Original research

Association between ultra-processed food consumption and lung cancer risk: a population-based cohort study

Kanran Wang ,[1,2] Junhan Zhao,[3] Dingyi Yang,[1,2] Mao Sun,[1,2] Wei Zhou,[1,2] Yongzhong Wu[1,2]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/thorax-2024-222100).

[1]Radiation Oncology Center, Chongqing Cancer Hospital, Chongqing, China
[2]Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing, China
[3]Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA

Correspondence to
Professor Yongzhong Wu; yongzhong.wu@cqu.edu.cn

Received 21 June 2024
Accepted 3 January 2025

ABSTRACT

Background The evidence on associations between ultra-processed foods (UPF) and lung cancer risk is limited and inconsistent.

Research question Are UPF associated with an increased risk of lung cancer, non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) ?

Methods Data of participants in this study were collected from the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Dietary intakes were assessed through a validated diet history questionnaire. These foods were categorised using the NOVA classification according to the degree of processing in the PLCO Cancer Screening Cohort. All cases of incident lung cancer were pathologically verified. Multivariable Cox regression was used to assess the association between consumption of UPF and lung cancer after adjustment for various potential confounders, including key risk factors related to lung cancer and overall diet quality.

Results A total of 1706 cases of lung cancer cases, including 1473 NSCLC and 233 SCLC, were identified during a mean follow-up of 12.2 years among 101 732 adults (mean age 62.5 years). After multivariable adjustments, individuals in the highest quarters for UPF consumption had a higher risk of lung cancer (HR=1.41, 95% CI 1.22 to 1.60), NSCLC (HR=1.37, 95% CI 1.20 to 1.58) and SCLC (HR=1.44, 95% CI 1.03 to 2.10) compared with those in the lowest quarter. These results remained statistically significant after a large range of subgroup and sensitivity analyses.

Conclusions Higher consumption of UPF is associated with an increased risk of lung cancer, NSCLC and SCLC. Although additional research in other populations and settings is warranted, these findings suggest the healthy benefits of limiting UPF.

WHAT IS ALREADY KNOW ON THIS TOPIC

⇒ The consumption of ultra-processed foods (UPF) has increased sharply during the past decades worldwide. Epidemiological studies have found associations between the consumption of UPF and a higher risk of all-cause mortality, all-cause cancer, irritable bowel disease, metabolic syndrome, obesity and hypertension.

⇒ The longitudinal effects of UPF on lung cancer and its subtypes remain unclear.

WHAT THIS STUDY ADDS

⇒ A higher consumption of UPF was independently associated with an increased hazard for lung cancer, non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) compared with those in the lowest quarter.

⇒ A non-linear dose–response pattern was shown for lung cancer and NSCLC, but not for SCLC.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE, OR POLICY

⇒ Limiting trends of UPF intake globally could contribute to reducing the burden of lung cancer.

INTRODUCTION

Lung cancer is the most frequent cancer and a leading cause of cancer-related death worldwide, with an estimated 2.20 million new cancer cases and 1.80 million deaths worldwide in 2020.[1] It is a major public health concern posing a substantial burden not only on patients but also on their families and national healthcare systems. The pathophysiological processes leading to lung cancer start many years before a clinically identifiable manifestation.[2] Consequently, prevention and early diagnosis are essential for increasing the survival from lung cancer. Although cigarette smoking has a dominant role in development of lung cancer, diet is widely recognised as a crucial driver of various chronic diseases.[3]

Ultra-processed foods (UPF) are formulations made by the food industry mostly from substances extracted from foods or obtained from the further processing of constituents of foods or through chemical synthesis with little or even no whole foods, and accounts for up to 60% of energy intake from foods and drinks, in other words constituting a staple of the diet.[4] These food products, such as processed meat, breads leavened without yeast, are convenient, hyper-palatable, energy-dense, highly profitable with low-cost ingredients and designed to replace all other food groups with attractive packaging and intensive marketing.[5] Processing can change the potential health of food by removing beneficial nutrients and naturally occurring bioactive components, introducing non-beneficial nutrients and food additives and modifying the physical structure.[6] Moreover, it might affect nutrient availability in the small intestine by altering the

To cite: Wang K, Zhao J, Yang D, *et al*. *Thorax* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/thorax-2024-222100

properties of the plant and animal cells in food. At the same time, it could also result in chronic low-titre inflammation, which may harm health.[7] Consumption of UPF has been associated with a higher risk of cerebrovascular diseases, dyslipidaemia, obesity, hypertension, diabetes, overall cancer and mortality in observational studies.[8–10] Previous cohort studies have shown that higher adherence to the Western dietary pattern characterised by high intakes of UPF consumption was associated with an increased risk of lung cancer.[11] In contrast, the European Prospective Investigation into Cancer and Nutrition (EPIC) Study showed that consumption of processed meat was not associated with the risk of lung cancer.[12] Meanwhile, to the best of our knowledge, limited studies have investigated the role of UPF in lung cancer and its subtypes in the community population.

Given the high burden of lung cancer and the growing consumption of UPF, a better characterisation of UPF and lung cancer risk will have a direct impact on informing cancer control and prevention through modifying health behaviours and developing tailored primary prevention. Hence, we aimed to evaluate whether UPF intake was associated with lung cancer risk using the data from both the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trials with longitudinal follow-up, and incorporated a detailed review of food intake and all medical records.

METHODS

Study population

The PLCO Cancer Screening Trial (NCT00002540, NCT01696968, NCT01696981 and NCT01696994) is a randomised, controlled trial of screening tests for prostate, lung, colorectal and ovarian cancers from 10 screening centres across the USA, and the design and methods of the PLCO trial have been reported in detail previously.[13] Approximately 155 000 participants aged 55–74 years were enrolled, randomised to the intervention (screening group) or control group between November 1993 and July 2001. Participants randomised to the intervention group received screening for prostate, lung, colorectal and ovarian cancers in the designated study years, whereas participants in the control group received usual care. Data were collected on cancer diagnoses through 2009 and mortality through 2018. This study complied with the Declaration of Helsinki and all participants provided written informed consent. The PLCO study protocol was approved by the National Cancer Institute, the National Institutes of Health Office of Protection from Research Risks (OH97-C-N041) and each of the 10 participating sites' institutional review boards. The current project and analysis using the PLCO data were approved by the National Cancer Institute Cancer Data Access System (PLCO-1547) and the Chongqing University Cancer Hospital Institutional Review Board (CZLS2024084-A).

For our research, participants who completed a baseline questionnaire (BQ) at study entry and a Diet History Questionnaire (DHQ) at study entry – for participants who were randomised after 1998 – or during the next follow-up after 1998 for those who were randomised before 1998, were included in the current analysis.[14] Furthermore, subjects who did not complete the BQ, had a personal history of lung cancer or any other cancer before completing the BQ, did not complete a valid DHQ (refer to dietary assessment section) or were not followed up after the enrollment were excluded. A flow chart is shown in figure 1 and details of population selection are shown in online supplemental eTable 1.

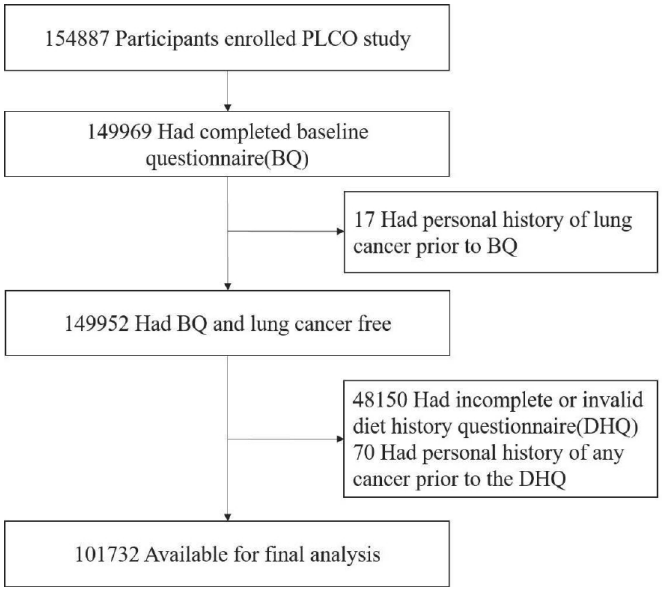

Figure 1 Flow chart of the study sample in the PLCO Study. PLCO, prostate, lung, colorectal and ovarian

Dietary assessment

Assessment of dietary variables was conducted at baseline using the DHQ to assess the frequency and portion size of food consumption and nutrient intake during the past year. The Healthy Eating Index-2015 (HEI-2015), a measure of diet quality, was calculated using the method described in the literature.[15] Food and nutrient intakes reported in the DHQ were validated against four 24-hour dietary recalls, indicating that the DHQ had good performance in estimating dietary intake (see supplement eMethod for timeline and details of the DHQ).[16]

Assessment of UPF consumption

All items of food and beverage of the Food Frequency Questionnaire food were categorised into one of the four mutually exclusive NOVA food groups by two trained dietitians: (1) unprocessed or minimally processed foods, (2) processed culinary ingredients, (3) processed foods and (4) UPF, which is a classification system according to the extent and purpose of the industrial processing they undergo.[17] A detailed description, including definition and example, for each group is available elsewhere.[18] The calculation of UPF and their energy is shown in the supplemental eMethod. We focused on UPF that include sour cream, cream cheese, ice cream, frozen yoghurt, fried foods, bread, cookies, cakes, pastries, salty snacks, breakfast cereals, instant noodles and soups, sauces, margarine, candy, soft drinks, artificially sweetened fruit drinks, restaurant/industrial hamburgers, hot dogs and pizza. Online supplemental eTable 2 shows the proportion of each item of the UPF. We adjusted UPF consumption for dietary energy intake using the residual method, which accounts for variations attributed to total energy intake.

Ascertainment of lung cancer

The endpoint was the incidence of lung cancer. Study participants self-reported lung cancer diagnoses through annual questionnaires. Patients with abnormal chest X-ray screening and cases notified by relatives were referred to their primary care provider for follow-up; those who had died were also noted. The diagnoses of lung cancer were all confirmed later through

medical record abstraction. The lung cancer histological subtypes derived from the International Classification of Diseases for Oncology, second edition (ICD-O-2) morphology, including non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). Of note, carcinoid lung cancer was not considered a target of lung cancer screening in the PLCO trial, thus, it was not confirmed as lung cancer in this study.

Assessment of other variables

Demographic, medical history and other risk factor information were extracted through the BQ. Hypertension was defined as systolic blood pressure >140 mm Hg, diastolic blood pressure >90 mm Hg, or the use of antihypertensive medications; diabetes mellitus was defined as fasting glucose >7 mmol/L or the use of antidiabetic therapy. Study arm, recruitment site, year of randomisation, sex, age at baseline, body mass index (BMI), smoking status, employment status, race, family history of any cancer, family history of lung cancer, physical activity, marital status and educational level (no high school degree, high school degree, some college, college degree or postgraduate) were assessed by the BQ.

Statistical analysis

Energy-adjusted UPF consumption was categorised into quarters and no data were missing for this variable of interest. Test of differences across consumption quarters was performed by one-way analysis of variance for continuous variables and X^2 test for categorical variables.

Follow-up was from the baseline examination to the time of the incident event or the date that the participant was last known to be lung cancer-free or dead, whichever came first. Multivariable Cox proportional hazard regression was used to estimate hazard ratios (HR) and 95% confidence intervals (CI) for the second to fourth quarters were estimated to assess the association between UPF consumption and lung cancer, with the lowest quarter as the reference category. To investigate linear trends across the quarters of energy-adjusted consumption of UPF, a median value was assigned to each category and considered the variable as being continuous in regression models. In multivariable adjustment, potential confounder selection was based on existing literature and deferred to statistical criteria. Based on these criteria, model 1 was adjusted for sex, age, race and family history of lung cancer; model 2 was further adjusted for prevalent hypertension, prevalent diabetes, smoking status, alcohol consumption, HEI-2015 score, employment status, marital status, physical activity status and BMI. The proportionality of hazards was verified using the Schoenfeld residuals method. Cluster robust standard errors were calculated to address the potential underestimation of standard errors for group-level variables.

In this study, the data were clustered across 10 screening centres. To assess how robust our results were to the potential unmeasured confounding, we calculated the E-value (https://www.evalue-calculator.com/), with an assumption of outcome prevalence of <15%. Kaplan-Meier curves were used with adjustment for confounding to describe lung cancer, NSCLC and SCLC risk according to energy-adjusted UPF consumption. Additionally, we conducted prespecified subgroup analyses to evaluate the robustness of our findings by age, sex, family history of lung cancer, BMI, smoking status and years of follow-up. Additionally, lunch meat, soft drinks and white bread/rolls were the main contributors to UPF intake; additional subgroup analyses were conducted, stratified by consumption levels. P for interaction was obtained through a likelihood ratio test, which compares the models with and without interaction terms.

Restricted cubic spline regression with four knots at the 5th, 35th, 65th and 95th percentiles was used to test the potential dose–response non-linear association between energy-adjusted consumption of UPF and lung cancer, with NSCLC and SCLC with 0 serving/day as the reference level. Tests for non-linearity used the likelihood ratio test to compare the model that comprised the linear term with the model that comprised both the linear and the cubic spline terms.

Analyses were performed using Stata statistical software, version 15 (Stata Corporation, College Station, Texas, USA) and SPSS 23.0 (SPSS Inc., Chicago, Illinois, USA). A two-sided p value <0.05 was considered statistically significant. All analyses were performed based on a predefined statistical analysis plan (available on request).

RESULTS

Characteristics of the study population

A total of 101 732 participants (50 187 men and 51 545 women) were included in the present study (figure 1). The mean age at baseline was 62.5 (SD 5.3) years and the mean (SD) energy-adjusted UPF consumption was 2.8 (2.4) servings/day, with lowest and highest intake quarters of 0.5 (0.4) and 6.0 (2.1) respectively. Online supplemental eTable 2 demonstrates the percentage of each food contributing to the total amount of UPF; the three types of foods with the highest proportion are lunch meat (11.1%), soft drinks (diet/caffeinated)(7.3%) and soft drinks (diet/decaffeinated)(6.6%). Table 1 and online supplemental eTable 3 show the baseline characteristics and recruitment sites of participants according to quarters of energy-adjusted total UPF consumption.

Lung cancer events

During a mean follow-up of 12.2 (IQR 10.5–13.6) years (1 213 533 person-years), a total of 1706 incident lung cancer events occurred, including 1473 (86.3%) cases of NSCLC and 233 (13.7%) of SCLC. Compared with participants consuming the least UPF, those with the highest intake had higher incidence rates of lung cancer (331/25 433 vs 495/25 434)(see table 2).

UPF and risk for lung cancer

When multivariable adjustments were made by Cox regression analysis, participants in the highest quarter of energy-adjusted UPF consumption had a 41% relatively higher hazard of lung cancer compared with those in the lowest quarter (HR=1.41, 95% CI 1.22 to 1.60), a 37% of NSCLC (HR=1.37, 95% CI 1.20 to 1.58) and a 44% of SCLC (HR=1.44, 95% CI 1.03 to 2.10)(see table 3).

Figure 2 shows the cumulative incidence curves for lung cancer, NSCLC and SCLC, stratified by groups with energy-adjusted UPF consumption over the median or not after adjusting for age and sex. When UPF consumption was expressed as a proportion of total energy intake, the initial results showed no significant changes (see online supplemental eTable 4).

Subgroup analyses were carried out by repeating the multivariable-adjusted Cox regression models in different scenarios comparing the highest with the lowest quarter of energy-adjusted UPF consumption. No significant interaction was found for predefined stratification factors (P for interaction >0.05)(see table 4 and online supplemental eTable 5).

In the whole study population, energy-adjusted UPF consumption was found to be associated with risks of lung cancer and

Table 1 Baseline characteristics of the study population according to energy-adjusted ultra-processed food consumption

Characteristic	Energy-adjusted ultra-processed foods consumption in each quarter (serving/day)				Full sample	P value
	Q1 (<1.0)	Q2 (1.0-<1.8)	Q3 (1.8-<3.7)	Q4 (≥3.7)		
No. of participants	25 433	25 433	25 432	25 434	101 732	
Energy-adjusted UPF consumption(serving/day)	0.5±0.4	1.5±0.3	3.0±0.5	6.0±2.1	2.8±2.4	<0.001
Intervention group, n (%)	13 227 (52.0)	13 043 (51.3)	12 962 (51.0)	12 958 (50.9)	52 190 (51.3)	0.058
Year of randomisation	1997.0±1.9	1997.0±1.9	1997.0±1.9	1997.0±1.9	1997.0±1.9	0.171
Age (years)	62.5±5.3	62.6±5.3	62.6±5.3	62.3±5.3	62.5±5.3	<0.001
Male, n (%)	12 385 (48.7)	12 488 (49.1)	12 227 (48.1)	13 087 (51.5)	50 187 (49.3)	<0.001
BMI (kg/m^2)	27.1±4.8	27.1±4.7	27.2±4.8	27.5±4.9	27.2±4.8	<0.001
Physical activity, n (%)						<0.001
Less than once a month	1882 (7.4)	1881 (7.4)	1958 (7.7)	2051 (8.0)	7772 (7.6)	
More than once a month	16 282 (64.0)	16 465 (64.7)	16 193 (63.7)	15 939 (62.7)	64 879 (63.8)	
Unknown	7269 (28.6)	7087 (27.9)	7281 (28.6)	7444 (29.3)	29 081 (28.6)	
Current or ever smoking, n (%)	13 246 (52.1)	13 282 (52.2)	13 214 (52.0)	13 738 (54.0)	53 480 (52.6)	<0.001
Prevalent diabetes mellitus, n (%)	1634 (6.4)	1607 (6.3)	1594 (6.3)	2047 (8.0)	6882 (6.8)	<0.001
Prevalent hypertension, n (%)	8147 (32.0)	8247 (32.4)	8346 (32.8)	8554 (33.6)	33 294 (32.7)	<0.001
Family history of any cancer, n (%)	14 222 (55.9)	14 338 (56.4)	14 251 (56.0)	14 319 (56.3)	57 130 (56.2)	0.702
Family history of lung cancer, n (%)	2641 (10.4)	2724 (10.7)	2774 (10.9)	2758 (10.8)	10 897 (10.7)	0.243
Employment, n (%)						0.131
Working	10 143 (39.9)	10 199 (40.1)	9892 (38.9)	10 158 (39.9)	40 392 (39.7)	
Retired	10 983 (43.2)	10 996 (43.2)	11 180 (44.0)	11 029 (43.4)	44 188 (43.4)	
Unemployed	4179 (16.4)	4128 (16.2)	4257 (16.7)	4128 (16.2)	16 692 (16.4)	
Unknown	128 (0.5)	110 (0.5)	103 (0.4)	119 (0.5)	460 (0.5)	
Education, n (%)						<0.001
No high school degree	1509 (5.9)	1501 (5.9)	1539 (6.1)	1651 (6.5)	6200 (6.1)	
High school degree	8981 (35.3)	8997 (35.4)	9361 (36.8)	9363 (36.8)	36 702 (36.1)	
Some college	5458 (21.5)	5369 (21.1)	5404 (21.2)	5551 (21.8)	21 782 (21.4)	
College degree or postgraduate	9485 (37.3)	9566 (37.6)	9128 (35.9)	8869 (34.9)	37 048 (36.4)	
Marital status, n (%)						0.047
Single	834 (3.3)	812 (3.2)	757 (3.0)	805 (3.2)	3208 (3.1)	
Married	19 732 (77.6)	19 934 (78.4)	20 079 (78.9)	19 999 (78.6)	79 744 (78.4)	
No longer married	4817 (18.9)	4645 (18.2)	4549 (17.9)	4586 (18.0)	18 597 (18.3)	
Unknown	50 (0.2)	42 (0.2)	47 (0.2)	44 (0.2)	183 (0.2)	
Race						<0.001
Non-Hispanic White	22 930 (90.2)	23 203 (91.2)	23 289 (91.6)	23 206 (91.3)	92 628 (91.1)	
Non-Hispanic Black	804 (3.2)	797 (3.1)	812 (3.2)	942 (3.7)	3355 (3.3)	
Hispanic	386 (1.5)	356 (1.4)	363 (1.4)	364 (1.4)	1469 (1.4)	
Others	1313 (5.1)	1077 (4.3)	968 (3.8)	922 (3.6)	4280 (4.2)	
Total caloric intake (cal/day)	1769.6±741.7	1718.7±716.9	1663.9±701.2	1794.7±755.5	1735.9±729.9	<0.001
Healthy Eating Index-2015	67.1±9.8	67.2±9.6	66.6±9.3	65.1±9.9	66.5±9.7	<0.001
Alcohol consumption (g/day)	10.5±25.8	9.7±24.3	8.6±22.0	8.5±22.5	9.3±23.6	<0.001
Macronutrients intake						
Carbohydrates (g/day)	225.4±92.3	219.7±89.3	212.8±87.5	228.0±94.8	221.4±91.1	<0.001
Protein (g/day)	67.9±30.9	66.0±29.5	63.9±29.1	68.6±31.2	66.6±30.2	<0.001
SFA (g/day)	20.1±12.0	19.5±11.5	19.1±11.1	20.9±12.1	19.9±11.7	<0.001
MUFA (g/day)	23.8±13.4	23.1±12.9	22.5±12.5	24.9±13.8	23.6±13.2	<0.001
PUFA (g/day)	14.2±7.7	13.9±7.5	13.5±7.3	14.7±7.9	14.1±7.6	<0.001
Food consumption						
Fruit (g/day)	284.8±228.7	280.1±216.4	264.7±203.6	265.0±215.1	273.2±215.9	<0.001

Continued

Table 1 Continued

Characteristic	Energy-adjusted ultra-processed foods consumption in each quarter (serving/day)				Full sample	P value
	Q1 (<1.0)	Q2 (1.0-<1.8)	Q3 (1.8-<3.7)	Q4 (≥3.7)		
Vegetable (g/day)	296.5±196.8	284.7±183.6	272.3±175.1	280.8±184.4	282.9±184.3	<0.001
Red meat (g/day)	61.3±52.8	59.4±50.9	58.8±49.4	66.7±55.5	61.6±52.2	<0.001
Dietary fibre (g/day)	18.6±8.9	18.1±8.3	17.3±8.1	18.0±8.4	18.0±8.4	<0.001
Added sugar (tsp/day)	12.4±9.1	12.1±8.8	11.9±8.4	13.6±10.4	12.5±9.3	<0.001
Mineral intake						
Sodium (mg/day)	2760.9±1226.1	2690.1±1180.1	2623.8±1160.9	2857.6±1275.8	2364.7±1150.2	<0.001
Potassium (mg/day)	3333.3±1299.1	3259.9±1232.4	3143.5±1206.8	3254.7±1264.6	3244.3±1248.6	<0.001
Calcium (mg/day)	770.8±424.8	752.8±401.9	720.6±388.8	754.0±399.6	748.7±402.8	<0.001
Magnesium (mg/day)	331.8±131.9	323.4±125.7	310.4±121.9	323.1±127.3	321.8±126.5	<0.001

Continuous variables were expressed as mean±SD and categorical variables as numbers and percentages.
BMI, body mass index; MUFA, monounsaturated fatty acids; PUFA, polyunsaturated fatty acid; SFA, saturated fatty acids; UPF, ultra-processed food.

NSCLC in a non-linear dose–response manner, but linear for SCLC (see figure 3).

DISCUSSION

In this large cohort with long-term follow-up, UPF consumption was linked to higher risks of lung cancer, including NSCLC and SCLC, independent of a variety of potential confounders including key risk factors related to lung cancer and overall diet quality. Subgroup analyses confirmed consistent results, supporting its role as a potential environmental risk factor.

Interpretation and comparison with other studies

UPF, classified by NOVA, are profitable owing to their extended shelf life but are nutritionally poor, with high energy density, low fibre, fewer micronutrients and excessive sugars, sodium, fats and additives.[19] Worse still, over the past two decades, the consumption of UPF has significantly increased worldwide, regardless of development or economic status.[20] The rise in UPF consumption may have driven global increases in obesity, cardiovascular disease, metabolic disorders, cancer and mortality, as these foods are confirmed risk factors for such conditions.[8 21]

A limited number of studies have explored the link between UPF and lung cancer. A UK Biobank cohort study of 416 588 participants with a median follow-up of 7.13 years found that high red meat and processed meat intake increased lung cancer risk. A healthy diet low in processed/red meats and refined grains, but high in fruits, vegetables, fish and whole grains, protects against lung cancer.[22] In addition, previous studies from several regions showed that higher adherence to the Western dietary pattern, characterised by high UPF consumption, was associated with an increased risk of lung cancer.[23] Replacing processed/UPF with minimally processed ones may reduce the risks of neck, colon and liver cancers, according to the European Prospective Investigation into Cancer and Nutrition (EPIC) Study.[24] Finally, a UK-based cohort study suggests that higher UPF consumption may be linked to an increased burden and mortality for overall and certain site-specific cancers, especially ovarian cancer in women.[25] However, the EPIC Study also indicated that neither red meat nor processed meat was significantly related to an increased risk of lung cancer.[12] Furthermore, a prospective study based on the Women's Health Initiative Observational Study found that high diet quality (low in red/processed meats, refined grains and sugary drinks) was not linked to overall lung cancer in postmenopausal women, but reduced squamous cell lung cancer risk.[26] To our knowledge, this study is one of the few large cohort study analyses that report an association between UPF consumption and an increased risk of lung cancer and its subtypes based on better capture of UPF intake using accurate intake assessment tools and a standardised diagnosis procedure for clinical outcomes over a relatively long follow-up period. In subgroup analysis, non-smokers had a greater risk from UPF, possibly due to their lower baseline lung cancer risk, making lifestyle factors, particularly diet, more influential.[3 27]

Several hypotheses could be put forward to explain our findings. First, the poor nutritional quality of UPF may drive the inverse association directly and indirectly. Several nutritional compounds, including high sodium, saturated fat, added sugars and low fibre and potassium, are known to harm cognitive

Table 2 Lung cancer of energy-adjusted ultra-processed foods consumption categories

Outcome	Energy-adjusted ultra-processed foods consumption in each quarter (serving/day)				Full sample	P value
	Q1 (<1.0)	Q2 (1.0-<1.8)	Q3 (1.8-<3.7)	Q4 (≥3.7)		
Lung cancer in follow-up, n	331/25 433	409/25 433	471/25 432	495/25 434	1706/101 732	<0.001
Non-small cell lung cancer, n (%)	289 (87.3)	354 (86.6)	407 (86.4)	423 (85.5)	1473 (86.3)	0.791
Small cell lung cancer, n (%)	42 (12.7)	55 (13.4)	64 (13.6)	72 (14.5)	233 (13.7)	
Lung cancer stage						0.349
Stage I and II, n (%)	118 (35.6)	142 (34.8)	181 (38.4)	158 (32.0)	599 (35.1)	
Stage III and IV, n (%)	171 (51.7)	212 (51.8)	226 (48.0)	265 (53.5)	874 (51.2)	
Small cell lung cancer, n (%)	42 (12.7)	55 (13.4)	64 (13.6)	72 (14.5)	233 (13.7)	

Categorical variables as numbers and percentages.

Table 3 Cox proportional hazard ratios for lung cancer of energy-adjusted ultra-processed foods consumption categories

Variables	Energy-adjusted ultra-processed foods consumption in each quarter (serving/day)				P for trend
	Q1 (<1.0)	Q2 (1.0–<1.8)	Q3 (1.8–<3.7)	Q4 (≥3.7)	
No. of participants	25 433	25 433	25 432	25 434	
Lung cancer					
No. of events	331	409	471	495	
Person-years	303 000.9	304 826.4	303 082.8	302 622.7	
Rate per 1000 person-years	1.1	1.3	1.6	1.6	
Unadjusted	1.00 (reference)	1.24 (1.09 to 1.41)	1.43 (1.26 to 1.62)	1.51 (1.34 to 1.71)	<0.001
Multivariable	1.00 (reference)	1.22 (1.05 to 1.37)	1.33 (1.17 to 1.54)	1.41 (1.22–1.60)	<0.001
Non-small cell lung cancer					
No. of events	289	354	407	423	
Person-years	302 595.5	304 353.3	302 476.6	301 917.9	
Rate per 1000 person-years	1.0	1.2	1.3	1.4	
Unadjusted	1.00 (reference)	1.23 (1.07 to 1.41)	1.41 (1.24 to 1.61)	1.48 (1.30 to 1.69)	<0.001
Multivariable	1.00 (reference)	1.21 (1.04 to 1.40)	1.35 (1.15 to 1.53)	1.37 (1.20 to 1.58)	<0.001
Small cell lung cancer					
No. of events	42	55	64	72	
Person-years	300 330.7	301 768.0	299 474.4	298 809.9	
Rate per 1000 person-years	0.1	0.2	0.2	0.2	
Unadjusted	1.00 (reference)	1.25 (0.92 to 1.78)	1.47 (1.02 to 2.10)	1.66 (1.17 to 2.36)	<0.001
Multivariable	1.00 (reference)	1.22 (0.81 to 1.80)	1.41 (0.95 to 2.01)	1.44 (1.03 to 2.10)	0.001

Values are hazard ratios (95% confidence intervals)
Multivariable model: adjusted for sex, age, race, family history of lung cancer, study arm, recruitment site, year of randomisation, prevalent hypertension, prevalent diabetes, smoking status, alcohol consumption, HEI-2015 score, employment status, marital status, physical activity status and BMI.
BMI, body mass index; HEI-2015, Healthy Eating Index-2015.

health.[22 28] Participants with high UPF intake had poor nutrition, reflected in a lower baseline HEI-2015. Additionally, low consumption of minimally processed foods like fruits, vegetables, fish and whole grains was linked to increased lung cancer risk.[22 29]

Second, UPF would affect satiety control and glycaemic responses. Indeed, a study showed that more processed foods lead to higher glycaemic response and lower satiety and may disrupt endocrine balance, increasing energy intake.[30] However, the associations between UPF consumption and lung cancer risk remained strong even after adjusting for energy intake and BMI, suggesting that other bioactive compounds in these foods may also play a role.

Third, a wide range of additives are used in UPF, which could have adverse effects on lung cancer. For instance, basic research suggests deregulated glutamate may play a role in lung cancer's pathogenesis and adverse outcomes.[31] Carrageenan, a food additive used for thickening, can cause intestinal inflammation in cell and animal models, leading to gastrointestinal issues and, when intestinal flora dysbiosis occurs, may contribute to lung cancer.[32 33]

Fourth, industrial processing alters the food matrix, affecting nutrient availability and absorption, while also generating harmful contaminants. Acrolein, found in grilled sausages and caramel candies, is a toxic component of cigarette smoke that contributes to lung cancer by damaging mitochondrial DNA,

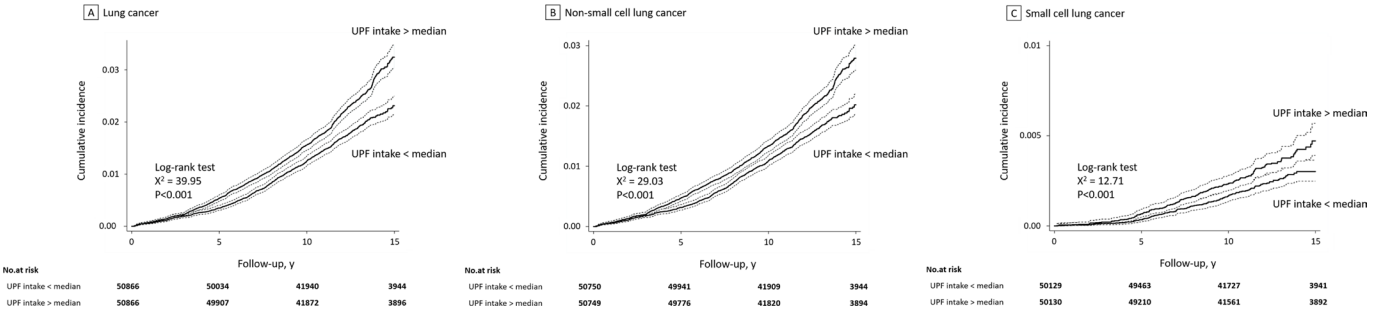

Figure 2 Adjusted cumulative incidence of lung cancer based on energy-adjusted ultra-processed food consumption. Data are for the cumulative incidence of (A) lung cancer, (B) non-small cell lung cancer and (C) small cell lung cancer among participants with and without UPF intake greater than the median. The dotted lines represent the upper and lower bounds of the 95% CI of curves. Adjustments were made for age and sex. UPF, energy-adjusted ultra-processed food.

Table 4 Subgroup analyses for association between consumption of energy-adjusted ultra-processed foods and lung cancer (highest vs lowest quarter of consumption)

	Lung cancer		Non-small cell lung cancer		Small cell lung cancer	
Subgroup variables	**HR (95% CI)**	**P for interaction**	**HR (95% CI)**	**P for interaction**	**HR (95% CI)**	**P for interaction**
Age (years)						
≥65	1.52 (1.31 to 1.77)	0.058	1.46 (1.23 to 1.73)	0.186	1.47 (1.06 to 2.12)	0.248
<65	1.22 (1.04 to 1.42)		1.20 (0.94 to 1.53)		1.22 (0.80 to 1.76)	
Sex						
Male	1.65 (1.34 to 2.02)	0.157	1.61 (1.27 to 2.01)	0.075	1.59 (1.28 to 2.35)	0.157
Female	1.28 (1.09 to 1.51)		1.22 (1.03 to 1.46)		1.24 (0.84 to 1.89)	
Family history of lung cancer						
Yes	1.54 (1.13 to 2.09)	0.488	1.45 (1.05 to 2.03)	0.548	1.50 (1.09 to 2.15)	0.845
No	1.36 (1.17 to 1.57)		1.36 (1.16 to 1.57)		1.37 (0.95 to 2.02)	
BMI (kg/m^2)						
≥25	1.54 (1.31 to 1.74)	0.257	1.47 (1.24 to 1.75)	0.467	1.47 (1.07 to 2.15)	0.784
<25	1.24 (1.01 to 1.49)		1.23 (0.99 to 1.55)		1.36 (0.95 to 2.02)	
Smoking						
Current or former	1.34 (1.15 to 1.54)	0.523	1.30 (1.08 to 1.47)	0.848	1.36 (0.96 to 2.03)	0.857
Never	1.44 (1.26 to 1.68)		1.43 (1.23 to 1.64)		1.69 (1.29 to 2.35)	
Years of follow-up						
≥2 years	1.43 (1.23 to 1.62)	0.756	1.38 (1.21 to 1.59)	0.578	1.45 (1.05 to 2.13)	0.178
< 2 years	1.40 (1.21 to 1.59)		1.37 (1.19 to 1.56)		1.43 (1.02 to 2.08)	

Adjusted for sex, age, race, family history of lung cancer, study arm, recruitment site, year of randomisation, prevalent hypertension, prevalent diabetes, smoking status, alcohol consumption, HEI-2015 score, employment status, marital status, physical activity status and BMI.
BMI, body mass index; HEI-2015, Healthy Eating Index-2015.

inducing mitochondrial fission and promoting mitophagy in human lung cells.[34 35]

Finally, UPF may be contaminated by packaging materials, like polychlorinated biphenyls, which can negatively affect lung cancer risk. Previous research found that high levels of polychlorinated biphenyls with oestrogenic activity could promote lung cancer cell proliferation in both non-neoplastic and neoplastic lung cells via oestrogen receptor beta.[36] Additionally, a population-based study found that serum concentrations of polychlorinated biphenyls were linked to lung cancer risk, even decades after their production and their use was banned.[37]

Strengths and limitations of this study

The major strengths of our study include its large, multicentre cohort design, standardised methods for food intake and outcome assessment, adjustment for various potential confounders and robust sensitivity analyses. Another strength of our longitudinal study is its novelty.

However, some limitations of the present study should be noted. First, due to the observational design, causality cannot be determined and residual confounding due to unmeasured risk factors for lung cancer or imprecision in the measure of included covariates cannot be excluded. In particular, adjustments

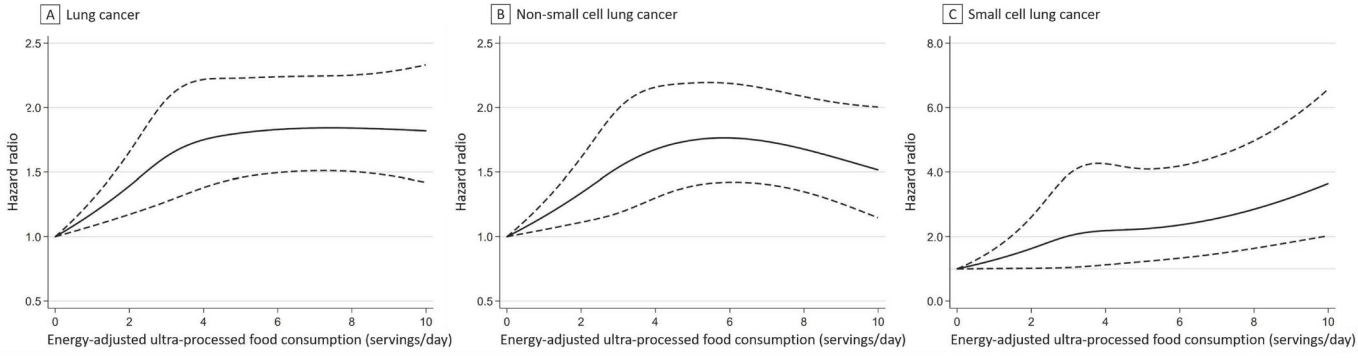

Figure 3 Adjusted dose–response associations between energy-adjusted ultra-processed food consumption and risk of lung cancer. Data are for the hazard ratio of (A) lung cancer, (B) non-small cell lung cancer and (C) small cell lung cancer. The reference level was set at 0 serving/day. The dotted lines represent the upper and lower bounds of the 95% CI of the hazard ratio. Multivariable adjustments were made for sex, age, race, family history of lung cancer, study arm, recruitment site, year of randomisation, prevalent hypertension, prevalent diabetes, smoking status, alcohol consumption, HEI-2015 score, employment status, marital status, physical activity status and body mass index .HEI-2015, Healthy Eating Index-2015.

were made for smoking status but not smoking intensity, which may introduce potential confounding factors. To limit residual confounding, we adjusted for a wide range of potential confounders regarding lifestyle, clinical and demographic variables collected within PLCO, and several sensitivity analyses showed the high stability of the results. Moreover, the E-value for lung cancer was 2.17, for NSCLC 2.02 and SCLC 2.24 in our study setting, indicating that an unmeasured confounder with an HR≥2.02 can explain away the observed association in our study. The possibility of the existence of such an unmeasured confounder seems to be low, as the HR for smoking, a strong lung cancer risk factor, was only 1.79 in our study.

Second, the limited number of incident events could lower the analytical power in these analyses, and risk estimates would have wide confidence intervals and should be interpreted with caution.

Third, inherent biases from self-reported DHQ might have affected our findings, including measurement errors and misclassification bias, which might have biased towards null. Hence, energy-adjusted food consumption was used in the present study and our analytic strategy capitalises on the ability of DHQ to rank individuals according to relative food intake. The dietary information was assessed only once between the baseline examination and outcome ascertainment, while the diets might change over time or in response to health advice, which could lead to misclassifications. The potential for selection bias and the fact that multiple comparisons were conducted in the regression, may result in false positives.

Fourth, the DHQ was not specifically designed to collect data about the NOVA classification of UPF consumption, so some items of UPF were missing since lacking information based on the Food Frequency Questionnaire.

Finally, the majority of participants were non-Hispanic White and the generalisability of our findings to other races/ethnicities is limited. Nevertheless, community-based cohorts are usually non-representative and therefore generalisation of these results must be based on biological mechanisms instead of statistical representativeness.

Conclusions and policy implications

In this population-based study, high consumption of UPF is associated with increased risks of lung cancer, NSCLC and SCLC, independent of multiple potential confounders. These findings need to be confirmed by other large-scale longitudinal studies in different populations and settings, considering the aforementioned limitations. If causality is established, limiting trends of UPF intake globally could contribute to reducing the burden of lung cancer. Finally, future studies should elucidate potential molecular mechanisms and increase understanding of the observed associations.

Acknowledgements We thank the National Cancer Institute (NCI) for access to their data collected by the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. We thank the PLCO screening center investigators, staff members and everyone involved in collecting data for the PLCO study. We thank the trial participants for their contributions that made this study possible. We also thank the Chongqing University Cancer Hospital. The statements contained herein are solely those of the authors and do not reflect the opinions or views of the NCI.

Contributors All authors read and approved the final manuscript. Study concept and design: all authors. Acquisition, analysis, or interpretation of data: KW. Drafting of the manuscript: KW. Critical revision of the manuscript for important intellectual content: WZ, and YW. Statistical analysis: KW and JZ. Obtained funding: KW and YW. Supervision: YW. YW is the guarantor.

Funding Support was provided by the Chongqing University Cancer Hospital, Chongqing, China, and grants from the Chongqing Talent Plan (No. cstc2022ycjh-bgzxm0208), Chongqing Shapingba District Technological Innovation Project(No. 2024161).

Competing interests None declared.

Patient and public involvement statement No participants were involved in setting the research question or the outcome measures, nor were they involved in developing plans for the design or implementation of the study. No participants were asked to advise on interpretation or writing up of results. We did not have access to patients or members of the public with the level of statistical or methodological expertise to analyze or interpret the present results. The results of this study will be disseminated to the participants and the public through the cohort website, public sessions, and a press release.

Patient consent for publication Consent obtained directly from patient(s)

Ethics approval The current project and analysis using the PLCO data were approved by the National Cancer Institute Cancer Data Access System (PLCO-1547) and the Chongqing University Cancer Hospital Institutional Review Board (CZLS2024084-A). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data described in the manuscript, code book and analytic code will not be made available because the authors are prohibited from distributing or transferring the data and code books on which their research was based to any other individual or entity under the terms of an approved NCI Research Proposal and Data and Materials Distribution Agreement through which the authors obtained these data.

Author note The authors of this manuscript declare that in the writing process of this work, no generative artificial intelligence (AI) or AI-assisted technologies were used to generate content, ideas, or theories. We used AI solely for enhancing readability and refining language. This use was under strict human oversight and control and the manuscript was thoroughly revised by human eyes and judgement.

ORCID iD

Kanran Wang http://orcid.org/0000-0002-6958-7677

REFERENCES

1 Thai AA, Solomon BJ, Sequist LV, *et al*. Lung cancer. *Lancet* 2021;398:535–54.
2 J. Saller J, Boyle TA. Molecular Pathology of Lung Cancer. *Cold Spring Harb Perspect Med* 2022;12.
3 Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. *Clin Chest Med* 2020;41:1–24.
4 Martí Del Moral A, Calvo C, Martínez A. Ultra-processed food consumption and obesity-a systematic review. *Nutr Hosp* 2021;38:177–85.
5 Allemandi L, Castronuovo L, Tiscornia MV, *et al*. Food advertising on Argentinean television: are ultra-processed foods in the lead? *Public Health Nutr* 2018;21:238–46.
6 Buckley JP, Kim H, Wong E, *et al*. Ultra-processed food consumption and exposure to phthalates and bisphenols in the US National Health and Nutrition Examination Survey, 2013-2014. *Environ Int* 2019;131:105057.
7 Vivi ACP, Azevedo-Silva TR, Neri D, *et al*. Association between ultraprocessed food intake and C-reactive protein levels in preterm and term infants. *Nutrition* 2022;99–100:111649.
8 Chen X, Zhang Z, Yang H, *et al*. Consumption of ultra-processed foods and health outcomes: a systematic review of epidemiological studies. *Nutr J* 2020;19:86.
9 Rico-Campà A, Martínez-González MA, Alvarez-Alvarez I, *et al*. Association between consumption of ultra-processed foods and all cause mortality: SUN prospective cohort study. *BMJ* 2019;365:l1949.
10 Srour B, Fezeu LK, Kesse-Guyot E, *et al*. Ultra-processed food intake and risk of cardiovascular disease: prospective cohort study (NutriNet-Santé). *BMJ* 2019;365:l1451.
11 Tu H, Heymach JV, Wen C-P, *et al*. Different dietary patterns and reduction of lung cancer risk: A large case-control study in the U.S. *Sci Rep* 2016;6:26760.
12 Linseisen J, Rohrmann S, Bueno-de-Mesquita B, *et al*. Consumption of meat and fish and risk of lung cancer: results from the European Prospective Investigation into Cancer and Nutrition. *Cancer Causes Control* 2011;22:909–18.

13 Prorok PC, Wright P, Riley TR, *et al*. Overall and Multiphasic Findings of the Prostate, Lung, Colorectal and Ovarian (PLCO) Randomized Cancer Screening Trial. *Rev Recent Clin Trials* 2018;13:257–73.
14 Hoyt M, Reger M, Marley A, *et al*. Vitamin K intake and prostate cancer risk in the Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) Screening Trial. *Am J Clin Nutr* 2019;109:392–401.
15 Krebs-Smith SM, Pannucci TE, Subar AF, *et al*. Update of the Healthy Eating Index: HEI-2015. *J Acad Nutr Diet* 2018;118:S2212-2672(18)30838-4:1591–602:.
16 Subar AF, Thompson FE, Kipnis V, *et al*. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires : the Eating at America's Table Study. *Am J Epidemiol* 2001;154:1089–99.
17 Monteiro CA, Cannon G, Levy RB, *et al*. Ultra-processed foods: what they are and how to identify them. *Public Health Nutr* 2019;22:936–41.
18 Orlich MJ, Sabaté J, Mashchak A, *et al*. Ultra-processed food intake and animal-based food intake and mortality in the Adventist Health Study-2. *Am J Clin Nutr* 2022;115:1589–601.
19 Elizabeth L, Machado P, Zinöcker M, *et al*. Ultra-Processed Foods and Health Outcomes: A Narrative Review. *Nutrients* 2020;12:1955.
20 Baker P, Machado P, Santos T, *et al*. Ultra-processed foods and the nutrition transition: Global, regional and national trends, food systems transformations and political economy drivers. *Obes Rev* 2020;21:e13126.
21 Cai Q, Duan MJ, Dekker LH, *et al*. Ultraprocessed food consumption and kidney function decline in a population-based cohort in the Netherlands. *Am J Clin Nutr* 2022;116:263–73.
22 Wei X, Zhu C, Ji M, *et al*. Diet and Risk of Incident Lung Cancer: A Large Prospective Cohort Study in UK Biobank. *Am J Clin Nutr* 2021;114:2043–51.
23 He F, Xiao RD, Lin T, *et al*. Dietary patterns, BCMO1 polymorphisms, and primary lung cancer risk in a Han Chinese population: a case-control study in Southeast China. *BMC Cancer* 2018;18:445.
24 Kliemann N, Rauber F, Bertazzi Levy R, *et al*. Food processing and cancer risk in Europe: results from the prospective EPIC cohort study. *Lancet Planet Health* 2023;7:e219–32.
25 Chang K, Gunter MJ, Rauber F, *et al*. Ultra-processed food consumption, cancer risk and cancer mortality: a large-scale prospective analysis within the UK Biobank. *EClinicalMedicine* 2023;56:101840.
26 Myneni AA, Giovino GA, Millen AE, *et al*. Indices of Diet Quality and Risk of Lung Cancer in the Women's Health Initiative Observational Study. *J Nutr* 2021;151:1618–27.
27 Jassem E, Szymanowska A, Siemińska A, *et al*. Smoking and lung cancer. *Pneumonol Alergol Pol* 2009;77:469–73.
28 Ko EA, Kim YW, Lee D, *et al*. Expression of potassium channel genes predicts clinical outcome in lung cancer. *Korean J Physiol Pharmacol* 2019;23:529–37.
29 Yan H, Jin X, Zhang C, *et al*. Associations between diet and incidence risk of lung cancer: A Mendelian randomization study. *Front Nutr* 2023;10:1149317.
30 Fardet A. Minimally processed foods are more satiating and less hyperglycemic than ultra-processed foods: a preliminary study with 98 ready-to-eat foods. *Food Funct* 2016;7:2338–46.
31 Song GQ, Wu P, Dong XM, *et al*. Elemene induces cell apoptosis via inhibiting glutathione synthesis in lung adenocarcinoma. *J Ethnopharmacol* 2023;311:116409.
32 Wu W, Zhou J, Xuan R, *et al*. Dietary κ-carrageenan facilitates gut microbiota-mediated intestinal inflammation. *Carbohydr Polym* 2022;277:118830.
33 Zhao Y, Liu Y, Li S, *et al*. Role of lung and gut microbiota on lung cancer pathogenesis. *J Cancer Res Clin Oncol* 2021;147:2177–86.
34 Hikisz P, Jacenik D. The Tobacco Smoke Component, Acrolein, as a Major Culprit in Lung Diseases and Respiratory Cancers: Molecular Mechanisms of Acrolein Cytotoxic Activity. *Cells* 2023;12:879.
35 Wang HT, Lin JH, Yang CH, *et al*. Acrolein induces mtDNA damages, mitochondrial fission and mitophagy in human lung cells. *Oncotarget* 2017;8:70406–21.
36 Recio-Vega R, Mendez-Henandez A, Gabriel APY, *et al*. Potentially estrogenic polychlorinated biphenyls congeners serum levels and its relation with lung cancer. *J Appl Toxicol* 2013;33:906–14.
37 Park EY, Park E, Kim J, *et al*. Impact of environmental exposure to persistent organic pollutants on lung cancer risk. *Environ Int* 2020;143:105925.