

Online Supplemental Material for:

Title: *Acute changes in lung function following controlled exposure to cookstove air pollution in the Subclinical Tests of Volunteers Exposed to Smoke (STOVES) Study.*

Authors: Kristen M Fedak; Nicholas Good; Ethan S. Walker; John Balmes; Robert D. Brook; Maggie L. Clark; Tom Cole-Hunter; Robert Devlin; Christian L'Orange; Gary Luckasen; John Mehaffy; Rhiannon Shelton; Ander Wilson; John Volckens; Jennifer L. Peel

Contents

Study Methods	3
Participants.....	3
Eligibility Criteria.....	3
Recruitment/Screening Process.....	3
Study Design	4
Treatments: Stove Descriptions and Generation of Exposures.....	4
Additional Characterization of Stove Emissions.....	4
Assignment of Treatment Sequences	4
Health and Additional Measurements.....	5
Data Analysis.....	6
Results	6
Participants, Exposures, and Health Measurement Times	6
Missing Data	6
Additional Pollutant Characterization	7
Health Measurement Timing	7
Model Results.....	7
Potential Confounders.....	7
Alternative Models/Sensitivity Analyses	14

Study Methods

Participants

Eligibility Criteria

Eligibility criteria were:

- 18 to 35 years old at the time of recruitment;
- never smokers;
- body mass index between 19 to 28 kg/m² with body weight greater than 50 kg;
- no history of heart disease, diabetes, kidney disease, systemic sclerosis, or any chronic inflammatory disease such as asthma, arthritis, or severe allergies;
- normal non-hypertensive blood pressure, normal electrocardiogram, spirometry values greater than 70% of the predicted value for the age/gender, and normal blood test results (including no evidence of iron-deficient anemia), as determined at the screening exam;
- not currently taking statins, anti-inflammatory medication, or other medications unless cleared by the study physician during the screening exam (cleared medications: oral contraceptives, some daily anti-depression/anxiety medications);
- no use of tetrahydrocannabinol or illicit drugs within the past three months;
- no ear or abdominal/thoracic surgery in the past month; no cancer (current or in remission for less than six months); no central intravenous line or port; never had a mastectomy;
- no pacemaker;
- not currently pregnant, breastfeeding, or planning a pregnancy within six months;
- not regularly exposed to smoke, dust, fumes, or solvents (occupationally or recreationally/at home), or regularly burned candles or incense within the last three months;
- no history of claustrophobia;
- no fear of needles;
- not planning to donate blood during the timeframe of participation;
- no latex allergy; and
- live within 20 miles of the study facility and not planning to move more than 20 miles away within six months.

Recruitment/Screening Process

We recruited participants via local and university newspaper articles, various university email list advertisements, and word-of-mouth. Interested individuals submitted a screening questionnaire to determine potential eligibility; those who appeared to meet eligibility criteria were invited to an in-person screening. At the screening meeting, medical staff reviewed the eligibility criteria with the individual, conducted a physical

exam, reviewed individual and family medical history, conducted an electrocardiogram, performed a spirometry test, and took a blood sample for analysis of complete blood count, comprehensive metabolic panel, lipid levels, and serum ferritin. The study physician considered the results of the screening exam to determine final eligibility for participation in the study. Informed consent was provided for all individuals who conducted a screening exam.

Individuals interested in participating received a tour of the study facility and reviewed the study process, expectations, and requirements for participation. Potential participants reviewed a written informed consent document with study staff, which described the study procedures, risks, benefits, and compensation for participation.

Study Design

Treatments: Stove Descriptions and Generation of Exposures

Stove makes/models were as follows:

1. Liquefied petroleum gas [LPG] stove: Classic Single Burner 25000 BTU, WokSmith, China
2. Gasifier: Ace 1 Gasifier, African Clean Energy (Pty) Ltd, Lesotho
3. Forced draft (fan-powered) rocket elbow: HomeStove, Biolite, USA
4. Natural draft rocket elbow: G3300, Envirofit International, USA
5. Traditional three stone fire: open fire, bricks in U-shape used to contain fuel

Detailed descriptions of how exposures were generated and pollution levels were monitored is provided in a previous publication (1). Briefly, stoves were operated within a total-capture fume hood and diluted with HEPA-filtered air before being drawn into the exposure chamber; PM_{2.5}, carbon monoxide (CO), oxygen, temperature, and humidity in the chamber were monitored in real time and adjusted using a dynamic control system.

Additional Characterization of Stove Emissions

We conducted additional characterization of the air from each treatment type (each stove emissions) after the end of the controlled human exposure study. The methods for these tests are described in detail elsewhere (1). Briefly, we collected samples on at least two occasions per treatment type and analyzed for PM_{2.5} mass, particle number size distributions (10 to 500 nm), elemental and organic carbon concentrations, nitrogen oxide and nitrogen dioxide, gas-phase carbonyls, and volatile organic compounds.

Assignment of Treatment Sequences

Detailed descriptions of the protocols for assigning treatment orders for participants are provided in a previous publication (1); brief notes are provided herein.

Each participant was scheduled to receive six exposure treatments (one filtered air control and five different stoves), with a washout period between treatments of two to six weeks by design. We used a modified Williams square to assign sequences to our 48 recruited participants across three study rounds (October 2016 to February 2017;

March to June 2017; August 2017 to January 2018); each round contained two sequence groups of eight individuals who completed their sessions on opposite weeks. Four participants completed their session on the same day. Opportunities to make up sessions that were missed due to illness or unforeseen scheduling conflicts were provided at the end of the study round, between 10 days to 14 weeks after the end of the round. Participants were not informed of the treatment they were receiving on the study day, however, complete blinding was not achievable due to

Health and Additional Measurements

We conducted a series of health measurements that included electrocardiogram readings via Holter monitor, blood pressure and pulse wave analysis, pulse wave velocity, spirometry, and a blood draw. The four rounds of health measurements occurred at approximately the same time of day across sessions; it took approximately one hour to complete the series of measurements. Spirometry measurements were conducted towards the end of the series (before the blood draw) and started approximately 30 minutes after the start of the series.

Pulmonary function tests were performed according to American Thoracic Society/European Respiratory Society (ATS/ERS) guidelines (2). We conducted a multi-flow calibration of the spirometer daily, at flow rates between 0.5 and 12 L/s using a three-liter syringe. Each test consisted of several trials of an expiratory-only maneuver, performed from a seated position with both feet on the ground, wearing a nose clip. Within each test, we required a minimum of three acceptable trials (e.g., free from artifacts due to cough, glottis closures, or obstructive mouthpieces and with full exhale) with the two largest FEV₁ and FVC values across trials within 150 mL (2). We allowed up to eight attempts in a single test. If a participant did not meet the requirement within eight attempts, we stopped the test. The health outcomes used from the pulmonary function test were forced vital capacity (FVC; the total volume of air exhaled in a forceful, complete expiration following maximum inspiration), the forced expiratory volume in one second (FEV₁; the volume of air exhaled in the first second of an FVC maneuver), the ratio of FEV₁ to FVC (FEV₁/FVC) and mid-expiratory flow (FEF₂₅₋₇₅; the mean flow rate between the 25th and 75th percent of the FVC). For all tests that met the quality criteria (2), we chose the largest FVC and FEF₂₅₋₇₅ values from the acceptable trials within that test and the FEV₁ and FEV₁/FVC value that came from the trial with the largest FVC. For tests that did not meet the minimum quality criteria, a board-certified pulmonologist reviewed the spirographs to determine which values, if any, could be used in analyses.

We administered surveys at each health measurement time point to assess other variables, including participants' recent exposures to medications, caffeine, smoke/fumes, and alcohol, mode of commute to the facility, and sleep duration, which potentially were confounders in the study. Hourly ambient PM_{2.5} and CO data from a local county monitoring site was downloaded via the U.S. EPA's Air Quality Data API (3). Ambient temperature data was obtained from Colorado State University

Atmospheric Science Department's Christman Field Weather Station located approximately four miles from the study location (4).

Data Analysis

We calculated descriptive statistics for the various potential confounders (alcohol use, medication use, caffeine intake, exposure to smoke/fumes, mode of commute to the facility, sleep duration, and ambient PM_{2.5}, CO, and temperature). We also conducted bivariate analyses of the potential confounder by treatment type to confirm no chance associations or imbalances due to missing data.

Results

Participants, Exposures, and Health Measurement Times

Missing Data

We recruited 48 participants. One participant was removed from the study after three sessions due to consistently low lung function at baseline. The participant had met eligibility criteria regarding lung function at screening (spirometry values greater than 70% of the predicted value for the age/gender), however, on study days was consistently achieving values between 65-75% at the baseline (pre exposure) measurements. As such, the study physician determined that removal of the participant from the study was advised. The data from this participant's three completed sessions was censored from the spirometry dataset.

Of the remaining 47 participants (22 male, 25 female), 81% contributed data to all six treatments, either in sequence or using the allotted makeup sessions. Of the 26 participants who missed scheduled study sessions, 12 missed only one session. Approximately half of the missed study sessions were due to scheduling conflicts that arose after a participant enrolled in the study; one quarter were due to illnesses on scheduled study dates, and one quarter were due to the participants enrolling in the study late, after the rest of their sequence cohort had completed the first study session. We attempted to schedule makeups for all missed sessions. However, three participants withdrew from the study for personal reasons prior to completing six study sessions, resulting in nine sessions that were not completed. Additionally, an error applying the exposure protocol and our exposure chamber operation resulted in the loss of data relevant to single sessions which were not repeated for three participants.

Within the sessions completed, six individual data points were not collected due to scheduling conflicts for participants that resulted in them leaving a study day without completing the three hour or 24 hour follow up time point. An additional nine data points were collected but censored from the dataset because they did not meet minimum quality criteria as established by the ATS/ERS and were not approved for use by the study pulmonologist.

Additional Pollutant Characterization

Please refer to the supplemental material of the previously publication on this study for results of the additional pollutant characterization tests (1). Results do not suggest any alternative explanations for the observed results in lung function related to non-PM_{2.5} pollutants.

Health Measurement Timing

Baseline pre-exposure measurements occurred on average 17 minutes before entering the exposure facility (range 3 to 54 min). Immediate post exposure measurements occurred on average 38 minutes (range 33 to 62 min) after exiting the facility. The average time of the three-hour post exposure measurements was 3 hr 33 min (range 3 hr 20 min to 3 hr 50 min) after exiting the exposure facility, and the average time of the 24 hour measurements was 24 hr 22 min (range 22 hr 18 min to 25 hr 50 min) after exiting the facility. While participants' start times were staggered, we attempted to keep each individual on the same timeline during each study session to maintain consistency. We calculated the maximum difference in the health measurement timing for each person at each time point across all of their study sessions; the mean maximum difference was 11 minutes for the baseline measurements, 13 minutes for both immediate post-exposure and three hour post-exposure measurements, and 39 minutes for the 24 hour post-exposure measurements.

The mean of the range in measurement times per person across the six study sessions were 13 minutes, 13 minutes, and 39 minutes for immediately, three hours, and 24 hours post-exposure, respectively.

Model Results

Potential Confounders

No confounding variables were included in the main analysis. There was no evidence of any meaningful associations between observed potential confounders and the treatments.

Alcohol, Caffeine, Medication Use, and Smoke Exposures

Participants were asked to avoid non-approved medications starting 72 hours before the start of each study session, and avoid alcohol, caffeine, and smoke exposures starting 24 hours before, and continue throughout the end of the 24-hour health measurements. Reported alcohol and caffeine consumption and medication use was low during the time period before each study session and during each study session (see Tables S1 and S2). Reported use of alcohol, caffeine, and medication occurred with relatively even distributions across the six treatments. Univariate models did not find statistically significant associations between alcohol use, medication use, or caffeine use and treatment type (not shown).

Table S1. Alcohol, Caffeine, Medication, and Smoke Exposures by treatments: 24 hours before session start.

Variable	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone	Total
Total responses	46	45	43	44	43	45	269
Consumed alcohol	1	1	1	1	3	1	8
Consumed caffeine	4	4	4	5	5	2	24
Used medications*	4	7	6	5	7	8	37
Exposed to smoke	0	1	0	0	1	0	2

*This includes some use of daily medications that were approved by the study physician, such as oral contraceptives.

Table S2. Alcohol, Caffeine, Medication, and Smoke Exposures by treatments: during the study session.

Variable	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone	Total†
Total responses	47	45	43	44	45	47	269
Consumed alcohol	2	1	0	0	0	0	3
Consumed caffeine	2	4	2	1	2	2	13
Used medications*	4	4	7	6	4	7	32
Exposed to smoke	0	0	1	2	0	0	3

*This includes some use of daily medications that were approved by the study physician, such as oral contraceptives.

†Total is lower than in Table S1 because some participants missed the 24-hour follow-up period.

Mode of Commute

Participants were asked to use the same mode of commute into the facility on each study day. Driving was the most common mode of commute (59% of all trips to the facility for the first study day and 56% of all trips for the second study day involved a

car), followed by bike (36% of all trips on the first study day and 31% on the second study day; see Tables S3 and S4).

Table S3. Mode of Commute to Facility by Treatments: Before Session Start.

Mode	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone	Total
Bike	15	16	17	16	18	14	96
Bike+walk	0	1	1	0	1	0	3
Bus	0	0	0	0	0	0	0
Bus+walk	0	0	1	0	1	0	2
Car	26	25	22	25	24	28	150
Car+walk	3	1	1	1	0	2	8
Walk	2	2	1	1	1	1	8
NA*	0	0	0	1	0	1	2
<i>total</i>	46	45	43	44	45	46	269

*Not applicable: The participant did not report.

Table S4. Mode of Commute to Facility by Treatments: Prior to the 24-Hour Health Measurements.

Mode	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone	Total
Bike	14	15	12	14	14	12	81
Bike+walk	0	0	1	0	1	1	3
Bus	1	1	0	2	0	0	4
Bus+walk	2	2	3	2	3	4	16
Car	25	24	23	23	21	28	145
Car+walk	2	1	1	0	2	0	6
Walk	1	1	0	1	2	1	6
NA*	1	1	3	2	2	0	9
<i>Total</i>	46	45	43	44	45	47	269

*Not applicable: The participant did not report or the participant was not present for the 24-hour measurements.

Sleep Quantity/Quality

Most participants reported getting an “average” amount of sleep (75% for the night before the study session began and 76% for the night before the second study day); the amount of people reporting below-average sleep was less for the second study day than the first (19% for the night before the study session began vs. 9% for the night before the second study day; see Tables S5 and S6).

Table S5. Sleep Quality by Treatment: Night Prior to Start of Study Session.

Sleep	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone	Total
Above average	3	3	3	0	5	3	17
Average	33	36	32	36	32	32	201
Below average	10	6	8	8	8	11	51
<i>Total</i>	46	45	43	44	45	47	269

Table S6. Sleep Quality by Treatment: Night Prior to the 24-Hour Health Measurements.

Sleep	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone	Total
Above average	6	7	6	8	6	6	39
Average	34	35	33	31	32	35	200
Below average	5	3	2	4	6	5	25
Not applicable*	1	0	2	1	1	0	5
<i>Total</i>	46	45	43	44	45	47	269

*Not applicable: Participant missed the 24-hour follow-up period/survey.

Ambient PM_{2.5}, CO, and Temperature

Mean ambient PM_{2.5} in the 24-hours prior to the start of a study day ranged from 5.3 µg/m³ (control) to 9.5 µg/m³ (fan rocket; see Table S7). Minimum recorded mean PM_{2.5} was 0.9 µg/m³ (three stone fire) and maximum recorded mean PM_{2.5} was 18.8 µg/m³ (LPG). However, the range of ambient PM_{2.5} overall was narrow.

Table S7. Ambient PM_{2.5} Levels* by Treatment: 24 Hours before Session Start.

PM_{2.5}	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone
mean	5.3	7.6	5.6	9.5	6.7	6.7
min	1.5	2.9	1.0	2.0	2.6	0.9
max	12.3	18.8	11.2	17.6	10.6	12.7

* 24-hour average in µg/m³

Mean ambient CO in the 24-hours prior to the start of a study day ranged from 0.26 ppm (rocket elbow) to 0.35 ppm (three stone fire; see Table S8). Minimum recorded mean CO was 0.13 ppm (LPG) and maximum recorded mean CO was 0.70 ppm (three stone fire). Ambient CO was significantly higher for the three stone fire and LPG compared to the control. However, the range of ambient CO overall was determined to be narrow enough to not include this variable in the main model.

Table S8. Ambient CO Levels* by Treatment: 24 Hours before Session Start.

CO	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone
mean	0.28	0.32	0.27	0.30	0.26	0.35
min	0.16	0.13	0.19	0.17	0.17	0.17
max	0.60	0.48	0.45	0.50	0.44	0.70

* 24-hour average in ppm

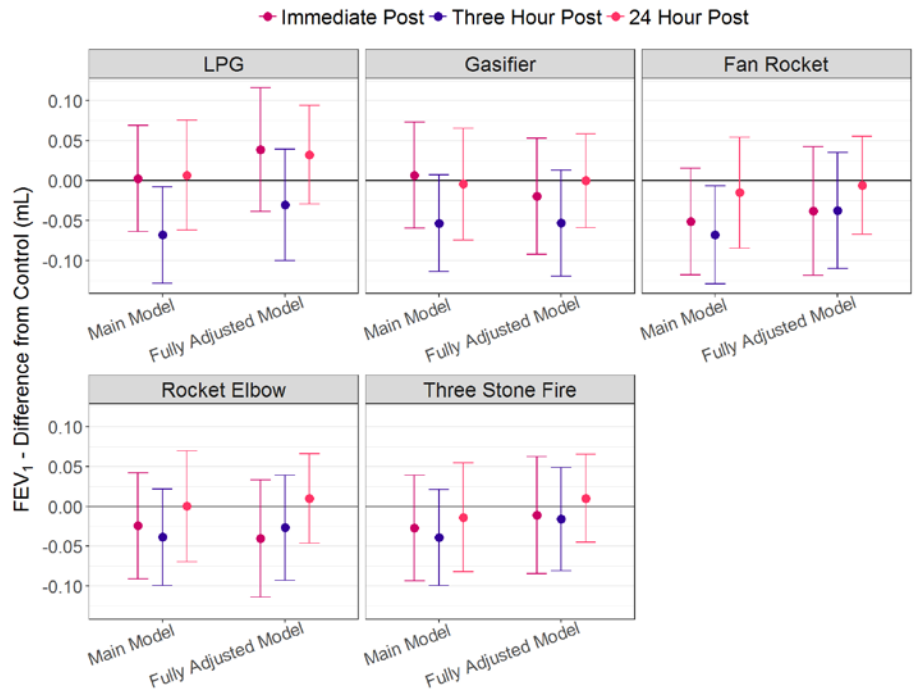
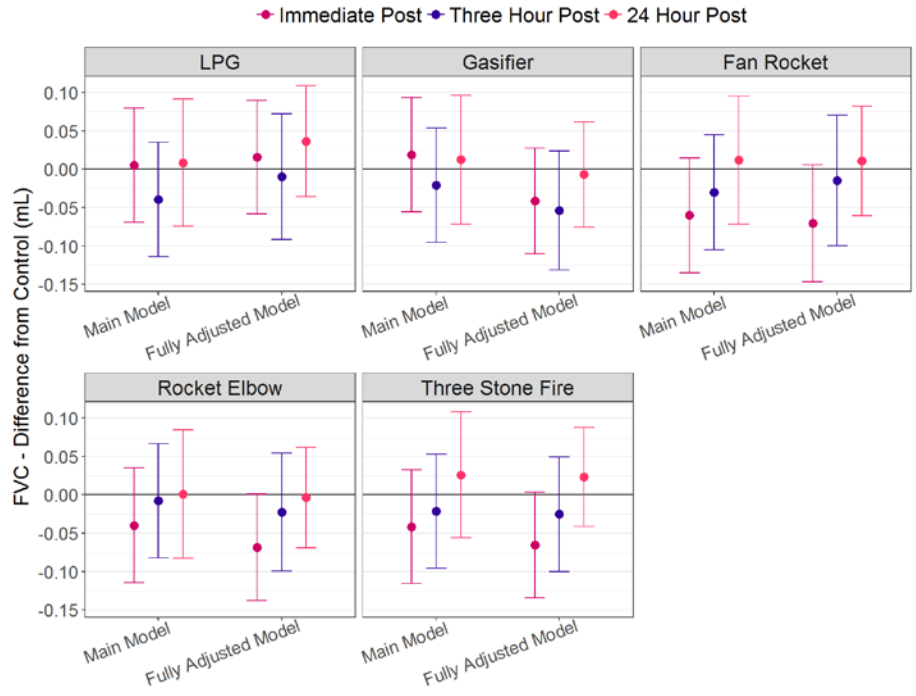
Mean temperature in the 24 hours prior to the start of a study day ranged from 6.4 °C (44 °F; three stone fire) to 15.4 °C (60 °F; fan rocket). Minimum recorded mean temperature was -8.5 °C (17 °F; rocket elbow) and maximum recorded mean temperature was 24.2 °C (76 °F; rocket elbow; see Table S9). However, the range of temperatures overall was determined to be narrow enough to exclude this variable in the main model.

Table S9. Mean Temperature (°C) by Treatment: 24 Hours before Study Session

Temp	Control	LPG	Gasifier	Fan rocket	Rocket elbow	Three stone
mean	7.1	10.5	7.9	15.4	13.2	6.4
min	-7.4	2.9	-6.2	4.6	-8.5	-3.1
max	20.0	23.9	16.8	22.3	24.2	15.3

Fully-Adjusted Model

The variables explored as potential confounders showed limited variation across the population and/or no associations with the treatment types, suggesting they were not confounders of the treatment-lung function association. However, to confirm that these variables did not impact our results, we ran a multivariable model that was equivalent to our main model but additionally included variables for alcohol use, caffeine use, medication use, sleep quantity, ambient PM_{2.5}, and ambient temperature. For models estimating lung function immediately and 3-hours post-exposure, we used the value of the binary variables that was reported for the 24 hours prior to the baseline measurement; for the 24-hour post-exposure measurement, we used the value reported for the time between the 3-hour post-exposure measurement and the 24-hour post-exposure measurement. For PM_{2.5}, and ambient temperature, we used the rolling average for the 24 hours prior to the measurement. Results of this model indicated that none of the added variables were significant predictors for lung function. Inclusion of the variables in the model did not meaningfully change the main effect estimates for treatment type (see Figure S1).



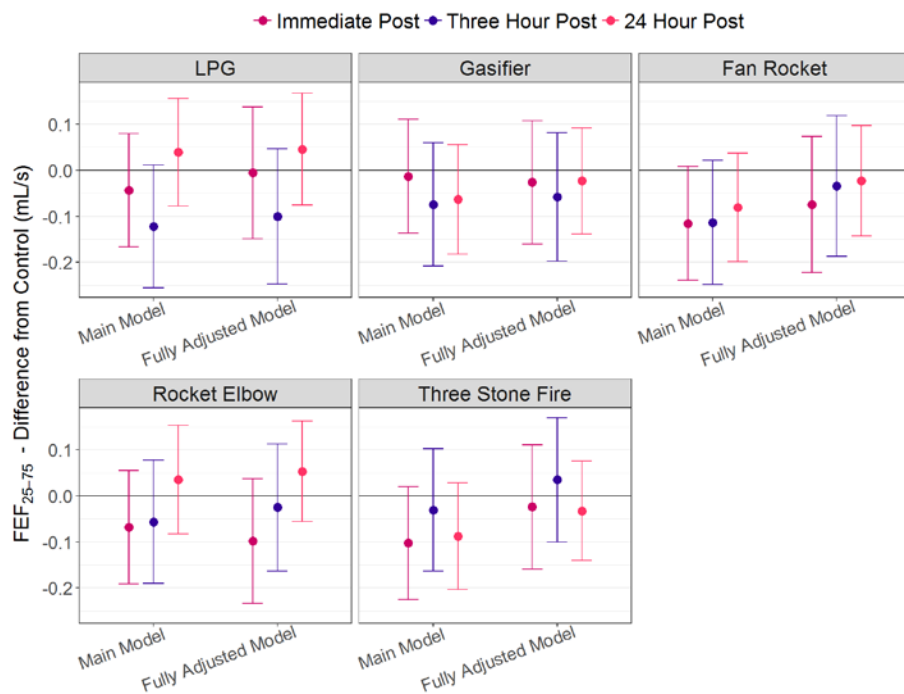


Figure S1. Effect estimates and 95% confidence intervals for mean difference in lung function for stove treatments compared to control for the main model versus the fully adjusted model. Fully adjusted model contains additional variables of alcohol consumption, caffeine consumption, medication use, sleep quantity, ambient PM_{2.5}, and ambient temperature. Top: FVC, Middle Panel: FEV₁, Bottom Panel: FEF₂₅₋₇₅.

Alternative Models/Sensitivity Analyses

Baseline values for each endpoint are shown in Table S10. Some differences occur across treatments, justifying the inclusion of a baseline term in the model.

Table S10. Mean Baseline (Pre-Exposure) Values for Spirometry Metrics.

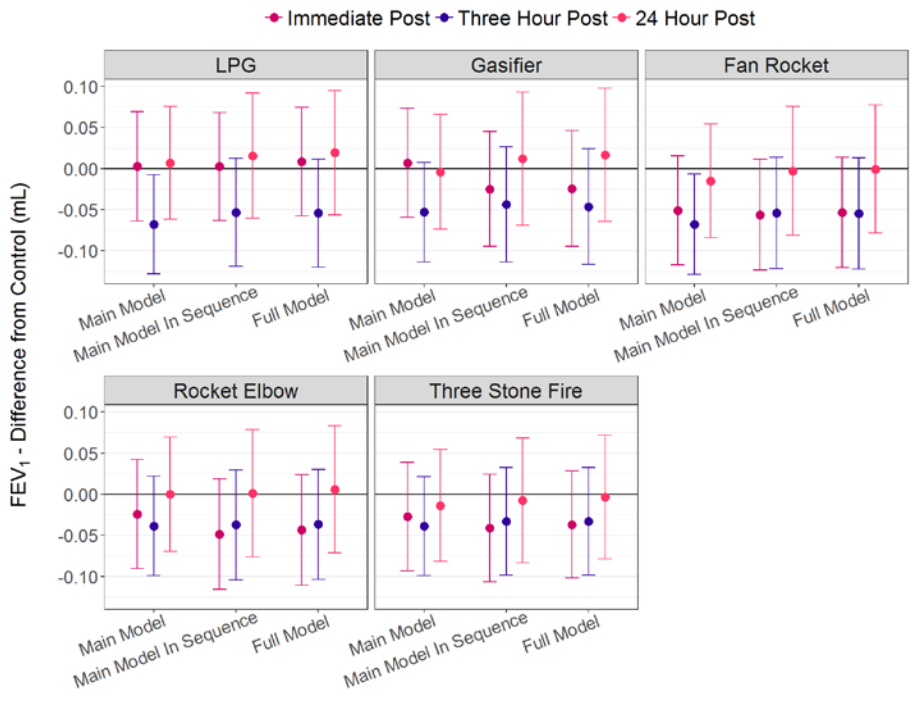
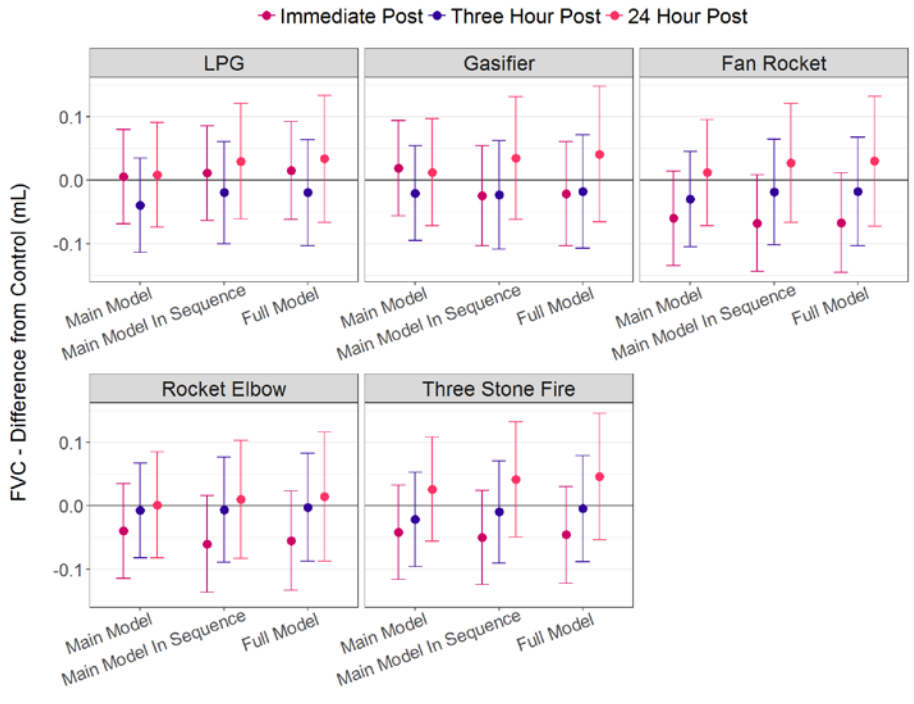
Treatment	Mean FVC (mL)	Mean FEV ₁ (mL)	Mean FEV ₁ /FVC (%)	Mean FEF ₂₅₋₇₅ (mL/s)
Control	4875 (1081)	3864 (787)	79.4 (6.6)	3832 (1101)
LPG	4854 (1024)	3860 (750)	79.8 (6.4)	3836 (987)
Gasifier	4867 (1043)	3873 (803)	79.4 (5.8)	3844 (1151)
Fan rocket	4879 (1148)	3873 (852)	79.4 (6.7)	3823 (1079)
Rocket elbow	4898 (1064)	3895 (816)	79.4 (6.4)	3862 (1148)
Three stone	4860 (1070)	3887 (793)	79.7 (6.8)	3915 (1126)

Structured model with study design parameters, in sequence data only (no makeups)

We developed a mixed-effect model that considered more structured study design parameters relevant to our Williams square than the main model. The model included terms for: baseline lung function, categorical stove treatment type, assigned sequence group, day of week (Monday vs. Wednesday), a sequence group/day interaction term, and a random person term. The model was run on a dataset that only included data collected within the intended sequence (e.g., we did not include data from makeup sessions), because the data from makeup sessions did not align with the sequence and day of week terms. Results of this model indicate no significance to the various fixed effect “design” terms (day of week, sequence group, or the sequence/day interaction term). There were no differences in main effect estimates compared to the main model with all data (Figure S2).

Main model, in sequence only (no makeups)

We ran the main model but on a data set that excluded data collected outside of the intended treatment sequence, to help compare between the main model and the structured model to understand differences due to the limited data set versus the different model type. Results of this model indicate no differences in main effect estimates compared to the main model with all data (Figure S2).



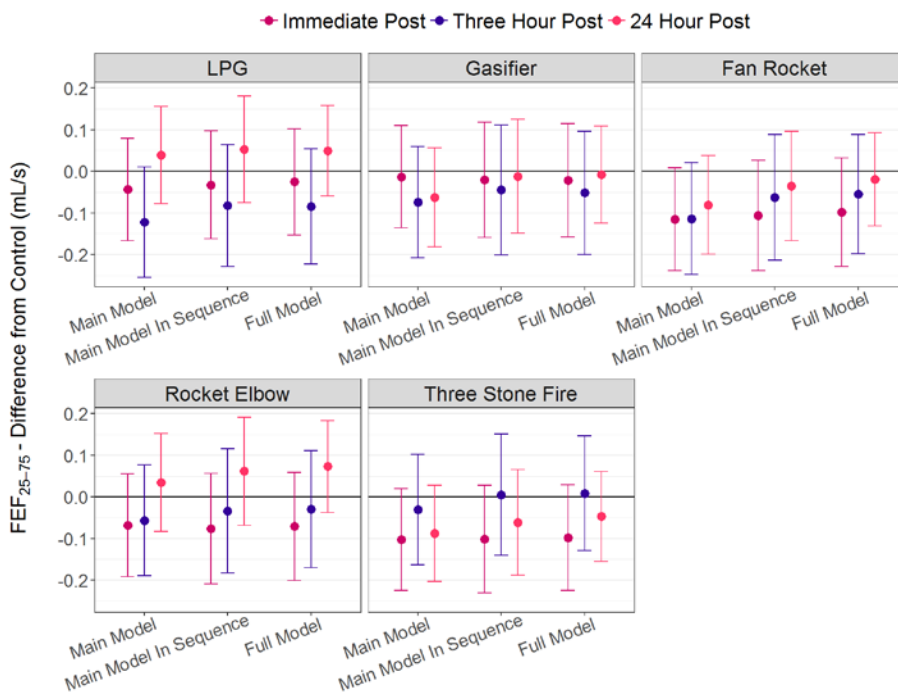


Figure S2. Effect Estimates and 95% Confidence Intervals for Mean Difference in Endpoint for Stove Treatments Compared to Control Across the Three Model Types. “Main Model” = as presented in main paper; “Main Model in Sequence” = the main model but with only data collected in sequence (makeups excluded), “Full Model” = the model containing all design variables, run only on data collected in sequence (makeups excluded). Top: FVC, Middle Panel: FEV₁, Bottom Panel: FEV₂₅₋₇₅.

Sensitivity analyses: Main model, remove C/D quality tests

We conducted sensitivity analyses wherein data was removed for measurements that did not meet an A or B quality rating (i.e., tests that did not meet the ATS/ERS quality criteria). This resulted in removing 21 data points (8% of the data) from the immediate-post exposure models, 25 data points (10% of the data) from the three-hour post-exposure models, and 17 data points (7% of the data) from the 24-hour post-exposure models. However, model results indicated no considerable differences between the estimates for the treatment effects (see Figure S3).

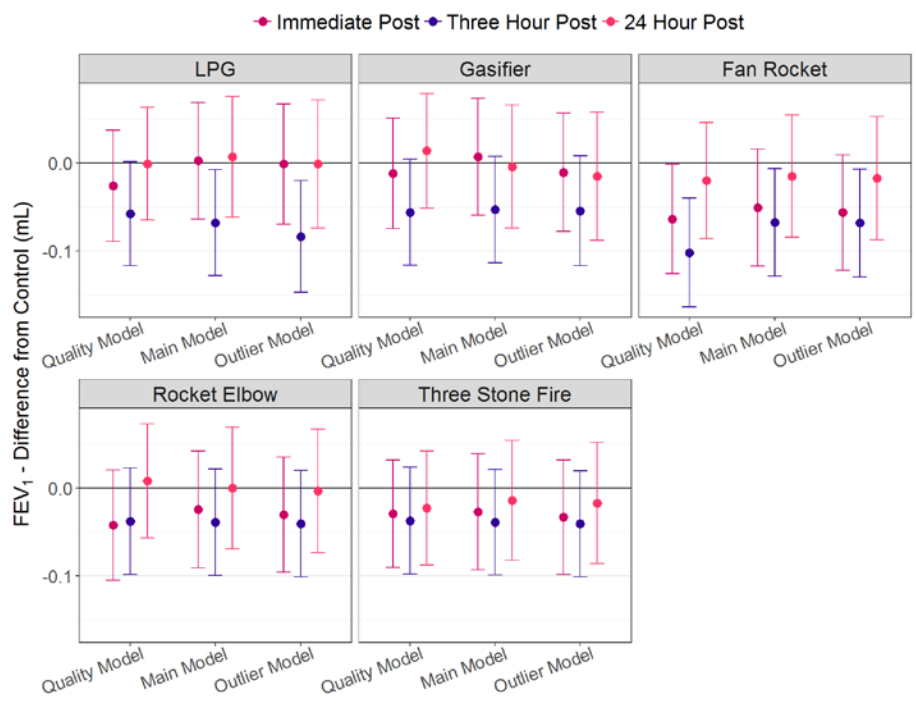
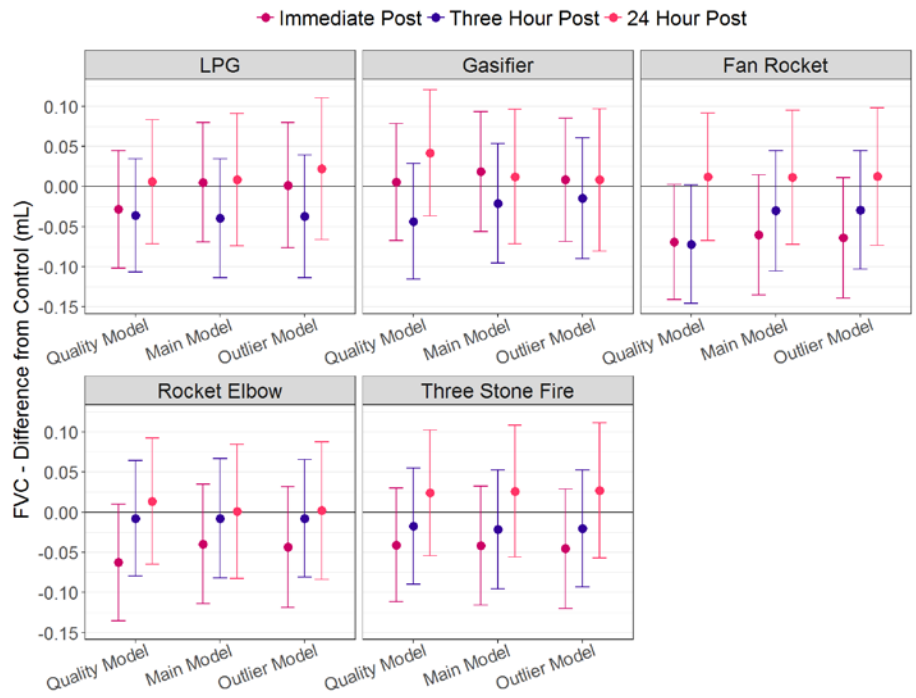
Sensitivity analyses: Main model, remove when exposure value outside narrow range of target

We ran the main model excluded data from study sessions where the exposure mean was outside of a narrowed range around the target value. The narrowed ranges were:

1. Control: less than 5 $\mu\text{g}/\text{m}^3$
2. LPG: 5-15 $\mu\text{g}/\text{m}^3$
3. Gasifier: 20-60 $\mu\text{g}/\text{m}^3$

4. Fan rocket: 75-125 $\mu\text{g}/\text{m}^3$
5. Rocket elbow: 175-300 $\mu\text{g}/\text{m}^3$
6. Three stone fire: 350-600 $\mu\text{g}/\text{m}^3$

Results indicated no considerable differences between the estimates for the treatment effects between this model and the main model (see Figure S3).



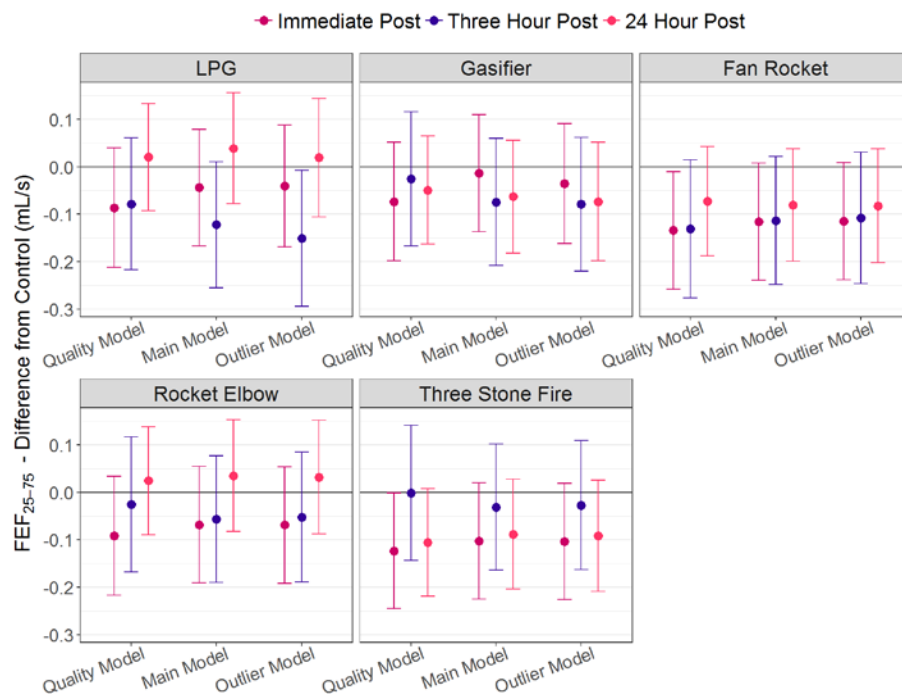


Figure S3. Effect Estimates and 95% Confidence Intervals for Mean Difference in Systolic Pressure (mmHg) for Stove Treatments Compared to Control: Comparison of Main Model to Models with C/D Quality Tests Removed (“Quality Model”) and Exposure Outliers Removed (“Outlier Model”). Top: FVC, Middle Panel: FEV₁, Bottom Panel: FEV₂₅₋₇₅.

References in Online Supplement:

1. Fedak KM, Good N, Walker E, Balmes J, et al. Acute effects on blood pressure following controlled exposure to cookstove air pollution in the Subclinical Tests of Volunteers Exposed to Smoke (STOVES) Study. *Journal of the American Heart Association* 2019; Under Review.
2. Miller MR, Hankinson J, Brusasco V, Burgos F, et al. Standardisation of spirometry. *The European respiratory journal* 2005; 26: 319-338.
3. U.S. Environmental Protection Agency. Air Quality Data API (<https://aq5.epa.gov/api>). [cited 2018 December 4]. Available from: <https://aq5.epa.gov/api>.
4. Colorado State University. Atmospheric Science Department’s Christman Field Weather Station (https://www.atmos.colostate.edu/fccwx/fccwx_data_form.php). [cited 2018 December 4].