

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

Supplemental Material

Constituents of Household Air Pollution and Risk of Lung Cancer among Never-Smoking Women in Xuanwei and Fuyuan, China

Roel Vermeulen, George S. Downward, Jinming Zhang, Wei Hu, Lützen Portengen, Bryan A. Bassig, S. Katharine Hammond, Jason Y.Y. Wong, Jihua Li, Boris Reiss, Jun He, Linwei Tian, Kaiyun Yang, Wei Jie Seow, Jun Xu, Kim Anderson, Bu-Tian Ji, Debra Silverman, Stephen Chanock, Yunchao Huang, Nathaniel Rothman, and Qing Lan

Table of Contents

Figure S1. Map of Geological Coal Deposit Layers in Xuanwei and Fuyuan Counties of Yunnan, China (A) and associated lung cancer odds ratios for the most frequently used deposit (B). Coal deposits are designated by numbers. Coal type is based on the Chinese State Standard coal classification.

Figure S2. Comparative histograms comparing distribution of predicted exposure metrics by case/control status.

Figure S3. Correlation heatmap showing spearman correlations between exposures for the full population (including both cases and controls). Blue colours indicate negative correlations, red colours positive correlations. Exposures were ordered according to cluster membership (clusters were derived from the exposure data of controls) from the hierarchical clustering and sign and strength of the point estimate from the regularized regression model.

Figure S4. Comparison of coefficients for lung cancer risk from the regularized regression model fitted to lifetime cumulative exposure data for the full study population to those from a model fitted on data from subjects using almost exclusively smoky coal (used smokeless coal and/or for less than 2 years during their life). The (rescaled) point estimate for retene was off the (vertical) scale and is therefore not shown in the graph.

Table S1. PM_{2.5}, PM_{2.5} absorbance, NO₂, and SO₂ by main fuel and stove.

Table S2. Geometric mean (and GSD) for indoor measurements of expanded PAH analysis including methylated PAHs.

Text S1. Determinant-based modelling of exposure to household air pollutants.

Table S3. Sensitivity analysis – model comparison against different levels of regularization.

Table S4. Distribution of eligible population among communes in Xuanwei and Fuyuan.

Table S5. Spearman correlations between cluster variables in the full study population.

Table S6. Spearman correlations between individual exposures and cluster variables in the full study population.

Table S7. Spearman correlations between cluster variables in subjects using smoky coal almost exclusively.

Table S8. Spearman correlations between individual exposures and cluster variables in subjects using smoky coal almost exclusively.

Table S9. Odds ratios (95% CI) for individual clusters in the whole population.

Table S10. Odds ratios (95% CI) for individual clusters in the smoky coal subpopulation.

Table S11. Odds ratios (95%CI) for individual constituents within identified clusters of the smoky coal subpopulation.

Table S12. Comparing deposit specific odds ratios (95% CI) for models with and without household air pollution (HAP) components.

Table S13. Comparing deposit specific odds ratios (95% CI) for models with and without household air pollution (HAP) components among the smoky coal subpopulation.

References

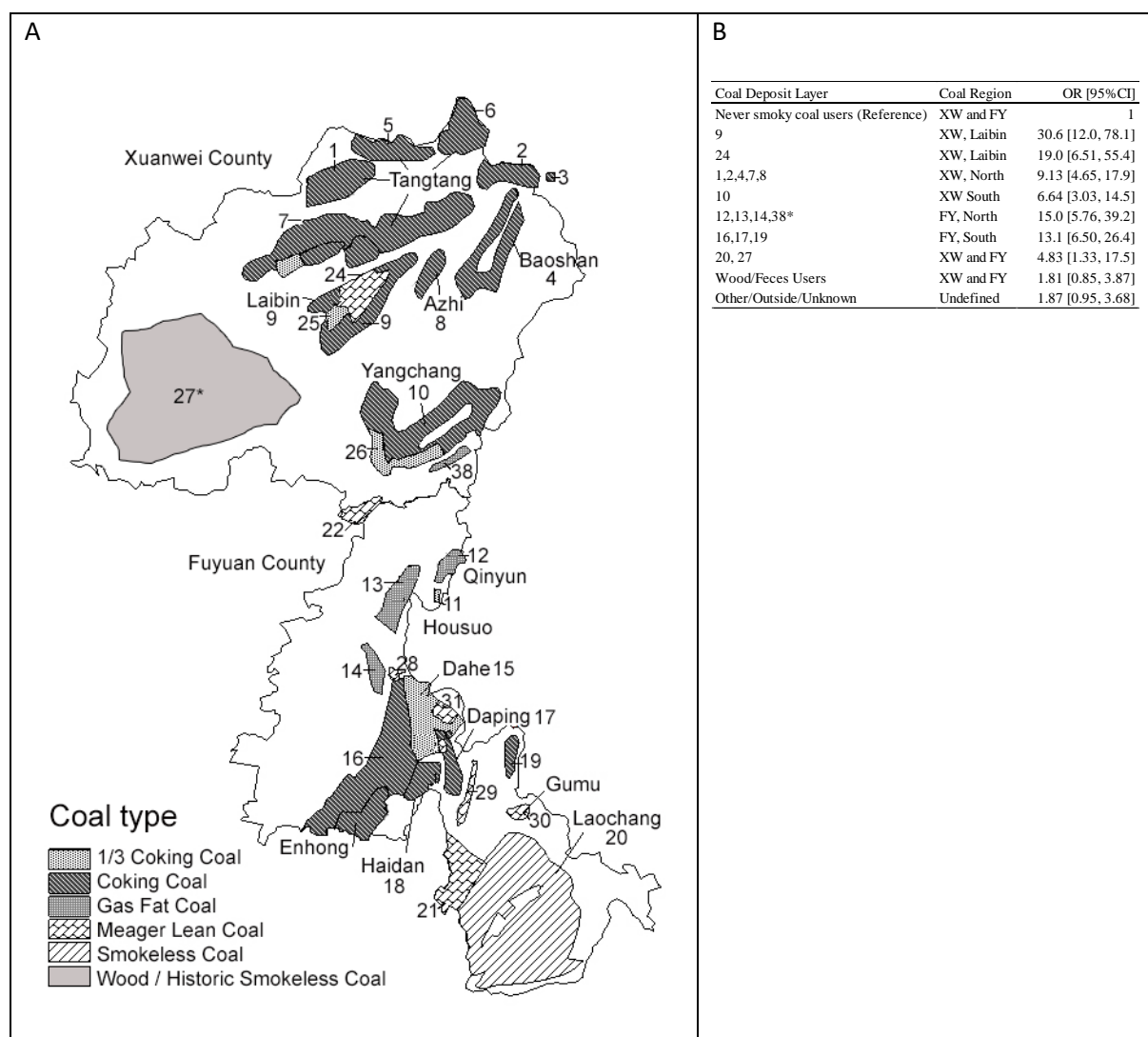
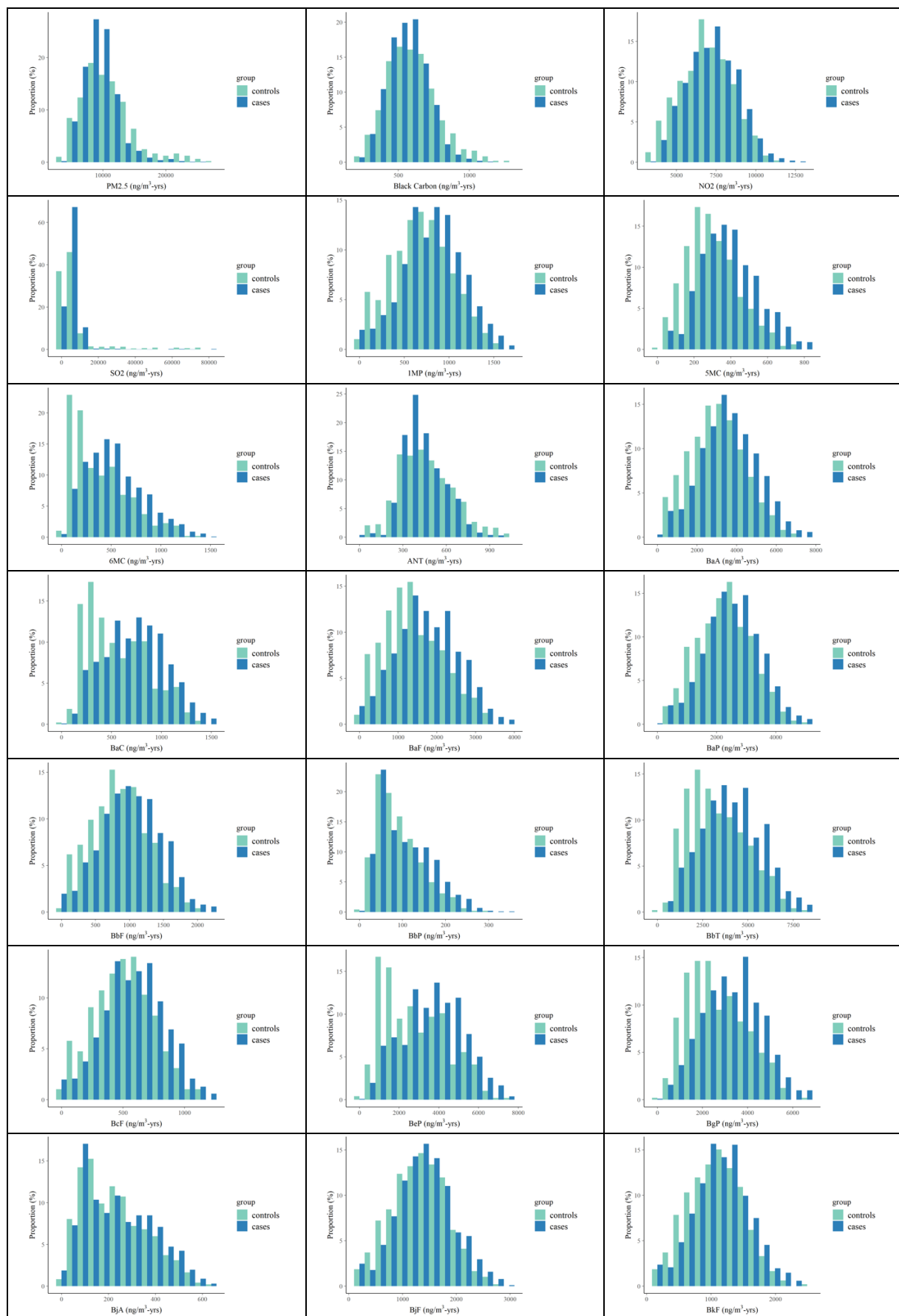


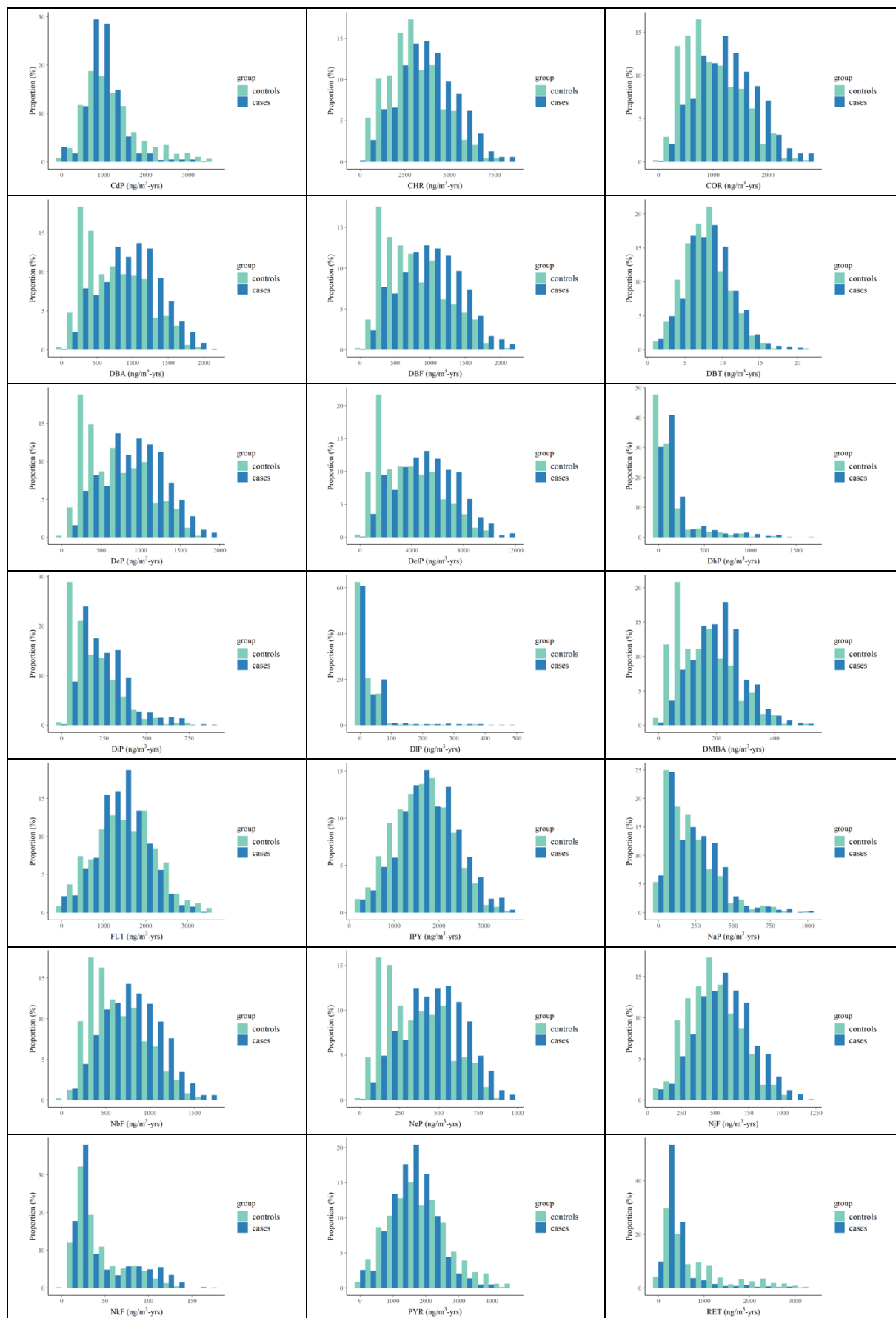
Figure S1. Map of Geological Coal Deposit Layers in Xuanwei and Fuyuan Counties of Yunnan, China (A) and associated lung cancer odds ratios for the most frequently used deposit (B). Coal deposits are designated by numbers. Coal type is based on the Chinese State Standard coal classification.

Odds ratios are derived from the same population contributing to the current publication and are adjusted for age and food sufficiency before marriage/20 years of age. Strata groupings were derived by combining coals of the same type and number of women using coal from each deposit. Deposits with sufficient users to allow separate analysis (9, 10, 24) were kept as such.

Map and odds ratios are reproduced with permission from Wong JYY, Downward GS, Hu W, et.al. *Lung cancer risk by geologic coal deposits: A case-control study of female never-smokers from Xuanwei and Fuyuan, China*. Int J Cancer. PMID: 30511435, <https://doi.org/10.1002/ijc.32034>. 2018.

*: Deposit 38 is in Xuanwei, the other deposits of this geological type (gas fat coal) are in Fuyuan.





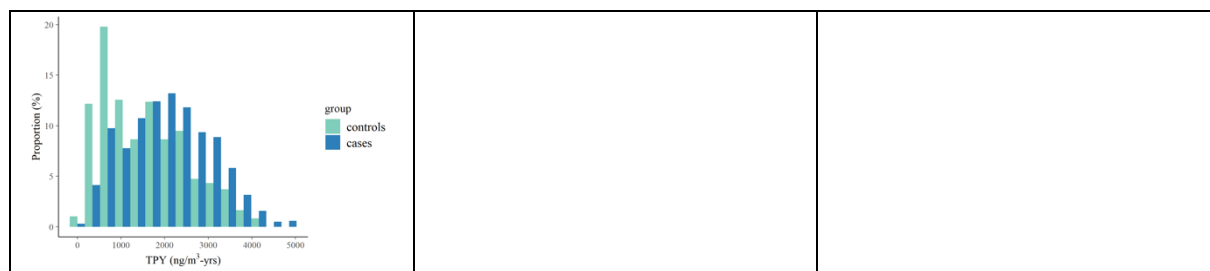


Figure S2. Comparative histograms comparing distribution of predicted exposure metrics by case/control status.

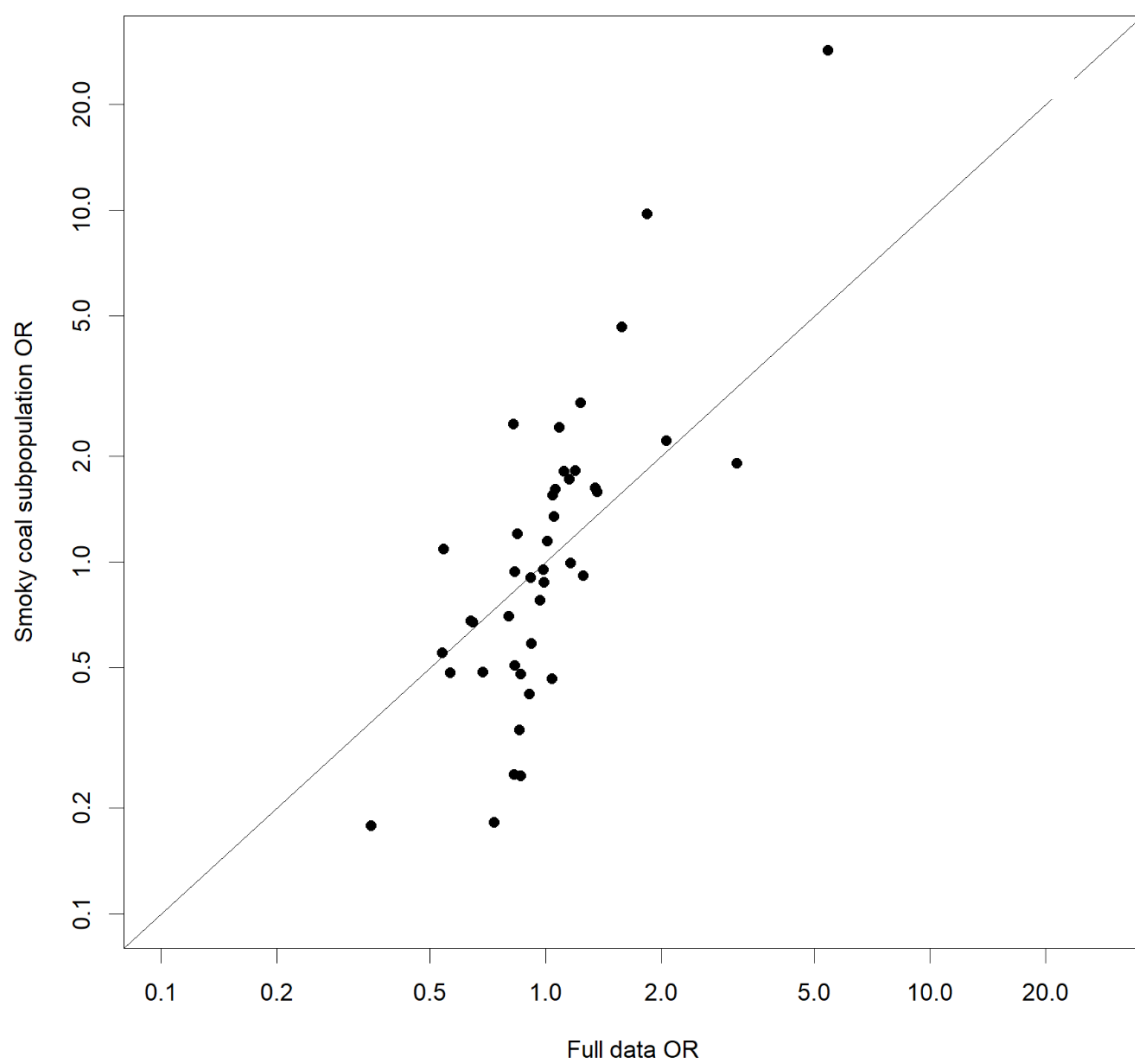


Figure S4. Comparison of coefficients for lung cancer risk from the regularized regression model fitted to lifetime cumulative exposure data for the full study population to those from a model fitted on data from subjects using almost exclusively smoky coal (used smokeless coal and/or for less than 2 years during their life). The (rescaled) point estimate for retene was off the (vertical) scale and is therefore not shown in the graph.

Table S1. PM_{2.5}, PM_{2.5} absorbance, NO₂, and SO₂ by main fuel and stove

| | PM _{2.5} | | PM _{2.5} absorbance | | NO ₂ | | SO ₂ | |
|----------------|-------------------|-------------|---------------------------------|-------------|-----------------|-------------|-----------------|---------------|
| | N | GM (GSD) | N | GM (GSD) | N | GM (GSD) | N | % Detected |
| Smoky Coal | 206 | 148 (1.9) | 202 | 12.0 (1.7) | 191 | 111 (1.4) | 191 | 40 |
| Vented | 110 | 134 (1.6) | 108 | 13.0 (1.6) | 105 | 102 (1.4) | 105 | 31 |
| Unvented | 8 | 233 (1.3) | 8 | 10.0 (1.4) | 6 | 113 (1.1) | 6 | 67 |
| Portable Stove | 22 | 143 (1.9) | 22 | 7.3 (2.0) | 19 | 118 (1.6) | 19 | 63 |
| Firepit | 15 | 277 (1.6) | 13 | 13.7 (1.4) | 15 | 132 (1.2) | 15 | 73 |
| Smokeless Coal | 47 | 115 (1.9) | 46 | 5.1 (1.7) | 42 | 132 (1.6) | 42 | 86 |
| Vented | 5 | 126 (2.0) | 5 | 6.1 (2.0) | 5 | 135 (1.5) | 5 | 80 |
| Unvented | 18 | 109 (2.1) | 17 | 4.8 (1.6) | 15 | 134 (1.5) | 15 | 93 |
| Portable Stove | 19 | 123 (1.9) | 19 | 5.1 (1.9) | 17 | 127 (1.7) | 17 | 88 |
| Firepit | 3 | 102 (1.3) | 3 | 3.7 (1.3) | 3 | 139 (1.9) | 3 | 100 |
| Wood | 24 | 289 (2.1) | 23 | 16.3 (1.4) | 24 | 110 (1.4) | 24 | 29 |
| Vented | 8 | 183 (1.9) | 8 | 13.5 (1.5) | 8 | 104 (1.4) | 8 | 38 |
| Unvented | 0 | - | 0 | - | 0 | - | 0 | - |
| Portable Stove | 6 | 320 (1.3) | 6 | 17.2 (1.2) | 5 | 90 (1.2) | 5 | 0 |
| Firepit | 10 | 392 (2.4) | 9 | 18.6 (1.4) | 10 | 128 (1.4) | 10 | 40 |

Values for "Mixed" fuels and stoves are not reported

N – number of samples collected

PM_{2.5} (in µg/m³) represents personal exposure from 163 women as reported by: Hu W, Downward GS, Reiss B, et.al. *Personal and indoor PM_{2.5} exposure from burning solid fuels in vented and unvented stoves in a rural region of China with a high incidence of lung cancer*. Environ Sci Technol 48:8456-8464, PMID: 25003800, <https://doi.org/10.1021/es502201s> (reprinted with permission)

PM_{2.5} absorbance (in 1 x 10⁻⁵ m⁻¹) represents personal exposure from 163 women as reported by: Downward GS, Hu W, Rothman N, et.al. *Outdoor, indoor, and personal black carbon exposure from cookstoves burning solid fuels*. Indoor Air 26:784-795, PMID: 26452237, <https://doi.org/10.1111/ina.12255> (reprinted with permission)

NO₂ (in µg/m³) and SO₂ (in % detected samples) represent indoor measurements from 163 households as reported by: Seow WJ, Downward GS, Wei H, et.al. *Indoor concentrations of nitrogen dioxide and sulfur dioxide from burning solid fuels for cooking and heating in Yunnan province, China*. Indoor Air 26:776-783, PMID: 26340585, <https://doi.org/10.1111/ina.12251> (reprinted with permission)

Table S2. Geometric mean (and GSD) for indoor measurements of expanded PAH analysis including methylated PAHs

| | Smoky Coal | | | | | Smokeless Coal | | | | | Wood | | | | |
|------|------------|------------|-------------|----------------|-------------|----------------|------------|------------|----------------|----------|------------|-------------|----------|----------------|------------|
| | Overall | Vented | Unvented | Portable Stove | Firepit | Overall | Vented | Unvented | Portable Stove | Firepit | Overall | Vented | Unvented | Portable Stove | Firepit |
| N | 37 | 21 | 3 | 8 | 5 | 20 | 6 | 4 | 9 | 1 | 8 | 3 | 0 | 1 | 4 |
| 1MP | 10.3 (4.4) | 6.6 (2.9) | 20.5 (3.3) | 12.8 (9.4) | 30.3 (3.2) | 0.4 (4.3) | 0.5 (6.6) | 0.6 (3.5) | 0.3 (4.1) | 0.3 (-) | 11 (3) | 21.8 (1.4) | - | 8.3 (-) | 7.1 (4) |
| 5MC | 6.5 (3.4) | 4.5 (2.5) | 13.2 (2.5) | 8 (6.4) | 15.6 (2.2) | 1.2 (3.2) | 2.6 (4.0) | 1.4 (2.3) | 0.8 (2.7) | 0.4 (-) | 3.7 (2.0) | 4.8 (1.9) | - | 2.4 (-) | 3.3 (2.3) |
| 6MC | 6.4 (4.2) | 4.4 (2.8) | 2.8 (10.6) | 9.9 (6.2) | 25.8 (2.0) | 0.7 (3.7) | 1 (6.8) | 0.5 (1.9) | 0.6 (3.3) | 1.4 (-) | 2.6 (2.2) | 3.1 (1.2) | - | 2.7 (-) | 2.3 (3.1) |
| ANT | 9.3 (2.9) | 7 (2.6) | 13.7 (2.8) | 12.2 (4) | 15.8 (2.4) | 0.9 (5.9) | 2.2 (5.7) | 0.9 (8.1) | 0.6 (4.6) | 0.1 (-) | 13.1 (2.4) | 29.2 (1.5) | - | 16.1 (-) | 6.8 (2) |
| BaA | 49.9 (3.5) | 34.9 (2.6) | 116 (3.4) | 56.7 (6.5) | 109.8 (2.1) | 4.3 (6.4) | 5.8 (8.8) | 6 (4.5) | 2.8 (7.5) | 8.3 (-) | 43.8 (2.5) | 85 (1.6) | - | 43.7 (-) | 26.7 (2.8) |
| BaC | 8.9 (3) | 5.7 (2.3) | 21.7 (2.9) | 12 (4.4) | 21.1 (1.8) | 2.7 (4) | 1.9 (4.6) | 2.4 (2.1) | 3.1 (4.9) | 9.1 (-) | 4.9 (2) | 7.4 (1.5) | - | 5.3 (-) | 3.5 (2.3) |
| BaF | 19.3 (5.1) | 11.4 (3.4) | 54 (6.1) | 23.3 (10.1) | 70.2 (3) | 1.1 (3.1) | 1.2 (4.4) | 1.3 (2.3) | 1 (3.2) | 1.5 (-) | 16.9 (2.8) | 30.1 (1.4) | - | 13.6 (-) | 11.6 (3.8) |
| BaP | 38 (2.9) | 28.7 (2.3) | 73.1 (2.6) | 41.8 (5) | 72.6 (2.1) | 7.8 (3.9) | 9.8 (6.1) | 7.3 (3.7) | 6.9 (3.6) | 7.9 (-) | 37.1 (2.3) | 73.3 (1.7) | - | 42.9 (-) | 21.4 (2.2) |
| BbF | 11.9 (5.1) | 7.4 (3.5) | 29.4 (7.3) | 14.1 (9.8) | 38.4 (3.2) | 0.9 (2.6) | 0.9 (3.7) | 1.1 (2) | 0.8 (2.7) | 1.1 (-) | 12.4 (2.8) | 23.3 (1.5) | - | 10.1 (-) | 8.1 (3.8) |
| BbP | 1.1 (3.7) | 1.1 (3.2) | 0.4 (1.1) | 1.1 (6.5) | 1.9 (3.5) | 0.5 (2.1) | 0.8 (2.7) | 0.6 (2.4) | 0.4 (1.1) | 0.4 (-) | 1.6 (2.8) | 2 (4.3) | - | 3.3 (-) | 1.1 (2.2) |
| BbT | 54 (3.0) | 35.4 (2.4) | 139 (2.5) | 74.6 (4.2) | 108.1 (1.9) | 17.3 (4.4) | 13 (5.5) | 19.7 (2.5) | 17.4 (5.4) | 55.3 (-) | 35.6 (2.2) | 62.6 (1.6) | - | 40.6 (-) | 22.5 (2.2) |
| BcF | 7.1 (4.9) | 4.5 (3.3) | 17 (5.8) | 8.3 (10.1) | 22.7 (3.1) | 0.3 (3.9) | 0.3 (5.6) | 0.4 (3) | 0.2 (4.1) | 0.3 (-) | 7.6 (3) | 15.1 (1.4) | - | 6.1 (-) | 4.8 (4) |
| BeP | 48.8 (3.1) | 31.1 (2.5) | 117.7 (2.2) | 67.6 (4.5) | 113 (1.7) | 10.2 (4.2) | 9.2 (6.1) | 9.9 (2.3) | 10.2 (4.8) | 23.8 (-) | 20.9 (2) | 31.4 (1.5) | - | 24.2 (-) | 14.9 (2.4) |
| BgP | 47.3 (2.9) | 32.5 (2.4) | 97.2 (2.2) | 60.4 (4.5) | 99.9 (1.9) | 10.1 (3.3) | 11.4 (4.8) | 8.6 (2.6) | 9.6 (3.3) | 13.1 (-) | 27.9 (2) | 47 (1.5) | - | 34.1 (-) | 17.9 (2) |
| BjA | 2.9 (5.9) | 2.1 (4.7) | 4.2 (8.9) | 3.3 (10.7) | 6.4 (5.2) | 0.4 (1.8) | 0.5 (2.8) | 0.4 (1.1) | 0.4 (1.5) | 0.4 (-) | 3.6 (4.5) | 4.3 (8.3) | - | 5.1 (-) | 2.9 (4.4) |
| BjF | 22.2 (2.9) | 17.4 (2.3) | 41.3 (2.6) | 23.6 (5) | 39.3 (2.1) | 4.8 (3.8) | 5.9 (5.8) | 4.5 (3.9) | 4.2 (3.5) | 4.5 (-) | 22.7 (2.3) | 45.1 (1.8) | - | 25.7 (-) | 13.1 (2.1) |
| BkF | 17.8 (2.9) | 14.1 (2.3) | 32.3 (2.9) | 18.6 (5) | 30.7 (2.1) | 3.9 (3.7) | 4.6 (5.4) | 3.9 (3.5) | 3.5 (3.6) | 4.5 (-) | 17.8 (2.3) | 36 (1.7) | - | 21.9 (-) | 10 (2.1) |
| CdP | 17 (4.2) | 13.8 (3.2) | 30.4 (4) | 14.2 (8.6) | 38.2 (3.1) | 1 (6.9) | 2 (8.9) | 1.9 (6.6) | 0.6 (5.8) | 0.1 (-) | 38.8 (4.3) | 113.4 (1.6) | - | 38.3 (-) | 17.4 (5.3) |
| CHR | 46.9 (3.3) | 30.1 (2.5) | 128.2 (3.2) | 60 (5.3) | 111.4 (2) | 6.3 (7.3) | 6.2 (8.8) | 9 (4) | 4.7 (10) | 24 (-) | 39.2 (2.4) | 70.6 (1.6) | - | 35.2 (-) | 26 (2.7) |
| COR | 16.8 (3) | 11.1 (2.5) | 35.8 (2.2) | 22.9 (4.5) | 37.2 (1.5) | 3.6 (3.4) | 3.7 (4.7) | 3.2 (2.7) | 3.5 (3.6) | 6.3 (-) | 9.9 (2.2) | 17.8 (1.4) | - | 12.7 (-) | 6 (2.2) |
| DBA | 11.8 (3.1) | 7.9 (2.4) | 26.7 (3.2) | 14.6 (4.6) | 27.3 (1.9) | 2.8 (4.1) | 2.3 (5.8) | 2.2 (2.2) | 3.1 (4.6) | 8.5 (-) | 5.8 (2) | 8.6 (1.5) | - | 7 (-) | 4.2 (2.3) |
| DBF | 12.1 (3.1) | 8.2 (2.5) | 26.9 (2.9) | 14.6 (4.8) | 28.1 (1.8) | 2.7 (3.8) | 2.6 (4.9) | 2.1 (2.5) | 2.8 (4.3) | 7.2 (-) | 6.9 (2.1) | 10.5 (1.5) | - | 8.7 (-) | 4.7 (2.4) |
| DBT | 0.1 (3.4) | 0.1 (1.6) | 0.4 (6.3) | 0.2 (6.7) | 0.3 (2.5) | 0.1 (2.6) | 0.1 (1.7) | 0.1 (1.9) | 0.1 (3.6) | 0.2 (-) | 0.1 (2.5) | 0.1 (2.3) | - | 0.1 (-) | 0.1 (3.3) |
| DeP | 12.2 (2.9) | 8 (2.1) | 26.3 (2.5) | 16.4 (4.7) | 28.9 (1.5) | 3.4 (2.9) | 3.4 (3.4) | 2.5 (1.5) | 3.7 (3.4) | 7.1 (-) | 4.6 (1.9) | 5.9 (1.4) | - | 6 (-) | 3.5 (2.4) |
| DeIP | 64.1 (3.2) | 39.9 (2.6) | 142.6 (2.1) | 89.8 (4.8) | 168.6 (1.4) | 11.6 (3.7) | 11.2 (5.3) | 9.1 (2.3) | 12.1 (4.1) | 28.2 (-) | 22.4 (2.2) | 33 (1.4) | - | 29.5 (-) | 15.6 (2.8) |
| DhP | 1 (8.1) | 0.9 (6.5) | 0.6 (16.2) | 0.6 (14.1) | 4.2 (4.6) | 0.1 (2.7) | 0.2 (6.4) | 0.1 (1.1) | 0.1 (1.1) | 0.1 (-) | 0.4 (6) | 0.4 (6.1) | - | 3.1 (-) | 0.3 (7.2) |

| | | | | | | | | | | | | | | | |
|------|------------|------------|------------|------------|------------|-----------|-----------|-----------|-----------|----------|------------|-------------|---|----------|-------------|
| DiP | 3.5 (5.3) | 3.5 (4.2) | 4 (8.9) | 1.9 (9.6) | 8.2 (3.8) | 0.7 (3.8) | 1.3 (5.3) | 0.3 (1.1) | 0.6 (4) | 0.3 (-) | 1.4 (5.2) | 0.9 (6.1) | - | 6.4 (-) | 1.4 (6) |
| DIP | 0.4 (9.8) | 0.3 (6.7) | 4.4 (26.9) | 0.7 (21.5) | 0.2 (1.9) | 1.3 (5.8) | 2.8 (4.9) | 0.7 (4.5) | 0.8 (7.2) | 4.2 (-) | 0.2 (3.7) | 0.1 (1.1) | - | 0.1 (-) | 0.3 (6.3) |
| DMBA | 3.1 (3.5) | 2.1 (2.6) | 5.1 (1.8) | 4 (6.9) | 7.6 (2.8) | 0.4 (3.2) | 0.8 (6.5) | 0.2 (1.1) | 0.3 (1.9) | 0.2 (-) | 1.2 (2.1) | 1.4 (1.3) | - | 1 (-) | 1.1 (2.9) |
| FLT | 17 (4.8) | 10.8 (3.1) | 39.9 (7.1) | 20.7 (9.5) | 49.6 (3.4) | 0.9 (2.3) | 0.9 (2.1) | 1.2 (2) | 0.8 (2.7) | 0.7 (-) | 37.8 (4.4) | 96.2 (1.5) | - | 24.9 (-) | 20.8 (6.6) |
| IPY | 26.6 (2.7) | 20.9 (2.2) | 50.2 (2.7) | 29.7 (4.3) | 42.2 (2.1) | 8.2 (3.1) | 8.7 (4.2) | 7.8 (3.1) | 7.8 (3.1) | 10.2 (-) | 29 (2.1) | 55.9 (1.7) | - | 33.5 (-) | 17.1 (1.7) |
| NaP | 2.2 (5.8) | 2.5 (4.9) | 1.1 (6.9) | 1.1 (9.8) | 5.2 (4) | 0.4 (1.6) | 0.5 (2.4) | 0.4 (1.1) | 0.4 (1.1) | 0.4 (-) | 3 (3.9) | 6.2 (1.2) | - | 7.3 (-) | 1.4 (5.3) |
| NbF | 10.6 (2.8) | 7.3 (2.1) | 22.7 (2.7) | 13.7 (4.4) | 21.9 (1.6) | 3.9 (2.8) | 3.6 (2.9) | 3.1 (2) | 4.2 (3.4) | 9 (-) | 6.2 (1.9) | 9.4 (1.6) | - | 7.4 (-) | 4.4 (2) |
| NeP | 6.2 (2.9) | 4.2 (2.2) | 12.8 (2.8) | 7.8 (4.9) | 14.1 (1.6) | 1.3 (3.1) | 1.3 (4.4) | 1.1 (1.6) | 1.2 (3.3) | 2.3 (-) | 2.6 (2.5) | 4.2 (1.5) | - | 3.6 (-) | 1.7 (3.3) |
| NjF | 8.5 (2.7) | 6.7 (2.1) | 14.3 (2.8) | 9.5 (4.7) | 15.1 (1.8) | 2.8 (2.3) | 3.4 (2.5) | 2.2 (2.4) | 2.7 (2.4) | 3.6 (-) | 7.1 (2) | 12 (1.7) | - | 8.5 (-) | 4.6 (1.8) |
| NkF | 0.7 (3.1) | 0.5 (2.2) | 1.6 (11.2) | 1 (4.4) | 0.6 (1.9) | 0.9 (3) | 2.1 (3.1) | 0.7 (3.2) | 0.6 (2.4) | 0.4 (-) | 1 (4.2) | 1.2 (8) | - | 0.3 (-) | 1.2 (3.4) |
| PYR | 19.7 (4.5) | 12.8 (2.9) | 49.3 (6.1) | 22.8 (9.3) | 56.8 (3.5) | 1 (2.6) | 1 (2.3) | 1.6 (2.7) | 0.9 (3) | 0.8 (-) | 46.7 (3.7) | 112.1 (1.5) | - | 31.2 (-) | 26.8 (5.2) |
| RET | 6.1 (4.3) | 4.4 (3.7) | 10.4 (2.7) | 7.6 (8.3) | 11.3 (2.7) | 1.3 (3.9) | 2.4 (4.5) | 1.4 (6.5) | 0.9 (2.9) | 0.6 (-) | 33.1 (8) | 54.4 (4) | - | 2.7 (-) | 42.5 (12.5) |
| TPY | 23.7 (3.7) | 13.7 (2.7) | 70.4 (2.3) | 33.6 (5.9) | 70.6 (1.8) | 2.6 (6.4) | 2.6 (8.5) | 3.2 (3.1) | 2 (8.3) | 10.4 (-) | 10.7 (2.1) | 15.2 (1.4) | - | 11.4 (-) | 8 (2.7) |

N – number of samples

1MP – 1-methylpyrene, 5MC – 5-methyl chrysene, 6MC – 6-methylchrysene, ANT – anthanthrene, BaA – benz(a)anthracene, BaC – benzo(a)chrysene, BaF – benzo(a)fluorene, BaP – benzo(a)pyrene, BbF – benzo(b)fluorene, BbP – benzo(b)perylene, BbT – benzo(b)fluoranthene, BcF – benzo(c)fluorene, BeP – benzo(e)pyrene, BgP – benzo(g,h,i)perylene, BjA – benz(j,e)aceanthrylene, BjF – benzo(j)fluoranthene, BkF – benzo(k)fluoranthene, CdP – cyclopenta(c,d) pyrene, CHR – chrysene, COR – coronene, DBA – dibenzo(a,h)anthracene, DBF – dibenzo(a,e)fluoranthene, DBT – dibenzothiophene, DeP – dibenzo(a,e,)pyrene, DelP – dibenzo(e,l)pyrene, DhP – dibenzo(a,h)pyrene, DiP – dibenzo(a,i)pyrene, DIP – dibenzo(a,l)pyrene, DMBA – 7,12-dimethylbenz.a.anthracene, FLT – fluoranthene, IPY – indeno (1,2,3)pyrene, NaP – naphtho(2,3,a)pyrene, NbF – naphtho(1,2,b)fluoranthene, NeP – naphtho(2,3,e)pyrene, NjF – naphtho(2,3,j)fluoranthene, NkF – naphtho(2,3,k)fluoranthene, PYR – pyrene, RET – retene, TPY – triphenylene

Concentrations in ng/m³

Samples were sourced from a subgroup of the population identified in Table S1 above. Selection was targeted towards representing stove and fuel variety in the study population

Text S1. Determinant-based modelling of exposure to household air pollutants

We enrolled 163 households and their never-smoking female heads, selected from 30 villages throughout Xuanwei and Fuyuan for an exposure survey in which multiple indoor and personal air samples were collected. Village and subject selection was targeted to represent the major geographical regions, solid fuels, and stove designs in use and to provide a population reflective of that within the case-control study. To that end, households that were at least ten years old and had not had any stove alterations undertaken in the past five years were preferentially selected.

Personal and indoor measurements were collected over 2 sequential 24-hour periods in 2008 and 2009. Samples of PM_{2.5}, black carbon (BC), quartz, trace elements, and particle bound PAHs and mPAHs were collected on 37mm teflon filters. Indoor measurements of NO₂ and SO₂ were collected using passively diffusing Ogawa filters. Radon was evaluated through the deployment of 90 day passively collecting detectors. Approximately 50% of subjects were visited again in a second season to allow seasonal adjustment of findings. Predictive linear mixed effect models were constructed for each pollutant, in which villages and individual subjects were assigned as random effects.(Downward 2015; Downward et al. 2016; Hu et al. 2014; Seow et al. 2016) Model construction was performed under a supervised step-wise procedure wherein variables were individually considered for model inclusion on the basis of goodness of fit and Akaike information criteria (AIC). A variety of potential predictors (including fuel type, fuel source, amount of fuel used, stove design, room size, number of windows/doors, the presence/absence of stairways, meteorological conditions, and season of measurement) were considered for inclusion as fixed effects. For all exposure variables (except for NO₂) the fuel type and/or source was identified as an important exposure predictor. The type of stove in use was also identified as a major predictor for exposure variables.

Quartz analyses of collected air samples resulted in a high proportion (89%) of non-detects with detected levels at or slightly above the limit of detection (LOD = 0.2 µg/m³). In addition, review of fibre characteristics by scanning electronic microscopy did not reveal asbestiform fibers.(Downward

et al. 2017) Radon was only detected in 50% of households with an average radon concentration of 3.4 pCi/L and there was no evidence of a difference in radon detection between homes using smoky vs. smokeless coal. Trace elements (Silicon, Aluminium, Arsenic, Bromine, Cadmium, Calcium, Chromium, Copper, Manganese, Nickel, Magnesium, Phosphorus, Lead, Sulphur, Selenium, Chlorine, Iron, Potassium, Titanium, Zinc, Vanadium, Yttrium, Molybdenum) were measured on filters using Energy Dispersive X-ray Fluorescence (EDXRF) and showed no evidence of difference between fuel or stove types.(Downward 2015)

The predictive models derived from the exposure survey were applied to the self-reported histories of stove and fuel use within the case-control dataset to assign annual geometric mean predicted exposure levels. For case-control subjects who reported using more than one fuel/stove in a given year, predictions from each permutation was combined based on the proportion of time spent using each permutation (e.g. a person using smoky coal and wood in equal amounts would have their final assignment calculated from a 50% smoky coal and 50% wood prediction).

References:

- Downward GS. 2015. Quantification and characterization of household air pollution exposure from the use of solid fuels; clues to the lung cancer epidemic in Xuanwei and Fuyuan, China:Utrecht University (PhD thesis), <https://dspace.library.uu.nl/handle/1874/312312>. ISBN: 978-90-393-6347-8.
- Downward GS, Hu W, Rothman N, Reiss B, Wu G, Wei F, et al. 2016. Outdoor, indoor, and personal black carbon exposure from cookstoves burning solid fuels. *Indoor Air* 26:784-795, PMID: 26452237, <https://doi.org/10.1111/ina.12255>.
- Downward GS, Hu W, Rothman N, Reiss B, Tromp P, Wu G, et al. 2017. Quartz in ash, and air in a high lung cancer incidence area in China. *Environ Pollut* 221:318-325, PMID: 27939206, <https://doi.org/10.1016/j.envpol.2016.11.081>.
- Hu W, Downward GS, Reiss B, Xu J, Bassig BA, Hosgood HD, 3rd, et al. 2014. Personal and indoor PM2.5 exposure from burning solid fuels in vented and unvented stoves in a rural region of China with a high incidence of lung cancer. *Environ Sci Technol* 48:8456-8464, PMID: 25003800, <https://doi.org/10.1021/es502201s>.
- Seow WJ, Downward GS, Wei H, Rothman N, Reiss B, Xu J, et al. 2016. Indoor concentrations of nitrogen dioxide and sulfur dioxide from burning solid fuels for cooking and heating in Yunnan province, China. *Indoor Air* 26:776-783, PMID: 2634058533, <https://doi.org/10.1111/ina.12251>.

Table S3. Sensitivity analysis – model comparison against different levels of regularization

| Cluster | Exposure | Population SD | Effective model size for Horseshoe analysis | | | Limited regularization model* |
|---------|----------|------------------|---|----------------------|----------------------|-------------------------------------|
| | | | 25 (main model) | 4 | 0.04 | - |
| WB7 | RET | 741 | 0.81 (0.24, 1.61) | 0.82 (0.24, 1.53) | 0.68 (0.25, 1.44) | 0.48 (0.01, 19.18) |
| WB7 | BC | 185 | 0.82 (0.41, 1.25) | 0.9 (0.49, 1.29) | 0.95 (0.53, 1.26) | 0.59 (0.25, 1.38) |
| WB7 | FLT | 731 | 0.83 (0.14, 2.46) | 0.84 (0.17, 2.17) | 0.82 (0.18, 1.94) | 1.91 (0.01, 417.29) |
| WB7 | ANT | 193 | 0.83 (0.28, 1.47) | 0.83 (0.29, 1.43) | 0.87 (0.32, 1.35) | 0.13 (0.01, 1.46) |
| WB7 | PYR | 906 | 0.86 (0.17, 2.37) | 0.84 (0.16, 2.08) | 0.88 (0.2, 1.96) | 2.71 (0.01, 809.28) |
| WB7 | CdP | 694 | 0.91 (0.24, 2.28) | 0.9 (0.25, 2.41) | 0.93 (0.32, 1.87) | 1.32 (0, 337.62) |
| WB7 | PM2.5 | 4263 | 1.05 (0.68, 1.85) | 1.02 (0.65, 1.62) | 1.01 (0.68, 1.6) | 1.23 (0.54, 2.85) |
| SO2 | SO2 | 12584 | 0.86 (0.58, 1.12) | 0.87 (0.59, 1.13) | 0.88 (0.6, 1.11) | 0.83 (0.49, 1.43) |
| PAH2 | NkF | 28 | 0.54 (0.27, 1.02) | 0.58 (0.29, 1.02) | 0.68 (0.31, 1.06) | 0.54 (0.19, 1.57) |
| PAH2 | DIP | 41 | 1.16 (0.95, 1.53) | 1.17 (0.95, 1.53) | 1.12 (0.96, 1.5) | 1.67 (1.07, 2.64) |
| PAH7 | NaP | 165 | 0.73 (0.24, 1.19) | 0.77 (0.27, 1.2) | 0.83 (0.31, 1.16) | 0.13 (0.02, 0.83) |
| PAH7 | BbP | 52 | 0.99 (0.59, 1.64) | 1 (0.63, 1.59) | 0.99 (0.66, 1.5) | 0.95 (0.18, 5.14) |
| PAH7 | 6MC | 285 | 1.01 (0.56, 2.01) | 1 (0.57, 1.88) | 1.01 (0.6, 1.66) | 2.15 (0.18, 27.49) |
| PAH7 | DMBA | 95 | 1.04 (0.43, 3.05) | 1.01 (0.42, 2.78) | 1.02 (0.51, 2.27) | 1.11 (0.03, 38.66) |
| PAH7 | DiP | 131 | 1.15 (0.7, 2.44) | 1.11 (0.68, 2.21) | 1.08 (0.73, 2.14) | 3.75 (0.5, 31.38) |
| PAH7 | DhP | 194 | 1.19 (0.82, 2.27) | 1.21 (0.83, 2.23) | 1.15 (0.85, 2.12) | 1.32 (0.37, 4.76) |
| PAH7 | BjA | 133 | 1.23 (0.68, 4.08) | 1.21 (0.69, 3.74) | 1.08 (0.66, 3.29) | 4.14 (0.38, 51.35) |
| PAH25 | BaC | 304 | 0.35 (0, 1.56) | 0.42 (0.01, 1.57) | 0.53 (0.01, 1.53) | 0.05 (0, 7.86) |
| PAH25 | BbF | 414 | 0.54 (0.05, 1.63) | 0.5 (0.04, 1.5) | 0.53 (0.05, 1.35) | 0.6 (0.01, 37.07) |
| PAH25 | NbF | 305 | 0.56 (0.02, 1.9) | 0.62 (0.03, 1.87) | 0.69 (0.04, 1.64) | 0.22 (0, 27.73) |
| PAH25 | DeP | 390 | 0.64 (0.04, 2.06) | 0.65 (0.04, 1.92) | 0.64 (0.05, 1.76) | 0.23 (0, 27.91) |
| PAH25 | BaF | 750 | 0.65 (0.06, 2.03) | 0.62 (0.06, 1.75) | 0.71 (0.07, 1.56) | 0.34 (0, 22.96) |
| PAH25 | DBA | 423 | 0.69 (0.04, 2.15) | 0.73 (0.06, 2.06) | 0.8 (0.1, 1.76) | 0.19 (0, 24.3) |
| PAH25 | BjF | 511 | 0.8 (0.11, 2.29) | 0.8 (0.11, 2.11) | 0.84 (0.14, 1.89) | 2.1 (0.02, 387.04) |
| PAH25 | IPY | 633 | 0.83 (0.12, 2.18) | 0.82 (0.13, 1.99) | 0.84 (0.14, 1.79) | 0.13 (0, 9.43) |

| | | | | | | |
|-------|------|------|------------------------|-----------------------|------------------------|--------------------------|
| PAH25 | DBT | 3 | 0.85 (0.33, 1.49) | 0.85 (0.34, 1.45) | 0.65 (0.23, 1.38) | 2.02 (0.3, 13.01) |
| PAH25 | BbT | 1570 | 0.86 (0.14, 2.35) | 0.84 (0.14, 2.09) | 0.87 (0.16, 1.92) | 0.14 (0, 10.73) |
| PAH25 | NjF | 191 | 0.91 (0.21, 2.82) | 0.94 (0.26, 2.81) | 0.93 (0.27, 2.12) | 1.5 (0.02, 128.2) |
| PAH25 | BkF | 422 | 0.92 (0.18, 2.89) | 0.89 (0.17, 2.61) | 0.85 (0.18, 2.13) | 1.25 (0.01, 163.41) |
| PAH25 | DBF | 425 | 0.97 (0.2, 3.88) | 0.98 (0.25, 3.51) | 0.91 (0.29, 2.76) | 1.15 (0.01, 136.81) |
| PAH25 | BaP | 917 | 0.99 (0.23, 4) | 0.97 (0.23, 3.81) | 0.99 (0.29, 2.95) | 3.99 (0.02, 1727.02) |
| PAH25 | BeP | 1577 | 1.04 (0.27, 4.99) | 1.02 (0.26, 4.24) | 1.02 (0.35, 3.1) | 0.56 (0, 77.6) |
| PAH25 | BcF | 239 | 1.06 (0.32, 4.9) | 1.03 (0.3, 4.09) | 1.01 (0.37, 3.2) | 1.33 (0.02, 116.62) |
| PAH25 | DeLP | 2321 | 1.09 (0.34, 5.88) | 1.07 (0.36, 4.76) | 1.08 (0.44, 3.64) | 2.92 (0.03, 455.85) |
| PAH25 | NeP | 197 | 1.12 (0.28, 8.24) | 1.08 (0.27, 6.72) | 1.06 (0.37, 4.65) | 1.24 (0.01, 225.49) |
| PAH25 | CHR | 1495 | 1.25 (0.37, 11.99) | 1.25 (0.4, 13.08) | 1.18 (0.46, 9.36) | 1.09 (0.01, 98.49) |
| PAH25 | 1MP | 331 | 1.34 (0.51, 11.34) | 1.32 (0.54, 10.6) | 1.81 (0.59, 11.54) | 0.52 (0.01, 25.06) |
| PAH25 | TPY | 962 | 1.36 (0.49, 16.28) | 1.36 (0.52, 14.62) | 1.24 (0.56, 13.23) | 2.73 (0.04, 292.08) |
| PAH25 | COR | 514 | 1.58 (0.58, 29.53) | 1.42 (0.55, 18.29) | 1.34 (0.64, 12.58) | 13.37 (0.12, 4739.15) |
| PAH25 | BgP | 1277 | 1.84 (0.52, 74.45) | 1.66 (0.54, 44.93) | 1.87 (0.59, 33.52) | 2.69 (0.02, 555.79) |
| PAH25 | BaA | 1373 | 3.15 (0.65, 134.58) | 4 (0.68, 218.69) | 3.21 (0.75, 186.93) | 2.5 (0.03, 272.31) |
| PAH25 | 5MC | 142 | 5.42 (0.94, 37.49) | 5.12 (0.91, 33.35) | 3.27 (0.86, 32.63) | 15.39 (1.15, 242) |
| NO2 | NO2 | 1638 | 2.06 (1.19, 3.49) | 1.9 (1.1, 3.24) | 1.81 (1.1, 3.1) | 2.52 (1.35, 4.68) |

* : Student t prior with 7 degrees of freedom and a scale of 2.5

All models were adjusted for age, self-reported exposure to second-hand smoke, and self-reported food-sufficiency before marriage

Table S4. Distribution of eligible population among communes in Xuanwei and Fuyuan

| County | Commune | N (percentage in county) |
|---------|------------|-----------------------------|
| Xuanwei | | 442,975 |
| | Rongcheng | 48,356 (11%) |
| | Laibin | 29,661 (7%) |
| | Geyi | 17,497 (4%) |
| | Tianba | 23,940 (5%) |
| | Yangchang | 20,445 (5%) |
| | Banqiao | 27,835 (6%) |
| | Tangtang | 22,472 (5%) |
| | Jingwuai | 5,696 (1%) |
| | Longchang | 17,014 (4%) |
| | Baoshan | 19,619 (4%) |
| | Puli | 11,932 (3%) |
| | Dongshan | 17,735 (4%) |
| | Haidai | 19,322 (4%) |
| | Luoshui | 13,330 (3%) |
| | Reshui | 25,247 (6%) |
| | Xize | 13,903 (3%) |
| | Wude | 15,155 (3%) |
| | Longtan | 19,165 (4%) |
| | Delu | 11,779 (3%) |
| Fuyuan | | 207,302 |
| | Zhongan | 39,613 (19%) |
| | Yingshang | 20,009 (10%) |
| | Huangnihe | 16,381 (8%) |
| | Housuo | 18,482 (9%) |
| | Dahe | 23,326 (11%) |
| | Mohong | 18,339 (9%) |
| | Zhuyuan | 13,313 (6%) |
| | Fucun | 24,015 (12%) |
| | Gugansuizu | 4,322 (2%) |
| | Yuwang | 17,755 (9%) |
| | Laochang | 11,747 (6%) |
| | Yangliu | 12,200 (3%) |
| | Shuanghe | 8,514 (2%) |
| | Lefeng | 14,615 (3%) |
| | Wenxing | 16,130 (4%) |
| | Adu | 11,413 (3%) |

Information derived from Fifth National Population Census of China

Table S5. Spearman correlations between cluster variables in the full study population.

| | WB7 | SO2 | PAH2 | PAH7 | PAH25 | NO2 |
|-------|-----|------|------|-------|-------|------|
| WB7 | - | 0.15 | 0.35 | 0.26 | 0.55 | 0.63 |
| SO2 | | - | 0.17 | 0.38 | 0.57 | 0.58 |
| PAH2 | | | - | -0.51 | 0.22 | 0.29 |
| PAH7 | | | | - | 0.58 | 0.47 |
| PAH25 | | | | | - | 0.64 |
| NO2 | | | | | | - |

Clusters were derived from the exposure data of controls.

Table S6. Spearman correlations between individual exposures and cluster variables in the full study population.

| | WB7 | SO2 | PAH2 | PAH7 | PAH25 | NO2 |
|--------------|------|-------|-------|-------|-------|------|
| RET | 0.76 | -0.05 | 0.52 | -0.24 | 0.17 | 0.28 |
| Black Carbon | 0.75 | 0.09 | -0.08 | 0.53 | 0.28 | 0.65 |
| FLT | 0.90 | 0.23 | 0.15 | 0.45 | 0.71 | 0.59 |
| ANT | 0.85 | 0.02 | 0.60 | 0.04 | 0.47 | 0.51 |
| PYR | 0.94 | 0.18 | 0.32 | 0.28 | 0.67 | 0.55 |
| CdP | 0.96 | 0.04 | 0.45 | 0.11 | 0.48 | 0.49 |
| PM2.5 | 0.89 | 0.30 | 0.31 | 0.28 | 0.44 | 0.79 |
| SO2 | 0.15 | 1.00 | 0.17 | 0.38 | 0.57 | 0.58 |
| NkF | 0.47 | 0.09 | 0.95 | -0.47 | 0.29 | 0.29 |
| DIP | 0.21 | 0.20 | 0.93 | -0.49 | 0.11 | 0.27 |
| NaP | 0.25 | 0.18 | -0.64 | 0.92 | 0.31 | 0.35 |
| BbP | 0.31 | 0.31 | -0.57 | 0.91 | 0.37 | 0.46 |
| 6MC | 0.22 | 0.55 | -0.28 | 0.83 | 0.74 | 0.49 |
| DMBA | 0.34 | 0.57 | -0.15 | 0.85 | 0.88 | 0.57 |
| DiP | 0.09 | 0.30 | -0.51 | 0.94 | 0.46 | 0.38 |
| DhP | 0.14 | 0.43 | -0.40 | 0.92 | 0.56 | 0.39 |
| BjA | 0.36 | 0.20 | -0.60 | 0.93 | 0.44 | 0.41 |
| BaC | 0.44 | 0.64 | 0.12 | 0.65 | 0.98 | 0.62 |
| BbF | 0.61 | 0.47 | 0.26 | 0.46 | 0.95 | 0.59 |
| NbF | 0.46 | 0.63 | 0.24 | 0.56 | 0.99 | 0.62 |
| DeP | 0.36 | 0.64 | 0.10 | 0.64 | 0.96 | 0.59 |
| BaF | 0.52 | 0.51 | 0.20 | 0.50 | 0.96 | 0.56 |
| DBA | 0.42 | 0.62 | 0.01 | 0.73 | 0.96 | 0.60 |
| BjF | 0.76 | 0.41 | 0.42 | 0.41 | 0.91 | 0.66 |
| IPY | 0.75 | 0.42 | 0.37 | 0.46 | 0.90 | 0.67 |
| DBT | 0.65 | 0.55 | 0.15 | 0.60 | 0.90 | 0.71 |
| BbT | 0.52 | 0.59 | 0.21 | 0.60 | 0.99 | 0.63 |
| NjF | 0.62 | 0.54 | 0.34 | 0.51 | 0.97 | 0.67 |
| BkF | 0.74 | 0.43 | 0.34 | 0.49 | 0.92 | 0.68 |
| DBF | 0.45 | 0.61 | 0.16 | 0.62 | 0.99 | 0.59 |
| BaP | 0.72 | 0.47 | 0.29 | 0.54 | 0.95 | 0.69 |
| BeP | 0.42 | 0.61 | 0.09 | 0.67 | 0.98 | 0.60 |
| BcF | 0.64 | 0.46 | 0.21 | 0.52 | 0.95 | 0.61 |
| DeIP | 0.37 | 0.61 | 0.19 | 0.56 | 0.97 | 0.55 |
| NeP | 0.40 | 0.62 | 0.10 | 0.65 | 0.97 | 0.59 |
| CHR | 0.58 | 0.55 | 0.29 | 0.50 | 0.99 | 0.62 |
| 1MP | 0.64 | 0.49 | 0.11 | 0.64 | 0.96 | 0.64 |
| TPY | 0.40 | 0.60 | 0.19 | 0.54 | 0.97 | 0.55 |
| COR | 0.49 | 0.58 | 0.26 | 0.55 | 0.98 | 0.59 |
| BgP | 0.50 | 0.59 | 0.17 | 0.65 | 0.98 | 0.63 |
| BaA | 0.64 | 0.51 | 0.31 | 0.51 | 0.98 | 0.64 |
| 5MC | 0.43 | 0.47 | 0.37 | 0.35 | 0.91 | 0.53 |
| NO2 | 0.63 | 0.58 | 0.29 | 0.47 | 0.64 | 1.00 |

Correlations with an absolute value over 0.75 are highlighted.

Clusters were derived from the exposure data of controls.

Table S7. Spearman correlations between cluster variables in subjects using smoky coal almost exclusively

| | retene | WB2 | PAH5+ | PAH2 | PAH29+ |
|--------|--------|------|-------|-------|--------|
| retene | - | 0.69 | -0.24 | 0.67 | 0.54 |
| WB2 | | - | 0.16 | 0.69 | 0.79 |
| PAH5 | | | - | -0.44 | 0.42 |
| PAH2 | | | | - | 0.38 |
| PAH29 | | | | | - |

Clusters were derived from the exposure data of controls.

Table S8. Spearman correlations between individual exposures and cluster variables in subjects using smoky coal almost exclusively.

| | retene | WB2 | PAH5+ | PAH2 | PAH29+ |
|--------------|--------|------|-------|-------|--------|
| RET | 1.00 | 0.69 | -0.24 | 0.67 | 0.54 |
| CdP | 0.69 | 0.97 | 0.24 | 0.61 | 0.86 |
| ANT | 0.64 | 0.97 | 0.06 | 0.76 | 0.67 |
| NaP | -0.39 | 0.05 | 0.96 | -0.53 | 0.33 |
| BbP | -0.19 | 0.04 | 0.96 | -0.49 | 0.39 |
| DiP | -0.30 | 0.14 | 0.98 | -0.43 | 0.34 |
| DhP | -0.21 | 0.23 | 0.91 | -0.33 | 0.46 |
| Black Carbon | 0.09 | 0.38 | 0.83 | -0.10 | 0.46 |
| BjA | -0.24 | 0.09 | 0.95 | -0.51 | 0.43 |
| DIP | 0.42 | 0.52 | -0.39 | 0.90 | 0.23 |
| NkF | 0.75 | 0.70 | -0.45 | 0.97 | 0.40 |
| BaC | 0.49 | 0.70 | 0.45 | 0.30 | 0.99 |
| FLT | 0.40 | 0.69 | 0.53 | 0.21 | 0.96 |
| BbT | 0.53 | 0.79 | 0.39 | 0.40 | 0.99 |
| NbF | 0.60 | 0.79 | 0.33 | 0.43 | 0.99 |
| DMBA | 0.21 | 0.57 | 0.75 | 0.01 | 0.87 |
| DBA | 0.41 | 0.64 | 0.56 | 0.19 | 0.97 |
| PYR | 0.55 | 0.79 | 0.35 | 0.40 | 0.97 |
| IPY | 0.56 | 0.88 | 0.35 | 0.48 | 0.94 |
| BbF | 0.61 | 0.72 | 0.26 | 0.41 | 0.96 |
| BkF | 0.54 | 0.88 | 0.38 | 0.46 | 0.95 |
| BaF | 0.60 | 0.68 | 0.27 | 0.37 | 0.95 |
| DeP | 0.52 | 0.69 | 0.43 | 0.30 | 0.98 |
| SO2 | 0.37 | 0.54 | 0.43 | 0.26 | 0.81 |
| BjF | 0.62 | 0.90 | 0.28 | 0.54 | 0.95 |
| DBF | 0.52 | 0.76 | 0.40 | 0.36 | 0.99 |
| BaP | 0.53 | 0.86 | 0.41 | 0.42 | 0.98 |
| NjF | 0.62 | 0.87 | 0.32 | 0.50 | 0.96 |
| CHR | 0.62 | 0.79 | 0.28 | 0.46 | 0.98 |
| 6MC | 0.14 | 0.27 | 0.70 | -0.17 | 0.69 |
| PM2.5 | 0.42 | 0.68 | 0.52 | 0.34 | 0.81 |
| DBT | 0.36 | 0.72 | 0.52 | 0.24 | 0.95 |
| BeP | 0.48 | 0.72 | 0.48 | 0.28 | 0.99 |
| TPY | 0.59 | 0.72 | 0.28 | 0.40 | 0.97 |
| BcF | 0.55 | 0.71 | 0.34 | 0.35 | 0.96 |
| 1MP | 0.43 | 0.70 | 0.50 | 0.25 | 0.97 |
| NeP | 0.53 | 0.73 | 0.44 | 0.32 | 0.99 |
| BaA | 0.60 | 0.84 | 0.31 | 0.47 | 0.98 |
| NO2 | 0.41 | 0.67 | 0.52 | 0.36 | 0.76 |
| DeIP | 0.57 | 0.76 | 0.31 | 0.41 | 0.98 |
| COR | 0.56 | 0.84 | 0.31 | 0.46 | 0.97 |
| BgP | 0.51 | 0.81 | 0.46 | 0.36 | 0.98 |
| 5MC | 0.81 | 0.81 | 0.05 | 0.61 | 0.87 |

Correlations with an absolute value over 0.75 are highlighted.

Clusters were derived from the exposure data of controls.

Table S9. Odds ratios (95% CI) for individual clusters in the whole population

| Cluster | OR (95% CI) |
|----------------|--------------------|
| WB7 | 0.24 (0.19, 0.31) |
| PAH2 | 1.14 (0.94, 1.5) |
| NO2 | 2.27 (1.48, 3.55) |
| PAH7 | 1.25 (0.95, 1.88) |
| PAH25 | 2.21 (1.67, 2.87) |
| SO2 | 0.55 (0.43, 0.7) |

Models were adjusted for age, self-reported exposure to second-hand smoke, and self-reported food-sufficiency before marriage

Table S10. Odds ratios (95% CI) for individual clusters in the smoky coal subpopulation

| Cluster | OR (95% CI) |
|----------------|--------------------|
| retene | 0.97 (0.8, 1.06) |
| WB2 | 1.04 (0.92, 1.33) |
| PAH5+ | 1.04 (0.93, 1.43) |
| PAH2 | 1.11 (0.98, 1.51) |
| PAH29+ | 1.13 (0.97, 1.54) |

Models were adjusted for age, self-reported exposure to second-hand smoke, and self-reported food-sufficiency before marriage

Table S11. Odds ratios (95%CI) for individual constituents within identified clusters of the smoky coal subpopulation

| Cluster | Exposure | OR (95% CI) |
|---------|--------------|-----------------------|
| retene | RET | 0 (0.00, 0.01) |
| WB2 | CdP | 0.42 (0, 33.76) |
| WB2 | ANT | 0.94 (0.14, 5.48) |
| PAH5+ | NaP | 0.18 (0.02, 1.05) |
| PAH5+ | BbP | 0.95 (0.29, 2.73) |
| PAH5+ | DiP | 1.72 (0.54, 12.32) |
| PAH5+ | DhP | 1.81 (0.8, 6.04) |
| PAH5+ | Black Carbon | 2.46 (0.82, 11.21) |
| PAH5+ | BjA | 2.83 (0.69, 36.28) |
| PAH2 | DIP | 0.99 (0.66, 1.52) |
| PAH2 | NkF | 1.09 (0.53, 2.75) |
| PAH29+ | BaC | 0.18 (0, 2.7) |
| PAH29+ | SO2 | 0.25 (0.02, 1.45) |
| PAH29+ | FLT | 0.25 (0, 5.08) |
| PAH29+ | PYR | 0.33 (0, 7.46) |
| PAH29+ | DMBA | 0.46 (0.02, 2.96) |
| PAH29+ | BbT | 0.48 (0.01, 4.15) |
| PAH29+ | NbF | 0.48 (0, 6.26) |
| PAH29+ | DBA | 0.49 (0.01, 4.67) |
| PAH29+ | IPY | 0.51 (0.01, 3.96) |
| PAH29+ | BbF | 0.55 (0.02, 4.05) |
| PAH29+ | BkF | 0.58 (0.01, 5.93) |
| PAH29+ | BaF | 0.67 (0.03, 5.17) |
| PAH29+ | DeP | 0.68 (0.01, 8.56) |
| PAH29+ | BjF | 0.7 (0.02, 8.39) |
| PAH29+ | DBF | 0.78 (0.03, 12.01) |
| PAH29+ | BaP | 0.87 (0.03, 14.72) |
| PAH29+ | NjF | 0.9 (0.06, 9.77) |
| PAH29+ | CHR | 0.91 (0.05, 11.66) |
| PAH29+ | 6MC | 1.15 (0.3, 6.01) |
| PAH29+ | DBT | 1.2 (0.27, 8.73) |
| PAH29+ | PM2.5 | 1.35 (0.3, 7.96) |
| PAH29+ | BeP | 1.54 (0.11, 84.29) |
| PAH29+ | TPY | 1.58 (0.16, 48.81) |
| PAH29+ | BcF | 1.61 (0.2, 60.23) |
| PAH29+ | 1MP | 1.62 (0.21, 47.26) |
| PAH29+ | NeP | 1.81 (0.13, 190.3) |
| PAH29+ | BaA | 1.91 (0.18, 108.67) |
| PAH29+ | NO2 | 2.21 (0.96, 5.71) |
| PAH29+ | DeIP | 2.4 (0.27, 197.88) |
| PAH29+ | COR | 4.64 (0.39, 814.39) |
| PAH29+ | BgP | 9.74 (0.34, 48789.29) |
| PAH29+ | 5MC | 28.46 (1.02, 772.39) |

Models were adjusted for age, self-reported exposure to second-hand smoke, and self-reported food-sufficiency before marriage

Table S12. Comparing deposit specific odds ratios (95% CI) for models with and without household air pollution (HAP) components

| Coal Deposit Layer | Coal deposit only [§] | Coal deposit + HAP [§] |
|------------------------|--------------------------------|---------------------------------|
| Never smoky coal users | 0.11 (0.05, 0.27) | 0.84 (0.38, 1.30) |
| 9 | 2.80 (1.17, 7.83) | 2.77 (1.12, 6.85) |
| 24 | 1.55 (0.81, 5.00) | 1.25 (0.72, 4.12) |
| 1,2,4,7,8 | 1.00 (0.60, 2.00) | 1.03 (0.70, 1.61) |
| 10 | 0.80 (0.41, 1.46) | 0.89 (0.45, 1.32) |
| 12,13,14,38 | 1.32 (0.76, 3.52) | 0.87 (0.36, 1.41) |
| 16,17,19 | 1.29 (0.82, 2.73) | 0.91 (0.53, 1.30) |
| 20,27 | 0.75 (0.24, 1.66) | 1.18 (0.68, 3.41) |
| Wood/Feces Users | 0.20 (0.10, 0.43) | 1.38 (0.87, 3.21) |

[§] Effect estimates for a categorical covariate in a Bayesian horseshoe model do not require reference coding, so these odds ratios are relative to an exposure-only model.

Deposit groupings were derived by combining coals of the same type and number of women using coal from each deposit. Deposits with sufficient users to allow separate analysis (9, 10, 24) were kept as such

Models were adjusted for age, self-reported exposure to second-hand smoke, and self-reported food-sufficiency before marriage

Table S13. Comparing deposit specific odds ratios (95% CI) for models with and without household air pollution (HAP) components among the smoky coal subpopulation

| Coal Deposit Layer | Coal deposit only [§] | Coal deposit + HAP [§] |
|--------------------|--------------------------------|---------------------------------|
| 9 | 2.55 (1.00, 7.45) | 1.58 (0.91, 4.11) |
| 24 | 1.15 (0.47, 4.24) | 1.17 (0.72, 3.05) |
| 1,2,4,7,8 | 1.06 (0.56, 2.16) | 0.74 (0.41, 1.12) |
| 10 | 0.80 (0.28, 1.52) | 0.79 (0.37, 1.21) |
| 12,13,14,38 | 1.16 (0.50, 3.90) | 1.23 (0.76, 3.26) |
| 16,17,19 | 0.93 (0.40, 1.73) | 1.07 (0.69, 1.82) |

[§] Effect estimates for a categorical covariate in a Bayesian horseshoe model do not require reference coding, so these odds ratios are relative to an exposure-only model.

Deposit groupings were derived by combining coals of the same type and number of women using coal from each deposit. Deposits with sufficient users to allow separate analysis (9, 24) were kept as such

Models were adjusted for age, self-reported exposure to second-hand smoke, and self-reported food-sufficiency before marriage