanian cohort. *Indoor Air*, *27*(1), 136–146. https://doi.org/10.1111/ina.12289

Wylie, Blair J., Singh, M. P., Coull, B. A., Quinn, A., Yeboah-Antwi, K., Sabin, L., Hamer, D. H., Singh, N., & MacLeod, W. B. (2015). Association between wood cooking fuel and maternal hypertension at delivery in central East India. *Hypertension in Pregnancy*, *34*(3), 355–368. https://doi.org/10.3109/10641955.2015.1046604

Xiao, Q., Zhang, H., Choi, M., Li, S., Kondragunta, S., Kim, J., Holben, B., Levy, R. C., & Liu, Y. (2016). Evaluation of VIIRS, GOCI, and MODIS Collection 6 AOD retrievals against ground sunphotometer observations over East Asia. *Atmospheric Chemistry and Physics*, *16*(3), 1255–1269. https://doi.org/10.5194/acp-16-1255-2016

Xiao, Qingyang, Chang, H. H., Geng, G., & Liu, Y. (2018). An Ensemble Machine-Learning Model To Predict Historical PM2.5 Concentrations in China from Satellite Data. *Environmental Science & Technology*, *52*(22), 13260–13269. https://doi.org/10.1021/acs.est.8b02917

Yin Peng, Brauer Michael, Cohen Aaron, Burnett Richard T., Liu Jiangmei, Liu Yunning, Liang Ruiming, Wang Weihua, Qi Jinlei, Wang Lijun, & Zhou Maigeng. (2017). Long-term Fine Particulate Matter Exposure and Nonaccidental and Cause-specific Mortality in a Large National Cohort of Chinese Men. *Environmental Health Perspectives*, *125*(11), 117002. https://doi.org/10.1289/EHP1673

Young, B. N., Clark, M. L., Rajkumar, S., Benka-Coker, M. L., Bachand, A., Brook, R. D., Nelson, T. L., Volckens, J., Reynolds, S. J., L'Orange, C., Good, N., Koehler, K., Africano, S., Osorto Pinel, A. B., & Peel, J. L. (2019). Exposure to household air pollution from biomass cookstoves and blood pressure among women in rural Honduras: A cross-sectional study. *Indoor Air*, *29*(1), 130–142. https://doi.org/10.1111/ina.12507

Yu, K., Qiu, G., Chan, K.-H., Lam, K.-B. H., Kurmi, O. P., Bennett, D. A., Yu, C., Pan, A., Lv, J., Guo, Y., Bian, Z., Yang, L., Chen, Y., Hu, F. B., Chen, Z., Li, L., & Wu, T. (2018). Association of Solid Fuel Use With Risk of Cardiovascular and All-Cause Mortality in Rural China. *JAMA*, *319*(13), 1351–1361. https://doi.org/10.1001/jama.2018.2151

Zeger S L, Thomas D, Dominici F, Samet J M, Schwartz J, Dockery D, & Cohen A. (2000). Exposure measurement error in time-series studies of air pollution: Concepts and consequences. *Environmental Health Perspectives*, *108*(5), 419–426. https://doi.org/10.1289/ehp.00108419

Zuk, M., Rojas, L., Blanco, S., Serrano, P., Cruz, J., Angeles, F., Tzintzun, G., Armendariz, C., Edwards, R. D., Johnson, M., Riojas-Rodriguez, H., & Masera, O. (2007). The impact of improved wood-burning stoves on fine particulate matter concentrations in rural Mexican homes. *Journal of Exposure Science & Environmental Epidemiology*, *17*(3), 224–232. https://doi.org/10.1038/sj.jes.7500499