STATISTICAL ANALYSIS PLAN

Exposure-Response Analysis for Blood Pressure in Older Adult Women
Manuscripts ER-2

Version 1.0
May 25, 2022

Household air pollution and health: A multi-country LPG stove intervention trial
As part of the Household Air Pollution Intervention Network (HAPIN) Trial

Trial Registration: NCT02944682
Protocol Version: HAPIN - Main Study Protocol v14.0 20January2022

Modification History: NA

1. INTRODUCTION

This document contains the statistical analysis plan (SAP) for blood pressure in the older adult women, one of four primary outcomes of the HAPIN trial. This SAP will be posted with the trial registration.

1.1. Background and Rationale

Globally, nearly 3 billion people rely on solid fuels for cooking and heating, the vast majority in low- and middle-income countries (LMICs). The resulting household air pollution is the most important environmental risk factor in the 2019 global burden of disease, accounting for an estimated 2.3 million premature deaths annually, largely among women and young children. Previous interventions have provided cleaner biomass-based cookstoves but have failed to reduce exposure to levels that produce meaningful health improvements. There have been no large-scale field trials with liquefied petroleum gas (LPG) cookstoves, likely the cleanest scalable intervention.

This study will provide evidence, including costs and implementation strategies, 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 household air pollution interventions globally.

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 LMIC settings (Tamil Nadu, India; Jalapa, Guatemala; Puno, Peru; and Kayonza, Rwanda) to deliver rigorous evidence regarding potential health benefits across the lifespan. Each intervention site will recruit 800 pregnant women (aged 18-<35 years, 9 to <20 weeks gestation), and will randomly assign half their households to receive LPG stoves and an 18-month supply of LPG. Controls will not receive the intervention at the commencement of the trial and are anticipated to continue cooking with solid biomass fuels; they will be compensated for their participation in the study. The mother will be followed along with her child until the child is 1 year old. In households with a second, non-pregnant older adult woman (aged 40 to <80 years) we will also enroll and follow her during the 18-month follow-up period in order to assess cardiopulmonary, metabolic, and cancer outcomes. To optimize intervention use, we will implement behavior change strategies. We will assess cookstove use, conduct repeated personal exposure assessments to HAP ($PM_{2.5}$, black carbon, carbon monoxide), and collect dried blood spots (DBS) and urinary samples for biomarker analysis and biospecimen storage on all participants at multiple time points. The primary outcomes are 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 study will address 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 is a randomized, 2-arm intervention trial with parallel assignment. Study sites in the four countries (Guatemala, India, Peru, Rwanda) have been selected and evaluated based on activities conducted in the formative research. HAPIN uses a rolling recruitment process whereby each International Research Center (IRC) will enroll 800 pregnant women (one per household) and an additional approximately 120 older adult women (this will vary by IRC) 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 (Clasen et al. 2020).

Recruitment and enrollment will occur over approximately 15 months at ~53 pregnant women/8 older adult women per month per IRC. All participants will be followed longitudinally for ~18 months (until the child is age 1 year).

2.2. Randomization

To ensure balance between arms, households have been randomly allocated to intervention or control arms as and when they consent to participate. 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 randomization list is stratified by the two study sites
- The Peru IRC randomization list is stratified by the six study 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, 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 sample size and power for the trial were based on the intention to treat analyses for the primary outcomes.

2.4. Trial Framework

HAPIN is a superiority trial. The intention-to-treat analysis and exposure-response analysis (described here) will be conducted per the original aims of the study.

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 are complete and the SAP has been approved and registered.

2.7. Timing of Outcome and Covariate Assessments

Each participating household is to be followed from enrollment until the index child reaches (or would have reached, assuming a live birth and continued vitality) his/her first birthday. For the participating older adult woman in a household the follow up time is approximately 18 months, including baseline measurements (prior to randomization and intervention), when the pregnant woman is at 24-28 and 32-36 weeks of gestation, and when the child is 3 months, 6 months, and 12 months of age (up to 6 total visits).

3. STATISTICAL PRINCIPLES

3.1. Confidence Intervals and P-Values

All confidence intervals will be presented at 95% confidence. Analysis of air pollution exposure-response associations will use an α-level of 0.05 to identify statistical significance (including effect modification). If the effect modifiers have more than two categories, simultaneous hypothesis tests will be used.

3.2. Adherence and Protocol Deviations

All homes in the intervention arm will be equipped with Stove Use Monitoring Systems (SUMS) on their traditional stoves, as well as a subset of approximately 80 homes in the control arm of each IRC. Compliance will be checked every two weeks when SUMS data is downloaded.

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

3.3. Analysis Populations

For each outcome, the analysis will include all valid outcome measurement (*complete-case*). We define loss to follow-up as any reason that contributes to a missing outcome value, including death of the study participant and withdrawal from study prior to measurement.

Secondary analysis may use various subsets of the study to examine effect modification.

4. TRIAL POPULATION

4.1. Eligibility

Pregnant women will be eligible to participate in the study if they fulfill the following inclusion and exclusion criteria at screening:

Inclusion criteria:

- Confirmed pregnancy (hCG positive blood or urine test)
- Aged 18 to <35 years (via self-report)
- Uses biomass stove predominantly
- Lives in study area
- 9 – <20 weeks gestation confirmed by ultrasound
- Singleton pregnancy (one fetus)
- Viable fetus with normal fetal heart rate (120-180 beats per minute) at time of ultrasound
- Continued pregnancy at the time of randomization confirmed by self-report
- Agrees to participate with informed consent

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 pregnant women live in the same household and are interested in participating, the one with the earliest gestational age will be chosen to participate.

An **older adult woman** in the same household will be eligible to participate in the study if they fulfill the following inclusion and exclusion criteria at screening:

Inclusion criteria:

- Aged 40 to <80 years (via self-report)

Exclusion criteria:

- Currently smokes cigarettes or other tobacco products
- Pregnant (via self-report)
- Plans to move out of her current household in the next 12 months
- Takes blood pressure medication at enrollment and/or any point during follow-up (will be included in sensitivity analysis, below)

If two or more older adult women live in the same household and are interested in participating, one woman will be randomly selected to participate (the one with the next birthday[month and day]).

4.2. Recruitment

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

- Reasons for exclusion when assessed for eligibility
- Participants determined to be ineligible after randomization
- Reasons for exits after randomization
    - Voluntary withdrawal
    - Withdrawn by study team
    - Moved away
- Reasons for exclusion due to missing data

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
- 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 soon as possible.

5. DATA ANALYSIS

In this section we provide the analysis approach for the exposure-response aim. The primary outcome is systolic blood pressure. We present the primary analysis along with secondary outcomes, subgroup (effect modification), sensitivity, and additional analyses.

5.1. Outcome Definitions

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

Following the 24-hour exposure assessment period, a nurse or trained field worker measured resting blood pressure in the right arm in triplicate (with at least 2 minutes between measurements) using an automatic monitor (model HEM-907XL; Omron®) at the participant's home. Before starting the measurement, the participant was instructed to sit on a chair in a quiet room for 5 min with legs uncrossed, their back supported by the chair, and their arm supported on a table. The participant also confirmed that she had not smoked, consumed alcohol, or caffeinated beverages (coffee, tea, or Coca-Cola), or cooked using biomass in the past 30 minutes. If she had done any of those activities in the 30 minutes prior to the measurement, she was asked to refrain from doing these activities for 30 minutes before proceeding with the measurements.

A participant with a measured systolic blood pressure ≥ 140 mmHg and/or a diastolic blood pressure ≥ 90 mmHg was checked again during the same visit. If the same result was observed on two measurements, the participant was referred to the nearest health center or hospital to receive age-appropriate treatment. If a participant had systolic blood pressure < 80 mmHg or diastolic blood pressure < 40 mmHg, she was also referred to the nearest health center or hospital. In analyses, the average of all three blood pressure measurements will be used. Systolic blood pressure values less than 70 mmHg and diastolic blood pressure values less than 35 mmHg will be excluded as implausible.

Pulse pressure is the difference between systolic blood pressure (SBP) and diastolic blood pressure (DBP) and mean arterial pressure is calculated as $DBP + (SBP - DBP)/3$.

.5.2. Exposure-Response Analysis

For each pollutant (PM2.5, black carbon and CO), time-weighted average exposures will be estimated using 24-hr personal measurements at baseline, and al follow-up visits (up to 5). For participants in the control group, an average will be calculated from all available measurements. For participants in the intervention group, baseline exposure levels will be weighted by days prior to LPG installation, while post baseline measurements will be weighted by the days after LPG installation. For the intervention group, if the baseline measurement is missing the mother will be excluded from the analysis.

The following two approaches will be used for both primary and secondary blood pressure outcomes. After excluding any women who reported using blood pressure medication at any visit, we will estimate the association between long-term exposure and the blood pressure outcome using the regression model given by

$$E[d_i] = \beta_0 + f(X_i) + \boldsymbol{\gamma Z}_i$$

where d is the change in outcome between the 6th follow-up measurement (B4 visit) and the baseline measurement, $f(X_i)$ is a function that uses the average exposure of interest across the follow-up period, defined above, and $\boldsymbol{Z}_i$ is the vector of confounders. This model evaluates the change in BP from baseline to end of follow-up, assessing the cumulative effect of exposure over the whole period.

Confounders and covariates to be included in the model are listed in Table 2, chosen a priori.

We will also use the following random-intercept model to analyze repeated measurements of blood pressure (separately in relation to each of our three pollutants) given by

$$E[Y_{ij}] = \beta_0 + \theta_i + \beta_1 X_i + \beta_2 (X_{ij} - X_i) + \boldsymbol{\gamma Z}_i$$

where y_{ij} denotes the outcome for woman i at time point j, θ_i is the woman-specific random intercept, and $(X_{ij} - X_i)$ is the woman-specific deviation for exposure at time point j. Here parameter β_1 reflects the 'long term' or average effect of exposure across the study period. The exposure variable X_i will be the average of all measurements as described above. The parameter β_2 describes the within-woman short-term association between exposure and outcome, adjusted for impacts of long-term exposures (both such effects have been found in studies of ambient $PM_{2.5}$) This model contains two special cases: (1) when $\beta_1 = \beta_2$, the model reduces to a random-intercept model that only uses the time-varying exposure directly, and (2) when $\beta_2 = 0$, the model reduces to a random-intercept model with only average exposure.

In the exposure-response analyses, all models will be adjusted for the covariates listed in Table 2. Confounder selection are based on conceptual directed acyclic graphs and from previous studies. Additional covariates included to explain variance in the outcome (e.g., age) to increase power.

Table 2. A priori covariate adjustments in exposure-response analyses

Parameter	**Type**	**Subgroup Definitions**
International Research Center	Categorical	Guatemala, India, Peru, Rwanda
Baseline: Age at baseline (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.
Baseline: Highest level of education completed	Categorical	• No formal education or some primary school • Primary school or some secondary school incomplete • Secondary school or vocational or university/college • Missing
Baseline: Body mass index (BMI)	Continuous	BMI calculated as the average weight (kg) divided by the average height squared (m^2)
Baseline: Household food insecurity score	Categorical	Categories (corresponding score): • Food secure (0) • Mild (1,2,3) • Moderate (4,5,6) / Severe (7,8) • Missing See http://www.fao.org/3/as583e/as583e.pdf
Baseline: Minimum diet diversity	Categorical	Categories (corresponding diet diversity score): • Low (< 4) • Medium (4-5) • High (>5) • Missing
Time of the blood pressure measurement; time-varying	Categorical	AM or PM

We will evaluate the shape of the relationship between blood pressure and exposures. Non-linear associations between blood pressure and exposure 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 most parsimonious model.

Subgroup Analysis. Subgroup (effect modification) analyses will be conducted using interaction terms between each of our three pollution variables and the effect modifiers. The list of pre-specified subgroup analyses for the exposure-response analysis is given in Table 3.

Table 3: Definition for variables for subgroup analysis for exposure-response analysis

Parameter	**Subgroup Definitions**
Age	Continuous and dichotomous (using median age)
International Research Center	Guatemala, India, Peru, Rwanda
BMI	Continuous and categorical (underweight, less than 18.5 and normal, 18.5-24.9, combined; overweight/obese, 25.0 and above). And also using 3 categories (underweight, normal, overweight/obese).
Study Arm	Categorical; intervention vs. control

Secondary Outcomes: We will evaluate diastolic blood pressure, pulse pressure, and mean arterial pressure as secondary outcomes.

Sensitivity Analysis. We will conduct an analysis including women taking blood pressure medication. We will also evaluate other potential confounders (physical activity, reported smokers in household, recent caffeine intake) by including them in the model.

Additional Analysis: We will consider extend the random intercept model to (1) include a linear time trend and its interactions with average exposure (X_i) for estimating associations between exposure and outcome trajectory, and (2) consider time-varying exposures defined as cumulative averages.

Missing Data. Our primary approach to missing outcome data will be a complete-case analysis by excluding participants with a missing baseline blood pressure measurement. All follow-up measurements will be used in the ITT analysis. It is anticipated that missing blood pressure will be balanced between intervention arms.

5.3. Analysis Replication Plan

Selected components of the 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 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 baseline characteristic table.
2. Replicate exposure-response analyses for primary and secondary outcomes according to models specified in Section 5.2