

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to [508 standards](#) due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehp508@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

Supplemental Material

Exposure Contrasts of Pregnant Women during the Household Air Pollution Intervention Network Randomized Controlled Trial

Michael Johnson, Ajay Pillarisetti, Ricardo Piedrahita, Kalpana Balakrishnan, Jennifer L. Peel, Kyle Steenland, Lindsay J. Underhill, Ghislaine Rosa, Miles A. Kirby, Anaité Díaz-Artiga, John McCracken, Maggie L. Clark, Lance Waller, Howard H. Chang, Jiantong Wang, Ephrem Dusabimana, Florian Ndagijimana, Sankar Sambandam, Krishnendu Mukhopadhyay, Katherine A. Kearns, Devan Campbell, Jacob Kremer, Joshua P. Rosenthal, William Checkley, Thomas Clasen, Luke Naeher, and the Household Air Pollution Intervention Network (HAPIN) Trial Investigators

Table of Contents

QA/QC: PM_{2.5} Sampling

QA/QC: Black Carbon

QA/QC: CO

IRC and pollutant-specific findings

Predicting missingness

Table S1. Gravimetric sample validity.

Table S2. Gravimetric ECM duplicate performance and fit metrics.

Table S3. Wearing compliance, defined as the fraction of time motion, was detected during daytime hours.

Table S4. BC duplicate performance and fit metrics.

Table S5. CO sample validity.

Table S6. CO duplicate performance and fit metrics.

Table S7. Modeling approaches.

Table S8. Baseline household and maternal characteristics in HAPIN by IRC and intervention arm.

Table S9. Additional household and exposure-related characteristics.

Table S10. Characteristics of households with and without measurements, by measurement round and arm.

Table S11. Repeated-measures model output of predictors of missing exposure data.

Table S12. Observed versus imputed PM_{2.5} data by study arm and measurement round.

Table S13. Observed and imputed model estimates of percent differences in PM_{2.5}.

Table S14. Tabular data for Figure 2 Trends in PM_{2.5} exposure.

Table S15. Adjusted p values from Dunn's test comparing pollutant values by study round.

Table S16. IRC-specific models of the impact (percent difference) of the intervention on PM_{2.5}.

Table S17. IRC-specific models of the impact (percent difference) of the intervention on BC.

Table S18. IRC-specific models of the impact (percent difference) of the intervention on CO.

Table S19. Comparison table for exposure comparisons with other studies. Exposure metrics have been filled in as available, with “-” indicating places where the metric was not presented.

Figure S1. ECM duplicate measures. Panels are individual IRCs. Points are paired measurements using collocated ECMs. The dashed line is 1:1. The blue line is a linear model fit to the data points.

Figure S2. Monitoring wearing compliance. Panels are individual IRCs. Bars are the number of measurements shown as wearing compliant for a given fraction of the day. Compliance is defined as the fraction of time motion was detected during daytime hours.

Figure S3. BC duplicate measures. Panels are individual IRCs. Points are paired measurements using collocated ECMs with filters analyzed via transmissometry for BC concentrations/exposures. The dashed line is 1:1. The blue line is a linear model fit to the data points.

Figure S4. CO duplicate measures. Panels are individual IRCs. Points are paired measurements using collocated Lascar CO monitors. The dashed line is 1:1. The blue line is a linear model fit to the data points. Axes are truncated.

Figure S5. Observed and Imputed Exposure Estimates by Study Arm. Panel A is the control arm; Panel B is the intervention arm. Imputation 0 is the observed data; imputations 1:10 are imputed data. Red lines are medians; the lower and upper box edges are the 25th and 75th percentiles, respectively; whiskers extend to 1.5 * the interquartile range. Dots represent individual data points; red are imputations, blue are observations.

Figure S6. HAPIN PM_{2.5}, Black Carbon (BC), and Carbon Monoxide (CO) exposures overall and by IRC. Red triangles and blue dots are per-country and study round samples in intervention and control households, respectively. Circles and triangles outlined in black are median values in control and intervention households, respectively. Lines are interquartile ranges. BL = baseline (9-20 weeks gestation), P1 = post-intervention visit 1 (24-28 weeks gestation), and P2 = post-intervention visit 2 (32-36 weeks gestation). The dotted line in the PM panels is the annual WHO Interim Target 1 guideline value (35 $\mu\text{g}/\text{m}^3$); the dashed line in the CO plots is the WHO guideline value of 6.11 ppm (7 mg/m^3).

Figure S7. HAPIN-wide and by country relationships between PM_{2.5} and CO by primary fuel for cooking. Both axes are Log₁₀ transformed. The solid lines are a linear model; the shaded areas are standard errors. “Traditional” panels include measurements made during baseline and during baseline and post-intervention 1 and 2 in control homes. “LPG” panels include measurements made post-intervention.

Figure S8. Country relationships between PM_{2.5} and black carbon by primary fuel for cooking. Both axes are log₁₀ transformed. The solid lines are a linear model; the shaded areas are standard errors. “Traditional” panels include measurements made during baseline and during baseline, post-intervention 1 and 2 in control homes. “LPG” panels include measurements made post-intervention.

Figure S9. Estimated impacts of the HAPIN LPG intervention on BC exposure. All linear mixed effects models had log transformed BC exposure as the dependent variable. Whiskers are 95% confidence intervals. The first panel (“Before and After”) uses data from both the control and intervention arms and compares the intervention period to the baseline period. The second panel (“Between Groups”) uses only data from the intervention period and contrasts the intervention arm with the control arm. The third panel (“Comparison of Changes”) uses all data from both study arms and both study periods; the model term of interest is the interaction between study arm and period, after controlling for each variable separately in the model. The “Overall” points consider an average post-intervention exposure; the Visit-specific points consider each post-randomization visit separately.

Figure S10. Estimated impacts of the HAPIN LPG intervention on CO exposure. All linear mixed effects models had log transformed CO exposure as the dependent variable. Whiskers are 95% confidence intervals. The first panel (“Before and After”) uses data from both the control and intervention arms and compares the intervention period to the baseline period. The second panel (“Between Groups”) uses only data from the intervention period and contrasts the intervention arm with the control arm. The third panel (“Comparison of Changes”) uses all data from both study arms and both study periods; the model term of interest is the interaction between study arm and period, after controlling for each variable separately in the model. The “Overall” points consider an average post-intervention exposure; the Visit-specific points consider each post-randomization visit separately.

Figure S11. Estimated impacts of the HAPIN LPG intervention on PM2.5 exposure for both imputed and observed data. All linear mixed effects models had log transformed PM2.5 exposure as the dependent variable. Whiskers are 95% confidence intervals. The first panel (“Before and After”) uses data from both the control and intervention arms and compares the intervention period to the baseline period. The second panel (“Between Groups”) uses only data from the intervention period and contrasts the intervention arm with the control arm. The third panel (“Comparison of Changes”) uses all data from both study arms and both study periods; the model term of interest is the interaction between study arm and period, after controlling for each variable separately in the model. The “Overall” points consider an average post-intervention exposure; the Visit-specific points consider each post-randomization visit separately. The red points (rightmost for each pair of points) are from the imputed data; the blue points (leftmost for each pair of points) are observed data.

Figure S12. IRC-specific models of the impact of the intervention on PM2.5. All linear mixed effects models had log transformed PM2.5 exposure as the dependent variable. Whiskers are 95% confidence intervals. The first panel (“Before and After”) uses data from both the control and intervention arms and compares the intervention period to the baseline period. The second panel (“Between Groups”) uses only data from the intervention period and contrasts the intervention arm with the control arm. The third panel (“Comparison of Changes”) uses all data from both study arms and both study periods; the model term of interest is the interaction between study arm and period, after controlling for each variable separately in the model. The “Overall” points consider an average post-intervention exposure; the Visit-specific points consider each post-randomization visit separately. The red points (leftmost for each pair of points) in panels B and D are from visit 1; the blue points (rightmost for each pair of points) in panels B and D are from visit 2.

Figure S13. IRC-specific models of the impact of the intervention on BC. All linear mixed effects models had log transformed BC exposure as the dependent variable. Whiskers are 95% confidence intervals. The first panel (“Before and After”) uses data from both the control and intervention arms and compares the intervention period to the baseline period. The second panel (“Between Groups”) uses only data from the intervention period and contrasts the intervention arm with the control arm. The third panel (“Comparison of Changes”) uses all data from both study arms and both study periods; the model term of interest is the interaction between study arm and period, after controlling for each variable separately in the model. The “Overall” points consider an average post-intervention exposure; the Visit-specific points consider each post-randomization visit separately. The red points in panels B and D (leftmost for each pair of points) are from visit 1; the blue points in panels B and D (rightmost for each pair of points) are from visit 2.

Figure S14. IRC-specific models of the impact of the intervention on CO. All linear mixed effects models had log transformed CO exposure as the dependent variable. Whiskers are 95% confidence intervals. The first panel (“Before and After”) uses data from both the control and intervention arms and compares the intervention period to the baseline period. The second panel (“Between Groups”) uses only data from the intervention period and contrasts the intervention arm with the control arm. The third panel (“Comparison of Changes”) uses all data from both study arms and both study periods; the model term of interest is the interaction between study arm and period, after controlling for each variable separately in the model. The “Overall” points consider an average post-intervention exposure; the Visit-specific points consider each post-randomization visit separately. The red points in panels B and D (leftmost for each pair of points) are from visit 1; the blue points in panels B and D (rightmost for each pair of points) are from visit 2.