



Full length article

Personal household air pollution exposure and SARS-CoV-2 antibody responses among women and infants: an analysis within the HAPIN trial

Lindsay J. Underhill^{a,*}, Maggie L. Clark^b, Dana Boyd Barr^c, Kalpana Balakrishnan^d, Naveen Puttaswamy^d, Ajay Pillarisetti^e, Anaite Diaz-Artiga^f, Lisa M. Thompson^m, John P. McCracken^{f,g}, Victor G. Davila-Roman^a, Lance A. Waller^k, Bonnie N. Young^b, Michael A. Johnson^h, Karthikeyan Rajamani^l, William Checkley^{i,j}, Thomas F. Clasen^c, Jennifer L. Peel^b, Kyle Steenland^c, On behalf of the HAPIN Investigators*

^a Cardiovascular Division and Global Health Center, WashU School of Medicine, St. Louis, MO, USA

^b Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO, USA

^c Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

^d ICMR Center for Advanced Research on Air Quality, Climate, and Health, Department of Environmental Health Engineering, Sri Ramachandra Institute for Higher Education and Research, Chennai, India

^e Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, CA, USA

^f Centro de Estudios en Salud, Universidad del Valle de Guatemala, Guatemala City, Guatemala

^g Epidemiology and Biostatistics Department, College of Public Health, University of Georgia, Athens, GA, USA

^h Berkeley Air Monitoring Group, Berkeley, CA, USA

ⁱ Division of Pulmonary and Critical Care, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

^j Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

^k Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

^l Centre for Global Health Research, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Thandalam, Chennai-602105, Tamil Nadu, India

^m Department of Family Health Care Nursing, School of Nursing, University of California, San Francisco, San Francisco, CA, USA

A B S T R A C T

Background: While extensive research has linked ambient air pollution to increased COVID-19 infection rates and severity, the role of household air pollution—a major public health exposure affecting nearly 3 billion people globally—remains understudied. Here, we evaluate associations between a clean cooking intervention, personal household air pollution exposure, and serologic evidence of prior SARS-CoV-2 infection using exploratory intention-to-treat and exposure–response analyses.

Methods: The Household Air Pollution Intervention Network (HAPIN) Trial randomized 3,200 households in Guatemala, India, Peru, and Rwanda to either continue using traditional solid fuel cookstoves or receive liquefied petroleum gas stoves with continuous fuel supply and behavioral messaging. The current ancillary study leveraged personal exposures to PM_{2.5} and carbon monoxide (CO) and dried blood spot (DBS) samples collected in the Guatemala, India, and Peru HAPIN sites to study the association between personal air pollution exposures and COVID-19; Rwanda was excluded due to low COVID-19 prevalence at the time of collection. SARS-CoV-2-specific antibodies were measured from DBS samples to define serology-based infection status and quantify immune response. Analyses included exploratory intention-to-treat and observational exposure–response analyses of associations between personal air pollution exposure and serology-based SARS-CoV-2 infection status and antibody levels.

Results: While the intervention effectively reduced household air pollution exposures, no significant differences in COVID-19 infection or antibody levels between intervention and control groups were found. However, exposure–response analyses were suggestive of potential inverse associations between PM_{2.5} exposure and antibody levels, diverging from most studies of ambient air pollution.

Conclusions: These findings suggest that household air pollution may influence SARS-CoV-2 infection and immune response through mechanisms distinct from those described for ambient air pollution.

* Corresponding author.

E-mail address: lindsayu@wustl.edu (L.J. Underhill).

<https://doi.org/10.1016/j.envint.2026.110180>

Received 31 December 2024; Received in revised form 26 February 2026; Accepted 27 February 2026

Available online 10 March 2026

0160-4120/© 2026 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Background

Since 2020, a growing body of evidence has linked higher air pollution exposure to increased COVID-19 infection rates and disease severity (Chen et al., 2022; Hernandez Carballo et al., 2022; Sheppard et al., 2023; Zang et al., 2022). However, the association between COVID-19 outcomes and household air pollution—which affects nearly 3 billion people worldwide—remains largely unexplored. Household air pollution exposure is a significant public health problem, with the largest burdens in low-resource communities in the Global South that rely on biomass, coal, and kerosene for cooking and heating. The incomplete combustion of these fuels generates household air pollution, including fine particulate matter (PM_{2.5}) and carbon monoxide (CO). Exposure to PM_{2.5} and CO has been associated with respiratory infections and cardiovascular disease, both of which are established risk factors for severe COVID-19 outcomes. Additionally, in observational studies focused on ambient air pollution, PM_{2.5} and CO have been associated with increased COVID-19 cases, severity, and mortality (English et al., 2022; Hernandez Carballo et al., 2022; Kogevinas et al., 2021; Yu et al., 2024). Proposed biological mechanisms include oxidative stress, lung tissue damage, and endothelial dysfunction, which render individuals more vulnerable to respiratory infections. Additionally, the upregulation of angiotensin-converting enzyme-2 (ACE-2) expression—a receptor for SARS-CoV-2—may also play a role in the exposure disease pathway (Yu et al., 2024). Despite these plausible links, there remains a critical gap in research on indoor air pollution, including household air pollution, and COVID-19. Given that millions of people are exposed to household air pollution daily, understanding its role in COVID-19 transmission and severity is crucial for informing public health strategies, particularly in low-resource communities where exposure is highest.

Most studies of air pollution and COVID-19 have been observational, using population-level data on infection rates, hospitalizations, and mortality, and area-level ambient air pollutant concentrations. These studies present substantial evidence of a positive association between increased air pollutant exposure and COVID-19. No studies to our knowledge have incorporated personal exposure assessments, which may provide more accurate measures of individual-level exposure to pollutants like PM_{2.5} and CO. Furthermore, most research has relied on acute infection markers, such as RT-qPCR or rapid antigen testing, ascertained through hospital records, self-report, or study-based testing. Few studies have investigated infection using serological measurement of SARS-CoV-2 antibodies (Kogevinas et al., 2021). While PCR and antigen tests are more accurate for detecting acute infection, serological testing provides insight into prior infections, including asymptomatic or undiagnosed cases, offering a complementary perspective on cumulative SARS-CoV-2 exposure. Furthermore, to date, no study has specifically focused on the relationship between HAP and COVID-19, leaving an important gap in our understanding of how indoor air pollutant exposure relates to infection risk.

The current study aims to address these gaps by leveraging data from the Household Air Pollution Intervention Network (HAPIN) Trial, a multi-country randomized controlled trial (RCT) conducted in Guatemala, India, Peru, and Rwanda. The aim of the HAPIN trial was to test the effects of replacing traditional solid fuel cookstoves with liquefied petroleum gas (LPG) stoves on health outcomes across the life course, including birthweight, infant stunting, infant severe pneumonia, and blood pressure in adult, non-pregnant women. The current analysis uses a subset of data and biospecimens from the HAPIN Trial. Using data from three of the four participating countries, where community transmission was highest, we evaluate whether serologic evidence of prior SARS-CoV-2 infection differs between intervention and control households and examine exposure–response relationships based on measured personal pollutant concentrations.

2. Methods

2.1. Parent study overview

The current study is based on a subset of longitudinal exposure and biomarker data from the HAPIN trial. HAPIN was an RCT of a liquefied petroleum gas (LPG) stove, continuous fuel distribution, and behavioral change intervention for approximately 18 months in 3200 households with pregnant women from Guatemala, India, Peru, and Rwanda. Households were randomized (1:1) to the intervention group—which received an LPG stove, continuous LPG fuel supply, and behavioral messaging to encourage exclusive LPG fuel use—or the control group, which did not receive the intervention. Control households were eligible to receive a free LPG stove and fuel supply (or equivalent compensation) at the end of study follow-up. Randomization began in March 2018 and was completed in February 2020. The HAPIN trial used a stratified randomization approach, first by country and then by study sites within the country, to ensure geographic balance across regions with distinct environmental and regional characteristics. Primary outcomes of the trial were birthweight, infant stunting, infant severe pneumonia, and blood pressure in the adult, non-pregnant women. Additional study design details and results for the primary outcomes have been published previously (Barr et al., 2020a; Clasen et al., 2020; Johnson et al., 2020).

To be eligible for the HAPIN trial, pregnant women had to be between 19 and 35 years of age, non-smoking, not planning to relocate, between 9 and 20 weeks of gestation with a confirmed singleton pregnancy (verified by ultrasound), and primarily cook with biomass. Additionally, non-pregnant, non-smoking women, aged 40–79 years, residing in a subset of pregnant woman households were also enrolled and are hereafter referred to as “other adult women” participants. All pregnant and other adult women participants provided written informed consent. Upon birth, the child of each enrolled pregnant woman was automatically included in the study. The HAPIN protocol included six scheduled outcome measurement visits per household: a pre-randomization baseline visit, followed by two prenatal visits approximately 1–3 and 3–5 months post-randomization, and three postnatal visits at roughly 9, 12, and 18 months post-randomization (when the infant was approximately 3–7 months, 6 months, and 12 months of age, respectively). The randomization and intervention delivery (for the intervention arm) occurred between the baseline and the first prenatal visit. At the baseline and all outcome measurement follow-up visits, 24-hour personal exposure sampling for PM_{2.5} and CO and the collection of urine and dried blood spot (DBS) samples were conducted. Special precautions were developed and undertaken to minimize infection risks from field activities during the COVID outbreak (Simkovich et al., 2021).

The HAPIN trial is registered with clinicaltrials.gov (NCT02944682). Study protocols and procedures were reviewed and approved by the institutional review boards or ethics committees of Emory University (00089799), Johns Hopkins University (00007403), Sri Ramachandra Institute of Higher Education and Research (IEC-N1/16/JUL/54/49), the Indian Council of Medical Research—Health Ministry Screening Committee (5/8/4–30/(Env)/Indo-US/2016-NCD-I), Universidad del Valle de Guatemala (146–08–2016), the Guatemalan Ministry of Health National Ethics Committee (11–2016), Asociación Beneficia PRISMA (CE2981.17), the London School of Hygiene and Tropical Medicine (11664–3), the Rwandan National Ethics Committee (No.016/RNEC/2018), and Washington University in St. Louis (201611159).

2.2. Assessment of personal exposure monitoring

The HAPIN exposure assessment protocol has been described in detail previously (Johnson et al., 2020). The 24-hour personal exposure measurements of PM_{2.5} and CO were collected using the Enhanced Children's MicroPEM (ECM), a validated, lightweight particulate sampler (RTI Inc, Research Triangle Park, NC, USA). The ECM uses a built-in internal 0.3 L/min pump to draw air through an impactor

attached to a cassette containing 15-mm Teflon® filters (PT15-AN-PF02; MTL Corporation), onto which the particulates deposit. Gravimetric measurements of PM_{2.5} were based on the pre- and post-weights of the filters, adjusted by sampling time. Real-time, continuous measures of PM_{2.5} were collected by an internal nephelometer (light scattering sensor) in the ECM. The ECM also monitored temperature, relative humidity, and filter-pressure drop. When a gravimetric sample was deemed invalid, due to missing or damaged filters or flow faults, nephelometer data was used to estimate PM_{2.5} concentrations. Real-time, continuous measures of CO were measured with EL-USB-300 monitors (Lascar Electronics), a lightweight, data-logging monitor. The Lascar instrument is approximately the size of a large marker, runs on ½ AA batteries, has a sensing range between 0 and 300 ppm, uses an electro-chemical cell to measure CO, and records the concentration at an interval of 2.5 s. During each 24-hour sampling period, the pregnant mothers and other adult women participants wore the ECM and Lascar in specially designed aprons or vests, created with participant input to ensure appropriateness for their settings. When unable to wear the monitors (e.g., while sleeping or bathing), they were instructed to place the monitors in a safe nearby location.

Since infants could not wear ECM or Lascar monitors, their personal exposures were estimated using an indirect reconstruction approach. During each sampling period, they were fitted with two coin-sized Bluetooth Beacon emitters (Model O, Roximity Inc., Denver, CO, USA). These emitters communicated with Beacon loggers (Berkeley Air Monitoring Group, Berkeley, CA, USA) co-located with the ECM and CO monitors worn by the mother and/or other adult women participants, as well as with additional ECM monitors placed in key household micro-environments (e.g., the infant's sleeping area, kitchen, or common areas). The beacon-logger system provided continuous proximity signals, indicating which monitored location the infant was closest to at each moment. Using these proximity data, the time infants spent near their mothers and in each monitored microenvironment was calculated. Time-weighted pollutant concentrations from these microenvironments were then combined to estimate each infant's 24-hour personal exposure. This indirect reconstruction method has been described in detail previously (Liao et al., 2020).

Exposure sampling was conducted for mothers at baseline, the 2 prenatal visits, and optionally during the 3 postnatal visits to support the infant's indirect exposure sampling; for other adult women at all visits; and for infants during the 3 postnatal visits.

2.3. HAPIN DBS collection

As part of the HAPIN trial, urine and DBS were collected to measure pre-specified biomarkers of household air pollution exposure and effects (Barr et al., 2020b). The original HAPIN consent and IRB approvals also permitted storage and secondary analyses of DBS samples for future health-related research. The DBS were collected at baseline and up to 3 additional timepoints in mother participants; all timepoints in other women participants; and the 3 postnatal visits in infants. The collection was conducted on the same day from all household participants at the end of the 24-hour exposure sampling periods. DBS were collected via finger prick by carefully allowing a drop of blood from a warmed, non-milked finger to fall onto a standard Guthrie DBS card. Each card was thoroughly dried before being packaged in a ziplock bag with a humidity indicator card and desiccant. The packaged DBS were kept at -20°C with the humidity indicators being checked periodically and desiccants being replaced as needed to keep the cards dry. Samples from the Guatemala and Peru locations were shipped to, and analyzed by, the Laboratory for Exposure Analysis and Development in Environmental Research (LEADER) at Rollins School of Public Health, Emory University, Atlanta, Georgia. All samples from the India site were analyzed at the laboratory at Sri Ramachandra Institute of Higher Education and Research (SRIHER), Chennai, India. As previously described (Barr et al., 2020b), these laboratories have conducted extensive cross-laboratory

validation and harmonization of analytic protocols to ensure comparability of biomarker measurements across sites. For the current SARS-CoV-2 substudy, both laboratories used the same commercial nucleocapsid-targeting ELISA platform described below and followed the manufacturer protocol, including duplicate testing and concurrent use of assay controls.

2.4. SARS-CoV-2 testing

For this ancillary analysis, we leveraged the longitudinal DBS samples collected as part of the parent HAPIN trial. All HAPIN participants (mothers, other adult women, and infants) from Guatemala, India, and Peru who were still enrolled and contributing DBS samples on or after March 15, 2020 (the estimated introduction of SARS-CoV-2 at the study sites) were eligible for inclusion. DBS samples collected before March 15, 2020 were not used in this ancillary study. Rwanda was excluded due to logistical and budgetary constraints, as well as very low community prevalence at the time of DBS collection.

For each eligible participant, the most recent DBS sample collected on or after March 15, 2020 was tested for SARS-CoV-2 antibodies. If this DBS sample was seropositive (based on the threshold described below), earlier DBS samples from the post-March 15 period were tested in reverse chronological order until the first seropositive sample was identified. This enabled estimation of the earliest detectable antibody response.

COVID-specific IgG antibodies were measured using the Epitope Diagnostics Novel Coronavirus COVID-19 IgG enzyme-linked immunosorbent assay (ELISA). These ELISA tests use a microplate-based technique that detects COVID-specific IgG antibodies, which are indicative of a long-term response (i.e., present and/or past infection). Since 10 µL of serum is required for the ELISA tests, a 6 mm punch from each DBS, equivalent to approximately 17–25 µL of serum, was used. Blood was eluted from the DBS punch with dilution buffer from the ELISA test kits. Assay controls, calibrants and diluted (1:100) blood (10 µL) samples were added to the wells of a microplate that was coated with COVID recombinant full length nucleocapsid protein. After incubation at physiologic temperature, the unbound protein matrix was washed away. A horseradish peroxidase (HRP) labeled polyclonal goat anti-human IgG tracer antibody was added to each well to form an immunocomplex of SARS COV-2 recombinant antigen, human anti-SARS COV-2 IgG antibody, and HRP labeled antihuman IgG tracer antibody when COVID-specific IgG antibody is present. After washing, an HRP substrate was added and incubated in a timed reaction to produce a chromophore that was measured via absorbance spectroscopy at 450 nm. The enzymatic activity of HRP in the immunocomplex is proportional to the amount of the anti-SARS COV-2 IgG antibody level in the serum. Concentrations were quantified using the Beer-Lambert Law. The limit of detection (LOD) of each assay was 5 IU/mL serum with relative standard deviations of <15%. The sensitivity and specificity of the ELISA were 98% and 99.8%, respectively.

For quality control, samples were analyzed in duplicate, and the average values were used for quantification. If the relative standard deviation (RSD) of the duplicate samples were >20%, the analysis was repeated. Both positive and negative controls were analyzed concurrently. Additionally, we conducted a pre-study validation assessment to ensure comparable data could be obtained from serum and DBS. For this assessment, we recruited 90 volunteers from Emory University (52 with prior COVID diagnosis and 38 without prior COVID diagnosis) and compared ELISA results from matched serum from a venous blood draw and DBS samples derived from a finger prick (Fig. S1). The DBS-based ELISA test demonstrated a sensitivity of 92% and specificity of 100% for detecting those who had been diagnosed.

2.5. SARS-CoV-2 outcome definition

For this ancillary study, SARS-CoV-2 antibody measurements were

used to define two outcomes: (1) a binary indicator of prior SARS-CoV-2 infection and (2) a continuous measure of antibody concentration. At the sample level, a DBS measurement was considered seropositive if the antibody concentration exceeded 5 IU/mL, based on the assay's analytical limit of detection (LOD) and the point of departure on the standard curve. At the participant level, evidence of prior SARS-CoV-2 infection ("infection positive") was defined as having at least one seropositive DBS sample across all tested time points; participants whose DBS samples all remained below this threshold were classified as infection negative. For the continuous outcome, antibody concentrations were summarized at the participant level using the mean of available DBS measurements and were used to characterize the magnitude of the immune response.

All instrument-derived concentrations that met quality control criteria were included in analyses without imputing values below the LOD (Buck Louis et al., 2016). Measurements below the LOD were retained but subject to greater uncertainty (Taylor, 2018).

2.6. Statistical analysis

Baseline participant and household characteristics were summarized as means and standard deviations for continuous variables and as counts and percentages for categorical variables, stratified by intervention arm. SARS-CoV-2 infection prevalence was defined as the proportion of participants with antibody concentrations above the assay's analytical limit of detection (5 IU/mL), interpreted as evidence of prior or current SARS-CoV-2 exposure. Prevalence was estimated by country and participant type. For observed prevalence, 95% CIs were calculated using the exact Clopper–Pearson binomial method, which is appropriate for proportions with small sample sizes (Clopper and Pearson, 1934; R Core Team, 2025). Assay-adjusted prevalence estimates were obtained using the Rogan–Gladen estimator to account for assay sensitivity (92%) and specificity (100%), and 95% CIs were derived by applying the Rogan–Gladen transformation to the corresponding exact Clopper–Pearson confidence limits (Rogan and Gladen, 1978). Assay performance (sensitivity and specificity) was evaluated in pre-study validation, as explained above; because the objective was to report observed infection prevalence based on measured antibody responses, estimates were not adjusted for test performance.

We conducted exploratory intention-to-treat (ITT) analyses to evaluate the association between intervention arm and ancillary SARS-CoV-2 serologic outcomes that were not pre-specified before randomization: binary infection (yes/no) and continuous antibody levels (u/mL), as defined above (Section 2.5). Although DBS samples were collected longitudinally, both outcomes were analyzed as participant-level aggregate measures, such that each individual contributed a single binary infection indicator and a single continuous antibody summary value to the regression models. In addition, we performed observational exposure–response analyses to evaluate the relationship between the same COVID-19 outcomes and long-term exposure to PM_{2.5} and CO. For all models, we used generalized estimating equations (GEE) employing an exchangeable correlation matrix to account for clustering within households. For models evaluating continuous antibody levels, a log link function was used with the Gaussian family to model the relationship between predictors and the logarithm of the antibody levels, to account for their right-skewed distribution (Fig. S2).

The primary ITT analysis for binary seropositivity is expressed as:

$$\text{logit}(P(\text{COVID}_i = 1)) = \beta_0 + \beta_1 \text{Arm}_i + \beta_2 \text{PType}_i + \beta_3 \text{IRC}_i + \beta_4 \text{Ppl}_i + \epsilon_i$$

In this model, COVID_i is a binary outcome indicating seropositivity for individual *i*, where 0 represents seronegative and 1 represents

seropositive. The intercept, β_0 , represents the baseline log odds of seropositivity. The coefficient for the intervention arm, β_1 , captures the difference in log odds of seropositivity between the control group (Arm_i = 0) and the intervention group (Arm_i = 1). The indicator variables PType_i and IRC_i were included to adjust for differences among participant types (mothers, children, and other adult women) and study sites (Guatemala, India, and Peru). Ppl_i is a continuous variable representing the number of people reported living in the household and ϵ_i represents the residual error term.

The secondary ITT analysis evaluating the association between intervention arm and continuous antibody levels (measured in u/ml) is specified below:

$$\text{AntibodyLevel}_i = \beta_0 + \beta_1 \text{Arm}_i + \beta_2 \text{PType}_i + \beta_3 \text{IRC}_i + \beta_4 \text{Ppl}_i + \epsilon_i$$

In this model, AntibodyLevel_i is a continuous outcome representing the antibody levels for individual *i*. The intercept β_0 represents the baseline mean antibody level, and β_1 captures the mean difference in antibody levels between the control and intervention groups. Like the primary analysis, this model includes the same covariates (PType, IRC, and Ppl) and an error term.

Exposure–response analyses evaluated the association of COVID-19 outcomes and exposure to PM_{2.5} and CO, using log-transformed exposure levels and exposure quartiles. The binary outcome (COVID-19 seropositivity) model is expressed as:

$$\text{logit}(P(\text{COVID}_i = 1)) = \beta_0 + \beta_1 \text{Pollutant}_i + \beta_2 \text{PType}_i + \beta_3 \text{IRC}_i + \beta_4 \text{Ppl}_i + \epsilon_i$$

For continuous antibody levels, the model is expressed as:

$$\text{AntibodyLevel}_i = \beta_0 + \beta_1 \text{Pollutant}_i + \beta_2 \text{PType}_i + \beta_3 \text{IRC}_i + \beta_4 \text{Ppl}_i + \epsilon_i$$

In these models, Pollutant_i is the primary exposure (PM_{2.5} or CO). Separate models for PM_{2.5} and CO were run due to the high to moderate correlations between them. In the HAPIN trial, exposure and biomarker measurements were collected concurrently at multiple timepoints for each participant. For the current ancillary study, we created a time-matched exposure metric by averaging all exposure measurements collected prior to the first seropositive DBS sample. For participants who remained seronegative, exposures were averaged across the entire follow-up period. Thus, each participant contributed a single COVID-19 outcome paired with a single measure of cumulative exposure, excluding any relevant post-infection exposure.

Covariate selection was guided by a directed acyclic graph (DAG) (Fig. S3) to identify minimally sufficient set of factors for adjustment, aiming to address potential confounders in the relationship between exposure and outcomes, and to ensure the chosen covariates did not function as colliders or mediators in the model. A 5% change-in-estimate approach, as described by Greenland, was also used to evaluate and finalize which covariates were included in the model (Greenland, 1989). Final primary models with all participants (mothers, children, and other adult women) included number of people in household, participant type, and study site as covariates.

Both models adjusted for the same covariates from the ITT models.

A multiple imputation model was used to reduce missing child exposure measurements for PM_{2.5} and CO. The model used predictive mean matching (PMM) implemented via chained equations using the mice package in R and the same predictors as the single imputation model. PMM selects observed exposure values from participants with similar covariate-predicted values, ensuring that imputed values remain plausible and within the original distribution. The multiple imputation procedure generated five distinct datasets, which were individually used in regression modeling. The results were pooled using Rubin's rules to account for both within- and between-imputation variability. To assess

Table 1
Baseline population characteristics, HAPIN, 2020–2021, Guatemala, Peru, India.

Variable	Control (N = 1109)	Intervention (N = 1104)
<i>Household characteristics</i>		
Household size (# people), mean (SD)	4.6 ± 2.1	4.6 ± 2.1
<i>Maternal characteristics</i>		
N (at baseline)	348	323
Age at baseline, y, mean (SD)	24.6 ± 4.2	24.4 ± 4.0
BMI, kg/m ² , mean (SD) [range]	21.6 ± 3.8	22.0 ± 4.1
<i>Mother's highest level of education</i>		
No formal education or primary school incomplete	139 (40%)	113 (35%)
Primary school complete or secondary school incomplete	118 (34%)	104 (32%)
Secondary school complete or vocational or some college or university	91 (26%)	106 (33%)
<i>Diet diversity, N (%)</i>		
Low	248 (71.3%)	230 (71.2%)
Medium	88 (25.3%)	77 (23.8%)
High	12 (3.4%)	16 (5.0%)
<i>Other Adult Woman characteristics</i>		
N (at baseline)	76	93
Age at baseline, y, mean (SD) [range]	52.4 ± 6.8	51.7 ± 8
BMI, kg/m ² , mean (SD) [range]	25.8 ± 4.8	26.1 ± 4.9
<i>OAW's highest level of education, N (%)</i>		
No formal education or primary school incomplete	60 (82.2%)	74 (82.2%)
Primary school complete or secondary school incomplete	13 (17.8%)	15 (16.7%)
Secondary school complete or vocational or some college or university	0 (0.0%)	1 (1.1%)
<i>Diet diversity, N (%)</i>		
Low	44 (57.9%)	60 (64.5%)
Medium	22 (28.9%)	25 (26.9%)
High	10 (13.2%)	8 (8.6%)

imputation performance, predicted values from both the single and multiple imputation models were compared against observed child exposure values.

Table 2
COVID-19 prevalence. HAPIN, 2020–2021, Guatemala, Peru, India.

Study site	Participant	N	SARS-CoV-2 infection prevalence		Antibody Level (u/ml) (Mean ± SD)
			Observed (95% CI)*	Adjusted (95% CI)*	
Guatemala	Child	537	7.8 (5.7, 10.4)	8.5 (6.2, 11.3)	2.4 ± 2.4
	Mother	312	7.4 (4.7, 10.9)	8.0 (5.1, 11.8)	2.3 ± 2.6
	OAW	59	6.8 (1.9, 16.5)	7.4 (2.0, 17.9)	1.8 ± 2.2
	Overall	908	7.6 (6.0, 9.5)	8.3 (6.5, 10.3)	2.4 ± 2.5
India	Child	334	0.0 (0.0, 1.1)	0.0 (0.0, 1.2)	0.1 ± 0.2
	Mother	334	2.7 (1.2, 5.1)	2.9 (1.3, 5.5)	1.2 ± 1.3
	OAW	51	68.6 (54.1, 80.9)	74.6 (58.8, 87.9)	6.4 ± 2.7
	Overall	719	6.1 (4.5, 8.1)	6.7 (4.9, 8.8)	1.1 ± 1.9
Peru	Child	502	12.2 (9.4, 15.3)	13.2 (10.2, 16.7)	2.6 ± 2.9
	Mother	25	28.0 (12.1, 49.4)	30.4 (13.1, 53.7)	3.5 ± 4.1
	OAW	59	8.5 (2.8, 18.7)	9.2 (3.1, 20.3)	2.3 ± 2.4
	Overall	586	12.5 (9.9, 15.4)	13.5 (10.8, 16.7)	2.7 ± 2.9

* Observed prevalence reflects the proportion of participants with detectable SARS-CoV-2 antibodies based on DBS testing. Adjusted prevalence was calculated using the Rogan–Gladen estimator to account for assay sensitivity (92%) and specificity (100%); confidence interval bounds were obtained by applying the same correction to the lower and upper limits of the observed 95% confidence intervals.

Subgroup analyses by study site (Guatemala, India, and Peru) and participant type (mother, older adult woman, and child) were conducted as sensitivity analyses.

All analyses were conducted in R (R Foundation for Statistical Computing).

3. Results

3.1. Participant characteristics

As of March 15, 2020, a total of 2397 households were actively enrolled in HAPIN across the Guatemala, Peru, and India study sites. Of this population, 2213 participants had valid SARS-CoV-2 serology measurements and were included in the current analytical population, including 671 mothers, 169 other adult women, and 1373 children. Baseline characteristics of the households and participants were generally similar by study arm (Table 1).

3.2. SARS-CoV-2 infection prevalence

Across study sites, the overall serology-based SARS-CoV-2 infection prevalence was 12.5% (CI: 9.9, 14.4) in Peru, followed by 7.6% (CI: 6.0, 9.55) in Guatemala and 6.1% (CI: 4.5, 8.1) in India (Table 2). Within Peru, mothers had the highest observed infection prevalence among participant types, while children and older adult women participants had lower but comparable prevalence. In Guatemala, infection prevalence was similar across participant types, with relatively consistent patterns observed for children, mothers, and older adult women. In contrast, India showed notable variation, with the non-pregnant older adult women exhibiting the highest observed infection prevalence, while mothers and children had markedly lower prevalence.

To contextualize the potential influence of vaccination on antibody measurements, we summarized the timing of DBS collection relative to national COVID-19 vaccine rollout dates (Table S1 and Fig. S4). Across study sites, 96% of all DBS samples used for COVID-19 determination were collected before participants were eligible to receive a COVID-19 vaccine based on their country- and age-specific rollout dates. This includes 86 and 92% of mothers and 92 and 98% of older adult women in Guatemala and Peru, respectively, and 92% of mothers and 78% of older adult women in India. All children's DBS were collected before vaccine eligibility. These data do not reflect whether participants had been vaccinated or which vaccine platform was used (e.g., whole-virus vs. non-whole virus).

Table 3

Mean exposure to PM_{2.5} (µg/m³) and CO (ppm) across all visits by control (CTL) and intervention (INT) arms. Household Air Pollution Intervention Network, 2020–2021, Guatemala, Peru, India.

Participant	Overall		CTL		INT	
PM _{2.5}	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD
Overall	2103	74.6 ± 87.4	1050	112.8 ± 102.6	988	34.6 ± 38.6
Child	1313	72.7 ± 96.2	652	113.8 ± 115.6	661	32.1 ± 43.7
Mother	630	78.2 ± 69.3	347	110.7 ± 77.2	283	38.4 ± 22.2
OAW	160	75.8 ± 74.2	76	114.0 ± 85.6	84	41.1 ± 37.2
CO	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD
Overall	2038	1.8 ± 3.0	1050	2.5 ± 3.2	988	1.1 ± 2.7
Child	1244	2.0 ± 3.5	627	2.8 ± 3.6	617	1.3 ± 3.2
Mother	647	1.3 ± 2.0	348	2.0 ± 2.5	299	0.6 ± 0.7
OAW	147	1.9 ± 2.5	75	2.1 ± 1.9	72	1.7 ± 2.9

3.3. Personal exposure measurements

The correlation between observed and imputed child PM_{2.5} and CO exposures using the mixed-effects model was 0.92 and 0.85, respectively. Following imputation, 110 (5%) of the 2213 participants with valid SARS-CoV-2 serology measurements were missing 24-hour personal PM_{2.5} exposure measurements, including 60 (4%) children, 41 (6%) mothers, and 9 (5%) other adult women. For CO exposure measurements, 175 (8%) participants were missing data, including 129 (9%) children, 24 (4%) mothers, and 22 (13%) other adult women.

Table 3 summarizes 24-hour average PM_{2.5} and CO exposure across the entire study period, including baseline, by arm and participant type. Mean PM_{2.5} exposure among all participants was 74.6 ± 87.4 µg/m³. PM_{2.5} exposure was higher in the control arm (112.8 ± 102.6 µg/m³) compared to the intervention arm (34.6 ± 38.6 µg/m³), and this trend was consistent across all participant types, including children, mothers, and other adult women. Mean CO exposure among all participants was 1.8 ± 3.0 ppm. Similar to PM_{2.5}, CO exposure was higher in the control arm (2.5 ± 3.2 ppm) compared to the intervention arm (1.1 ± 2.7 ppm), with this pattern observed consistently across all participant types. Detailed exposure results for the full HAPIN sample are reported elsewhere (Johnson et al., 2022).

3.4. Intention-to-treat analysis

Table 4 summarizes the ITT analyses comparing SARS-CoV-2 infection status and antibody levels between intervention and control groups. Infection prevalence was 7.7% in the control group and 9.1% in the

Table 4

Results of exploratory intention-to-treat analyses testing differences in SARS-CoV-2 infection status and antibody levels between intervention and control arms (GEE). HAPIN, 2020–2021, Guatemala, Peru, India.

Outcome	Estimate (95% CI)	p-value
SARS-CoV-2 infection (yes/no)	OR: 1.10 (0.78, 1.54)	0.59
Antibody level (u/ml)	Relative Mean: 1.03 (0.92, 1.15)	0.58

Based on 2213 participants with valid SARS-CoV-2 serology measurements. Models were fit using Generalized Estimating Equations (GEE) to account for participant clustering within households (exchangeable correlation structure). The seropositivity model used a binomial family with a logit link to estimate odds ratios, while the antibody level model used a Gaussian family with a log link to estimate relative mean differences. Both models controlled for study site (Guatemala, Peru, India) and participant type (mother, other adult woman, or child). Odds ratios are reported for participant-level SARS-CoV-2 infection status (yes/seropositive vs no/seronegative), and relative means are reported for antibody levels.

Table 5

Results of exposure–response analyses for the association between COVID-19 seropositivity and antibody concentration and exposure (PM_{2.5} and CO) in HAPIN, 2020–2021, Guatemala, Peru, India.

COVID-19 outcome	PM _{2.5} (ug/m ³)			CO (ppm)		
Infection (0/1)	Estimate	95% CI	p-value	Estimate	95% CI	p-value
Log linear	−0.15	(−0.40, 0.09)	0.21	−0.04	(−0.15, 0.07)	0.44
Quartile (ref. quartile 1)						
Quartile 2	−0.08	(−0.65, 0.48)	0.77	−0.24	(0.81, 0.34)	0.42
Quartile 3	−0.43	(−1.04, 0.17)	0.16	−0.19	(−0.68, 0.30)	0.45
Quartile 4	−0.41	(−0.95, 0.12)	0.13	−0.29	(−0.87, 0.31)	0.31
Antibody level (u/ml)	Estimate	95% CI	p-value	Estimate	95% CI	p-value
Log linear	−0.06	(−0.13, 0.01)	0.11	−0.01	(−0.05, 0.03)	0.60
Quartile (ref. quartile 1)						
Quartile 2	−0.03	(−0.23, 0.18)	0.78	−0.04	(−0.22, 0.13)	0.59
Quartile 3	−0.10	(−0.30, 0.10)	0.33	−0.07	(−0.23, 0.10)	0.43
Quartile 4	−0.17	(−0.35, −0.001)	0.049	−0.04	(−0.21, 0.15)	0.75

Exposure-response analyses assessing the association between air pollution exposure (PM_{2.5} and CO) and two outcomes: COVID-19 seropositivity (binary outcome, yes/no) and antibody concentration (continuous outcome). Generalized Estimating Equation (GEE) models with an exchangeable correlation matrix were used to account for clustering by household. Models for infection status (binomial family with logit link) and antibody concentration (Gaussian family with log link) produced coefficients on the log scale. Positive coefficients indicate that higher exposure is associated with increased odds of infection or higher antibody levels, respectively, and vice versa, in either log-linear models or relative to the first quartile in categorical models. Models were adjusted for participant type (mother, other adult woman, or child), number of people in household, and study site (Guatemala, Peru, or India). A total of 2103 and 2038 complete case observations were included in the PM_{2.5} and CO models, respectively.

intervention group. However, the ITT analysis did not detect a significant difference in the odds of infection between groups (OR = 1.1, 95% CI: 0.78, 1.54). The mean antibody level was 2.0 ± 2.4 u/ml in the control group and 2.1 ± 2.6 u/ml in the intervention group, with no significant difference in mean antibody levels between groups (relative mean: 1.03; 95% CI: 0.92, 1.15). Sensitivity analyses by study site and participant type demonstrated similar trends, with no statistically significant findings (Table S2).

3.5. Exposure-response analysis

In both log-linear and quartile-based analyses, associations between PM_{2.5} exposure and SARS-CoV-2 infection were consistently inverse, suggesting lower odds of infection with higher exposures; however, confidence intervals were wide and uniformly included the null (Table 5).

For antibody levels, log-linear models suggested an inverse association with PM_{2.5} exposure, although confidence intervals included the null (β = −0.06; 95% CI: −0.13, 0.01). In quartile analyses, antibody concentrations were progressively lower in higher PM_{2.5} exposure quartiles relative to the lowest quartile, with confidence intervals excluding the null only for the highest exposure quartile (Q4 vs Q1: β = −0.17; 95% CI: −0.35, −0.001). Consistent with this pattern, a trend

analysis across PM_{2.5} quartiles was also significant ($\beta = -0.058$; 95% CI: $-0.114, -0.003$).

For CO exposure, associations with SARS-CoV-2 infection and antibody levels were also directionally inverse across log-linear and quartile-based analyses; however, effect sizes were small, confidence intervals uniformly included the null, and there was no evidence of a monotonic or trend-based exposure–response relationship.

The direction of association in subgroup analyses by study site for both COVID-19 infection (Table S3) and antibody levels (Table S4) was generally consistent with overall findings.

4. Discussion

In this ancillary analysis of HAPIN trial data, the LPG stove and fuel intervention was not associated with differences in SARS-CoV-2 infection status or antibody levels during the study period. Exposure-response analyses similarly did not provide clear evidence of an association between personal PM_{2.5} exposure and infection status. For antibody concentrations, estimates were generally directionally inverse, and participants in the highest PM_{2.5} exposure quartile had significantly lower antibody levels compared with those in the lowest quartile. These findings were contrary to our original hypothesis, which was informed by extensive evidence linking ambient air pollution to increased COVID-19 risk, severity, and mortality. However, unlike prior studies, this analysis was the first to leverage longitudinal personal exposure monitoring and SARS-CoV-2-specific antibody measurements collected within a randomized trial, offering new insight into how household air pollution may influence COVID-19 infection and immune response in high-exposure settings.

These exploratory findings are consistent with prior reports from HAPIN showing no intervention effects on pneumonia (McCollum et al., 2024), and suggest that reductions in household air pollution from the intervention did not translate into detectable differences in serology-based SARS-CoV-2 outcomes in these high-exposure settings. The lack of significant differences in COVID-19 outcomes between intervention and control groups may reflect the timing and duration of data collection within the trial. Although implemented before the pandemic, the intervention substantially reduced PM_{2.5} and CO exposure, the pandemic's rapidly evolving nature and community-level transmission dynamics may have limited our ability to detect potential impacts of exposure reduction on infection or antibody response. Additionally, COVID-19 outcomes were not prespecified in the parent trial, and the short follow-up period may not have been sufficient to capture long-term impacts of exposure reductions on infection or immune response.

Our findings differ from most studies investigating the relationship between air pollution and COVID-19. Most have focused on ambient air pollution, relying on area-level exposure estimates that may not accurately reflect individual exposures and are subject to exposure misclassification and residual confounding. In addition, the outcomes in these studies have been restricted to COVID-19 cases, severity (e.g., hospitalizations), and mortality, rather than serological measures of infection (Kogevinas et al., 2021), as was used in this study. A systematic review of 116 observational studies in Europe and North America found that higher ambient PM_{2.5} exposure was associated with increased COVID-19 spread, hospitalizations, and mortality in more than half of the studies (Hernandez Carballo et al., 2022). Similarly, a meta-analysis of 35 studies from high- and low- and middle-income countries reported that long-term PM_{2.5} exposure was associated with increased COVID-19 incidence (1.06 per 1 $\mu\text{g}/\text{m}^3$, 95% CI: 1.04–1.07) and mortality (1.05 per 1 $\mu\text{g}/\text{m}^3$, 95% CI: 1.03–1.07) (Zang et al., 2022, p. 35). However, these associations were derived from studies using ambient air pollution exposure, which differs substantially from the predominantly indoor exposures assessed in the present study.

To date, only one study, to our knowledge, has examined serological markers of COVID-19 infection in relation to air pollution exposure. In the population-based cohort study of 9,605 adults in Catalonia, Spain,

no statistically significant association between PM_{2.5} exposure and SARS-CoV-2 infection was found (aRR: 1.04, 95% CI: 0.94, 1.14) (Kogevinas et al., 2021). However, among those infected, higher PM_{2.5} exposure was associated with increased IgG antibody levels, suggesting that air pollution may influence the immune response to infection. These findings align with mechanistic studies suggesting that air pollution can modulate immune function, potentially through pathways such as oxidative stress, inflammation, and altered expression of angiotensin-converting enzyme-2 (ACE-2), a key receptor for SARS-CoV-2 (Lai et al., 2021).

In contrast, the directionally inverse associations between PM_{2.5} exposure and antibody levels in our study warrant cautious interpretation. One potential explanation is immune dysregulation associated with long-term household air pollution exposure. Chronic exposure to air pollution has been linked to impaired immune responses, which could weaken antibody production after infection (Kogevinas et al., 2023). Under this scenario, reduced humoral response in the more highly exposed could lead to infections that result in lower peak antibody titers that fall below the assay detection threshold. Such biologically-mediated misclassification could bias exposure–response estimates toward the null or even produce an apparent inverse association, making harmful exposures appear protective. Additionally, long-term PM_{2.5} exposure may still contribute to more severe illness, as suggested in previous literature, despite lower observed antibody levels.

Another consideration is the timing of infection relative to antibody measurement. Since we restricted our sample to DBS collected on or after March 15, 2020—a conservative estimate of local epidemic onset—participants infected early in epidemic waves may have had several weeks or months between infection and their first available sample. If higher PM_{2.5} exposure was associated with increased susceptibility, more highly exposed participants may have been infected earlier, allowing additional time for antibody levels to decline below the assay detection threshold before the first DBS collection in the follow-up window. This could lead to under-ascertainment of prior infection among more highly exposed participants and bias exposure–response estimates toward the null or even produce a spurious inverse association.

Another possible explanation is that heavy exposure to indoor smoke may influence the physical environment in ways that affect viral transmission. Dense particulate smoke could alter airflow patterns or aerosol dynamics, including droplet aggregation or deposition. This could potentially cause respiratory droplets to settle more quickly, reducing the likelihood of viral inhalation. While speculative, these mechanisms highlight the need for further research to disentangle the complex interactions between indoor air pollution, immune response, and infection dynamics.

Differences in serology-based infection status across participant types and study sites should be interpreted cautiously. This is particularly true in India where infection prevalence for the other adult women was unexpectedly much higher than for the mothers and infants. These patterns may reflect unmeasured differences in mobility, social interaction, or household roles that influence viral exposure. For example, older adult women may have had greater responsibility for activities outside the home, such as shopping or community engagement, leading to higher exposure risk. In some households, older adult women also did not fully co-reside with mother–child dyads, potentially resulting in unique exposure environments. Age-related differences in immune response may also contribute. Children and some mothers frequently had low but detectable antibody levels that did not meet the seropositivity threshold, suggesting that variation in antibody magnitude could influence classification. Importantly, even in households without older adult women, children generally had lower antibody levels than mothers, indicating that age-related differences in immune response or exposure patterns may partially explain observed within-household differences.

A key strength of this COVID-19 analysis is that it leveraged data

collection within an RCT. Even though COVID-19 outcomes were not pre-specified in the original HAPIN trial and only a subset of participants were used in the current analysis, the HAPIN randomization was conducted before the COVID-19 epidemic. This provided a strong basis for an exploratory ITT, which helps reduce bias and enhances the validity of causal inferences about the effects of exposure on COVID-19 outcomes compared to observational studies. Additionally, the household-level exposure contrasts were largely generated by the randomized intervention (gas vs. biomass stove), which would likely reduce confounding in the exposure–response analyses compared to observational studies where stove or fuel type would be self-selected.

Another strength was the use of directly measured personal exposures (PM_{2.5} and CO) in exposure–response analyses, which minimizes exposure misclassification compared with area-level estimates. Although the exposure–response analyses were observational, the household-level exposure contrasts were largely generated by the randomized intervention (gas vs. biomass stove), reducing uncontrolled confounding relative to studies where stove or fuel type is self-selected. Finally, the use of serology from DBS provided an objective marker of prior infection, avoiding reliance on self-reported testing or symptoms. Infection status was also ascertained from DBS samples collected during trial follow-up rather than symptom-triggered clinical testing, which helps limit selection related to healthcare-seeking or testing access.

However, several limitations should be considered. As noted, the current analysis included only participants with valid DBS samples from three of the four HAPIN countries. This led to small sample sizes for certain subgroups, particularly the older adult women. Additionally, the restriction to an analytic subset can lessen the protection normally afforded by randomization and introduce some residual confounding. While selection bias is possible, it is less likely given that the analytic sample was determined by factors such as DBS availability, timing, and operational constraints, which are not likely related to participants' COVID-19 exposure or susceptibility. The relatively short follow-up may have limited the ability to observe longer-term intervention effects as the pandemic evolved. In addition, several important contextual factors were unmeasured and unaccounted for in analyses, including community-level COVID-19 prevalence or behavioral patterns that would lead to increased viral exposure (e.g., time spent outside the home).

The serological assessment of COVID-19 infection may also introduce uncertainty. Antibody responses vary across individuals and may wane over time, reducing sensitivity for detecting very recent or remote infections. The documented availability of inactivated whole-virus vaccines could have introduced the possibility that some anti-nucleocapsid positivity was due to vaccination rather than natural infection; however, misclassification was likely minimal given that 96% of DBS samples used in infection classification were collected before the first day of vaccine eligibility, and immediate uptake was unlikely given documented delays in the predominantly rural and resource-limited study settings. The nucleocapsid-targeting assay further reduced potential misclassification from vaccination in those that received non-whole-virus vaccines. However, some misclassification remains possible, particularly for participants with low antibody titers or waning immunity.

Serologic analyses were conducted in two laboratories, which could introduce inter-laboratory variability. However, both laboratories operated within the HAPIN Biomarker Center infrastructure and used harmonized protocols, the same commercial ELISA platform, identical manufacturer procedures, duplicate testing, and concurrent assay controls. Laboratory assignment was determined by study site rather than individual participant characteristics or PM_{2.5} exposure levels, and study site was included as a covariate in all primary models. Thus, any systematic laboratory differences would largely align with site and are partially addressed analytically. Any remaining measurement variability would most likely be nondifferential with respect to exposure and bias estimates toward the null.

Finally, interpretation of infection status based on serology is constrained by the absence of a universally applicable antibody threshold for SARS-CoV-2. Because this ancillary analysis relied on DBS samples not originally collected for COVID-19 serology, our objective was to identify detectable SARS-CoV-2 antibodies rather than to establish or validate a diagnostic threshold linked to clinical outcomes. Seropositivity was defined using an analytically derived limit of detection (5 IU/mL), validated in a pre-study assessment of paired serum and DBS samples to distinguish measurable assay signal from background. This validation established analytical performance only and did not define a clinical threshold for infection or protective immunity. The ELISA exhibited high specificity (100%) but imperfect sensitivity for DBS specimens (92% relative to matched serum); therefore, it is possible that some infections with low or waning antibody levels may have been missed. To reduce dependence on a single threshold and minimize information loss, antibody concentrations were also analyzed as a continuous measure.

Overall, while these limitations temper interpretation, the combination of a design with randomized exposures, objective serology, and direct personal exposure measurement strengthens causal inference relative to fully observational studies.

4.1. Conclusion

In this ancillary analysis of HAPIN trial data from India, Guatemala, and Peru, the LPG stove and fuel intervention was not associated with differences in SARS-CoV-2 infection status or antibody levels. Exploratory exposure–response analyses suggested a potential inverse association between personal PM_{2.5} exposure and SARS-CoV-2 antibody response, highlighting potentially complex relationships between household air pollution and immune function that warrant further investigation.

5. HAPIN investigators

Thomas F. Clasen, Jennifer L. Peel, William Checkley, Vigneshwari Aravindalochanan, Kalpana Balakrishnan, Glorioso Bankundiye, Dana Boyd Barr, Vanessa Burrowes, Alejandra Bussalleu, Devan Campbell, Eduardo Canuz, Adly Castañaza, Howard H. Chang, Yunyun Chen, Marilú Chiang, Maggie L. Clark, Carmen Lucia Contreras, Rachel Craik, Victor G. Davila-Roman, Lisa de las Fuentes, Oscar De León, Anaité Diaz-Artiga, Priya D'Souza, Ephrem Dusabimana, Lisa Elon, Juan Gabriel Espinoza, Irma Sayury Pineda Fuentes, Sarada S. Garg, Ahana Ghosh, Dina Goodman-Palmer, Savannah Gupton, Sarah Hamid, Stella M. Hartinger, Steven A. Harvey, Mayari Hengstermann, Ian Hennessee, Phabiola Herrera, Shakir Hossen, Marjorie Howard, Penelope P. Howards, Shirin Jabbarzadeh, Michael A. Johnson, Katherine Kearns, Miles A. Kirby, Jacob Kremer, Margaret A. Laws, Grace Lee, Pattie Lenzen, Jiawen Liao, Amy E. Lovvorn, Jane Mbabazi, Eric D. McCollum, John P. McCracken, Julia N. McPeck, Rachel Meyers, J. Jaime Miranda, Erick Mollinedo, Libny Monroy, Alexie Mukeshimana, Krishnendu Mukhopadhyay, Moses Mutabazi, Bernard Mutariyani, Luke P. Naeher, Durairaj Natesan, Florian Ndagijimana, Laura Nicolaou, Azhar Nizam, Jean de Dieu Ntivuguruzwa, Parinya Panuwet, Aris T. Papageorgiou, Ricardo Piedrahita, Ajay Pillarisetti, Naveen Puttaswamy, Elisa Puzolo, Karthikeyan Dharmapuri Rajamani, Sarah Rajkumar, Usha Ramakrishnan, Rengaraj Ramasami, Alexander Ramirez, Ghislaine Rosa, Joshua Rosenthal, P. Barry Ryan, Sudhakar Saidam, Sankar Sambandam, Suzanne Simkovich, Sheela S. Sinharoy, Kirk R. Smith, Kyle Steenland, Damien Swearing, Gurusamy Thangavel, Lisa M. Thompson, Ashley Toenjes, Lindsay J. Underhill, Viviane Valdes, Amit Verma, Lance A. Waller, Jiantong Wang, Megan Warnock, Kendra N. Williams, Wenlu Ye, Bonnie N. Young, Ashley Younger.

CRediT authorship contribution statement

Lindsay J. Underhill: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Maggie L. Clark:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Dana Boyd Barr:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Kalpna Balakrishnan:** Writing – review & editing. **Naveen Puttaswamy:** Writing – review & editing, Investigation, Formal analysis. **Ajay Pillariseti:** Writing – review & editing, Data curation. **Anaite Diaz-Artiga:** Writing – review & editing. **Lisa M. Thompson:** Writing – review & editing. **John P. McCracken:** Writing – review & editing. **Victor G. Davila-Roman:** Writing – review & editing. **Lance A. Waller:** Writing – review & editing, Data curation. **Bonnie N. Young:** Writing – review & editing. **Michael A. Johnson:** Writing – review & editing, Data curation. **Karthikeyan Rajamani:** Writing – review & editing, Investigation. **William Checkley:** Writing – review & editing. **Thomas F. Clasen:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Jennifer L. Peel:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Kyle Steenland:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Funding

The HAPIN trial was funded by the U.S. National Institutes of Health (cooperative agreement 1UM1HL134590) in collaboration with the Bill & Melinda Gates Foundation [OPP1131279].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The investigators would like to thank the members of the advisory committee – Drs. Patrick Breyse, Donna Spiegelman, and Joel Kaufman – for their valuable insight and guidance throughout the implementation of the trial. We also wish to acknowledge all research staff and study participants for their dedication to and participation in this important trial.

A multidisciplinary, independent Data and Safety Monitoring Board (DSMB) appointed by the National Heart, Lung, and Blood Institute (NHLBI) monitored the quality of the data and protected the safety of patients enrolled in the HAPIN trial. The DSMB consisted of: Catherine Karr (Chair), Nancy R. Cook, Stephen Hecht, Joseph Mllum, Nalini Sathiakumar (deceased), Paul K. Whelton, and Gail Weinmann and Thomas Croxton (Executive Secretaries). Program Coordination: Gail Rodgers, Bill & Melinda Gates Foundation; Claudia L. Thompson, National Institute of Environmental Health Sciences; Mark J. Parascandola, National Cancer Institute; Marion Koso-Thomas, Eunice Kennedy Shriver National Institute of Child Health and Human Development; Joshua P. Rosenthal, Fogarty International Center; Concepcion R. Nierras, NIH Office of Strategic Coordination – The Common Fund; Katherine Kavounis, Dong-Yun Kim, Barry S. Schmetter (deceased), and Antonello Punturieri, NHLBI.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2026.110180>.

Data availability

Data will be made available on request.

References

- Barr, D.B., Puttaswamy, N., Jaacks, L.M., Steenland, K., Rajkumar, S., Gupton, S., Ryan, P.B., Balakrishnan, K., Peel, J.L., Checkley, W., Clasen, T., Clark, M.L., (HAPIN Investigative Team). Design and Rationale of the Biomarker Center of the Household Air Pollution Intervention Network (HAPIN). *Environ Health Perspect* Apr;128 (4):47010.
- Barr, Dana Boyd, Puttaswamy, N., Jaacks, L.M., Steenland, K., Rajkumar, S., Gupton, S., Ryan, P.B., Balakrishnan, K., Peel, J.L., Checkley, W., Clasen, T., Clark, M.L., (HAPIN Investigative Team), 2020b. Design and rationale of the biomarker center of the household air pollution intervention network (HAPIN) trial. *Environ. Health Perspect.* 128, 47010. <https://doi.org/10.1289/EHP5751>.
- Buck Louis, G.M., Barr, D.B., Kannan, K., Chen, Z., Kim, S., Sundaram, R., 2016. Paternal exposures to environmental chemicals and time-to-pregnancy: overview of results from the LIFE study. *Andrology* 4, 639–647. <https://doi.org/10.1111/andr.12171>.
- Chen, C., Wang, J., Kwong, J., Kim, J., van Donkelaar, A., Martin, R.V., Hystad, P., Su, Y., Lavigne, E., Kirby-McGregor, M., Kaufman, J.S., Benmarhnia, T., Chen, H., 2022. Association between long-term exposure to ambient air pollution and COVID-19 severity: a prospective cohort study. *CMAJ* 194, E693–E700. <https://doi.org/10.1503/cmaj.220068>.
- Clasen, T., Checkley, W., Peel, J.L., Balakrishnan, K., McCracken, J.P., Rosa, G., Thompson, L.M., Barr, D.B., Clark, M.L., Johnson, M.A., Waller, L.A., Jaacks, L.M., Steenland, K., Miranda, J.J., Chang, H.H., Kim, D.-Y., McCollum, E.D., Davila-Roman, V.G., Papageorgiou, A., Rosenthal, J.P., Investigators, H.A.P.I.N., 2020. Design and rationale of the HAPIN Study: a multicountry randomized controlled trial to assess the effect of liquefied petroleum gas stove and continuous fuel distribution. *Environ. Health Perspect.* 128, 47008. <https://doi.org/10.1289/EHP6407>.
- Clopper, C.J., Pearson, E.S., 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 26, 404–413. <https://doi.org/10.1093/biomet/26.4.404>.
- English, P.B., Von Behren, J., Balmes, J.R., Boscardin, J., Carpenter, C., Goldberg, D.E., Horiuchi, S., Richardson, M., Solomon, G., Valle, J., Reynolds, P., 2022. Association between long-term exposure to particulate air pollution with SARS-CoV-2 infections and COVID-19 deaths in California. *U.S.A. Environ. Adv.*, 100270 <https://doi.org/10.1016/j.envadv.2022.100270>.
- Greenland, S., 1989. Modeling and variable selection in epidemiologic analysis. *Am. J. Public Health* 79, 340–349. <https://doi.org/10.2105/ajph.79.3.340>.
- Hernandez Carballo, I., Bakola, M., Stuckler, D., 2022. The impact of air pollution on COVID-19 incidence, severity, and mortality: a systematic review of studies in Europe and North America. *Environ. Res.* 215, 114155. <https://doi.org/10.1016/j.envres.2022.114155>.
- Johnson, M., Pillariseti, A., Piedrahita, R., Balakrishnan, K., Peel, J.L., Steenland, K., Underhill, L.J., Rosa, G., Kirby, M.A., Diaz-Artiga, A., McCracken, J., Clark, M.L., Waller, L., Chang, H.H., Wang, J., Dusbimana, E., Ndagijimana, F., Sambandam, S., Mukhopadhyay, K., Kearns, K.A., Campbell, D., Kremer, J., Rosenthal, J.P., Checkley, W., Clasen, T., Naeher, L., the Household Air Pollution Intervention Network (HAPIN) Trial Investigators, 2022. Exposure contrasts of pregnant women during the household air pollution intervention network randomized controlled trial. *Environ. Health Perspect.* 130, 97005. <https://doi.org/10.1289/EHP10295>.
- Johnson, M.A., Steenland, K., Piedrahita, R., Clark, M.L., Pillariseti, A., Balakrishnan, K., Peel, J.L., Naeher, L.P., Liao, J., Wilson, D., Sarnat, J., Underhill, L.J., Burrows, V., McCracken, J.P., Rosa, G., Rosenthal, J., Sambandam, S., de Leon, O., Kirby, M.A., Kearns, K., Checkley, W., Clasen, T., Investigators, H.A.P.I.N., 2020. Air pollutant exposure and stove use assessment methods for the household air pollution intervention network (HAPIN) trial. *Environ. Health Perspect.* 128, 047009. <https://doi.org/10.1289/EHP6422>.
- Kogevinas, M., Castaño-Vinyals, G., Karachaliou, M., Espinosa, A., de Cid, R., Garcia-Aymerich, J., Carreras, A., Cortés, B., Pleguezuelos, V., Jiménez, A., Vidal, M., O'Callaghan-Gordo, C., Cirach, M., Santano, R., Barrios, D., Puyol, L., Rubio, R., Izquierdo, L., Nieuwenhuijsen, M., Dadvand, P., Aguilár, R., Moncunill, G., Dobaño, C., Tonne, C., 2021. Ambient air pollution in relation to SARS-CoV-2 infection, antibody response, and COVID-19 disease: a cohort study in Catalonia, Spain (COVICAT Study). *Environ. Health Perspect.* 129, 117003. <https://doi.org/10.1289/EHP9726>.
- Kogevinas, M., Karachaliou, M., Espinosa, A., Aguilár, R., Castaño-Vinyals, G., Garcia-Aymerich, J., Carreras, A., Cortés, B., Pleguezuelos, V., Papantoniou, K., Rubio, R., Jiménez, A., Vidal, M., Serra, P., Parras, D., Santamaría, P., Izquierdo, L., Cirach, M., Nieuwenhuijsen, M., Dadvand, P., Straif, K., Moncunill, G., de Cid, R., Dobaño, C., Tonne, C., 2023. Long-term exposure to air pollution and covid-19 vaccine antibody response in a general population cohort (COVICAT Study, Catalonia). *Environ. Health Perspect.* 131, 47001. <https://doi.org/10.1289/EHP11989>.
- Lai, A., Chang, M.L., O'Donnell, R.P., Zhou, C., Sumner, J.A., Hsiai, T.K., 2021. Association of COVID-19 transmission with high levels of ambient pollutants: Initiation and impact of the inflammatory response on cardiopulmonary disease. *Sci. Total Environ.* 779, 146464. <https://doi.org/10.1016/j.scitotenv.2021.146464>.
- Liao, J., McCracken, J.P., Piedrahita, R., Thompson, L., Mollinedo, E., Canuz, E., De León, O., Diaz-Artiga, A., Johnson, M., Clark, M., Pillariseti, A., Kearns, K., Naeher, L., Steenland, K., Checkley, W., Peel, J., Clasen, T.F., HAPIN investigators, 2020. The use of bluetooth low energy Beacon systems to estimate indirect personal

- exposure to household air pollution. *J. Exposure Sci. Environ. Epidemiol.* 30, 990–1000. <https://doi.org/10.1038/s41370-019-0172-z>.
- McCollum, E.D., McCracken, J.P., Kirby, M.A., Grajeda, L.M., Hossen, S., Moulton, L.H., Simkovich, S.M., Goodman-Palmer, D., Rosa, G., Mukeshimana, A., Balakrishnan, K., Thangavel, G., Garg, S.S., Castañaza, A., Thompson, L.M., Diaz-Artiga, A., Papageorgiou, A.T., Davila-Roman, V.G., Underhill, L.J., Hartinger, S.M., Williams, K.N., Nicolaou, L., Chang, H.H., Lovvorn, A.E., Rosenthal, J.P., Pillarsetti, A., Ye, W., Naeher, L.P., Johnson, M.A., Waller, L.A., Jabbarzadeh, S., Wang, J., Chen, Y., Steenland, K., Clasen, T.F., Peel, J.L., Checkley, W., 2024. Liquefied petroleum gas or biomass cooking and severe infant pneumonia. *N. Engl. J. Med.* 390, 32–43. <https://doi.org/10.1056/NEJMoa2305681>.
- R Core Team, 2025. Documentation for R package ‘stats’ version 4.6.0: binom.test: Exact Binomial Test. In: *The R Project for Statistical Computing*. Vienna.
- Rogan, W.J., Gladen, B., 1978. Estimating prevalence from the results of a screening test. *Am. J. Epidemiol.* 107, 71–76. <https://doi.org/10.1093/oxfordjournals.aje.a112510>.
- Sheppard, N., Carroll, M., Gao, C., Lane, T., 2023. Particulate matter air pollution and COVID-19 infection, severity, and mortality: a systematic review and meta-analysis. *Sci. Total Environ.* 880, 163272. <https://doi.org/10.1016/j.scitotenv.2023.163272>.
- Simkovich, S.M., Thompson, L.M., Clark, M.L., Balakrishnan, K., Bussalleu, A., Checkley, W., Clasen, T., Davila-Roman, V.G., Diaz-Artiga, A., Dusabimana, E., de Fuentes, L., Harvey, S., Kirby, M.A., Lovvorn, A., McCollum, E.D., Mollinedo, E.E., Peel, J.L., Quinn, A., Rosa, G., Underhill, L.J., Williams, K.N., Young, B.N., Rosenthal, J., HAPIN Investigators, 2021. A risk assessment tool for resumption of research activities during the COVID-19 pandemic for field trials in low resource settings. *BMC Med. Res. Method.* 21, 68. <https://doi.org/10.1186/s12874-021-01232-x>.
- Taylor, J.K., 2018. *Quality Assurance of Chemical Measurements*. Routledge, New York, 10.1201/9780203741610.
- Yu, K., Zhang, Q., Wei, Y., Chen, R., Kan, H., 2024. Global association between air pollution and COVID-19 mortality: a systematic review and meta-analysis. *Sci. Total Environ.* 906, 167542. <https://doi.org/10.1016/j.scitotenv.2023.167542>.
- Zang, S.-T., Luan, J., Li, L., Yu, H.-X., Wu, Q.-J., Chang, Q., Zhao, Y.-H., 2022. Ambient air pollution and COVID-19 risk: evidence from 35 observational studies. *Environ. Res.* 204, 112065. <https://doi.org/10.1016/j.envres.2021.112065>.