STATISTICAL ANALYSIS PLAN

Household Air Pollution Exposure and Severe Infant Pneumonia: Exposure-Response Analysis

Version 1
November 4, 2024

Household air pollution and health: A multi-country LPG stove intervention trial (HAPIN)

Trial Registration: NCT02944682
Protocol Version: HAPIN - Main Study Protocol v17.0

Modification History: NA

Version	Date	Changes
1.0	November 4, 2024	

1. INTRODUCTION

This document contains the statistical analysis plan (SAP) for the exposure-response analysis of household air pollution (HAP) exposure and infant severe pneumonia. HAP exposure was measured during 24-hour periods repeatedly during pregnancy and infancy. HAP indicator pollutants include air concentrations of final particles, black carbon, and carbon monoxide. Severe pneumonia is one of the four primary outcomes of the HAPIN trial. This SAP covers the exposure-response analysis of severe pneumonia and other respiratory endpoints of the trial. The pneumonia outcomes analyzed were detected by active surveillance, record review, and verbal autopsy.

1.1. Background and Rationale

Globally, nearly 3 billion people rely on solid fuels for household cooking and heating, the vast majority in low- and middle-income countries (LMICs).[3] The emissions from solid fuel use comprise a major source of household air pollution (HAP), which is recognized as an important environmental risk factor for morbidity and mortality, accounting for an estimated 2.3 million premature deaths globally. Previous large-scale interventions have provided cleaner cookstoves but have failed to reduce exposure to levels that might be required to produce meaningful or detectable health improvements. There have been no large-scale field trials with liquefied petroleum gas (LPG) cookstoves, likely the cleanest scalable intervention.

The HAPIN trial will provide evidence to inform national and global policies on scaling up LPG stoves among vulnerable populations. Ultimately, this will facilitate deeper policy-level discussions as well as identify requirements for initiating and sustaining HAP interventions globally.

A large burden of severe pneumonia is attributed to HAP in young children in low- and middle-income countries, who often have high prenatal and postnatal HAP exposure. Previous cleaner biomass stove and clean fuel intervention trials have failed to provide unequivocal evidence of intervention effects on pneumonia incidence in infants, despite observational evidence that exposure to biomass smoke and other sources of combustion-generated pollution cause childhood pneumonia. Failure to provide evidence of the expected effects in intervention trials might hinder promotion of cleaner cooking stoves for the health benefits. A Exposure-response analysis of the causal relationship between HAP and pneumonia may further our understanding of the expected magnitude of effects from interventions.

1.2. HAPIN Study Overview

The aim of the HAPIN study is to conduct a randomized controlled trial of LPG stove and fuel distribution in 3200 households in four LMICs (India, Guatemala, Peru, and Rwanda) to deliver rigorous evidence regarding potential health benefits across the lifespan. Each intervention site recruited 800 pregnant women (aged 18-<35 years, 9 to <20 weeks' gestation), and randomly assigned half their households to receive LPG stoves and an 18-month supply of LPG. Controls did not receive the intervention at the commencement of the trial and were generally expected to continue cooking with solid biomass fuels; the control group was compensated for

their participation in the study. The mother was followed along with her child until the child was 1 year old. In households with a second, non-pregnant older adult woman (aged 40 to <80 years) we also enrolled and followed her during the 18-month follow-up period to assess cardiopulmonary, metabolic, and cancer outcomes. To optimize intervention use, we implemented behaviour change strategies. We assessed cookstove use, conducted repeated personal exposure assessments to HAP ($PM_{2.5}$, black carbon, carbon monoxide), and collected dried blood spots (DBS) and urinary samples for biomarker analysis and biospecimen storage on all participants at multiple time points. The primary outcomes were birth weight, severe pneumonia, and stunting at age 1 year in the child, and blood pressure in the older adult woman.

1.3. Study Objectives

The HAPIN trial will address the following specific aims: (1) using an intention-to-treat analysis, determine the effect of a randomized LPG stove and fuel intervention on health in four diverse LMIC populations using a common protocol; (2) determine the exposure-response relationships for HAP and four primary health outcomes (birthweight, stunting and severe pneumonia in infants, and blood pressure in older adults); and (3) determine relationships between the LPG intervention and both targeted and exploratory biomarkers of exposure and health. This SAP addresses a subset of the second aim, examining an exposure-response relationship between HAP exposure and severe infant pneumonia.

2. STUDY METHODS

2.1. Trial design

HAPIN was a randomized, two-arm intervention trial with parallel assignment. Study sites in the four countries (Guatemala, India, Peru, Rwanda) were selected and evaluated based on activities conducted in the formative research. HAPIN used a rolling recruitment process whereby each International Research Center (IRC) enrolled 800 pregnant women (one per household) and additionally up to 120 older adult women from the same households who meet inclusion/exclusion criteria (Section 4.1). Key characteristics of each study site are given in Table 2 of the HAPIN design publication.[4] Recruitment and enrollment occurred over approximately 15 months at ~53 pregnant women/8-12 older adult women per month per IRC. All participants were followed longitudinally for ~18 months (until the child was age 1 year).

2.2. Randomization

To ensure balance between arms, households have been randomly allocated to intervention or control arms after consent was given and baseline data are collected. To maintain balance of treatment assignments within each study site at the IRCs, a total of 10 randomization strata are implemented as follows.

- The India IRC stratified randomization across two sites.
- The Peru IRC stratified randomization across six sites.
- Guatemala and Rwanda have one site each.

Separate randomization lists have been generated for each field team conducting randomization at each IRC. Two randomization lists are produced for each of those field teams: one for households that include an older adult woman (OAW), and one for households that do not. Additional details on randomization of households can be found in the HAPIN protocol.

2.3. Sample Size Considerations

The study population size was decided based on sample size calculation for a test of the relative risk of pneumonia associated with the randomized LPG intervention. This separate ITT analysis has been reported.[1] We assumed a 1-year cumulative incidence rate among the control group as $p_1 = 0.06$ [a weighted average from Mackenzie et al. 2014, Gupta et al. 2010, Mortimer et al. 2017, Broor et al. 2007, Farooqui et al. 2015].[5-9] Multiple events per child is expected for some children over the follow-up period. We assume a 10% attribution over the 1-year follow-up period, resulting in an effective sample size of $n = 1440$ per treatment arm for pneumonia. There is no sample size calculation for the exposure-response analysis, but the power should be

like the ITT. Power of E-R models is expected to be slightly greater than the ITT if the exposure measures are complete and valid enough.

2.4. Trial Framework

HAPIN is a superiority trial. While the primary intention-to-treat analysis (described in a separate document) is a test of statistical significance to evaluate whether there is a difference in the incidence of severe infant pneumonia between the intervention and control arms, the **exposure-response analysis** is a test of the association between personal exposures to household air pollutants and severe infant pneumonia.

2.5. Statistical Interim Analyses and Stopping Guidance

No interim analysis will be conducted.

2.6. Timing of Analysis

All analysis will be conducted once data collection and cleaning are complete, and the SAP has been approved and registered. Exposure-response analysis between severe infant pneumonia and exposure will be conducted as a standalone analysis.

2.7. Timing of Outcome and Covariate Assessments

Participating households were followed from enrollment until the index child reaches (or would have reached, assuming a live birth and continued vitality) their first birthday. For the purposes of this analysis plan, severe pneumonia follow-up is through the first year of life.

HAP exposure was assessed prenatally and postnatally. The measurement periods were 24 hours. Prenatal exposures include one measurement at baseline (<20 weeks, denoted as the BL measurement in our trial) and two post-randomization (at 24-28 weeks' gestation denoted as P1; and at 32-36 weeks' gestation denoted as P2). We measured child postnatal exposures indirectly as a reconstructed average of maternal personal exposures and two area concentrations three times at ages 3 (denoted as the B1 measurement), 6 (denoted as B2) and at 12 months (denoted as B4). A validation subset of 10% of participants had double the numbers of measures, with extra measures at time points spaced between those of the main study.

Time-invariant covariates that are potential confounders and effect modifiers were collected at baseline and birth. If effects of time-varying exposures are considered, covariates that could be time-varying confounders were collected at quarterly household visits during infancy (e.g. vaccination, breast feeding).

3. STATISTICAL PRINCIPLES

3.1. Confidence Intervals

Confidence intervals will be presented at 95% confidence for the primary exposure-response analyses and for any subgroup analyses. For the purposes of comparability of findings with those from the intention-to-treat analyses, we will also present 98.75% confidence intervals, i.e., a familywise adjustment to account for four primary outcomes, in the Online Supplement.

3.2. Analysis Populations

The **exposure-response** analyses will include all follow-up time after birth and before death in study children whose households continue participating in HAPIN and who are at risk of a new severe pneumonia episode. No person-time is contributed to the cases of death of the mother prior to birth, miscarriage, stillbirth, and withdrawal from study prior to birth. Otherwise, person-time at risk was counted from the date of birth until the earliest date among the following: age 1 year, death of child, withdrawal from the study, or migration from the study area.

Missingness may affect exposure-response relationships. If the proportion missing exposure or confounder data is <10%, our study population for primary exposure-response analyses will be restricted to those

participants with complete data. If the proportion of exposure or confounder data missing is ≥10%, we will consider using multiple imputation methods to reduce missing data bias and loss of power.

4. STUDY POPULATION

4.1. Eligibility

Pregnant women were eligible to participate in the study if they fulfilled the following inclusion and exclusion criteria at screening:

Inclusion criteria:

- Offspring of a confirmed pregnancy (hCG positive blood or urine test)
- Offspring of a pregnant woman aged 18 to less than 35 years (via self-report)
- Offspring pf a pregnant woman 9 to less than 20 weeks' gestation confirmed by ultrasound
- Offspring of a singleton pregnancy (one fetus confirmed by ultrasound)
- Viable fetus with normal fetal heart rate (120-180 beats per minute) at time of ultrasound
- Lives in study area
- Uses biomass stove for cooking predominantly
- Agrees to participate with informed consent
- Continued pregnancy at the time of randomization confirmed by self-report

Exclusion criteria:

- Currently smokes cigarettes or other tobacco products
- Plans to move permanently outside study area in the next 12 months
- Uses LPG stove predominantly, or is likely to use LPG predominantly, in the near future

If two eligible pregnant women lived in the same household and were interested in participating, the one with the earliest gestational age was chosen to participate.

The severe pneumonia exposure-response analysis includes all live-born children in households that continued participating in the HAPIN trial through trial exit.

4.2. Recruitment

The following information was included in the CONSORT flow diagram. All counts will be reported as total and by IRC.

- Reasons for exclusion when assessed for eligibility
  - Not pregnant/no viable fetus
  - Mother outside of age range
  - Does not/will not primarily cook with biomass
  - Planned to move/moved away
  - Unwilling to participate
  - Gestational age out of range
  - Not a singleton
  - Smoker
  - Not in study area
  - Withdrawn by study team/not pursued further
- Participants determined to be ineligible after randomization
- Reasons for exits after randomization
  - Voluntary withdrawal prior to birth and after birth
  - Withdrawn by study team prior to birth and after birth
  - Moved away prior to birth and after birth
  - Pregnancy loss (termination/miscarriage/stillbirth)

4.3. Withdrawal/follow-up

The study recorded reasons for exit classified into several categories:

- Not eligible
- Participant voluntary withdrawal
- Withdrawn by study team
- Moved away from study area (periods of intermittent moves outside study area will be excluded)
- Deceased
- Lost to follow up
- Mother abortion/miscarriage/stillbirth
- Child death
- Other

For exits due to eligibility, voluntary withdrawal and withdrawal by study team, several pre-specified reasons were used, as well as the option to fill in other reasons. The last completed visit was also recorded. Reasons for withdrawal and loss to follow-up were ascertained as soon as possible.

If reason for study exit is voluntary withdrawal

- Procedures too intrusive
- Procedures too time-consuming
- Do not see value in the study
- Family does not want me to participate
- Do not want to be in assigned group
- Other

If reason for study exit is withdrawn by study team:

- Repeated resistance to study procedures
- Danger to study personnel
- Other

5. KEY VARIABLE DEFINITIONS AND DATA COLLECTION

In this section, we describe the data used for the exposure-response analysis of HAP exposure and severe infant pneumonia, as defined by HAPIN and WHO.

5.1 Overview of Key Analysis Variables for Manuscript

Personal HAP exposures were measured directly in pregnant women/mothers both prenatally and postnatally, including fine particulate matter ($PM_{2.5}$), black carbon (BC), and carbon monoxide (CO). Postnatally, we measured area concentrations in the kitchen and infant's bedroom. Personal exposures to $PM_{2.5}$ and CO in infants will be assessed indirectly. Specifically, we will use logging transponders on the mother and in the kitchen and bedroom, and the infant will wear emitters. We have developed and validated an interpolation algorithm to estimate indirect exposures in infants;[10] however, we will consider different approaches to estimate personal exposure in infants if missingness of emitter/transponder logging data is ≥10%. We summarize HAP exposures in the trial in Table 1. Per protocol, exposure variables used in exposure-response models were calculated as weighted averages of subsets of these assessments (Section 6.1). However, we also explore performance of alternative predictions of longer-term exposure using our exposure validation sub-study.

Table 1. HAP exposures measured in the trial

Measurement time	Notation	Personal exposure in pregnant woman	Personal exposure in mother	Area concentration in kitchen	Area concentration in bedroom	Indirect measurement in children
9 – 19 weeks' gestation	BL	$PM_{2.5}$, BC, CO				
24 – 28 weeks' gestation	P1	$PM_{2.5}$, BC, CO				

32 – 36 weeks' gestation	P2	$PM_{2.5}$, BC, CO				
Child is 3 months of age	B1		$PM_{2.5}$, BC, CO	$PM_{2.5}$, BC, CO	$PM_{2.5}$, BC, CO	$PM_{2.5}$, CO
Child Is 6 months of age	B2		$PM_{2.5}$, BC, CO	$PM_{2.5}$, BC, CO	$PM_{2.5}$, BC, CO	$PM_{2.5}$, CO
Child is 12 months of age	B4		$PM_{2.5}$, BC, CO	$PM_{2.5}$, BC, CO	$PM_{2.5}$, BC, CO	$PM_{2.5}$, CO

The main outcome for this analysis is the same as in the intention-to-treat analysis (infant severe pneumonia). We will also analyze a set of predefined alternative definitions of pneumonia that cover a range of sensitivity, specificity, and disease severity. There were 175 HAPIN severe pneumonia during the follow-up period, among 2984 infants (6%), and there were 267 severe WHO pneumonias (9%)[2]. These respiratory endpoints are summarized in Table 2:

Table 2. Pneumonia and other respiratory outcomes

Parameter	Definition
(HAPIN) severe infant pneumonia	• Presence of cough and/or difficult breathing and at least one general danger sign and primary endpoint pneumonia on a lung ultrasound or chest x-ray • Presence of cough and/or difficult breathing and hypoxemia (measured via pulse oximetry, SpO_2) • Children who died but their death is attributed to pneumonia by verbal autopsy General danger signs for all participants include any of the 5 signs of: • Unable to drink or breastfeed • Vomiting everything Convulsions • Lethargy or unconscious • Stridor at rest • Severe acute malnutrition (weight-for-length z-score < -3 or weight-for-age < -3 if length is missing) There are 4 additional general danger signs for participants <2 months of age (neonatal danger signs): • Unable to feed • Grunting • Not moving at all or moves with stimulation only • Severe chest indrawing
IMCI WHO severe infant pneumonia	Cough and/or difficult breathing and • any general danger sign (unable to drink or breastfeed, vomiting everything, convulsions, lethargic or unconscious) or • stridor at rest or • severe acute malnutrition or • HIV infection or exposure (if chest indrawing also) or • Hypoxemia (SpO_2 <90%)

5.2. Identification of Confounders

To assess conditions for identifiability of causal effects of HAP exposure on pneumonia, we gathered expert views on the causal relations among variables related to exposure, pneumonia, or the HAPIN study design. Confounder selection was based on conceptual directed acyclic graphs as shown in Appendix figures and code.

A key feature of this exposure-response analysis for causal inference is that it is conducted within the population of a randomized intervention trial that created a strong exposure contrast by arm. Nevertheless, HAP exposure is determined by many other factors and it is likely that some of these factors are common

causes of severe infant pneumonia or share a common cause with severe infant pneumonia. Since exposures are qualitatively and quantitatively different during gestational and postnatal periods, we have included a time-varying exposure that can take one value during pregnancy (HAP1_T) and another after birth (HAP2_T). The DAG can be analyzed for adjustement sets for estimating effects of prenatal exposure, postnatal exposure, or both of these exposures.

A general assumption in the DAG in appendix Figure 1 is that the HAPIN trial intervention can have direct effects on respiratory disease or effects that are not pathways going through HAP exposure. In particular, the the provision of free fuel can have a large impact on the household economy and influence behaviors. For example, we have assumed the intervention has direct effects on breastfeeding, nutrition, healthcare seeking, maternal vaccination, and pneumonia. However, there was not a consensus in the research group about the pathway between intervention and severe pneumonia and the downside to conditioning on intervention could be reducing study power by reducing residual exposure variability (i.e.. greater variance of exposure-response estimates). The primary analysis will not condition on intervention or types of fuels and stoves.

We also considered measurement error as a source of bias. The DAG includes nodes for measured (e.g. exposures HAP1_M and HAP2_M, and outcomes Pneumonia_M) and true values of HAP exposure and severe pneumonia. These allow representation of structures that can be important for understanding information bias. For example, the HAPIN intervention can influence the probability that a caregiver with a sick child seeks care at a surveillance site or that the healthcare seeking event results in screening by HAPIN staff. Moreover, the spatiotemporal and dispersion properties of HAP would suggest that error in indirect exposure assessment is likely to be dependent on intervention arm and other determinants of exposure.

Based on our DAG, we have identified the following variables, for which we have data, as the minimal adjustment set to identify the effect of prenatal HAP exposure on severe infant pneumonia: crowding (number sleeping in house), intervention IRC, nutrition history of breastfeeding,, SES, sex, birthweight, and a data for small for gestational age (z-score). As noted, intervention status will not be included for either pre or post-natal exposure as it will be superceded by measured exposure level.

Birthweight would need to be included to reduce bias for estimating effects of postnatal exposure, but conditioning on birth weight could bias estimates of prenatal exposure because it is suspected of being on the causal pathway, so caution must be exercised here.

We will conduct sensitivity analyses to assess variability between estimates with alternative sets of appropriate covariates. Confounders selected from adjustment sets are found in the HAPIN codebook.

5.3. Baseline Participant Characteristics

Baseline characteristics will be summarized, and completeness assessed as previously described.[2]

5.4. Data Collection

The exposure assessment methods have been described previously in detail.[11] As noted above, we measured direct personal exposures in pregnant women once at baseline and twice during the trial period and measured indirect personal exposures in infants three times, at 3, 6 and 12 months. Results of this exposure assessment and the effects of the intervention have been reported separately.[12,13] HAP exposure will be summarized in terms of timing and concentration, separately by each IRC according to subgroups defined in Table 5.

The person-time intervals per child can be considered as one continuous period, multiple weekly periods, or intermittent with interruptions for migration and recent cases. We will include the same respiratory outcomes reported in an intention-to-treat analysis and described in more detail in the published SAP and supplemental material for the severe pneumonia intention-to-treat analysis[1]. Here we only briefly describe each respiratory endpoint, including data collection and processing.

The pneumonia case definitions are in Table 2. The primary endpoint of the trial included three criteria: 1) respiratory signs (cough or difficulty breathing), 2) evidence of severity (hypoxemia or danger signs), and 3) an objective diagnosis (hypoxemia, imaging, verbal autopsy, or advanced respiratory care).

Lung ultrasound was the primary imaging modality obtained and where logistical or regulatory barriers existed we obtained a chest x-ray. We defined tachypnea based on current World Health Organization thresholds ($\geq$60 breaths/minute for age 0-2 months and $\geq$50 breaths/minute for age 3-11 months). We will determine hypoxemia based on the physiologic threshold of $\leq$92% for altitudes <2,500 meters above sea level and $\leq$86% for altitudes $\geq$2,500 meters above sea level. Children are also considered hypoxemic, regardless of oxygen saturation, and assumed to have cough and/or difficult breathing if they are receiving advanced respiratory support at the time of evaluation, which includes any of the following: intubation and mechanical ventilation, non-invasive ventilation with continuous positive airway pressure support (CPAP), non-invasive ventilation with bi-level positive airway pressure support (BIPAP), or high-flow nasal cannula oxygen.

Children who died but were determined to have had pneumonia or symptoms and/or signs of a respiratory illness consistent with pneumonia by verbal autopsy were also considered a case of severe pneumonia. Adjudication panels determined the final interpretation of lung ultrasounds, chest radiographs, and verbal autopsies.

Screenings that are not complete enough for case ascertainment will be considered missing outcome data. If a substantial proportion of screenings (≥10%) are incomplete, we will consider imputation or weighting to reduce potential for bias.

The variables described in Table 3 were used to classify cases according to the two case definitions. Pneumonia endpoint variables were constructed (calculated) from individual data collection items across multiple data sources. Since these data processes are highly intricate, a panel of three data coders replicated the creation of the analysis dataset. This work was performed by the HAPIN Pneumonia Working Group members in collaboration with DMC. The final code for calculation of pneumonia outcome variables was provided by the HAPIN Pneumonia Working Group (PWG) and is available through the DMC.

Table 3. Severe pneumonia case characteristics to be reported

Variables	**Type**	**Definition/Assessment Methods**
Clinical signs		
Temperature >38^{o}C, n (%)	Binary	
Average Heart rate	Continuous	Average of measurements when multiple measurements available Beats per minute
Average Respiratory rate	Continuous	Average of measurements when multiple measurements available Breaths per minute
Average SpO_2	Continuous	Average of measurements when multiple measurements available
Hypoxemia, n (%)	Binary	Categorized as SpO_2 <93% for Guatemala, India, and/or Rwanda. <87% for Peru. Yes/No
Wheeze +/- crackles, n (%)	Binary	Yes/No
At least one respiratory danger sign, n (%)	Binary	Any of the following: chest indrawing, severe chest indrawing, head nodding, persistent nasal flaring, grunting, stridor when calm, audible wheeze, tracheal tugging, intercostal recessions Yes/No
Chest indrawing, n (%)	Binary	Yes/No
Head nodding, n (%)	Binary	Yes/No
Persistent nasal flaring, n (%)	Binary	Yes/No

Grunting, n (%)	Binary	Yes/No
Stridor when calm, n (%)	Binary	Yes/No
Audible wheeze, n (%)	Binary	Yes/No
Tracheal tugging, n (%)	Binary	Yes/No
Intercostal retractions, n (%)	Binary	Yes/No
At least one general danger sign, n (%)	Binary	Any of the following: unable to drink or breastfeed, vomiting everything, convulsions, lethargy or unconscious, unable to feed, not moving at all or moves with stimulation only Yes/No
Unable to drink or breastfeed, n (%)	Binary	Yes/No
Vomiting everything, n (%)	Binary	Yes/No
Convulsions, n (%)	Binary	Yes/No
Lethargy or unconscious, n (%)	Binary	Yes/No
At least one neonatal danger sign, n (%)	Binary	Any of the following: unable to feed well, not moving at all or moves with stimulation only, grunting, severe chest wall indrawing Yes/No
Unable to feed well, n (%)	Binary	Applies to <2 months only. Yes/No
Not moving at all or moves only, n (%)	Binary	Applies to <2 months only. Yes/No
Grunting, n (%)	Binary	Applies to <2 months only. Yes/No
Severe chest indrawing, n (%)	Binary	Applies to <2 months only. Yes/No
Lung Imaging		
Lung ultrasound, n (%)	Binary	Primary endpoint pneumonia Yes/No
Chest radiography, n (%)	Binary	Primary endpoint pneumonia Yes/No
Clinical care		
Hospitalized, n (%)	Binary	Yes/No
Oxygen treatment, n (%)	Binary	Yes/No
Advanced respiratory supportive care, n (%)	Binary	Any of the following: High flow, NIV (CPAP/BiPAP), mechanical ventilation Yes/No
Outcome		
Mortality, n (%)	Binary	Death <30 days since diagnosis and/or Verbal Autopsy positive Yes/No

Children were permitted to meet the above case definition for severe pneumonia multiple times if the repeat event occurred either >14 days after the hospital discharge date or, if the hospital discharge data was not available or the child was not hospitalized, >30 days after the date the child met criteria for severe pneumonia. Implausible values and outliers identified from the variables that comprise the case definition will be considered missing. In addition to reporting the incidence rates of each endpoint, we will present the completeness of case ascertainment (e.g. proportion of reported healthcare visits with a screening, proportion of eligible cases with ultrasound assessment).

6. DATA ANALYSIS

6.1 Model specification

Our main aim is to estimate the exposure-response relationship between prenatal and postnatal HAP exposures and severe pneumonia incidence during the first year of life. The unit of observation in this analysis is infant-quarter. Occurrence of HAPIN severe pneumonia is the primary endpoint.

We will use log-binomial regression to estimate a risk ratio for the association between $PM_{2.5}$ exposure and severe infant pneumonia on any given quater. Exposure-response relationships will be modelled as follows:

$$\log E(y_{ij}) = \beta_0 + f(X_{ij}) + \sum_{k=1}^{m} \gamma_k \times Z_{ij}$$

where y_{ij} is the presence or absence of a severe pneumonia episode for the i^{th}infant and in the j^{th} quarter. The function $f(X_{ij})$ represents a range of functions to capture the relationship between HAPIN severe infant pneumonia rates and HAP exposures. We will consider different types of functions for $f(X_{ij})$, including linear and polynomials including regression splines, and functional smooths using generalized additive models (GAMs). We will conduct visual assessment of observed versus predicted values to assess goodness-of-fit and use metrics like the Akaike Information Criteria or Root Mean Squared Error to compare models. The Z_{ij} represent the m potential confounders (Section 5.2) some of which are fixed (Z_i) and others time-dependent (Z_{ij}). We will either use generalized estimating equations (GEE) with a compound symmetry matrix and a robust variance estimator to account for serial correlation by child or generalized linear mixed models with random effects to account for heterogeneity among children.

We will consider different functional forms for $f(X_{ij})$ to model pollutant exposures ($PM_{2.5}$ and CO) separately for prenatal and post-natal exposures, i.e., $f_{prenatal}(X_{ij}) + \ f_{post-natal}(X_{ij})$. For prenatal exposures, we will use the mean of the mother's three pre-natal measurements (at baseline, at 24-28 weeks, and 32-36 weeks). Not all mothers had all three measurements, although most ded. We will use all available measurements to calculate a mean. It is possible that the effect of pre-natal exposures wanes over time. We will compare goodness-of-fit between models that consider the average (power value of 0) and those at different power values (a decay at a power value of 1, 2 or 3).

As for post-natal values of X_{ij}, we will consider two approaches for measuring averages. The first will use means of values for **X** for the j^{th} quarter using all values of personal exposures to $PM_{2.5}$ that fall in the first quarter defined in Table 4, or the average of any prior exposures thereafter. We will consider unweighted means of the $PM_{2.5}$ (or CO) exposures, and weighted mean of the values for **X** considering the distances in time between the ages when personal exposures to $PM_{2.5}$ were measured and the age at the start of the j^{th} quarter. The third approach uses a regression model to estimate infant's mean personal exposures. As children's data are missing for about 50% of post-natal measurements, we will use the mother's personal exposure measured at the same timepoint when the infants's measurement is missing. The unadjusted correlation of maternal and infant measurements is about 0.89. Prior work has shown that when using a regression model to predict time-specific infant exposures using mother's post-natal exposure and adjusted for season, weekend, and IRC, has correlation of 0.95 when compared with the infant's observed personal exposure. Using the assumption that maternal and child personal exposures measurements are similar, we anticipate that we will only have ~10% missing child exposure data.We will select the best approach for modeling exposure from the approach that leads to the least amount error in estimation of exposure (based on a root mean squared error in validation study).

Table 4. Values of $PM_{2.5}$ to use for exposure-response modeling.		
t_{ij}	Values for X_{ij} (prenatal)*	Values for X_{ij} (post-natal)
$t_{ij} \leq 3$	$\boldsymbol{X} = \{PM_{2.5}(t = BL), PM_{2.5}(t = P1), PM_{2.5}(t = P2)\}$	$\boldsymbol{X} = \{PM_{2.5}(t = B1)\}$
$3 < t_{ij} \leq 6$		$\boldsymbol{X} = \{PM_{2.5}(t = B1), PM_{2.5}(t = B2)\}$
$6 < t_{ij} \leq 12$		$\boldsymbol{X} = \{PM_{2.5}(t = B1), PM_{2.5}(t = B2), PM_{2.5}(t = B4)\}$

The error term for our regression model is a binomial distribution, and the canonical link is the logarithm function. Exponentiation of the difference between $f(X_{ij})$ at different exposure values provides the risk ratio associated with that change in exposure. If we identify estimation problems when using the logarithm function, we will revert to use a logit function where exponentiation at different exposure values provides an odds ratio associated with that change in exposure. Since there are few cases (175 severe pneumonias) compared to the overall number of infant-quarters (~12,000), the odds ratio of a logistic regression should approximate the relative risk relatively well. We also do not expect zero-inflation to be a problem when using log or logit canonical links.The same log-linear (or logit) regression models will be used when using CO as the exposure, and in subgroup analyses and secondary outcomes, unless otherwise indicated.

A priori time-independent confounders will be crowding (number sleeping in house), IRC, SES, infant's sex and birthweight (or a z-score for birthweight for gestational age of child instead of birthweight), An SES score will be measured using a SES wealth index, but mayalso incoporate mother's education and nutrition data, and has been developed by our group and used in previous publications. A priori time-dependent confounders for the jth quarter include age and season (winter or no), vaccination status, breastfeeding, and an indicator for pre vs post COVID-19 pandemic (i.e., 1 after March 15, 2020, and 0 otherwise).

As a secondary analysis we will conduct a time-to-event analysis using the Anderson-Gill method counting process method, which allows for repeated events for individuals.[14] The start and stop times of the counting process will be defined according to the dates of exposure assessment for each child (i.e., birth to B1, B2 to B2, and B2 to B4), and the status will be classified as either pneumonia (where the outcome is 1) or no pneumonia (where the outcome is 0). Given that the underlying model based on Cox proportional hazards, we will test for proportionality of hazards assumption. To account for serial correlation of repeated time-to-event states for each child, we will use a sandwich estimator or random effects as defined by Amorim and Cai [14]. The estimand for this model is a hazard ratio.

We will conduct subgroup analyses to identify potential effect modifiers: International Research Center (i.e., site), and known risk factors for severe infant pneumonia in LMICs (Table 5):

Table 5. Subgroup analysis (effect modifiers)		
Variables	**Type**	**Definition/Assessment Methods**
International Research Center	Categorical	Guatemala, India, Peru, Rwanda
Child age	Continuous	
Up-to-date vaccination status	Fixed and time-varying categorical	Receipt of 3 respiratory vaccines by one year of life, including 3 doses of pneumococcal conjugate vaccine, 3 doses of *Haemophilus influenzae* vaccine, one dose of measles vaccine (yes/no/missing)
COVID-19 pandemic	Categorical	Before and after March 15, 2020

6.2. Analysis Replication Plan

Selected components of the exposure-response analyses will be replicated by a second analyst. Secondary analyses of any outcome related to sensitivity analyses (i.e., alternative health model specifications, alternative covariate specification) will not be replicated.

The replication team will receive the following from the Data Management Core (DMC).

1. A cleaned analytic dataset. The dataset will also include baseline variables, covariates for subgroup analysis and covariates to include in the exposure-response analyses.
2. A table summarizing baseline characteristics (overall and by IRC).
3. The set of outcomes (primary and secondary).
4. For the exposure-response analysis only, the list of pre-specified covariates to be included in the regression models and forms of the exposure-response function.

Specific replication tasks include:

1. Replicate summary statistics (e.g., mean, standard deviation, percentages, proportion missing) in the baseline characteristic table.
2. Replicate exposure-response analyses for primary and secondary outcomes according to modelling strategies specified in Section 6.1.

REFERENCES

2. McCollum ED, McCracken JP, Kirby MA, et al. Liquefied Petroleum Gas or Biomass Cooking and Severe Infant Pneumonia. *N Engl J Med.* 2024;390(1):32-43.
3. Proportion of population with primary reliance on clean fuels and technologies for cooking (%) https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-phe-primary-reliance-on-clean-fuels-and-technologies-proportion. Accessed Accessed April 26, 2022.
4. Clasen T, Checkley W, Peel JL, et al. Design and Rationale of the HAPIN Study: A Multicountry Randomized Controlled Trial to Assess the Effect of Liquefied Petroleum Gas Stove and Continuous Fuel Distribution. *Environmental health perspectives.* 2020;128(4):47008.
5. Mackenzie GA, Bottomley C, van Hoek AJ, et al. Efficacy of different pneumococcal conjugate vaccine schedules against pneumonia, hospitalisation, and mortality: re-analysis of a randomised trial in the Gambia. *Vaccine.* 2014;32(21):2493-2500.
6. Gupta M, Kumar R, Deb AK, et al. Multi-center surveillance for pneumonia & meningitis among children (<2 yr) for Hib vaccine probe trial preparation in India. *Indian J Med Res.* 2010;131:649-658.
7. Mortimer K, Ndamala CB, Naunje AW, et al. 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. *Lancet.* 2017;389(10065):167-175.
8. Broor S, Parveen S, Bharaj P, et al. A prospective three-year cohort study of the epidemiology and virology of acute respiratory infections of children in rural India. *PLoS One.* 2007;2(6):e491.
9. Farooqui H, Jit M, Heymann DL, Zodpey S. Burden of Severe Pneumonia, Pneumococcal Pneumonia and Pneumonia Deaths in Indian States: Modelling Based Estimates. *PLoS One.* 2015;10(6):e0129191.
10. Liao J, McCracken JP, Piedrahita R, et al. The use of bluetooth low energy Beacon systems to estimate indirect personal exposure to household air pollution. *J Expo Sci Environ Epidemiol.* 2020;30(6):990-1000.
11. Johnson MA, Steenland K, Piedrahita R, et al. Air Pollutant Exposure and Stove Use Assessment Methods for the Household Air Pollution Intervention Network (HAPIN) Trial. *Environmental health perspectives.* 2020;128(4):47009.
12. Johnson M, Pillarisetti A, Piedrahita R, et al. Exposure Contrasts of Pregnant Women during the Household Air Pollution Intervention Network Randomized Controlled Trial. *Environ Health Perspect.* 2022;130(9):97005.
13. Pillarisetti A, Ye W, Balakrishnan K, et al. Post-birth exposure contrasts for children during the Household Air Pollution Intervention Network randomized controlled trial. *medRxiv.* 2023.
14. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. Int J Epidemiol. 2015 Feb;44(1):324-33. doi: 10.1093/ije/dyu222.

APPENDIX Figures

Figure A1. Directed acyclic graph (DAG) for the effect of prenatal and postnatal exposures to household air pollution (HAP1_T and HAP2_T, respectively) on infant severe pneumonia by one year of age (Pneumonia_T).