Environment International 159 (2022) 107021

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Multinational prediction of household and personal exposure to fine particulate matter ($PM_{2.5}$) in the PURE cohort study

Matthew Shupler[a,b,*], Perry Hystad[c], Aaron Birch[a], Yen Li Chu[a], Matthew Jeronimo[a], Daniel Miller-Lionberg[d], Paul Gustafson[e], Sumathy Rangarajan[f], Maha Mustaha[f], Laura Heenan[f], Pamela Seron[f], Fernando Lanas[g], Fairuz Cazor[g], Maria Jose Oliveros[g], Patricio Lopez-Jaramillo[h], Paul A. Camacho[i], Johnna Otero[j], Maritza Perez[k], Karen Yeates[l,m], Nicola West[k], Tatenda Ncube[m], Brian Ncube[m], Jephat Chifamba[m], Rita Yusuf[n], Afreen Khan[n], Zhiguang Liu[o], Shutong Wu[p], Li Wei[p], Lap Ah Tse[q], Deepa Mohan[r], Parthiban Kumar[r], Rajeev Gupta[s], Indu Mohan[t], KG Jayachitra[u], Prem K. Mony[u], Kamala Rammohan[v], Sanjeev Nair[v], P.V.M. Lakshmi[w], Vivek Sagar[w], Rehman Khawaja[x], Romaina Iqbal[x], Khawar Kazmi[x], Salim Yusuf[f], Michael Brauer[a], on behalf of the PURE-AIR study investigators

[a] *School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada*
[b] *Department of Public Health, Policy and Systems, University of Liverpool, Liverpool, United Kingdom*
[c] *College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, United States*
[d] *Access Sensors Technologies, Fort Collins, CO, United States*
[e] *Department of Statistics, University of British Columbia, Vancouver, British Columbia, Canada*
[f] *Population Health Research Institute, Hamilton Health Sciences, McMaster University, Hamilton, Ontario, Canada*
[g] *Universidad de La Frontera, Temuco, Chile*
[h] *Universidad de Santander (UDES), Bucaramanga, Colombia*
[i] *Fundación Oftalmológica de Santander (FOSCAL), Floridablanca, Colombia*
[j] *Universidad Militar Nueva Granada, Bogota, Colombia*
[k] *Pamoja Tunaweza Research Centre, Moshi, Tanzania*
[l] *Department of Medicine, Queen's University, Kingston, Ontario, Canada*
[m] *Department of Biomedical Sciences, University of Zimbabwe, Harare, Zimbabwe*
[n] *School of Life Sciences, Independent University, Dhaka, Bangladesh*
[o] *Beijing An Zhen Hospital of the Capital University of Medical Sciences, China*
[p] *Medical Research & Biometrics Center, National Center for Cardiovascular Diseases, Fuwai Hospital, Chinese Academy of Medical Sciences, China*
[q] *Jockey Club School of Public Health and Primary Care, the Chinese University of Hong Kong, HKSAR, China*
[r] *Madras Diabetes Research Foundation, Chennai, India*
[s] *Eternal Heart Care Centre & Research Institute, Jaipur, India*
[t] *Mahatma Gandhi University of Medical Sciences and Technology, Jaipur, India*
[u] *St. John's Medical College & Research Institute, Bangalore, India*
[v] *Health Action By People, Government Medical College, Trivandrum, India*
[w] *Post Graduate Institute of Medical Education and Research, Chandigarh, India*
[x] *Department of Community Health Science, Aga Khan University Hospital, Karachi, Pakistan*

ARTICLE INFO

Handling Editor: Xavier Querol

Keywords:
Household air pollution
$PM_{2.5}$
Kitchen concentrations
Personal exposures

ABSTRACT

Introduction: Use of polluting cooking fuels generates household air pollution (HAP) containing health-damaging levels of fine particulate matter ($PM_{2.5}$). Many global epidemiological studies rely on categorical HAP exposure indicators, which are poor surrogates of measured $PM_{2.5}$ levels. To quantitatively characterize HAP levels on a large scale, a multinational measurement campaign was leveraged to develop household and personal $PM_{2.5}$ exposure models.

Methods: The Prospective Urban and Rural Epidemiology (PURE)-AIR study included 48-hour monitoring of $PM_{2.5}$ kitchen concentrations (n = 2,365) and male and/or female $PM_{2.5}$ exposure monitoring (n = 910) in a

* Corresponding author.
E-mail address: mshupler@mail.ubc.ca (M. Shupler).

https://doi.org/10.1016/j.envint.2021.107021
Received 23 August 2021; Received in revised form 30 November 2021; Accepted 2 December 2021
Available online 13 December 2021
0160-4120/

Predictive modeling
Bayesian hierarchical modeling

subset of households in Bangladesh, Chile, China, Colombia, India, Pakistan, Tanzania and Zimbabwe. PURE-AIR measurements were combined with survey data on cooking environment characteristics in hierarchical Bayesian log-linear regression models. Model performance was evaluated using leave-one-out cross validation. Predictive models were applied to survey data from the larger PURE cohort (22,480 households; 33,554 individuals) to quantitatively estimate $PM_{2.5}$ exposures.

Results: The final models explained half (R^2 = 54%) of the variation in kitchen $PM_{2.5}$ measurements (root mean square error (RMSE) (log scale):2.22) and personal measurements (R^2 = 48%; RMSE (log scale):2.08). Primary cooking fuel type, heating fuel type, country and season were highly predictive of $PM_{2.5}$ kitchen concentrations. Average national $PM_{2.5}$ kitchen concentrations varied nearly 3-fold among households primarily cooking with gas (20 μg/m^3 (Chile); 55 μg/m^3 (China)) and 12-fold among households primarily cooking with wood (36 μg/m^3 (Chile)); 427 μg/m^3 (Pakistan)). Average $PM_{2.5}$ kitchen concentration, heating fuel type, season and secondhand smoke exposure were significant predictors of personal exposures. Modeled average $PM_{2.5}$ female exposures were lower than male exposures in upper-middle/high-income countries (India, China, Colombia, Chile).

Conclusion: Using survey data to estimate $PM_{2.5}$ exposures on a multinational scale can cost-effectively scale up quantitative HAP measurements for disease burden assessments. The modeled $PM_{2.5}$ exposures can be used in future epidemiological studies and inform policies targeting HAP reduction.

1. Introduction

Approximately 3.8 billion people residing in low- and middle-income countries (LMICs) use polluting cooking fuels (e.g. wood, charcoal, animal dung, coal) in traditional stoves (e.g. open fires, mud stoves) (Health Effects Institute, 2020). Exposure to household air pollution (HAP) from incomplete combustion of polluting cooking fuels in inefficient stoves has several adverse health and environmental consequences. In epidemiological studies, exposure to elevated concentrations of fine particulate matter ($PM_{2.5}$), a pollutant of primary health concern found in HAP, has been associated with respiratory infections in children (Bates et al., 2013; Ezzati and Kammen, 2001; Upadhyay et al., 2015), lung cancer (Kurmi et al., 2012), chronic obstructive pulmonary disease (COPD) (Kurmi et al., 2010; Salvi and Barnes, 2010), cataracts (Pokhrel, 2004), adverse pregnancy outcomes (Amegah et al., 2014; Thompson et al., 2011; Alexander et al., 2018), hypertension (Alexander et al., 2017; Arku et al., 2018; Baumgartner et al., 2014; Baumgartner et al., 2011; Burroughs Pena et al., 2015; Clark et al., 2013; Norris et al., 2016) and cardiovascular diseases (CVD) including ischemic heart disease (IHD) and stroke (Kephart et al., 2020; Alam et al., 2012; Yu et al., 2018). HAP contributes up to one-third of all global anthropogenic emissions of black carbon (Rehman et al., 2011; Bond et al., 2013; Grieshop et al., 2011), a component of $PM_{2.5}$ that has the second largest radiative forcing, behind only carbon dioxide (Grieshop et al., 2011; Ramanathan and Carmichael, 2008). HAP is also a major source of ambient $PM_{2.5}$ pollution (Liu et al., 2016; Chafe et al., 2014).

1.1. Quantitative household AIR pollution exposure modeling

Quantitative $PM_{2.5}$ exposure measurements are needed for more accurate assessment of the health risks from cooking with polluting fuels (Burnett and Cohen, 2020; Burnett et al., 2018), but require significant resource, time and financial investment, precluding large-scale HAP monitoring in many LMICs. Thus, there is substantial uncertainty in the exposure–response relationship in the range of $PM_{2.5}$ levels typically found in HAP for various diseases. To reduce HAP exposure misclassification from use of categorical exposure indicators, such as primary cooking fuel type or 'clean versus polluting' fuel (Smith et al., 2014), quantitative exposure estimation of $PM_{2.5}$ levels is needed to facilitate larger-scale HAP exposure assessment with reduced air monitoring (Balakrishnan et al., 2013;12(1):77.; Baumgartner et al., 2011). HAP predictive models leverage the association of $PM_{2.5}$ measurements with household characteristics that are more easily characterized via surveys. This requires strategically collecting survey data on factors that affect household $PM_{2.5}$ concentrations. Previous HAP predictive models have linked survey data on the cooking environment (e.g. primary cooking fuel type, heating fuel, ventilation) and socioeconomic status (SES) (e.g. income, education) to $PM_{2.5}$ measurements at a national level in several countries, including India and China (Carter et al., 2016; Gurley et al., 2013; Jin et al., 2005; Massey et al., 2012; Ni et al., 2016).

1.2. Modeling household AIR pollution levels on a global scale

With limited existing multinational HAP measurement studies, global HAP risk assessments such as the Global Burden of Disease (GBD) study have historically relied on a compilation of available $PM_{2.5}$ exposure measurements from published studies within the *WHO Global HAP database* (https://www.who.int/data/gho/data/themes/air-pollution/hap-measurement-db) (Shupler et al., 2018). A previous HAP modeling study aggregated $PM_{2.5}$ kitchen concentration and female exposure measurements, descriptive survey data (primary fuel type, season (wet vs. dry)) and an index of country-level SES from 44 published measurement studies conducted from 1996 to 2016 and available data in the WHO database to predict country and primary fuel-specific $PM_{2.5}$ levels (Shupler et al., 2018). A constraint of this global modeling study was the limited number of households (range: 2–470 households; median = 17) in each study and minimal predictors available for modeling due to differential reporting of population characteristics in the publications. Studies included in the WHO database also have diverse study designs, monitoring technology, analytic methods and measurement periods, which introduced measurement bias into the quantitative $PM_{2.5}$ model estimates.

This study uses household survey data and personal and kitchen $PM_{2.5}$ measurements from a single study, the Prospective Urban and Rural Epidemiology (PURE)-AIR study. PURE-AIR included 48-hour air monitoring of kitchen concentrations among approximately 2,500 households and simultaneous male and/or female exposure monitoring (among a subset of households; n ~ 1,000) across 120 rural communities in eight countries (Shupler et al., 2020). The PURE-AIR measurements presented in Shupler et al (Shupler et al., 2020) were leveraged in this study to achieve several additional aims that were previously infeasible due to logistical challenges associated with extensive HAP exposure assessments. The specific aims of this analysis were: (1) to use machine learning methods to understand the most important drivers of $PM_{2.5}$ exposure variations due to HAP on a global scale and (2) to develop multinational predictive models of $PM_{2.5}$ kitchen concentrations and personal exposure measurements from PURE-AIR (~2,500 households) to scale up HAP exposures to the larger PURE cohort (~25,000 households). The results from objective 1 can inform data to be prioritized for collection in future national global health surveys (e.g. WHO Harmonized Survey, Demographic Health surveys) when aiming to explain multinational variation in HAP exposures. The quantitative $PM_{2.5}$ estimates generated from objective 2 can uncover the quantitative effect (and the variability around the effect estimates) of different factors on $PM_{2.5}$ exposures and will have utility in future epidemiological studies using health outcome data from the PURE cohort.

2. Methods

2.1. Study design

The PURE study, initiated in 2003, is a multinational cohort designed to identify risk factors for cardiovascular disease across low-, middle-, and high-income counties (LMICs). Approximately 190,000 participants have been enrolled from around 800 rural and urban communities within 'sub-national regions' (defined as urban centers around which rural and urban communities were clustered) in 27 countries (Teo et al., 2009). Communities represent neighborhoods in urban areas and villages in rural areas. Within PURE communities, participants are representative of the age and sex distribution of adults aged 35–70 (Corsi et al., 2013). At baseline, all households completed a PURE Household questionnaire that contained questions related to household energy usage, including primary cooking fuel type, cooking location (indoors/outdoors), ventilation (presence of windows, chimney) and heating fuel type. The male and female heads of household also completed PURE individual questionnaires regarding their socioeconomic status (SES) (e.g. education, occupation).

The PURE-AIR study, nested within the PURE cohort, integrated 48-hour kitchen $PM_{2.5}$ monitoring among a subset of 2,541 PURE households within 120 rural communities of eight LMICs (Bangladesh, Chile, China, Colombia, India, Pakistan, Tanzania, Zimbabwe) where cooking with polluting fuels exceeded 10% prevalence at baseline (Shupler et al., 2020; Arku et al., 2018). In a subset of 20% of PURE-AIR households, 48-hour male and/or female personal monitoring was conducted alongside household monitoring among 951 participants. Stratified random sampling was used to select PURE-AIR households within each rural community, proportional to baseline percentage of primary cooking fuel type, with polluting fuels oversampled to capture a wide-ranging exposure distribution (Shupler et al., 2020). Prior to and after monitoring, a PURE-AIR household survey was administered, which contained cooking environment questions identical to a baseline PURE household survey (see Supplement of Arku et al. 2018 for PURE baseline survey), and additional questions on stove type, secondary cooking fuel type, daily cooking time, time spent in the cooking area and years using the current primary cooking fuel. Detailed information on the PURE-AIR sampling strategy and measurement protocol is documented elsewhere (Shupler et al., 2020; Arku et al., 2018).

2.2. Variable importance

Variables were selected from (1) the PURE-AIR survey and (2) the PURE baseline household and individual surveys that were hypothesized *a priori* to be associated with HAP kitchen concentrations and personal exposures. A unique set of variables for kitchen concentration and personal exposure models were selected with two separate goals: (1) to use PURE-AIR survey data to determine the maximum predictive power of cooking environment variables and (2) to use cooking environment variables available in PURE baseline surveys to develop the most accurate quantitative exposure coefficients to apply to the PURE cohort (for whom only PURE baseline survey was available).

To evaluate cooking environment characteristics that explained the largest percent of between- and within-country variability in $PM_{2.5}$ kitchen concentrations and personal exposure among rural PURE communities, machine learning models, developed using random forests via the *randomForest* package in R, were built for all measurements as well as separate models for China, India, South America (Chile/Colombia), other South Asia (Bangladesh/Pakistan) and Africa (Tanzania/Zimbabwe). For each model, variable importance was evaluated using the *interpretable machine learning (iml)* package in R (Molnar, 2021; Fisher et al., 2019), which ranked model variables according to the lowest mean absolute error (MAE). Before running all models, continuous variables were grouped into tertiles when possible to avoid bias in variable importance calculations (see Table S3 for a list of variables); a similar approach was followed in a previous modeling study of primary cooking fuel switching in PURE (Shupler et al., 2019). Due to small size, some heating fuel types were condensed to increase the power of the analysis; heated coal beds ("kang"), commonly used as a heating fuel in Northern China, were grouped with 'coal open fires', and households reporting using animal dung and agricultural waste in open fires for heating were grouped with wood to form a 'biomass open fire' heating category.

2.3. Bayesian modeling

While machine learning methods are advantageous to this study in their ability to find the best relationship between predictors and $PM_{2.5}$ concentrations, they are limited by their lack of interpretability (Rudin, 2019). This is because machine learning models are a 'black box' of complicated functions of variables, rendering it infeasible to understand how the predictors are jointly related to each other. As such, machine learning results cannot be easily applied to external datasets for prediction. Therefore, Bayesian hierarchical models were built using the PURE-AIR sample to generate quantitative $PM_{2.5}$ exposure coefficients and variance around the estimates for predicting quantitative $PM_{2.5}$ exposures for the entire PURE cohort.

A Bayesian approach also had the advantage of application of primary cooking fuel-specific 'priors' to the predictive models, using data from a previous global WHO Global Database modeling study (Shupler et al., 2018). The Bayesian log-linear regression models accounted for the clustered sampling within PURE-AIR (households nested in communities nested in sub-national regions), with fixed effects added in the order of their importance determined from the random forest modeling. Households (n = 19, 0.8% of sample) and participants (n = 9, 1% of sample) sampled in Tanzania were excluded from Bayesian modeling of $PM_{2.5}$ kitchen concentrations due to low sample size that precluded model convergence. Separate Bayesian models were built for $PM_{2.5}$ kitchen concentrations (n = 2,384 out of 2,541 (6% missing data) and $PM_{2.5}$ personal exposures (n = 910 out of 951 (4% missing)) using only variables available in PURE baseline surveys.

Using the *brms* package in R (Bürkner, 2017), Bayesian models were run with two chains and model convergence was monitored via visual chain inspection. A total of 8,000 posterior estimates were retained for use in variance calculations for the modeled exposures. Model selection was based on simultaneously optimizing the coefficient of determination (R^2) and reducing the leave-one-out information criterion (LOOIC) (Gelman et al., 2017). Given differences in sex-specific average 48-hour $PM_{2.5}$ exposures between countries detected descriptively from PURE-AIR measurements (Shupler et al., 2020), a sex*country interaction was evaluated in the personal exposure model. Bayesian model validation was conducted using leave-one-out (LOO) cross validation via an approximation technique (Pareto smoothed importance sampling) (Vehtari et al., 2016).

Bayesian models included 'power priors' of cooking fuel-specific $PM_{2.5}$ concentrations and exposures from the WHO Global Database modeling (Shupler et al., 2018). Priors were normally distributed and centered at the mean $PM_{2.5}$ level obtained from the previous global HAP models (see Table 1 in Shupler et al., 2018). Power priors (Table 1) accounted for the different study designs of publications included in the WHO Global Database modeling and PURE-AIR by employing a hyperparameter to quantify the heterogeneity (Ibrahim and Chen, 2000).

2.4. Predicting fine particulate matter exposures for PURE cohort

We applied coefficients from the final Bayesian models to PURE baseline survey data to quantitatively estimate baseline $PM_{2.5}$ kitchen

Table 1
Power priors used in Bayesian hierarchical models of PURE-AIR measurements.

Model type	Cooking fuel type	Mean $PM_{2.5}$ estimate (μg/m^3)[1]	Power prior (log scale) (mean, standard deviation)[2]
Kitchen concentration	Gas	104	~Normal(4.65, 0.48)
	Electricity	104	~Normal(4.65, 0.48)
	Wood	395	~Normal(5.98, 0.48)
	Coal	319	~Normal(5.77, 0.48)
	Animal dung	958	~Normal(6.87, 0.48)
Personal exposure	Gas	42	~Normal(3.75, 0.48)
	Electricity	42	~Normal(3.75, 0.48)
	Wood	161	~Normal(5.08, 0.48)
	Coal	130	~Normal(4.87, 0.48)
	Animal dung	391	~Normal(5.97, 0.48)

1. Estimates obtained from Table 1 of WHO Global Database modeling study (Shupler et al., 2018).
2. Standard deviation of 0.48 is a power prior obtained by multiplying largest fuel-specific standard deviation from WHO Global Database modeling study (0.06) by 'fudge factor' of 8 to account for heterogeneity in study design between WHO Global Database measurements and PURE-AIR measurements.

and personal exposures in the PURE cohort. We restricted prediction to 22,480 households and 33,554 individuals in rural communities of eight PURE countries with > 20% polluting cooking fuel use at baseline (see Fig. 2 in Arku et al. 2018 for baseline prevalence of primary cooking fuel use in each country).

To estimate average annual $PM_{2.5}$ concentrations based on a single 48-hour HAP measurement, all countries were assumed to have equal (50/50) wet (April-September; summer in northern hemisphere countries) and dry (October-March; winter) seasons, as done in the previous WHO global HAP modeling study (Shupler et al., 2018). 'Prediction intervals' around point estimates were calculated using the posterior distribution from the Bayesian models.

2.5. Comparing PURE-AIR measurements to ambient air pollution concentrations

Due to lower anticipated $PM_{2.5}$ emissions from clean fuels (Shen et al., 2018) and the infiltration of outdoor air pollution to the indoor environment (Krebs et al., 2021), the utility of using average 48-hour $PM_{2.5}$ kitchen concentrations in households cooking exclusively with clean fuels as a surrogate measure of ambient air pollution concentrations was assessed. Spearman correlation coefficients (r) were calculated to quantify the relationship between modeled average annual ambient $PM_{2.5}$ levels (at $0.1^{\circ}\times0.1^{\circ}$ resolution; approximately 11 × 11 km resolution at the equator) in 2018 (spanning the time of PURE-AIR sampling). The ambient air pollution estimates were derived from a Bayesian modeling study that incorporated satellite and ground $PM_{2.5}$ measurements (Shaddick et al., 2018). These estimates were map-matched (via GPS coordinates) at a sub-national regional level to average kitchen and personal exposure measurements from PURE-AIR. Due to a lack of PURE-AIR measurements among households using electric stoves in most countries, only kitchen measurements from households cooking with gas were compared to modeled average annual ambient $PM_{2.5}$ concentrations. Due to a focus on cooking environmental predictors of $PM_{2.5}$ concentrations, a lack of resolution in average annual ambient air pollution concentrations at a household-level within PURE-AIR communities precluded its inclusion as a predictor in PURE-AIR modeling.

All statistical analyses were conducted in R version 3.5.1 (R Core Team. R, 2017). Initial and ongoing ethics approvals for the PURE cohort were obtained from the Institutional Review Board at McMaster University, Hamilton, Ontario and from each PURE study country using local Institutional Review Boards. Ethics approval for the PURE-AIR study was obtained from the University of British Columbia's Behavioral Research Ethics Board.

3. Results

3.1. Characteristics of PURE baseline sample

Due to recruitment of PURE households in different waves, 16% (n = 4,932) of households recruited in the first wave in India did not have information on heating fuel type. Due to the importance of heating fuel type as a $PM_{2.5}$ concentration determinant, these households were excluded from the modeling to ensure maximum predictive accuracy. Additionally, PURE households using kerosene as their primary cooking fuel at baseline (n = 363, 1% of total sample) were excluded as PURE-AIR did not include monitoring of households cooking with kerosene. The final analytic sample was comprised of 22,480 rural households and 33,554 individuals living in the households. A sensitivity analysis that examined characteristics between the full and analytic baseline sample revealed that households in India with missing data on heating fuel type at baseline tended to be lower SES (lower income tertile, higher percent of income spent on food) than the average household in India (Table S10). Thus, predicted average $PM_{2.5}$ kitchen concentrations among rural communities in India at PURE baseline may be slightly attenuated due to an inverse association between higher household SES and lower average $PM_{2.5}$ kitchen concentrations found among PURE-AIR households (Shupler et al., 2020).

PURE baseline households were predominantly from China (49%; n = 15,163) and India (29%; n = 9,051). The most frequently used primary cooking fuels at PURE baseline were wood (37%; n = 11,334) and gas (24%; n = 7,385) (Table S1). Coal (19%; n = 5,722) was almost exclusively used for cooking in China. Approximately two-thirds of households heated their homes during the cold season; a biomass open fire was the most common heating method (outside of China) (27%; n = 8,351). A coal open fire was used by 31% (n = 9,686) of the PURE sample at baseline, with nearly all coal heating occurring in China (97%; n = 9,382). Less than 10% of rural PURE households heated their homes with clean fuels (electricity (7%; n = 2,119), central heating (1%; n = 454), gas furnace (1%; n = 350)).

Approximately half (49%) of households had 3–7 rooms; the proportion of households in Bangladesh (52%; n = 487) and Pakistan (46%; n = 398) with only 1–2 rooms was over twice the overall average (18%; n = 5,521) (Table S1). The proportion of households with natural roof materials (e.g. thatch, wood) in Pakistan (32%; 270) and Zimbabwe (29%; n = 186) was four times the average of the PURE sample (8%; n = 2,401). Nearly half (48%; n = 9,618) of female study participants indicated their primary occupation as a homemaker, compared with less than 10% (n = 1,326) of males (Table S2). Approximately half (48%) of male study participants earned a secondary or university/trade school degree, compared with only one-third (34%) of female participants. Over half of rural PURE participants were ranked in the lowest one-third of household income in their country. Characteristics of the sub-sample of households and individuals selected to receive HAP monitoring as part of the PURE-AIR study are described elsewhere (Shupler et al., 2020; Arku et al., 2018).

3.2. Predictors of $PM_{2.5}$ kitchen concentrations

The PURE-AIR $PM_{2.5}$ kitchen concentration model performed slightly better ($R^2 = 0.45$) than the kitchen concentration model using only PURE baseline survey data ($R^2 = 0.37$), which did not include information on stove type and secondary cooking fuel (Table 2). The decrease in R^2 in the PURE baseline model was largely due to the absence of stove type, which was the most important variable for characterizing the variability in $PM_{2.5}$ concentrations across countries (Fig. 1). PURE baseline models explained three times as much of the variability in kitchen concentrations in India ($R^2 = 0.31$) and Chile/Colombia ($R^2 = 0.32$), compared with China ($R^2 = 0.11$), Bangladesh/Pakistan ($R^2 = 0.05$) and Tanzania/Zimbabwe ($R^2 = 0.13$) (Table S7).

Country, primary cooking fuel type, heating fuel type and season

Table 2
Machine learning model performance for predicting average 48-hour $PM_{2.5}$ kitchen concentrations and personal exposures among PURE-AIR participants.

Model	N	R^2	RMSE[1]	NormalizedRMSE[2](%)	Important variables (1–5)
Kitchen concentrations (PURE-AIR survey variables)[3]	2,384	0.45	130	18%	(1) stove type (2) country (3) heating fuel (4) winter/summer season (5) number of household members
Kitchen concentrations (PURE baseline survey variables only)[4]	2,384	0.37	172	23%	(1) country (2) primary cooking fuel (3) heating fuel (4) winter/ summer season (5) household income
Personal exposures (PURE-AIR survey variables)[3]	910	0.33	83	13%	(1) $PM_{2.5}$ kitchen concentration (2) stove type (3) winter/summer season (4) heating fuel (5) time spent in kitchen
Personal exposures (PURE-AIR survey variables without kitchen $PM_{2.5}$ concentration)[3]	910	0.23	92	14%	(1) stove type (2) winter/summer season (3) country (4) heating fuel (5) secondhand smoke exposure
Personal exposures (PURE baseline survey variables only)4	910	0.20	99	15%	(1) roof material (2) primary cooking fuel (3) country (4) heating fuel (5) winter/summer season

1. RMSE = root mean squared error.
2. Normalized RMSE = normalized root mean squared error (obtained by dividing the RSME by the range of $PM_{2.5}$ concentrations in each region or country).
3. PURE-AIR survey variables = variables available in PURE-AIR survey, including: primary cooking fuel type, secondary cooking fuel type, stove type, heating fuel type, hours spent cooking during monitoring, roof material, number of household members, highest household level of education, household income, percent of income spent on food, season, country, hours spent in kitchen during monitoring, indoor/outdoor kitchen, chimney in the kitchen, window in the kitchen, smoking inside the home, years using current fuel.
4. PURE Baseline only = variables only present in PURE baseline survey, including: primary cooking fuel type, heating fuel type, roof material, number of household members, highest household level of education, household income, percent of income spent on food, season, country, indoor/outdoor kitchen, chimney in the kitchen, window in the kitchen.

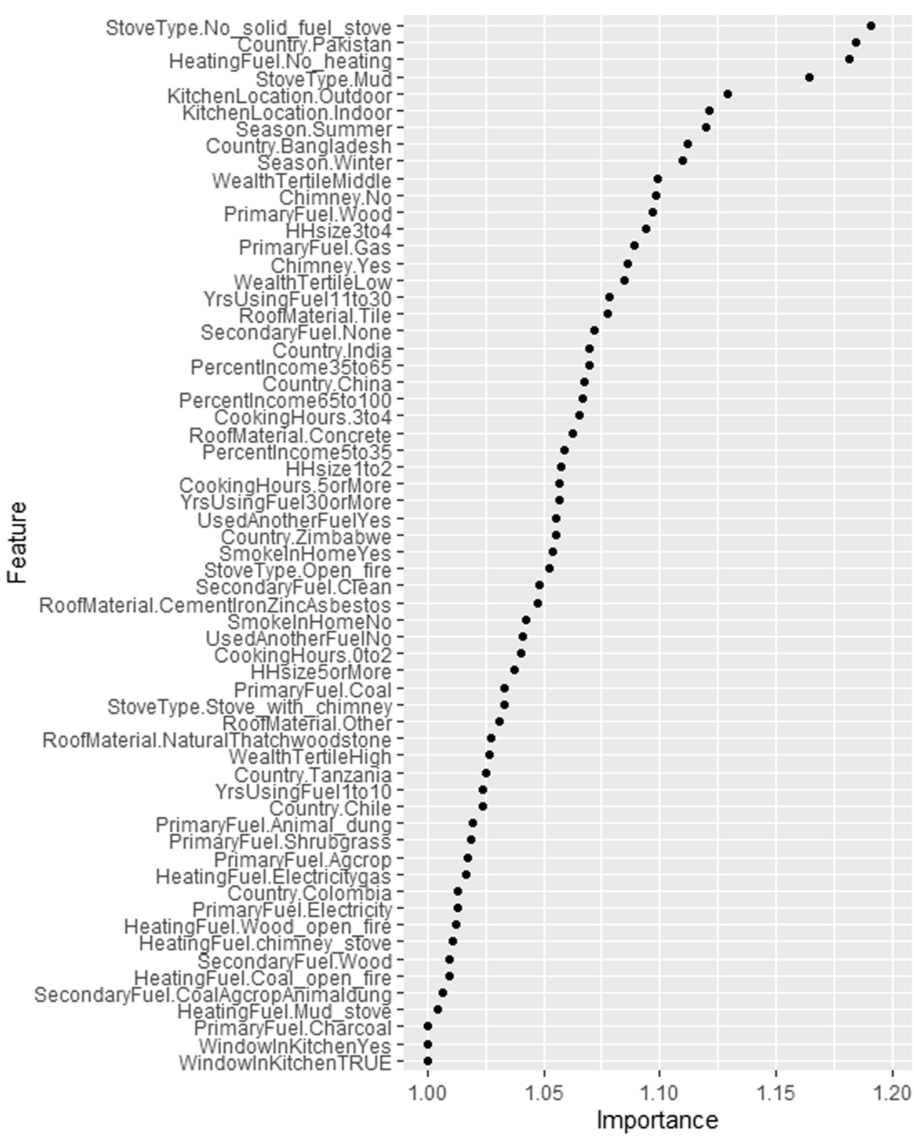

Fig. 1. Variable importance of PURE-AIR survey data for average $PM_{2.5}$ kitchen concentration among all study households. Note: Used Another Fuel = household previously used a different primary cooking fuel (yes/no). Yrs Using Fuel Cat = self-reported number of years using current primary fuel. See Table S3 in Supplement for all variable categories. Note: stove type excluded from India model due to all cooking fuel types used in the same stove across households. Presence of window and kitchen location excluded from China model due to nearly all households having windows and located indoors. Presence of chimney excluded from Africa and Other South Asia models due to all households not having a chimney.

(winter/summer) were the most significant determinants of $PM_{2.5}$ kitchen concentrations in the model using only PURE baseline variables (Table 2). The importance of heating fuel type in the overall model is partially attributed to its significance as a determinant of $PM_{2.5}$ kitchen concentrations in China (which makes up 50% of the PURE sample) (Fig. S1). In China and India, sub-national region was the second most important predictor of $PM_{2.5}$ kitchen concentrations (Table S7).

3.3. Predictors of $PM_{2.5}$ personal exposures

The explained variability of the overall PURE-AIR $PM_{2.5}$ personal exposure model ($R^2 = 0.33$) was lower than that of the kitchen concentration model ($R^2 = 0.45$). Average 48-hour $PM_{2.5}$ kitchen concentration was the best predictor of average male and female $PM_{2.5}$ exposures (Table 1); removal of average 48-hour $PM_{2.5}$ kitchen concentration from the overall $PM_{2.5}$ personal exposure model reduced the coefficient of determination by 30% ($R^2 = 0.23$). Similar to the $PM_{2.5}$ kitchen concentration model, stove type, season and heating fuel type were significant drivers of $PM_{2.5}$ exposure differences (Fig. 2). Kitchen ventilation characteristics, including presence of a chimney and location (indoor versus indoor) were stronger predictors of variations in $PM_{2.5}$ kitchen concentrations (Fig. 1) than personal exposures. However, time spent in the kitchen was highly predictive of average 48-hour $PM_{2.5}$ personal exposure among PURE-AIR participants. Notably, exposure to secondhand smoke was a slightly more important predictor of average $PM_{2.5}$ personal exposures than primary cooking fuel type.

Secondhand smoke was the second most important predictor of $PM_{2.5}$ personal exposures in South American and African PURE countries (Fig. S2), and less important in South Asian countries. In Bangladesh, Pakistan, Tanzania and Zimbabwe, average $PM_{2.5}$ kitchen

Fig. 2. Variable importance of PURE-AIR survey data for average $PM_{2.5}$ personal exposures. Note: $PM_{2.5}$ kitchen concentration not shown on this figure to improve figure scale/visibility. Used Another Fuel = household previously used a different primary cooking fuel (yes/no). Yrs Using Fuel Cat = self-reported number of years using current primary fuel. See Table S3 in Supplement for all variable categories.

concentration was less important in explaining variation in average $PM_{2.5}$ personal exposures than in China, India and South American countries (Fig. S2).

3.4. Modeled $PM_{2.5}$ kitchen concentrations from the PURE-AIR sample

Among 2,384 PURE-AIR households, the model with the best fit (Bayesian $R^2 = 0.53$) (Table S8) included primary cooking fuel type, primary heating fuel type, roof material, season, age group, household income tertile, presence of a chimney in the kitchen and country (Equation (1)). Random intercepts were included for sub-national region (intraclass correlation coefficient (ICC) = 0.01) and community (ICC = 0.24), along with a random slope for wet versus dry season to account for different meteorological effects on HAP concentrations between study countries (Table S8).

Eq. (1): Kitchen concentration model

$$\begin{aligned} \boldsymbol{log(PM_{2.5})_{ijk}} = & (\beta_0+\beta_{0j}+\beta_{0k}) + (\beta_{1*season}+\beta_{1j}+\beta_{1k})*\boldsymbol{season_i} + \beta_2(\boldsymbol{Gas})_i \\ & + \beta_2(\boldsymbol{Electricity})_i + \beta_3(\boldsymbol{Wood})_i + \beta_4(\boldsymbol{Coal})_i + \beta_5(\boldsymbol{Charcoal})_i \\ & + \beta_6(\boldsymbol{Ag/crop\ \ residue})_i + \beta_7(\boldsymbol{Animal\ \ dung})_i \\ & + \beta_8(\boldsymbol{Shrubs/grass})_i + \beta_9(\boldsymbol{Heating\ \ fuel})_i \\ & + \beta_{10}(\boldsymbol{Roof\ \ material})_i + \beta_{11}(\boldsymbol{Season})_i + \beta_{12}(\boldsymbol{Age\ \ group})_i \\ & + \beta_{13}(\boldsymbol{Household\ \ income\ \ tertile})_i + \beta_{14}(\boldsymbol{Chimney})_i \\ & + \beta_{15}(\boldsymbol{Country})_i + \boldsymbol{e_{ijk}} \end{aligned} \tag{1}$$

log$(PM_{2.5})_{ijk}$ is natural logarithm of mean 48-hour $PM_{2.5}$ kitchen concentration of *ith* household in community *j* in sub-national region *k*. β_0 is overall intercept, B_{0j} is random intercept for the *jth* community in sub-national region k, B_{0k} is random intercept for the *kth* sub-national region. $(\beta_{1*season} + \beta_{1j} + \beta_{ik})*$ $season_i$ represents the random slope for season in community j within sub-national region k. e_{ijk} is the leftover error.

Primary cooking fuel, heating fuel, having a chimney in the kitchen and country were statistically evident in the final kitchen concentration model (i.e. the 95% credible interval excluded a null value) (Table 3); households cooking primarily with polluting cooking fuels had significantly higher average 48-hour $PM_{2.5}$ kitchen concentrations compared to households cooking with gas or electricity. Households cooking with wood in China had 32 μg/m^3 (95 %CI: [25,40] higher average $PM_{2.5}$ kitchen concentrations (75 μg/m^3; 95 %CI:[68,83]) than households cooking with gas (43 μg/m^3; 95 %CI: [31,60]).

Households in China using gas for cooking and coal for heating their homes had a significantly (14 μg/m^3) higher average $PM_{2.5}$ kitchen concentration (57 μg/m^3; 95 %CI:[45,74]) than households cooking with gas and not using any heating fuel (43 μg/m^3; 95 %CI: [31,60]). Families in China cooking with wood in a kitchen with a chimney (66 μg/m^3; 95 %CI:[61,72]) had 15 μg/m^3 lower average kitchen concentrations than those cooking with wood in a kitchen without a chimney (81 μg/m^3; 95 %CI:[73,90]).

3.5. Predicting household concentrations for the PURE baseline cohort

Modeled average annual kitchen concentrations among all PURE participants varied four-fold among primary cooking fuel types, ranging from 47 μg/m^3 (95 %CI: [47,47] (gas) to 204 μg/m^3 (95 %CI:[195,213]) (animal dung) (Fig. 3).

Only 4% (n = 1,000) of average $PM_{2.5}$ kitchen concentrations at PURE baseline were below the WHO Interim-1 Target (35 μg/m^3). Nearly 90% of the households in compliance with the WHO target were using clean cooking fuels (gas: 85%; electricity: 2%); only 2% of households using wood (all using improved chimney stoves in Chile) met the WHO target (Table 4). With the exception of 2% (n = 70) of households cooking with gas in China, South America was the only continent in which modeled $PM_{2.5}$ kitchen concentrations among PURE

Table 3
Fixed effect coefficients from final Bayesian hierarchical model of $PM_{2.5}$ kitchen concentrations.

Variable	Mean	Std. Dev	Stat. Sig. ($p < 0.05$)
Primary cooking fuel			
Intercept (gas)	3.84	0.18	
Electricity	−0.04	0.07	
Ag/crop residue	0.37	0.10	*
Charcoal	0.61	0.31	*
Coal	0.51	0.10	*
Wood	0.55	0.05	*
Shrub/grass	0.79	0.17	*
Animal dung	0.81	0.16	*
Heating fuel type			
No heating	REF		
Gas/electric	0.24	0.10	*
Wood (open fire or chimney stove)	0.12	0.06	*
Coal open fire	0.28	0.12	*
Season			
Wet (summer)	REF		
Dry (winter)	0.14	0.11	
Roof material			
Concrete	REF		
Tile	0.01	0.04	
Wood/thatch	0.10	0.07	
Zinc/iron/asbestos	0.11	0.08	
Other	0.0	0.08	
Age group			
35–54	REF		
55–64	−0.02	0.04	
65–85	−0.02	0.04	
Chimney in kitchen			
Yes	−0.20	0.05	*
Household income tertile			
Lowest	REF		
Middle	−0.04	0.04	
Highest	−0.01	0.05	
Country/Region			
China	REF		
India	−1.13	0.40	*
Zimbabwe	1.12	0.49	*
South America (Chile, Colombia)	−0.85	0.36	*
South Asia (Bangladesh, Pakistan)	1.30	0.36	*

Stat. sig. = statistically significant at alpha = 0.95 level
REF = reference group.

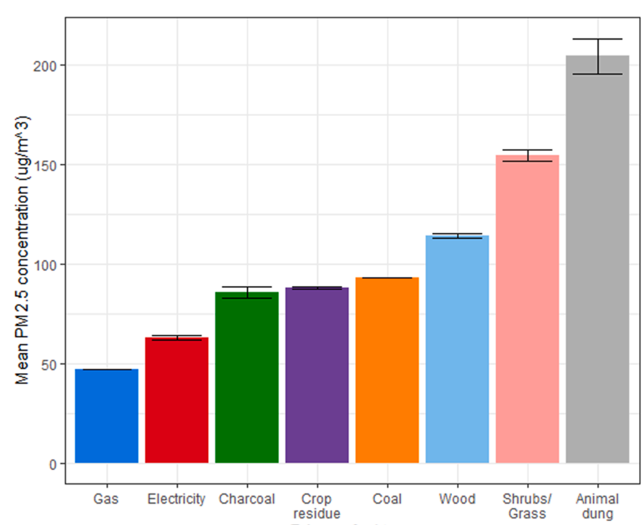

Fig. 3. Modeled average $PM_{2.5}$ kitchen concentrations (μg/m^3) (95% CI) in rural PURE communities at baseline by primary cooking fuel type.

Table 4
Average $PM_{2.5}$ kitchen concentrations (μg/m^3) by primary cooking fuel type from PURE modeling.

Primary cooking fuel	Summer Mean (95% CI)	Winter Mean (95% CI)	Annual Mean (95% CI)	% under WHO interim-1
Gas	44 (44, 45)	50 (50, 50)	47 (46, 48)	14
Electricity	59 (58, 60)	67 (66, 68)	61 (62, 64)	4
Charcoal	83 (81, 86)	88 (85, 91)	86 (83, 88)	0
Ag/crop	84 (83, 84)	92 (91, 92)	88 (87, 88)	0
Coal	86 (86, 86)	101 (100, 101)	93 (93, 93)	0
Wood	106 (105, 107)	122 (121, 124)	114 (113,115)	2
Shrubs/grass	148 (146, 151)	160 (157, 163)	154 (151, 157)	0
Animal dung	195 (186, 203)	214 (205, 223)	204 (195, 213)	0

households met the WHO Interim-1 target. Average modeled $PM_{2.5}$ kitchen concentrations during summer months (wet season) were consistently 5–15 μg/m^3 lower than average winter $PM_{2.5}$ kitchen concentrations across all cooking fuel types (Table 4). At a country level, the difference in average $PM_{2.5}$ kitchen concentrations between wet and dry seasons varied, with the largest seasonal disparity in South Asian countries (Fig. S3). Seasonal $PM_{2.5}$ concentration differences remained in some countries (e.g. India) among households not using any heating fuels during the winter (Fig. S4).

There was substantial within-country variation in $PM_{2.5}$ levels among households using the same primary cooking fuel type. Among the PURE baseline sample, the distribution of predicted average annual $PM_{2.5}$ kitchen concentrations spanned >100 μg/m^3 among households using gas for cooking in China and India (41–148 μg/m^3 and 51–171 μg/m^3, respectively) (Fig. 4). Among households predominantly cooking with wood, average 48-hour $PM_{2.5}$ kitchen concentrations also differed by up to 100 μg/m^3 in China (50 to 149 μg/m^3) and Pakistan (403 to 501 μg/m^3) and over 75 μg/m^3 in India (77 to 159 μg/m^3) (Fig. 5).

Country-level modeled average $PM_{2.5}$ kitchen concentrations varied by a factor of 2.5 among households primarily cooking with gas (20 μg/m^3 (95 %CI: [20,20]) in Chile to 55 μg/m^3 (95 %CI: [55,55]) in China) and 12-fold among households primarily cooking with wood (36 μg/m^3 (95 %CI: [36,36]) in Chile to 427 μg/m^3 (95 %CI:[425,429]) in Pakistan) (Table S4).

In China, the average modeled 48-hour $PM_{2.5}$ kitchen concentration was approximately 30% higher (72 μg/m^3 (95 %CI:[72,73])) among households cooking with wood (chimney stoves were predominantly used), compared with those using gas (55 μg/m^3 (95 %CI: [55,55])) or electricity (58 μg/m^3 (95 %CI: [58,59])) (Fig. 6). In India, average concentrations among households cooking with wood (primarily mud stoves) (89 μg/m^3 (95 %CI:[88,90])) were over double that of households cooking primarily with gas (41 μg/m^3 (95 %CI: [40,41])).

The average $PM_{2.5}$ kitchen concentration among households cooking with gas in Jaipur, India (79 μg/m^3) were approximately 50 μg/m^3 higher than average kitchen concentration among households using gas in Trivandrum (30 μg/m^3) (Table S6). In China, modeled average $PM_{2.5}$ kitchen concentrations among households cooking with gas varied by approximately 35 μg/m^3 across sub-national regions (40 μg/m^3 in Jiangxi to 76 μg/m^3 in Shaanxi) (Table S6). At the sub-national regional level, average $PM_{2.5}$ kitchen concentrations among households cooking primarily with gas were highly correlated with average annual outdoor $PM_{2.5}$ concentrations (Fig. S5).

3.6. *Modeled $PM_{2.5}$ personal exposures from the PURE-AIR sample*

Among 903 participants (499 females; 404 males), the best fitting $PM_{2.5}$ personal exposure model (Bayesian $R^2 = 0.48$) included modeled average 48-hour kitchen concentration (dependent variable from Equation (1)), primary cooking fuel type, primary heating fuel type, roof material, summer/winter season, sex, age group, exposure to second-hand smoke, country and a sex*country interaction term (Equation (2)). In a sensitivity analysis, excluding average 48-hour kitchen concentration from the personal modeling, the explained variability of $PM_{2.5}$ personal exposures decreased by 9% ($R^2 = 0.39$). A random intercept was included for sub-national region (ICC = 0.23) (the personal exposure model did not converge with a community-level random intercept) (Table S9).

Eq. (2): Personal exposure model

$$\begin{aligned} log(PM_{2.5})_{ij} = & (\beta_0 + \beta_{0j}) + \beta_1(Gas)_i + \beta_2(Electricity)_i + \beta_3(Wood)_i \\ & + \beta_4(Coal)_i + \beta_5(Charcoal)_i + \beta_6(Ag/crop\ residue)_i \\ & + \beta_7(Animal\ dung)_i + \beta_8(Shrubs/grass)_i + \beta_9(Age)_i \\ & + \beta_{10}(average\ 48-hour\ PM_{2.5}\ kitchen\ concentration)_i \\ & + \beta_{11}(Roof\ material)_i + \beta_{12}(Secondhand\ smoke\ exposure)_i \\ & + \beta_{13}(Season)_i + \beta_{14}(Sex)_i + \beta_{15}(Country)_i \\ & + \beta_{16}(Sex*Country)_i + e_{ij} \end{aligned} \quad (2)$$

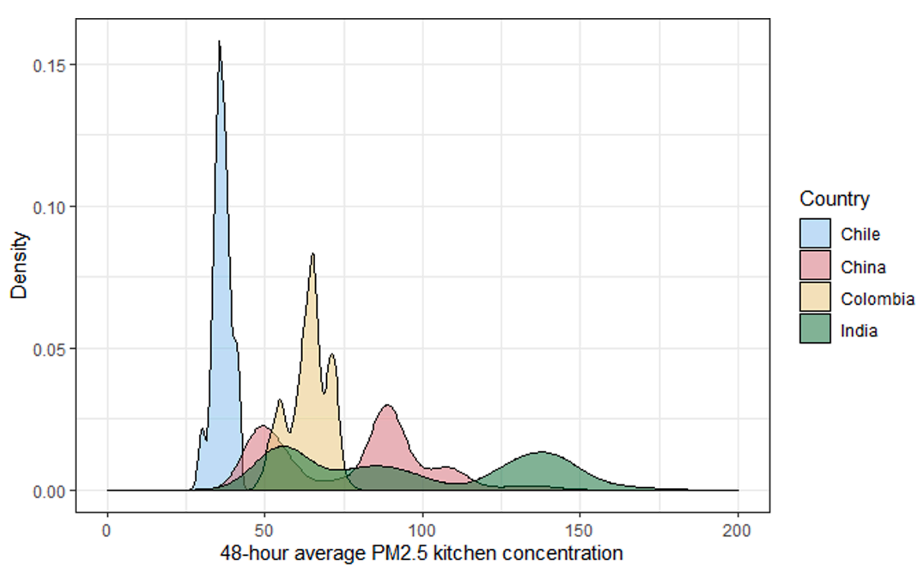

Fig. 4. Distribution of modeled baseline $PM_{2.5}$ kitchen concentrations (μg/m^3) among households cooking primarily with gas in rural PURE communities by country.

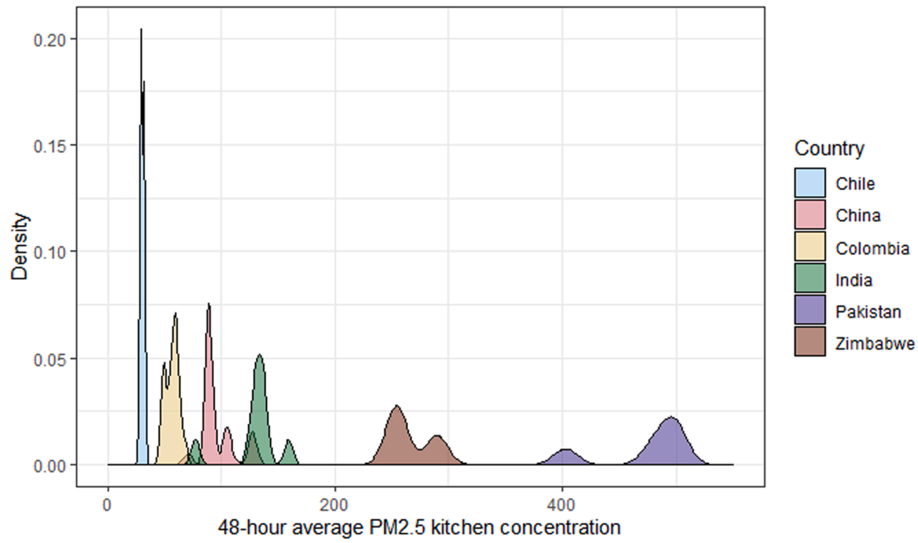

Fig. 5. Distribution of modeled baseline $PM_{2.5}$ kitchen concentrations (μg/m^3) among households cooking primarily with wood in rural PURE communities by country.

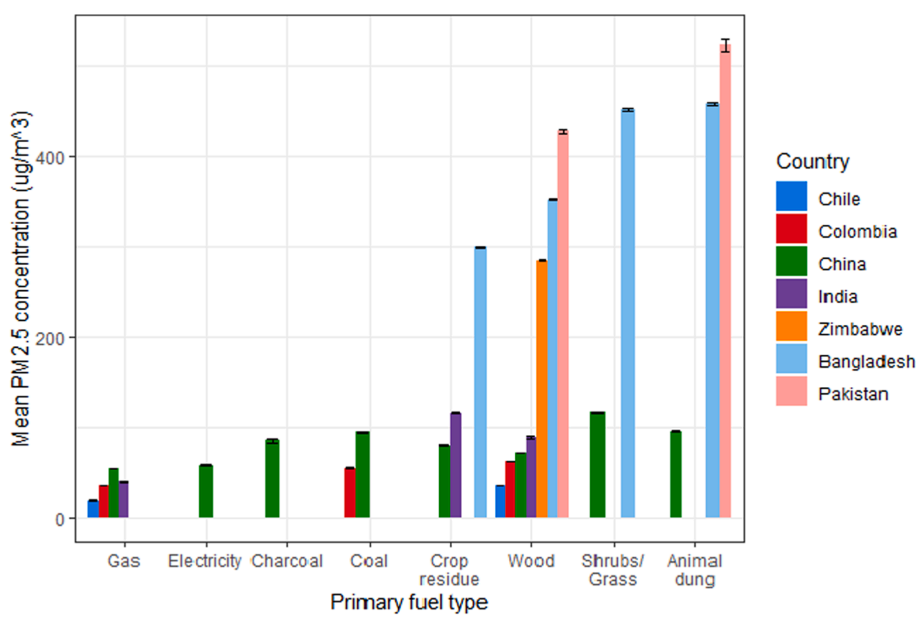

Fig. 6. Modeled PURE baseline average $PM_{2.5}$ kitchen concentrations (μg/m^3) (95% CI) in rural communities by country and primary cooking fuel type (only fuel types with n > 30 in a country shown).

$log(PM_{2.5})_{ij}$ is natural logarithm of mean 48-hour $PM_{2.5}$ personal exposure of *ith* individual in sub-national region *j*. β_0 is overall intercept, B_{0j} is random intercept for the *jth* sub-national region. e_{ij} is the leftover error after accounting for all fixed and random effects.

Average 48-hour kitchen concentration, season (summer/winter), exposure to secondhand smoke, and the interaction between sex and country were statistically evident in the final model (Table 5). A 1 ug/m^3 increase in average 48-hour $PM_{2.5}$ kitchen concentration was associated with a 0.54 ug/m^3 (95 %CI:[0.51,0.58]) increase in average 48-hour personal exposure. Individuals cooking with gas in China and exposed to secondhand smoke in their home had an average 48-hour $PM_{2.5}$ exposure (46 ug/m^3 95 %CI: [42,50]) that was 5 ug/m^3 higher than individuals cooking with gas and not exposed to secondhand smoke (41 ug/m^3 95 %CI: [37,45]).

In sensitivity analyses that included occupation as a predictor among a subset of 912 participants with available data, occupation type was not strongly associated with variations in average $PM_{2.5}$ personal exposures (Fig. S7). When stratifying by sex, occupation was a slightly more important predictor of $PM_{2.5}$ exposure differences among PURE male participants than female participants (Fig. S8).

While ambient $PM_{2.5}$ levels were not examined at an individual level, precluding their inclusion in the modeling, measured personal exposures among PURE-AIR participants primarily cooking with gas were compared to average annual outdoor $PM_{2.5}$ concentrations in each sub-national region (Fig. S6). A high correlation was found between average annual outdoor $PM_{2.5}$ concentrations and average male ($r = 0.78$) and female ($r = 0.68$) $PM_{2.5}$ exposures.

3.7. Predicting personal exposures for the PURE baseline cohort

Among the PURE baseline sample, the distribution of predicted average annual $PM_{2.5}$ personal exposures spanned approximately 70 μg/

Table 5
Fixed effect coefficients from final Bayesian hierarchical model of $PM_{2.5}$ personal exposures.

Variable	Mean	Std. Dev	Stat. Sig. ($p < 0.05$)
Primary cooking fuel			
Intercept (gas)	2.05	0.18	
Electricity	0.01	0.09	
Ag/crop residue	0.22	0.13	
Charcoal	N/A		
Coal	0.05	0.15	
Wood	−0.03	0.07	
Shrub/grass	−0.03	0.17	
Animal dung	0.01	0.14	
Heating fuel type			
No heating	REF		
Gas/electric	0.10	0.10	
Wood (open fire or chimney stove)	0.23	0.08	*
Coal open fire	N/A		
Season			
Wet (summer)	REF		
Dry (winter)	0.22	0.08	*
Roof material			
Concrete	REF		
Tile	0.03	0.06	
Wood/thatch	0.06	0.09	
Zinc/iron/asbestos	0.03	0.11	
Other	0.12	0.11	
Age group			
35–54	REF		
55–64	0.07	0.05	
65–85	−0.03	0.06	
$PM_{2.5}$ kitchen conc (log)	0.39	0.03	*
Sex			
Female	REF		
Male	−0.02	0.04	
Secondhand smoke exposure			
Yes	0.11	0.05	*
Country			
China	REF		
Bangladesh	0.46	0.40	
Chile	−0.26	0.45	
Colombia	−0.19	0.38	
India	0.16	0.21	
Pakistan	0.32	0.41	
Zimbabwe	0.56	0.39	
Country*Sex			
China*Male	REF		
Bangladesh*Male	−0.29	0.21	
Chile*Male	0.37	0.34	
Colombia*Male	0.22	0.19	
India*Male	0.04	0.10	
Pakistan*Male	−0.56	0.24	*
Zimbabwe*Male	−0.44	0.20	*

Stat. sig. = statistically significant at alpha = 0.95 level.
REF = reference group.

m^3 among individuals using gas for cooking in China and India (25–88 $\mu g/m^3$ and 29–97 $\mu g/m^3$, respectively) (Fig. 7). Among households predominantly cooking with wood, average 48-hour $PM_{2.5}$ kitchen concentrations differed by 100 $\mu g/m^3$ in China (40 to 139 $\mu g/m^3$), 120 $\mu g/m^3$ in India (50–169 $\mu g/m^3$) and 200 $\mu g/m^3$ in Pakistan (353–568 $\mu g/m^3$) (Fig. 8).

Modeled average male $PM_{2.5}$ exposures were higher than female exposures among households primarily cooking with gas and charcoal (Fig. 9); average male exposures were significantly lower than female exposures among the majority of polluting fuels (wood, crop waste, shrubs, animal dung). On a country-level, male participants had approximately 50% lower exposures than that of females across all primary cooking fuel types in Pakistan and Zimbabwe (Fig. 10). Although there was not a significant difference in $PM_{2.5}$ exposures between sexes in other countries, modeled $PM_{2.5}$ male exposures were higher than females across all primary fuel types in South American countries (Chile, Colombia) and India (Fig. 10).

4. Discussion

The quantitative $PM_{2.5}$ models developed in this study, based on measurements obtained under a single study protocol across a variety of cooking environments and socioeconomic conditions in seven countries, represent one of the largest multinational HAP exposure assessments conducted to-date. The PURE-AIR models uncovered a gradient of increasing average 48-hour $PM_{2.5}$ kitchen concentrations (Fig. 3) and personal exposures (Fig. 9) when moving from clean to polluting primary cooking fuels. Aside from primary cooking fuel type, stove type, heating fuel type, and presence of a chimney in the kitchen significantly contributed to variation in $PM_{2.5}$ kitchen concentrations across rural PURE communities of LMICs (Table 3). The PURE-AIR quantitative exposure estimates also revealed that average $PM_{2.5}$ male exposures were as high as, or greater than, that of females in some countries (e.g. Chile, Colombia, China, Chile) (Fig. 10), despite females spending an average of nearly thrice as much time in the cooking area (1.9 versus 0.7 h) during the 48-hour monitoring (Shupler et al., 2020).

4.1. Contextual factors

The statistical significance of the country indicator in the final PURE-AIR kitchen concentration model (Table 3) demonstrates that HAP personal exposures and kitchen concentrations are greatly impacted by location-specific factors. This is evidenced by the distribution of $PM_{2.5}$ kitchen concentrations (Figs. 4 and 5) and personal exposures (Figs. 7 and 8) varying substantially by country among households cooking with the same primary cooking fuel; for instance, average 48-hour $PM_{2.5}$ kitchen concentrations ranged from 20 $\mu g/m^3$ in Chile to 55 $\mu g/m^3$ in China among households primarily using gas and from 36 $\mu g/m^3$ in Chile to 427 $\mu g/m^3$ in Pakistan among households primarily cooking with wood.

Moreover, the significant interaction between country and sex in the personal exposure model (Table 5) signals that the impact of contextual factors on $PM_{2.5}$ exposures varies between males and females. In machine learning models stratified by sex, occupation was more predictive of average 48-hour $PM_{2.5}$ male exposures than female exposures (Fig. S8), which echoes findings from *peri*-urban India (Sanchez et al., 2020). In addition, ambient $PM_{2.5}$ modeled estimates from Shaddick et al (Shaddick et al., 2018) were more highly correlated with average male $PM_{2.5}$ exposures ($r = 0.78$) than female exposures ($r = 0.68$) (Fig. S6), signaling that outdoor sources of PM, including occupational exposures, may have a disproportionate impact on male exposure levels due to more time spent outside the home.

Some $PM_{2.5}$ exposure variation can also likely be attributed to other indoor sources of $PM_{2.5}$ including secondhand smoke and kerosene lighting. While PURE surveys did not contain information on primary lighting source, precluding its inclusion in the modeling, kerosene lighting has been shown to increase indoor $PM_{2.5}$ concentrations (Muyanja et al., 2017; Lam et al., 2012). Additionally, other factors including SES characteristics (e.g. roof material, education) were more predictive of personal exposures than cooking environment characteristics in $PM_{2.5}$ exposure models for Tanzania/Zimbabwe, Bangladesh/Pakistan and Chile/Colombia (Table S7).

Sub-national region explained more variation in average $PM_{2.5}$ kitchen concentration in China than primary cooking fuel type (Table S6). Additionally, there was a high correlation ($r = 0.79$) between average annual outdoor $PM_{2.5}$ concentrations and PURE-AIR average $PM_{2.5}$ kitchen measurements among households cooking primarily with gas at a regional-level (Fig. S5). These results suggest that regional differences in ambient air pollution (due to emissions from agricultural burning, traffic, industry) (Zhang et al., 2016; Zhuang et al., 2013;

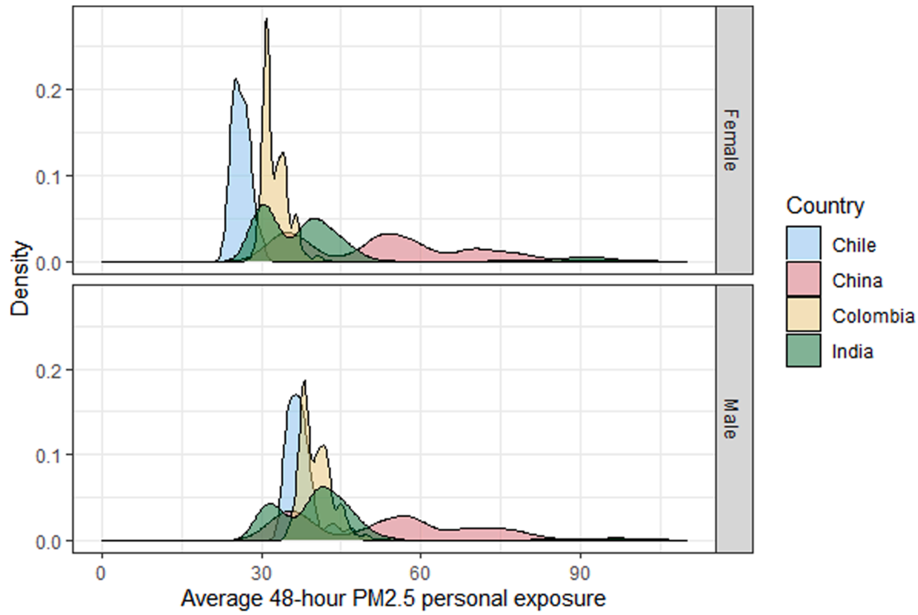

Fig. 7. Distribution of modeled baseline $PM_{2.5}$ personal exposures ($\mu g/m^3$) among households cooking primarily with gas in rural PURE communities by country.

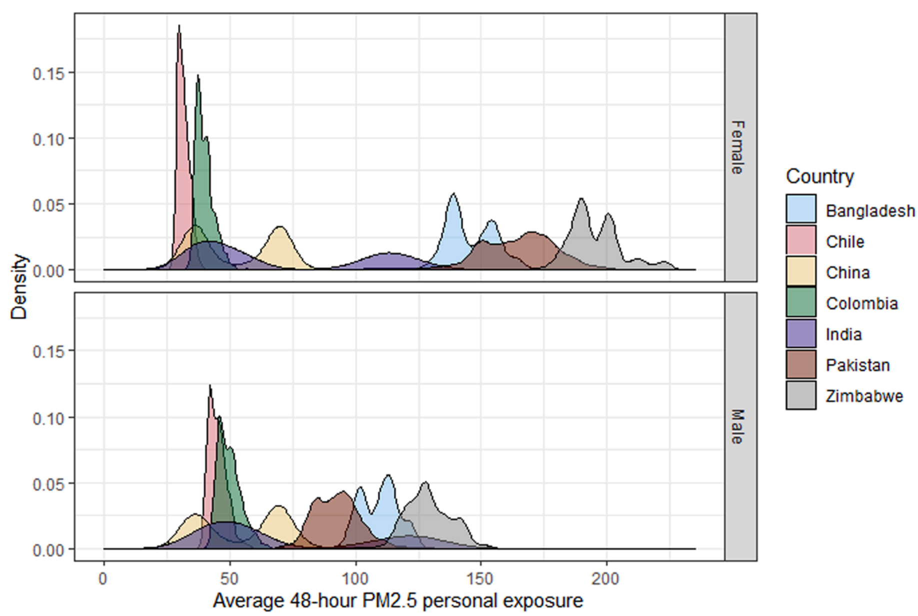

Fig. 8. Distribution of modeled baseline $PM_{2.5}$ female and male personal among households cooking primarily with wood in rural PURE communities by country.

Conibear et al., 2021) are greatly affecting indoor $PM_{2.5}$ levels. A HAP measurement study in rural China similarly found that distance to the highway, a proxy for the level of exposure to vehicular emissions, significantly impacted participants' average $PM_{2.5}$ exposures, additional to residential exposure from cooking with polluting fuels (Baumgartner et al., 2014). The large range (35 $\mu g/m^3$) of modeled average $PM_{2.5}$ kitchen concentrations among households cooking with gas across sub-national regions in China, from an average of 40 $\mu g/m^3$ in Jiangxi to 76 $\mu g/m^3$ in Shaanxi) (Table S5) further indicates the probable existence of other $PM_{2.5}$ sources. Hence, ambient air pollution is likely a key factor in the WHO Interim-1 target (35 $\mu g/m^3$) being exceeded by 98% of modeled average 48-hour kitchen concentrations among rural PURE households cooking with clean fuels in India and China. Accordingly, a transition to clean cooking fuels alone in these two rapidly developing countries is likely not sufficient to reach $PM_{2.5}$ concentrations that meet the WHO Interim-1 target.

4.2. Comparing male and female household air pollution exposures

Minor differences between modeled average 48-hour $PM_{2.5}$ male and female exposures across primary cooking fuel types in rural PURE communities of China, India, Chile and Colombia (Fig. 10) may be due to greater exposure to ambient sources of pollution among males, as

China, India and Temuco, Chile have some of the highest levels of outdoor pollution globally (Conibear et al., 2021; Archer-Nicholls et al., 2016; Díaz-Robles et al., 2008). The higher correlation of average 48-hour $PM_{2.5}$ male exposure measurements ($r = 0.78$) with kitchen concentrations compared with female exposures ($r = 0.68$) (Fig. S6) and the significance of exposure to ambient $PM_{2.5}$ pollution while traveling to work in the predictive model for PURE communities in Chile/Colombia (Table S7) indicates that male participants likely had higher ambient $PM_{2.5}$ exposures than females in these areas. It is feasible that males in PURE communities were more highly exposed to ambient air pollution

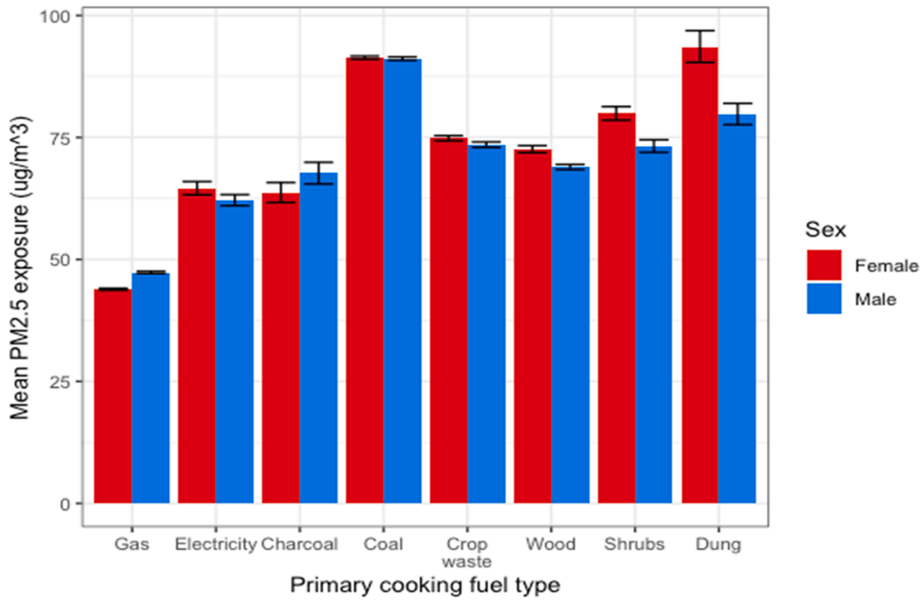

Fig. 9. Modeled average annual $PM_{2.5}$ male and female exposures by primary fuel type.

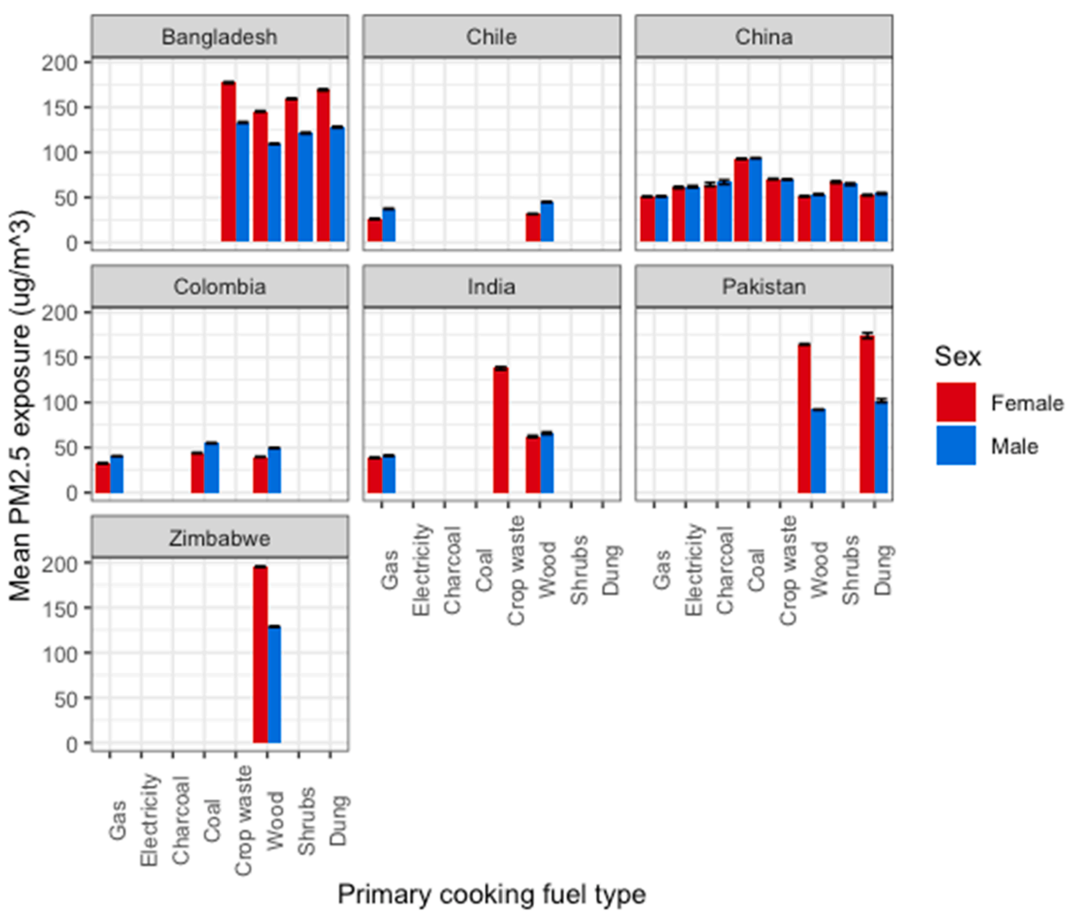

Fig. 10. Modeled average annual $PM_{2.5}$ male and female exposures by primary fuel type and country.

since they were more likely to travel outside the home to their place of employment (Shupler et al., 2020). With cooking emissions possibly playing a less important role in overall $PM_{2.5}$ exposures in these rapidly developing countries, where industrial air pollution sources may dominate, these modeling results are contrary to previous studies that suggest that women have disproportionately higher $PM_{2.5}$ exposures in LMICs where HAP is common. The swiftly changing environmental landscape in several urbanizing LMICs requires that increased attention be placed on mitigating exposure to industrial pollution alongside residential emissions to improve public health.

Contrastingly, female participants in Pakistan, Bangladesh and Zimbabwe had 100–200 μg/m^3 higher modeled average 48-hour $PM_{2.5}$ exposure than males (significant interaction term in personal exposure model (Table 5). Additionally, hours spent in the kitchen during the 48-hour monitoring was an important determinant of average $PM_{2.5}$ exposures in Tanzania/Zimbabwe and Bangladesh/Pakistan (Table S7), indicating that HAP was possibly a more dominant $PM_{2.5}$ exposure source in these settings. Consequently, policies that increase use of clean cooking fuels may achieve greater health benefits from higher reductions in female $PM_{2.5}$ exposures in rural communities in these countries, compared with China, India and South American countries (Shupler et al., 2021; Gould and Urpelainen, 2018; Astuti et al., 2019). As the proportion of households in Bangladesh (52%) and Pakistan (46%) with only 1–2 rooms was more than twice the overall average (18%) (Table S1), higher female $PM_{2.5}$ exposures in these locations may partially be due to lower dispersion rates because of smaller household size.

The findings of similar male and female exposures are consistent with other studies from China (Lee et al., 2021) and *peri*-urban India (Sanchez et al., 2020) that also showed minimal differences in $PM_{2.5}$ exposures between sexes. Both studies also found smoking status to be a significant driver of personal exposures, which is consistent with our finding of secondhand smoke being positively associated with $PM_{2.5}$ personal exposures (Table 5).

As HAP measurement studies have historically prioritized collection of kitchen and living room concentrations or female (main cook) personal measurements, the variation in the difference between modeled average 48-hour $PM_{2.5}$ male and female exposures at a country-level emphasizes the need for additional monitoring of male $PM_{2.5}$ exposures in future HAP measurement studies. Direct $PM_{2.5}$ male exposure measurements will prevent the need for use of personal:kitchen exposure ratios from external studies, thereby reducing biases in male $PM_{2.5}$ exposure estimation in risk assessments.

4.3. Seasonal variation in household air pollution levels

Use of heating fuels was a statistically significant driver of $PM_{2.5}$ kitchen concentrations (Table 3) and personal exposures (Table 5) among PURE communities. However, heating fuel type only partially explained the higher average $PM_{2.5}$ kitchen concentrations in winter, as elevated average winter concentrations (~15 μg/m^3 higher) remained in India among households that reported no use of heating fuels (Fig. S4). Thus, temporal HAP exposure assessment is warranted to minimize misclassification of average annual $PM_{2.5}$ concentrations that can occur when using a 24-hour or 48-hour measurement in a single season.

As studies have shown large within-individual variability in average $PM_{2.5}$ personal exposures in similar *peri*-urban settings (Sanchez et al., 2020; Lee et al., 2021), the single 48-hour exposure measurements in this study may not reflect average annual $PM_{2.5}$ levels. However, the PURE study provides a large, geographically diverse sample, and included collection of measurements across different seasons in some communities. Therefore, season-specific and annual $PM_{2.5}$ levels were derived from the modeling as they may be useful for post-hoc adjustment of short-term HAP measurements in other studies that similarly do not collect seasonal measurements. Using the season-specific estimates presented in this study or others (Shupler et al., 2018) may potentially allow other researchers to more accurately estimate annual $PM_{2.5}$ exposures by factoring in the effects of seasonality, in the absence of other reference data.

4.4. Evaluating model performance

The overall kitchen concentration (Bayesian $R^2 = 0.54$) and personal exposure (Bayesian $R^2 = 0.48$) models moderately explained the large variation in $PM_{2.5}$ levels among rural PURE communities in the seven study countries, and were similar to the performance in other *peri*-urban settings in India (Sanchez et al., 2020) and Kenya (Johnson et al., 2021). Taken together, these studies demonstrate the utility of using quantitative exposure estimation (e.g. $PM_{2.5}$ levels) as opposed to categorical indicators (e.g. primary cooking fuel type) to more accurately capture the range of $PM_{2.5}$ exposures for use in national and multinational models. Similar studies that conduct limited HAP monitoring alongside larger-scale survey collection in the future can generate improved quantitative $PM_{2.5}$ exposure datasets for spatially resolved, large-scale exposure assessment, with HAP sampling becoming increasingly less expensive and resource intensive with the advancement of air monitoring technology (Piedrahita et al., 2014; Amegah, 2018).

Among rural PURE-AIR households in African countries (Tanzania/ Zimbabwe) and Bangladesh/Pakistan, a smaller sample size and more polluted kitchen environments in Tanzania/Zimbabwe and Bangladesh/ Pakistan led to increased variability in $PM_{2.5}$ concentrations, which lessened the predictive power of the corresponding models (Table S6). This finding is partially due to minimal variation in primary cooking fuel used (nearly all households cooked with wood at baseline) and rare use of heating fuels; in a previous HAP predictive modeling study conducted in *peri*-urban Kenya among households cooking with gas and polluting fuels, primary cooking fuel type was the most important exposure predictor (Johnson et al., 2021).

4.5. Strengths and limitations

Exposure determinants included in the PURE-AIR modeling study were not constrained by data reported in publications, and therefore minimized bias that occurs when combining $PM_{2.5}$ measurements across studies with different equipment, measurement techniques and monitoring periods. As 'rural' households recruited into the PURE cohort were typically within a 45-minute drive of urban centers for biological sample storage, the modeled $PM_{2.5}$ estimates in this study may not be nationally representative (Corsi et al., 2013). Nonetheless, as HAP measurement studies typically recruit households in rural communities that have the highest prevalence of polluting cooking fuel use, PURE-AIR modeling was conducted among a unique demographic of communities. The predictive models developed in this study were thus able to examine $PM_{2.5}$ exposure differences in rapidly changing communities, as evidenced by the high rate of primary cooking fuel switching among PURE households over the last two decades (Shupler et al., 2019).

Although PURE-AIR households may not be representative of the full PURE baseline sample, the main goal of the predictive modeling was to establish a diverse $PM_{2.5}$ exposure profile, across a range of primary cooking fuel types, to enable a sufficient sample size for assigning $PM_{2.5}$ exposures to all individuals and households cooking with various cooking technologies. The modeling was able to achieve this goal due to use of a stratified sampling design for HAP monitoring in the PURE-AIR study.

This modeling study did not include direct measurement of ambient air pollution levels in each community. Therefore, while we conclude that air pollution is likely affecting HAP exposures, the relative effect of localized ambient air pollution as opposed to other region-specific factors (e.g. housing type, food choices, time-activity patterns) cannot be quantified. Future multinational studies should aim to collect information on these characteristics to better assess how factors aside from cooking and heating fuel type can alter average $PM_{2.5}$ exposures. These studies can uncover additional changes, beyond a transition to clean cooking fuels, that will be needed in order to meet WHO-interim target levels, which were only achieved by 14% of households cooking primarily with gas at PURE baseline (Table 4).

Because $PM_{2.5}$ measurements and survey data collected in the PURE-AIR study in 2017–2019 were used to assign kitchen and personal levels to PURE participants at baseline (~2005–2010), this modeling study assumed the relationship among the household environment, external factors (e.g. ambient air pollution) and HAP exposures were not substantially altered over the follow up period. As ambient pollution levels and cooking environments may have changed during PURE follow up, particularly in rapidly developing countries like China, applying measurements from 2017 to 2019 to estimate HAP levels a decade prior may have introduced bias in predicted baseline $PM_{2.5}$ exposures. Nevertheless, the quantitative $PM_{2.5}$ exposure estimates obtained in PURE remain the most accurate for a multinational study of this size due to the detailed household-level information included in the predictive modeling.

4.6. Conclusion

Cooking environment characteristics (e.g. stove type, heating fuel and presence of a chimney) partially explained the variation of $PM_{2.5}$ kitchen concentrations and personal exposures among PURE households cooking with the same primary cooking fuel in different countries. Collecting this information in global health surveys (e.g. National Censuses, WHO Harmonized Survey, Demographic Health surveys) can therefore be useful to reasonably quantifying global variations in $PM_{2.5}$ concentrations and exposures due to HAP. Integration of ambient air pollution measurements into $PM_{2.5}$ exposure models may further increase their accuracy.

The heterogenous modeled $PM_{2.5}$ exposures derived in the PURE study can be combined with longitudinal and cross-sectional health data collected among the PURE cohort in epidemiological models. These multinational models will elucidate the shape of $PM_{2.5}$ exposure–response for several respiratory and cardiovascular health outcomes (Burnett and Cohen, 2020; Burnett et al., 2018). Improved estimation of global HAP-related morbidity and mortality obtained using the exposure assessment conducted in PURE can affect how HAP is prioritized, relative to other environmental risk factors, on the global health agenda. This modeling study can therefore benefit policymakers tasked with allocation of finite resources and funding to efficiently alleviate pressing global health problems.

Author contributions

MS coordinated the PURE-AIR study, managed, cleaned, analysed, and interpreted all data, and wrote the first and final drafts of the article. PH and MB designed and supervised the conduct of the PURE-AIR study, supervised the data analysis and interpretation of the data, acquired the funding and reviewed and commented on all drafts and the final article. AB led PURE-AIR protocol development and assisted with study logistics. DM-L assisted with the monitoring equipment, data quality control and provided input on the final article. MJ oversaw laboratory analysis of the data. YLC assisted with data management and study logistics. PG assisted with interpretation of the data. SY designed and supervised the conduct of the PURE study and reviewed and commented on the final Article. MM and LH assisted with data management and study logistics. SR coordinated the worldwide study and reviewed and commented on the final article. All other authors coordinated the study in their respective countries and commented on the final Article.

Declaration of Competing Interest

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

Acknowledgements

The authors would like to acknowledge the PURE-AIR study coordinators and field teams that collected thousands of air pollution measurements across the study countries from 2017-2019. Their hard work and persistence enabled a study of this size to be completed in a timely manner.

The authors also acknowledge Dr. Casey Quinn for his assistance with deploying the monitoring equipment used in the PURE-AIR study.

The PURE-AIR study is funded by Canadian Institutes for Health Research (CIHR) [grant #136893]; and by the Office of The Director, National Institutes of Health (NIH). [Award Number DP5OD019850]. The content is solely the responsibility of the authors and does not necessarily represent the official views of CIHR or the NIH.

The PURE study is an investigator-initiated study that is funded by the Population Health Research Institute, the Canadian Institutes of Health Research, Heart and Stroke Foundation of Ontario, Support from Canadian Institutes of Health Research's Strategy for Patient Oriented Research, through the Ontario SPOR Support Unit, as well as the Ontario Ministry of Health and Long-Term Care and through unrestricted grants from several pharmaceutical companies [with major contributions from AstraZeneca (Canada), Sanofi-Aventis (France and Canada), Boehringer Ingelheim (Germany and Canada), Servier, and GlaxoSmithKline], and additional contributions from Novartis and King Pharma and from various national or local organisations in participating countries. Further details on the funding and participating countries and institutions, and on collaborating staff, are shown below.

Funding/Support

Dr S Yusuf is supported by the Mary W Burke endowed chair of the Heart and Stroke Foundation of Ontario.

The PURE study is an investigator-initiated study that is funded by the Population Health Research Institute, Hamilton Health Sciences Research Institute (HHSRI), the Canadian Institutes of Health Research, Heart and Stroke Foundation of Ontario, Support from Canadian Institutes of Health Research's Strategy for Patient Oriented Research, through the Ontario SPOR Support Unit, as well as the Ontario Ministry of Health and Long-Term Care and through unrestricted grants from several pharmaceutical companies [with major contributions from AstraZeneca (Canada), Sanofi-Aventis (France and Canada), Boehringer Ingelheim (Germany and Canada), Servier, and GlaxoSmithKline], and additional contributions from Novartis and King Pharma and from various national or local organisations in participating countries.

These include: **Argentina:** Fundacion ECLA **(Estudios Clínicos Latino America)** ; **Bangladesh**: Independent University, Bangladesh and Mitra and Associates; **Brazil:** Unilever Health Institute, Brazil; **Canada:** This study was supported by an unrestricted grant from Dairy Farmers of Canada and the National Dairy Council (U.S.), Public Health Agency of Canada and Champlain Cardiovascular Disease Prevention Network; **Chile:** Universidad de La Frontera [DI13-PE11]; **China:** National Center for Cardiovascular Diseases and ThinkTank Research Center for Health Development; **Colombia:** Colciencias (grant 6566-04-18062 and grant 6517-777-58228); **India:** Indian Council of Medical Research; **Malaysia:** Ministry of Science, Technology and Innovation of Malaysia (grant number: 100-IRDC/BIOTEK 16/6/21 [13/2007], and 07-05-IFN-BPH 010), Ministry of Higher Education of Malaysia (grant number: 600-RMI/LRGS/5/3 [2/2011]), Universiti Teknologi MARA, Universiti Kebangsaan Malaysia (UKM-Hejim-Komuniti-15-2010); **occupied Palestinian territory:** the United Nations Relief and Works Agency for Palestine Refugees in the Near East, occupied Palestinian territory; International Development Research Centre, Canada; **Philippines:** Philippine Council for Health Research and Development; **Poland:** Polish Ministry of Science and Higher Education (grant number: 290/W-PURE/2008/0), Wroclaw Medical University; **Saudi Arabia:** Saudi Heart Association. Saudi Gastroenterology Association. Dr. Mohammad Alfagih Hospital. The Deanship of Scientific Research at King Saud University, Riyadh, Saudi Arabia (Research group number: RG -1436-013); **South Africa:** The North-West University, SA and Netherlands Programme for Alternative Development, National Research Foundation, Medical Research Council of South Africa, The South Africa Sugar Association, Faculty of Community and Health Sciences; **Sweden:** Grants from the Swedish state under the Agreement concerning research and education of doctors; the Swedish Heart and Lung Foundation; the Swedish Research Council; the Swedish Council for Health, Working Life and Welfare, King Gustaf V:s and Queen Victoria Freemason's Foundation, AFA Insurance; **Turkey:** Metabolic Syndrome Society, AstraZeneca, Sanofi Aventis; **United Arab Emirates:** Sheikh Hamdan Bin **Role of Sponsor:** The external Rashid Al Maktoum Award For Medical Sciences and Dubai Health Authority, Dubai.

funders and sponsors had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; in the preparation, review, or approval of the manuscript; or in the decision to submit the manuscript for publication.

PURE Project Office Staff, National Coordinators, Investigators, and Key Staff:

Project office (Population Health Research Institute, Hamilton Health Sciences and McMaster University, Hamilton, Canada): S Yusuf* (Principal Investigator).

S Rangarajan (Program Manager); K K Teo, S S Anand, C K Chow, M O'Donnell, A Mente, D Leong, A Smyth, P Joseph, M Duong, O Kurmi, R D'Souza, M Walli-Attaei, B Balaji, R Naito, S Islam (Statistician), W Hu (Statistician), C Ramasundarahettige (Statistician), P Sheridan (Statistician), S Bangdiwala, L Dyal, M Dehghan (Nutrition Epidemiologist), A Aliberti, A Reyes, A Zaki, B Connolly, B Zhang, D Agapay, D Krol, E McNeice, E Ramezani, F Shifaly, G McAlpine, I Kay, J Rimac, J Swallow, M Di Marino, M Jakymyshyn, M(a) Mushtaha, M(o) Mushtaha, M Trottier, N Aoucheva, N Kandy, P Mackie, R Buthool, R Patel, R Solano, S Gopal, S Ramacham, S Trottier.

Core Laboratories: G Pare, M McQueen, S Lamers, J Keys (Hamilton), X Wang (Beijing, China), A Devanath (Bangalore, India).

Argentina: R Diaz*, A Orlandini, P Lamelas, M L Diaz, A Pascual, M Salvador, C Chacon; **Bangladesh:** O Rahman*, R Yusuf*, S A K S. Ahmed, T Choudhury, M Sintaha, A Khan, O Alam, N, Nayeem, S N Mitra, S Islam, F Pasha; **Brazil:** A Avezum*, C S Marcilio, A C Mattos, G B Oliveira; **Canada:** K Teo*, S Yusuf*, Sumathy Rangarajan, A Arshad, B Bideri, I Kay, J Rimac, R Buthool, S Trottier, G Dagenais, P Poirier, G Turbide, AS Bourlaud, A LeBlanc De Bluts, M Cayer, I Tardif, M Pettigrew, S Lear, V de Jong, A N Saidy, V Kandola, E Corber, I Vukmirovich, D Gasevic, A Wielgosz, A Pipe, A Lefebvre, A Pepe, A Auclair, A Prémont, A S Bourlaud; **Chile:** F Lanas*, P Serón, M J Oliveros, F Cazor, Y Palacios; **China:** Liu Lisheng*, Li Wei*, Chen Chunming#, Zhao Wenhua. Hu Bo, Yin Lu, Zhu Jun, Liang Yan, Sun Yi, Wang Yang, Deng Qing, Jia Xuan, He Xinye, Zhang Hongye, Bo Jian, Wang Xingyu, Liu Xu, Gao Nan, Bai Xiulin, Yao Chenrui, Cheng Xiaoru, Wang Chuangshi, Li Sidong, Liu Weida, Lang Xinyue, Liu Xiaoyun, Zhu Yibing, Xie Liya, Liu Zhiguang, Ren Yingjuan, Dai Xi, Gao Liuning, Wang Liping, Su yuxuan, Han Guoliang, Song Rui, Cao Zhuangni, Sun Yaya, Li Xiangrong, Wang Jing, Wang Li, Peng Ya, Li Xiaoqing, Li Ling, Wang Jia, Zou Jianmei, Gao Fan, Tian Shaofang, Liu Lifu, Li Yongmei, Bi Yanhui, Li Xin, Zhang Anran, Wu Dandan, Cheng ying, Xiao Yize, Lu Fanghong, Li Yindong, Hou Yan, Zhang Liangqing, Guo Baoxia, Liao Xiaoyang, Chen Di, Zhang Peng, Li Ning, Ma Xiaolan, Lei Rensheng, Fu Minfan, Liu Yu, Xing Xiaojie, Yang Youzhu, Zhao Shenghu, Xiang Quanyong, Tang Jinhua, Liu Zhengrong, Qiang Deren, Li Xiaoxia, Xu Zhengting, Aideeraili.Ayoupu, Zhao Qian; **Colombia:** P Lopez-Jaramillo*, P A Camacho-Lopez, M Perez, J Otero-Wandurraga, D I Molina, C Cure-Cure, JL Accini, E Hernandez, E Arcos, C Narvaez, A Sotomayor, F Manzur, H Garcia, G Sanchez, F Cotes, A Rico, M Duran, C Torres; **India: Bangalore -** P Mony *, M Vaz*, S Swaminathan, AV Bharathi, K Shankar, A V Kurpad, K G Jayachitra, H A L Hospital, AR Raju, S Niramala, V Hemalatha, K Murali, C Balaji, A Janaki, K Amaranadh, P Vijayalakshmi, **Chennai** - V Mohan*, R M Anjana, M Deepa, K Parthiban, L Dhanasekaran, SK Sundaram, M Rajalakshmi, P Rajaneesh, K Munusamy, M Anitha, S Hemavathy, T Rahulashankiruthiyayan, D Anitha, R. Dhanasekar, S. Sureshkumar, D Anitha, K Sridevi, **Jaipur** - R Gupta, R B Panwar, I Mohan, P Rastogi, S Rastogi, R Bhargava, M Sharma, D Sharma, **Trivandrum -** V Raman Kutty, K Vijayakumar, S Nair, Kamala R, Manu MS, Arunlal AR, Veena A, Sandeep P Kumar, Leena Kumari, Tessi R, Jith S, K Ajayan, G Rajasree, AR Renjini, A Deepu, B Sandhya, S Asha, H S Soumya, **Chandigarh**- R Kumar, M Kaur, P V M Lakshmi, V Sagar J S Thakur, B Patro, R Mahajan, A Josh, G Singh, K Sharma, P Chaudary, **Iran:** R Kelishadi*, A Bahonar, N Mohammadifard, H Heidari, **Kazakhstan:** K Davletov*, B Assembekov, B Amirov; **Kyrgyzstan:** E Mirrakhimov*, S Abilova, U Zakirov, U Toktomamatov; **Malaysia: UiTM -** K Yusoff*, T S Ismail, K Ng, A Devi, N Mat-Nasir, AS Ramli, MNK Nor-Ashikin, R Dasiman, MY Mazaouspavina, F Ariffin, M Miskan, H Abul-Hamid, S Abdul-Razak, N Baharudin, NMN Mohd-Nasir, SF Badlishah-Sham, M Kaur, M Koshy, F A Majid, N A Bakar, N Zainon, R Salleh, SR Norlizan, NM Ghazali, M Baharom, H Zulkifli, R Razali, S Ali, CWJCW Hafar, F Basir; **UKM** - Noorhassim Ismail, M J Hasni, M T Azmi, M I Zaleha, R Ismail, K Y Hazdi, N Saian, A Jusoh, N Nasir, A Ayub, N Mohamed, A Jamaludin, Z Rahim; **Occupied Palestinian Territory:** R Khatib*, U Khammash, R Giacaman; **Pakistan:** R Iqbal*, R Khawaja, I Azam, K Kazmi; **Peru:** J Miranda*, A Bernabe Ortiz, W Checkley, R H Gilman, L Smeeth, R M Carrillo, M de los Angeles, C Tarazona Meza**; Philippines:** A Dans*, H U Co, J T Sanchez, L Pudol, C Zamora-Pudol, L A M Palileo-Villanueva, M R Aquino, C Abaquin, SL Pudol, K Manguiat, S Malayang; **Poland:** W Zatonski*, A Szuba, K Zatonska, R Ilow#, M Ferus, B Regulska-Ilow, D Różańska, M Wolyniec; **Saudi Arabia:** KF AlHabib*, M Alshamiri, HB Altaradi, O Alnobani, N Alkamel, M Ali, M Abdulrahman, R Nouri; **South Africa:** L Kruger*, A Kruger#, P Bestra, H Voster, A E Schutte, E Wentzel-Viljoen, FC Eloff, H de Ridder, H Moss, J Potgieter, A Roux, M Watson, G de Wet, A Olckers, J C Jerling, M Pieters, T Hoekstra, T Puoane, R Swart*, E Igumbor, L Tsolekile, K Ndayi, D Sanders, P Naidoo, N Steyn, N Peer, B Mayosi#, B Rayner, V Lambert, N Levitt, T Kolbe-Alexander, L Ntyintyane, G Hughes, J Fourie, M Muzigaba, S Xapa, N Gobile , K Ndayi, B Jwili, K Ndibaza, B Egbujie; **Sweden** A Rosengren*, K Bengtsson Boström, A Rawshani, A Gustavsson, M Andreasson, L Wirdemann; **Tanzania:** K Yeates*, M Oresto, N West **Turkey:** A Oguz*, N Imeryuz, Y Altuntas, S Gulec, A Temizhan, K Karsidag, K B T Calik, A K Akalin, O T Caklili, M V Keskinler, K Yildiz; **United Arab Emirates:** A H Yusufali, F Hussain, M H S Abdelmotagali, D F Youssef, O Z S Ahmad, F H M Hashem, T M Mamdouh, F M AbdRabbou, S H Ahmed, M A AlOmairi, H M Swidan, M Omran, N A Monsef ; **Zimbabwe:** J Chifamba*, T Ncube, B Ncube, C Chimhete, G K Neya, T Manenji, L Gwaunza, V Mapara, G Terera, C Mahachi, P Murambiwa, R Mapanga, A Chinhara.

*National Coordinator.

Deceased.

PURE Country Institution Names:

	Institution
South Africa	Faculty of Health Science North-West University Potchefstroom Campus University of the Western Cape Department of Dietetics and Nutrition Private Bag X17, 7535 Bellville, South Africa
Zimbabwe	University of Zimbabwe College of Health Sciences Physiology Department Harare, Zimbabwe
Tanzania	Pamoja Tunaweza Women Center, Moshi, Tanzania Division of Nephrology, Department of Medicine Queen's University
China	National Centre for Cardiovascular Diseases Cardiovascular Institute & Fuwai Hospital Chinese Academy of Medical Sciences 167, Bei Li Shi Lu, Beijing, China Fuwai Hospital 167 Beilishi Rd. Xicheng District Beijing. 100037 China
Philippines	University of Philippines, Section of Adult Medicine & Medical Research Unit, Manila, Philippines
Pakistan	Department of Community Health Sciences and Medicine Aga Khan University

(*continued on next page*)

(*continued*)

	Institution
	Stadium Road, P.O Box 3500 Karachi Pakistan
India, Bangalore	St John's Medical College and Research Institute Bangalore 560034, India
India, Chennai	Madras Diabetes Research Foundation & Dr. Mohan's Diabetes Specialities Centre, Chennai
India Jaipur	Eternal Heart Care Centre and Research Institute, Jaipur
India, Trivandrum	Health Action by People, Thiruvananthapuram, Kerala, 695011 INDIA
India, Chandigarh	School of Public Health, Post Graduate Institute of Medical Education & Research, Chandigarh (India)
Bangladesh	Independent University, Bangladesh Bashundhara, Dhaka Bangladesh
Malaysia	Universiti Teknologi MARA, Sungai Buloh, Selangor, Malaysia AND UCSI University, Cheras, Selangor, Malaysia Department of Community Health. Faculty of Medicine. University Kebangsaan Malaysia. Kuala Lumpur. Malaysia
Poland	Wroclaw Medical University Department of Internal Medicine; Department of Social Medicine Borowska 213 street; 50–556 Wroclaw, Poland Department of Epidemiology, The Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology 02–034 Warsaw, 15B Wawelska str. Poland
Turkey	Istanbul Medeniyet University Istanbul, Turkey
Sweden	Sahlgrenska Academy University of Gothenburg Sweden
Iran	Isfahan Cardiovascular Research Center, Isfahan Research Institute Isfahan University of Medical Sciences, Isfahan, Iran
UAE	Dubai Medical University, Hatta Hospital, Dubai Health Authority, Dubai, United Arab Emirates
Saudi Arabia	Department of Cardiac Sciences, King Fahad Cardiac Center College of Medicine King Saud University Riyadh, Saudi Arabia
Palestine	Institute of Community and Public Health, Birzeit University, Ramallah, occupied Palestinian territory
Canada	Université Laval Institut universitaire de cardiologie et de pneumologie de Québec, Quebec Canada G1V 4G5 Simon Fraser University, Dept. of Biomedical Physiology & Kinesiology, BC, Canada Department of Medicine, University of Ottawa, Ottawa, Canada Population Health Research Institute, McMaster University, Hamilton Health Sciences, Hamilton, Ontario, Canada
Argentina	Estudios Clinicos Latinoamerica ECLA Rosario, Santa Fe Argentina Department of Chronic Diseases South American Center of Excellence for Cardiovascular Health (CESCAS) Institute for Clinical Effectiveness and Health Policy (IECS)
Brazil	Dante Pazzanese Institute of Cardiology; Hospital Alemao Oswaldo Cruz Sao Paulo, SP Brazil
Colombia	Facultad de Ciencias de la Salud, Universidad de Santander (UDES), Bucaramanga, Santander, Fundacion Oftalmologica de Santander (FOSCAL) Floridablanca-Santander, Colombia
Chile	Universidad de La Frontera Temuco, Chile
Ecuador	DECANO Facultad de Ciencias de la Salud Eugenio Espejo Universidad Tecnológica Equinoccial Dirección: Av. Mariscal Sucre S/N y Av. Mariana de Jesús, Quito Ecuador
Peru	CRONICAS Centro de Excelencia en Enfermedades Crónicas | www.cronicas-upch.pe Universidad Peruana Cayetano Heredia | www.upch.edu.pe Av. Armendáriz 497, Miraflores, Lima
Russia	Research Institute for Complex Issues of Cardiovascular Diseases, Kemerovo, Russia Institute For Medical Education, Yaroslav-the-Wise Novgorod State University Ministry of Education and Science of the Russian Federation Russia, Saint-Petersburg, 197022, Karpovka river emb., Bld.13, office 28
Kazakhstan	Research Institute of Cardiology & Internal Diseases, Almaty, Kazakhstan
Kyrgyzstan	Kyrgyz Society of Cardiology, National Center of Cardiology and Internal Disease, Bishkek, Kyrgyzstan

(*continued on next column*)

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2021.107021.

References

Alam, D.S., Chowdhury, M.A.H., Siddiquee, A.T., Ahmed, S., Hossain, M.D., Pervin, S., et al., 2012. Adult Cardiopulmonary Mortality and Indoor Air Pollution. Global Heart. 7 (3), 215–221.

Alexander, D., Northcross, A., Wilson, N., Dutta, A., Pandya, R., Ibigbami, T., et al., 2017. Randomized Controlled Ethanol Cookstove Intervention and Blood Pressure in Pregnant Nigerian Women. Am. J. Respir. Crit. Care Med. 195 (12), 1629–1639.

Alexander, D.A., Northcross, A., Karrison, T., Morhasson-Bello, O., Wilson, N., Atalabi, O. M., et al., 2018. Pregnancy outcomes and ethanol cook stove intervention: A randomized-controlled trial in Ibadan, Nigeria. Environ. Int. 111, 152–163.

Amegah, A.K., 2018. Proliferation of low-cost sensors. What prospects for air pollution epidemiologic research in Sub-Saharan Africa? Environ. Pollut. 241, 1132–1137.

Amegah, A.K., Quansah, R., Jaakkola, J.J.K., 2014. Household Air Pollution from Solid Fuel Use and Risk of Adverse Pregnancy Outcomes: A Systematic Review and Meta-Analysis of the Empirical Evidence. PLoS ONE 9 (12). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4252082/.

Archer-Nicholls, S., Carter, E., Kumar, R., Xiao, Q., Liu, Y., Frostad, J., et al., 2016. The Regional Impacts of Cooking and Heating Emissions on Ambient Air Quality and Disease Burden in China. Environ. Sci. Technol. 50 (17), 9416–9423.

Arku, R.E., Ezzati, M., Baumgartner, J., Fink, G., Zhou, B., Hystad, P., et al., 2018. Elevated blood pressure and household solid fuel use in premenopausal women: Analysis of 12 Demographic and Health Surveys (DHS) from 10 countries. Environ. Res. 160, 499–505.

Arku, R.E., Birch, A., Shupler, M., Yusuf, S., Hystad, P., Brauer, M., 2018. Characterizing exposure to household air pollution within the Prospective Urban Rural Epidemiology (PURE) study. Environ. Int. 114, 307–317.

Astuti, S.P., Day, R., Emery, S.B., 2019. A successful fuel transition? Regulatory instruments, markets, and social acceptance in the adoption of modern LPG cooking devices in Indonesia. Energy Res. Social Sci. 1 (58), 101248.

Balakrishnan, K., Ghosh, S., Ganguli, B., Sambandam, S., Bruce, N., Barnes, D.F., et al., 2013. State and national household concentrations of PM2. 5 from solid cookfuel use: results from measurements and modeling in India for estimation of the global burden of disease. Environ Health. 12 (1), 77.

Bates, M.N., Chandyo, R.K., Valentiner-Branth, P., Pokhrel, A.K., Mathisen, M., Basnet, S., et al., 2013. Acute Lower Respiratory Infection in Childhood and Household Fuel Use in Bhaktapur, Nepal. Environ. Health Perspect. 121 (5), 637–642.

Baumgartner, J., Schauer, J.J., Ezzati, M., Lu, L., Cheng, C., Patz, J.A., et al., 2011. Indoor Air Pollution and Blood Pressure in Adult Women Living in Rural China. Environ. Health Perspect. 119 (10), 1390–1395.

Baumgartner, J., Schauer, J.J., Ezzati, M., Lu, L., Cheng, C., Patz, J., et al., 2011. Patterns and predictors of personal exposure to indoor air pollution from biomass combustion among women and children in rural China: Biomass smoke exposure in rural China. Indoor Air 21 (6), 479–488.

Baumgartner, J., Zhang, Y., Schauer, J.J., Huang, W., Wang, Y., Ezzati, M., 2014. Highway proximity and black carbon from cookstoves as a risk factor for higher blood pressure in rural China. Proc. Natl. Acad. Sci. 111 (36), 13229–13234.

Bond, T.C., Doherty, S.J., Fahey, D.W., Forster, P.M., Berntsen, T., DeAngelo, B.J., et al., 2013. Bounding the role of black carbon in the climate system: A scientific assessment: BLACK CARBON IN THE CLIMATE SYSTEM. J. Geophys. Res. Atmosph. 118 (11), 5380–5552.

Bürkner, P.-C., 2017. brms : An R Package for Bayesian Multilevel Models Using Stan. J. Stat. Softw. 80 (1). http://www.jstatsoft.org/v80/i01/.
Burnett, R., Cohen, A., 2020. Relative Risk Functions for Estimating Excess Mortality Attributable to Outdoor PM2.5 Air Pollution: Evolution and State-of-the-Art. Atmosphere 11 (6), 589.
Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C.A., et al., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. PNAS 29, 201803222.
Burroughs Pena, M., Romero, K.M., Velazquez, E.J., Davila-Roman, V.G., Gilman, R.H., Wise, R.A., et al., 2015. Relationship Between Daily Exposure to Biomass Fuel Smoke and Blood Pressure in High-Altitude Peru. Hypertension 65 (5), 1134–1140.
Carter, E., Archer-Nicholls, S., Ni, K., Lai, A.M., Niu, H., Secrest, M.H., et al., 2016. Seasonal and Diurnal Air Pollution from Residential Cooking and Space Heating in the Eastern Tibetan Plateau. Environ. Sci. Technol. 50 (15), 8353–8361.
Chafe, Z.A., Brauer, M., Klimont, Z., Van Dingenen, R., Mehta, S., Rao, S., et al., 2014. Household Cooking with Solid Fuels Contributes to Ambient PM2.5 Air Pollution and the Burden of Disease. Environ. Health Perspect. http://ehp.niehs.nih.gov/1206340.
Clark, M.L., Bachand, A.M., Heiderscheidt, J.M., Yoder, S.A., Luna, B., Volckens, J., et al., 2013. Impact of a cleaner-burning cookstove intervention on blood pressure in Nicaraguan women. Indoor Air 23 (2), 105–114.
Conibear, L., Reddington, C.L., Silver, B.J., Chen, Y., Knote, C., Arnold, S.R., et al., 2021. Statistical Emulation of Winter Ambient Fine Particulate Matter Concentrations From Emission Changes in China. GeoHealth. 5 (5) e2021GH000391.
Corsi, D.J., Subramanian, S.V., Chow, C.K., McKee, M., Chifamba, J., Dagenais, G., et al., 2013. Prospective urban rural epidemiology (pure) study: baseline characteristics of the household sample and comparative analyses with national data in 17 countries. Am. Heart J. 166 (4), 636–646.e4.
Díaz-Robles, L.A., Ortega, J.C., Fu, J.S., Reed, G.D., Chow, J.C., Watson, J.G., et al., 2008. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. Atmosph. Environ. 42 (35), 8331–8340.
Ezzati, M., Kammen, D.M., 2001. Indoor air pollution from biomass combustion and acute respiratory infections in Kenya: an exposure-response study. Lancet. 358 (9282), 619–624.
Fisher, A., Rudin, C., Dominici, F., 2019. All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. arXiv:180101489 [stat] [Internet]. 2019 Dec 23 [cited 2020 Dec 14]; Available from: http://arxiv.org/abs/1801.01489.
Gelman, A., Goodrich, B., Gabry, J., Ali, I., 2017. R-squared for Bayesian regression models.
Gould, C.F., Urpelainen, J., 2018. LPG as a clean cooking fuel: Adoption, use, and impact in rural India. Energy Policy. 1 (122), 395–408.
Grieshop, A.P., Marshall, J.D., Kandlikar, M., 2011. Health and climate benefits of cookstove replacement options. Energy Policy. 39 (12), 7530–7542.
Gurley, E.S., Salje, H., Homaira, N., Ram, P.K., Haque, R., Petri, W.A., et al., 2013. Seasonal concentrations and determinants of indoor particulate matter in a low-income community in Dhaka, Bangladesh. Environ. Res. 121, 11–16.
Health Effects Institute. State of Global Air 2020; 2020, p. 28.
Ibrahim, J.G., Chen, M.-H., 2000. Power prior distributions for regression models. Statist Sci. 15 (1), 46–60.
Jin, Y., Zhou, Z., He, G., Wei, H., Liu, J., Liu, F., et al., 2005. Geographical, Spatial, and Temporal Distributions of Multiple Indoor Air Pollutants in Four Chinese Provinces. Environ. Sci. Technol. 39 (24), 9431–9439.
Johnson, M., Piedrahita, R., Pillarisetti, A., Shupler, M., Menya, D., Rossanese, M., et al., 2021. Modeling approaches and performance for estimating personal exposure to household air pollution: A case study in Kenya. Indoor Air [Internet]. [cited 2021 Mar 3];n/a(n/a). Available from: http://onlinelibrary.wiley.com/doi/abs/10.1111/ina.12790.
Kephart, J.L., Fandiño-Del-Rio, M., Koehler, K., Bernabe-Ortiz, A., Miranda, J.J., Gilman, R.H., et al., 2020. Indoor air pollution concentrations and cardiometabolic health across four diverse settings in Peru: a cross-sectional study. Environ Health. 19 (1), 59.
Krebs, B., Burney, J., Zivin, J.G., Neidell, M., 2021. Using Crowd-Sourced Data to Assess the Temporal and Spatial Relationship between Indoor and Outdoor Particulate Matter. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.0c08469.
Kurmi, O.P., Semple, S., Simkhada, P., Smith, W.C.S., Ayres, J.G., 2010. COPD and chronic bronchitis risk of indoor air pollution from solid fuel: a systematic review and meta-analysis. Thorax 65 (3), 221–228.
Kurmi, O.P., Arya, P.H., Lam, K.-B.-H., Sorahan, T., Ayres, J.G., 2012. Lung cancer risk and solid fuel smoke exposure: a systematic review and meta-analysis. Eur. Respir. J. 40 (5), 1228–1237.
Lam, N.L., Smith, K.R., Gauthier, A., Bates, M.N., 2012. Kerosene: A Review of Household Uses and their Hazards in Low- and Middle-Income Countries. J. Toxicol. Environ. Health, Part B 15 (6), 396–432.
Lee, M., Carter, E., Yan, L., Chan, Q., Elliott, P., Ezzati, M., et al., 2021. Determinants of personal exposure to PM2.5 and black carbon in Chinese adults: A repeated-measures study in villages using solid fuel energy. Environ. Int. 1 (146), 106297.
Liu, J., Mauzerall, D.L., Chen, Q., Zhang, Q., Song, Y., Peng, W., et al., 2016. Air pollutant emissions from Chinese households: A major and underappreciated ambient pollution source. PNAS 113 (28), 7756–7761.
Massey, D., Kulshrestha, A., Masih, J., Taneja, A., 2012. Seasonal trends of PM10, PM5.0, PM2.5 & PM1.0 in indoor and outdoor environments of residential homes located in North-Central India. Build. Environ. 47, 223–231.
Molnar, C., 2021. Interpretable Machine Learning [Internet]. [cited 2021 Jan 14]. Available from: https://christophm.github.io/interpretable-ml-book/.
Muyanja, D., Allen, J.G., Vallarino, J., Valeri, L., Kakuhikire, B., Bangsberg, D.R., et al., 2017. Kerosene lighting contributes to household air pollution in rural Uganda. Indoor Air 27 (5), 1022–1029.
Ni, K., Carter, E., Schauer, J.J., Ezzati, M., Zhang, Y., Niu, H., et al., 2016. Seasonal variation in outdoor, indoor, and personal air pollution exposures of women using wood stoves in the Tibetan Plateau: Baseline assessment for an energy intervention study. Environ. Int. 94, 449–457.
Norris, C., Goldberg, M.S., Marshall, J.D., Valois, M.-F., Pradeep, T., Narayanswamy, M., et al., 2016. A panel study of the acute effects of personal exposure to household air pollution on ambulatory blood pressure in rural Indian women. Environ. Res. 147, 331–342.
Piedrahita, R., Xiang, Y., Masson, N., Ortega, J., Collier, A., Jiang, Y., et al., 2014. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. Atmos. Meas. Tech. 7 (10), 3325–3336.
Pokhrel, A.K., 2004. Case-control study of indoor cooking smoke exposure and cataract in Nepal and India. Int. J. Epidemiol. 34 (3), 702–708.
R Core Team, 2017. R: A language and environment for statistical computing. [Internet]. Vienna, Austria: R Foundation for Statistical Computing. Available from: http://www.R-project.org/.
Ramanathan, V., Carmichael, G., 2008. Global and regional climate changes due to black carbon. Nat. Geosci. 1 (4), 221–227.
Rehman, I.H., Ahmed, T., Praveen, P.S., Kar, A., Ramanathan, V., 2011. Black carbon emissions from biomass and fossil fuels in rural India. Atmos. Chem. Phys. 11 (14), 7289–7299.
Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. 1 (5), 206–215.
Salvi, S., Barnes, P.J., 2010. Is exposure to biomass smoke the biggest risk factor for COPD globally? Chest 138 (1), 3–6.
Sanchez, M., Milà, C., Sreekanth, V., Balakrishnan, K., Sambandam, S., Nieuwenhuijsen, M., et al., 2020. Personal exposure to particulate matter in peri-urban India: predictors and association with ambient concentration at residence. J. Expo Sci. Environ. Epidemiol. 30 (4), 596–605.
Shaddick, G., Thomas, M.L., Amini, H., Broday, D., Cohen, A., Frostad, J., et al., 2018. Data Integration for the Assessment of Population Exposure to Ambient Air Pollution for Global Burden of Disease Assessment. Environ. Sci. Technol. 52 (16), 9069–9078.
Shen, G., Hays, M.D., Smith, K.R., Williams, C., Faircloth, J.W., Jetter, J.J., 2018. Evaluating the Performance of Household Liquefied Petroleum Gas Cookstoves. Environ. Sci. Technol. 52 (2), 904–915.
Shupler, M., Godwin, W., Frostad, J., Gustafson, P., Arku, R.E., Brauer, M., 2018. Global estimation of exposure to fine particulate matter (PM2.5) from household air pollution. Environ. Int. 120, 354–363.
Shupler, M., Balakrishnan, K., Ghosh, S., Thangavel, G., Stroud-Drinkwater, S., Adair-Rohani, H., et al., 2018. Global household air pollution database: Kitchen concentrations and personal exposures of particulate matter and carbon monoxide. Data Brief. 21, 1292–1295.
Shupler, M., Hystad, P., Gustafson, P., Rangarajan, S., Mushtaha, M., Jayachtria, K.G., et al., 2019. Household, community, sub-national and country-level predictors of primary cooking fuel switching in nine countries from the PURE study. Environ. Res. Lett. 14 (8), 085006.
Shupler, M., Hystad, P., Birch, A., Miller-Lionberg, D., Jeronimo, M., Arku, R.E., et al., 2020. Household and personal air pollution exposure measurements from 120 communities in eight countries: results from the PURE-AIR study. Lancet Plane. Health. 4 (10), e451–e462.
Shupler, M., Mangeni, J., Tawiah, T., Sang, E., Baame, M., Cuevas, R.A., et al., 2021. Beyond Household Socioeconomic Status: Multilevel Modeling of Supply-Side Determinants of LPG Consumption among 5,500 Households in Sub-Saharan Africa. Nature Portfolio. https://doi.org/10.21203/rs.3.rs-154082/v1.
Smith, K.R., Bruce, N., Balakrishnan, K., Adair-Rohani, H., Balmes, J., Chafe, Z., et al., 2014. Millions Dead: How Do We Know and What Does It Mean? Methods Used in the Comparative Risk Assessment of Household Air Pollution. Annu. Rev. Public Health 35 (1), 185–206.
Teo, K., Chow, C.K., Vaz, M., Rangarajan, S., Yusuf, S., 2009. The Prospective Urban Rural Epidemiology (PURE) study: Examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. Am. Heart J. 158 (1), 1–7.e1.
Thompson, L.M., Bruce, N., Eskenazi, B., Diaz, A., Pope, D., Smith, K.R., 2011. Impact of Reduced Maternal Exposures to Wood Smoke from an Introduced Chimney Stove on Newborn Birth Weight in Rural Guatemala. Environ. Health Perspect. 119 (10), 1489–1494.
Upadhyay, A.K., Singh, A., Kumar, K., Singh, A., 2015. Impact of indoor air pollution from the use of solid fuels on the incidence of life threatening respiratory illnesses in children in India. BMC Public Health 15 (1). http://www.biomedcentral.com/1471-2458/15/300.
Vehtari, A., Gelman, A., Gabry, J., 2016. loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. Available from: https://CRAN.R-project.org/package=loo.
Yu, K., Qiu, G., Chan, K.-H., Lam, K.-B.-H., Kurmi, O.P., Bennett, D.A., et al., 2018. Association of Solid Fuel Use With Risk of Cardiovascular and All-Cause Mortality in Rural China. JAMA 319 (13), 1351–1361.
Zhang, L., Liu, Y., Hao, L., 2016. Contributions of open crop straw burning emissions to $PM_{2.5}$ concentrations in China. Environ. Res. Lett. 11 (1), 014014.
Zhuang, Y., Chen, D., Li, R., Chen, Z., Cai, J., He, B., et al., 2018. Understanding the Influence of Crop Residue Burning on PM2.5 and PM10 Concentrations in China from 2013 to 2017 Using MODIS Data. Int. J. Environ. Res. Public Health 15 (7), 1504.