Long-term Fine Particulate Matter Exposure on Lung Cancer Incidence and Mortality in Chinese Nonsmokers

Meng Zhu, PhD[1,2,3†], Yuting Han, PhD[4†], Yuanlin Mou, MD[1†], Xia Meng, PhD [5†], Chen Ji, MD[1], Xia Zhu, MD[1], Canqing Yu, PhD[4,6,7], Dianjianyi Sun, PhD[4,6,7], Ling Yang, PhD[8,9], Qiufen Sun, PhD[1], Yiping Chen, DPhil[8,9], Huaidong Du, PhD[8,9], Juncheng Dai, PhD[1,2], Zhengming Chen, PhD[9], Zhibin Hu, PhD[1,2], Jun Lv, PhD[4,6,7,10], Guangfu Jin, PhD[1,2,3], Hongxia Ma, PhD[1,2,3], Haidong Kan, PhD[5*], Liming Li, MPH[4,6,7*], Hongbing Shen, PhD[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 211166, China.

[2] Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Medicine and China International Cooperation Center for Environment and Human Health, Nanjing Medical University, Nanjing 211166, China.

[3] Wuxi Medical Center, Nanjing Medical University, Nanjing 211166, China.

[4] Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing 100191, China.

[5] School of Public Health, Key Laboratory of Public Health Safety of the Ministry of Education, National Health Commission Key Laboratory of Health Technology Assessment, Integrated Research on Disaster Risk, International Centers of Excellence on Risk Interconnectivity and Governance on Weather/Climate Extremes Impact and Public Health, Fudan University, Shanghai, China.

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

[7] Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing 100191, China.

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

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

[10] State Key Laboratory of Vascular Homeostasis and Remodeling, Peking University, Beijing 100191, China.

Corresponding authors:

Prof Hongbing Shen, Ph.D., Department of Epidemiology, School of Public Health, Nanjing Medical University, 101 Longmian Avenue, Nanjing 211166, China. Tel./Fax: +86-25-86868437, E-mail: hbshen@njmu.edu.cn.

Prof Liming Li, MD, MPH, Department of Epidemiology and Biostatistics, Peking University Health Science Center, 38 Xueyuan Road, Beijing 100191, China, Tel./Fax: +86-10-82801528, E-mail: lmlee@bjmu.edu.cn.

[†] Authors share co-first authorship

[*] Senior authors who jointly directed the study

Authors' Contributions:

HS, LL, and HK supervised the entire project and designed the work. MZ, YH, YM, and XM contributed to the data analysis, data interpretation, and writing of the report. CJ, XZ, CY, DS, LY, QS, YC, HD, JD, ZC, ZH, JL, GJ, and HM contributed to the discussion, data interpretation and revised the manuscript. All authors reviewed or revised the manuscript and approved the final draft for submission. The authors declare no conflicts of interest.

Sources of Support: This research was supported by the National Natural Science Foundation of China (82473708, 82388102, 82273714, 82192900, 81390540, 91846303, 81941018); and the Excellent Youth Foundation of Jiangsu Province (BK20220100). The CKB baseline survey and the first re-survey were supported by the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up has been supported by Wellcome grants to Oxford University (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z) and grants (2016YFC0900500) from the National Key R&D Program of China, and Chinese Ministry of Science and Technology (2011BAI09B01). We thank all the study participants and research staff for their contributions and commitment to the present study.

Running Title: Fine Particulate Matter Increased Lung Cancer Risk in Nonsmokers

Descriptor Number: 6.01 Air Pollution: Epidemiology; 6.06 Gene-Environment Interaction; 6.12 Lung Cancer Epidemiology: Occupational and Environmental Factors; 9.28 Lung Cancer: Epidemiology

Word Count: 3485

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org

At a Glance Commentary

Scientific Knowledge on the Subject: Ambient fine particulate matter ($PM_{2.5}$) and genetic factors are established risk factors for lung cancer. However, the association between long-term $PM_{2.5}$ exposure and lung cancer in non-smokers (LCINS) remains inconsistent. Most previous studies have been conducted in settings with relatively low $PM_{2.5}$ exposure. Given China's high $PM_{2.5}$ levels, potential genetic differences, and significant LCINS burden, there is a critical need for high-quality evidence on the relationship between $PM_{2.5}$ exposure, genetic risk, and LCINS in China.

What This Study Adds to the Field: Based on the China Kadoorie Biobank (CKB) data, we confirmed a significant association between long-term $PM_{2.5}$ exposure and LCINS, identifying a nonlinear concentration-response function. We demonstrated a significant additive interaction, with the highest LCINS risk and LCINS-related mortality risk observed among individuals with high genetic risk (top 30%) and high $PM_{2.5}$ ($\geq$50.9 $\mu g/m^3$) exposure. Additionally, this study highlighted the mediating effect of chronic obstructive pulmonary disease (COPD) between $PM_{2.5}$ exposure and LCINS at high $PM_{2.5}$ levels. We also estimated the exposure-response relationship between long-term $PM_{2.5}$ exposure and the population-attributable fraction (PAF) to emphasize the public health importance of improving air quality in China.

Abstract

Rationale: The association between fine particulate matter ($PM_{2.5}$) and lung cancer incidence in non-smokers (LCINS) remains inconsistent.

Objectives: To investigate the association between long-term $PM_{2.5}$ exposure and LCINS in a Chinese population and to assess the modifying effect of genetic factors.

Methods: Time-dependent Cox proportional hazard models were employed to evaluate the hazard ratios (HR) and 95% confidence interval (CI) of $PM_{2.5}$ with LCINS risk and LCINS-related mortality. The polygenic risk score (PRS) was constructed to further explore the interactions between genetic risk and $PM_{2.5}$ exposure. Additionally, the population attributable fraction (PAF) of $PM_{2.5}$ to lung cancer risk and mortality was calculated.

Measurements and Main Results: The results demonstrated significant associations between $PM_{2.5}$ exposure and LCINS incidence (HR: 1.10, 95% confidence interval [CI]: 1.04-1.17, per 10 $\mu g/m^3$) and mortality (HR: 1.17, 95% CI: 1.08-1.27, per 10 $\mu g/m^3$). Compared to the lowest-risk group, individuals exposed to the high $PM_{2.5}$ level ($\geq$50.9 $\mu g/m^3$) and high genetic risk (top 30%) exhibited the highest LCINS incidence (HR: 2.01, 95% CI: 1.39-2.87) and mortality (HR: 2.30, 95% CI: 1.38-3.82). A significant additive interaction between $PM_{2.5}$ and genetic risk on LCINS incidence was observed. Approximately 33.6% of LCINS cases and 48.5% of LCINS-related deaths in China could be prevented if $PM_{2.5}$ concentrations were reduced to meet WHO guidelines.

Conclusion: Long-term exposure to outdoor $PM_{2.5}$ increases LCINS risk and LCINS-related mortality, especially in populations with high genetic risk. Strengthening air pollution control

measures in China has the potential to significantly reduce the burden of LCINS.

Abstract word count: 248

Keywords: air pollution; genetic susceptibility; lung cancer in non-smokers; additive interaction; population attributable fraction

Introduction

Lung cancer (LC) is the most common cancer and the leading cause of cancer-related death worldwide (1). Although tobacco smoking remains the main cause of LC, LC in individuals who have never smoked (LCINS) is on the rise and is estimated to be the fifth most common cause of cancer-related deaths worldwide in 2023 (2). The proportion of LCINS varies across countries, at <20% in men and women in the USA and the UK, while >40% in men and >85% in women in China (3-5). Therefore, it urgently needs to provide reliable evidence regarding the cause of LCINS among Chinese nonsmokers.

In 2013, the International Agency for Research on Cancer (IARC) classified ambient air pollution, especially fine particulate matter ($PM_{2.5}$), as a known human carcinogen (6). China has experienced increasing air pollution during the last two decades, with an estimated annual mean $PM_{2.5}$ concentration of 54.3 μg/m³ in 2013, compared to 10.7 μg/m³ in the USA for the same year (7). Previous cohort studies have reported positive associations between ambient $PM_{2.5}$ exposure and LCINS-related mortality (8, 9), but the results about LCINS incidence remain controversial (10, 11). Moreover, evidence on the risk of LC in relation to ambient $PM_{2.5}$ was mainly performed in low-exposure settings, but limited studies were conducted in highly polluted regions (8-11). Although Li and colleagues attempted to evaluate the chronic effects of high ambient $PM_{2.5}$ exposure (31.17 to 96.96 μg/m³) on LC risk based on the China-PAR (Prediction for Atherosclerotic Cardiovascular Disease Risk in China) project, it did not focus on LCINS and failed to adjust household air pollution and other confounders, which might not reflect the actual effects of ambient $PM_{2.5}$ (12). Therefore, more studies are needed to clarify

the $PM_{2.5}$-LC response pattern in high-exposure settings among nonsmokers in China.

Genome-wide association studies (GWASs) have identified more than 60 susceptibility loci of LC in different ethnicities and subsequently developed polygenic risk scores (PRSs) that were proved to be effective at measuring individuals' genetic risk of LC (13, 14). It has been widely accepted that environmental factors can modify the genetic risk, and significant gene-environment interaction has been observed for ambient $PM_{2.5}$ at low exposure settings among European populations (10). However, evidence has shown ethnic differences in genetic factors of lung cancer between European and Chinese populations (15), which highlighted the need for further studies to explore the gene-environment interactions of high $PM_{2.5}$ exposure in Chinese populations.

There is evidence for the specific linkage process between $PM_{2.5}$, chronic obstructive pulmonary disease (COPD), and LC, which showed a mediation proportion of 17.56% at low $PM_{2.5}$ settings (16). However, it remains unclear whether the proportion is universal in settings with high $PM_{2.5}$ exposure. Hence, based on the China Kadoorie Biobank (CKB), we aimed to investigate the main effects and gene-environment interactions of high ambient $PM_{2.5}$ on LCINS incidence and mortality in Chinese populations; in addition, we also intended to explore the mediation by COPD at high $PM_{2.5}$ exposure on LCINS risk and estimate the population attributable fraction (PAF) of $PM_{2.5}$ exposure for nonsmokers in China.

Methods

Study population

The CKB study, which has been detailed previously (17), served as the data source for this study. In brief, the CKB recruited over 500,000 participants aged 30-79 between 2004 and 2008 from ten regions across China, including five urban areas and five rural areas. Baseline data was collected by trained interviewers using laptop-based questionnaires designed with quality control measures. Participants also underwent physical measurements and provided blood samples. Written informed consent was obtained from all participants before data collection. The CKB study received approval from the Ethical Review 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, United Kingdom).

Figure 1 outlines the participant selection process and workflow for this study. We excluded participants who reported a history of cancer at baseline (n=2,954). Since the $PM_{2.5}$ concentrations in Henan and Haikou are significantly different from other regions, to eliminate spurious correlations caused by this difference, we excluded these two regions from the main analysis (**Figure E1**). To focus on nonsmokers, we excluded participants who had ever smoked (n=161,862). Finally, 255,829 participants with complete data were included in the primary analysis, and 49,587 participants with genetic data were included in the genetic risk analysis.

Assessment of ambient $PM_{2.5}$ exposure

A satellite-based model was employed to estimate daily $PM_{2.5}$ concentrations at a 1×1 km spatial resolution from 2005 to 2018. The details of this model are documented in previous studies (18-20). Briefly, the model used the random forest machine learning algorithm and Multi-Angle Implementation of Atmospheric Correction aerosol optical depth values from

NASA to predict $PM_{2.5}$ concentrations. It was trained with daily $PM_{2.5}$ concentration data from over 1,300 ground monitors between 2013 and 2017, incorporating meteorological parameters, land use, population density, and Modern-Era Retrospective Analysis for Research and Applications (MERRA) $PM_{2.5}$ products. Historical $PM_{2.5}$ concentrations from 2005 to 2012 were retrieved using this model. External validation was performed by comparing the $PM_{2.5}$ measurements not included in model development (21). The results showed good consistency between the measurements and predictions, with an R^2 of 0.70 and an RMSE of 20.9 $\mu g/m^3$, indicating the model's effectiveness in capturing temporal variations in historical data.

The annual $PM_{2.5}$ concentrations for each participant were determined based on the location of the nearest clinic where they were recruited. In the primary analysis, $PM_{2.5}$ was treated as a time-varying variable on a 1-year scale to account for temporal variations during the study period.

PRS calculation

Genotyping, quality control, and imputation for the participants in the CKB have been reported previously (22). To assess individuals' genetic susceptibility to LC, we generated a PRS for LC using 19 genetic variants (PRS-19) according to our previous study (14). The construction of PRS-19 was performed in accordance with the tutorial on PRS analyses (23) and is detailed in the **online data supplement**. Detailed information on the 19 SNPs for PRS calculation is provided in **Table E1**. Participants were categorized into low (bottom 30%), intermediate (30%-70%), and high (top 30%) genetic risk groups based on the distribution of PRS among the non-cases.

Ascertainment of outcome

LC events in the CKB were identified through linkage to the national health insurance database and local disease and death registries. Annual active follow-up was conducted for participants not linked to the national health insurance database to collect their disease and vital status information. As of December 31, 2018, less than 1% of participants were lost to follow-up due to moving out of their recruitment areas. All diseases in the CKB were coded using the International Classification of Diseases, 10th version (ICD-10), with LC coded as C33-C34 (24). LC deaths were defined as individuals who died from LC. Participants were followed from the enrollment date or January 1, 2005, whichever was later, until the date of LC diagnosis, LC death, or censoring. The censoring date was defined as the time of non-lung cancer death, loss to follow-up, or the end of the follow-up period, whichever occurred first.

Covariates

Potential confounders were identified based on existing literature and analyzed using directed acyclic graphs (**Figure E2**) (25). We adjusted for a series of potential confounders from the baseline questionnaires, including sociodemographic characteristics, lifestyle factors, passive smoking, household air pollution, and self-rated health. Sociodemographic characteristics included age, sex, education, household income, and living areas. Lifestyle factors encompassed red meat intake and fresh fruit consumption. Additionally, we accounted for passive smoking, measured as hours of second-hand smoke exposure per week, and household air pollution, which included fuel types for cooking and years of exposure to solid fuels. BMI was calculated as weight (kg) divided by the square of standing height (m). The

prevalent COPD was defined as individuals with hospital admission records of COPD prior to the baseline assessment, self-reported COPD, or those with FEV1/FVC<0.70 at the baseline assessment. Detailed definitions and categorizations of covariates are provided in the **online data supplement**.

Statistical analyses

Cox proportional hazard models were employed to calculate the hazard ratio (HR) and 95% confidence interval (CI) for long-term $PM_{2.5}$ exposure and LCINS incidence and LCINS-related mortality. Annual average $PM_{2.5}$ concentrations were included as a time-varying variable (on a yearly scale), with follow-up time as the underlying time scale. Schoenfeld residuals were used to test the proportional hazards assumption, and no violations were observed. Restricted cubic spline analysis assessed the potential nonlinear relationship between $PM_{2.5}$ concentrations and LCINS. Subgroup analyses were performed according to age, sex, BMI, family cancer history, living area (urban/rural), cooking fuel types, heating fuel types, second-hand smoke exposure hours per week, and years of exposure to solid fuels. The heterogeneity of associations among groups was tested using Cochran's Q test.

The additive interaction between $PM_{2.5}$ and genetic risk was evaluated by relative excess risk due to interaction (RERI) and attributable proportion due to interaction (AP) (26). The 95% CIs for RERI and AP were generated by bootstrapping 3000 times, with values not including 0 indicating the presence of an additive interaction (27). To evaluate the role of COPD in the pathogenesis of LCINS due to $PM_{2.5}$ exposure, mediation analyses were performed. Prevalent COPD at baseline was considered as a mediator, and mediation proportions were calculated by

comparing estimates from models with and without the hypothesized mediator. Additionally, we generated the concentration-response curve of $PM_{2.5}$ to the PAF (**online data supplement**), considering two guideline target thresholds from China (35 μg/m³) and the World Health Organization (10 μg/m³).

Several sensitivity analyses were conducted to validate the robustness of our results. First, we extended the exposure time window by using 3-year or 5-year average $PM_{2.5}$ concentrations as the time-varying variable. Second, we used the average concentration of $PM_{2.5}$ from 2005 to 2018 as the time-unvarying variable. Third, we excluded participants with incident LC or deaths due to LC in the first year or first three years. Fourth, we reincluded participants from Henan and Haikou in our association analyses. Statistical significance was set at $P < 0.05$ with a 2-sided test. All statistical analyses were performed using R, version 4.3.1 (R Foundation for Statistical Computing).

Results

The mean $PM_{2.5}$ concentration in the CKB from 2005 to 2018 was 51.63 ± 10.55 μg/m³. **Figure E1** illustrates the distribution of average $PM_{2.5}$ concentration and annual trends from 2005 to 2018 across the 10 regions, showing a slight increase from 2005 to 2013, followed by a decrease after 2013. **Table 1** presents the characteristics of participants, categorized by median annual average $PM_{2.5}$ concentration.

Over a follow-up of 3,041,403 person-years (with a median follow-up of 12.17 years), 2,571 cases of LCINS and 1,416 LCINS-related deaths were documented. After adjusting for

potential confounders, a nonlinear relationship was found between long-term $PM_{2.5}$ exposure and both LCINS incidence and LCINS-related mortality (*P* for a nonlinear trend < 0.05) (**Figure 2**). $PM_{2.5}$ concentration was significantly associated with LCINS incidence (HR: 1.10, 95% CI: 1.04-1.17, per 10 μg/m³). Compared to the lowest quartile of $PM_{2.5}$ (30.2-42.3 μg/m³) exposure, the HRs for the second (42.3-50.4 μg/m³), third (50.4-54.1 μg/m³), and highest (54.2-63.1 μg/m³) quartiles were 1.50 (95% CI: 1.21-1.85), 1.80 (95% CI: 1.48-2.20), and 1.91 (95% CI: 1.55-2.36), respectively (*P* for trend < 0.01, **Figure 2**). Similarly, a positive association and nonlinear concentration-response relationship were observed for LCINS-related mortality. Higher $PM_{2.5}$ exposure increased the risk of LCINS-related death (HR: 1.17, 95% CI: 1.08-1.27, per 10 μg/m³, **Figure 2**). Stratified analysis indicated that the effects of long-term $PM_{2.5}$ exposure on LCINS incidence were more significant among participants living in rural areas, and those using solid fuels for cooking or heating (**Figure E3**). Interestingly, we also observed a reduction in the 5-year absolute risk for participants in the CKB during the 2013-2018 period compared to 2004-2013, which aligns with the reduction in $PM_{2.5}$ levels after 2013 (**Figure E4**).

Sensitivity analyses confirmed the robustness of our findings (**Tables E2-E4**). Expanding the $PM_{2.5}$ exposure window to 3-year or 5-year periods as time-varying variables or using the average $PM_{2.5}$ concentrations from 2005 to 2018 as time-invariant variables, resulted in increased estimated associations between $PM_{2.5}$ exposure and both LCINS incidence and mortality. Excluding LC cases diagnosed within the first year or first three years of follow-up did not significantly alter the results. Additionally, re-including participants from Henan and Haikou, who were initially excluded, did not change our findings.

In the CKB study, we observed a significant association between $PM_{2.5}$ exposure and an

increased risk of COPD at baseline (**Table E5**). The prevalent COPD was significantly associated with LCINS incidence and LCINS-related mortality **(Table E6)**. Additionally, COPD was found to mediate the relationship between $PM_{2.5}$ and LCINS incidence and LCINS-related mortality, with a mediation proportion of 22.8% (95% CI: 8.2%-34.7%) for LCINS incidence and 18.5% (95% CI: 0.6%-34.1%) for LCINS-related mortality (**Figure E5**).

Table E1 lists the SNPs used to construct the PRS-19. Individuals who developed LCINS had higher PRS compared to those who did not (**Figure E6**). We found a significant joint effect of $PM_{2.5}$ concentration and genetic risk on LCINS incidence, showing a dose-response relationship (**Figure 3**). Participants with the high PRS (top 30%) and high $PM_{2.5}$ ($\geq$50.9 $\mu g/m^3$) exposure had the greatest risk of LCINS compared to those with the low PRS (bottom 30%) and low $PM_{2.5}$ (<50.9 $\mu g/m^3$) exposure (HR: 2.01, 95% CI: 1.39-2.87). Similarly, the highest risk of LCINS-related death was observed in individuals with high PRS and high $PM_{2.5}$ exposure (HR: 2.30, 95% CI: 1.38-3.82). Positive synergistic effects were observed in the additional interaction scale between $PM_{2.5}$ and PRS on LCINS risk (**Figure 3**). Specifically, participants with high levels of $PM_{2.5}$ and high PRS had a relative excess risk of 0.64 (95% CI: 0.21-1.22) due to additive interaction, accounting for 32% (95% CI: 11%-60%) of the LCINS incidence. Although the confidence intervals were not significant, similar RERI and AP were found in LCINS mortality analyses. Similar additive interactions were noted when participants were reclassified using the tertiles of $PM_{2.5}$ concentration **(Figure E7)**.

The concentration-response relationship between $PM_{2.5}$ and the PAF is illustrated in **Figure 4**. For individuals exposed to a median annual average $PM_{2.5}$ concentration of 52.3 $\mu g/m^3$, the PAF was 12.1% for LCINS incidence and 16.1% for LCINS mortality if we take the

35 μg/m³ as the reference. Furthermore, with the WHO target guideline of 10 μg/m³, the preventable proportion rises to 33.6% for LCINS incidence and 48.5% for LCINS mortality. This indicates that about one-third of LCINS and nearly half of LCINS-related deaths could be prevented if $PM_{2.5}$ concentration in China were reduced to meet WHO guidelines.

Discussion

In this large prospective cohort study of Chinese adults, we confirmed that prolonged $PM_{2.5}$ exposure is associated with an increased risk of LCINS incidence and mortality. Notably, we identified a nonlinear concentration response between $PM_{2.5}$ and LCINS, validated the synergistic effect of $PM_{2.5}$ exposure and genetic predisposition on LCINS, and revealed COPD partially mediates the adverse effects of $PM_{2.5}$ on LCINS in a high-exposure setting. Lastly, we presented the PAF curves to illustrate the impact of $PM_{2.5}$ on LCINS in China, offering a clear perspective on its public health implications.

Previous studies have shown a significant association between $PM_{2.5}$ exposure and LCINS-related death (8, 28, 29), but results regarding LCINS incidence have been inconsistent. Results from the Nurses' Health Study showed that a 10-μg/m³ increase of $PM_{2.5}$ was associated with a 37% increased risk of LCINS (HR, 1.37; 95% CI, 1.06-1.77) (30). However, other studies from North America (31, 32) and Europe (10, 11) did not support the significant associations. Given the much higher $PM_{2.5}$ concentrations and relatively high incidence of LCINS in China, the associations and magnitudes of the effects still need to be evaluated (15). Li and colleagues found that each 10 μg/m³ increase in $PM_{2.5}$ exposure was linked to a 19% higher risk of LC

incidence and 20% higher risk of LC mortality (12). Our study comprehensively estimated the associations of $PM_{2.5}$ exposure with LCINS using a larger sample size (250,000 *vs.* 85,000) and adjusting more comprehensive confounding factors. We found slightly lighter effects of $PM_{2.5}$ for LCINS incidence and LCINS mortality. These findings offer robust evidence of the association between $PM_{2.5}$ exposure and LCINS in China, which may be applicable to other low- and middle-income countries experiencing severe air pollution.

Understanding the exposure-response relationship is essential for estimating the LC burden caused by $PM_{2.5}$ and for developing effective public health strategies. However, inconsistent exposure-response curves have been reported. In low-exposure settings, nonlinear exposure-response curves for LC mortality have been observed, with steep increases in risk at $PM_{2.5}$ concentrations between 0-20 μg/m³ and flattening slopes at higher concentrations (33, 34). However, Li and colleagues highlighted nonlinear exposure-response associations for LC incidence and mortality, showing steeper slopes at higher $PM_{2.5}$ levels in the range of 31.17-96.96 μg/m³ (12). Our findings revealed nonlinear exposure-response relationships for LCINS incidence and mortality, showing steeper slopes between 30.2-52.3 μg/m³ and flattening slopes at higher concentrations. Our result is consistent with the findings from the ChinaHEART project, which demonstrated a nonlinear exposure-response relationship between $PM_{2.5}$ and LC mortality with an inflection point at 50.4 μg/m³ (35). Our results suggest that a significant reduction in the LCINS burden could be achieved with improvements in air quality at the current stage in China.

Subgroup analysis indicated greater effects of $PM_{2.5}$ on LCINS among participants living in rural areas and those using solid fuels for cooking or heating, two groups that largely overlap

277 in China. Solid fuel combustion can emit $PM_{2.5}$ with abundant organic compounds and heavy
278 metals, especially under poor combustion conditions (36). Previous evidence indicated that
279 household solid fuel use for heating and cooking was a major source of $PM_{2.5}$ in China,
280 contributing 34%-47% of $PM_{2.5}$ between 1990 and 2005 (37). The greater effects might be
281 explained by the synergistic impact of indoor $PM_{2.5}$ derived from solid fuels and ambient $PM_{2.5}$
282 exposure on LCINS risk in rural areas of China. Moreover, the differences in $PM_{2.5}$ sources and
283 components between urban and rural areas might also related to the heterogeneity (38).

284 Chronic inflammation associated with COPD can lead to repeated damage to airway
285 epithelial cells and increased cell turnover rates, which results in DNA errors and amplifies the
286 carcinogenic effects of $PM_{2.5}$ (39). We evaluated the mediation effects of COPD on the
287 relationship between $PM_{2.5}$ and LCINS incidence and found that COPD mediated 17.5%-18.7%
288 of the effect in a low-exposure setting (16). In a high-exposure setting, we further validated this
289 mediation effect and observed that approximately 22.8% of LCINS incidence and 18.5% of
290 LCINS mortality were mediated by COPD. These findings suggest that actively preventing and
291 treating COPD could partially alleviate the link between $PM_{2.5}$ and LC across varying levels of
292 air pollution. However, considering the similar mediation proportions of COPD at both high
293 and low $PM_{2.5}$ settings, we can speculate that $PM_{2.5}$ exerts multiple carcinogenic effects
294 independently of COPD, such as the induction of DNA damage (40), activation of tumor
295 suppressor genes and inactivation of oncogenes (41, 42), inhibition of apoptosis (43), and
296 modulation of the lung inflammatory microenvironment (44).

297 Our previous study demonstrated an additive interaction between high genetic risk and
298 $PM_{2.5}$ exposure on LC incidence in a low-exposure setting among Europeans (RERI: 0.37, 95%

CI: 0.24-0.49; AP: 21%, 95% CI: 14%-29%) (10). In this study, we further validated the additive interaction and observed even greater effects (RERI: 0.64, 95% CI: 0.21-1.22; AP: 32%, 95% CI: 11%-60%) in a high-exposure setting among Chinese adults. These results indicate that individuals with high genetic risk should be especially cautious about air pollution exposure. While genetic factors are unavoidable, our study highlights the critical public health importance of reducing $PM_{2.5}$ exposure to lessen the burden of LC incidence and mortality. Specifically, our findings suggest that up to 33.6% of LCINS incidence and 48.5% of LCNIS mortality could be prevented if $PM_{2.5}$ levels in China were reduced to meet the WHO guidelines.

Our study has several limitations. First, participants' $PM_{2.5}$ exposure was assigned based on the nearest clinic, potentially causing misclassification. Second, we did not account for address changes during the study, but less than 1% of participants moved out of the baseline city. Third, we did not consider the composition of $PM_{2.5}$ due to a lack of information. Fourth, categorizing participants may introduce residual confounding and lead to potentially unclear results, particularly given the existing measurement errors. Finally, more than 90% of the non-smokers included in this study were women, which may not adequately represent the overall non-smoking population in China.

In conclusion, this large prospective cohort study of Chinese adults demonstrates that long-term exposure to $PM_{2.5}$ significantly increases the risk of lung cancer incidence and mortality among non-smokers. Additionally, it shows that $PM_{2.5}$ exposure can amplify the impact of genetic susceptibility in high-exposure environments. Our findings expand the current understanding of the harmful effects of $PM_{2.5}$ on lung cancer in non-smoking populations. As China continues to strengthen air pollution control measures, a significant reduction in lung

cancer incidence and mortality can be anticipated.

Acknowledgments

We thank all the study participants and research staff for their contributions and commitment to the present study.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2024; 74: 229-263.

2. Lam DC, Liam CK, Andarini S, Park S, Tan DSW, Singh N, Jang SH, Vardhanabhuti V, Ramos AB, Nakayama T, Nhung NV, Ashizawa K, Chang YC, Tscheikuna J, Van CC, Chan WY, Lai YH, Yang PC. Lung Cancer Screening in Asia: An Expert Consensus Report. *J Thorac Oncol* 2023; 18: 1303-1322.

3. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, McCullough ML, Patel AV, Ma J, Soerjomataram I, Flanders WD, Brawley OW, Gapstur SM, Jemal A. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J Clin* 2018; 68: 31-54.

4. Brown KF, Rumgay H, Dunlop C, Ryan M, Quartly F, Cox A, Deas A, Elliss-Brookes L, Gavin A, Hounsome L, Huws D, Ormiston-Smith N, Shelton J, White C, Parkin DM. The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015. *Br J Cancer* 2018; 118: 1130-1141.

5. Islami F, Chen W, Yu XQ, Lortet-Tieulent J, Zheng R, Flanders WD, Xia C, Thun MJ, Gapstur SM, Ezzati M, Jemal A. Cancer deaths and cases attributable to lifestyle factors and infections in China, 2013. *Ann Oncol* 2017; 28: 2567-2574.

6. Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K. The carcinogenicity of outdoor air pollution. *Lancet Oncol* 2013; 14: 1262-1263.

7. Brauer M, Freedman G, Frostad J, van Donkelaar A, Martin RV, Dentener F, van Dingenen R, Estep K, Amini H, Apte JS, Balakrishnan K, Barregard L, Broday D, Feigin V, Ghosh S, Hopke PK, Knibbs LD, Kokubo Y, Liu Y, Ma S, Morawska L, Sangrador JL, Shaddick G, Anderson HR, Vos T, Forouzanfar MH, Burnett RT, Cohen A. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. *Environ Sci Technol* 2016; 50: 79-88.

8. Turner MC, Krewski D, Pope CA, 3rd, Chen Y, Gapstur SM, Thun MJ. Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers. *Am J Respir Crit Care Med* 2011; 184: 1374-1381.

9. Jerrett M, Burnett RT, Beckerman BS, Turner MC, Krewski D, Thurston G, Martin RV, van Donkelaar A, Hughes E, Shi Y, Gapstur SM, Thun MJ, Pope CA, 3rd. Spatial analysis of air pollution and mortality in California. *Am J Respir Crit Care Med* 2013; 188: 593-599.

10. Huang Y, Zhu M, Ji M, Fan J, Xie J, Wei X, Jiang X, Xu J, Chen L, Yin R, Wang Y, Dai J, Jin G, Xu L, Hu Z, Ma H, Shen H. Air Pollution, Genetic Factors, and the Risk of Lung Cancer: A Prospective Study in the UK Biobank. *Am J Respir Crit Care Med* 2021; 204: 817-825.

11. Raaschou-Nielsen O, Andersen ZJ, Beelen R, Samoli E, Stafoggia M, Weinmayr G, Hoffmann B, Fischer P, Nieuwenhuijsen MJ, Brunekreef B, Xun WW, Katsouyanni K, Dimakopoulou K, Sommar J, Forsberg B, Modig L, Oudin A, Oftedal B, Schwarze PE, Nafstad P, De Faire U, Pedersen NL, Ostenson CG, Fratiglioni L, Penell J, Korek M, Pershagen G, Eriksen KT, Sørensen M, Tjønneland A, Ellermann T, Eeftens M, Peeters

PH, Meliefste K, Wang M, Bueno-de-Mesquita B, Key TJ, de Hoogh K, Concin H, Nagel G, Vilier A, Grioni S, Krogh V, Tsai MY, Ricceri F, Sacerdote C, Galassi C, Migliore E, Ranzi A, Cesaroni G, Badaloni C, Forastiere F, Tamayo I, Amiano P, Dorronsoro M, Trichopoulou A, Bamia C, Vineis P, Hoek G. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). *Lancet Oncol* 2013; 14: 813-822.

12. Li J, Lu X, Liu F, Liang F, Huang K, Yang X, Xiao Q, Chen J, Liu X, Cao J, Chen S, Shen C, Yu L, Lu F, Wu X, Zhao L, Wu X, Li Y, Hu D, Huang J, Zhu M, Liu Y, Shen H, Gu D. Chronic Effects of High Fine Particulate Matter Exposure on Lung Cancer in China. *Am J Respir Crit Care Med* 2020; 202: 1551-1559.

13. Hung RJ, Warkentin MT, Brhane Y, Chatterjee N, Christiani DC, Landi MT, Caporaso NE, Liu G, Johansson M, Albanes D, Marchand LL, Tardon A, Rennert G, Bojesen SE, Chen C, Field JK, Kiemeney LA, Lazarus P, Zienolddiny S, Lam S, Andrew AS, Arnold SM, Aldrich MC, Bickeböller H, Risch A, Schabath MB, McKay JD, Brennan P, Amos CI. Assessing Lung Cancer Absolute Risk Trajectory Based on a Polygenic Risk Model. *Cancer Res* 2021; 81: 1607-1615.

14. Dai J, Lv J, Zhu M, Wang Y, Qin N, Ma H, He YQ, Zhang R, Tan W, Fan J, Wang T, Zheng H, Sun Q, Wang L, Huang M, Ge Z, Yu C, Guo Y, Wang TM, Wang J, Xu L, Wu W, Chen L, Bian Z, Walters R, Millwood IY, Li XZ, Wang X, Hung RJ, Christiani DC, Chen H, Wang M, Wang C, Jiang Y, Chen K, Chen Z, Jin G, Wu T, Lin D, Hu Z, Amos CI, Wu C, Wei Q, Jia WH, Li L, Shen H. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations.

393 *Lancet Respir Med* 2019; 7: 881-891.
394 15. Zhu M, Lv J, Huang Y, Ma H, Li N, Wei X, Ji M, Ma Z, Song C, Wang C, Dai J, Tan F, Guo
395 Y, Walters R, Millwood IY, Hung RJ, Christiani DC, Yu C, Jin G, Chen Z, Wei Q, Amos
396 CI, Hu Z, Li L, Shen H. Ethnic differences of genetic risk and smoking in lung cancer:
397 two prospective cohort studies. *Int J Epidemiol* 2023; 52: 1815-1825.
398 16. Zhu M, Huang Y, Ma H, Shen H. Reply to Ward et al.: Air Pollution as a Risk Factor for
399 Lung Cancer: Potential Mechanisms. *Am J Respir Crit Care Med* 2022; 205: 367-368.
400 17. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, Li L. China Kadoorie Biobank of 0.5
401 million people: survey methods, baseline characteristics and long-term follow-up. *Int J*
402 *Epidemiol* 2011; 40: 1652-1666.
403 18. Liu Z, Xiao Q, Li R. Full Coverage Hourly PM(2.5) Concentrations' Estimation Using
404 Himawari-8 and MERRA-2 AODs in China. *Int J Environ Res Public Health* 2023; 20.
405 19. Xiao Q, Chen H, Strickland MJ, Kan H, Chang HH, Klein M, Yang C, Meng X, Liu Y.
406 Associations between birth outcomes and maternal PM(2.5) exposure in Shanghai: A
407 comparison of three exposure assessment approaches. *Environ Int* 2018; 117: 226-236.
408 20. Sun D, Liu C, Zhu Y, Yu C, Guo Y, Sun D, Pang Y, Pei P, Du H, Yang L, Chen Y, Meng X,
409 Liu Y, Zhang J, Schmidt D, Avery D, Chen J, Chen Z, Lv J, Kan H, Li L. Long-Term
410 Exposure to Fine Particulate Matter and Incidence of Esophageal Cancer: A Prospective
411 Study of 0.5 Million Chinese Adults. *Gastroenterology* 2023; 165: 61-70.e65.
412 21. Shi S, Wang W, Li X, Hang Y, Lei J, Kan H, Meng X. Optimizing modeling windows to
413 better capture the long-term variation of PM(2.5) concentrations in China during 2005-
414 2019. *Sci Total Environ* 2023; 854: 158624.

415 22. Walters RG, Millwood IY, Lin K, Schmidt Valle D, McDonnell P, Hacker A, Avery D, Edris
416 A, Fry H, Cai N, Kretzschmar WW, Ansari MA, Lyons PA, Collins R, Donnelly P, Hill
417 M, Peto R, Shen H, Jin X, Nie C, Xu X, Guo Y, Yu C, Lv J, Clarke RJ, Li L, Chen Z.
418 Genotyping and population characteristics of the China Kadoorie Biobank. *Cell Genom*
419 2023; 3: 100361.
420 23. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score
421 analyses. *Nat Protoc* 2020; 15: 2759-2772.
422 24. ICD-10: there's a code for that. *Lancet* 2015; 386: 1420.
423 25. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research.
424 *Epidemiology* 1999; 10: 37-48.
425 26. Li R, Chambless L. Test for additive interaction in proportional hazards models. *Ann*
426 *Epidemiol* 2007; 17: 227-236.
427 27. Assmann SF, Hosmer DW, Lemeshow S, Mundt KA. Confidence intervals for measures of
428 interaction. *Epidemiology* 1996; 7: 286-290.
429 28. Turner MC, Andersen ZJ, Baccarelli A, Diver WR, Gapstur SM, Pope CA, 3rd, Prada D,
430 Samet J, Thurston G, Cohen A. Outdoor air pollution and cancer: An overview of the
431 current evidence and public health recommendations. *CA Cancer J Clin* 2020.
432 29. Hamra GB, Guha N, Cohen A, Laden F, Raaschou-Nielsen O, Samet JM, Vineis P,
433 Forastiere F, Saldiva P, Yorifuji T, Loomis D. Outdoor particulate matter exposure and
434 lung cancer: a systematic review and meta-analysis. *Environ Health Perspect* 2014; 122:
435 906-911.
436 30. Puett RC, Hart JE, Yanosky JD, Spiegelman D, Wang M, Fisher JA, Hong B, Laden F.

Particulate matter air pollution exposure, distance to road, and incident lung cancer in the nurses' health study cohort. *Environ Health Perspect* 2014; 122: 926-932.

31. Gharibvand L, Shavlik D, Ghamsary M, Beeson WL, Soret S, Knutsen R, Knutsen SF. The Association between Ambient Fine Particulate Air Pollution and Lung Cancer Incidence: Results from the AHSMOG-2 Study. *Environ Health Perspect* 2017; 125: 378-384.

32. Tomczak A, Miller AB, Weichenthal SA, To T, Wall C, van Donkelaar A, Martin RV, Crouse DL, Villeneuve PJ. Long-term exposure to fine particulate matter air pollution and the risk of lung cancer among participants of the Canadian National Breast Screening Study. *Int J Cancer* 2016; 139: 1958-1966.

33. Pinault LL, Weichenthal S, Crouse DL, Brauer M, Erickson A, Donkelaar AV, Martin RV, Hystad P, Chen H, Finès P, Brook JR, Tjepkema M, Burnett RT. Associations between fine particulate matter and mortality in the 2001 Canadian Census Health and Environment Cohort. *Environ Res* 2017; 159: 406-415.

34. Bowe B, Xie Y, Yan Y, Al-Aly Z. Burden of Cause-Specific Mortality Associated With PM2.5 Air Pollution in the United States. *JAMA Netw Open* 2019; 2: e1915834.

35. Li W, Tian A, Shi Y, Chen B, Ji R, Ge J, Su X, Pu B, Lei L, Ma R, Wang Q, Ban J, Song L, Xu W, Zhang Y, He W, Yang H, Li X, Li T, Li J. Associations of long-term fine particulate matter exposure with all-cause and cause-specific mortality: results from the ChinaHEART project. *Lancet Reg Health West Pac* 2023; 41: 100908.

36. Sun J, Yu J, Niu X, Zhang X, Zhou L, Liu X, Zhang B, He K, Niu X, Ho KF, Cao J, Shen Z. Solid fuel derived PM(2.5) induced oxidative stress and according cytotoxicity in A549 cells: The evidence and potential neutralization by green tea. *Environ Int* 2023;

171: 107674.

37. Chafe ZA, Brauer M, Klimont Z, Van Dingenen R, Mehta S, Rao S, Riahi K, Dentener F, Smith KR. Household cooking with solid fuels contributes to ambient PM2.5 air pollution and the burden of disease. *Environ Health Perspect* 2014; 122: 1314-1320.

38. Srivastava D, Xu J, Vu TV, Liu D, Li L, Fu P, Hou S, Moreno Palmerola N, Shi Z, Harrison RM. Insight into PM2.5 sources by applying positive matrix factorization (PMF) at urban and rural sites of Beijing. *Atmos Chem Phys* 2021; 21: 14703-14724.

39. Raviv S, Hawkins KA, DeCamp MM, Jr., Kalhan R. Lung cancer in chronic obstructive pulmonary disease: enhancing surgical options and outcomes. *Am J Respir Crit Care Med* 2011; 183: 1138-1146.

40. Quezada-Maldonado EM, Sánchez-Pérez Y, Chirino YI, García-Cuellar CM. Airborne particulate matter induces oxidative damage, DNA adduct formation and alterations in DNA repair pathways. *Environ Pollut* 2021; 287: 117313.

41. Li R, Zhou R, Zhang J. Function of PM2.5 in the pathogenesis of lung cancer and chronic airway inflammatory diseases. *Oncol Lett* 2018; 15: 7506-7514.

42. Holme JA, Vondráček J, Machala M, Lagadic-Gossmann D, Vogel CFA, Le Ferrec E, Sparfel L, Øvrevik J. Lung cancer associated with combustion particles and fine particulate matter (PM(2.5)) - The roles of polycyclic aromatic hydrocarbons (PAHs) and the aryl hydrocarbon receptor (AhR). *Biochem Pharmacol* 2023; 216: 115801.

43. Aghaei-Zarch SM, Nia AHS, Nouri M, Mousavinasab F, Najafi S, Bagheri-Mohammadi S, Aghaei-Zarch F, Toolabi A, Rasoulzadeh H, Ghanavi J, Moghadam MN, Talebi M. The impact of particulate matters on apoptosis in various organs: Mechanistic and

therapeutic perspectives. *Biomed Pharmacother* 2023; 165: 115054.

44. Hill W, Lim EL, Weeden CE, Lee C, Augustine M, Chen K, Kuan FC, Marongiu F, Evans EJ, Jr., Moore DA, Rodrigues FS, Pich O, Bakker B, Cha H, Myers R, van Maldegem F, Boumelha J, Veeriah S, Rowan A, Naceur-Lombardelli C, Karasaki T, Sivakumar M, De S, Caswell DR, Nagano A, Black JRM, Martínez-Ruiz C, Ryu MH, Huff RD, Li S, Favé MJ, Magness A, Suárez-Bonnet A, Priestnall SL, Lüchtenborg M, Lavelle K, Pethick J, Hardy S, McRonald FE, Lin MH, Troccoli CI, Ghosh M, Miller YE, Merrick DT, Keith RL, Al Bakir M, Bailey C, Hill MS, Saal LH, Chen Y, George AM, Abbosh C, Kanu N, Lee SH, McGranahan N, Berg CD, Sasieni P, Houlston R, Turnbull C, Lam S, Awadalla P, Grönroos E, Downward J, Jacks T, Carlsten C, Malanchi I, Hackshaw A, Litchfield K, DeGregori J, Jamal-Hanjani M, Swanton C. Lung adenocarcinoma promotion by air pollutants. *Nature* 2023; 616: 159-167.

Figure Legends

Figure 1. Participants selection and workflow.

Definition of abbreviations: CKB, China Kadoorie Biobank; $PM_{2.5}$, fine (≤2.5 μm) particulate matter; PRS, polygenic risk score; RCS, restricted cubic spline; BMI, body mass index; COPD, chronic obstructive pulmonary disease.

Figure 2. Association of LCINS with long-term $PM_{2.5}$ exposure.

Definition of abbreviations: HR, hazard ratio; CI, confidence interval; $PM_{2.5}$, fine (≤2.5 μm) particulate matter; Q1-Q4, increasing quartiles of $PM_{2.5}$ concentrations.

The concentration-response relationships between long-term $PM_{2.5}$ exposure and LCINS are shown on the left side of the figure, while the association analyses are depicted on the right side; the outcomes are (A) LCINS incidence and (B) LCINS-related mortality.

* Estimate: Adjusted for age, sex, BMI, education, household income, living areas, weekly red meat consumption, weekly fresh fruit intake, second-hand smoke exposure hours per week, fuel types used in cooking at the baseline residence, years exposed to solid fuels, and self-rated health degree.

† Adjusted for age, sex, household income, physical activities (MET), weekly red meat consumption, fuel types used in cooking at the baseline residence, fuel types used in heating at the baseline residence, and BMI.

Figure 3. Risk of LCINS and LCINS-related death according to air pollution and genetic categories.

Definition of abbreviations: $PM_{2.5}$, fine (≤2.5 μm) particulate matter; CI, confidence interval; RERI: relative excess risk due to the interaction; AP, attributable proportion due to the interaction.

The outcomes are (A) LCINS incidence and (B) LCINS-related mortality.

* Estimate: Adjusted for age, sex, BMI, education, household income, living areas, weekly red meat consumption, weekly fresh fruit intake, second-hand smoke exposure hours per week, fuel types used in cooking at the baseline residence, years exposed to solid fuels, and self-rated health degree.

† Adjusted for age, sex, household income, physical activities (MET), weekly red meat consumption, fuel types used in cooking at the baseline residence, fuel types used in heating at the baseline residence, and BMI.

‡ Defined by PRS: low (bottom 30%), intermediate (30%-70%), and high (top 30%).

§ Defined by the median average concentration of $PM_{2.5}$ in all monitors: low (<50.9 $\mu g/m^3$) and high (≥50.9 $\mu g/m^3$).

‖ To estimate RERI and AP, the low pollution category and the lowest genetic risk (low PRS) groups were the reference categories.

Figure 4. Concentration-response relationship between long-term $PM_{2.5}$ exposure and population attributable fraction of $PM_{2.5}$ to LCINS.

Definition of abbreviations: $PM_{2.5}$, fine (≤2.5 μm) particulate matter; PAF, population attributable fraction.

The outcomes are (A) LCINS incidence and (B) LCINS-related mortality.

Table 1: Baseline characteristics of participants by median of average $PM_{2.5}$ concentration from 2005 to 2018

		Annual average concentration of $PM_{2.5}$	
	Total (n=255829)	<52.3 μg/m^3 (n=128159)	≥52.3 μg/m^3 (n=127670)
Age, years, mean (SD)	51.4 (10.5)	51.1 (10.5)	51.7 (10.5)
Age, <60 years, n (%)	202266 (79.1)	102349 (79.9)	99917 (78.3)
Sex, female, n (%)	232188 (90.8)	116992 (91.3)	115196 (90.2)
BMI, kg/m^2, mean (SD)	23.7 (3.3)	23.3 (3.2)	24.2 (3.4)
BMI, n (%)			
Normal (<25 kg/m^2)	173044 (67.6)	93382 (72.9)	79662 (62.4)
Overweight (25-29.9 kg/m^2)	71582 (28.0)	30997 (24.2)	40585 (31.8)
Obesity (≥30 kg/m^2)	11203 (4.4)	3780 (2.9)	7423 (5.8)
Education, n (%)			
No formal school	62753 (24.5)	32769 (25.6)	29984 (23.5)
Primary school	77662 (30.4)	45200 (35.3)	32462 (25.4)
Middle school	64168 (25.1)	30388 (23.7)	33780 (26.5)
High school or above	51246 (20.0)	19802 (15.5)	31444 (24.6)
Household income, n (%)			
<10,000 RMB/y	68263 (26.7)	48728 (38.0)	19535 (15.3)
10,000-19,999 RMB/y	69140 (27.0)	37367 (29.2)	31773 (24.9)
>20,000 RMB/y	118426 (46.3)	42064 (32.8)	76362 (59.8)
Areas, urban, n (%)	122449 (47.9)	35846 (28.5)	86603 (66.6)
Daily consumer of red meat, n (%)	77661 (30.3)	36033 (28.7)	41628 (32.0)
Daily consumer of fresh fruit, n (%)	65475 (25.6)	21322 (17.0)	44153 (33.9)
Second-hand smoke exposure hours per week, mean (SD)	5.42 (8.2)	5.17 (7.6)	5.68 (8.7)
Cooking fuel types, n (%)			
Clean fuels	122104 (47.7)	39736 (31.0)	82368 (64.5)

Rarely or no cooking facility	26923 (10.5)	11901 (9.3)	15022 (11.8)
Solid fuels	106266 (41.5)	76095 (59.4)	30171 (23.6)
Other fuels type	536 (0.2)	427 (0.3)	109 (0.1)
Heating fuel types, n (%)			
Clean fuels	62906 (24.6)	15450 (12.3)	47456 (36.5)
Solid fuels	84952 (33.2)	61090 (48.6)	23862 (18.3)
Other fuels type	1229 (0.5)	1142 (0.9)	87 (0.1)
No heating in winter	106742 (41.7)	48037 (38.2)	58705 (45.1)
Years of exposure to solid fuels, mean (SD)	26.8 (19.3)	30.7 (19.3)	22.9 (18.4)
Self-rated good health, yes, n (%)	71751 (28.0)	34274 (26.7)	37477 (29.4)
Total physical activity, *MET-h/d*, mean (SD)	6.69 (3.3)	6.72 (3.2)	6.65 (3.5)

Definition of abbreviations: $PM_{2.5}$, fine (≤2.5 μm) Particulate Matter; BMI, body mass index; RMB, Chinese renminbi; MET, metabolic equivalent of task.