AIM: Airflow Improvements during Meal-prep

NCT04464720

Statistical Analysis Plan
Version 4, 2/12/24

Stephanie Holm, MD PhD MPH

Table of Contents

Administrative Information

This SAP uses the framework for observational studies recommended by Hiemstra and colleagues.(1)

Revisions to SAP

v2-

- At this time the processing of the exposure data, consolidation of the dataset, and cleaning of the dataset have occurred. Some descriptive statistics have been calculated and visualized. None of the analyses listed in the analysis section have been run yet.
- Roles have been updated to reflect that Dr. Holm took over the analysis from Dr. Bueno de Mesquita after his post-doc was completed.
- In the Intervention-Pollutant Outcomes analysis, we initially described that we would perform the analyses relative to CO2 emissions, however as the time integrated emissions was difficult to identify for CO2, we will instead do this analysis relative to the Mean prominence of carbon dioxide during cooking-linked pollutant events. This has been updated in the relevant section below.
- In the Intervention-Lung Outcomes analysis, we had referred to the “variables listed below” and failed to list those. These have now been added, and more explicitly stated in the pollutant-respiratory outcome section of the analyses as well.
- In reviewing basic summaries of the data, we remembered that there are two households that had range hood replacements indicated, but that did not get them. We have added a sensitivity analysis excluding those participants to the sensitivity analyses section below.

V3-

- Following the posting of v2, we began to run GEE analyses as previously described, but ran in to difficulties executing a beta GEE for the proportion/percent outcomes. In discussion between Drs. Holm and Kang-Dufour, we realized that due to the simplifications that were made to the study design over the course of the pandemic (effectively meaning that there are only two time points per person) and the small final sample size, the analyses could more simply be done using a Wilcoxon signed rank for differences between the pre-post intervention analyses. Also due to the small sample size, and the need to control for multiple potential confounders in the pollution-outcome analyses, that third set of analyses will not be performed and instead a simple visualization of the relationship will be performed as an exploration.

V4- As we are completing analyses, we realized that the changes made in versions2-3 necessitate a couple additional edits that we had not foreseen. (1) We have edited the statistical framework section to note that we will use a p cutoff of <0.05 to determine statistical significance; (2) multiple imputation will not be attempted due to the small sample size; (3) we added a note that initial processing of the continuously monitored data was done in python; and (4) sensitivity analyses that relied on the addition of a potential confounding variable are not feasible so we will instead look for graphical patterns.

Roles and Responsibilities

Stephanie Holm, MD PhD MPH was an epidemiology PhD candidate when she wrote the grant, when this study was funded, planned and data collection began. She is responsible for developing this project idea, initial drafting of this analysis plan, completing the analysis and writing and revising the results for publication. She has selected the statistical models for use and discussed with other members of the team.

John Balmes, MD is the PI of the AIM study grant. Dr. Balmes has overseen all aspects of this work, and will contribute to interpretation of results, discussions with the team and revising of written drafts.

Brett Singer, PhD is an expert in indoor air quality, air pollutant emissions associated with natural gas burners and cooking, and kitchen ventilation. He has been integral to the planning of this project since prior to funding and co-led, with Drs. Holm and Balmes, the exposure assessment plan, including pollutant measurement protocols and specification of equipment use and exposure outcome metrics.

Mi-Suk Kang Dufour, PhD is a biostatistician who has assisted with statistical planning for the project.

Jacob Bueno de Mesquita, PhD is a post-doctoral researcher at Lawrence Berkeley Lab (working with Dr. Singer). He has been performing field data collection since the fall of 2021, including interactions with families, calibration of instruments and coordination with other folks at LBL. He was involved in processing and synthesis of study data, and may participate in drafting of papers as his interest and time allow.

William Delp, PhD is a mechanical engineer and expert in air pollutant emissions and controls, with specific expertise in gas burners, cooking and kitchen ventilation. Dr. Delp contributed to the protocols for kitchen ventilation assessment and joined for the analysis portion of the project. His role will be to identify and characterize cooking, range hood and particle emission events and to analyze the data to obtain the metrics noted.

Introduction

Background and rationale

In the past decade, there has been increasing concern regarding indoor pollution and human health as humans spend 80% or more of their time indoors.(2) In addition, pollutants such as fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2), can have higher levels indoors.(3) Cooking activities, especially use of gas stoves, can be a substantial contributor to emissions of multiple air pollutants.(4–7) A review of the literature on indoor sources of pollution reported that for some pollutants (such as NO_2) cooking exposures were the primary sources, with heating and other sources only adding minor amounts.(7) This contribution is especially important because as many as 65% of California households have gas stoves.(8)

Mechanistic studies of $PM_{2.5}$-related health effects point to multiple modes of action in the body, including the induction of oxidative stress(9,10) and inflammation.(9) These mechanisms can directly increase respiratory symptoms and decrease lung function.(9) Given the connection between particulate matter, NO_2, and respiratory symptoms generally, it is not surprising that both gas cooking and specifically household NO_2 levels have been associated with wheezing in children.(11–13) Studies of school-aged children have reported an association of cooking with cough,(14) daily NO_2 exposure with nighttime inhaler use,(15) and 72-hour NO_2 exposure with cough or nighttime symptoms.(16) In a large cross-sectional, nationally representative survey there was an association between lack of cooking ventilation and higher odds of childhood cough,(17) asthma,(18) wheezing (with or without asthma diagnosis),(18) and bronchitis.(18) A recent study by our group reported that cooking exposures were an important contributor to indoor $PM_{2.5}$ in a sample of urban households with asthmatic children, and specifically that households that used ventilation during cooking had lower $PM_{2.5}$ levels.(19)

Objectives

This is a pilot study to explore whether an intentional cooking ventilation intervention could improve exposures and health outcomes in the homes of children.

Study Methods

Study Design

This is a before-after trial in homes in the San Francisco Bay Area that have both a school-aged child, a gas stove and a venting range hood or over-the-range microwave and hood (OTR). The trial assessed a cooking ventilation intervention consisting of: (1) ensuring that the home has a working range hood that meets both airflow and sound performance standards and (2) education about the hazards of cooking pollutants and the benefits of using the range hood whenever cooking occurs. The intervention started with measurements of the range hood airflow and sound level. If the existing range hood or OTR did not meet the targeted combination of a minimum airflow of 100 cubic feet per minute (cfm) and sound pressure (loudness) no greater than 60 A-weighted decibels (dBA) at a distance of 2 m, a contractor assessed the feasibility of replacing the device. When feasible, the participant (and building owner, if the participant was a renter) was offered a replacement hood or OTR that met the specifications. All study participants received education regarding use of the range hood during all cooking events. We installed stove and range hood sensors in participant homes for the entire data collection period, allowing for objective measurements of cooking intervals and range hood use both before and after the intervention. In addition to a working range hood, we provided all families an incentive at each visit (totaling $100 per family). Outcomes include measurements of estimated $PM_{2.5}$ and NO_2 in the household, and both objective and subjective measures of respiratory health in the children (spirometry, FeNO, ISAAC Asthma questionnaire and ISAAC Environmental Questionnaire for all children; the asthma control test as well for children with asthma). The trial is registered with clinicaltrials.gov, #NCT04464720.

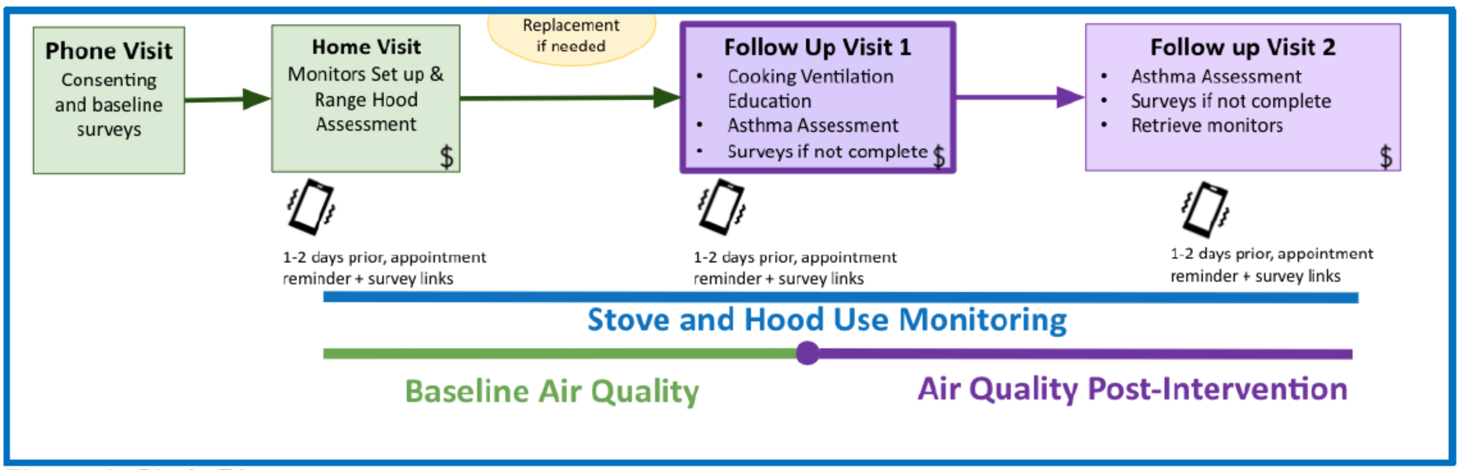


Figure 1. Study Diagram

The educational intervention consisted of a video instructing the child and their guardian on cooking practices to decrease pollution exposure. The Richmond youth council has provided crucial feedback during development of the video. Their involvement helped to ensure that the information is presented in a way that will resonate with the local community and also provided the youth with exposure to developing health education tools. The family will also receive a printed copy of the accompanying infographic as a reminder of the material in the video. The video features a woman of color who is the cook in her own home, Dr. Stephanie Holm, and a diverse group of children at a playground. The video explains that: cooking creates air pollution (even though you can't always see it), that pollutant levels from cooking can reach the same levels achieved by wildfire smoke, that you can decrease pollution by doing 3 things (1) running the range hood every time you cook, (2) cooking on the furthest back burner and (3) moving other cooking appliances closer to the range hood.

Due to COVID-related obstacles with recruitment, as well as delays related to IRB review and shifting safety guidance, we modified the initial study plan during the summer of 2021, which would have been a pilot stepped-wedge randomized trial, meaning that households would have been randomized to receive the intervention at one of two possible times. The intervention would have been provided to those households randomized to group 1 after one week of baseline data collection and to group 2 after 2 weeks. Due to the recruitment difficulties associated with the ongoing pandemic, the study interval was shortened to two weeks (to ease participant burden), with all families receiving the intervention after the first week. The study was also opened to all households with children in the appropriate age range at that time, rather than restricting only to homes that had a child with asthma. In the spring of 2022, we extended the study interval to two weeks between the initial home visit and follow up visit one, and two weeks between follow ups 2 and 3. In total:

- 2 households completed the protocol with 4 total visits & roughly one week between each pair of visits
- 5 households completed the protocol with 3 total visits & roughly one week between each pair of visits
- 7 households completed the protocol with 3 total visits & roughly two weeks between each pair of visits

In total, the range of days between first visit and the educational intervention is 5-26 days (mean 13.1 days) and the range between the educational intervention and last visit is 5-21 days (mean 11.5 days).

Data Collection Procedures

Baseline Phone Visit

At the baseline data phone visit, study staff conducted interviews to complete the baseline questionnaires: the modified ISAAC Core Questionnaire (a four-page questionnaire about asthma and associated diseases), the ISAAC Environmental Questionnaire (a one-page excerpt of a questionnaire regarding environmental exposures at home) and the baseline cooking survey (a two-page Questionnaire about cooking behaviors in the household and some demographic information about the parent). Portions of the ISAAC Core questionnaire assess general medical history (whether a child has ever had symptoms consistent with asthma, allergies or eczema), these questions were asked of all children. The questions specific to an asthma diagnosis were only asked of the children with asthma.

Home Visit #1

2 days prior to the home visit an appointment reminder was sent, with a link to complete the visit surveys via REDCap: the childhood Asthma Control Test (cACT, only if the child has asthma, a one-page symptom questionnaire) and the follow up cooking survey.

At the visit, an assessment of the cooking ventilation present in all participant's homes was performed. If a hood was present, functional and exhausting to the outdoors, a measurement of the airflow through that hood was performed by an experienced range hood assessor who is also a licensed contractor. All other procedures were performed concurrently by a study staff member. A Digisense Data Logging Vane Anemometer was affixed to the existing range hood. This collected and logged real-time air flow data to monitor the frequency of ventilation use prior to the educational intervention. Real-time stove use data was also collected using Lascar Easylog Thermocouple Data Loggers and Hobo temperature loggers, which use temperature data to derive cooking intervals. The thermocouples were fixed in place on the stove with either magnets or a small amount of heat-resistant tape. Concentrations of $PM_{2.5}$, NO_2, CO_2, and other parameters, were measured in the primary living area of each home (a location which can well-represent the exposure received by the occupants) and typically logged every 1 minute in each household, starting at the time of this visit. The air contaminants were measured along with temperature and relative humidity using an eLichens Indoor Air Quality Pro Monitor, which had/has sensors for $PM_{2.5}$, NO_2, CO_2, volatile organic compounds and other environmental parameters . Time resolved NO_X and NO_2 were measured during each interval at each house using an Ogawa passive sampler, which was placed with the eLichens real-time monitor.

The child (regardless of asthma status) performed two breathing tests (spirometry using an EasyOne Portable Spirometer and fractional exhaled Nitric Oxide (FeNO) using the Niox VERO device) with assistance from study staff.

Visit #2

The parent and child completed the same pre-visit surveys, and at the visit the child performed the same lung function testing. They were given the educational intervention, consisting of watching the video and receiving a reminder card and magnet. The equipment in place at the home was checked (anemometer, thermocouples, eLichens). The Ogawa passive sampler was exchanged for a new one. If air flow in the range hood (or OTR) was less than 100 cfm or noise levels are intolerable (>60 dBA) for settings >100 cfm, the range hood was replaced with a functioning range hood model during or prior to this visit, if feasible within the study timeframe and budget.

Visit #3

The parent and child completed the same pre-visit surveys, and at the visit the child performed the same lung function testing. The equipment in place at the home was collected (anemometer, thermocouples, eLichens, Ogawa).

Power Considerations

This is a pilot study, and with COVID-related delays and adjustments, we were not able to complete the study with enough households to be fully powered for either exposure analyses or respiratory health outcomes (especially as the study was modified to include all children, rather than only those with asthma). However, the initial power calculation is retained here, to document the intended recruitment and that we acknowledge that we are underpowered.

In THE AQUA study the adjusted mean difference in $PM_{2.5}$ between homes that used the range hood (ever) and those that never did was 5 $\mu g/m^3$ (standard deviation 8.24).(19) Using Cohen's equations for t-tests (assuming $p = 0.05$ and $\beta=0.2$), if we expect to see at least a 7 $\mu g/m^3$ difference with an intentional cooking ventilation intervention, we would expect to be able to detect that difference with 23 households. Belanger and colleagues(21) showed an average NO_2 level in homes of kids with asthma in Massachusetts was roughly 10 ppb, and low-income homes that use gas stoves a lot (>1h/day) have NO_2 levels 2-5 times those that do not.(6) In a study out of Lawrence Berkeley Lab, in homes expected to have high NO_2, NO_2 levels were roughly 10 ppb higher in homes with gas cooking. Thus we estimate that a ventilation intervention in homes with gas cooking could optimistically decrease NO_2 7 ppb. To detect a difference of this size requires 28 households.

An intentional use of kitchen ventilation may decrease the $PM_{2.5}$ levels by 10 $\mu g/m^3$ or more, which could decrease FeNO in asthmatic children by 7 ppb based on work by Mar et , showing a decrease in FeNO based on prior day $PM_{2.5}$ averages.(23) The minimum detectable change in FeNO for a group size of 30 is 3.3 ppb. Stanojevic et al (24) report that the coefficient of variation for FEV1 in children with asthma in around this age is roughly 10%, suggesting a standard deviation of roughly 10. For 30 children, the minimum detectable effect size will be 0.296, or a minimum detectable FEV1 difference of 2.9 percent of predicted.

Framework

Though our preference is to report estimates and confidence intervals, because we have pivoted to the use of Wilcoxon-Signed Rank tests (which do not generate an estimate), we will report *p* values for statistical tests.

Interim Analysis and stopping guidance

The only interim analyses to be performed include summary statistics about basic cohort demographics, and discussion of the free-text comments from participants in their end of study survey. All other analyses have been held until data collection is complete.

Timing of Analysis and Outcome Assessments

While field data collection was initially planned for summer 2020, it was delayed until summer 2021, due to the onset of the COVID-19 pandemic. In the summer of 2021, a surge in cases related to the delta variant of SARS-CoV-2 made recruitment difficult. We made protocol changes as described in the study design section. All data collection has been completed prior to the finalization of the statistical analysis plan, though no analyses have been done (other than a few descriptive statistics for this SAP).

Statistical Principles

Confidence Intervals and p-values

P-values less than 0.05 will be used as the threshold for significance.

Adherence and Protocol Deviations.

Not applicable

Analysis Populations

All households for which we collected data for 3 or more cooking events both prior to and following the intervention will be included in exposure analyses. Children who successfully completed lung function testing both before and after intervention will be included in respiratory health outcomes. For households with data on more than one child, the child who has the most complete data (most spirometry and FeNO results that meet grading criteria for acceptable tests) will be used in the analyses.

Study Population

Cohort description

We recruited children ages 6-12, living in Richmond/ San Pablo/ El Cerrito/ Albany/ Berkeley/ Oakland/ Alameda. They were eligible for this study if the parent reported that they had both a gas stove and a venting range hood, i.e. one that extracted air from the kitchen to outdoors. They were excluded from the study if they lived with a smoker who smoked indoors, if they knew they would not have stable housing for the period of the study or if they were not fluent in English.

Recruitment occurred through East bay pediatric clinics where children are seen, using contacts previously established by the investigators. These clinics include Contra Costa Health System (CCHS) Clinics, UCSF Benioff Children's Physicians-Hilltop Pediatrics Clinic, Lifelong-Jenkins clinic, Kiwi pediatrics and UCSF Benioff Children's Hospital Oakland (BCHO) Outpatient clinics. We posted recruitment fliers in the clinics and had study information cards that physicians and staff could give to children with asthma. At a few clinics postcard mailings were sent to potentially eligible patients.

Due to the paucity of in-person visits happening at the Contra Costa County Health Clinics and Hilltop, recruitment at those sites was supplemented in the spring of 2021 with a postcard mailing to all patients aged 6-12 with a diagnosis of asthma. Recruitment fliers and cards were also posted in community spaces where kids and families may spend time (e.g. local library), and at other clinics where no staff are participating in any recruitment activities (not handing out cards or otherwise). Digital flyers are also distributed to school districts and to other community organizations that regularly distribute materials to families, for inclusion in their outreach (newsletters, digital bulletin boards). We have also been advertising the study on our webpage and through social media.

Potential confounding covariates

Participant age, sex, BMI, maternal educational level, presence of a smoker (who smokes outside, as households with an indoor smoker will be excluded), as well as history of asthma, eczema and allergies will be assessed using the Isaac Questionnaires,(25) as these could be potential confounders of the range hood use and asthma symptom relationship. Geographic factors which could influence the air quality such as proximity of the home to major roadways, railways and industrial sites will be assessed by geocoding the home address of the family. Outdoor $PM_{2.5}$ and NO_2 concentrations will be estimated from regulatory data and/or local low-cost sensor network(s).

Analysis

Environmental Assessment and Appliance Use

Calibration, Initial QA of the Data

We performed multiple sets of instrument co-locations of the e-Lichens air quality monitors for cross calibration, and comparison to reference monitors. Those colocations occurred in June 2021 before any field deployments, in the fall (Oct-Nov) of 2021 and again in June of 2022, after the last field deployment.

Quality assurance of the pollutant data will include:

- Preliminary visual review of data
- Confirm consistent units for temperatures (some of which were recorded in F and some in C)

For quality assurance, thermocouples were tested simultaneously during cooking events to ensure that all showed similar response rates to temperature changes (which they did).

Compiling, Aligning and QA of Continuous Data

Data will be compiled from our various data collection sources to create a master environment/appliance use dataset for each household in the study. The data include the following:

- Main living area measurements
    - eLichens; 1 min interval measurements (Pull from eLichen API with Python script): NO_2, Total VOC, CO_2, $PM_{2.5}$, Temperature, Relative humidity, Sound intensity
    - Ogawa samplers: (Pull from spreadsheet directly to R): Time-integrated NO_2 and NO_X over periods of roughly 1-week each
- Kitchen measurements (pull from Box uploads directly into R):
    - Burner-adjacent temperature measurements from thermocouples; 1 min interval measurements
    - Temperature above cooktop and and at oven outlet, measured by HOBOs; 1 min interval measurements
    - Airflow through hood exhaust, measured by vane anemometer; 1 min interval measurements
- Start and stop times for sampler/data logger deployment (Pull from REDCap download directly to R)
    - Pre-intervention data logging start and stop date/time (end of initial visit-time of Ogawa placement at intervention visit)
    - Post-intervention data logging start and stop date/time (time of new Ogawa placement-removal of Ogawa at final visit)
    - Variables for date of intervention and date of range hood replacement/work
- Outdoor air pollution data
    - $PM_{2.5}$, NO_2 (Pull from regulatory data, create inverse distance weighted surface- may also be able to access data from BEACO2N supplemental, high-resolution network)

Prior to calculating any summary measures, all data will be visually reviewed to screen for obvious quality assurance issues including the following: missing data; instrument malfunction (e.g. values out of range of typical residential conditions); data points that are inconsistent with temporally adjacent values (e.g. a one-minute value of 500 ppb NO_2 in the middle of a steady progression from a low baseline to a peak that is below 100 ppb); shifts in baseline values that suggest that the device has not equilibrated after deployment, 1-h time shifts in some but not all data following spring or fall daylight savings time changes; and inconsistencies between temperature sensors at the stovetop and above the stove. Any issues identified and any adjustments (e.g. shifting of one or more time series by 1 hour to fix the time-change error) will be documented.

We will add flags for periods of missing data or data points removed based on the visual screening. When data are missing over short intervals and interpolation based on neighboring data is feasible, interpolated values will be imputed into the dataset using the simplest method that is consistent with the pattern of the surrounding data, and the imputation will be documented. As an example, a single missing data point in a steadily increasing concentration profile may be estimated by linear interpolation, but an interval of 5 missing data points during a first order decay may be estimated based on the fitted decay. Intervals of missing data that are not clearly interolatable will be left as missing values.

Air quality measurements recorded by the eLichens monitors may be adjusted to achieve consistency across devices, based on analyses of instrument co-locations for cross calibration. Any adjustments will be documented. For situations where multiple devices were present (e.g., duplicate sensors deployed) we will calculate all metrics with each device and average the results by event/metric.

At this point there will be a temporally aligned dataset of all continuous data from each home.

Time-integrated NO_X and NO_2 samples

The analysis of time-integrated NO_X and NO_2 samples, collected with Ogawa passive samplers, require temperature (T) and relative humidity (RH) values to obtain estimates of the concentrations over the sample period. We will compute mean T and RH for each Ogawa deployment period based on measurements of the co-located eLichens monitors. If T and RH data are unavailable for any Ogawa samples, we will impute the average conditions from other homes sampled as close in time to these data as possible.

Identification and analysis of self-reported cooking data

The survey responses about cooking will be pulled from the electronic data collection system for the 1 or more weeks prior to the intervention and 1 or more weeks after the intervention. (Pull from REDCap download directly to R)

- ○ How often was the gas stove or oven used in your home over the last week: Breakfast, lunch, dinner, other times (never, 1-2 days, 3-4 days, 5+ days)
- If food was being cooked, how often was it: Fried, deep fried, grilled, steamed/boiled, toasted, electric pot cooker, baked, microwaved
- Think about the burner you used the most this week. Is it one of the burners in the front or back? (front/back/my stove doesn’t have front and back burners)
- In the last week, how often was the stove hood fan used during cooking? (never, rarely, occasionally, frequently)

Identification and analysis of events associated with monitored cooking burner and range hood use

Since the intervention was designed to increase range hood use and reduce in-home concentrations of pollutants resulting from cooking (leading to overall pollutant exposure reduction), the analysis of environment and range hood use outcomes will start with the identification of cooking burner usage events. These events will be identified by analysis of the temperatures measured by thermocouples adjacent to cooktop burners, T/RH sensors above the oven vent, and T/RH sensors mounted above the cooktop, on/in the range hood. Events of range hood usage will be identified by analysis of data collected by the anemometer mounted in the path of airflow for the range hood. And the magnitudes of air pollutant concentration increases resulting from cooking will be quantified by analysis of the time series data from the eLichens monitor. The pollutant event magnitudes will also be referenced to the amount of cooking burner use.

When burners were used, and in some cases, also from cooking that occurred in the homes, it is expected that concentrations of air pollutants increased at the monitoring location, which was typically 1-2 rooms away from the kitchen. The concentration profile should be related to the emissions characteristics (total mass emitted, emission duration, quality or pattern of emissions), the volume of all the spaces into which the emissions mixed (starting from the kitchen), and the rate of removal processes such as mixing into farther rooms, outdoor air ventilation, filtration, deposition etc. Since many emission events occur over time periods that are shorter than removal rates, a common pattern in the kitchen and often also closely connected spaces, is for a sharp or gradual increase of pollutant concentrations, reaching a distinct peak value, followed by a decay that reflects a combination of additional mixing and removal processes. For a consistent set of transport and removal processes, the peak height will be proportional to the magnitude of total emissions. When emissions occur over periods longer than removal processes, the pollutant concentration may achieve a roughly steady value, or “steady state”. Again, with removal processes constant, the magnitude of steady state would be proportional to emissions, and inversely proportional to the aggregate rate of removal processes. Conversely, differences in peak height for a short term emission event and differences in steady state levels for longer emission events, could result from different emissions, different removal rates, or both.

Another common feature of indoor air pollutant concentration patterns in the room with the source, and sometimes an adjacent room, is a faster initial decay from the peak that reflects mixing from the room to other rooms, or within a very large room (e.g., a “great room” of an open floor plan home), followed by a slower first order decay that reflects removal from the space, e.g. by outdoor air ventilation, deposition, etc.

Kitchen ventilation can reduce exposures related to cooking by (1) removing the pollutants at the source, before they mix into the home, and by (2) increasing the removal rate through higher overall ventilation.

Identify cooking events

The analysis will start with identification and characterization of cooking *burner activity* and *cooking events*.

- Individual instances of potential *burner activity* will be identified algorithmically and checked / adjusted based on visual review. For each incidence of an individual burner temperature signal increasing in a manner that suggests potential use the analysis will identify the start time of the temperature increase, the time when temperatures start to decrease back to baseline (i.e. accounting for potential resetting of the burner to a lower temperature for steady cooking), peak temperature above the baseline prior to the increase, and mean rate of temperature increase (fitted).
- If the thermocouple was placed closely adjacent to the burner, the start of a burner at any setting of medium or above should provide a signal of temperature increasing much more rapidly than would occur from even the most extreme environmental drivers (e.g. direct sunlight hitting the site through a window, or room air heated by heating system). The temperature increase will continue as the pot on the burner is heated, and gradually reduces its own rate of heat absorption from the flame and increases its own contribution to the adjacent thermocouple by radiation and increasing convective heat transfer. A pot being held at constant temperature is expected to provide a relatively stable temperature at the adjacent thermocouple. Adjusting the burner to a smaller flame may result in a decrease in thermocouple temperature, to a lower plateau. A rapid decrease in temperature is assumed to represent a burner being turned off or adjusted to a much lower setting. Differentiating these two events is done based on whether the temperature returns to baseline or to a new, lower plateau.
- Recognizing that heat generated at one burner causes temperatures to increase at adjacent burners (confirmed by controlled exploratory testing on several cooking ranges), the individual potential burner activity instances will be visually reviewed and coded as (a) confirmed distinct activity, (b) unlikely distinct (i.e., caused by transfer from another burner), or (c) indeterminate, i.e. possibly distinct individual burner activity..
    - If the timing of the temperature increase and decrease for one burner closely correlates or slightly lags with that of an adjacent burner, and the overall peak temperature is substantially lower for one or more burners with this temporal alignment, the event for the lower/lagging burner typically will be assumed to be caused by the burner with higher temperature and steeper change with time. All complex burner events will be reviewed by a second researcher and disagreements will be discussed and resolved.
    - For each confirmed burner activity we will record the number of minutes, providing a burner activity metric of burner-minutes.
- Distinct individual burner activities that overlap within 10 min will be combined into *cooking events*.
- Each cooking event will be characterized by recording or calculating the following:
    - Event start time, starting with activity of the first burner.
    - Event end time, when the temperature of the last burner in use starts to decrease with no subsequent plateau.
    - Event duration (min).

    - Number of burners used, with two separate uses of the same burner, e.g. 10 min of burner activity, 4 min off, 8 min of burner activity, would count as two distinct burner activities but fall within a single cooking event.
    - Number of burner minutes for each burner for the event.
    - Locations of burners used: oven, front of cooktop, back of cooktop.
    - Cooking events will also be identified and characterized using the temperature measured by the HOBO above, in the range hood.

Identify and characterize range hood use events

- Identify range hood use intervals and for each specify start and stop times, duration. Note speed/setting if discernible from data and calibration is available.
- Each range hood event will be characterized by recording or calculating the following:
    - Event start time
    - Event end time
    - Event duration (min).
    - Estimated speed setting for event if possible

Identification and analysis of pollutant events including those that occur coincident with cooking events ('cooking-linked pollution events')

Important notes about pollutant events:

- If the monitor was in the kitchen with the cooking burners, the NO_2 and CO_2 is expected to start increasing noticeably within a few minutes of the start of the cooking event. With the eLichens monitor in a separate room, the delay between the start of the cooking event and the start of pollutant events will differ between homes, owing to the different mixing or movement of air from the kitchen to the location of the monitor. The delay also could vary across events within a single home, owing to different room-to-room mixing conditions occurring throughout a day and with varying outdoor conditions, resulting in different directional air flows within the home.
- The timing of CO_2 and NO_2 increases from cooking burner use should be closely aligned, as the pollutants will be transported together. However, since CO_2 has a higher background and is also emitted by occupants, the NO_2 increase may appear earlier and more clearly.
- $PM_{2.5}$ is produced in significant quantities in only some cooking events (especially frying), and may not be produced steadily throughout the events with emissions. For many cooking events that produce $PM_{2.5}$, the emissions occur as a burst, sometimes within a very short interval.
- There are other sources of $PM_{2.5}$ and CO_2 and possibly, though less likely, NO_2.

Identify and characterize pollutant events

- Identify events based on moderately to sharply increasing NO_2, CO_2, or $PM_{2.5}$ up to a peak or plateau when the start of the increase closely follows the start of a cooking event. To be considered as a cooking-linked pollutant event, the pollutant concentration must start to increase within 10 min if only one room away and within 20 min if 2 or more rooms away, and end with similar or shorter delay relative to the end of the cooking event. (These times are imprecise and based on the experience of the research team reviewing data from homes, with relevant examples provided in Singer et al. 2017). The start time is the last point before the increase. The end time of emissions is the peak or the last point before concentration starts to drop, following a roughly steady, elevated level. Event / peak analysis starts with the identification and subtraction of the baseline or background contribution from long-duration sources including infiltration of outdoor pollutants (for $PM_{2.5}$ and NO_2) and occupant-generated CO_2. We use the LOESS routine within the Python pyBaselines package to estimate the undisturbed values (baselines) for each pollutant and also for burner-adjacent and above the cooktop temperature series. This routine is referred to as RBE, robust baseline estimator, and is based on a locally estimated scatterplot algorithm with reweighting. The width of the rolling window is set to eight hours. The estimated baseline is subtracted from the time series of measured values to produce a time series of incremental concentrations resulting from events and activities. We use the Find Peaks routine within the Python SciPy package to identify events based on prominence over the baseline / background. The routine determines the peaks, and the associated start and end times.

- Calculate CO_2, NO_2 and $PM_{2.5}$ event magnitudes (for all events whether cooking-linked or not) based on the prominence of the highest 10-minute average (compared to pre-event baseline) and integration of concentration above baseline, using manual and/or algorithmic detections of peaks. Integrated metric will have units of concentration * time, (ppm-min for CO_2, ppb-min for NO_2, ug/m3-min for $PM_{2.5}$).
  - Peaks that occur within 10 min of each other (also an imprecise threshold based on experience of the research team), with intervening concentrations not dropping to <50% of the level of the first peak, shall be considered as being part of the same event.
  - Determine an event baseline based on the fitted baseline / background.
  - Events for which algorithmic classification provides nonsensical or undetermined results will be flagged and evaluated visually by at least two researchers.

For each pollutant event ($PM_{2.5}$, NO_2, CO_2), we will calculate the following statistics:

- Event statistics:
  - start time and end time of *increasing* concentrations (i.e., not including the decay)
  - duration of emissions event
  - time of return to baseline (based on fitted baseline)
  - end of decay interval.
- Pollutant statistics:
  - prominence (highest ten-minute average above baseline) of each pollutant
  - integrated concentration of each pollutant above baseline (ug/m3-min)

Notes:

- There should be few or no NO_2 events that are not linked to cooking.
- There are unlikely to be any sharp CO_2 events not linked to cooking; but there will be CO_2 events linked to occupancy.
- There could be many $PM_{2.5}$ events not linked to cooking.

All events identified by algorithms (Burners, cooking, range hood, $PM_{2.5}$, NO_2, CO_2 increases) will be visually reviewed and any that are not deemed to be clearly correct will be reviewed and discussed with a second researcher.

Calculate Summary Statistics for each home during each of the pre- and post-intervention periods:

- Primary summary statistics of cooking and range hood events
  - Number of discrete events of individual burner use (confirmed and confirmed + possible) for cooktops, ovens, and any burner.
  - Number of cooking events.
  - Number and fraction of cooking events with substantial particle emissions.
  - Total burner minutes.
  - Mean (SD) number of burners used and burner minutes per cooking event (as an indicator of similarity).
  - Delay from start of cooking to start of range hood use.
  - Fraction of total burner minutes that reflects usage of the front burners and fraction that reflects usage of the back burners.
  - Fraction of cooking events with any range hood use.
  - Fraction of cooking events with range hood used for >80% of cooking burner use.
  - Fraction of cooking-linked pollution events that resulted in any range hood use and use for >80% of active burner time.
- Secondary summary statistics of self-reported cooking behaviors:
  - mean frequency of gas stove or oven use
  - mean frequency of the queried cooking methods (fried, deep fried, grilled, steamed/boiled, toasted, electric pot cooker, baked, microwaved)

    - most common burner use (front or back)
    - mean frequency of range hood fan use
- Primary summary statistics of pollution events ($PM_{2.5}$, NO_2 and CO_2). We will aggregate across all cooking-linked pollution events and all non-cooking-linked pollution events, both in the pre-intervention and post-intervention periods.
    - Means and SD of the:
        - prominences
        - pollutant concentrations above baseline
    - Number and fraction of pollutant events occurring without cooking burner use, including which have hood use
    - Total magnitude of emissions (total aggregated pollutant during cooking-linked pollution events divided by burner minutes)
- Secondary statistics
    - Time-integrated NO_2 and NO_X by Ogawa passive sampler
    - Time-integrated $PM_{2.5}$ by eLichens (if discernible from the data)

Outcome Definitions

At the baseline phone visit, health information was collected using the validated Isaac Questionnaires developed for use with 6-7 year-olds.(25) The core Isaac questionnaire contains 8 questions on asthma history, 6 on allergy history and 7 regarding history of eczema.

Children with asthma (with their parent) completed an asthma control test prior to or during each study visit. Study staff were trained on administering spirometry and FeNO tests for all participants, which occurred at follow up visits initially and later in the study was added to the initial visit to give children time to learn the techniques. These measures of respiratory status will be continuous variables.

The study used an EasyOne Spirometer, which was brought to the homes. Under the supervision of study staff, each child performed spirometry to complete three acceptable efforts (maximum eight attempts), in accordance with standard ATS/ERS performance criteria.(27) Each effort will be graded by a trained physician, in accordance with the 2019 ATS/ERS guidelines. If there are two or more efforts with acceptable quality and reproducible FEV_1 and FVC (each $\leq$ 0.15L), the best FEV_1 and best FVC will be used. If two or more efforts had usable quality and reproducible FEV_1, but unacceptable FVC, only the best FEV_1 will be used and FVC will be set to missing.

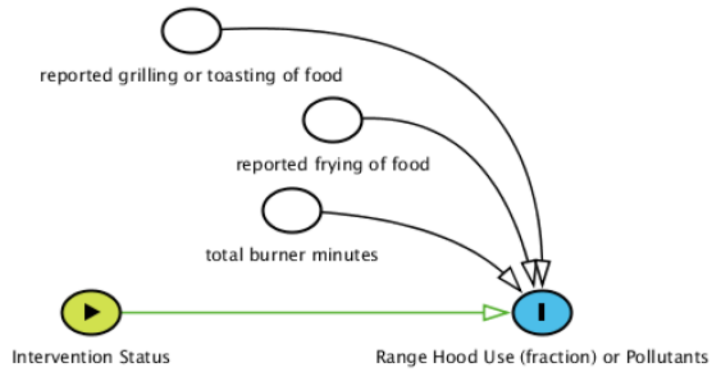
Each child also had a FeNO measured in accordance with ATS/ERS criteria,(28) by blowing a steady sustained exhalation into a NIOX Vero device to complete 2 measurements with exhalation duration greater than 10 seconds, and within 10% of each other (maximum eight attempts). If there are two measurements within 10%, this testing session will be graded as acceptable and the larger value will be used. If there are one or multiple FeNO measurements of adequate duration (but none within 10%), this testing session will be graded as usable and these values will be averaged to create the FeNO value for analysis.

Analysis Methods

Determine if a cooking ventilation intervention (not receiving anything, receiving education only, or receiving education and range hood replacement) is associated with changes in (1) ventilation, (2) front vs back burner use, (3) household particulate matter, (4) NO_2 and (5) CO_2 levels. We will use a

Wilcoxon Signed Rank test to compare the pre-intervention data to the post-intervention data to assess the relationship between intervention status and the following variables.

1. Range hood use, characterized 4 ways:
    a. fraction of cooking events with any range hood use (this will be primarily estimated using the vane anemometer data, but can also be assessed with self-reported (weekly) range hood use frequency
    b. fraction of cooking events with ≥80% range hood use
    c. fraction of cooking-linked pollutant events with substantial range hood use
    d. mean delay between start of cooking and start of range hood use



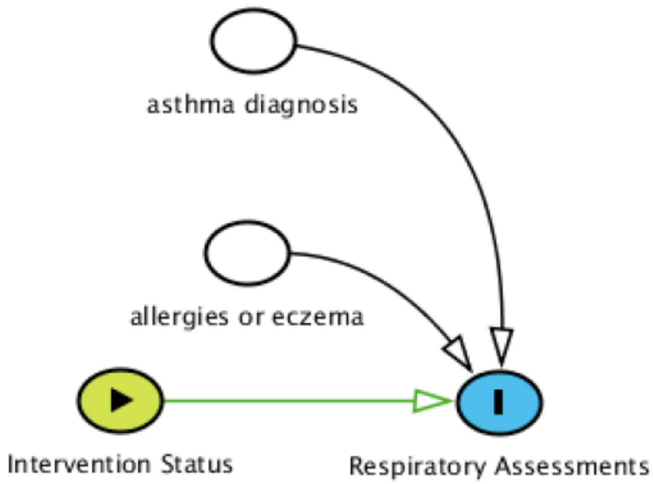
2. Front versus back burner use, characterized 2 ways
    a. fraction of total burner minutes that reflects usage of the front burners and fraction that reflects usage of the back burners.
    b. self-reported most common burner use
3. $PM_{2.5}$ levels, characterized four ways:
    a. mean prominence (increase above baseline) of $PM_{2.5}$ during cooking-linked pollutant events
    b. mean ug/m3-min $PM_{2.5}$ aggregated during cooking-linked pollutant events
    c. total magnitude of emissions (aggregated $PM_{2.5}$ during cooking-linked pollutant events divided by burner minutes)
    d. time-integrated $PM_{2.5}$ (if discernible from the data)
4. Nitrogen dioxide levels, characterized four ways:
    a. mean prominence (increase above baseline) of NO_2 during cooking-linked pollutant events
    b. mean ppb-min NO_2 aggregated during cooking-linked pollutant events
    c. total magnitude of emissions (aggregated NO_2 during cooking-linked pollutant events divided by burner minutes)
    d. time-integrated NO_2
5. Mean prominence of carbon dioxide during cooking-linked pollutant events

In general, the covariates that will vary over time (eg. total burner use minutes) are not related to the intervention status and thus do not need to be controlled in order to prevent bias. Given the small final sample size, these will not be adjusted in the analysis (but original DAG maintained for completeness).

Legend for the Three Directed Acyclic Graph Figures

- exposure
- outcome
- ancestor of exposure
- ancestor of outcome
- ancestor of exposure *and* outcome
- adjusted variable
- unobserved (latent)
- other variable
- causal path
- biasing path

Determine if a cooking ventilation intervention is associated with changes in airways inflammation and lung function in children

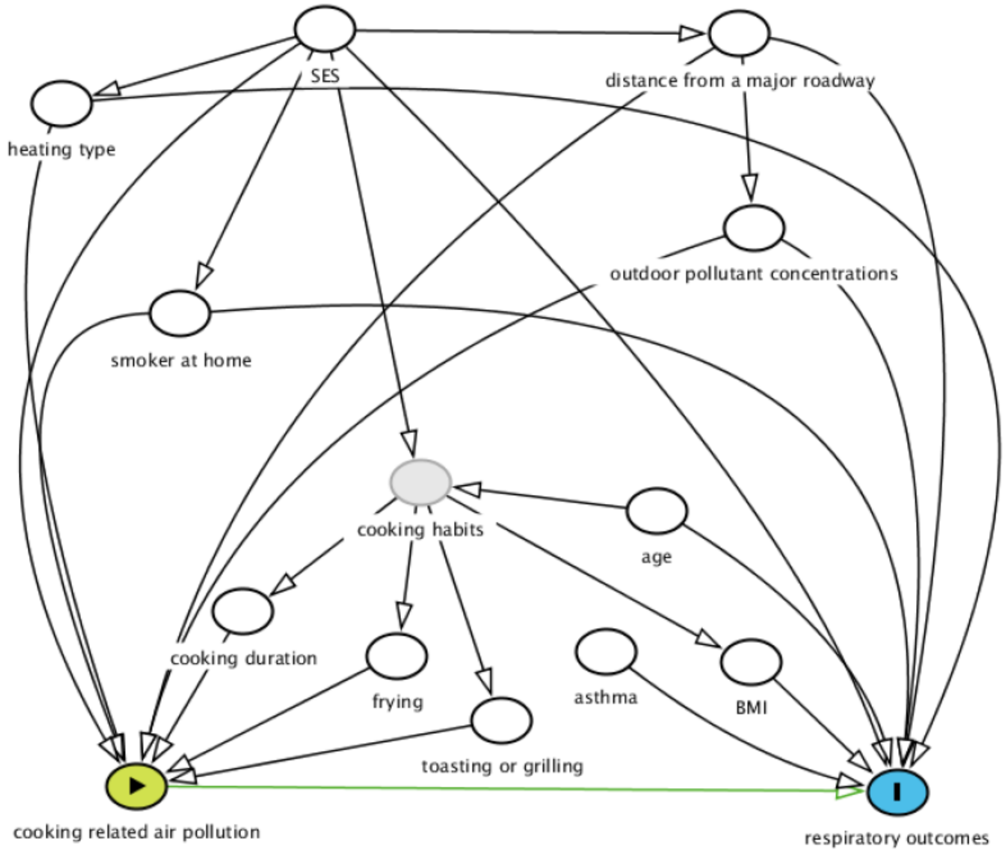


We will also use Wilcoxon Signed Rank tests to compare the pre-intervention data to the post-intervention data to assess the relationship between intervention status and the following respiratory outcome variables. Given the small final sample size, these will not be adjusted in the analysis (but original DAG maintained for completeness). Though we had initially intended to control for the number of visits with our team, only 2 of 14 participants had one additional visit with our team prior to intervention, meaning that there is not enough variability in this variable to control for it.

1. Airways inflammation based on measured fractional exhaled nitric oxide (FeNO)
2. Airways obstruction based on measured forced expiratory volume in 1 second (FEV1) and the ratio of FEV1 to the Forced Vital Capacity (FEV1/FVC)
3. Self-reported lung health as measured by the asthma control test score (ACT)

Assess whether cooking-related pollution is associated with respiratory outcomes in children.
We will use a simple visualization to do a rough exploration of these data. Due to the small sample size, we would be unable to control for all necessary confounders. But the original DAG is maintained for completeness
These will occur for the following sets of variables:

- Respiratory Health: ACT score, FeNO and FEV1
- Pollutants: $PM_{2.5}$, NO_2 and CO_2, using measures a,b, and d listed above



Covariates
see analysis methods above for each set of models

Sensitivity Analyses
Because there were some variations in the data collection protocol due to the pandemic, different households had different numbers of visits before/after receiving the intervention. As a sensitivity analysis, we will explore whether there were patterns related to the number of visits with our team prior to the intervention.

Two households had range hood replacements indicated that did not receive them (one due to landlord refusal, another due to cost/feasibility limits for the contractor). We will perform a sensitivity analysis removing the two households that did not receive the indicated replacements.

In addition, the study was initially designed to be in the homes of children with asthma. Four of the households that completed the study had a child with asthma, meaning they may have been particularly motivated to make changes. We will observe all our calculated statistics in the subgroup of households that have an asthmatic child, to look for any suggestion of a larger effect in this group (recognizing that this group may simply be too small to see anything).

Missing Data

For exposure data that are correlated (e.g. different burner thermocouple data) regression modeled predicted values will be used to estimate missing values. For all other covariates and exposure variables, a complete case analysis was used due to sample size limitations that would preclude us from performing multiple imputation as initially planned.

Statistical Software

The R statistical programming language, implemented in R studio and with associated R packages, will be used to perform all epidemiologic analyses. Cooking, range hood and pollution data were processed using python.

References

1. Hiemstra B, Keus F, Wetterslev J, Gluud C, van der Horst ICC. DEBATE-statistical analysis plans for observational studies. BMC Med Res Methodol. 2019 Dec;19(1):233.

2. Klepeis NE, Nelson WC, Ott WR, Robinson JP, Tsang AM, Switzer P, et al. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J Expo Sci Environ Epidemiol. 2001 Jul;11(3):231–52.

3. Vardoulakis S, Giagloglou E, Steinle S, Davis A, Sleeuwenhoek A, Galea KS, et al. Indoor Exposure to Selected Air Pollutants in the Home Environment: A Systematic Review. Int J Environ Res Public Health. 2020 Dec 2;17(23).

4. Singer BC, Pass RZ, Delp WW, Lorenzetti DM, Maddalena RL. Pollutant concentrations and emission rates from natural gas cooking burners without and with range hood exhaust in nine California homes. Building and Environment. 2017 Sep 1;122:215–29.

5. Habre R, Coull B, Moshier E, Godbold J, Grunin A, Nath A, et al. Sources of indoor air pollution in New York City residences of asthmatic children. J Expo Sci Environ Epidemiol. 2014 May;24(3):269–78.

6. Baxter LK, Clougherty JE, Laden F, Levy JI. Predictors of concentrations of nitrogen dioxide, fine particulate matter, and particle constituents inside of lower socioeconomic status urban homes. J Expo Sci Environ Epidemiol. 2007 Aug;17(5):433–44.

7. Heinrich J. Influence of indoor factors in dwellings on the development of childhood asthma. International Journal of Hygiene and Environmental Health. 2011 Jan;214(1):1–25.

8. Klug V, Lobscheid A, Singer B. Cooking Appliance Use in California Homes – Data Collected from a Web-Based Survey. Berkeley, CA: Lawrence Berkeley National Laboratory; 2011.

9. Anderson JO, Thundiyil JG, Stolbach A. Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. J Med Toxicol. 2012 Jun 1;8(2):166–75.

10. Kelly FJ, Fussell JC. Air pollution and public health: emerging hazards and improved understanding of risk. Environ Geochem Health. 2015 Aug 1;37(4):631–49.

11. Lin W, Brunekreef B, Gehring U. Meta-analysis of the effects of indoor nitrogen dioxide and gas cooking on asthma and wheeze in children. International Journal of Epidemiology. 2013 Dec 1;42(6):1724–37.

12. Vieira SE, Stein RT, Ferraro AA, Pastro LD, Pedro SSC, Lemos M, et al. Urban Air Pollutants Are Significant Risk Factors for Asthma and Pneumonia in Children: The Influence of Location on the Measurement of Pollutants☆. Arch Bronconeumol. :7.

13. Sharma HP, Hansel NN, Matsui E, Diette GB, Eggleston P, Breysse P. Indoor Environmental Influences on Children's Asthma. Pediatric Clinics of North America. 2007 Feb;54(1):103–20.

14. Behrens T, Maziak W, Weiland SK, Rzehak P, Siebert E, Keil U. Symptoms of Asthma and the Home Environment. The ISAAC I and III Cross-Sectional Surveys in Münster, Germany. Int Arch Allergy Immunol. 2005;137(1):53–61.

15. Paulin LM, Williams D 'Ann L, Peng R, Diette GB, McCormack MC, Breysse P, et al. 24-h Nitrogen dioxide concentration is associated with cooking behaviors and an increase in rescue medication use in children with asthma. Environmental Research. 2017 Nov;159:118–23.

16. Hansel NN, Breysse PN, McCormack MC, Matsui EC, Curtin-Brosnan J, Williams DL, et al. A Longitudinal Study of Indoor Nitrogen Dioxide Levels and Respiratory Symptoms in Inner-City Children with Asthma. Environmental Health Perspectives. 2008 Oct;116(10):1428–32.

17. Coker ES, Smit E, Harding AK, Molitor J, Kile ML. A cross sectional analysis of behaviors related to operating gas stoves and pneumonia in U.S. children under the age of 5. BMC Public Health. 2015 Dec;15(1):77.

18. Kile ML, Coker ES, Smit E, Sudakin D, Molitor J, Harding AK. A cross-sectional study of the association between ventilation of gas stoves and chronic respiratory illness in U.S. children enrolled in NHANESIII. Environmental Health. 2014 Sep 2;13(1):71.

19. Holm SM, Balmes J, Gillette D, Hartin K, Seto E, Lindeman D, et al. Cooking behaviors are related to household particulate matter exposure in children with asthma in the urban East Bay Area of Northern California. PLoS One [Internet]. 2018 Jun 6 [cited 2021 Mar 5];13(6). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5991365/

20. California Energy Commission. 2008 Building Energy Efficiency Standards [Internet]. [cited 2018 Feb 1]. Available from: https://www.energy.ca.gov/2008publications/CEC-400-2008-001/CEC-400-2008-001-CMF.PDF

21. Belanger K, Holford TR, Gent JF, Hill ME, Kezik JM, Leaderer BP. Household Levels of Nitrogen Dioxide and Pediatric Asthma Severity: Epidemiology. 2013 Mar;24(2):320–30.

22. Garcia Algar O, Pichini S, Basagana X, Puig C, Vall O, Torrent M, et al. Concentrations and determinants of NO2 in homes of Ashford, UK and Barcelona and Menorca, Spain. Indoor Air. 2004 Aug;14(4):298–304.

23. Mar TF, Jansen K, Shepherd K, Lumley T, Larson TV, Koenig JQ. Exhaled Nitric Oxide in Children with Asthma and Short-Term $PM_{2.5}$ Exposure in Seattle. Environ Health Perspect. 2005 Dec;113(12):1791–4.

24. Stanojevic S, Wade A, Stocks J, Hankinson J, Coates AL, Pan H, et al. Reference Ranges for Spirometry Across All Ages: A New Approach. Am J Respir Crit Care Med. 2008 Feb;177(3):253–60.

25. Wong GWK, Brunekreef B, Ellwood P, Anderson HR, Asher MI, Crane J, et al. Cooking fuels and prevalence of asthma: a global analysis of phase three of the International Study of Asthma and Allergies in Childhood (ISAAC). Lancet Respir Med. 2013 Jul;1(5):386–94.

26. Cabana MD, Kunselman SJ, Nyenhuis SM, Wechsler ME. Researching asthma across the ages: Insights from the National Heart, Lung, and Blood Institute's Asthma Network. Journal of Allergy and Clinical Immunology. 2014 Jan;133(1):27–33.

27. Graham BL, Steenbruggen I, Miller MR, Barjaktarevic IZ, Cooper BG, Hall GL, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. Am J Respir Crit Care Med. 2019 Oct 15;200(8):e70–88.

28. ATS/ERS Recommendations for Standardized Procedures for the Online and Offline Measurement of Exhaled Lower Respiratory Nitric Oxide and Nasal Nitric Oxide, 2005. Am J Respir Crit Care Med. 2005 Apr 15;171(8):912–30.

29. Harel O, Mitchell EM, Perkins NJ, Cole SR, Tchetgen Tchetgen EJ, Sun B, et al. Multiple Imputation for Incomplete Data in Epidemiologic Studies. American Journal of Epidemiology. 2018 Mar 1;187(3):576–84.