**STATISTICAL ANALYSIS PLAN [ER-4]**

**Exposure-Response analysis on stunting, wasting, and underweight in infants 12 months of age in a multi-country LPG stove intervention trial (HAPIN)**

Version 1.2
February 14, 2024

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

Trial Registration: NCT02944682
Protocol Version: HAPIN - Main Study Protocol v15.0 17 September 2022

Modification History:

| Version | Date |
|---|---|
| 1.0 | 11/16/2023 |
| 1.1 | 01/26/2024 |
| 1.2 | 02/14/2024 |
| 1.3 | |
| 1.4 | |

## 1. INTRODUCTION

This document contains the statistical analysis plan (SAP) for the assessment of anthropometric growth at 12 months of age among infants enrolled in the HAPIN trial, including stunting, one of four primary outcomes of the HAPIN trial. This SAP will be posted with the trial registration.

Infant *stunting* at 1 year of age, defined as a length for age Z-score (LAZ) that is 2 standard deviations below the median of the growth standard, is one of the four primary outcomes of the trial. This SAP describes the exposure-response analysis for this outcome. Other secondary outcomes will be assessed at 12 months only, including: 1) length-for-age, 2) severe stunting (LAZ < -3 SDs), 3) *underweight*, assessed by weight for age Z-score (WAZ), a measure of chronic malnutrition resulting in linear growth retardation; 4) *wasting*, assessed by weight for length Z-score (WLZ), a measure of severe famine. Except for WLZ, these measures are age-adjusted, and all are sex-specific. Binary outcomes will be evaluated using the moderate cut-off, defined as more than 2 standard deviations (or 3 SDs for severe stunting) below a median international reference population, the 2006 WHO Multicenter Growth Reference Standards (MRGS).[1]

### 1.1. Background and Rationale

Nearly 30% of the world's population is exposed to HAP and rely on solid fuels for cooking and heating, most of whom reside in low-resourced countries. An estimated 2.4 million premature deaths are caused by household exposure to fine particulate matter (PM2.5), with women and children most heavily exposed during cooking.[2] HAP causes about 4% of the total global disease burden.[3] Previous interventions have provided "improved" biomass-based cookstoves that either vent smoke outside or combust fuel more efficiently, but many have failed to reduce HAP exposures to levels that demonstrate beneficial health outcomes. There have been no multi-country randomized field trials conducted across rural settings that have used liquefied petroleum gas (LPG) cookstoves at scale. At present, LPG is scalable in rural settings that don't have reliable sources of electricity and is thus considered the cleanest immediately available and scalable cooking intervention.

### 1.2. HAPIN Study Overview

The aim of the HAPIN study was to conduct a randomized controlled trial of an LPG stove and fuel distribution intervention in 3200 rural households in four LMICs (India, Guatemala, Peru, and Rwanda) to deliver evidence regarding potential health benefits across the lifespan. Each intervention site recruited 800 pregnant women (aged 18 to <35 years, 9 to <20 weeks gestation), half of whom were randomly assigned to receive LPG

stoves, an 18-month supply of free LPG, and behavioral reinforcement messaging by the study team to encourage LPG use. Controls were asked to continue cooking with solid biomass fuels and were compensated for their participation during and/or at the end of the study.[4] The mother and her infant were followed until the infant was 1 year old. If a non-pregnant older adult woman (aged 40 to <80 years) resided in the household, there were invited to enroll in the study and were followed during the 18-month follow-up period to assess cardiopulmonary, metabolic, and cancer outcomes. We assessed cookstove use, conducted repeated personal exposure measurements of 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 infant outcomes are birth weight, severe pneumonia, and stunting at 1 year of age. The fourth outcome is blood pressure changes in older adult women. Protocols for the trial have been described elsewhere.[5–7]

### 1.3. Study Objectives

The HAPIN study addressed the following specific aims: (1) using an intent-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 health outcomes; and (3) determine relationships between LPG intervention and both targeted and exploratory biomarkers of exposure/health effects.

## 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 an additional approximately 120 older adult women from the same households who met inclusion/exclusion criteria (Section 4.1). Key characteristics of each study site are given in Table 2 of the HAPIN design publication.[5]

Recruitment and enrollment occurred over approximately 15 months at ~53 pregnant women/8 older adult women per month per IRC. All participants were followed longitudinally for ~18 months (until the infant reached 12 months), or until the participant exited the study (e.g., voluntary withdrawal, death).

### 2.2. Randomization

To ensure balance between arms, households were randomly allocated to intervention or control arms after they consented to participate. To maintain balance of treatment assignments within each IRC, 10 randomization strata were implemented as follows.

- The India IRC randomization list was stratified by the two study sites
- The Peru IRC randomization list was stratified by the six study sites
- Guatemala and Rwanda had one site each

Separate randomization lists were generated for each field team conducting randomization at each IRC. Two randomization lists were produced: 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 trial protocol.

### 2.3. Sample Size Considerations

The sample size and power for the trial were based on the intention to treat analyses for the primary outcomes. There is no sample size calculation for the exposure-response analyses in the HAPIN trial.

### 2.4. Trial Framework
HAPIN was a superiority trial. The primary intention-to-treat analysis is a test of statistical significance to evaluate whether the outcome data are consistent with the assumption of there being no difference between the intervention and control arms (see SAP for intention to treat for stunting). Exposure-response analysis between anthropometric outcomes at the one-year follow-up period is described here as a separate analysis per the original aims of the study (as described here).

### 2.5. Statistical Interim Analyses and Stopping Guidance
No interim analysis will be conducted.

### 2.6. Timing of Analysis
All analyses will be conducted once data collection is complete, and the SAP has been approved and registered.

### 2.7. Timing of Outcome and Covariate Assessments
Each participating household was followed from enrollment until the index infant reached (or would have reached, assuming a live birth and continued vitality) their first birthday. For the purposes of this analysis anthropometry will be assessed at 12 months. In addition to baseline assessments conducted at recruitment, personal exposures to household air pollution were conducted during pregnancy at 24-28 and 32-36 weeks of gestation, and three times during the first year of life: at <3 months, 6 months, and 12 months of age, for a total of 6 measurements, with 5 measured after baseline.

## 3. STATISTICAL PRINCIPLES

### 3.1. Confidence Intervals
Analyses of air pollution exposure-response associations will report 95% confidence intervals for the purposes of estimation of health effects. Analyses for subgroups and other secondary outcomes will also report 95% confidence intervals.

### 3.2. Adherence and Protocol Deviations
All homes in the intervention arm used Stove Use Monitoring Systems (SUMS) on their traditional stoves, and a subset of ~ 20% also had their gas stoves monitored. In a subset of up to 20% of control households, all stoves used more than once per week were monitored. Compliance was checked every two weeks when SUMS data were downloaded.

Behavioral reinforcements (messages and materials) were delivered when intervention households showed any use of their traditional stoves. We flagged households that were using their traditional stove one or more times over the previous two-week monitoring period. After flagging these households, we probed members of the participating household to ascertain reasons for non-compliance and intervene as necessary. At all behavioral reinforcement visits, a brief questionnaire was conducted to identify barriers to LPG stove use in the household and document the messages and materials used to address those barriers. Once specific reasons/factors were determined, personalized behavior change reinforcements were delivered.

### 3.3. Analysis Populations
For each anthropometric outcome, the analysis will include all children who have a valid height, and/or weight (*complete-case analysis*). We define loss to follow-up as any reason that contributes to a missing outcome value, including death or withdrawal of the infant prior to the first year of life. We will also use subsets of the study to examine effect modification.

## 4. TRIAL POPULATION

**4.1. Eligibility**
Children 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
- Offspring of pregnant women aged 18 to less than 35 years (via self-report)
- Offspring of pregnant women 9 to less than 20 weeks gestation confirmed by ultrasound
- Offspring is a singleton pregnancy (one fetus)
- Was a viable fetus with normal fetal heart rate (120-180 beats per minute) at time of ultrasound
- Mother continued pregnancy at the time of randomization confirmed by self-report
- Mother agreed to participate with her informed consent
- Household uses biomass stove predominantly
- Lives in study area

*Exclusion criteria*:
- Mother currently smoked cigarettes or other tobacco products
- Mother planned to move permanently outside study area in the next 12 months
- Mother used LPG stove predominantly, or was likely to use LPG predominantly, in the near future

We enrolled only 1 infant per household.

**4.2. Recruitment**
The following will be included in the CONSORT flow diagram. All counts will be reported as total and by IRC.

- Reasons for exclusion when women were assessed for study eligibility:
  - Not pregnant/no viable fetus
  - Woman is outside of age range
  - Woman does not/will not primarily cook with biomass
  - Woman plans to move/moved away
  - Woman 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
- Maternal follow-up period
  - Number of children with at least one valid measure of PM2.5, BC, or CO during post-baseline visit (24-38 or 32-36 week prenatal visits)
  - Reasons for exits after randomization but before live birth
    - Voluntary withdrawal
    - Withdrawn by study team
    - Moved away
    - Pregnancy loss (termination/miscarriage/stillbirth)
- Infant follow-up period
  - Number of children with at least one valid measure of PM2.5, BC, and CO during infancy (3, 6, or 12-month visits)
  - Reasons for exits after live birth
    - Voluntary withdrawal
    - Withdrawn by study team
    - Moved away
    - Infant death

- Reasons for exclusion due to missing/implausible data
    - Erroneous anthropometric measurements (e.g., length at 12 months is < than length at 9 months)
    - Missing anthropometric measurements
- Number of children with at least one valid pre-natal and one valid post-natal exposure measure and at least one anthropometric measurement at 12 months of age.

**4.3. Withdrawal/follow-up**

The study will record reasons for exit classified into several categories:

- Not eligible
- Participant voluntary withdrawal
- Withdrawn by study team
- Moved away from study area
- Deceased (mother and/or infant)
- Mother abortion/miscarriage/stillbirth/infant death
- Lost to follow up
- Other

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

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

Reason for study exit is withdrawn by study team:

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

**4.4. Participant Characteristics**

For the exposure-responses analysis, participant characteristics will be summarized separately by each IRC as defined by Table 1. Medians and interquartile ranges will be calculated for continuous variables and counts/percentages will be calculated for categorical variables. Missing data will be reported as a separate category.

**Table 1. Participant characteristics to be reported**

| Variables | Type | Definition/Assessment Methods |
|---|---|---|
| **Maternal Factors** | | |
| Mother's age (years) | Continuous | Calculated as the date at baseline minus the date of birth. Date at baseline is assigned by the date of visit if not missing. |
| Mother height | Continuous | Average height calculated from two closest heights measurements |
| Mother's body mass index (BMI) | Continuous | BMI calculated as the average weight (kg) divided by the average height squared ($m^2$) |
| Mother's highest level of education completed | Categorical | • No formal education or some primary school<br>• Primary school or some secondary school incomplete |

| | | |
|---|---|---|
| | | • Secondary school or vocational or university/college<br>• Missing |
| Mother's minimum diet diversity, at baseline and at infant age (12 months) | Categorical | (asked at baseline and B4)<br>Categories (corresponding diet diversity score):<br>• Low (< 4)<br>• Medium (4-5)<br>• High (>5)<br>• Missing |
| **Pregnancy Factors** | | |
| Nulliparous (Never having given birth before) | Categorical | If A1 = 1 or (A1 = 0 and A5 = 0 and A6 = 0) then nulliparity = 1;<br>else if A1 ne . then nulliparity = 0;<br>else if A1 eq . then nulliparity = .;<br>A1 = Is this your first pregnancy?<br>A5 = How many of your children were born alive?<br>A6 = How many of your children were stillborn?<br>Yes / No / Missing |
| Gestational age at baseline (weeks) | Continuous | Calculated as the date at baseline minus the date of screening ultrasound plus gestational age at screening, and then divided by 7 |
| **Child factors** | | |
| Gestational age at birth | Continuous | Weeks |
| Preterm | Categorical | • Born before 37 weeks of gestation<br>• Born >= 37 weeks of gestation<br>• Missing |
| Infant sex | Categorical | • Male<br>• Female<br>• Missing |
| Breastfeeding, ever | Categorical | (asked at B4)<br>C32: Q1 Has *(**NAME OF INDEX CHILD**)* ever been breastfed? |
| Breastfeeding at 3, 6, 9, and 12 months, duration | Categorical | (asked at B1, B2, B3, and B4)<br>C32: Q2 Was *(**NAME OF INDEX CHILD**)* breastfed yesterday during the day or at night? |
| **Household factors** | | |
| Household food insecurity score, at baseline and at infant age (12 months) | Categorical | (asked at baseline and B4)<br>Categories (corresponding score):<br>• Food secure (0)<br>• Mild (1,2,3)<br>• Moderate (4,5,6) / Severe (7,8)<br>• Missing<br>See http://www.fao.org/3/as583e/as583e.pdf |
| Number of people who sleep in this house | Continuous | |
| Second-hand smoking | Categorical | Whether someone other than the pregnant woman in household smokes (smoking of the pregnant mother was an exclusion criteria) (yes/no/missing) |
| Assets | Categorical | Responses for each of the following 5 items: TV, radio, mobile phone, bicycle, and bank account. (Yes / No / Missing) |

## 5. DATA ANALYSIS

In this section we provide the analytic approach for the exposure-response aims related to infant anthropometrics at 12 months.

### 5.1. Outcome Definitions

This section describes each outcome, including data collection approaches and calculations for derived outcomes.

Infant length and weight were assessed quarterly (3, 6, 9, and 12 months of age) using standardized procedures.

Recumbent length was measured to the nearest 0.1 cm using a measuring board (seca 417). If the first and second length measurements differed by >0.7 cm, a third measurement was taken. The two closest measurements will be averaged.

Weight was measured using a calibrated scale (seca 876). Using the TARE function, naked or lightly dressed infants were weighed in their mother's arms. Two weights were recorded to the nearest 0.1 kg. If the 2 weight measurements differ by more than 0.1 kg, then a third weight measurement was taken. The two closest measurements will be averaged.

Head circumference was measured twice at the maximal circumference to the nearest 0.1 cm using Gulick II tapes with a tensioning device. If the 2 head circumference measurements differ by more than 0.5 cm, then a third head circumference measurement should be taken.

Stunting (LAZ), underweight (WAZ), and wasting (LWZ) Z-scores will be calculated based on the 2006 WHO Multicenter Growth Reference Standards (MRGS), which formulated prescriptive growth curves based on a multi-country cohort of healthy infants who were optimally fed.[8] Stunting, underweight, and wasting are defined as a Z-score that is < -2 SD below the median reference population. Implausible values (e.g. infant's length or head circumference at 12 months is less than measurement at 9 months) and outliers (identified by Z-scores falling outside of range (see below for defined parameters) will be flagged and excluded.[9]

- length-for-age Z-score falling outside of (-6, +6)
- weight-for- age Z-score falling outside of (-6, +5)
- weight-for-length Z-score falling outside of (-5, 5)

**Research Question:**
Are there exposure–response associations between prenatal and postnatal HAP exposures ($PM_{2.5}$, BC, and CO) on the following outcomes of infant growth at 12 months?

**Primary analysis**

1. Stunting at 12 months (LAZ < -2 SDs below the median WHO MGRS), dichotomous

**Secondary analyses**

2. Length-for-age Z-score at 12 months (LAZ), continuous
3. Severe stunting (LAZ < -3 SDs) at 12 months, dichotomous
4. Weight for age Z-score (WAZ), continuous
5. Underweight at 12 months (WAZ < -2 SDs below the median WHO MGRS), dichotomous
6. Wasting at 12 months, assessed by Weight for length z-score (WLS), continuous

### 5.2. Exposure-Response Analysis

For each pollutant ($PM_{2.5}$, CO, and black carbon), we will estimate the following exposure metrics:

(1) Post-natal average exposures for the children. After birth, we reconstructed children's exposure using a microenvironmental measurement approach and statistical models (max n = 3).

(2) Prenatal average maternal exposures weighted by time until randomization. For participants in the control group, the average exposure will be estimated using available prenatal maternal measures (max n = 3). For the intervention group, gestational days prior to LPG installation will be assigned the maternal baseline measurement value (n = 1); gestational days following intervention with LPG will be assigned the average of post-randomization prenatal maternal measurements (n = 2).

#### 5.2.1 Dichotomous outcomes

The 3 dichotomous outcomes of interest include stunting, severe stunting, and underweight. For the dichotomous outcomes of interest (e.g., LAZ <-2 SD) we will use log-binomial regression to characterize relative risk of the outcomes given the personal $PM_{2.5}$/BC/CO exposure. The general base model specification is as follows:

$$log(Y_i) = \beta_0 + \beta_1 f(PrenatalExposure_i) + \beta_2 f(PostnatalExposureC_i) + \beta_3 SES_i$$

where $Y_i$ is the expected occurrence of the outcome of subject $i$, $\beta_0$ is the population intercept, $\beta_1$ and $\beta_2$ are the exposure coefficients for the average prenatal and postnatal exposures respectively and $f(\text{PrenatalExposure}_i) + f(\text{PostnatalExposure}_i)$ are functions (i.e., linear, log linear, or categorical in quartiles) that use the average exposure of interest for the prenatal and postnatal periods of exposure (as described above), $\beta_3$ is the coefficient for the socioeconomic index used to adjust the model for the total effect of the gestational and postnatal exposure, as shown in the DAG in Figure 1. Results will be expressed as relative risk of occurrence with 95% confidence interval (CI) per unit (or IQR) increase in single pollutant models for $PM_{2.5}$/BC/CO exposures.

#### 5.2.2 Continuous outcomes

We will analyze the associations between long-term exposures and final growth outcomes at 12 months of age. We will estimate the association between long-term exposure and the 3 continuous anthropometric outcomes using the regression model given by:

$$E[d_i] = \beta_0 + \beta_1 f(\text{PrenatalExposure}_i) + \beta_2 f(\text{PostnatalExposure}_i) + \beta_3 \text{SES}_i$$

where $\boldsymbol{d}$ is the change in outcome between the follow-up measurement at the B4 visit and the baseline measurement at birth, $\beta_1$ and $\beta_2$ are the exposure coefficients for the average prenatal and postnatal exposures respectively and $f(\text{PrenatalExposure}_i) + f(\text{PostnatalExposure}_i)$ are functions (i.e., linear, log linear, or categorical in quartiles) that use the average exposure of interest for the prenatal and postnatal periods of exposure, as defined above, and $\beta_3$ is the coefficient for the socioeconomic index. This model evaluates the change in the continuous anthropometric measurements from baseline to end of follow-up, assessing the relative effect of our exposure metrics on the outcomes at 12 months.

We will evaluate the shape of the relationship between the anthropometric measures and the three air pollutants in separate models. Non-linear associations between anthropometric measures and exposures will be evaluated via (1) log transformation of the exposure, (2) exposure categories based on quartiles, (3) regression splines of varying order and number of internal knots (we will start with a knot selection based on quartiles), and (4) penalized smoothing splines. Model comparison will be based on traditional goodness-of-fit

(e.g., plotting observed and predicted values, use of residual plots and added variable plots, use of $R^2$) and information criteria to measure prediction error (i.e., AIC) to identify the best-fitting most parsimonious model. Residual normality and potential outliers will be considered to assess the model assumptions.

**Confounders and covariates to be included in the model**

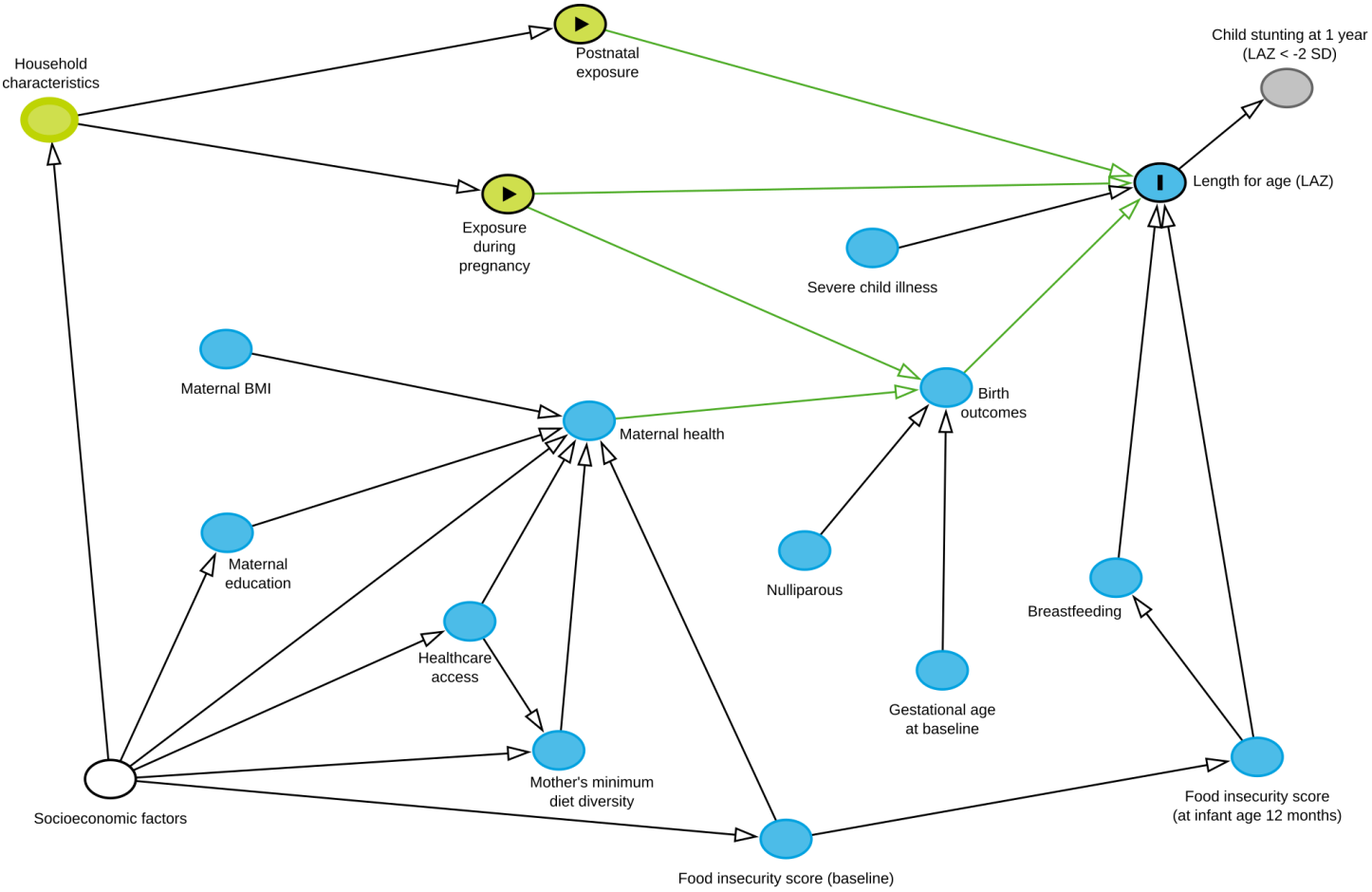
Confounder selection is based on conceptual directed acyclic graphs (DAG) and from previous studies.[9] Additional covariates are included to explain variance in the outcome (e.g., maternal height). **Figure 1** below shows the DAG used when considering potential confounders and sources of bias. The figure shows the DAG only for the LAZ outcome, but it is equivalent to the DAG for WAZ, WLZ, and the dichotomized variables derived from those measures.

Following this DAG, the models testing for total effect of the gestational and postnatal exposure should be adjusted by socioeconomic factors. Additionally possible effect modifications should be assessed for IRC, maternal characteristics (maternal height, Maternal minimum diet diversity, Household food insecurity at baseline), pregnancy factors (gestational age at baseline) an birth outcomes (infant sex, small for gestational age at birth). These subgroup analyses will be conducted using interaction terms between each of our exposure variables and the effect modifiers. The covariates considered for pre-specified subgroup analyses are described in more detail in Table 2.

***Sensitivity Analyses:*** We will perform sensitivity analyses for the models described above by adjusting for the variables listed under *Variables considered for sensitivity analyses* in Table 2.

**Figure 1:** Directed acyclic graph for the total effect of exposure to the air pollutants on the standardized

length for age.



| Table 2. A priori covariate adjustments in exposure-response analyses | | |
|---|---|---|
| **Parameter** | **Type** | **Subgroup Definitions** |
| ***Adjustment to calculate the total effect of the exposure on the outcomes*** | | |
| Baseline: Household SES index (Socioeconomic factors) | Continuous | SES index based on ownership of 24 selected household assets, water and sanitation quality, access to electricity, number of people in the household, food insecurity, participant's education level, and floor, wall, and roofing material, as calculated for HE-04. |
| ***Variables considered to evaluate effect modification*** | | |
| International Research Center | Categorical | Guatemala, India, Peru, Rwanda |
| Maternal height at Baseline | Continuous | Height in meters |
| Baseline: Household food insecurity score | Categorical | Categories (corresponding score):<br>Food secure (0)<br>Mild (1,2,3)<br>Moderate (4,5,6) / Severe (7,8)<br>Missing<br>See http://www.fao.org/3/as583e/as583e.pdf |
| Baseline: Maternal minimum diet diversity | Categorical | Categories (corresponding diet diversity score):<br>• Low (< 4)<br>• Medium (4-5)<br>• High (>5)<br>• Missing |
| Gestational age at baseline | Continuous | As calculated for HE-01 |
| Infant sex | Categorical | Male / Female |
| Small for gestational age at birth | Categorical | As calculated for HE-04 |
| ***Variables considered for sensitivity analyses*** | | |
| At 12 months: Infant currently breastfeeding | Categorical | Yes/No (no question was used to evaluate weaning age) |
| At 12 months: Household food insecurity score | Categorical | Categories (corresponding score):<br>Food secure (0)<br>Mild (1,2,3)<br>Moderate (4,5,6) / Severe (7,8)<br>Missing<br>See http://www.fao.org/3/as583e/as583e.pdf |
| At 12 months: Maternal minimum diet diversity | Categorical | Categories (corresponding diet diversity score):<br>• Low (< 4)<br>• Medium (4-5)<br>• High (>5)<br>• Missing |
| Severe child illness | Categorical | Whether the child experienced during the first year of life any severe illness episode that required hospitalization (including pneumonia or diarrhea) |
| Maternal education (Highest education level achieved) | Categorical | 1. No formal education<br>2. Primary school incomplete<br>3. Primary school complete<br>4. Secondary school incomplete (e.g. high school)<br>5. Secondary school complete (e.g. high school)<br>6. Vocational<br>7. Some college or university |

| | | |
|---|---|---|
| Preterm birth | Categorical | •Born before 37 weeks of gestational<br>•Born >= 37 weeks of gestational<br>•Missing |
| Second-hand smoking | Categorical | Whether someone other than the pregnant woman in household smokes (smoking of the pregnant mother was an exclusion criteria) (yes/no/missing) |

***Missing Data.*** A complete-case analysis will be carried out by excluding participants without an anthropometric measurement record. Missing confounder information will be addressed with the use of a missing categorical variable for each covariate (i.e., the missing by indication approach). In the exposure-response analysis, participants without time-weighted pollutant exposures as defined in Section 5.4 will be excluded in a complete-case analysis.

Based on HE-4 (published, January 2024), ~20% of anthropometric data are missing at 12 months due to COVID pandemic, therefore we will explain the large proportion of missingness as a limitation of this analysis. However, missingness in length measurements were balanced between intervention arms.

**5.4. Analysis Replication Plan**

Exposure-response analyses will be replicated by an independent analyst. Sensitivity analyses will not be replicated.

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

1. A cleaned analytic dataset where exclusions have been applied following the CONSORT diagram. The dataset will also include maternal characteristics at baseline, covariates for subgroup analysis and covariates to include in the exposure-response analyses.
2. The set of outcomes (primary and secondary) and subgroup analysis to be replicated.
3. 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 participant characteristic table.
2. Replicate exposure-response analyses according to models specified in Section 5.2.

**References**

1. de Onis M, Blössner M. The World Health Organization Global Database on Child Growth and Malnutrition: methodology and applications. *Int J Epidemiol*. 2003;32(4):518-526. doi:10.1093/ije/dyg099

2. Proportion of population with primary reliance on clean fuels and technologies for cooking (%). Accessed April 26, 2022. https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-phe-primary-reliance-on-clean-fuels-and-technologies-proportion

3. Bennitt FB, Wozniak SS, Causey K, Burkart K, Brauer M. Estimating disease burden attributable to household air pollution: new methods within the Global Burden of Disease Study. *Lancet Glob Health*. 2021;9:S18. doi:10.1016/S2214-109X(21)00126-1

4. Quinn A. Compensating control participants when the intervention is of significant value: experience in Guatemala, India, Peru and Rwanda. Presented at: ISEE; 2019.

5. 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. *Environ Health Perspect*. 2020;128(4):047008. doi:10.1289/EHP6407

6. Barr Dana Boyd, Puttaswamy Naveen, Jaacks Lindsay M., et al. Design and Rationale of the Biomarker Center of the Household Air Pollution Intervention Network (HAPIN) Trial. *Environ Health Perspect*. 128(4):047010. doi:10.1289/EHP5751

7. 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. *Environ Health Perspect*. 2020;128(4):047009. doi:10.1289/EHP6422

8. de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull*. 2004;25(1 Suppl):S27-36. doi:10.1177/15648265040251S104

9. WHO | Recommendations for data collection, analysis and reporting on anthropometric indicators in children under 5 years old. WHO. Accessed February 17, 2023. http://www.who.int/nutrition/publications/anthropometry-data-quality-report/en/

10. Kim R, Mejía-Guevara I, Corsi DJ, Aguayo VM, Subramanian SV. Relative importance of 13 correlates of child stunting in South Asia: Insights from nationally representative data from Afghanistan, Bangladesh, India, Nepal, and Pakistan. *Soc Sci Med*. 2017;187:144-154. doi:10.1016/j.socscimed.2017.06.017