Household air pollution and risk of incident lung cancer in urban China: a prospective cohort study

Chen Ji [1, 2, †], Jun Lv [3, 4, †], Jing Zhang [1, 2, †], Meng Zhu [1, 2], Canqing Yu [3, 4], Hongxia Ma [1, 2], Guangfu Jin [1, 2], Yu Guo [5], Pei Pei [4], Ling Yang [6, 7], Yiping Chen [6, 7], Huaidong Du [6, 7], Zhengming Chen [7], Zhibin Hu [1, 2], Liming Li [3, 4 *], Hongbing Shen [1, 2, *], on behalf of the China Kadoorie Biobank Collaborative Group [‡]

[1] Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China

[2] Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, China

[3] Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing, China

[4] Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing 100191, China

[5] Fuwai hospital Chinese Academy of Medical Sciences, Beijing, China

[6] Medical Research Council Population Health Research Unit at the University of Oxford, Oxford, United Kingdom

[7] Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, United Kingdom

[†] Chen Ji, Jun Lv and Jing Zhang contributed equally to this work.

[‡] The members of steering committee and collaborative group are listed in the online-only supplementary material.

[*] **Corresponding authors:** Hongbing Shen, Department of Epidemiology, School of Public Health, Nanjing Medical University, 101 Longmian Avenue, Nanjing 21116, China, Phone: 86-25-86868439, Email: hbshen@njmu.edu.cn or Liming Li, Department of Epidemiology and Biostatistics, Peking University Health Science Center, 38 Xueyuan Road, Beijing 100191, China, Phone: 86-10-82801528, Email: lmlee@bjmu.edu.cn

Abstract

Household air pollution (HAP) is associated with the development of lung cancer, yet few studies investigated the exposure patterns and joint associations with tobacco smoking. In this study, we included 224,189 urban participants from China Kadoorie Biobank (CKB), 3,288 of which diagnosed with lung cancer during the follow-up. Exposure to four HAP sources (solid fuels for cooking/heating/stove, and environmental tobacco smoke exposure) was assessed at baseline. Distinct HAP patterns and their associations with lung cancer were examined through latent class analysis (LCA) and multivariable Cox regression. A total of 76.1% of the participants reported regular cooking and 52.2% reported winter heating, of which 9% and 24.7% used solid fuels, respectively. Solid fuel heating increased lung cancer risk (Hazards ratio [HR]: 1.25, 95% confidence interval [CI]: 1.08-1.46). LCA identified three HAP patterns; the "clean fuel cooking & solid fuel heating" pattern significantly increased lung cancer risk (HR: 1.25, 95%CI: 1.10-1.41), compared with low HAP pattern. An additive interaction was observed between heavy smoking and "clean fuel cooking & solid fuel heating" (relative excess risk [RERI]: 1.32, 95%CI: 0.29-2.47, attributable proportion [AP]: 0.23, 95%CI: 0.06-0.36). Cases resulting from solid fuel account for ~4% of total cases (population attribute fraction [PAF] $_{\text{overall}}$: 4.31%, 95% CI: 2.16%-6.47%, PAF $_{\text{ever smokers}}$: 4.38%, 95%CI: 1.54%-7.23%). Our results suggest that in urban China, solid fuel heating increased the risk of lung cancer, particularly among heavy smokers. The whole population could benefit from cleaner indoor air quality by reducing using solid fuels, especially smokers.

Keywords: Household air pollution; Solid fuel; Additive interaction; Lung cancer

Novelty and Impact

Household air pollution (HAP) is associated with lung cancer. Here, based on China Kadoorie Biobank (CKB), the authors evaluated potential HAP exposure patterns and their association with lung cancer risk, as well as the joint effects and additive interactions with tobacco smoking. Results showed solid fuel heating increased the risk of lung cancer, particularly among heavy smokers. The whole population would have health benefits by stopping using solid fuels, especially smokers.

Abbreviations

AIC, Akaike Information Criteria; AP, attributable proportion; BIC, Bayesian Information Criteria; BMI, body mass index; CI, confidence interval; CKB, China Kadoorie Biobank; ETS, environmental tobacco smoke exposure; GATS, Global Adult Tobacco Survey; GBD, global burden of disease; HAP, Household air pollution; HR, hazard ratio; IARC, International Agency for Research on Cancer; ICD, International Classification of Diseases; IQR, interquartile range; LCA, latent class analysis; LMICs, low and middle-income countries; PAF, population attribute fraction; PAHs, polycyclic aromatic hydrocarbons; PM, particulate matter; $PM_{2.5}$, fine particulate matter with diameter <2.5 μm; RERI, relative excess risk; SD, standard deviation; TAHS, Tasmanian Longitudinal Health Study.

1. Introduction

Household air pollution (HAP) from the use of solid fuels for cooking or heating is one of the major risk factors for the global burden of disease (GBD). The damage caused by HAP is unevenly distributed across countries in the world, with almost all deaths due to HAP occurring in low and middle-income countries (LMICs)[1]. The GBD study [2]estimated that 2.3 million deaths in 2019 were associated with HAP, including over 0.3 million in China [3]. Thus, controlling HAP is critical to reducing the burden of disease, especially in developing countries like China.

Emissions of HAP from the incomplete solid fuels burning were proved to contain high concentrations of carcinogenic volatile substances, such as benzo[a]pyrene, polycyclic aromatic hydrocarbons (PAHs), and particulate matter (PM) with a diameter ≤ 2.5 μm [4], which had been linked to many different diseases, especially respiratory diseases. In terms of lung cancer, some epidemiologic studies [5-7], mostly cross-sectional or case-control designed studies, have reported the relationship between HAP and lung cancer risk. A retrospective cohort study [8] based on 27,310 participants in Xuanwei, China also found that compared with anthracite, bituminous coal users had a significantly higher risk of lung cancer, along with 36 times higher for men and 99 times higher for women. However, few prospective cohort studies based on large-scale populations have examined the health effects of HAP, limited by the scale and geographical diversity.

Although the Chinese economy is developing rapidly, there are over 450 million individuals in China who still rely on solid fuels [9]. Despite this fact, few previous studies have investigated the association between HAP and lung cancer based on the urban population. Besides, the diverse fuel types and combustion processes may induce different exposure patterns [10], detailed evidence on different exposure patterns of HAP is still inadequate [11]. A cohort study [12] based on the Tasmanian Longitudinal Health Study (TAHS) examined the relationship between exposure to heating/cooking facilities, smoking, and mold and respiratory health in middle-aged adults. The study identified seven HAP profiles and found that exposure to wood heating, gas cooking and heating, and tobacco smoke increased the risks of persistent asthma. However, it considered only selected sources of HAP, ignoring some other sources that are common in developing countries, such as biomass fuel and coal. Given the heterogeneity in the uses of solid fuels in different countries, the exposure patterns of the Chinese population to HAP and their effect on health remains unclear.

In China, coal and wood are the main solid fuels [13]. A variety of behaviors (such as cooking and heating) are associated with solid fuel use. However, previous studies have often failed to fully consider the complex patterns of air pollution exposure caused by multiple behaviors. Besides, concerning participants with different smoking status, previous studies usually ignored the potential joint effects and interactions with smoking. Thus, in this large prospective study based on the China Kadoorie Biobank (CKB), we investigated: (i) the effect of separate HAP sources on risks of incident lung cancer; (ii) potential HAP exposure patterns and their association with lung cancer risk; (iii) possible interactions between HAP exposure patterns and smoking status in relation to lung cancer and (iv) the benefit of cessation of solid fuels in urban areas of China.

2. Materials and Methods

2.1 Study populations

Details of the CKB have been described elsewhere [14, 15]. Briefly, the objective of this cohort is to explore different influential factors such as genetic factors, environmental factors, and lifestyle factors on the occurrence and development of several chronic diseases. The CKB study recruited participants aged from 30 to 79 years through ten geographically diverse regions across China between 2004 and 2008, including 5 urban regions (Harbin, Qingdao, Suzhou, Liuzhou, and Haikou) and 5 rural regions (Gansu, Henan, Sichuan, Zhejiang, and Hunan). In each of these regions, residents without a major disability were identified and invited to participate. Every participant answered an electronic questionnaire and undertook baseline measurements carried out by trained workers following standardized procedures to collect information on sociodemographic characteristics, health-related questions and physical measurements.

2.2 Assessment of HAP sources and tobacco smoking

In the survey, participants were asked to provide information about their residences at baseline and two previous residences where they had lived for at least 1 year. Questions were asked to recall the years of living, cooking frequency (no cooking facility/ never/ rarely, monthly, weekly, daily), cooking fuel (gas/ coal/ wood/ electricity/ other unspecified), whether heat the house in winter (yes/ no), heating fuel (central heating/ gas/ coal/ wood/ electricity/ other unspecified), ownership of stove always kept under slow burning throughout the day (always/ sometimes/ no), fuels used in stove always kept under slow burning (smokeless coal/ smoky coal/ coal brick or coalite/ other unspecified)and environmental tobacco smoke exposure (ETS, also called involuntary or passive smoking, was divided into never/ occasionally/ 1-2 days per week/ 3-5 days per week/ daily). If more than one fuel type was used at any residence, only the one used most frequently was recorded, and further details are available from the study website: http://www.ckbiobank.org.

In accordance with the previous epidemiologic studies [16], fuel types, along with stove condition and ETS exposure were used as proxy indicators of HAP exposure. Participants who cooked less than monthly were considered as noncooking. Those who claimed to expose to ETS daily were classified as exposed to ETS while others were classified as not. Cooking and heating fuels were divided into two types according to their degree of pollution. Solid fuels comprised coal or wood, whilst clean fuels included electricity or gas (central heating additionally for heating fuels) because they tended to generate much less air pollution than solid fuels.

For tobacco smoking, questions included self-reported details related to smoking, including frequency, amount, age started smoking, age stopped, and the key reason for stopping. The smoking category included never smokers, former smokers, and current smokers. Participants were classified as never smokers if they reported never or occasionally smoking (i.e., <100 cigarettes throughout their lifetime) at baseline assessment. For smoking status, all the participants were categorized according to pack-years of smoking which were calculated by multiplying the number of years of smoking by the number of cigarettes smoked per day. For never-smokers, pack-year of smoking was defined as 0. Light smokers referred to people whose pack-years were greater than 0 and less than 30. Heavy smokers referred to people whose pack-years were greater than or equal to 30. In the present study, HAP exposure and smoking status were considered together to investigate the potential joint effects and interactions on lung cancer incidence. Since ambient air pollution and HAP is constantly exchanged [17], outdoor air pollution will influence the concentration of HAP and may confound the association in our assessment [18]. Thus, we also assessed the annual exposure of outdoor $PM_{2.5}$ from 2005 to 2008 using the 4-year mean value in the analysis. The $PM_{2.5}$ exposure was accessed based on satellite aerosol optical depth (AOD) data and ground monitoring station data, and then multiple fillings were used to fill the missing data, which had been described elsewhere [19].

2.3 Assessment of Outcome

The outcome of this study was incident lung cancer. Participants were followed up until the date of a lung cancer diagnosis (classified by the 10th Revision of International Classification of Diseases, ICD-10: codes C33-C34), loss to follow-up, death, or 31 December 2017, whichever came first. Lung cancer cases were ascertained via the Chinese national health insurance claim database, established chronic disease registries, and local death registries semi-annually, supplemented by active confirmation of cancer diagnosis by trained staff who were blinded to baseline data.

2.4 Statistical analysis

Figure S1 shows the construction and analysis process of the cohort. Our analyses were restricted to urban participants in CKB (*n* = 226,193). We first excluded participants with self-reported doctor-diagnosed cancer before the baseline survey (*n*=1,425) and those with missing HAP data (*n*=579), leaving 224,189 participants in the final study. Baseline characteristics of the study populations were presented to identify potential confounders (Table 1).

We used the latent class analysis (LCA)[20] to determine the analysis model based on HAP patterns. Four different types of HAP exposure (main cooking fuel, main heating fuel, main fuel used in stove always kept under slow burning, ETS exposure) were added into the latent class model to specify different HAP patterns. We fit models ranging from two to ten classes and then determined the best fitting models based on a set of criteria [21] : i) the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Log-likelihood were used to select the best model, favoring the one for which the first two values were smaller and the last one smaller in absolute value; ii) the number of participants assigned to each class, where models with adequate sample size per class were preferred and iii) interpretability of each class. Following model selection, participants were assigned to the subgroups in which they had the highest probability of correct assignment. Probabilities of exposure of different HAP exposure derived by LCA were used to characterize different latent classes. These latent class assignments served as the independent variable for subsequent regression analyses. Details of LCA and the process of selecting the final analysis model were provided in supplement materials. Cox proportional-hazards regression models were used to estimate multivariable hazard ratios (HRs) and 95% CIs of incident risks associated with lung cancer, stratified by age-at-risk (five-year intervals), sex, and region (when appropriate), and adjusted for body mass index (BMI, categorized as normal, overweighted, obesity), household income (<¥20,000, ¥20,000-34,999, ≥¥35,000), education level (middle school or lower, high school or higher), family history of cancer (yes, no, missing data), smoking category (never smoker, former smoker, current smoker), pack-years of smoking (continuous), physical activity (MET-h/d; continuous), mean value of outdoor $PM_{2.5}$ from 2005 to 2008 (continuous) and cookstove ventilation (all stoves had chimney or extractor, no/ not all stove), where appropriate. Fuller details of the assessment of covariates for adjustment were provided in supplement materials. Schoenfeld residuals were used to test the proportional hazards assumption.

We estimated the joint effects of different HAP patterns and different smoking status. The additive interactions were also tested by 2 epidemic indexes [22]: the relative excess risk because of the interaction (RERI), along with the attributable proportion because of the interaction (AP). The 95% CIs of the RERI and AP were generated by drawing 5000 bootstrap samples from the estimation dataset. Additive interaction would not be considered if the 95%CIs of the RERI and AP included 0.

To calculate the proportion of incident lung cancer that could be prevented if solid fuels were to be eliminated from the study populations during a certain time interval, we calculated the population attribute fraction (PAF) [23] of never-smokers and ever-smokers among participants. The 5-year PAF calculated the proportion of incident lung cancer in the estimated population attributable to solid fuel in the cohort over a 5-year follow-up period [24].

We conducted subgroup analyses by baseline characteristics (age-at-risk, sex, BMI, smoking status, educational level, household income, outdoor air pollution, cookstove ventilation). We also carried out several sensitivity analyses to examine the potential impact of residual confounding by: (i) excluding participants without complete covariates data (n = 9,777); (ii) excluding participants who were diagnosed with lung cancer within the first year of follow-up (n = 162); (iii) stratified analysis based on the number of years living in the most recent residence, iv) LCA conducted based on the fuel used in the past three residences surveyed in the baseline questionnaire, and (v) excluding participants with unreliable years of residence information (i.e. participants with a difference of more than 1 year between the sum of years in three surveyed residences and baseline age) (n = 115).

In our statistical analysis, all *P*-values were two-sided, and $P<0.05$ was considered statistically significant. All analyses were conducted using R software version 4.0.2. The latent class analysis, Cox regression, and the calculation of PAF were performed using R packages, including the poLCA, survival, and AF packages.

3. Results

Among the 224,189 urban participants, 3288 lung cancer events were observed during the follow-up. Table 1 shows the distribution of the baseline characteristics of the participants. The mean (SD) age was 52.8(10.9) years old and 59.6% of participants were female. Compared to non-lung cancer participants, participants with lung cancer were older, with lower household income, more likely to be current smokers or former smokers with longer pack-years of smoking and to use solid heating fuels and smoky coal.

As surveyed through the baseline questionnaire, we tested the risk of lung cancer exposed to different HAP sources separately (Figure S2). A significant association was observed between the use of solid fuel heating and lung cancer (HR: 1.25, 95% CI: 1.08-1.46), but not in solid fuel cooking (HR: 1.16, 95% CI: 0.98-1.37). Using smokeless coal or smoky coal in stove always kept under slow burning did not significantly increase the risk of incident lung cancer ($HR_{\text{smokeless coal}}$: 1.17, 95% CI: 0.96-1.42; $HR_{\text{smoky coal}}$: 1.19, 95% CI: 0.85-1.65). Besides, we also observed an inverse association between ETS exposure and risk of lung cancer. As described in the method section, we added four HAP sources into the LCA to investigate the exposure pattern of HAP. Probabilities of various types of exposure derived by LCA and fit Statistics for the LCA models were presented in Table 2 and Figure S3. Details of the selection of models could be seen in supplementary methods. Considering several fitting criteria, we finally selected the 3-class model. According to the maximum response probability of LCA (Table 2), there was no difference between the subgroups of the 3-class models on the characteristic of ETS exposure and stove always kept under slow burning. Therefore, three different HAP patterns identified by LCA were named "clean fuel cooking & clean fuel heating " (class 3), "solid fuel cooking & never heating " (class 2), and "clean fuel cooking & solid fuel heating " (class 1) according to cooking and heating fuels. When compared to the "clean fuel cooking & clean fuel heating" pattern, the "clean fuel cooking & solid fuel heating" pattern significantly increased the risk of lung cancer (HR: 1.25, 95% CI: 1.10-1.41). However, we did not observe an association between the "solid fuel cooking & never heating" pattern and lung cancer risk (Table 3). Stratified analyses showed that the effects of exposure to different HAP patterns on lung cancer were generally similar across different subgroups (Table S1).

In the sensitivity analyses, we found no significant difference in the HRs of different HAP patterns with lung cancer before and after the exclusion of participants with incomplete covariate data (Table S2). When excluding incident cases during the first year of follow-up, the association was robust (Table S3). Analysis was also performed based on the strategy that all the participants could be stratified by year of living in the most recent residence (Table S4). For participants with year of living < 5, using "Clean fuel cooking & clean fuel heating" patterns as reference, HR for "Solid fuel cooking & never heating" was 0.90 (95% CI: 0.55-1.48); HR for "Clean fuel cooking & solid fuel heating" was 1.19 (95% CI: 0.85-1.67); for participants with year of living ⩾ 5, HR for "Solid fuel cooking & never heating" was 1.22 (95% CI: 1.02-1.45); HR for "Clean fuel cooking & solid fuel heating" was 1.25 (95% CI: 1.09-1.43).

The association was not significant when the analysis was restricted to participants whose years of living <5, while for participants who live longer than 5 years in the most recent residence, the association was still significant.

As described in the method section, we categorized all the participants into different smoking status according to their pack-years and calculated their risk of lung cancer (Table S5). For heavy smokers, the risk of lung cancer was nearly 4.5 times higher than that for nonsmokers. When combining the HAP exposure and smoking status, we used those who had never smoked and always used clean fuels for cooking and heating as the reference group, and observed a joint effect on lung cancer risk that presented in a dose-response manner (Fig. 1). Heavy smokers with the "clean fuel cooking & solid fuel heating" pattern had the highest risk of incident lung cancer (HR= 5.87, 95% CI:4.79-7.21). Furthermore, it was worth noting that for heavy smokers, the RERI was 1.32 (95% CI: 0.29-2.47) suggested there would be 1.32 relative excess risk because of the additive interaction, accounting for 23% (6%- 36%) of the risk of lung cancer in heavy smokers who were exposed to the "clean fuel cooking & solid fuel heating" HAP pattern (Table 4).

To estimate the health benefit of cessation of solid fuels in this cohort, we estimated the 5-year PAF of using solid fuels (including solid fuel heating and cooking) in the cohort. Cases resulting from solid fuel use account for 4.31% of total cases ($PAF_{overall}$: 4.31%, 95% CI: 2.16%-6.47%), while 4.38% of incident lung cancer could have been reduced by stopping the use of solid fuels among ever smokers ($PAF_{ever\ smokers}$: 4.38%, 95%CI: 1.54%-7.23%) (Table

5), indicating that the population would benefit from stopping using solid fuels, especially for ever smokers.

To examine the effect of the switch of fuels, we further considered the exposure to four HAP sources (solid fuels for cooking/ heating/ stove fuels, and ETS) in the three most recent dwellings surveyed in the baseline questionnaire. Fit statistics for the new LCA models were shown in Figure S3. It could be derived from the fit statistics that there was no substantial improvement for HAP exposure patterns over the 5-class model. Probabilities of different HAP exposure by 5-class HAP patterns obtained by LCA were presented in Table S6. Similarly, we named the 5-class pattern according to the feature of cooking and heating fuels. Compared with the "never cooking & never heating" pattern (class 3), the "solid to clean cooking fuel & solid to clean heating fuel" pattern (class 1) was still significantly associated with lung cancer risk (HR = 1.22, 95%CI: 1.00 1.48) (Table S7). However, we did not observe a significant association for participants who had the switch of solid fuel for cooking or heating alone (i.e., "solid to clean cooking fuel & never heating" group and "never cooking & solid to clean heating fuel" group). The observed associations did not change significantly after the exclusion of participants lacking reliable year of residence information (Table S8).

4. Discussion

In this study, we found that the use of solid fuel for heating was associated with a higher (~ 25%) risk of lung cancer compared with the use of clean fuel. Using the LCA, the complex sources of HAP were classified into 3 distinct groups, namely "clean fuel cooking & clean fuel heating", "solid fuel cooking & never heating", and "clean fuel cooking & solid fuel heating". Compared with HAP exposure pattern with clean fuels for cooking and heating, the "clean fuel cooking & solid fuel heating" pattern significantly increased the risk of lung cancer, with consistent results across a range of population subgroups. When examining the joint effects between HAP exposure patterns and smoking status, the greatest relative increase of risk was observed among heavy smokers with a "clean fuel cooking & solid fuel heating" exposure pattern. We also found an additive interaction between the "clean fuel cooking &solid fuel heating" pattern and heavy smoking. These findings suggested that solid fuel heating played an important role in the development of lung cancer. The exposure pattern of "clean fuel cooking & solid fuel heating" could also enhance the negative effect of heavy smoking.

In 2010, the International Agency for Research on Cancer (IARC) classified indoor emissions from household combustion of coal and biomass as Group 1 carcinogen and Group 2A probable carcinogen respectively [25]. Findings of several studies, including some meta-analysis [26, 27] and prospective studies [8, 28, 29] have shown an increased risk of lung cancer among solid fuel users, which accord with our current study findings. Nevertheless, previous studies usually focused on a single source of HAP, ignoring that the exposure components usually included other sources, such as cooking fuels, heating fuels, stove fuels and ETS. Increasing evidence from HAP studies [30] also indicated the complexity and heterogeneity in exposure patterns. In the present study, we investigated 4 different types of HAP exposure sources and confirmed the associations between only solid fuel heating and lung cancer risk through a large prospective study (HR: 1.25, 95% CI: 1.08-1.46). The HAP exposure patterns analysis further showed the association of the "clean fuel cooking & solid fuel heating" pattern with the risk of lung cancer (HR: 1.25, 95% CI: 1.10-1.41). These results indicated that solid fuel heating is relevant to lung cancer. A possible contributing factor is that participants using solid fuels for heating were more than participants exposed to other HAP sources, which may influence the statistical power of the analysis. Another explanation is that different exposure patterns may account for the results [16]. While solid fuel cooking produces HAP lasting for a short period, several times a day, solid fuel heating usually results in a higher duration of HAP exposure throughout the day. Moreover, people tend to take insulation measures to close doors and windows when heating, resulting in high concentrations of HAP. Thus, HAP from solid fuel heating may result in a higher cumulative inhaled contaminants dose. A meta-analysis [31] of solid fuel exposure and risk of lung cancer showed that the estimated odds ratio of exposure to solid fuel associated with lung cancer was 1.70 (95%CI: 1.50 1.94). The effect of solid fuel in our study (~25% increased risk) was lower than that in the meta-analysis, the possible reasons may be as follows: (i) previous studies tend to consider all sources of exposure uniformly, while our study considered different HAP sources separately, which may lead to a lower observed effect; (ii) according to the survey data of China Family Panel Studies (CFPS) [32], the proportion of solid fuel used in China has been declining while that of clean fuel has been increasing, which may contribute to improvements of indoor air quality, leading to the underestimate of the effect in our study.

Previous studies [33] found a positive association between lung cancer and ETS. However, we observed an inverse association in the present study. A previous study [34] based on CKB found a significant inverse association between airflow obstruction and passive smoking. Besides, the Global Adult Tobacco Survey (GATS)[35] conducted in 2018 found that nearly half of the population was exposed to ETS, while 35% of the participants in the present study reported exposure to ETS, indicating that exposure to ETS in China may be hard to define. Thus, we supposed the present association to be affected by some residual confounding or the current ETS exposure assessments may not be accurate. Besides, a 2~5-fold increased risk was found among smokers considering the risk between active smoking and lung cancer, which was in accordance with the effect observed among Asian population in previous studies [36], and was lower than that in European and American populations (9.4-23.2 fold increased risk), suggesting racial differences in smoke-related lung cancer risk [37], and we expect further discussion of these differences in subsequent studies.

To the best of our knowledge, our study is the first to assess the joint effects and possible additive interactions between HAP exposure patterns and smoking status in China. Results observed that the joint effects presented in a dose-response manner and revealed the "clean fuel cooking & solid fuel heating" exposure pattern and active smoking synergistically increased the risk of lung cancer. This was in accordance with previous studies [38, 39] claiming that solid fuel users had an increased risk of lung cancer, especially among cigarette smokers. The observed results are plausible, given that both combustions of tobacco and solid fuel share similar carcinogenic constituents [40], which may increase the risk of incident lung cancer to a greater extent.

The main strengths of this study lay in the scale and diversity of our participants, covering over 220,000 adults across five geographically diverse urban areas of China. Moreover, we simultaneously examined four latent sources of HAP exposure (cooking fuel, heating fuel, fuels used in stoves kept under slow burning, and ETS exposure) using the LCA which was based on the posterior probability thus helping avoid the subjectivity of classification [41], and for the first time, explored the interactions between HAP exposure patterns and smoking status. This study also examined the longitudinal relationship between solid fuel using and lung cancer considering self-reported fuel types used in the three most recent residences, which involved participants who were using clean fuel in the baseline residence but had used solid fuel in one or more earlier residence(s). Additionally, it is worth mentioning that our study also provided quantitative data about the health benefit of cessation of solid fuels for never-smokers and ever-smokers among participants. These results have public health implications, supporting the benefit of improving indoor air quality for the entire population, and urging for the promotion of clean fuels. Meanwhile, heavy smokers should pay more attention to HAP exposure. Further studies will be needed to confirm our findings.

Our study also has several limitations. First, solid fuel use was self-reported although this referred to many other previous studies [30, 42], which might cause recall bias. Second, we lacked detailed information on secondary types of fuel use and the ventilation of the whole house, as well as the effectiveness of ventilation devices, which might cause the misclassification of different HAP exposure. Besides, the fuel type may have shifted during the follow-up period; however, we don't have data on changes in fuel use during the follow-up. Thus, the current results should be considered with caution. Third, although we carefully adjusted for several covariates related to demographics status, socioeconomic status, and outdoor air pollution status, residual confounding by unmeasured (e.g., occupational exposure) or unknown factors remains possible.

In summary, our findings demonstrated that household solid fuel use, especially solid fuel heating was associated with an increased risk of incident lung cancer in urban China. There was an additive interaction between heavy smoking and the "clean fuel cooking & solid fuel heating" exposure pattern. These findings reinforced the urgency to control indoor air quality by improving access to clean fuels and reducing the use of solid fuels in urban China, especially for smokers.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science of China (9194330001), the Excellent Youth Foundation of Jiangsu Province (BK20220100), the National Natural Science of China (81820108028, 81973123, 82273724, 82273714), and Research Unit of Prospective Cohort of Cardiovascular Diseases and Cancer, Chinese Academy of Medical Sciences (2019RU038).

CONFLICT OF INTEREST

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

AUTHOR CONTRIBUTIONS

Chen Ji: Methodology, Formal analysis, Writing – original draft. **Jun Lv**: Writing – review & editing, Data curation. **Jing Zhang**: Formal analysis, Writing – original draft. **Meng Zhu**: Conceptualization, Writing – review & editing, Data curation. **Canqing Yu**: Data curation. **Hongxia Ma**: Conceptualization, Writing – review & editing. **Guangfu Jin**: Writing – review & editing, Data curation. **Yu Guo**: Data curation. **Pei Pei**: Data curation. **Ling Yang**: Data curation. **Yiping Chen**: Data curation. **Huaidong Du**: Data curation. **Zhengming Chen**: Data curation. **Zhibin Hu**: Writing – review & editing. **Liming Li**: Conceptualization, Resources, Supervision. **Hongbing Shen**: Conceptualization, Resources, Supervision, Funding acquisition, Writing – review & editing. The work reported in the paper has been performed by the authors, unless clearly specified in the text.

DATA AVAILABILITY STATEMENT

Requests for data should be submitted to the China Kadoorie Biobank (CKB) Data Access Committee. Details of how to access China Kadoorie Biobank data and details of the data release schedule are available from https://www.ckbiobank.org/site/Data+Access. Further information is available from the corresponding author upon request.

ETHICS STATEMENT

The CKB study was approved by the Ethical Committee of the Chinese Center for Disease Control and Prevention (Beijing, China) and the Oxford Tropical Research Ethics Committee at the University of Oxford (Oxford, UK). Written informed consent was obtained from all participants.

References

1. World Health Organization. Burden of disease from household air pollution for 2016: World Health Organization, 2018.

2. Fuller R, Landrigan PJ, Balakrishnan K, Bathan G, Bose-O'Reilly S, Brauer M, Caravanos J, Chiles T, Cohen A, Corra L, Cropper M, Ferraro G, et al. Pollution and health: a progress update. The Lancet. Planetary Health 2022;6:e535-e47.

3. Institute for Health Metrics and Evaluation. 2019.

4. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. Nat Rev Cancer 2007;7:778-90.

5. Liu ZY, He XZ, Chapman RS. Smoking and other risk factors for lung cancer in Xuanwei, China. Int J Epidemiol 1991;20:26-31.

6. Raspanti GA, Hashibe M, Siwakoti B, Wei M, Thakur BK, Pun CB, Al-Temimi M, Lee YC, Sapkota A. Household air pollution and lung cancer risk among never-smokers in Nepal. Environ Res 2016;147:141-5.

7. Liu Y, Tian Z, He X, Wang X, Wei H. Short-term effects of indoor and outdoor air pollution on the lung cancer morbidity in Henan Province, Central China. Environ Geochem Health 2021.

8. Barone-Adesi F, Chapman RS, Silverman DT, He X, Hu W, Vermeulen R, Ning B, Fraumeni JF, Jr., Rothman N, Lan Q. Risk of lung cancer associated with domestic use of coal in Xuanwei, China: retrospective cohort study. BMJ 2012;345:e5414.

9. Chan KH, Bennett DA, Kurmi OP, Yang L, Chen Y, Lv J, Guo Y, Bian Z, Yu C, Chen X, Dong C, Li L, et al. Solid fuels for cooking and tobacco use and risk of major chronic liver disease mortality: a prospective cohort study of 0.5 million Chinese adults. International Journal of Epidemiology 2020;49:45-55.

10. Chen Y, Shen G, Liu W, Du W, Su S, Duan Y, Lin N, Zhuo S, Wang X, Xing B. Field measurement and estimate of gaseous and particle pollutant emissions from cooking and space heating processes in rural households, northern China. Atmospheric environment 2016;125:265-71.

11. Chan KH, Lam KBH, Kurmi OP, Guo Y, Bennett D, Bian Z, Sherliker P, Chen J, Li L, Chen Z. Trans-generational changes and rural-urban inequality in household fuel use and cookstove ventilation in China: A multi-region study of 0.5 million adults. Int J Hyg Environ Health 2017;220:1370-81.

12. Dai X, Bui DS, Perret JL, Lowe AJ, Frith PA, Bowatte G, Thomas PS, Giles GG, Hamilton GS, Tsimiklis H, Hui J, Burgess J, et al. Exposure to household air pollution over 10 years is related to asthma and lung function decline. Eur Respir J 2021;57.

13. Tang X, Liao H. Energy poverty and solid fuels use in rural China: Analysis based on national population census. Energy for Sustainable Development 2014;23:122-29.

14. Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, Linksted P, Peto R. Cohort profile: the Kadoorie Study of Chronic Disease in China (KSCDC). Int J Epidemiol 2005;34:1243-9.

15. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, Li L. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol 2011;40:1652-66.

16. Li J, Qin C, Lv J, Guo Y, Bian Z, Zhou W, Hu J, Zhang Y, Chen J, Cao W, Yu C, Li L. Solid Fuel Use and Incident COPD in Chinese Adults: Findings from the China Kadoorie Biobank. Environ Health Perspect 2019;127:57008.

17. Wallace L, Williams R. Use of personal-indoor-outdoor sulfur concentrations to estimate the infiltration factor and outdoor exposure factor for individual homes and persons. Environ Sci Technol 2005;39:1707-14.

18. !!! INVALID CITATION !!! 19.

19. Xiao Q, Chen H, Strickland MJ, Kan H, Chang HH, Klein M, Yang C, Meng X, Liu Y. Associations between birth outcomes and maternal PM exposure in Shanghai: A comparison of three exposure assessment approaches. Environment International 2018;117:226-36.

20. Vermunt JK, Magidson J. "Latent Class Analysis" in The Sage encyclopedia of social science research methodsed.: Sage Publications, Inc, 2004.

21. Naldi L, Cazzaniga S. Research Techniques Made Simple: Latent Class Analysis. J Invest Dermatol 2020;140:1676-80.e1.

22. Piper M, Mair W, Partridge L. Comment by Matthew Piper, William Mair, Linda Partridge on Min, K.J., Flatt, T., Kulaots, I., Tatar, M. (2006) "Counting calories in Drosophila dietary restriction"Exp. Gerontology, doi:10.1016/j.exger.2006.10.009. Exp Gerontol 2007;42:253-55.

23. Mansournia MA, Altman DG. Population attributable fraction. BMJ 2018;360:k757.

24. Laaksonen MA, Knekt P, Härkänen T, Virtala E, Oja H. Estimation of the population attributable fraction for mortality in a cohort study using a piecewise constant hazards model. Am J Epidemiol 2010;171:837-47.

25. Iarc Working Group on the Evaluation of Carcinogenic Risks to Humans. Household use of solid fuels and high-temperature frying. IARC Monogr Eval Carcinog Risks Hum 2010;95:1-430.

26. Hosgood HD, 3rd, Boffetta P, Greenland S, Lee YC, McLaughlin J, Seow A, Duell EJ, Andrew AS, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, Lissowska J, et al. In-home coal and wood use and lung cancer risk: a pooled analysis of the International Lung Cancer Consortium. Environ Health Perspect 2010;118:1743-7.

27. Kurmi OP, Arya PH, Lam KB, Sorahan T, Ayres JG. Lung cancer risk and solid fuel smoke exposure: a systematic review and meta-analysis. Eur Respir J 2012;40:1228-37.

28. Lan Q, Chapman RS, Schreinemachers DM, Tian L, He X. Household stove improvement and risk of lung cancer in Xuanwei, China. Journal of the National Cancer Institute 2002;94:826-35.

29. Cheng ES, Chan KH, Weber M, Steinberg J, Young J, Canfell K, Yu XQ. Solid Fuel, Second-Hand Smoke, and Lung Cancer Mortality: A Prospective Cohort of 323,794 Chinese Never-Smokers. American Journal of Respiratory and Critical Care Medicine 2022.

30. Clark ML, Peel JL, Balakrishnan K, Breysse PN, Chillrud SN, Naeher LP, Rodes CE, Vette AF, Balbus JM. Health and household air pollution from solid fuel use: the need for improved exposure assessment. Environ Health Perspect 2013;121:1120-8.

31. Kurmi OP, Arya PH, Lam K-BH, Sorahan T, Ayres JG. Lung cancer risk and solid fuel smoke exposure: a systematic review and meta-analysis. The European Respiratory Journal 2012;40:1228-37.

32. Chen Y, Shen H, Zhong Q, Chen H, Huang T, Liu J, Cheng H, Zeng EY, Smith KR, Tao S. Transition of household cookfuels in China from 2010 to 2012. Applied Energy 2016;184:800-09.

33. Kim AS, Ko H-J, Kwon J-H, Lee J-M. Exposure to Secondhand Smoke and Risk of Cancer in Never Smokers: A Meta-Analysis of Epidemiologic Studies. International Journal of Environmental Research and Public Health 2018;15.

34. Smith M, Li L, Augustyn M, Kurmi O, Chen J, Collins R, Guo Y, Han Y, Qin J, Xu G, Wang J, Bian Z, et al. Prevalence and correlates of airflow obstruction in approximately 317,000 never-smokers in China. Eur Respir J 2014;44:66-77.

35. Chinese Center for Disease Control and Prevention, World Health Organization, US Centers for Disease Control and Prevention. Global Adult Tobacco Survry Fact Sheet China 2018, 2019.

36. Jung KJ, Jeon C, Jee SH. The effect of smoking on lung cancer: ethnic differences and the smoking paradox. Epidemiol Health 2016;38:e2016060.

37. Haiman CA, Stram DO, Wilkens LR, Pike MC, Kolonel LN, Henderson BE, Le Marchand L. Ethnic and racial differences in the smoking-related risk of lung cancer. N Engl J Med 2006;354:333-42.

38. Hosgood HD, Boffetta P, Greenland S, Lee Y-CA, McLaughlin J, Seow A, Duell EJ, Andrew AS, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, Lissowska J, et al. In-home coal and wood use and lung cancer risk: a pooled analysis of the International Lung Cancer Consortium. Environmental Health Perspectives 2010;118:1743-47.

39. Tang L, Lim WY, Eng P, Leong SS, Lim TK, Ng AW, Tee A, Seow A. Lung cancer in Chinese women: evidence for an interaction between tobacco smoking and exposure to inhalants in the indoor environment. Environ Health Perspect 2010;118:1257-60.

40. Sood A. Indoor fuel exposure and the lung in both developing and developed countries: an update. Clin Chest Med 2012;33:649-65.

41. Sinha P, Calfee CS, Delucchi KL. Practitioner's Guide to Latent Class Analysis: Methodological Considerations and Common Pitfalls. Crit Care Med 2021;49:e63-e79.

42. Liu T, Chen R, Zheng R, Li L, Wang S. Household Air Pollution From Solid Cooking Fuel Combustion and Female Breast Cancer. Frontiers In Public Health 2021;9:677851.

Figure captions

Figure 1. Joint effects of HAP patterns and smoking status.

Definition of abbreviations: HAP, household air pollution; HR, hazards ratio; CI, confidence interval;

Subgroups were defined by pack-years as never smokers referred to people whose pack-year =0; light smokers (0<pack-years <30); heavy smokers (pack-years ≥30). The hazard ratios for lung cancer according to the joint effect of air pollution patterns and pack years of smoking were estimated using Cox proportional hazard models with adjustment for age at risk, sex, region(stratified), BMI, household income, education level, family history of cancer, smoking category, pack-years of smoking, physical activity, outdoor $PM_{2.5}$, cookstove ventilation.

Table 1. Population characteristics included in the study

Characteristics	Lung cancer(*n*=3,288)	Non-lung cancer(*n*=220,901)	Overall(*n*=224,189)
Sex, %			
Male	56.7(*n*=1,863)	40.1(*n*=88,651)	40.4(*n*=90,514)
Female	43.3(*n*=1,425)	59.9(*n*=132,250)	59.6(*n*=133,675)
Age(years), mean (SD)	61.2(9.5)	52.7(10.8)	52.8(10.9)
Age, %			
<60 years	40.9(*n*=1,346)	72.5(*n*=160,142)	72.0(*n*=161,488)
≥60 years	59.1(*n*=1,942)	27.5(*n*=60,759)	28.0(*n*=62,701)
BMI (kg/m2), mean (SD)	23.9(3.5)	24.3(3.4)	24.3(3.4)
BMI, %			
Normal (<25 kg/m^2)	62.3(*n*=2,047)	59.9(*n*=132,255)	60.0(*n*=134,302)
Overweight (25 to 29.9 kg/m^2)	33.1(*n*=1,088)	34.7(*n*=76,681)	34.7(*n*=77,769)
Obesity (≥30 kg/m^2)	4.7(*n*=153)	5.4(*n*=11,965)	5.4(*n*=12,118)
Education level, %			
High school or higher	25.8(*n*=847)	36.0(*n*=79,578)	35.9(*n*=80,425)
Middle school or lower	74.2(*n*=2,441)	64.0(*n*=141,323)	64.1(*n*=143,764)
Household income, %			
<¥20,000	50.5(*n*=1,661)	42.7(*n*=94,316)	42.8(*n*=95,977)
¥20,000-34,999	29.7(*n*=975)	31.9(*n*=70,406)	31.8(*n*=71,381)
≥¥35,000	19.8(*n*=652)	25.4(*n*=56,179)	25.3(*n*=56,831)
Family history of cancer, %			
Yes	21.5(*n*=708)	19.0(*n*=41,863)	19.0(*n*=42,571)
No	73.8(*n*=2,427)	76.7(*n*=169,414)	76.7(*n*=171,841)
Missing	4.7(*n*=153)	4.4(*n*=9,624)	4.4(*n*=9,777)
Smoking category, %			
Never	46.0(*n*=1,514)	70.4(*n*=155,453)	70.0(*n*=156,967)
Former	12.4(*n*=408)	6.5(*n*=14,332)	6.6(*n*=14,740)
Current	41.5(*n*=1,366)	23.1(*n*=51,116)	23.4(*n*=52,482)
Pack years of smoking, mean (SD)	18.5(22.9)	7.2(14.8)	7.4(15.1)
Physical activity (MET-h per day) [a]**, median (IQR)**	11.2(11.83)	15.0(17.71)	15.0(17.61)
Cooking fuel, %			
Never (<Monthly)	29.5(*n*=970)	23.8(*n*=52,589)	23.9(*n*=53,559)
Clean fuel	63.6(*n*=2,090)	69.4(*n*=153,248)	69.3(*n*=155,338)
Solid fuel	6.9(*n*=228)	6.8(*n*=15,064)	6.8(*n*=15,292)
Heating fuel, %			
Never	39.7(*n*=1,304)	47.9(*n*=105,914)	47.8(*n*=107,218)
Clean fuel	42.4(*n*=1,395)	39.2(*n*=86,669)	39.3(*n*=88,064)
Solid fuel	17.9(*n*=589)	12.8(*n*=28,318)	12.9(*n*=28,907)
Stove always kept under slow burning, %			
Never/Sometimes	95.4(*n*=3,136)	96.9(*n*=214,012)	96.9(*n*=217,148)
Smokeless coal/Coalite	3.5(*n*=115)	2.5(*n*=5,469)	2.5(*n*=5,584)

Smoky coal	1.1(n=37)	0.6(n=1,420)	0.6(n=1,457)
Cookstove ventilation, %			
Yes	74.7(n=2,455)	74.6(n=164,732)	74.6(n=167,187)
No/ Not all stoves	25.3(n=833)	25.4(n=56,169)	25.4(n=57,002)
Environmental tobacco smoke exposure			
No	67.2(n=2,208)	64.8(n=143,098)	64.8(n=145,306)
Yes	32.8(n=1,080)	35.2(n=77,803)	35.2(n=78,883)
PM$_{2.5}$			
Low pollution	47.4(n=1,560)	49.9(n=110,133)	49.8(n=111,693)
High pollution	52.6(n=1,728)	50.1(n=110,768)	50.2(n=112,496)

Definition of abbreviations: BMI, body mass index; SD, standard deviation; IQR, interquartile range; PM$_{2.5}$, fine particulate matter with diameter <2.5 μm.

[a] MET-h per day=metabolic equivalent of task h per day. 1 MET-h is defined as 1 kcal/kg per h

Table 2. Probabilities of HAP exposure by different patterns derived by LCA

Item		Latent class 1	Latent class 2	Latent class 3
Prevalence		0.1291	0.0659	0.8050
Cooking fuel	Never (<Monthly)	0.2135	0.2409	0.2431
	Clean fuel	**0.7394**	0.3386	**0.7569**
	Solid fuel	0.0470	**0.4204**	1.26E-46
Heating fuel	Never	0.0048	**0.9440**	0.4686
	Clean fuel	0.0145	0.0500	**0.5313**
	Solid fuel	**0.9807**	0.0060	2.28E-05
Stove fuel	Never/Sometimes	**0.8167**	**0.9579**	**0.9982**
	Smokeless coal/Coalite	0.1336	0.0420	0.0018
	Smoky coal	0.0497	0.0001	1.60E-123
Environmental tobacco smoke exposure	No	**0.6515**	**0.6271**	**0.6518**
	Yes	0.3485	0.3729	0.3482

Definition of abbreviations: LCA: Latent class analysis.

Table 3. Risk of lung cancer according to different HAP patterns

HAP exposure profiles	No. cases/ person years	Model 1[a]		Model 2[b]	
		HR (95%CI)	*P* value	HR (95%CI)	*P* value
Clean fuel cooking & clean fuel heating	2493/1938732	Ref.		Ref.	
Solid fuel cooking & never heating	206/158111	1.19(1.02-1.39)	0.026	1.16(0.99-1.37)	0.072
Clean fuel cooking & solid fuel heating	589/310110	1.47(1.30-1.66)	5.66E-10	1.25(1.10-1.41)	4.69E-04

Definition of abbreviations: HAP, household air pollution; HR, hazards ratio; CI, confidence interval;

[a] Adjusted for age at risk, sex, region(stratified)

[b] Adjusted for age at risk, sex, region(stratified), BMI, household income, education level, family history of cancer, smoking category, pack-years of smoking, physical activity, outdoor $PM_{2.5}$, cookstove ventilation

Table 4. RERI and AP for additive interaction between smoking status and HAP exposure profiles [a]

HAP exposure profiles	Smoking status [c]			
	Light smokers		Heavy smokers	
	RERI [a](95%CI)	AP [b] (95%CI)	RERI [a] (95%CI)	AP [b] (95%CI)
Solid fuel cooking& never heating	0.39(-0.46-1.32)	0.16(-0.26-0.39)	0.83(-0.52-2.37)	0.16(-0.13-0.35)
Clean fuel cooking & solid fuel heating	0.13(-0.38-0.67)	0.06(-0.19-0.24)	**1.32(0.29-2.47)**	**0.23(0.06-0.36)**

Definition of abbreviations: HAP, household air pollution; RERI, relative excess risk due to interaction; AP, attributable proportion due to interaction; CI, confidence interval;

[a] Adjusted for age at risk, sex, region(stratified), BMI, household income, education level, family history of cancer, smoking category, pack-years of smoking, physical activity, outdoor $PM_{2.5}$, cookstove ventilation

[b] To estimate RERI and AP, the lowest pollution exposure profile(clean fuel cooking & clean fuel heating) and the lowest pack-year group(pack year=0) were the reference categories.

[c] Defined by pack-years of smoking: light(0<pack-years <30); heavy(pack-years ⩾30)

Table 5. 5-year PAF between people with different smoking status

Smoking status [b]	Attribute cases(95%CI)	Total cases	PAF%(95%CI)	*P* value [a]
Never smoker(*n*=157,043)	21(4-37)	501	4.12(0.79-7.44)	0.015
Ever Smoker(*n*=67,146)	27(10-45)	626	4.38(1.54-7.23)	0.003
Overall(*n*=224,189)	49(24-73)	1127	4.31(2.16-6.47)	8.85E-05

Definition of abbreviations: PAF, population attributable fraction; CI, confidence interval;

[a] Adjusted for age at risk, sex, region(stratified), BMI, household income, education level, family history of cancer, smoking category, pack-years of smoking, physical activity, outdoor $PM_{2.5}$, cookstove ventilation

[b] Defined by pack-years of smoking: never smokers (pack-year =0); ever smokers (pack-years >0)