

Sex disparity of lung cancer risk in non-smokers: a multicenter population-based prospective study based on China National Lung Cancer Screening Program

Zheng Wu^{1,2}, Fengwei Tan³, Zhuoyu Yang^{1,2}, Fei Wang^{1,2}, Wei Cao^{1,2}, Chao Qin^{1,2}, Xuesi Dong^{1,2}, Yadi Zheng^{1,2}, Zilin Luo^{1,2}, Liang Zhao^{1,2}, Yiwen Yu^{1,2}, Yongjie Xu^{1,2}, Jiansong Ren^{1,2}, Jufang Shi^{1,2}, Hongda Chen¹, Jiang Li^{1,2}, Wei Tang⁴, Sipeng Shen^{5,6}, Ning Wu^{4,7}, Wanqing Chen¹, Ni Li^{1,2}, Jie He³

¹Office of Cancer Screening, National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China;

²Key Laboratory of Cancer Data Science, Chinese Academy of Medical Sciences, Beijing 100021, China;

³Department of Thoracic Surgery, National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China;

⁴PET-CT Center, National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China;

⁵Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu 211166, China;

⁶Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing, Jiangsu 211166, China;

⁷Department of Diagnostic Radiology, National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China.

Abstract

Background: Non-smokers account for a large proportion of lung cancer patients, especially in Asia, but the attention paid to them is limited compared with smokers. In non-smokers, males display a risk for lung cancer incidence distinct from the females—even after excluding the influence of smoking; but the knowledge regarding the factors causing the difference is sparse. Based on a large multicenter prospective cancer screening cohort in China, we aimed to elucidate the interpretable sex differences caused by known factors and provide clues for primary and secondary prevention.

Methods: Risk factors including demographic characteristics, lifestyle factors, family history of cancer, and baseline comorbidity were obtained from 796,283 Chinese non-smoking participants by the baseline risk assessment completed in 2013 to 2018. Cox regression analysis was performed to assess the sex difference in the risk of lung cancer, and the hazard ratios (HRs) that were adjusted for different known factors were calculated and compared to determine the proportion of excess risk and to explain the existing risk factors.

Results: With a median follow-up of 4.80 years, 3351 subjects who were diagnosed with lung cancer were selected in the analysis. The lung cancer risk of males was significantly higher than that of females; the HRs in all male non-smokers were 1.29 (95% confidence interval [CI]: 1.20–1.38) after adjusting for the age and 1.38 (95% CI: 1.28–1.50) after adjusting for all factors, which suggested that known factors could not explain the sex difference in the risk of lung cancer in non-smokers. Known factors were 7% (1.29–1.38/1.29) more harmful in women than in men. For adenocarcinoma, women showed excess risk higher than men, contrary to squamous cell carcinoma; after adjusting for all factors, 47% (1.30–1.16/[1.30–1]) and 4% (7.02–6.75/[7.02–1]) of the excess risk was explainable in adenocarcinoma and squamous cell carcinoma. The main causes of gender differences in lung cancer risk were lifestyle factors, baseline comorbidity, and family history.

Zheng Wu and Fengwei Tan contributed equally to this work.

Correspondence to: Dr. Ni Li, Office of Cancer Screening, National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College; Key Laboratory of Cancer Data Science, Chinese Academy of Medical Sciences, No. 17 Panjiayuananli, Chaoyang District, Beijing 100021, China

E-Mail: nli@cicams.ac.cn;

Dr. Jie He, National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 17 Panjiayuananli, Chaoyang District, Beijing 100021, China

E-Mail: hejie@cicams.ac.cn

Copyright © 2022 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2022;135(11)

Received: 08-12-2021; Online: 13-07-2022 Edited by: Peifang Wei

Access this article online

Quick Response Code:



Website:
www.cmj.org

DOI:
10.1097/CM9.0000000000002161

Conclusions: Significant gender differences in the risk of lung cancer were discovered in China non-smokers. Existing risk factors did not explain the excess lung cancer risk of all non-smoking men, and the internal causes for the excess risk still need to be explored; most known risk factors were more harmful to non-smoking women; further exploring the causes of the sex difference would help to improve the prevention and screening programs and protect the non-smoking males from lung cancers.

Keywords: Lung cancer; Non-smoker; Sex disparity; Excess risk; Risk factor

Background

As the leading cause of cancer mortalities worldwide, lung cancer accounts for 1.80 million deaths in 2020,^[1] pressuring people to comprehend its adverse consequence and pay attention to its prevention and control.

Although the incidence density is lower than that in smokers,^[2] a high proportion of lung cancers were seen in non-smokers.^[3] In Asia, women accounted for the vast majority of non-smoking people, and the number of non-smoking women with lung cancer was large^[4]; although non-smoking men only accounted for nearly one-third of all males in Asia,^[5,6] the incidence density of lung cancer in them was dozens of times of non-smoking males in the United States,^[7] with a definite disease burden of lung cancer.

Smoking was considered as the highest risk factor by the lung cancer screening guidelines,^[8-12] closely related to gender, and considered the main reason why higher risk of lung cancer was observed in men than in women^[13]; however, despite excluding the influence of smoking, enough attention was not given to the existing large sex differences in lung cancer in non-smokers. Recent studies reported that the incidence density of lung cancer in non-smoking men was two-fold of that in non-smoking women in parts of Asia.^[5] But the knowledge regarding the incidence of lung cancer difference in non-smoking males and females, especially in China, remained sparse.

Based on this scenario, this study aimed to clarify the impact proportion of the sex difference on the lung cancer risk, and identify the risk factors, aiming to improve the existing screening strategy, and strive for due health benefits for non-smokers, based on a large multicenter prospective cancer screening cohort.

Methods

Data source and study population

The study was based on the China National Lung Cancer Screening (NLCS) Program. A multicenter population-based cancer screening study aimed to investigate the five most common types of cancers in the urban areas since October 2012. The study was approved by the Ethics Committee of China National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College (No. 15-070/997). All participants provided written informed consent. Ethical approval was obtained for all the data collection.

This study used the data of the enrolled subjects from 2013 to 2018. The data from 458 communities in 12 cities, 8 provinces (Beijing, Zhejiang, Jiangsu, Anhui, Hunan,

Liaoning, Guangxi, and Henan Provinces) involved in the program were selected based on the following criteria: (1) complete cancer registration data, (2) complete vital statistics data (e.g., age and sex), and (3) minor migration representing a relatively stable population. These participants were ineligible: patients who have been diagnosed with tumor, or have other serious diseases (for example, severe organ dysfunction and mental illness) under treatment.

Data acquisition

Eligible participants who provided the written informed consent completed a baseline questionnaire about their exposure to risk factors and they were evaluated for their lung cancer risk using the NLCS protocol.^[14] Forms including the paper-based and computer-based documentation (epidemiological questionnaire, low-dose computed tomography report, follow-up information, and pathology report) were used in the data collection for the screening program. Each participant had an identification code for management and traceability.

The following factors were collected: age (continuous variable), the education level (low category was defined as the primary school or below education; medium category was defined as the intermediate junior school to high school education; the high class was defined as the college or above education), body mass index (BMI; categorized as: underweight, <18.5 kg/m²; normal, 18.5–23.9 kg/m²; overweight, 24.0–27.9 kg/m²; obese, ≥28.0 kg/m²), occupational exposure (yes or no), passive smoking years (categorized as: none, 0–19.9 years, 20.0–39.9 years, ≥40.0 years), frequent exercise (yes or no), air pollution exposure (yes or no), cooking oil fumes exposure (none, low, medium, or high; it is difficult to quantitatively measure the exposure in real-world screening programs, so it can only be based on participants' self-reported information), drinking (non-drinker, current drinker, or former drinker), tea drinking (non-drinker, current drinker, or former drinker), family history of lung cancer (yes or no), family history of any cancer (yes or no), number of relatives with lung cancer (0, 1, ≥2), chronic respiratory diseases (yes or no), gastrointestinal diseases (yes or no), hepatobiliary diseases (yes or no), hypertension (yes or no), hyperlipidemia (yes or no), and diabetes (yes or no). Occupational exposure was defined as exposure to asbestos, dust, or other factors harmful to respiratory system; frequent exercise was defined as >3 times a week, and >30 min for each time; air pollution and cooking oil fumes exposure was self-reported by participants; drinking was defined as at least one time in a week for >6 consecutive months; and tea drinking was defined as at least three times in a week for >6 consecutive months. Chronic respiratory and other diseases must be the definite diagnosis by regular medical institutions.

The information regarding outcomes, including time of diagnosis, clinical stage, and pathological type from January 2013 to June 2021 was obtained through the clinical follow-up and cross-verified in the local tumor registration system and death surveillance and almost no loss of follow-up occurred. The end event was defined as the diagnosis of lung cancer, which was coded as C34 in the International Classification of Diseases (the 10th revision). The follow-up person-year was defined as the time from enrollment to the diagnosis of lung cancer; for a participant who was not diagnosed or the date was missing, it was defined as the time from enrollment to earlier of the death date or the last follow-up date.

Quality control

All questionnaires were logically checked by field staff in the process of collecting; 2% of the questionnaires were randomly selected for review, and the consistency rate of each item after review shall not be <90%. Data were saved and analyzed using the National Cancer Prevention and Control Network at the National Cancer Center of China through a web-based management system. When entering the system, the data would be further logically verified, reviewed by professionals, and cross-validated in the cancer registration system and medical records from hospital information systems.

Statistics analysis

Contingency table analyses were used for the description of categorical data. For approximately normally distributed continuous data, Student's *t*-test was used. Standardized mean difference was used to further compare the differences in the distribution of risk factors. Cox regression analysis was performed, because our study was a dynamic cohort with censored data and different follow-up time. The hazard ratios (HRs) of sex were calculated for the assessment of the gender difference in risk for lung cancer. The parameters were categorized into several types: demographic characteristics including age, education level, and the BMI; lifestyle factors including occupational exposure, passive smoking years, frequent exercise, air pollution exposure, cooking oil fumes exposure, drinking, and tea drinking; family history, including family history of lung cancer, family history of any cancer, and the number of relatives with lung cancer; baseline comorbidities including chronic respiratory diseases, gastrointestinal diseases, hepatobiliary diseases, hypertension, hyperlipidemia, and diabetes. All analyses were completed in R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria).

The incidence density was obtained by dividing the number of cases by the follow-up time of the population. To evaluate the effects of risk factors on the gender difference in lung cancer risk, we compared the HRs of sex before and after adjusting the aiming factors. One or more variables were adjusted to determine their impact on gender-specific lung cancer risk. If decreasing HRs were observed after adjusting the selected variables, these variables could explain the excess risk. The proportion of the excess risk that can be interpreted by the selected factors was expressed

as $(HR_{\text{age-adjusted}} - HR_{\text{fully adjusted}}) / (HR_{\text{age-adjusted}} - 1) \cdot |(HR_{\text{age-adjusted}} - HR_{\text{fully adjusted}}) / (HR_{\text{age-adjusted}})|$ represented the different sensitivity of genders to this factor. Similar methods were used to assess the impact of the covariates in previous studies.^[15,16]

For example, the HR of sex (males *vs.* females) was 2.00 after adjusting age, which meant that men had twice the risk of lung cancer as women; and after adjusting age and another known factor, the HR became 1.50, and the risk of lung cancer in males was 50% higher than that in females, 50% less than age adjusted only. The reduced risk was explained by the adjusting of the known factor. At this time, the risk of men decreased by 25% relatively (2.00 to 1.50), indicating that the harm of this factor is 25% greater in men than in women.

Results

Population characteristics

With a median follow-up of 4.80 years, 1,016,740 participants received risk assessment, and among which 794,283 were non-smokers, including 247,901 males and 546,382 females, who were retained for further analyses; 3351 participants were diagnosed with lung cancer, and among them, 1247 were males and 2104 were females [Figure 1].

The crude incidence density in the males was higher than that in the females (105.03 *vs.* 78.71 cases per 100,000 person-years), and the adjusted incidence density was (102.02 *vs.* 76.38 cases per 100,000 person-years) after adjusting the age and the education level. In the analysis, 177 (14.2%) males and 53 (2.5%) females were diagnosed with squamous cell carcinoma; 422 (33.8%) males and 1198 (56.9%) females were diagnosed with adenocarcinoma [Table 1].

With the statistical differences observed, men and women showed diverse manifestations with the distribution of demographic characteristics and potential risk factors of lung cancer (all *P* values < 0.001). The males tend to have lower passive smoking years and fewer number of relatives with lung cancer compared with females. The rates of having a family history of lung cancer, history of occupational exposure, air pollution exposure, cooking oil fumes exposure, chronic respiratory diseases, hepatobiliary diseases, hypertension, hyperlipidemia, and diabetes were lower in males than in females. The rate of ever-drinking and tea drinking, the frequency of exercise, and the prevalence of gastrointestinal diseases were higher in the males [Table 2].

Sex disparity: The high lung cancer excess risk in male participants

Most known factors were not the causes for the excess risk in males, but more dangerous for the females. The HRs of sex (males *vs.* females) in all non-smokers were 1.29 (95% confidence interval [CI]: 1.20–1.38) after adjusting for the age and 1.38 (95% CI: 1.28–1.50) after overall adjustments. Adjusting for the demographic factors showed

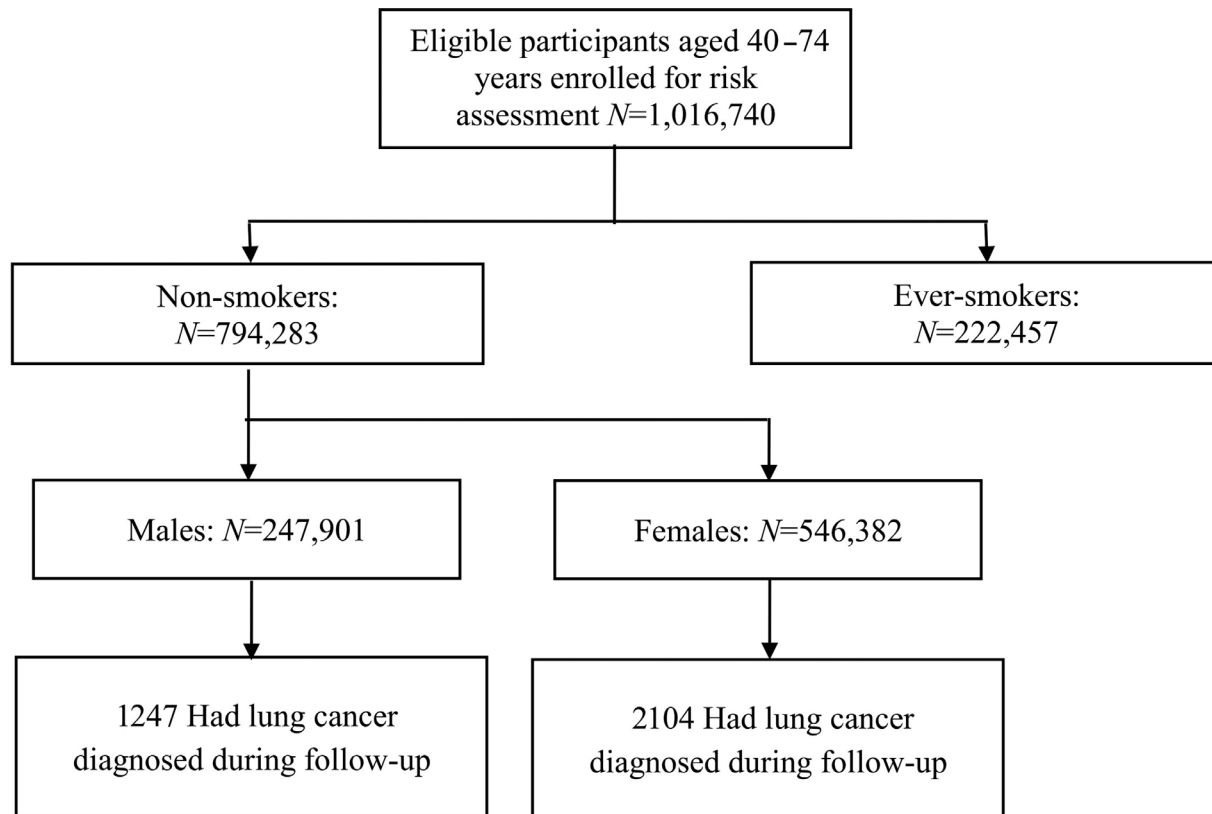


Figure 1: Flowchart of including and excluding participants for the study of sex disparity of lung cancer risk in non-smokers.

Table 1: Lung cancer incident cases in different sexes in non-smokers.

| Items | Male | Female |
|------------------------------------|------------|-------------|
| Overall | 247,901 | 546,382 |
| Incident cases | 1247 | 2104 |
| Incidence density (crude rate)* | 105.03 | 78.71 |
| Incidence density (adjusted rate)† | 102.02 | 76.38 |
| Histological type (%) | | |
| Adenocarcinoma | 422 (33.8) | 1198 (56.9) |
| Squamous cell carcinoma | 177 (14.2) | 53 (2.5) |
| Small-cell carcinoma | 45 (3.6) | 36 (1.7) |
| Others | 22 (1.8) | 26 (1.2) |

* Rate was the number of cases per 100,000 person-years. † Adjusted for age (continuous), and education.

limited explanation on the excess risk in the males (HR = 1.28, 95% CI: 1.19–1.37). Adjusting for the lifestyle factors (HR = 1.33, 95% CI: 1.23–1.43), the baseline comorbidity (HR = 1.37, 95% CI: 1.27–1.48), and the family history (HR = 1.37, 95% CI: 1.27–1.48) could not explain but increased the excess risk, which suggested that the known factors could not reduce the excess risk of lung cancer in males effectively but they were more closely related to the lung cancer in females who never smoked, accounting for relatively 7% (|1.29–1.38|/1.29) higher harm to women than men. The relative harm of the lifestyle factors (3%, |1.29–1.33|/1.29), baseline comorbidity (6%, |1.29–1.37|/1.29), and family history

(6%, |1.29–1.37|/1.29) was greater in women than in men [Figure 2A].

Similar results were obtained from the univariate analysis. Except for air pollution (HR = 1.28, 95% CI: 1.19–1.37), most of the known risk factors cannot explain the excess risk of men; a series of known risk factors represented by the family history of any cancer (HR = 1.37), family history of lung cancer (HR = 1.35), passive smoking (HR = 1.33), number of relatives with cancer (HR = 1.35), gastrointestinal diseases (HR = 1.32), hyperlipidemia (HR = 1.32), were relatively <6% (|1.29–1.37|/1.29) more harmful in women [Figure 3A].

High adenocarcinoma excess risk in female participants

Conversely, non-smoking women showed a significant excess risk compared with non-smoking men in terms of adenocarcinoma risk. For adenocarcinoma, the HRs of the sex (females vs. males) were 1.30 (95% CI: 1.17–1.46) after adjusting for the age, and 1.16 (95% CI: 1.02–1.31) after adjusting for all factors, which suggested that 47% (|1.30–1.16|/[1.30–1]) of excess risk can be explained by known factors. After adjusting for the demographic factors, the excess risk was increased in the females (HR = 1.33, 95% CI: 1.19–1.49). Adjusting for the lifestyle factors (HR = 1.27, 95% CI: 1.13–1.43), the baseline comorbidity (HR = 1.16, 95% CI: 1.03–1.31), and the family history of cancer (HR = 1.14, 95% CI: 1.01–1.28) explained the excess risk of lung cancer in males. Known risk factors caused about 11% (|1.30–1.16|/1.30) higher harm in women than in men [Figure 2B].

Table 2: Baseline characteristics of the participants receiving lung cancer screening.

| Items | Male (N = 247,901) | Female (N = 546,382) | P values | Statistics (t/χ ²) | SMD |
|--------------------------------------|--------------------|----------------------|----------|--------------------------------|-------|
| Demographic characteristics | | | | | |
| Age (years) | 55.98 ± 8.69 | 56.16 ± 9.35 | <0.001 | 8.19 | 0.012 |
| Education | | | <0.001 | 8561.65 | 0.227 |
| Low | 34,656 (13.98) | 115,573 (21.15) | | | |
| Medium | 162,436 (65.52) | 352,153 (64.45) | | | |
| High | 50,809 (20.50) | 78,656 (14.40) | | | |
| BMI (kg/m ²) | | | <0.001 | 4657.11 | 0.169 |
| <18.5 | 3169 (1.28) | 12,105 (2.22) | | | |
| 18.5–23.9 | 130,014 (52.45) | 299,843 (54.88) | | | |
| 24.0–27.9 | 100,531 (40.55) | 187,473 (34.31) | | | |
| ≥28.0 | 13,866 (5.59) | 46,261 (8.47) | | | |
| NA | 321 (0.13) | 700 (0.13) | | | |
| Lifestyle factors | | | | | |
| Occupational exposure | | | <0.001 | 9546.99 | 0.251 |
| No | 229,474 (92.57) | 462,451 (84.64) | | | |
| Yes | 18,427 (7.43) | 83,931 (15.36) | | | |
| Passive smoking years (years) | | | <0.001 | 54,005.70 | 0.638 |
| None | 219,959 (88.73) | 351,944 (64.41) | | | |
| <20.0 | 13,768 (5.55) | 46,390 (8.49) | | | |
| 20.0–39.9 | 12,567 (5.07) | 123,066 (22.52) | | | |
| >40.0 | 1580 (0.64) | 24,433 (4.47) | | | |
| NA | 27 (0.01) | 549 (0.10) | | | |
| Frequent exercise | | | <0.001 | 47.98 | 0.017 |
| No | 131,920 (53.21) | 295,327 (54.05) | | | |
| Yes | 115,981 (46.79) | 251,055 (45.95) | | | |
| Air pollution exposure | | | <0.001 | 4008.54 | 0.152 |
| No | 84,533 (34.10) | 148,181 (27.12) | | | |
| Yes | 163,358 (65.90) | 398,163 (72.87) | | | |
| NA | 10 (0.00) | 38 (0.01) | | | |
| Cooking oil fumes exposure | | | <0.001 | 15,262.33 | 0.318 |
| None | 45,635 (18.41) | 67,775 (12.40) | | | |
| Low | 187,181 (75.51) | 398,367 (72.91) | | | |
| Medium | 13,536 (5.46) | 66,218 (12.12) | | | |
| High | 1538 (0.62) | 13,958 (2.55) | | | |
| NA | 11 (0.00) | 64 (0.01) | | | |
| Drinking | | | <0.001 | 18,232.11 | 0.309 |
| Non-drinker | 201,631 (81.34) | 501,363 (91.76) | | | |
| Current drinker | 39,953 (16.12) | 38,541 (7.05) | | | |
| Former drinker | 6317 (2.55) | 6478 (1.19) | | | |
| Tea drinking | | | <0.001 | 2949.73 | 0.131 |
| Non-drinker | 155,871 (62.88) | 368,976 (67.53) | | | |
| Current drinker | 81,279 (32.79) | 147,597 (27.01) | | | |
| Former drinker | 10,751 (4.33) | 29,809 (5.46) | | | |
| Family history of cancer | | | | | |
| Family history of lung cancer | | | <0.001 | 14,359.26 | 0.321 |
| No | 240,321 (96.94) | 485,043 (88.77) | | | |
| Yes | 7580 (3.06) | 61,339 (11.23) | | | |
| Family history of any cancer | | | <0.001 | 32,501.97 | 0.477 |
| No | 222,585 (89.79) | 390,488 (71.47) | | | |
| Yes | 25,310 (10.21) | 155,878 (28.53) | | | |
| NA | 6 (0.00) | 16 (0.00) | | | |
| Number of relatives with lung cancer | | | <0.001 | 14,379.81 | 0.322 |
| 0 | 240,321 (96.94) | 485,043 (88.77) | | | |
| 1 | 6945 (2.80) | 54,699 (10.01) | | | |
| ≥2 | 635 (0.26) | 6640 (1.22) | | | |
| Baseline comorbidity | | | | | |
| Chronic respiratory diseases | | | <0.001 | 16,881.45 | 0.340 |
| No | 229,850 (92.72) | 445,189 (81.48) | | | |

Table 2
(continued).

| Items | Male (N=247,901) | Female (N=546,382) | P values | Statistics (t/χ ²) | SMD |
|---------------------------|------------------|--------------------|----------|--------------------------------|-------|
| Yes | 18,051 (7.28) | 101,192 (18.52) | | | |
| NA | 0 (0.00) | 1 (0.00) | | | |
| Gastrointestinal diseases | | | <0.001 | 21,395.86 | 0.380 |
| No | 219,730 (88.64) | 404,986 (74.12) | | | |
| Yes | 28,171 (11.36) | 141,396 (25.88) | | | |
| Hepatobiliary diseases | | | <0.001 | 19,974.01 | 0.365 |
| No | 218,664 (88.21) | 405,171 (74.16) | | | |
| Yes | 29,237 (11.79) | 141,209 (25.84) | | | |
| NA | 0 (0.00) | 1 (0.00) | | | |
| Hypertension | | | <0.001 | 5031.30 | 0.201 |
| No | 190,187 (76.72) | 397,590 (72.77) | | | |
| Yes | 28,767 (11.60) | 100,179 (18.33) | | | |
| NA | 28,947 (11.68) | 48,613 (8.90) | | | |
| Hyperlipidemia | | | <0.001 | 7795.90 | 0.245 |
| No | 200,647 (80.94) | 417,303 (76.38) | | | |
| Yes | 18,297 (7.38) | 80,442 (14.72) | | | |
| NA | 28,957 (11.68) | 48,637 (8.90) | | | |
| Diabetes | | | <0.001 | 1859.92 | 0.143 |
| No | 209,463 (84.49) | 462,952 (84.73) | | | |
| Yes | 9477 (3.82) | 34,801 (6.37) | | | |
| NA | 28,961 (11.68) | 48,629 (8.90) | | | |

Data are presented as mean ± SD or n (%). BMI: Body mass index; NA: Not available; SD: Standard difference; SMD: Standard mean difference.

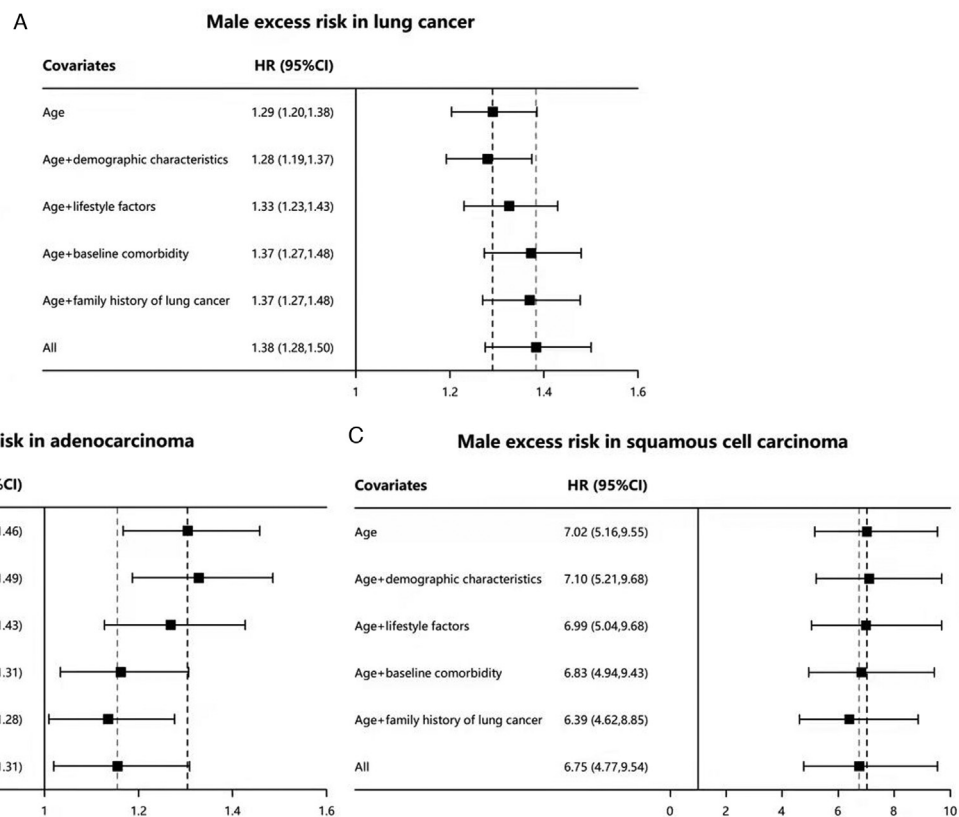


Figure 2: Lung cancer excess risk explanation in multivariate-adjusted analysis. CI: Confidence interval; HR: Hazard ratio.

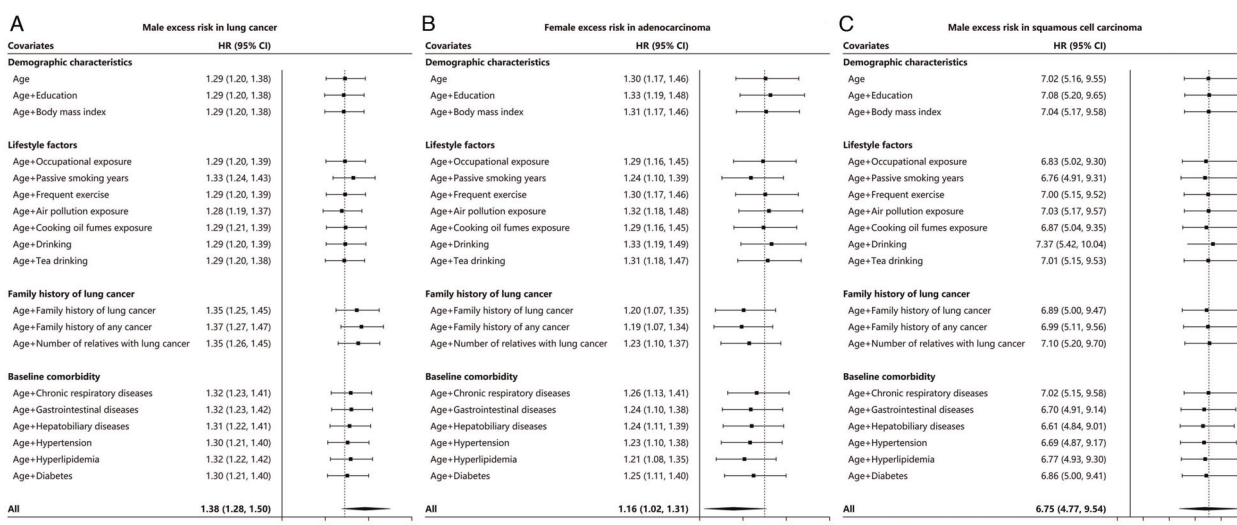


Figure 3: Lung cancer excess risk explanation in univariate adjusted analysis. CI: Confidence interval; HR: Hazard ratios.

In the univariate analysis, we further confirmed that passive smoking, family history, and baseline comorbidity were the main source of additional risk for women. Known factors like the family history of any cancer (HR = 1.19), family history of lung cancer (HR = 1.20), passive smoking (HR = 1.24), number of relatives with cancer (HR = 1.23), gastrointestinal diseases (HR = 1.24), hyperlipidemia (HR = 1.21), could explain about up to 37% ([1.30–1.19]/[1.30–1]) of the excess risk in females, with up to 8% ([1.30–1.19]/1.30) higher relative harm to females than males [Figure 3B].

High squamous cell carcinoma excess risk in male participants

The male excess risk was relatively higher in squamous cell carcinoma than in adenocarcinoma, characterized by the age-adjusted HR (7.02, 95% CI: 5.16–9.55). Adjusting for the demographic factors (HR = 7.10, 95% CI: 5.21–9.68), the lifestyle factors (HR = 6.99, 95% CI: 5.04–9.68), and the baseline comorbidity (HR = 6.83, 95% CI: 4.94–9.43) caused little influence between the different sexes; only about 3% maximum excess risk could be explained by anyone of them; and they were not the main reason for excess risk. The family history of lung cancer (HR = 6.39, 95% CI: 4.82–8.85) could elucidate 10% ([7.02–6.39]/[7.02–1]) of the excess risk; however, a large proportion of excess risk was not recognized. The all factors-adjusted HR (6.75, 95% CI: 4.77–9.54) indicated the existence of unknown risk factors; only 4% ([7.02–6.75]/[7.02–1]) of the risk could be explained [Figure 2C].

In the univariate analysis, we found that most of the known risk factors of squamous cell carcinoma were less harmful in women. Known factors like hepatobiliary diseases (HR = 6.61), hypertension (HR = 6.69), gastrointestinal diseases (HR = 6.70), and passive smoking (HR = 6.76) were up to 6% ([7.02–6.61]/7.02) more harmful among males than females. However, drinking (HR = 7.37) was 5% ([7.02–7.37]/7.02) more dangerous in females [Figure 3C].

Discussion

The epidemiological risk factors and lung cancer risk for the non-smokers in China were described in detail in this study; and the excess lung cancer risk in males was described for the first time based on a large sample-size prospective screening program. The male non-smokers had a higher lung cancer risk than that in the female non-smokers, even after adjusting the existing risk factors.

In the further exploration of the causes of the male excess risk, we found that existing risk factors are more dangerous for women, and our understanding of the excess risk in the non-smoking men was still insufficient. Air pollution may explain the excess risk in males to a certain extent; however, no such effect was observed with the other covariates. The known factors including lifestyle factors, family history of lung cancer, and baseline comorbidity could not explain the excess risk in non-smoking males.

In the subgroup of adenocarcinoma, a statistically significant excess risk in the non-smoking women was observed instead, and half of the excess risk could be elucidated by factors including lifestyle factors, family history of lung cancer, and baseline comorbidity. Our research showed that currently recognized risk factors help to explain the risk of lung cancer, especially adenocarcinoma, in women and benefited them. However, this discrepancy between the HRs of all-type lung cancer and adenocarcinoma suggested that in other subtypes of lung cancer, the excess risk in men remained significant.

In the subgroup of squamous cell carcinoma, few factors were significantly related to the excess risk. Previous elucidations proved that the adenocarcinoma was caused by multiple factors including air pollution and cigarette smoking, and the squamous cell carcinoma was only associated with smoking.^[17] However, after excluding the influence of smoking, the incidence rate of squamous cell carcinoma was still significantly higher in the males than the females, and the discrepancy was more significant than

that of the adenocarcinoma in our analysis. This result suggested that the factors other than smoking were strongly associated with squamous cell carcinoma and resulted in the gender differences of the lung cancer risk, contrasting to our previous understanding.

Based on the characteristics of the Chinese population, further exploration of the risk factors and reducing the excess risk were critical to better formulating the screening strategies and protecting the health of the non-smokers. In the Western population, smoking was a serious threat to public health with strong associations with cancers^[18,19]; further, the risk of non-smoking men was quite low and not higher than that of women.^[3] Consequently, most of the existing lung cancer prediction models and screening guidelines from Europe and America reported the risk factors mainly based on smoking in men.^[20] However, in Asia, due to the difference in population characteristics, lung cancer cases in the non-smokers were seen at a high level and higher than that in America.^[6] Similarly, the same principles for lung cancer screening may not be efficient in the Asian population, especially in non-smokers. The internal causes for the sex differences still need to be explored in the Asian population to protect the non-smokers.

Some studies provided the right directions for us to explore the risk in the non-smoking cohort. Molecular genetics in the area of lung cancer may find associations for the non-smoking population^[21]; a whole gene sequencing analysis in Europe found that there were differences in the genotypes among non-smokers and smokers^[22]; in contrast, some studies proved that after adjusting the imaging parameters, the risk difference between the sexes in an Asian population was no longer significant.^[23-25] Therefore, more studies about biomarkers and imaging diagnosis might be needed to understand the sex difference of non-smokers.

Our research also had some limitations. Our data were cross-validated with different medical registration systems; in this process, a large number of data of the specific pathological classification of lung cancer cannot be cross-referenced and were deleted. The missing data made our research only mirror the real situation to a certain extent in particular subtypes of lung cancers. Meanwhile, our study was in the short term, which leads to a limited number of lung cancer cases. Our study did not analyze the causes and mechanisms of the sex difference from the perspective of genetic susceptibility. Finally, we failed to give specific variables and screening strategies to effectively reduce the men's excess risk. Despite the above shortcomings, our research still reflects the current situation of the sex difference in an Asian population and warrants higher requirements for the current screening criteria.

Conclusions

Our study revealed the current situation of lung cancer in the non-smoking population in China, discovered the important excess risk in non-smokers, and drew attention to it. Improving the screening strategy and favoring the non-smoking population would help to further control the

risk of lung cancers and reduce the disease burden in the non-smoking population.

Funding

This study was supported by the grants from the National Key Research and Development Program of China, Non-profit Central Research Institute Fund of China (No. 2018YFC1315000); National Natural Science Foundation of China (No. 8187102812); and Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (Nos. 2020PT330001, 2019PT320027, 2019PT320023, 2018RC320010, and 3332019005).

Conflicts of interest

None.

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209–249. doi: 10.3322/caac.21660.
- Barta JA, Powell CA, Wisnivesky JP. Global epidemiology of lung cancer. *Ann Glob Health* 2019;85:8. doi: 10.5334/aogh.2419.
- Kerpel-Fronius A, Tammemägi M, Cavic M, Henschke C, Jiang L, Kazerooni E, *et al.* Screening for lung cancer in individuals who never smoked: an international association for the study of lung cancer early detection and screening committee report. *J Thorac Oncol* 2022;17:56–66. doi: 10.1016/j.jtho.2021.07.031.
- Corrales L, Rosell R, Cardona AF, Martín C, Zatarain-Barrón ZL, Arrieta O. Lung cancer in never smokers: the role of different risk factors other than tobacco smoking. *Crit Rev Oncol Hematol* 2020;148:102895. doi: 10.1016/j.critrevonc.2020.102895.
- Park B, Kim Y, Lee J, Lee N, Jang SH. Sex difference and smoking effect of lung cancer incidence in Asian population. *Cancers (Basel)* 2020;13:113. doi: 10.3390/cancers13010113.
- Yang D, Liu Y, Bai C, Wang X, Powell CA. Epidemiology of lung cancer and lung cancer screening programs in China and the United States. *Cancer Lett* 2020;468:82–87. doi: 10.1016/j.canlet.2019.10.009.
- Wakelee HA, Chang ET, Gomez SL, Keegan TH, Feskanich D, Clarke CA, *et al.* Lung cancer incidence in never smokers. *J Clin Oncol* 2007;25:472–478. doi: 10.1200/JCO.2006.07.2983.
- US Preventive Services Task Force, Krist AH, Davidson KW, Mangione CM, Barry MJ, Cabana M, *et al.* Screening for lung cancer: US preventive services task force recommendation statement. *JAMA* 2021;325:962–970. doi: 10.1001/jama.2021.1117.
- Oudkerk M, Devaraj A, Vliegenthart R, Henzler T, Prosch H, Heussel CP, *et al.* European position statement on lung cancer screening. *Lancet Oncol* 2017;18:e754–e766. doi: 10.1016/S1470-2045(17)30861-6.
- National Comprehensive Cancer Network. Lung Cancer Screening, Version 1.2021. Available from: <https://www.nccn.org/>. [Accessed on February 27, 2021].
- Expert Panel on Thoracic Imaging; Donnelly EF, Kazerooni EA, Lee E, Henry TS, Boiselle PM, *et al.* ACR appropriateness criteria lung cancer screening. *J Am Coll Radiol* 2018;15:S341–S346. doi: 10.1016/j.jacr.2018.09.025.
- Mazzone PJ, Silvestri GA, Souter LH, Caverly TJ, Kanne JP, Katki HA, *et al.* Screening for lung cancer: CHEST guideline and expert panel report. *Chest* 2021;160:e427–e494. doi: 10.1016/j.chest.2021.06.063.
- GBD 2019 Respiratory Tract Cancers Collaborators. Global, regional, and national burden of respiratory tract cancers and associated risk factors from 1990 to 2019: a systematic analysis for the global burden of disease study. *Lancet Respir Med* 2021;9:1030–1049. doi: 10.1016/S2213-2600(21)00164-8.
- Li N, Tan F, Chen W, Dai M, Wang F, Shen S, *et al.* One-off low-dose CT for lung cancer screening in China: a multicentre,

- population-based, prospective cohort study. *Lancet Respir Med* 2022;10:378–391. doi: 10.1016/S2213-2600(21)00560-9.
15. Wang X, O'Connell K, Jeon J, Song M, Hunter D, Hoffmeister M, *et al.* Combined effect of modifiable and non-modifiable risk factors for colorectal cancer risk in a pooled analysis of 11 population-based studies. *BMJ Open Gastroenterol* 2019;6:e000339. doi: 10.1136/bmjgast.
 16. Niedermaier T, Heisser T, Gies A, Guo F, Amitay EL, Hoffmeister M, *et al.* To what extent is male excess risk of advanced colorectal neoplasms explained by known risk factors? Results from a large German screening population. *Int J Cancer* 2021;149:1877–1886. doi: 10.1002/ijc.33742.
 17. Pesch B, Kendzia B, Gustavsson P, Jöckel KH, Johnen G, Pohlabein H, *et al.* Cigarette smoking and lung cancer - relative risk estimates for the major histological types from a pooled analysis of case-control studies. *Int J Cancer* 2012;131:1210–1219. doi: 10.1002/ijc.27339.
 18. Cokkinides V, Bandi P, McMahon C, Jemal A, Glynn T, Ward E. Tobacco control in the United States - recent progress and opportunities. *CA Cancer J Clin* 2009;59:352–365. doi: 10.3322/caac.20037.
 19. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin* 2021;71:7–33. doi: 10.3322/caac.21654.
 20. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, *et al.* Variations in lung cancer risk among smokers. *J Natl Cancer Inst* 2003;95:470–478. doi: 10.1093/jnci/95.6.470.
 21. Subramanian J, Govindan R. Molecular genetics of lung cancer in people who have never smoked. *Lancet Oncol* 2008;9:676–682. doi: 10.1016/S1470-2045(08)70174-8.
 22. Zhang T, Joubert P, Ansari-Pour N, Zhao W, Hoang PH, Lokanga R, *et al.* Genomic and evolutionary classification of lung cancer in never smokers. *Nat Genet* 2021;53:1348–1359. doi: 10.1038/s41588-021-00920-0.
 23. Chen XB, Yan RY, Zhao K, Zhang DF, Li YJ, Wu L, *et al.* Nomogram for the prediction of malignancy in small (8-20 mm) indeterminate solid solitary pulmonary nodules in Chinese populations. *Cancer Manag Res* 2019;11:9439–9448. doi: 10.2147/CMAR.S225739.
 24. Dong J, Sun N, Li J, Liu Z, Zhang B, Chen Z, *et al.* Development and validation of clinical diagnostic models for the probability of malignancy in solitary pulmonary nodules. *Thorac Cancer* 2014;5:162–168. doi: 10.1111/1759-7714.12077.
 25. Yang L, Zhang Q, Bai L, Li TY, He C, Ma QL, *et al.* Assessment of the cancer risk factors of solitary pulmonary nodules. *Oncotarget* 2017;8:29318–29327. doi: 10.18632/oncotarget.16426.
-
- How to cite this article:** Wu Z, Tan F, Yang Z, Wang F, Cao W, Qin C, Dong X, Zheng Y, Luo Z, Zhao L, Yu Y, Xu Y, Ren J, Shi J, Chen H, Li J, Tang W, Shen S, Wu N, Chen W, Li N, He J. Sex disparity of lung cancer risk in non-smokers: a multicenter population-based prospective study based on China National Lung Cancer Screening Program. *Chin Med J* 2022;135:1331–1339. doi: 10.1097/CM9.00000000000002161