Supplementary Online Content

Yu K, Qiu G, Chan K-H, et al. Association of solid fuel use with risk of cardiovascular and all-cause mortality in rural China. *JAMA*. doi:10.1001/jama.2018.2151

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. Supplementary Methods

Assessment of Solid Fuel Use

At baseline, all participants were asked about the duration (in years) they lived in the present/previous/before previous residence. A small proportion (n=2,132) of participants reported a greater total number of years living in the three residences than their baseline age and were excluded from all analyses.

For cooking, participants were asked “In the present/previous/before previous house, how often did you cook at home?” (Single choice allowed; daily, weekly, monthly, never/rarely, and no cooking facility). The participants who selected the first three categories were further asked “What is the main cooking fuel used?” (Single choice allowed; gas, electricity, coal, wood/charcoal, and other unspecified fuels). In this study, those who cooked monthly or less were grouped as “not cooking regularly”, and participants who reported cooking weekly or daily were classified according to the primary fuel type reported. The participants were also asked “In your present house/ previous/before previous house, do your cook-stove(s) all have a chimney/extractor?” (Single choice allowed; yes, not all stoves, no) following the questions on cooking fuels. The participants who selected the first two categories were defined as having at least some ventilated cookstoves.

For heating, participants were asked “In winter, do you normally heat your present/previous/before previous house?” (Yes or no). The participants who answered yes were further asked “If yes, what was the main heating fuel used in your house?” (Single choice allowed; gas, central heating, electricity, coal, wood/charcoal, and other unspecified fuels). Central heating refers to heating by piped water or steam transported from a remote boiler to individual households. Participants from Zhejiang (n=56,813) were excluded from heating-related analyses, as heating was rarely reported (n=341, 0.6%) among them.

In each residence, if the participant used more than one type of fuel for cooking or heating, the fuel used most frequently was recorded as the primary fuel. Gas, electricity, and central heating were considered as clean fuels while coal and wood were considered as solid fuels. Participants who reported using other unspecified fuel were excluded from all analyses (n=2,401 and n=1,793, respectively, excluded from cooking- and heating-related analyses).

Duration of Solid Fuel Use

The duration of solid fuel use for cooking and heating was calculated based on the duration the participant reported living in each residence. It was assumed that the primary fuel used had not changed during the time period lived in each residence. Based on this assumption, if a participant reported using solid fuels in his or her present residence, the duration of solid fuel use for cooking was calculated by aggregating the time (in years) lived in consecutive residences during which this fuel was used as the primary cooking fuel. The duration of solid fuel use for heating was calculated the same way.

For example, assume that a participant reported having lived in his three most recent residences for 20, 15 and 10 years, respectively. The corresponding cooking frequencies were weekly, daily and weekly when living in these three residences, the cooking fuels were coal, wood and coal, and the heating fuels were wood, coal and coal. For cooking, the duration of solid fuel use equals to 20 + 15 + 10 = 45 years. For heating, the duration of solid fuel use equals to 20 + 15 + 10 = 45 years.

Weighted Duration of Solid Fuel Use

The duration of solid fuel use for cooking was weighted on cooking frequency, following the steps outlined below:

1. A weight coefficient was assigned to each cooking frequency group, namely 0.5 to 'weekly' and 1.0 to 'daily'.
2. The weighted duration of solid fuel use for cooking was derived using the self-reported total number of years lived in each residence multiplied by the assigned weight coefficient. For example, assume that a participant reported having lived in his three most recent residences for 20, 15 and 10 years, respectively. The corresponding cooking frequencies were weekly, daily and weekly when living in these three residences, and the corresponding cooking fuels were coal, wood and coal. In this case, the duration of solid fuel use would equal to 0.5*20 + 1.0*15 + 0.5*10 =30 years.

The duration of solid fuel use for heating was weighted according to the average number of months per year during which the average temperature in each study area was under 8 degrees Celsius, following the steps outlined below:

1. A weight coefficient was assigned to each study area, by aggregating the number of months during which the average temperature was under 8 degrees Celsius from 1999-2013, then dividing the aggregate number by 180 months (the total number of months during these 15 years). The calculated weight coefficients were 0.19 for Sichuan, 0.42 for Gansu, 0.27 for Henan, and 0.18 for Hunan. In this calculation, the monthly mean temperatures of the provincial capital in each study area was used during 1999-2013, which was available from the "Statistical yearbook of China" published by the State Statistical bureau (http://www.stats.gov.cn/tjsj/ndsj/).
2. The weighted duration of solid fuel use for heating was derived using the self-reported total number of years in each residence multiplied by the assigned weight coefficient. Assume that a participant from Gansu reported having lived in his three most recent residences for 20, 15 and 10 years, respectively. The corresponding heating fuels were wood, coal and coal. The duration of solid fuel use would equal 0.41*20 + 0.41*15 + 0.41*10 = 18.5 years.

Detailed Statistical Analysis

After excluding the participants who had self-reported physician-diagnosed coronary heart disease, stroke or transient ischemic attack (n=8,578), those who reported total residence time living in the three most recent residences longer than baseline age (n=2,132) and those who were aged <35 years (n=5,891) at baseline, 271,217 remained (1,113 participants met multiple exclusion criteria). In the analyses of fuel use for cooking, 91,265 participants not cooking regularly and 2,401 participants who reported unspecified fuel use were further excluded. In the analyses of fuel use for heating, 50,522 non-heating participants, 1,793 participants who reported unspecified fuel use, and participants from Zhejiang (n=56,813) where heating was rarely reported (0.6%) were further excluded. As a result, the cooking-related analyses involved 177,551 participants, and the heating-related analyses involved 162,089 participants. Weighted Kappa was calculated to evaluate the reproducibility of cooking and heating exposure categories. Baseline characteristics were described as means and standard deviations or percentages by categories of cooking and heating exposure, with adjustment for age, sex and study areas as appropriate, via direct standardization.

Adjusted hazard ratios (HRs) for the first four years and for the subsequent years of follow-up

were compared, and no evidence of departure from the proportional hazards assumption was observed for cardiovascular and all-cause mortality. Mortality rate per 100,000 person-years were adjusted for age, sex and study areas, and the absolute rate difference (ARD) was calculated as the adjusted mortality rate of the exposed group minus that of the reference group. Cox proportional hazards regression models were used to calculate HRs and 95% confidence intervals (CIs) for cardiovascular (overall and subtypes) and all-cause mortality in association with baseline solid fuel use as compared with clean fuel use for cooking and for heating separately. Cox models were stratified by age-at-risk (in 5-year intervals), sex, and study areas (5 groups for cooking-related analyses; 4 groups for heating-related analyses), and further adjusted for established mortality risk factors including smoking status, passive smoking, alcohol consumption, body-mass index (continuous), and physical activity level (metabolic equivalent of tasks hours/day [MET-hour/day]; continuous), variables related to socio-economic status including education, household income, and diet (consumption of preserved vegetables, fresh fruit, meat, fish, dairy, rice, poultry, and eggs), and cookstove ventilation (yes and no). In addition, cooking and heating exposures were mutually adjusted for in the final models. All analyses were performed using reported fuel use at baseline.

Sensitivity analyses were conducted by 1) further adjusting for potential confounders including survey season, occupation, self-reported health status, family history of cardiovascular disease; 2) excluding participants on anti-hypertensive treatments at baseline (n=8,972); 3) excluding participants who used clean fuels for <10 years (Cooking: n=11,651; heating: n=6,403); and 4) excluding events that occurred within the first two years of follow-up, respectively. Additional analysis on the weighted duration of continuous exposure to solid fuel use for cooking and heating was conducted, as discussed in the previous session. Two-sided *P* values were used and $P < .05$ denotes statistical significance. All analyses were performed using SAS 9.3 (SAS Institute Inc, Cary, NC), and graphs were plotted using R 3.4.2.

eTable 1. Baseline Fuel Types According to Study Areas.

Areas	N	Cooking (%)				Heating (%) [a]			
		Clean Fuels [b]	Solid Fuels [c]	Not Cooking Regularly [d]	Other unspecified Fuels	Clean Fuels[b]	Solid Fuels[c]	No Heating	Other unspecified Fuels
All Rural	271,217	26,559 (9.8)	150,992 (55.7)	91,265 (33.7)	2,401 (0.9)	14,817 (6.9)	147,272 (68.7)	50,522 (23.6)	1,793 (0.8)
Hunan	56,374	1,214 (2.2)	36,000 (63.9)	18,876 (33.5)	284 (0.5)	3,457 (6.1)	52,420 (93.0)	483 (0.9)	14 (0.0)
Henan	58,860	800 (1.4)	36,231 (61.6)	20,020 (34.0)	1,809 (3.1)	742 (1.3)	48,295 (82.1)	9,763 (16.6)	60 (0.1)
Gansu	45,381	642 (1.4)	27,616 (60.9)	17,056 (37.6)	67 (0.1)	435 (1.0)	43,138 (95.1)	229 (0.5)	1,579 (3.5)
Sichuan	53,789	7,555 (14.0)	30,792 (57.2)	15,241 (28.3)	201 (0.4)	10,183 (18.9)	3,419 (6.4)	40,047 (74.5)	140 (0.3)
Zhejiang	56,813	16,348 (28.8)	20,353 (35.8)	20,072 (35.3)	40 (0.1)	314 (0.6)	4 (0.0)	56,495 (99.4)	0 (0.0)
All Urban [e]	207,764	143,990 (69.3)	14,806 (7.1)	48,902 (23.5)	66 (0.0)	78,065 (37.6)	26,663 (12.8)	102,946 (49.5)	90 (0.0)

[a] Participants from Zhejiang were excluded from the calculation of overall distribution of heating exposure categories (N=56,813).

[b] Clean fuels refer to electricity, gas, or central heating (for heating only).

[c] Solid fuels refer to coal, wood, or charcoal.

[d] Not cooking regularly refers to cooking monthly or less.

[e] Solid fuel use was reported with a small prevalence (7.1% for cooking and 12.8% for heating) among urban participants, and statistical power would be limited if urban participants were included in analyses of the present study, which were stratified by study areas. As a result, this study focused only on rural urban participants in the present study.

eTable 2. Baseline Characteristics of Participants at Baseline Survey and Resurvey.

Characteristics	Baseline Survey	Resurvey
N	271,217	10,892
Age, means (SD), year	51.0 (10.2)	50.9 (10.2)
Women, %	58.6	59.8
Education, %		
No formal education	24.0	24.6
Primary school	41.8	41.2
Middle school or higher	34.3	34.0
Household income, yuan/year, %		
<10,000	39.2	43.4
10,000-19,999	28.8	30.4
≥20,000	32.0	26.3
Regular smoking, %		
Never	65.7	67.0
Previous	2.5	2.3
Current	31.8	30.7
Regular alcohol intake, %		
Never	82.1	82.2
Previous	4.2	4.0
Current	13.7	13.9
Physical activity, means (SD), MET-hour/d [a]	23.4 (14.4)	22.8 (15.0)
Body mass index, means (SD), kg/m^2	23.2 (3.3)	23.2 (3.3)

[a] MET-h/d denotes Metabolic Equivalent of Tasks - hours per day

eTable 3. Reproducibility of Reported Fuel Use for Cooking and for Heating Between Baseline and Resurvey (N = 10,892). [a]

Baseline	**Resurvey**			
	Clean fuels	**Solid fuels**	**No cooking/heating**	**Total**
Cooking [b]				
Clean fuels	**627 (63.2%)**	240	125	**992 (9.3%)**
Solid fuels	637	**4,835 (79.1%)**	641	**6,113 (57.3%)**
Not cooking regularly	238	433	**2,886 (81.1%)**	**3,557 (33.4%)**
Total	**1,502 (14.1%)**	**5,508 (51.7%)**	**3,652 (34.3%)**	**10,662 (100%)**
Weighted Kappa: 0.62 (95% CI: 0.61-0.64).				
Heating [c]				
Clean fuels	**159 (44.0%)**	65	137	**361 (4.3%)**
Solid fuels	209	**5,388 (88.6%)**	487	**6,084 (72.0%)**
No heating	179	264	**1,565 (77.9%)**	**2,008 (23.8%)**
Total	**547 (6.5%)**	**5,717 (67.6%)**	**2,189 (25.9%)**	**8,453 (100%)**
Weighted Kappa: 0.64 (95% CI: 0.62-0.66).				

[a] Overall, 8,348 participants (78.3%) reported the same category of exposure from cooking in the resurvey as at baseline; 7,112 participants (84.1%) reported the same category of exposure from heating in the resurvey as at baseline.

Clean fuels refer to gas and electricity, along with central heating in analyses of heating fuels; Solid fuels refer to coal, wood, or charcoal; Not cooking regularly refers to cooking for monthly or less; no heating refers to not using heating in winter.

[b] Participants who reported using other unspecified fuel were excluded from cooking related analyses.

[c] Participants who reported using other unspecified fuel and those from Zhejiang were excluded from heating related analyses.

eTable 4. Adjusted Hazard Ratios for Cardiovascular and All-Cause Mortality According to Baseline Fuel Use for Cooking After Additional Adjustments or Further Exclusions.

Mortality	No. of deaths (%)	Main model [a]	With extended adjustments [b]	Excluding participants on anti-hypertensive treatment	Excluding participants using clean fuels for <10 years	Excluding events which occurred within the first two years of follow-up
Cardiovascular mortality						
Total cardiovascular death						
Clean fuel	180 (0.7)	Reference	Reference	Reference	Reference	Reference
Solid fuel	2,957 (2.0)	1.20 (1.02-1.41)	1.21 (1.16-1.27)	1.19 (1.14-1.25)	1.27 (1.22-1.33)	1.19 (1.13-1.24)
Ischemic heart disease death						
Clean fuel	42 (0.2)	Reference	Reference	Reference	Reference	Reference
Solid fuel	899 (0.6)	1.12 (0.81-1.56)	1.13 (1.04-1.22)	1.05 (0.96-1.15)	1.28 (1.19-1.39)	1.12 (1.02-1.22)
Stroke death						
Clean fuel	107 (0.4)	Reference	Reference	Reference	Reference	Reference
Solid fuel	1,612 (1.0)	1.16 (0.94-1.44)	1.18 (1.11-1.25)	1.19 (1.12-1.27)	1.21 (1.14-1.28)	1.12 (1.04-1.19)
Other cardiovascular death						
Clean fuel	31 (0.2)	Reference	Reference	Reference	Reference	Reference
Solid fuel	446 (0.3)	1.42 (0.98-2.13)	1.49 (1.00-2.21)	1.39 (0.93-2.07)	1.50 (0.92-2.42)	1.50 (0.98-2.30)
All-cause mortality						
Clean fuel	855 (3.2)	Reference	Reference	Reference	Reference	Reference
Solid fuel	7,955 (5.3)	1.11 (1.03-1.20)	1.12 (1.09-1.15)	1.11 (1.08-1.14)	1.13 (1.10-1.16)	1.09 (1.05-1.12)

[a] Hazard ratios were stratified according to age-at-risk, sex, and study areas, and adjusted for education level, income, alcohol consumption, smoking status, passive smoking, physical activity, body-mass index, diet (consumption of fresh fruit, preserved vegetables, meat, fish, dairy, rice, poultry, eggs), cookstove ventilation, and heating exposure.

[b] Additionally adjusted for survey season, occupation, self-rated health status, and family history of cardiovascular disease.

eTable 5. Adjusted Hazard Ratios for Cardiovascular and All-Cause Mortality According to Baseline Fuel Use for Heating After Additional Adjustments or Further Exclusions.

Mortality	No. of deaths (%)	Main model [a]	With extended adjustment [b]	Excluding participants on anti-hypertensive treatment	Excluding participants using clean fuels for <10 years	Excluding events which occurred within the first two years of follow-up
Cardiovascular mortality						
Total cardiovascular death						
Clean fuel	148 (1.0)	Reference	Reference	Reference	Reference	Reference
Solid fuel	3,809 (2.6)	1.29 (1.06-1.55)	1.33 (1.22-1.46)	1.25 (1.13-1.38)	1.26 (1.14-1.39)	1.20 (1.08-1.32)
Ischemic heart disease death						
Clean fuel	45 (0.3)	Reference	Reference	Reference	Reference	Reference
Solid fuel	1,280 (0.9)	1.13 (0.82-1.59)	1.19 (1.03-1.37)	1.17 (0.99-1.38)	1.04 (0.88-1.24)	1.12 (0.95-1.32)
Stroke death						
Clean fuel	78 (0.4)	Reference	Reference	Reference	Reference	Reference
Solid fuel	2,049 (1.4)	1.45 (1.12-1.88)	1.50 (1.31-1.71)	1.39 (1.21-1.61)	1.45 (1.25-1.68)	1.32 (1.13-1.53)
Other cardiovascular death						
Clean fuel	25 (0.2)	Reference	Reference	Reference	Reference	Reference
Solid fuel	480 (0.4)	1.00 (0.63-1.59)	1.04 (0.66-1.65)	0.90 (0.57-1.45)	1.03 (0.59-1.79)	0.93 (0.57-1.53)
All-cause mortality						
Clean fuel	628 (4.2)	Reference	Reference	Reference	Reference	Reference
Solid fuel	8,780 (6.0)	1.14 (1.03-1.26)	1.16 (1.10-1.23)	1.13 (1.06-1.20)	1.07 (1.00-1.14)	1.14 (1.07-1.22)

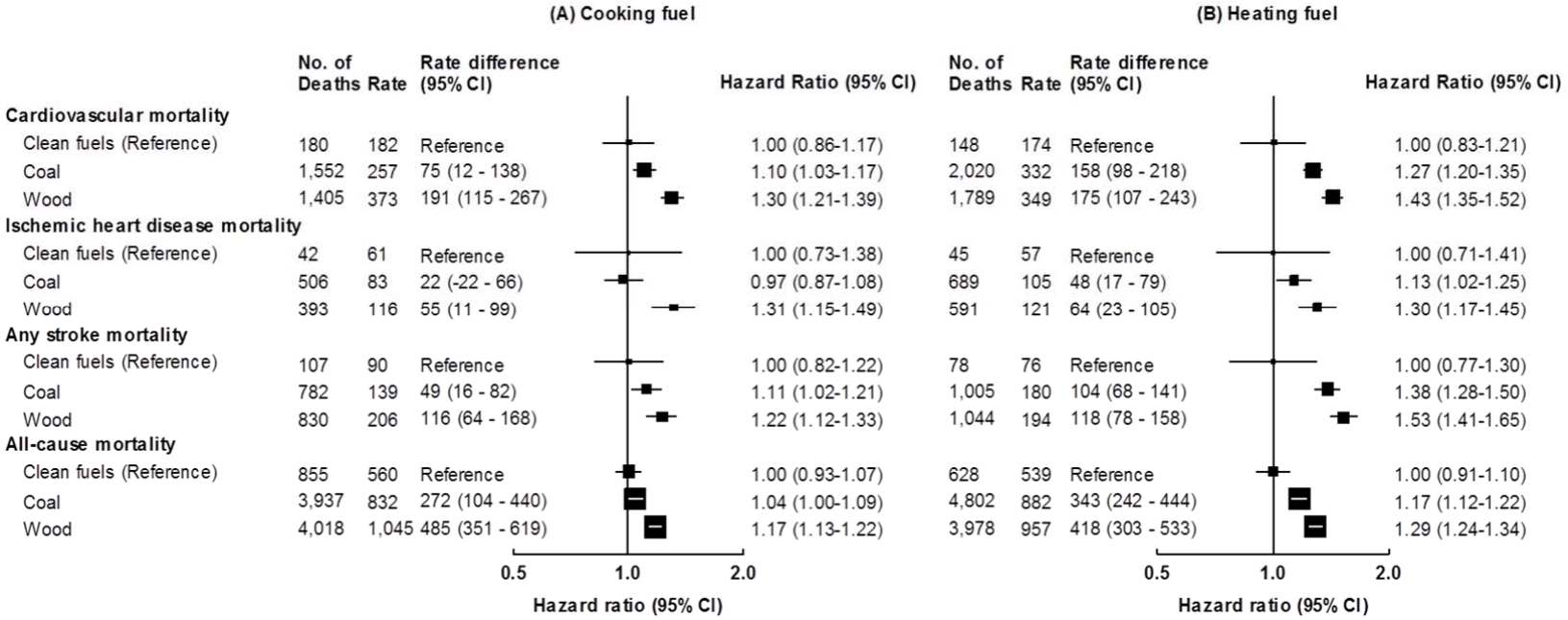
[a] Hazard ratios were stratified according to age-at-risk, sex, and study areas, and adjusted for education level, income, alcohol consumption, smoking status, passive smoking, physical activity, body-mass index, diet (consumption of fresh fruit, preserved vegetables, meat, fish, dairy, rice, poultry, eggs), cookstove ventilation, and cooking exposure.

[b] Additionally adjusted for survey season, occupation, self-rated health status, and family history of cardiovascular disease.

eFigure 1. Locations of the 5 Rural Study Areas of the China Kadoorie Biobank Study in Mainland China.

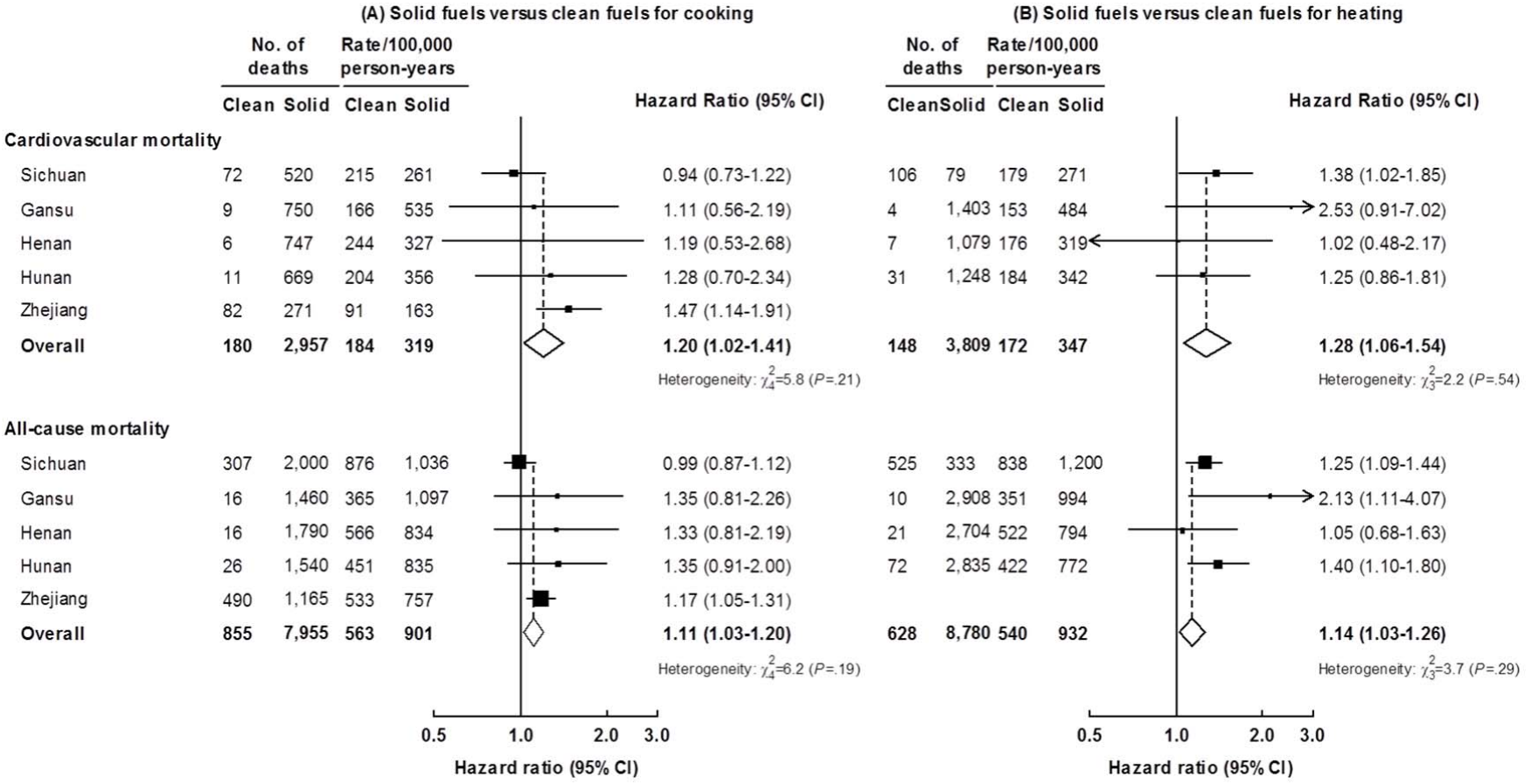
Red circles indicate the five rural areas and number of participants recruited at baseline in each site is shown in brackets. The time window of the baseline survey was presented under the name of each study area. (Adopted from Chen et al 2011)[1] Solid fuel use was reported with a small prevalence (7.1% for cooking and 12.8% for heating) among urban participants, and statistical power would be limited if urban participants were included in analyses of this study, which were stratified by study areas. As a result, this study focused only on rural participants.

eFigure 2. Adjusted Hazard Ratios for Cardiovascular and All-Cause Mortality by Baseline Cooking and Heating Fuel Types.



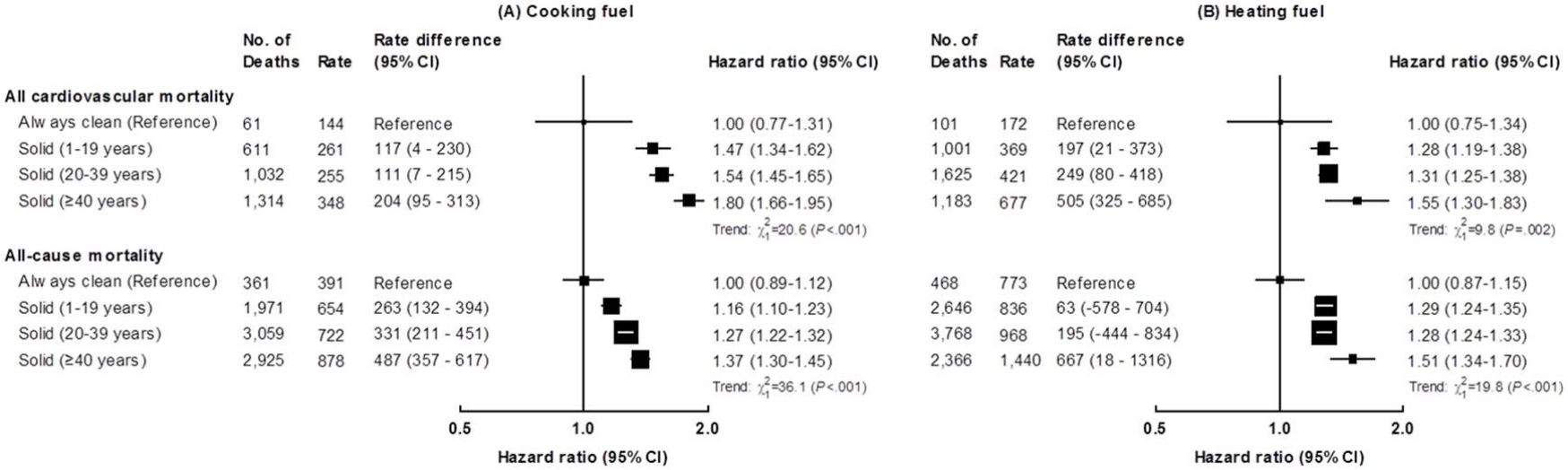
Clean fuels refer to electricity, gas, or central heating (for heating only); "Wood" refers to wood or charcoal. Number of participants (and person-years) for clean fuels, coal and wood users included in the cooking-related analyses were, respectively: 26,559 (195,392), 89,066 (652,022), 61,926 (445,963); in the heating-related analyses were, respectively: 14,817 (103,860), 84,942 (612,471), and 62,330 (447,688). Mortality rate per 100,000 person-years were adjusted for age, sex and study areas, and the rate difference was calculated as the adjusted mortality rate of the exposed group minus that of the reference group. Hazard ratios were stratified according to age-at-risk, sex, and study areas, and adjusted for education level, income, alcohol consumption, smoking status, passive smoking, physical activity, body-mass index, diet (consumption of fresh fruit, preserved vegetables, meat, fish, dairy, rice, poultry, eggs), cookstove ventilation, and cooking and heating exposures. The floating-absolute-risk method provides the variance of the logarithm of the hazard ratio for each category (including the reference category) to facilitate comparisons across the different exposure categories.[2] The boxes represent hazard ratios, with the size inversely proportional to the variance of the logarithm of the hazard ratio, and the horizontal lines represent 95% confidence intervals.

eFigure 3. Adjusted Hazard Ratios for Cardiovascular and All-Cause Mortality by Baseline Cooking and Heating Fuel Types, Stratified by Study Areas.



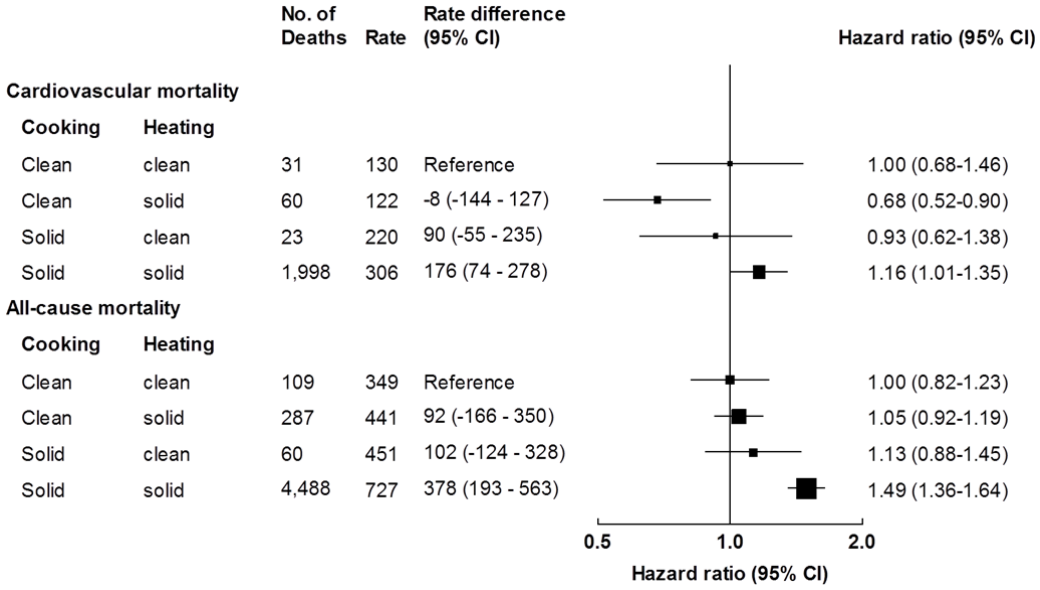
“Clean” refers to clean fuels (i.e., electricity, gas, or central heating [for heating only]); “Solid” refers to solid fuels (i.e., coal, wood, or charcoal). Mortality rate per 100,000 person-years was adjusted for age, sex, and study areas. Hazard ratios were stratified according to age-at-risk and sex, and adjusted for education level, income, alcohol consumption, smoking status, passive smoking, physical activity, body-mass index, diet (consumption of fresh fruit, preserved vegetables, meat, fish, dairy, rice, poultry, eggs), cookstove ventilation, and cooking and heating exposures. The numbers of participants (and person-years) included for clean and solid fuels users in the five areas (from Sichuan to Zhejiang) for cooking-related analyses were, respectively: 7,555 (52,857) and 30,792 (222,249), 642 (4,948) and 27,616 (200,565), 800 (5,357) and 36,231 (263,375), 16,348 (123,655) and 20,353 (149,965), and 1,214 (8,575) and 36,000 (261,831); and those in the four areas (from Sichuan to Hunan) for heating-related analyses were, respectively: 10,183 (70,141) and 3,419 (23,851), 435 (3,361) and 43,138 (307,346), 742 (4,913) and 48,295 (351,102), and 3,475 (25,445) and 52,420 (377,860). The floating-absolute-risk method provides the variance of the logarithm of the hazard ratio for each category (including the reference category) to facilitate comparisons across the different exposure categories.[2] The boxes represent hazard ratios, with the size inversely proportional to the variance of the logarithm of the hazard ratio.

eFigure 4. Adjusted Hazard Ratios for Cardiovascular and All-Cause Mortality According to Weighted Long-term Assessments of Fuel Use for Cooking and for Heating.



"Clean" refers to clean fuels (i.e., electricity, gas, or central heating [for heating only]); "Solid" refers to solid fuels (i.e., coal, wood, or charcoal). The duration of solid fuel use was weighted based on self-reported cooking frequency (daily: 1.0; weekly:0.5) for cooking-related analyses and the mean proportion of months with average temperature under 8 degrees Celsius during 1999-2013 in each study area (Sichuan: 0.19; Gansu: 0.42; Henan: 0.27; Hunan: 0.18). Mortality rate per 100,000 person-years were adjusted for age, sex and study areas, and the rate difference was calculated as the adjusted mortality rate of the exposed group minus that of the reference group. Hazard ratios were stratified according to age-at-risk, sex, and study areas, and adjusted for education level, income, alcohol consumption, smoking status, passive smoking, physical activity, body-mass index, diet (consumption of fresh fruit, preserved vegetables, meat, fish, dairy, rice, poultry, eggs), cook-stove ventilation, and cooking and heating exposures. The numbers of participants (and person-years) for the four exposure categories (from always clean to solid [≥40 years]) included for the cooking-related analyses were, respectively: 11,084 (82,206), 46,606 (344,019), 77,931 (567,653), and 26,455 (186,313); for the heating-related analyses were, respectively: 9,060 (61,981), 65,216 (474,459), 61,401 (441,828), and 20,655 (143,872). The floating-absolute-risk method provides the variance of the logarithm of the hazard ratio for each category (including the reference category) to facilitate comparisons across the different exposure categories.[2] The boxes represent hazard ratios, with the size inversely proportional to the variance of the logarithm of the hazard ratio.

eFigure 5. Association of Combined Solid Fuel Use for Cooking and for Heating With Cardiovascular and All-Cause Mortality.



"Clean" refers to clean fuels (i.e. electricity, gas, or central heating [for heating only]); "Solid" refers to solid fuels (i.e. coal, wood, or charcoal). Mortality rate per 100,000 person-years were adjusted for age, sex and study areas, and the rate difference was calculated as the adjusted mortality rate of the exposed group minus that of the reference group. Hazard ratios were stratified according to age-at-risk, sex, and study areas and adjusted for education level, income, alcohol consumption, smoking status, passive smoking, physical activity, body-mass index, diet (consumption of fresh fruit, preserved vegetables, meat, fish, dairy, rice, poultry, eggs), and cookstove ventilation. The number of participants (and person-years) included in the analysis were, from clean fuels for both cooking and heating to solid fuels for both: 3,631 (25,545), 6,458 (45,462), 1,950 (13,816), and 92,289 (669,897). The relative excess risk due to interaction was 0.17 (95% CI, 0.06-0.27) for cardiovascular mortality, and 0.40 (95% CI, 0.06-0.73) for all-cause mortality. The floating-absolute-risk method provides the variance of the logarithm of the hazard ratio for each category (including the reference category) to facilitate comparisons across the different exposure categories.[2] The boxes represent hazard ratios, with the size inversely proportional to the variance of the logarithm of the hazard ratio, and the horizontal lines represent 95% confidence intervals.

eReferences

1. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011;40(6):1652-66.

2. Plummer M. Improved estimates of floating absolute risk. *Stat Med.* 2004;2(1)3:93-104.