ENVIRONMENTAL RESEARCH
HEALTH

OPEN ACCESS

RECEIVED
9 May 2025

REVISED
9 October 2025

ACCEPTED FOR PUBLICATION
24 October 2025

PUBLISHED
5 November 2025

PAPER

Evaluating associations between the transition to cleaner cooking energy use and hypertension in India

Priyanka deSouza[1,2,*], Jenny J Lee[3,4], Jeremy Németh[1], Sunil Mani[5], Abhishek Jain[5], Abhishek Kar[5], Jennifer Peel[6], Sadeer Al-Kindi[8], Patrick L Kinney[9], Ajay Pillarisetti[10], Wenlu Ye[10], Rockli Kim[11,12], SV Subramanian[13,14], Michelle L Bell[12,15] and Eric A F Simoes[16,17,18]

[1] Department of Urban and Regional Planning, University of Colorado Denver, Denver, CO 80202, United States of America
[2] CU Population Center, University of Colorado Boulder, Boulder, CO 80302, United States of America
[3] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, United States of America
[4] Department of Data Science, Ewha Womans University, Seoul, Republic of Korea
[5] Council on Energy, Environment and Water, New Delhi, India
[6] Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO, United States of America
[7] Harrington Heart and Vascular Institute, University Hospitals Cleveland Medical Center, Cleveland, OH, United States of America
[8] Case Western Reserve University School of Medicine, Cleveland, OH, United States of America
[9] Boston University School of Public Health, Boston, MA 02118, United States of America
[10] Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA, United States of America
[11] Division of Health Policy and Management, College of Health Sciences, Korea University, Seoul, Republic of Korea
[12] Interdisciplinary Program in Precision Public Health, Department of Public Health Sciences, Graduate School of Korea University, Seoul, Republic of Korea
[13] Harvard Center for Population and Development Studies, Cambridge, MA, United States of America
[14] Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA, United States of America
[15] School of the Environment, Yale University, 195 Prospect Street, New Haven, CT 06511, United States of America
[16] Department of Pediatrics, University of Colorado School of Medicine, Aurora, CO, United States of America
[17] Children's Hospital Colorado, Aurora, CO, United States of America
[18] Department of Epidemiology, Colorado School of Public Health, Aurora, CO, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: priyanka.desouza@ucdenver.edu

Keywords: clean energy transition, India, pollution, cardiovascular disease, hypertension

Supplementary material for this article is available online

Abstract

Sustainable Development Goal 7 aims to ensure access to clean and affordable energy for all. India has implemented large-scale policies to improve access to clean fuels (CFs) such as liquefied petroleum gas. However, the impacts of such policies on public health outcomes have been understudied. This study utilized data from India's National Family Health Surveys (NFHSs) conducted in 2016 and 2021 to evaluate the associations between changes in the prevalence of non-CF (non-CF) usage, driven by India's clean energy policies, and hypertension, a significant cardiovascular disease risk factor, on a national level. Specifically, a precision-weighted method that accounts for survey design and sampling variability was implemented to estimate the prevalence of hypertension and a wide range of socio-economic (SES) factors for the years 2010–2015 (NFHS-4) and 2015–2020 (NFHS-5) at the district level for the 2019/20 NFHS-5. Associations between the change in solid-fuel use, also termed as non-CF at the district level between 2015–2020 and 2010–2015, and the change in the prevalence of hypertension after controlling for various SES were investigated using regression analyzes. A significant decrease in hypertension prevalence: 0.41% (95% CI: 0.07%, 0.75%) was associated with a 10% point decrease in the prevalence of non-CF use at the district level. Subpopulation-specific analyzes revealed substantial variation in associations by gender, region, urban/rural designation, socioeconomic groups, and age. Specifically, these associations were highest among older populations and residents of North India.

1. Introduction

Cardiovascular disease (CVD), primarily ischemic heart disease and stroke, is a leading cause of death and disability worldwide (Mensah *et al* 2019, Vaduganathan *et al* 2022). India alone accounts for one-fifth of all CVD-related deaths. The Global Burden of Disease estimates that the age-standardized CVD death rate is 314 (95% CI: 265–362) per 100 000 people in India, whereas the global average is 240 (95% CI: 221–255) per 100 000 people (WHO 2016). CVD is diagnosed in Indian individuals at least a decade earlier than in Western populations; 52% of CVD-related deaths in India occur before 70 years of age, compared with 23% in developed economies (Joshi *et al* 2007, Prabhakaran *et al* 2016). Risk factors for CVD include environmental exposures, like household air pollution; pre-existing conditions, like diabetes; and socioeconomic indicators and lifestyle factors, including an appropriate diet, physical inactivity, smoking, and/or tobacco chewing (Roth *et al* 2020).

Household air pollution (HAP) from the use of solid fuels is a risk factor for CVD morbidity and mortality (Smith 2002, Al-Kindi *et al* 2020). India has undertaken large-scale government policies towards universal electrification (e.g. the Saubhagya scheme, announced in 2017) and clean-cooking access (Pradhan Mantri Ujjwala Yojana (PMUY), announced in 2016) to address the HAP burden (Goldemberg *et al* 2018). For both policies, tremendous improvements in access to cleaner technologies have occurred: Saubhagya electrified 214 million homes (GoI 2024), and PMUY enabled approximately 90 million households to access liquefied petroleum gas (LPG) at a reduced cost (GoI 2024). Previous modeling suggests that India's clean fuel (CF) policies may yield improved birth weight and blood pressure and reduce pneumonia (Rosenthal *et al* 2018, Steenland *et al* 2018). However, little empirical evidence exists on the scale of exposure reductions that have occurred or on subsequent health benefits of these policies.

This research uses publicly available data from repeated National Family Health Survey (NFHS) datasets to assess associations between the change in the prevalence of CF use spurred by clean energy policies in India and cardiovascular risk factors. A precision-weighted methodology that accounts for NFHS survey design and sampling variability was used to estimate the prevalence of an important CVD risk factor: hypertension, and a wide range of socioeconomic status (SES) factors for 2010–2015 and 2015–2020 at the district level. The change in the prevalence of solid, non-CFs (non-CFs) at the district level between 2015–2021 and 2010–2015 was then estimated. We hypothesize differences in LPG use at the district level due to differences in fuel affordability, policy implementation, and cooking practices. The associations between the change in the prevalence of non-CF use and the change in the prevalence of hypertension were evaluated. Separately, the change in the district-level prevalence of LPG alone on hypertension is reported, as enabling the transition to LPG is a key aim of PMUY.

2. Methods and materials

2.1. NFHS data

The NFHS-4 survey conducted between January 2015 and November 2016 (NFHS 2017) and the NFHS-5 survey conducted between June 2019 and April 2021 (NFHS 2021) were used. NFHS are nationally representative household sample surveys measuring indicators of SES, demographics, health, and nutrition, with special emphasis on maternal and child health. They have a two-stage design, in which some clusters (villages in rural areas and census enumeration blocks in urban areas) are first selected from each district (640 districts at the time of the 2011 census for NFHS-4; 707 districts as of 2017 for NFHS-5), and then 25–30 households are selected by equal probability systematic sampling in each of the selected clusters, and women of reproductive age (15–49 years) and men (15–54 years) are then selected from those households for in-depth surveys.

The eligibility criteria used were: 1) Information on geographic coordinates for each cluster was present, 2) Individuals had lived in the same residence for at least five years at the time of the interview, and 3) Women were not pregnant during the interview. Overall, we had 683 105 individuals (542 590 (84.1%) women and 102 262 (15.9%) men) interviewed as part of the NFHS-4, and 644 852 individuals (587 054 (85.9%) women and 96 051 (14.1%) men) as part of NFHS-5. Individuals with missing information were excluded when evaluating district-level proportions for each variable from the above cohorts. The number and proportion of missing data for each variable considered are listed in table S1.1. Overall, we observe that $<$10% of the data is missing for all the main variables considered; only occupation had $>$50% of data missing for the NFHS-4 and 5 cohorts.

2.1.1. Hypertension outcome

Blood pressure was measured for participants using an OMRON Blood Pressure monitor as part of the NFHS. Blood pressure readings were taken on three separate occasions with an interval of 5 min between readings by trained health technicians (IIPS 2017, Sciences 2020). The mean of the second and third measurements was used to record systolic and diastolic blood pressure. Hypertension was defined based on the cut-offs provided by the Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC7) Guideline 2003. Specifically, an individual was categorized as having hypertension if their systolic blood pressure was $\geqslant$140 mm Hg or diastolic blood pressure was $\geqslant$90 mm Hg (Chobanian *et al* 2003). In a sensitivity analysis, we defined an individual as hypertensive if their systolic blood pressure was $\geqslant$140 mm Hg or diastolic blood pressure was $\geqslant$90 mm Hg or if they reported in the survey using antihypertensive medication.

Supplementary health outcomes, diabetes, and heart disease, which include self-reported information, are defined in section S2 in supplementary information. As there are significant socioeconomic inequalities in awareness, treatment, and control of diseases among adults in India, which could bias results when using outcomes that incorporate self-reported information (Maiti *et al* 2023), we evaluate these outcomes in secondary analyzes alone.

2.1.2. CF use

Individuals were categorized as CF-users if they reported using electricity, LPG, or natural gas/biogas; and not-CF-users if they responded they used coal, charcoal, wood, kerosene, straw/shrubs/grass, agricultural crop or animal dung as their primary cooking fuel (Pope *et al* 2017, 2021). This study separately considers the adoption of LPG alone, as it is the focus of the PMUY program.

2.1.3. Covariates

In district-level analyzes, the following covariates were used: the prevalence of individuals in the lowest wealth quintile (poor individuals), the prevalence of individuals with no formal education, the prevalence of obese (BMI $\geqslant$ 30 kg m^{-2}) individuals, the prevalence of smokers, the prevalence of individuals who drink alcohol, the prevalence of non-vegetarians, the prevalence of households with improved sanitation (if the household has access to water pipes into dwelling, yard, or plot, public tap or standpipe, tube well or borehole, protected well or spring, rainwater, and bottled water), the prevalence of individuals belonging to different job-categories: manual, agricultural, professional, services, sales, clerical, not-working, and the prevalence of scheduled castes (SCs), scheduled tribes (STs) and other backward classes (OBCs), the prevalence of Muslims. SCs, STs, and OBCs are officially designated groups of people among the most socioeconomically disadvantaged in India (deSouza *et al* 2023).

2.2. Ambient $PM_{2.5}$ exposure

Because India lacks a dense network of surface $PM_{2.5}$ monitoring sites, this study relied on satellite-derived annual averaged $PM_{2.5}$ estimates (Hammer *et al* 2020), which have been well-validated and widely used (Boing *et al* 2022, deSouza *et al* 2022a, 2022b).

Specifically, satellite aerosol optical depths (AODs) were combined from multiple satellite products: MISR, MODIS Dark Target, MODIS and SeaWiFS Deep Blue, and MODIS MAIAC, with simulation-based results based on their relative uncertainties. These AODs were related to near-surface monthly $PM_{2.5}$ concentrations at a 0.01° × 0.01° (~1 km × 1 km at the equator) resolution over the globe using the ratio of simulated AOD and $PM_{2.5}$ from the GEOS-Chem model. On an annual scale, the $PM_{2.5}$ estimates are highly consistent with globally distributed ground monitors (R^2 = 0.90–0.92). To maintain the privacy of respondents, NFHS randomly displaces the location of each cluster by a maximum of 5 km from the true location for rural areas and 2 km in urban areas, with a further 5% of all clusters displaced by 10 km. Thus, mean $PM_{2.5}$ levels in the 2 km/5 km buffer were assigned to urban and rural clusters, respectively.

2.3. Statistical models

2.3.1. Deriving district-level estimates

Precision-weighted predicted probabilities of the health outcomes, exposure, and covariates were derived at the district level from individual and household-level data using four-level multilevel models. As there are differences in the districts used in the NFHS-4 and NFHS-5, the prevalence of each outcome and SES variable was estimated using the 707 districts defined in the NFHS-5 survey. NFHS-4 clusters were mapped onto NFHS-5 districts based on the NFHS-5 district in which the center point of each NFHS-4 cluster fell.

The four levels of geographic units are individuals (or households) at level 1 (i), clusters at level 2 (j), districts at level 3 (k), and states at level 4 (l). The model is presented below:

$$\text{logit}(\text{NFHS variable}_{ijkl}) = \alpha_0 + u_{0jkl} + v_{0kl} + f_{0l} \quad (1)$$

where α_0 is the constant, and represents the median log odds of each covariate across all of India; u_{0jkl}, v_{0kl}, and f_{0l} are the residuals at the cluster, district, and state levels, respectively. The residuals are assumed to be normally distributed with a mean of 0 and a variance of σ^2_{u0}, σ^2_{v0}, and σ^2_{f0}, respectively. These variance terms can be interpreted as within-district between-cluster variation (σ^2_{u0}), within-state between-district variation (σ^2_{v0}), and between-state variation (σ^2_{f0}).

From the model described in equation (1), the district-specific logit values were converted to probabilities: $\exp(\alpha_0 + v_{0kl} + f_{0l}) / (1 + exp(\alpha_0 + v_{0kl} + f_{0l}))$.

The lme4 package in R (Bates *et al* 2009) was used to evaluate these probabilities for every NFHS-5 district for the periods: 2016–2017 and 2019–2021. Small pairwise correlations between all potential confounding variables, the exposure, and the outcome were observed (figure S4.1). Estimates of district-level prevalences of all variables were also derived for individuals disaggregated by sex, age category (15–19, 20–29, 30–39, 40–49, 50–54 yrs), caste (SC, ST, OBC, Other), and urban/rural designation.

2.3.2. District-level regression models

Associations between the changes in the prevalence of non-CF use between NFHS-4 and NFHS-5 surveys and the differences in the prevalence of the health outcomes at the district level were estimated using:

$$\Delta\text{Outcome}_{\text{district}} = \beta_0 + \beta_1 \times \Delta\text{NonCF use}_{\text{district}} + \sum_{\text{covariates}} \beta_{\text{covariate}} \times \Delta\text{covariates}_{\text{district}} + \beta_2 \times \Delta\text{PM}_{2.5} + \text{error} \quad (2)$$

where β_1 represents the association of interest between the change in prevalence of non-CF use and the change in prevalence of the health outcome (hypertension in the main analysis). Because of the potential for a lack of statistical independence between districts in the same state, clustered standard errors by state were estimated for all models. We report associations as the percentage change in hypertension for a 10% point increase in non-CF use.

The robustness of the associations to the inclusion of different covariates was evaluated by including different sets of covariates in a stepwise manner. The regression model represented in equation (2) was rerun using district-level population weights. The assumption of linearity in the relationship between the Δ use of non-CF and the health outcomes considered was relaxed using generalized additive models (GAMs; for more details, refer to section S2).

The regression model represented in equation (2) was rerun using the prevalence of health outcomes for different subpopulations (disaggregated by gender, urban/rural, caste, and age group) to evaluate associations between district-level Δ non-CF use and Δ subpopulation-specific health outcomes using the fully adjusted model specification. To assess region-specific associations, the Δ non-CF use was interacted with a dummy variable for region: **North** (comprising of the states, Jammu & Kashmir, Ladakh, Himachal Pradesh, Punjab, Chandigarh, Uttarkahand, Haryana, Rajasthan, Delhi), **Central** (comprising of the states, Uttar Pradesh, Odisha, Chhattisgarh, Madhya Pradesh), **East** (comprising of the states, Bihar, West Bengal, Jharkhand), **West** (comprising of the states, Gujarat, Maharashtra, Goa), **South** (comprising of the states, Andhra Pradesh, Karnataka, Lakshadweep, Kerala, Tamil Nadu, Puducherry, Andaman and Nicobar Islands, Telangana), and **North-East** (comprising of the states, Sikkim, Arunachal Pradesh, Nagaland, Manipur, Mizoram, Tripura, Meghalaya, Assam). We conducted these subpopulation-specific analyzes since the literature on the cooking energy transition has shown that the uptake of clean cooking fuels and the choice of cooking fuels at the household level depend on the complex interplay among economic, social, demographic, and institutional factors and that the health benefits are likely, not uniform (Roy 2024).

In the main subpopulation analysis, we chose to use district-level Δ non-CF use as the key exposure of interest, as it is likely that a net change in area-level solid fuel use will impact the air pollution and thus the health of everyone in a given district. However, in supplementary analyzes, regression models were run using the district-level prevalence of subpopulation-specific health outcomes, exposure, *and* covariate information, and subpopulation-specific associations were also reported.

Cross-sectional regressions for each of the NFHS-4 and 5 time periods were also run to assess time period-specific associations between the prevalence of non-CF use and hypertension. Finally, this study also reports associations between the absence/presence of solid fuels and hypertension from individual-level analyzes using data from the NFHS-4 and 5 surveys, respectively (for more details, refer to section

S3). As it is impossible to know if the same individuals were interviewed in the NFHS-4 and 5 surveys, we could not evaluate associations between the change in solid fuel use spurred by the CF transition and the health outcomes considered using the individual-level analysis.

CF Use and Ambient Air Pollution

Associations between district-level Δ non-CF use and ambient air pollution were also investigated. Specifically, the association between district-specific Δ non-CF and the change in mean district-level ambient $PM_{2.5}$ between 2010–2015 and 2015–2020 was evaluated using an unadjusted model, and then via an adjusted model using the covariates listed.

All statistical analyzes were performed using R 4.2.0 (Team *et al* 2013). p-values were two-sided with a significance threshold of 0.05.

3. Results

Descriptive statistics of district-level prevalence in hypertension, exposure, and covariates in 2016 (NFHS-4) and 2021 (NFHS-5) are displayed in table 1. The prevalence (expressed as a %) of residents with hypertension increased between 2016 and 2021 from 11.3% to 11.8%. Hypertension is highest in clusters in South, North, and North-East India (figure 1). non-CF use was highest in poorer areas in Central and North-East India. The mean prevalence of non-CF use decreased from 64.7% to 46.1% between 2016 and 2021. There was a decrease in the prevalence of non-CF use in almost every district between the two time periods, suggesting that individuals are transitioning away from solid fuels (figure 1). This transition may have been spurred by India's clean energy policies. Non-CF use appears to decrease most in areas that did not have the highest baseline non-CF use in South India, particularly in the state of Karnataka. The district-level prevalence of LPG use increased from an average percentage of 33.5% in 2016 to 51.9% in 2021. Maps of the district-level prevalence of LPG for 2016 and 2021 are displayed in figure S4.2. The changes in LPG use correspond closely to the changes in non-CF use.

Figure 2 shows the difference in the prevalence of covariates (poverty levels, obesity, smoking status, education levels, diet (vegetarian/nonvegetarian), sanitation, caste/tribal status, and religion. Figure S4.3 presents the prevalence of each covariate in 2016 and 2021; we note stable distributions of covariates over time. Individuals in Central India had the highest prevalence of being in the lowest wealth quintile, having the least formal education, and actively smoking tobacco. The highest prevalence of obese individuals is in South India; prevalence increased from 18.5% to 22.9% overall between 2016 and 2021. The prevalence of households with improved sanitation was highest in Central India.

The prevalence of individuals in different occupations in 2016 and 2021 and the difference in prevalence between the two time periods are displayed in figures S4.4–5. The most striking patterns observed are the increase in the prevalence of manual laborers and professional workers in South India, the increase in unemployed workers in Central India, and the increase in the number of service industry workers in North-East India between 2016 and 2021.

We observed a 0.41% (95% CI: 0.07%, 0.75%) decrease in the prevalence (expressed as a percentage) of district-level hypertension for every 10% point decrease in the prevalence of non-CF use in fully-adjusted regression models (table 2). Results from population-weighted regression analyzes yielded similar results (table S1.2). An increase in the prevalence of LPG use by 10% points in a district was significantly associated with a decrease in the prevalence of hypertension corresponding to: −0.41% (95% CI: −0.75%, −0.07%) (table S1.3).

The sensitivity of the association between the change in the prevalence of district-level hypertension and the change in the prevalence of non-CF use to the inclusion of covariates was evaluated by evaluating changes in the association observed in unadjusted, partially adjusted, and fully adjusted models (table S1.2). In unadjusted models, a 10% point decrease in the prevalence of non-CF use was associated with an increase of 0.22% (95% CI: −0.21%, 0.65%) in the prevalence of hypertension (table S1.2). When controlling for the prevalence of key risk factors of hypertension and SES (obesity, smoking, alcohol consumption, non-vegetarian lifestyle, education status, and access to improved sanitation), the association observed was 0.45% (95% CI: 0.00%, 0.89%), suggesting that the association observed was sensitive to the covariates considered. In fully adjusted models that further controlled for the prevalence of individuals in different occupations, the association remained largely unchanged: 0.40% (95% CI: 0.09%, 0.71%) (table S1.2). The results for the supplementary outcomes are described in section S1.4.

Subpopulation analyzes revealed that a 10% point decrease in the prevalence of non-CF use was associated with a decrease in hypertension among men of 0.67% (95% CI: 0.16%, 1.18%), compared with 0.40% (95% CI: 0.08%, 0.72%) for women (table 3). Similar associations were observed in urban and rural areas, corresponding to 0.44% (95% CI: 0.12%, 0.76%) and 0.41% (95% CI: 0.08%, 0.74%), respectively (table 3). The association among the ST population is higher than for SC, OBC, or non-SC,

Table 1. Descriptive statistics of the prevalence (unit: %) of the main health outcome, hypertension, exposures, and covariates at the district level using data from the NFHS-4 and NFHS-5 surveys. Note that there were three districts in Maharashtra (1) and Telangana (2) for which we did not have data on occupation in the year 2016. Mean (standard deviation) is reported.

	NFHS-4 (%)	NFHS-5 (%)	Change in prevalence between NFHS-5 and NFHS-4	Percentage change in prevalence between NFHS-5 and NFHS-4: $100\times$ ($value_{nfhs-5}-value_{nfhs-4}$)/$value_{nfhs-4}$
Outcome				
Hypertension	11.3% (4.0%)	12.1% (3.6%)	0.8% (3.6%)	13.6% (35.4%)
Exposure				
Non-CF use	64.7% (28.7%)	46.1% (28.7%)	−18.5% (12.7%)	−34.3% (26.5%)
LPG use	33.4% (27.8%)	51.9% (28.6%)	18.6% (12.7%)	128.8% (140.3%)
Covariates				
Obese	18.5% (8.7%)	22.9% (10.4%)	4.4% (5.0%)	28.4% (32.3%)
Drinks alcohol	5.1% (6.4%)	3.2% (3.9%)	−1.9% (3.5%)	−32.0% (25.3%)
Non-vegetarian	64.6% (30.9%)	66.2% (30.9%)	1.6% (6.7%)	6.8% (25.4%)
Smokes	14.1% (12.1%)	9.0% (9.9%)	−5.1% (4.8%)	−38.0% (20.6%)
Poor	16.3% (19.2%)	17.7% (18.9%)	1.4% (9.0%)	68.2% (225.1%)
No formal education	27.2% (14.1%)	22.0% (12.2%)	−5.2% (5.0%)	−18.4% (18.4%)
Improved household sanitation	58.4% (28.6%)	82.7% (14.8%)	24.3% (18.4%)	92.6% (137.4%)
Caste and religion				
Schedule caste (SC)	13.8% (9.1%)	15.4% (10.3%)	1.6% (5.2%)	19.9% (58.2%)
Scheduled tribe (ST)	17.2% (30.1%)	17.2% (30.4%)	−0.0% (4.9%)	52.5% (203.6%)
Other backward classes (OBC)	38.9% (23.9%)	38.3% (24.4%)	−0.6% (9.1%)	6.0% (68.2%)
Muslim	6.7% (17.3%)	6.5% (17.5%)	−0.2% (3.2%)	13.0% (89.8%)
Occupation				
Manual labor	13.8% (5.7%)	13.6% (5.1%)	−0.2% (6.2%)	10.4% (52.6%)
Agricultural	20.8% (10.1%)	22.3% (11.9%)	1.5% (7.1%)	10.9% (4.0%)
Professional	3.7% (1.2%)	3.5% (1.7%)	−0.2% (1.1%)	−6.7% (24.8%)
Sales	4.4% (1.1%)	4.3% (1.1%)	−0.1% (0.9%)	−0.7% (18.6%)
Services	4.4% (1.6%)	4.0% (1.6%)	−0.4% (1.7%)	−3.7% (37.2%)
Clerical	1.1% (0.5%)	0.9% (0.4%)	−0.1% (0.5%)	−4.2% (41.5%)
Not working	43.5% (7.7)	41.1% (7.9%)	−2.3% (5.7%)	−4.8% (13.2%)
Ambient $PM_{2.5}$ ($\mu g\,m^{-3}$)	61 (31)	62 (31)	1 (2)	0.016 (0.04)

ST, or OBC populations, corresponding to 0.61% (95% CI: 0.24%, 0.98%) (table 3). Associations were highest among the older age groups of 40–49 and 50–54 yr of age, corresponding to a decreased risk of 0.79% (95% CI: 0.23%, 1.36%), and 1.63% (95% CI: 0.57%, 2.70%), respectively, for a 10% point decrease in non-CF use; this is likely because older adults are more vulnerable to the impact of pollution from CF use than younger individuals. All study participants in the 50–54 year age range were men, as the age of women in our sample ranged from 15 to 49 yr (table 3).

Non-CF use was most pronounced in North India (figure 1). Associations between the change in non-CF use and the change in hypertension correspond to 1.20% (95% CI: 0.11%, 2.29%) and 0.56% (95% CI: 0.01%, 1.11%) in North and Central India, respectively (table 3). Associations in the North-East and West regions are not statistically significant, although the same general trend of associations is observed; while in the East and South, the general trend of associations was reversed (table 3). Region-specific differences could arise from different lifestyles, cooking, activity patterns, and housing structures that can modify the impact of non-CF use on health outcomes.

Subpopulation-specific analyzes were repeated using the prevalence of the outcome, exposure, and covariates for the respective subpopulations (table S1.4). In these analyzes, the associations reported correspond to the change in the prevalence of subgroup-specific health outcomes for a 10% point change in the prevalence of subgroup-specific non-CF use. In general, the associations reported in table S1.4

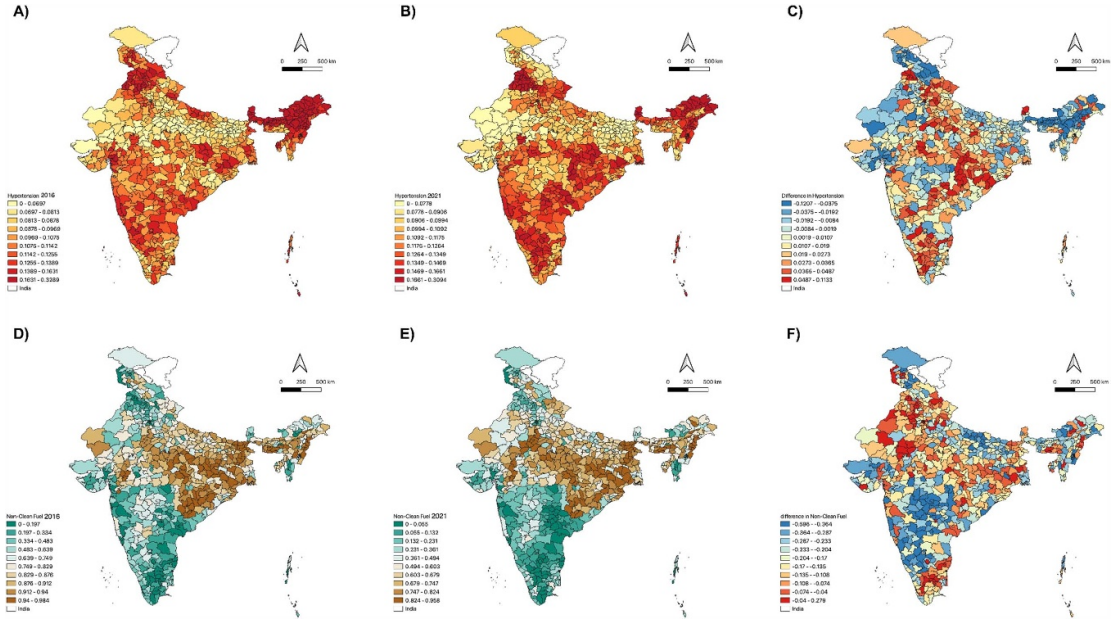

Figure 1. Prevalence (unit: fraction) of the two main outcomes: (A) hypertension in 2016 from NFHS-4, (B) hypertension in 2021 from NFHS-5, (C) difference in the prevalence of Hypertension between the two time periods, (D) non-clean fuel use in 2016 from NFHS-4, and main exposure: (E) non-clean fuel use in 2021 from NFHS-5, (F) difference in non-clean fuel use between the two time periods, classified into deciles.

were smaller than those in table 3. Some associations (for urban areas and non-SC/ST/OBC) disappeared completely. The difference in associations reported is likely because, in table S1.4, the change in subpopulation-specific solid fuel use may be small compared to that of the overall district. Therefore, although we do not see an association between the change in subpopulation-specific solid fuel use and health outcomes, we see that the overall district-level change in solid fuel use is associated with a change in health status.

Cross-sectional regressions revealed significant associations between the prevalence of non-CF use and the prevalence of hypertension in 2016: 0.37% (95% CI: 0.10%, 0.64%) (table S1.5). We did not observe significant associations between the change in non-CF use and the change in district-level $PM_{2.5}$ concentrations, although the general trend of the association observed: -0.12 $\mu g\ m^{-3}$ (95% CI: -0.34 $\mu g\ m^{-3}$, 0.11 $\mu g\ m^{-3}$) suggests that $PM_{2.5}$ concentrations decrease as non-CF use increases (table S1.6). A visual determination shows that the changes in $PM_{2.5}$ concentrations (figure S4.6) are dissimilar to the change in prevalence of non-CF use (figure 1), suggesting that other factors likely drove the changes in $PM_{2.5}$ levels observed.

Large uncertainties in the exposure-response curve were observed when relaxing the assumption of linearity in the relationship between the Δ use of non-CF and the health outcomes (For more details, refer to section S2). No significant associations were obtained between the presence/absence of non-CF and the main health outcomes in individual-specific analyzes (for more details on why this is the case, refer to section S3).

4. Discussion

We evaluated associations between the change in prevalence of non-CF use spurred by India's clean energy policies on hypertension, a key CVD-risk factor, in India. Overall, a significant decrease in the prevalence of hypertension at the district level was observed, corresponding to 0.41% (95% CI: 0.07%, 0.75%) for every 10% point decrease. The prevalence of hypertension in India is 11.3% (8.8% among women and 13.8% among men) (Gupta *et al* 2019). Thus, even a 10% point decrease in non-CF use could reduce the prevalence of hypertension by ∼4%. Importantly, it appeared that most of the change in non-CF use in India was driven by the increased adoption of LPG, part of which is due to the PMUY program. Our results thus suggest that India's CF policies have likely resulted in substantial benefits to public health and are playing a key role in achieving Sustainable Development Goal (SDG)-7, which aims to ensure affordable, clean energy to all.

Table 2. Associations between a 10% point increase in non-CF use and the change in prevalence of hypertension between 2021 and 2016 in fully-adjusted models.

	Δ Hypertension
	Adjusted for difference in prevalence of obesity, smoking, drinking alcohol, nonvegetarians, poverty, no formal education, improved sanitation, SC, ST, OBC, Muslim, Manual, Agricultural, Professional, Services, Sales, Clerical, Not working
Δ Non-CF use (std error clustered by state)	0.41%* (0.07%, 0.75%)
Δ Non-CF use (population^-weighted; std errors clustered by state)	0.40%* (0.09%, 0.71%)

^Populations were derived by aggregating population counts for men and women between the ages 15–54 yr and 15–45 yr, respectively to the NFHS-5 district level, using age- and sex-specific population data from The Gridded Population of the World, Version 4 (GPWv4) Basic Demographic Characteristics, Revision 11 for the year 2010 at a ∼ 1 km resolution for India.

Table 3. Associations between a 10% increase in non-CF use at the district level and the change in subpopulation-specific health outcomes. All models were adjusted for the differences in prevalence of obesity, smoking, individuals who drink alcohol, non-vegetarians, individuals from poor backgrounds, those with no formal education, improved sanitation, SC, ST, OBC, Muslim, manual, agricultural, professional, service, sales, clerical, and non-working backgrounds. Robust standard errors are reported, clustered by state.

	Δ Hypertension
Gender	
Women	0.40%* (0.08%, 0.72%)
Men	0.67%* (0.16%, 1.18%)
Urban/Rural	
Urban	0.44%* (0.12%, 0.76%)
Rural	0.41%* (0.08%, 0.74%)
Caste	
SC	0.40%* (0.11%, 0.69%)
ST	0.61%* (0.24%, 0.98%)
OBC	0.47%* (0.15%, 0.80%)
Not SC, ST, OBC	0.42%* (0.14%, 0.70%)
Age Category	
1 (15–19 yr)	0.13% (−0.01%, 0.27%)
2 (15–19 yr)	0.18% (0.00%, 0.35%)
3 (20–29 yr)	0.26% (−0.01%, 0.53%)
4 (30–40 yr)	0.62%* (0.26%, 0.98%)
5 (40–49 yr)	0.79%* (0.23%, 1.36%)
6 (50–54 yr)	1.63%* (0.57%, 2.70%)
Region	
Central	0.56%* (0.01%, 1.11%)
East	−0.47% (−1.09%, 0.15%)
North	1.2%* (0.11%, 2.29%)
North-East	0.12% (−0.91%, 1.16%)
South	−0.20% (−0.57%, 0.18%)
West	0.06% (−0.27%, 0.39%)

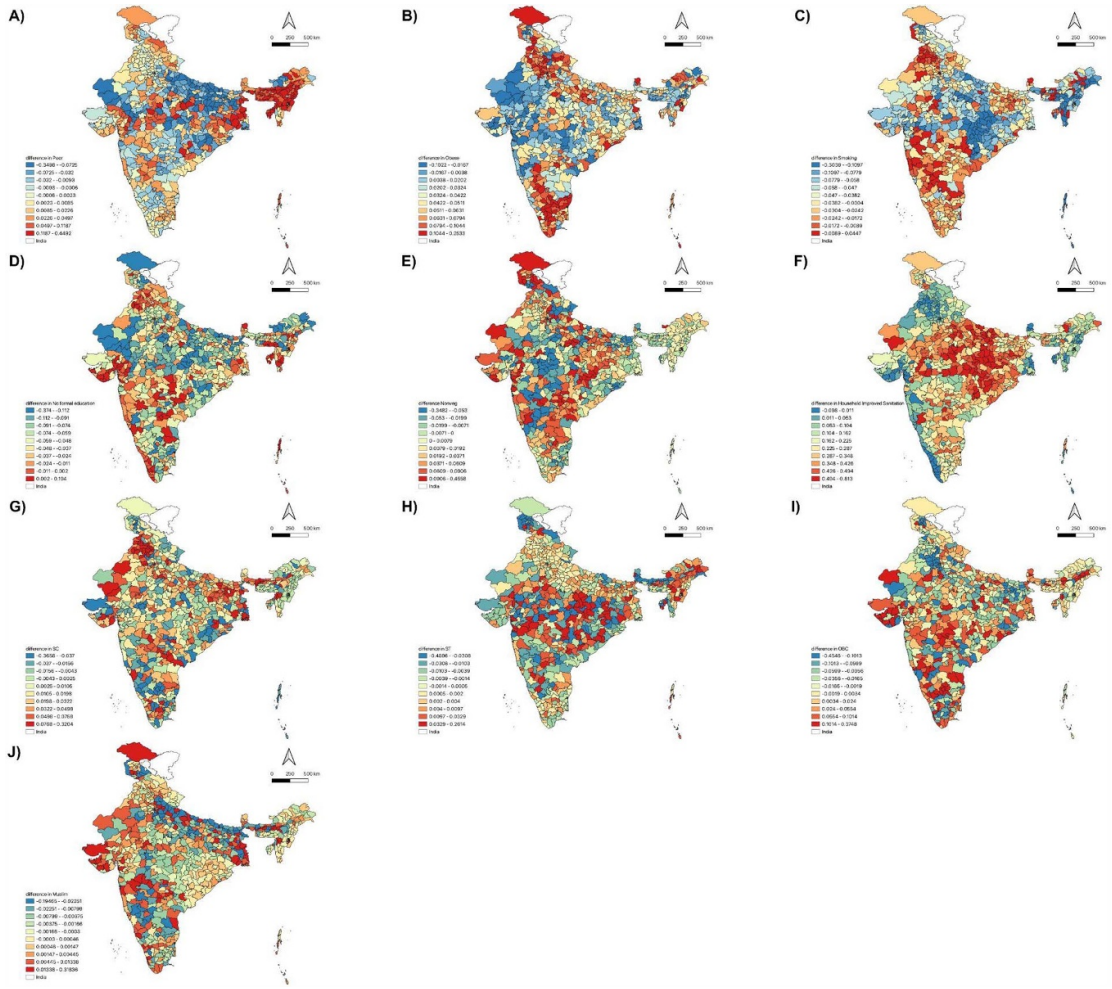

Figure 2. Difference between the prevalence (unit: fraction) of the following individuals in 2021 compared to 2016: (A) poor, (B) obese, (C) smokers, (D) Nn formal education, (E) nonvegetarian, (F) have improved sanitation facilities, (G) SC, (H) ST, (I) OBC, (J) Muslim, classified into deciles.

Region and age-group-specific analyzes revealed heterogeneity in associations between non-CF use and the prevalence of hypertension, likely due to the differences in populations, housing types, behavioral patterns, and cooking styles that impact exposure and vulnerability to hypertension (Roy 2024). Specifically, reducing the prevalence of non-CF use had the largest association with change in the prevalence of hypertension among older populations, an important vulnerable population for CVD. The strong associations between the change in the prevalence of non-CF use and hypertension among older populations is likely driving the higher association observed among men than women aged 15–54 yr: 0.58% (95% CI: −0.08%, 1.23%) vs. 0.23% (95% CI: −0.07%, 0.53%). We also note that the number of men in our sample is smaller, which explains the larger confidence intervals observed for men than women. In addition, the change in the prevalence of non-CF use was associated with a larger positive change in the health status of residents of North India compared to other regions. Such a difference could be observed because the populations in these regions are very different, region-specific policies are different, and thus baseline hypertension rates and vulnerabilities to exposure to indoor smoke are different; Moreover, cooking practices are different, which can impact exposures. More work needs to be conducted to evaluate why this was the case. Finally, associations between subpopulation-specific non-CF use and subpopulation-specific hypertension were smaller than between overall non-CF use and subpopulation-specific health hypertension, likely because the general reduction in the prevalence of non-CF use can lower overall pollution concentrations, which benefits all subpopulations living in a district. Although not significant, we observed that for every 10% increase in the prevalence of non-CF use at the district level, $PM_{2.5}$ levels were estimated to decrease by -0.12 μg m^{-3} (95% CI: -0.34 μg m^{-3}, 0.11 μg m^{-3}).

The results from this study are consistent with previous findings that have reported associations between indoor air pollution and the risk of hypertension. Specifically, a systematic review in 2020 based

on 11 studies (two of which were based in India) reported that the use of household solid fuel was significantly associated with an increased risk of hypertension, OR: 1.52 (95% CI: 1.26, 1.85) (Li *et al* 2020). That systematic review also identified substantial heterogeneity across the studies reviewed. The previous studies considered have been comprised of relatively small samples from small geographic regions.

Associations evaluated in this study between the change in non-CF use and the difference in district-level ambient $PM_{2.5}$ concentrations between 2015 and 2020, although not significant, were suggestive of a reduction in ambient concentrations. Other studies have reported that mitigating emissions from household energy sources can substantially improve India's ambient air quality levels via modeling approaches (Chowdhury *et al* 2019).

Our study has some limitations.

(1) This study used the prevalence of district-level non-CF use as a proxy for indoor air pollution. However, it is likely that due to differences in cooking patterns and housing conditions, among other factors, a reduction in non-CF use may not correspond to the same decreases in indoor pollution everywhere. This study attempts to account for likely differences in indoor pollution exposures by reporting results from stratified analyzes by region and for different subpopulations.
(2) This study considers the changes in the prevalence of primary fuel use as documented in the NFHS. However, previous research has suggested that fuel-stacking, or the use of multiple fuels in households, is common in India (Gurley *et al* 2013, Gould and Urpelainen 2018, Shupler *et al* 2020). If fuel stacking practices are similar across districts and unrelated to the outcome, hypertension, our exposure of interest is a noisy proxy for the true reduction in solid fuel exposure, and the estimated association is biased towards the null. If the degree of fuel stacking differs systematically across districts, then the bias direction is unpredictable. Future work should incorporate the impacts of fuel stacking that could lead to heterogeneous effects of the change in non-CF use.
(3) When evaluating associations between the change in the prevalence of non-CF use and the change in the prevalence of hypertension in our first difference models, this study makes the reasonable assumption that the population does not change substantially in the 5 years between the two NFHS surveys. We acknowledge, however, that other time-varying factors could impact our results.
(4) The $PM_{2.5}$ dataset used here is a satellite-derived estimate that has uncertainties, especially in rural areas where the scarcity of ground-based monitors precludes extensive evaluation.
(5) The NFHS-5 survey was conducted during the height of the COVID-19 pandemic in India, a period marked by widespread social and economic disruptions. These disruptions likely influenced household energy use—many people spent more time at home and cooked more frequently, while supply chain interruptions may have led to increased fuel stacking. As a result, the exposure variable used in this study, which captures the primary cooking fuel, may not accurately reflect true household exposure during the pandemic. If the correspondence between primary fuel type and actual exposure to household air pollution weakened during this period, the estimated associations are likely attenuated toward the null. However, if the degree of exposure misclassification varied across districts, the direction of bias becomes unpredictable. Additionally, reduced healthcare access and pandemic-related lifestyle changes—such as elevated stress levels and altered health behaviors—could have affected hypertension prevalence. For instance, if districts that continued to rely on solid fuels also experienced greater socioeconomic stress, resulting in higher hypertension rates, the observed associations would again be biased toward the null. The methodology that we have developed in this article can be used with later rounds of the NFHS to continue to assess the persistence of India's clean energy transition impacts on health outcomes.
(6) All district-level estimates used were derived from multi-level models and have some amount of uncertainty. Our study did not account for measurement errors in the modeled estimates.
(7) Our analysis is at the district level. Ecological regression analyzes are unable to adjust for individual-level risk factors, and thus, this approach leaves us unable to make conclusions regarding individual-level associations. Nevertheless, ecological regression analyzes still allow us to make conclusions at the area level, which can be useful for policy-making. Specifically, we observed that changes in non-CF use and the rising prevalence of LPG as the primary cooking fuel reported have likely resulted in significant decreases in the prevalence of hypertension in India.

Data Availability Statement.

NFHS data is available on submitting a request via the DHS website https://dhsprogram.com/.

The $PM_{2.5}$ data are freely available from https://sites.wustl.edu/acag/datasets/surface-pm2-5/.

The data cannot be made publicly available upon publication because they contain sensitive personal information. The data that support the findings of this study are available upon reasonable request from the authors.

Supplementary data 1 available at https://doi.org/10.1088/2752-5309/ae1755/data1.

Acknowledgments

R Kim was supported by the National Research Foundation of Korea(NRF) Grant No. RS-2023-00219289 funded by the Korean government (MSIT).

Conflict of interests

The authors declare no competing interests.

Author contributions

Priyanka deSouza 0000-0002-2618-4050
Conceptualization (lead), Data curation (lead), Formal analysis (lead), Investigation (lead), Methodology (lead), Project administration (lead), Resources (lead), Software (lead), Validation (lead), Visualization (lead), Writing – original draft (lead), Writing – review & editing (lead)

Jenny J Lee
Methodology (supporting), Writing – review & editing (equal)

Jeremy Németh
Writing – review & editing (equal)

Sunil Mani
Writing – review & editing (equal)

Abhishek Jain
Writing – review & editing (equal)

Abhishek Kar
Writing – review & editing (equal)

Jennifer Peel
Writing – review & editing (equal)

Sadeer Al-Kindi
Writing – review & editing (equal)

Patrick L Kinney
Writing – review & editing (equal)

Ajay Pillarisetti 0000-0003-0518-2934
Writing – review & editing (equal)

Wenlu Ye 0000-0002-2212-6160
Writing – review & editing (equal)

Rockli Kim
Conceptualization (supporting), Writing – review & editing (equal)

SV Subramanian
Conceptualization (supporting), Writing – review & editing (equal)

Michelle L Bell 0000-0002-3965-1359
Conceptualization (supporting), Writing – review & editing (equal)

Eric A F Simoes
Conceptualization (supporting), Writing – review & editing (equal)

References

Al-Kindi S G, Brook R D, Biswal S and Rajagopalan S 2020 Environmental determinants of cardiovascular disease: lessons learned from air pollution *Nat. Rev. Cardiol.* **17** 656–72

Bates D, Maechler M, Bolker B, Walker S, Christensen R H B, Singmann H, Dai B, Scheipl F and Grothendieck G 2009 Package 'lme4' URL Httplme4 R-Forge R-Proj. Org

Boing A F, deSouza P, Boing A C, Kim R and Subramanian S V 2022 Air pollution, socioeconomic status, and age-specific mortality risk in the United States *JAMA Netw. Open* **5** e2213540

Chobanian A and (the National High Blood Pressure Education Program Coordinating Committee) 2003 The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressurethe JNC 7 report *JAMA* **289** 2560–71

Chowdhury S, Dey S, Guttikunda S, Pillarisetti A, Smith K R and Di Girolamo L 2019 Indian annual ambient air quality standard is achievable by completely mitigating emissions from household sources *Proc. Natl Acad. Sci.* **116** 10711–6

deSouza P N *et al* 2023 An environmental justice analysis of air pollution in India *Sci. Rep.* **13** 16690

deSouza P N, Dey S, Mwenda K M, Kim R, Subramanian S V and Kinney P L 2022a Robust relationship between ambient air pollution and infant mortality in India *Sci. Total Environ.* **815** 152755

deSouza P N, Hammer M, Anthamatten P, Kinney P L, Kim R, Subramanian S V, Bell M L and Mwenda K M 2022b Impact of air pollution on stunting among children in Africa *Environ. Health* **21** 128

Government of India 2024 (available at: https://powermin.gov.in/en/content/saubhagya)

Goldemberg J, Martinez-Gomez J, Sagar A and Smith K R 2018 Household air pollution, health, and climate change: cleaning the air *Environ. Res. Lett.* **13** 030201

Gould C F and Urpelainen J 2018 LPG as a clean cooking fuel: adoption, use, and impact in rural India *Energy Policy* **122** 395–408

Gupta R, Gaur K and Ram C V 2019 Emerging trends in hypertension epidemiology in India *J. Hum. Hypertens.* **33** 575–87

Gurley E S *et al* 2013 Seasonal concentrations and determinants of indoor particulate matter in a low-income community in Dhaka, Bangladesh *Environ. Res.* **121** 11–16

Hammer M S *et al* 2020 Global estimates and long-term trends of fine particulate matter concentrations (1998–2018) *Environ. Sci. Technol.* **54** 7879–90

IIPS 2017 National family health survey (NFHS-4): 2014–15: India (Mumbai International Institute Population Sciences)

International Institute for Population Sciences (IIPS) & ICF 2017 National family health survey (NFHS-4), 2015–16: India (Fact Sheet No. FR339) (available at: https://dhsprogram.com/pubs/pdf/fr339/fr339.pdf) (Accessed 29 October 2025)

International Institute for Population Sciences (IIPS) & ICF 2021 National family health survey (NFHS-5), 2019–21: India – volume I (Final Report No. FR375) (Mumbai IIPS) (available at: https://dhsprogram.com/pubs/pdf/FR375/FR375.pdf) (Accessed 29 October 2025

Joshi P *et al* 2007 Risk factors for early myocardial infarction in south asians compared with individuals in other countries *JAMA* **297** 286–94

Li L, Yang A, He X, Liu J, Ma Y, Niu J and Luo B 2020 Indoor air pollution from solid fuels and hypertension: a systematic review and meta-analysis *Environ. Pollut.* **259** 113914

Maiti S, Akhtar S, Upadhyay A K and Mohanty S K 2023 Socioeconomic inequality in awareness, treatment and control of diabetes among adults in India: evidence from National family health survey of India (NFHS), 2019–2021 *Sci. Rep.* **13** 2971

Mensah G A, Roth G A and Fuster V 2019 The global burden of cardiovascular diseases and risk factors *J. Am. Coll. Cardiol.* **74** 2529–32

Pope D *et al* 2021 Are cleaner cooking solutions clean enough? A systematic review and meta-analysis of particulate and carbon monoxide concentrations and exposures *Environ. Res. Lett.* **16** 083002

Pope D, Bruce N, Dherani M, Jagoe K and Rehfuess E 2017 Real-life effectiveness of 'improved' stoves and clean fuels in reducing $PM_{2.5}$ and CO: systematic review and meta-analysis *Environ. Int.* **101** 7–18

Prabhakaran D, Jeemon P and Roy A 2016 Cardiovascular diseases in India *Circulation* **133** 1605–20

Rosenthal J, Quinn A, Grieshop A P, Pillarisetti A and Glass R I 2018 Clean cooking and the SDGs: integrated analytical approaches to guide energy interventions for health and environment goals *Energy Sustain. Dev.* **42** 152–9

Roth G A *et al* 2020 Global burden of cardiovascular diseases and risk factors, 1990–2019 *J. Am. Coll. Cardiol.* **76** 2982–3021

Roy K 2024 Transition to cooking with clean fuels in rural households of India: studying the effect of policy and other factors *Energy Sustain. Dev.* **80** 101456

Sciences I I for P 2020 National family health survey-5 2019–21 *Ministry Health Family Welfare National* vol 361 p 2

Shupler M *et al* 2020 Household and personal air pollution exposure measurements from 120 communities in eight countries: results from the PURE-AIR study *Lancet Planet. Health* **4** e451–62

Smith K R 2002 Indoor air pollution in developing countries: recommendations for research† *Indoor Air* **12** 198–207

Steenland K, Pillarisetti A, Kirby M, Peel J, Clark M, Checkley W, Chang H H and Clasen T 2018 Modeling the potential health benefits of lower household air pollution after a hypothetical liquified petroleum gas (LPG) cookstove intervention *Environ. Int.* **111** 71–79

Team R C *et al* 2013. R: a language and environment for statistical computing

Vaduganathan M, Mensah G A, Turco J V, Fuster V and Roth G A 2022 The global burden of cardiovascular diseases and risk *J. Am. Coll. Cardiol.* **80** 2361–71

who 2016 Global health estimates 2015: deaths by cause, age, sex, by country and by region, 2000–2015 (Geneva World Health Organ)