

1 **Ultra-processed food consumption and risk of lung cancer: Results from a Prospective**  
2 **Study**

3 Tefera Chane Mekonnen<sup>1,5</sup>, Yohannes Adama Melaka<sup>1,2,3</sup>, Zumin Shi<sup>4</sup>, Tiffany K Gill<sup>1</sup>

4 <sup>1</sup> Adelaide Medical School, The University of Adelaide, South Australian Health and Medical  
5 Research Institute, North Terrace, Adelaide SA 5000, Australia.

6 <sup>2</sup>Flinders Health and Medical Institute, Flinders University, Adelaide 5001, South Australia.

7 <sup>3</sup> Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia

8 <sup>4</sup>Human Nutrition Department, College of Health Sciences, Qatar University, Qatar

9 <sup>5</sup>School of Public Health, College of Medicine and Health Science, Wollo University, Dessie  
10 1145, Ethiopia.

11 **Competing interest:** The authors declare no conflicts of interest.

12 **Corresponding Author:** Name: Tefera Chane Mekonnen, mailing address: Adelaide  
13 Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Level 7,  
14 SAHMRI North Tce, Adelaide SA 5000, Email: [tefera.mekonnen@adelaide.edu.au](mailto:tefera.mekonnen@adelaide.edu.au)

15

16 **Abbreviations:** BMI, Body Mass Index; BQ, Baseline Questionnaire; EPIC, European  
17 Prospective Investigation into Cancer and Nutrition; CVD, Cardiovascular disease; DHQ,  
18 Dietary History Questionnaire; FFQ, Food Frequency Questionnaire; LC: Lung Cancer;  
19 NCDs, Non-communicable Diseases; NCI, National Cancer Institute; PLCO, Prostate, Lung,  
20 Colorectal and Ovarian Cancer; UPF, Ultra-processed Food; UK, United Kingdom.

21

22

23

24

25

26 **Abstract**

27 **Background and Aims:** There is limited evidence on the link between ultra-processed food  
28 (UPF) intake and the risk of lung cancer (LC). This study examined the association between  
29 UPF and LC risk using data from the Prostate, Lung, Colorectal, and Ovarian (PLCO) cancer  
30 trial.

31 **Methods:** This study involved PLCO participants (n = 96,607, aged  $\geq 55$  years) who were  
32 followed between 1998 and 2009. Food items were categorized based on the NOVA  
33 classification. Cox regression models with inverse probability of censoring weighting  
34 (IPCW) were utilized to estimate the association between UPF intake and LC risk. The joint  
35 effect of UPF and diabetes was explored using additive hazard models to calculate the  
36 additional number of LC cases.

37 **Results:** During a median follow-up period of 9.4 years, 1,596 incident LC cases were  
38 identified. UPF consumption (in %gram/day) showed no significant association with the  
39 overall risk of LC. However, adults with diabetes in the highest quintile of UPF intake had a  
40 significantly higher risk of LC (HR = 2.44; 95% CI: 1.27, 4.67) compared to participants  
41 without diabetes. A small excess risk due to the interaction between UPF and diabetes (0.13;  
42 95% CI -0.32, 0.58) was observed, resulting in an additional 201 cases of LC per  $10^5$  person-  
43 years (95% CI: 70, 332) attributed to the highest UPF intake and diabetes interaction.  
44 Furthermore, a 10% increment in UPF intake (%kcal/day) increased the risk of LC by 32%.

45 **Conclusions:** While UPF, in terms of weight contribution, is associated with a higher risk of  
46 LC in participants with diabetes, UPF (in %kcal/day), is associated with an increased risk of  
47 LC in all participants. Lowering UPF intake may help reduce the risk of LC in both diabetic  
48 patients and the general population.

49 Key words: Ultra-processed food; incidence; lung cancer; older adult

## 50 **Introduction**

51 Lung cancer (LC) is one of the most common and deadliest types of cancer worldwide and its  
52 incidence has been increasing in many countries over the past few decades [1]. It is estimated  
53 that approximately 2.3 million people are diagnosed with LC each year, over 2 million people  
54 die from the disease annually and LC accounts for 45.9 million disability-adjusted life years  
55 (DALYs) in both sexes [2, 3]. LC is the leading cause of cancer death, accounting for 23% of  
56 all cancer deaths in America [4]. While smoking is the primary cause of LC, there is growing  
57 concern that dietary factors may also play a role in the development of the disease [5, 6].

58 Ultra-processed food (UPF) is typically highly processed and contain a large number of  
59 additives, such as artificial flavours, colours, and preservatives. The consumption of UPF in  
60 the modern diet has increased dramatically in recent years, and this has raised concerns about  
61 their potential impact on health [7].

62 There is limited scientific evidence directly linking the consumption of UPF to an increased  
63 risk of LC. However, there is a growing body of research suggesting that a diet high in UPF  
64 may increase the risk of other types of cancer, as well as other chronic diseases [8-10]. A  
65 recent (2023) systematic review and meta-analysis, which did not include LC, found that a  
66 higher proportion of UPF intake increases the risk of overall cancer by 13%, breast cancer by  
67 11%, colorectal cancer by 30% and pancreatic cancer by 49% [11]. Consistent  
68 epidemiological findings from population-based studies (UK biobank, the European  
69 Prospective Investigation into Cancer and Nutrition (EPIC) study, the French NutriNet-Sante  
70 cohort and US-based cohorts) have reported a link between UPF consumption and overall  
71 cancer, ovarian, colorectal, breast and pancreatic cancer [12-17]. Only two of these studies  
72 evaluated the association of UPF and LC. These two studies have investigated the association  
73 between UPF and LC using UK biobank and EPIC study and shown no association. One

74 study found that a diet high in UPF was associated with an increased risk of all types of  
75 cancer, including LC. However, the study did not specifically examine the relationship  
76 between ultra-processed food consumption and LC risk [18].

77 While the relationship between UPF consumption and LC risk is not yet fully understood,  
78 there is some evidence to suggest that a diet high in ultra-processed foods may be linked to an  
79 increased risk of LC [19-24]. Examining the association and understanding the potential link  
80 between UPF consumption and LC risk is important for developing effective strategies to  
81 prevent this disease. Hence, the current study was aimed to investigate the association  
82 between proportion of UPF intake and LC among older adults in United States of America  
83 (USA).

## 84 **Methods**

### 85 **Study Population and Study Design**

86 The study used data from the Prostate, Lung, Ovarian and Colorectal Cancer trial (PLCO), a  
87 cancer screening trial in 10 study centers across the USA. The trial was designed to evaluate  
88 whether screening examinations reduces cancer mortality by randomly allocating  
89 approximately 155,000 older adults to the control arm (usual care) or to the intervention arm  
90 (screening examination). The randomization began in 1993. Eligible study subjects were  
91 aged between 55 to 74 years at enrolment and free from prostate, lung, colorectal and ovarian  
92 cancer, with additional criteria documented elsewhere [25]. Initial information on  
93 demographic, lifestyle, medical history, family history and medication used was collected  
94 through the baseline questionnaire (BQ). The dietary history questionnaire (DHQ) was  
95 administered to both arms commencing in 1998 and a total of 113,000 (77%) participants  
96 provided a complete response with a 3-year median time to the study. Study subjects were  
97 monitored approximately for 12 years to obtain data related to the diagnosis of LC.

98 Overall, 101,732 participants were eligible for dietary history and LC analysis (53,155 study  
99 participants were excluded due to invalid BQ and DHQ). After examining the distribution of  
100 total energy, we further excluded 5,125 participants who were in the bottom and top 1% of  
101 energy intakes, with 96,607 participants remaining in the final analysis (**Supplementary Fig**  
102 **1**). Ethics approval for the trial was obtained from the US National Cancer Institute (NCI)  
103 Review Boards and written informed consent was obtained from each participant. The data  
104 for dietary history and lung cancer analysis were obtained using the Cancer Data Access  
105 System (CDAS) after obtaining approval from NCI (project registration: PLCO-982).

#### 106 **Data Collection and UPF Assessment**

107 Baseline data on demographic characteristics (age, sex, race, marital status, education,  
108 occupation, family income and study arm), types and date of diagnosis for cancer, treatment,  
109 self-reported medical history, personal lifestyle features (smoking, alcohol drinking, and  
110 body-mass index (BMI)), family history of lung cancer, medication use, physical activity and  
111 dietary habits were obtained from the BQ, DHQ, supplementary questionnaire (SQ) and brief  
112 survey questionnaire (BSQ). Dietary data were collected using a self-administered food  
113 frequency questionnaire (FFQ), a validated assessment tool which provides better nutrient  
114 estimate[26]. All food items in the DHQ were grouped into one of four categories from the  
115 NOVA classification, with an emphasis on UPF as described previously by Monteiro [27]. In  
116 case of uncertainty regarding which category the food or beverage belongs to, consensus was  
117 reached among the researchers (**Supplementary Table 1**).

118 From a total of 275 foods or beverages, 145 were categorized as UPF. UPF in this study  
119 includes carbonated drinks, savory packaged snacks; ice cream, chocolate, confectionery;  
120 breads and buns; margarines and spreads; cookies, pastries, and cakes; breakfast cereals,  
121 cereal and energy bars; flavoured milk drinks; cocoa drinks; sweet desserts made from fruit  
122 with added sugars, artificial flavours and texturizing agents; cooked seasoned vegetables with

123 ready-made sauces; meat and chicken extracts and instant sauces; health and slimming  
124 products such as powdered or fortified meal and dish substitutes; ready to heat products;  
125 poultry and fish nuggets and sticks, sausages, burgers, cold cuts, hot dogs, and other  
126 reconstituted meat products, and instant soups, and noodles. These food items were further  
127 regrouped into sugary drinks; processed meats; milk dessert, yogurt, and soy products;  
128 cookies, pies, and pastries; margarine and dressings; sweets and other condiments; salty  
129 snacks; quick bread; and others (**Supplementary Table 1**).

130 UPF was measured by calculating the weight ratio of all food items classified as UPF to the  
131 total weight of all food items consumed by individuals per day (% grams/day). This weight  
132 ratio was chosen instead of an energy ratio because it considers UPF with no or low-calorie  
133 content, such as artificially sweetened beverages, as well as non-nutritional factors associated  
134 with food processing (such as neoformed contaminants, additives, and alterations to the  
135 structure of raw foods)[12, 13, 28]. Additionally, the caloric contribution of UPF was  
136 determined by calculating the fraction of energy from UPF to the total energy from all food  
137 items consumed per day.

### 138 **Outcome Assessment**

139 The outcome of this study was incidence of LC. LC diagnosis was made by self-reported  
140 responses from participants through annual follow-up questionnaires. Cases were  
141 documented from abnormal x-ray results, death certificates and relative reports. LC diagnosis  
142 confirmed from medical record abstraction as per histopathologic type derived from the  
143 International Classification of Diseases-Oncology, 2<sup>nd</sup> edition (ICD-O-2) morphology that  
144 includes either non-small cell lung cancer (NSCLC) (n=1,464) or small cell lung cancer  
145 (SCLC) (n=242). Lung carcinoid tumour was not considered confirmed LC during the trial  
146 and was excluded in this analysis.

### 147 **Statistical Analysis**

148 We utilized an inverse probability of censoring weighting (IPCW) estimator to reduce bias  
149 caused by censoring, specifically dependent censoring [29, 30]. This method has been  
150 employed in right-truncated data [31] and can address the issue of censored subjects by  
151 assigning greater weight to subjects with similar characteristics who are not censored [32].  
152 By making the positivity assumption, this approach enables us to consistently estimate the  
153 effects of covariates and avoids the need to estimate baseline hazards [31].

154 We applied Cox regression models with IPCW to estimate the hazard ratio (HR) and 95%  
155 confidence interval (CI) for the association between proportion of UPF consumption and LC  
156 risk. The follow-up time was defined as the time between completion of DHQ and date of  
157 diagnosis of LC, death, drop-out, or the end of the study up to 31 December 2009. We built  
158 five models in incremental steps by adjusting for potential confounders, selected from  
159 literature and prior knowledge [12, 13, 28, 33]. These covariates considered for the model  
160 adjustment were age (continuous), sex (men, women) , marital status (married, widowed,  
161 divorced, separated/never married), education (up to grade 12 completion, post high school  
162 and some college, and under and postgraduate), ethnicity (Hispanic or non-Hispanic ), study  
163 arm (control, intervention), BMI (<18.5, 18.5-24.99, 25-29.99,  $\geq 30+$ ), cigarette smoking  
164 status (never, current, former), alcohol intake(g/day), total energy (kcal/day), aspirin use (yes,  
165 no), history of baseline comorbidity (yes, no), and family history of LC (yes, no). Models  
166 were also adjusted for average family income (<\$50,000, \$50,000-\$99,000 and  $\geq$  \$100,000)  
167 and physical activity measured in total time spent (in minutes) during each session doing  
168 moderate-to-strenuous exercise, as recorded in the self-reported SQ. Since cigarette smoking  
169 is a well-established risk factor for LC, we further refined the models by considering the  
170 intensity and duration of smoking for both current and former smokers. Smokers were  
171 classified into the following groups: never smokers, current smokers consuming 1-10  
172 cigarettes per day, current smokers consuming 11-20 cigarettes per day, current smokers

173 consuming 21 or more cigarettes per day, former smokers who quit within the past 10 years,  
174 former smokers who quit 11-20 years ago, and former smokers who quit more than 20 years  
175 ago [13, 34]. We also consider cigar and pipe smoking in the analyses.

176 Cox regression with restricted cubic splines was performed to examine the relationship  
177 between per 10% increment in UPF intake and LC risk with three knots (10<sup>th</sup>, 50<sup>th</sup>, and 90<sup>th</sup>),  
178 selected based on Akaike Information Criterion (AIC). In addition to using IPCW, we  
179 conducted competing risk regression analyses to consider the impact of deaths that occurred  
180 before the diagnosis of LC on the observed relationship between the quintile of UPF  
181 consumption and the risk of LC.

182 Subgroup analysis was carried out by stratifying the analysis by sex, age group, race,  
183 cigarette smoking, BMI, family history of LC, history of hypertension, diabetes, chronic  
184 bronchitis, and emphysema to determine whether associations were modified by  
185 multiplicative interactions of covariates and avoid potentially misleading subgroup  
186 differences.

187 Effect modification was assessed using Cox regression to determine the relative excess risk  
188 due to interaction (RERI:  $RR_{11} - RR_{10} - RR_{01} + 1$ ), attributable proportion (AP:  $RERI/RR_{11}$ ),  
189 and synergy index (SI:  $(RR_{11} - 1)/(RR_{10} + RR_{01} - 1)$ ) on an additive scale, with 95%  
190 confidence intervals [35].  $RR_{11}$  represents the relative risk (RR) in the groups with high UPF  
191 exposure and diabetes exposure,  $RR_{10}$  represents the RR in the groups with high UPF intake  
192 and no diabetes, and  $RR_{01}$  represents the RR in the groups with low UPF intake and diabetes.

193 Interactions between UPF consumption and diabetes status were analysed using the surrogate  
194 measures mentioned earlier. A value of zero for RERI and AP indicates no additive  
195 interaction. A positive value suggests a super-additive interaction, while a negative value  
196 indicates a sub-additive interaction [36, 37]. A value of 1 for SI implies no interaction or  
197 exactly additivity;  $SI > 1$  indicates positive interaction or more than additivity;  $SI < 1$



198 suggests negative interaction or less than additivity [37]. The study also examined the  
199 interaction between diabetes status (yes/no) and the quintile of UPF exposure to compare the  
200 outcomes of both types of exposure.

201 In Cox models, deviations from additivity can be assessed using the surrogate measures of  
202 additive interaction obtained from the multiplicative models mentioned above. However,  
203 these measures may sometimes be counter-intuitive and invalid. An alternative approach is to  
204 use additive hazard models, which directly estimates the absolute magnitude of the deviation  
205 from additivity [38, 39]. Therefore, we used additive hazard models to estimate the number  
206 of additional lung cancer incidents per 100,000 person-years of observation. We considered  
207 the quintiles or binary categories of UPF intake (low, below the mean (<31.2%), and high,  $\geq$   
208 the mean ( $\geq 31.2\%$ )), diabetes status, and their interaction using the R package "timereg" [38].  
209 Finally, to check the stability of the findings, we performed the following sensitivity  
210 analyses: a) multiple imputation for covariates with missing values including physical  
211 activity (23,002; 24%) and family income (26,629; 28%); b) excluding participants with  
212 history of hypertension, diabetes, chronic bronchitis, emphysema and obese at baseline; c)  
213 excluding LC cases that occurred during the first five years of follow-up; d) Adjusting for  
214 processed meat, artificial sweetening beverages, fruits and vegetables, fish consumption and  
215 nutrients such as fibers, omega-3, total fat, protein, carbohydrates, and sodium; e) comparing  
216 the findings with the proportion of UPF (%kcal) of the total energy; g) considering  
217 histopathological subtypes of LC (NSCLC and SCLC) and h) removing putative confounders  
218 (smoking, alcohol intake, and BMI) of lung cancer in the final model to see their impact on  
219 the observed association between UPF intake and LC risk. All statistical analyses were  
220 performed using R-software version 4.2.3 and Stata version 18 (College Station, TX, USA).

## 221 **Results**

### 222 **Baseline Characteristics of Study Participants**

223 Among the total participants, 50,803(52.6%) were female and the mean (SD) age was  
224 65.58(5.74) years at baseline for all participants. The proportion of women was consistently  
225 lower across the quintiles UPF consumption from 66.5% to 43.5% compared to men. A  
226 slightly higher pattern of UPF consumption was observed in younger adults (mean age of  
227 64.6 vs 66 years). In addition, compared to participants in the first quintile of UPF  
228 consumption, the number of participants in the fifth quintile was higher among those  
229 classified as obese (18.1% vs 27.8%), current cigarette smokers (6.3% vs 12.8%), current  
230 alcohol drinkers (13.3% vs 18.7%), history of diabetes (6.2% vs 7.8%), and hypertensive  
231 (30.1% vs 34.6 (**Table 1**).

232 The average (SD) proportion of UPF consumption in the total diet (gm/day) was 31.2%  
233 (14.0%) (**Supplementary Fig 2**) and the mean percentage of UPF consumption (%kcal) was  
234 37.13% (**Supplementary Fig 3**). Sugary drinks, milk dessert and processed meats  
235 contributed higher proportion to the total diet in the study participants (**Fig 1a**). The mean  
236 proportion of UPF contribution was relatively higher in men than women 33.45% vs 29.2%.  
237 The predominant UPF groups consumed by both men and women were sugary drinks, milk  
238 dessert, yogurts, and soy products, processed (white and red) meat, cakes, cookies, pies and  
239 pastries, quick breads, and sugar, condiments, and sweets (**Fig 1 b, c**).

#### 240 **UPF Consumption and Lung Cancer Incidence**

241 A total of 1,596 incident cases of LC occurred during a median [IQR] follow up of 9.4[8.02,  
242 10.11) years and mean (SD) follow-up time 8.82(1.93) years. Over a total 853,968.16  
243 person-years, the overall incidence rate of LC was found to be 18.8 (95% CI: 17.7, 19.6)  
244 cases per 10,000 individuals. The incidence rate is higher in study subjects in the highest  
245 quintile of UPF consumption compared to lowest (21 vs 18 per 10,000 observations)  
246 (**Supplementary Table 2**).

247 Following adjustment for all covariates, the risk of LC incidence had no significant  
248 association with a higher consumption of UPF (HR = 0.98; 95% CI: 0.83,1.15, p-for trend  
249 =0.32). Results from competing risk regression did not indicate significant association (HR =  
250 0.95; 95% CI: 0.81,1.12). Furthermore, restricted cubic spline analysis did not demonstrate a  
251 significant relationship between UPF consumption and LC incidence (HR = 0.98; 95% CI:  
252 0.85,1.13; p-nonlinear =0.44) (**Table 2, Fig 2**).

253 Subgroup analysis demonstrated that the magnitude of the observed association between risk  
254 of LC and UPF consumption stratified by most covariates was unchanged (p-interaction >  
255 0.05). However, when comparing the lowest quintile of UPF consumption to the highest  
256 quintile among adults with diabetes, those in the highest quintile had significantly higher risk  
257 of LC (HR<sub>quintile5 vs 1</sub> = 2.44; 95%CI: 1.27, 4.67) compared to those without diabetes (HR<sub>quintile5</sub>  
258 vs 1 = 0.92; 95% CI: 0.42, 2; P<sub>interaction</sub> = 0.002) (**Fig 3**).

259 A further analysis was conducted to examine the relationship between high and low intake of  
260 UPF and diabetes status in relation to the risk of LC. Among individuals without diabetes,  
261 those with high UPF intake had a risk of LC of 0.99 (95% CI: 0.89, 1.10) compared to those  
262 with low UPF intake. However, for individuals with both high UPF intake and diabetes, the  
263 hazard ratio for the risk of LC was 1.17 (95% CI: 0.79, 1.75). There was evidence of a  
264 positive interaction on the additive scale (supra-additivity) with RERI = 0.13 (95% CI: -0.32,  
265 0.58), AP = 0.10 (95% CI: -0.22, 0.42), and SI = 1.67 (95% CI: 0.22, 12.46). However, there  
266 was no statistically significant interaction between high UPF intake and diabetes on the  
267 multiplicative scale (p-value = 0.12). Nevertheless, a significant multiplicative interaction  
268 was observed between quintiles of UPF intake and diabetes (p-value = 0.007) (**Table 3**).  
269 Using additive hazards regression, it was estimated that 6 (95% CI: -13, 26) cases per  
270 100,000 person-years of LC were attributable to high UPF consumption. For diabetes alone,  
271 14 (95% CI: -50, 77) cases of LC were estimated. Due to the joint effect of high UPF intake

272 and diabetes (both binary exposures), an additional 35 (95% CI: -50, 120) cases of LC were  
273 estimated. Among participants with diabetes only, 186 (95% CI: 59, 313) cases per 100,000  
274 person-years of LC were attributable to the highest quintile of UPF exposure. An additional  
275 201 (95% CI: 70, 332) cases of LC were estimated due to the combined interaction between  
276 the highest UPF intake and diabetes (**Table 4**).

277 Sensitivity analyses were undertaken by considering different scenarios. First, we undertook  
278 multiple imputation for observations with missing values. The associations remained  
279 unchanged (i.e., similar with the complete-cases results). The second scenario was used to  
280 rule out reverse causation by excluding participants with diseases at baseline such as  
281 hypertension, diabetes, obesity, and LC cases that occurred during the first five years of  
282 follow-up. The results generally unchanged but participants in the highest quintile of UPF  
283 intake compared to the lowest had a 5% of higher risk of LC incidence when LC cases  
284 occurred in the first five years of follow-up were excluded (HR = 1.05; 95% CI: 0.83, 1.32;  
285 p-trend = 0.047) (**Supplementary Table 3**). The association did not also differ by the  
286 histopathological subtypes of LC. However, a 10% increase in the UPF intake (in %  
287 kcal/day) increased risk of LC by 32% (HR = 1.32; 1.15,1.52), a linear dose-response  
288 relationship (p-nonlinear = 0.82) (**Supplementary Table 3, Supplementary Fig 4**).  
289 Furthermore, when adjusted for artificial sweetening beverages, consumption of UPF was  
290 significantly increased the risk of LC by 53% (HR =1.53; 95% CI: 1.26, 1.86; p-trend  
291 =0.032). However, the risk of lung cancer was decreased by 2% among participants in the  
292 highest quintile of UPF intake compared to lowest when adjusted for dietary fiber intake (HR  
293 = 0.98; 95% CI: 0.83, 1.15; p-trend =0.864) (**Supplementary Table 3**).

## 294 **Discussion**

295 In this multi-center prospective study, we investigated the association between UPF intake in  
296 percentage weight (%g/day) and the risk of LC incidence among adults participating in

297 PLCO trial in the US. The finding of this study did not demonstrate link between a higher  
298 proportion of UPF in the total diet (%g/day) and risk of LC across all participants. However,  
299 during subgroup analysis, the study showed that the risk of LC was twice as high among  
300 adults in the highest quintile of UPF intake compared to study participants without diabetes.  
301 Further analysis using Cox proportional model on additive scale showed the effect of  
302 interaction between UPF and diabetes is, in fact super-additive (where both RERI and AP are  
303 found to be greater than zero and SI is greater than one). Quantification of the public health  
304 burden attributed to the interaction between quintiles of UPF consumption and diabetes  
305 revealed a total of 201 additional incident LC cases per 100,000 person-years due to these  
306 joint effects.  
307 Additionally, the risk of LC was 21% higher in participants within the highest quintile of  
308 UPF intake according to total energy (% kcal/day) consumption compared to participants in  
309 the lowest quintile.

310

311 A few studies from EPIC, UK Biobank, and a meta-analysis have investigated a positive  
312 association between UPF consumption and overall cancer risk [11-13]. However, the  
313 association between UPF consumption and the risk of LC has not yet been well established,  
314 despite LC being the second most commonly diagnosed cancer globally, following breast  
315 cancer [40]. A national representative study from the UK demonstrated a borderline  
316 statistically non-significant higher risk (25%) of LC among people at the highest level of  
317 UPF intake compared to lowest [12]. In contrast, a study from European countries, EPIC  
318 study showed that there was a non-significant inverse association between UPF consumption  
319 and LC, with a 4% lower risk of LC among subjects at the highest level of UPF consumption  
320 [13].

321 The current study revealed a positive association between LC incidence and a higher  
322 consumption of UPF (% gm/day) only among participants who had diabetes at baseline.  
323 When UPF consumption was measured in terms of energy ratio (% kcal/day), the risk of LC  
324 increased among participants with highest level of UPF consumption.

325 The underlying mechanism for the observed association may be explained by the interplay  
326 between hyperglycaemia and central obesity-cancer linkage, resulted from insulin-resistance  
327 and insulin-like growth factor 1, and adipokine pathophysiology and systemic inflammation  
328 from circulating pro-inflammatory cytokines interleukin-6,8 and 1 $\beta$ , tumour necrosis factor-  
329  $\alpha$ , vascular endothelial growth factor, chemokine, ligand 2 and interferon [41-43]. A growing  
330 body of evidence supports the idea that abdominal obesity, which can be measured by waist  
331 circumference or waist-to-hip ratio, and metabolic dysregulation, such as hyperglycaemia,  
332 insulin resistance, and dyslipidaemia, may be the underlying biological mechanisms that explain  
333 the link between consumption of UPF and the risk of LC. National representative prospective  
334 studies from Korea, the USA and the UK have emphasized the significance of considering  
335 metabolic status and markers when it comes to the primary prevention of lung cancer and the  
336 identification of high-risk populations for lung cancer screening [44-46].

337 The direct association of UPF consumption and risk of obesity, type 2 diabetes, and CVD are  
338 sufficiently well documented [8, 9, 47] and type 2 diabetes increases the risk of LC [48-50].  
339 Studies showed that a 10 cm increase in waist circumference had 10% higher risk of LC [51,  
340 52]. Central adiposity affects overall survival in LC and adiposity is inversely associated  
341 with inflammatory genes which may downregulate the anti-tumour immune response[53, 54].  
342 Metabolic effect of central obesity encompasses deranged insulin signalling, increased steroid  
343 hormone signalling, increased glucose utilization, fatty acid utilization and aberrant adipokine  
344 signalling [55]. Generally, dietary patterns with higher UPF proportion are nutritionally  
345 inferior and high in energy, fat, hydrogenated fat, free sugars, and low dietary fiber, that all

346 promote pro-inflammatory response, oxidative stress, metabolic dysregulation, and  
347 carcinogenicity [27, 44, 45]. This study highlighted that the observed association may not be  
348 affected by residual confounding, mainly by a history of cigarette smoking status, intensity,  
349 pipe, and cigar smoking. No significant subgroup difference was detected between UPF  
350 intake and risk of lung cancer by smoking status, BMI, family history of lung cancer and  
351 subtypes of lung cancer.

352 Moreover, during food processing alteration in food matrices contribute to degradation of  
353 essential nutrients involved in promoting health and change in microbiota [12, 21, 56].

354 Though evidence remains inconclusive, the mechanism for lung carcinogenesis in relation to  
355 UPF intake may be implicated due to metabolic disorders such as hyperglycaemia, insulin-  
356 resistance, dyslipidaemia, and inflammation due to altered levels of insulin-like growth  
357 factors, adipokines, myokines and sex hormones [57-59].

358 This study also determined whether the observed association between UPF intake (in %g/day,  
359 and %kcal/day and the risk of LC is altered by certain confounders such as age, sex, details of  
360 smoking status, baseline comorbidity, BMI, alcohol consumption, histological subtype,  
361 follow-up length, dietary risk factors such as fruits and vegetables, processed meat, artificial  
362 sweetening beverages, fish, fiber, sodium, total fats and other nutrients not largely changed.

363 The risk of LC was higher among participants with NSCLC subtype and after excluding LC  
364 cases developed during the first five years of follow-up with level of UPF consumption  
365 increases. Adjusting for dietary fiber reduced the risk of LC, non-statistically significant  
366 while adjusting for artificial sweetening beverages significantly increased the risk. The  
367 association was not largely changed for other covariates.

368 Our finding using the weight contribution of UPF consumption showed no significant  
369 association with the risk of LC across all study participants, contrary to the association  
370 between caloric contribution and LC. This could be due to fact that the energy contribution of

371 UPF can directly impact metabolic regulations and inducing oxidative stress than the weight  
372 contribution of UPF which comprises non-energy yielding constituents such as water, and  
373 artificial sweeteners [13, 33, 44, 45]. A well-designed, longitudinal study with adequate  
374 number of cases that incorporate biomarkers for central obesity, oxidative stress,  
375 inflammation, and additives and other non-nutritional factors, may help in reaching a  
376 definitive conclusion and provide a greater understanding of the biological mechanism(s)  
377 underlying the UPF consumption and LC association.

378 This study has important limitations. The NOVA grouping was performed using a single  
379 dietary data up to 25 years ago as a measure for usual dietary exposure. Since then, foods in  
380 the market have been continuously changed and UPF has become more dominate than  
381 minimally or unprocessed food groups in the American food system. Thus, the caloric and  
382 weight contribution of UPF may be underestimated and could not represent the current UPF  
383 consumption pattern in America. Recall bias from dietary assessment and misclassification  
384 bias from categorization of food items could not be ruled out even if a validated FFQ and  
385 NOVA food classification were used. The findings of the current study could only be inferred  
386 for the general aged 55 years and over and the follow-up duration is relatively short (12  
387 years). The study did not investigate potential biomarkers that may mediate the observed  
388 association between UPF and LC such as measures of body fat composition, inflammatory  
389 factors and non-nutritional factors associated with UPF consumption (additives, preservatives  
390 and neofomed contaminants) and potential confounding factors related to environmental or  
391 occupational exposures.

## 392 **Conclusions**

393 In summary, our findings showed that adults with diabetes and a higher consumption of UPF  
394 had a disproportionately higher risk of LC. In addition, a higher caloric contribution of UPF  
395 was positively associated with a higher risk of LC among all study participants. Given the



396 accumulating evidence relating to the adverse effect of UPF intake on overall cancer risk,  
397 consideration of the degree of food processing and integrating public health policies that  
398 promote healthy foods may be critical in reducing the modifiable burden of LC. Limiting a  
399 higher level of UPF consumption and regulation of metabolic biomarkers could prevent a  
400 higher number of LC cases. Additionally, further research is needed to better understand the  
401 relationship between UPF consumption and the incidence of LC.

#### 402 **Acknowledgements**

403 TCM is grateful to thank for support provided by the Australian Government Research  
404 Training Program Scholarship. The authors thank the National Cancer Institute for access to  
405 NCI's data collected by the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening  
406 Trial.

#### 407 **Statement of authors' contributions to manuscript.**

408 T.C.M: involved in conceptualization, data curation, formal analysis, investigation,  
409 methodology, project administration, resources, software, validation, visualization, writing  
410 the original draft, review & editing of the manuscript. Y.A.M, Z.S and T.K.G involved in  
411 conceptualization; methodology, supervision, validation, visualization, review and editing of  
412 the final version of the manuscript. All authors have read and approved the final manuscript.

413 Data Sharing. The PLCO trial data are available upon request to The National Cancer  
414 Institute at [Access to PLCO Data, Images, and Biospecimens - Learn - PLCO - The Cancer](#)  
415 [Data Access System](#).

416 **Funding:** This research did not receive any specific grant from funding agencies in the  
417 public, commercial, or not-for-profit sectors.

418

#### 419 **References**

- 420 1. Zhang, J., et al., *Global burden of lung cancer: implications from current evidence*. *Annals of*  
421 *Cancer Epidemiology*, 2021. **5**.
- 422 2. Ebrahimi, H., et al., *Global, regional, and national burden of respiratory tract cancers and*  
423 *associated risk factors from 1990 to 2019: a systematic analysis for the Global Burden of*  
424 *Disease Study 2019*. *The Lancet Respiratory Medicine*, 2021. **9**(9): p. 1030-1049.
- 425 3. Fitzmaurice, C., et al., *Global, Regional, and National Cancer Incidence, Mortality, Years of*  
426 *Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups,*  
427 *1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study*. *JAMA Oncol*,  
428 2019. **5**(12): p. 1749-1768.
- 429 4. Prevention, C.f.D.C.a. *An Update on Cancer Deaths in the United States*. Atlanta, GA: US  
430 *Department of Health and Human Services, Centers for Disease Control and Prevention,*  
431 *Division of Cancer Prevention and Control*. 2022; Available from:  
432 <https://www.cdc.gov/cancer/dcpc/research/update-on-cancer-deaths/index.htm>.
- 433 5. *GBD 2015 Risk Factors Collaborators*. *Global, regional, and national comparative risk*  
434 *assessment of 79 behavioural, environmental and occupational, and metabolic risks or*  
435 *clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study*  
436 *2015*. *Lancet*, 2016. **388**(10053): p. 1659-1724.
- 437 6. Lv, D., et al., *Fish Intake, Dietary Polyunsaturated Fatty Acids, and Lung Cancer: Systematic*  
438 *Review and Dose-Response Meta-Analysis of 1.7 Million Men and Women*. *Nutr Cancer*,  
439 2022. **74**(6): p. 1976-1985.
- 440 7. Juul, F., et al., *Ultra-processed food consumption among US adults from 2001 to 2018*. *Am J*  
441 *Clin Nutr*, 2022. **115**(1): p. 211-221.
- 442 8. Suksatan, W., et al., *Ultra-Processed Food Consumption and Adult Mortality Risk: A*  
443 *Systematic Review and Dose-Response Meta-Analysis of 207,291 Participants*. *Nutrients*,  
444 2021. **14**(1).

- 445 9. Schnabel, L., et al., *Association Between Ultraprocessed Food Consumption and Risk of*  
446 *Mortality Among Middle-aged Adults in France*. *JAMA Intern Med*, 2019. **179**(4): p. 490-498.
- 447 10. Liu, F.-H., et al., *Dietary Inflammatory Index and Health Outcomes: An Umbrella Review of*  
448 *Systematic Review and Meta-Analyses of Observational Studies*. *Frontiers in Nutrition*, 2021.  
449 **8**.
- 450 11. Isaksen, I.M. and S.N. Dankel, *Ultra-processed food consumption and cancer risk: A*  
451 *systematic review and meta-analysis*. *Clinical Nutrition*, 2023. **42**(6): p. 919-928.
- 452 12. Chang, K., et al., *Ultra-processed food consumption, cancer risk and cancer mortality: a large-*  
453 *scale prospective analysis within the UK Biobank*. *eClinicalMedicine*, 2023. **56**.
- 454 13. Kliemann, N., et al., *Food processing and cancer risk in Europe: results from the prospective*  
455 *EPIC cohort study*. *The Lancet Planetary Health*, 2023. **7**(3): p. e219-e232.
- 456 14. Wang, L., et al., *Association of ultra-processed food consumption with colorectal cancer risk*  
457 *among men and women: results from three prospective US cohort studies*. *BMJ*, 2022. **378**: p.  
458 e068921.
- 459 15. Fiolet, T., et al., *Consumption of ultra-processed foods and cancer risk: results from NutriNet-*  
460 *Santé prospective cohort*. *BMJ*, 2018. **360**: p. k322.
- 461 16. Zhong, G.-C., et al., *Ultra-processed food consumption and the risk of pancreatic cancer in*  
462 *the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial*. *International Journal of*  
463 *Cancer*, 2023. **152**(5): p. 835-844.
- 464 17. Romaguera, D., et al., *Consumption of ultra-processed foods and drinks and colorectal,*  
465 *breast, and prostate cancer*. *Clin Nutr*, 2021. **40**(4): p. 1537-1545.
- 466 18. Bradbury, K.E., N. Murphy, and T.J. Key, *Diet and colorectal cancer in UK Biobank: a*  
467 *prospective study*. *International Journal of Epidemiology*, 2019. **49**(1): p. 246-258.
- 468 19. Martínez Steele, E., et al., *Association between dietary contribution of ultra-processed foods*  
469 *and urinary concentrations of phthalates and bisphenol in a nationally representative sample*  
470 *of the US population aged 6 years and older*. *PLOS ONE*, 2020. **15**(7): p. e0236738.

- 471 20. Nerín, C., M. Aznar, and D. Carrizo, *Food contamination during food process*. Trends in Food  
472 Science and Technology, 2016. **48**: p. 63-68.
- 473 21. Viennois, E., et al., *Dietary Emulsifier-Induced Low-Grade Inflammation Promotes Colon*  
474 *Carcinogenesis*. Cancer Res, 2017. **77**(1): p. 27-40.
- 475 22. Clarke, R.E., et al., *Dietary Advanced Glycation End Products and Risk Factors for Chronic*  
476 *Disease: A Systematic Review of Randomised Controlled Trials*. Nutrients, 2016. **8**(3): p. 125.
- 477 23. NTP (National Toxicology Program). *Report on Carcinogens, Fifteenth Edition*. Research  
478 Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service.  
479 <https://ntp.niehs.nih.gov/go/roc15> (EndNote XML) DOI: <https://doi.org/10.22427/NTP->  
480 [OTHER-1003](https://doi.org/10.22427/NTP-OTHER-1003). 2021.
- 481 24. Banerjee, A., S. Mukherjee, and B.K. Maji, *Worldwide flavor enhancer monosodium*  
482 *glutamate combined with high lipid diet provokes metabolic alterations and systemic*  
483 *anomalies: An overview*. Toxicol Rep, 2021. **8**: p. 938-961.
- 484 25. Prorok, P.C., et al., *Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer*  
485 *Screening Trial*. Control Clin Trials, 2000. **21**(6 Suppl): p. 273s-309s.
- 486 26. Subar, A.F., et al., *Comparative validation of the Block, Willett, and National Cancer Institute*  
487 *food frequency questionnaires : the Eating at America's Table Study*. Am J Epidemiol, 2001.  
488 **154**(12): p. 1089-99.
- 489 27. Monteiro, C.A., et al., *Ultra-processed foods: what they are and how to identify them*. Public  
490 Health Nutr, 2019. **22**(5): p. 936-941.
- 491 28. Thibault, F., et al., *Consumption of ultra-processed foods and cancer risk: results from*  
492 *NutriNet-Santé prospective cohort*. BMJ, 2018. **360**: p. k322.
- 493 29. Robins, J.M. and D.M. Finkelstein, *Correcting for noncompliance and dependent censoring in*  
494 *an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests*.  
495 Biometrics, 2000. **56**(3): p. 779-88.

- 496 30. Willems, S., et al., *Correcting for dependent censoring in routine outcome monitoring data by*  
497 *applying the inverse probability censoring weighted estimator*. *Stat Methods Med Res*, 2018.  
498 **27**(2): p. 323-335.
- 499 31. Vakulenko-Lagun, B., M. Mandel, and R.A. Betensky, *Inverse probability weighting methods*  
500 *for Cox regression with right-truncated data*. *Biometrics*, 2020. **76**(2): p. 484-495.
- 501 32. Dong, G., et al., *The inverse-probability-of-censoring weighting (IPCW) adjusted win ratio*  
502 *statistic: an unbiased estimator in the presence of independent censoring*. *J Biopharm Stat*,  
503 2020. **30**(5): p. 882-899.
- 504 33. Llaveró-Valero, M., et al., *Ultra-processed foods and type-2 diabetes risk in the SUN project:*  
505 *A prospective cohort study*. *Clinical Nutrition*, 2021. **40**(5): p. 2817-2824.
- 506 34. Cordova, R., et al., *Consumption of ultra-processed foods and risk of multimorbidity of cancer*  
507 *and cardiometabolic diseases: a multinational cohort study*. *Lancet Reg Health Eur*, 2023. **35**:  
508 p. 100771.
- 509 35. Rothman, K.J., S. Greenland, and T.L. Lash, *Modern epidemiology*. Vol. 3. 2008: Wolters  
510 Kluwer Health/Lippincott Williams & Wilkins Philadelphia.
- 511 36. Correia, K. and P.L. Williams, *Estimating the Relative Excess Risk Due to Interaction in*  
512 *Clustered-Data Settings*. *American Journal of Epidemiology*, 2018. **187**(11): p. 2470-2480.
- 513 37. Knol, M.J., et al., *Estimating measures of interaction on an additive scale for preventive*  
514 *exposures*. *European Journal of Epidemiology*, 2011. **26**(6): p. 433-438.
- 515 38. Rod, N.H., et al., *Additive Interaction in Survival Analysis: Use of the Additive Hazards Model*.  
516 *Epidemiology*, 2012. **23**(5).
- 517 39. Turner, M.C., et al., *Interactions between cigarette smoking and ambient PM2.5 for*  
518 *cardiovascular mortality*. *Environmental Research*, 2017. **154**: p. 304-310.
- 519 40. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and*  
520 *Mortality Worldwide for 36 Cancers in 185 Countries*. *CA: A Cancer Journal for Clinicians*,  
521 2021. **71**(3): p. 209-249.

- 522 41. Brennan, P., et al., *Obesity and cancer: Mendelian randomization approach utilizing the FTO*  
523 *genotype*. Int J Epidemiol, 2009. **38**(4): p. 971-5.
- 524 42. Robado de Lope, L., et al., *Tumour-adipose tissue crosstalk: fuelling tumour metastasis by*  
525 *extracellular vesicles*. Philos Trans R Soc Lond B Biol Sci, 2018. **373**(1737).
- 526 43. Samet, J.M., *Lung Cancer, Smoking, and Obesity: It's Complicated*. JNCI: Journal of the  
527 National Cancer Institute, 2018. **110**(8): p. 795-796.
- 528 44. Sin, S., et al., *Metabolic Syndrome and Risk of Lung Cancer: An Analysis of Korean National*  
529 *Health Insurance Corporation Database*. J Clin Endocrinol Metab, 2020. **105**(11).
- 530 45. Li, M., et al., *Association of Metabolic Syndrome With Risk of Lung Cancer: A Population-*  
531 *Based Prospective Cohort Study*. CHEST, 2024. **165**(1): p. 213-223.
- 532 46. Yuan, Z., et al., *Association between metabolic overweight/obesity phenotypes and*  
533 *readmission risk in patients with lung cancer: A retrospective cohort study*. eClinicalMedicine,  
534 2022. **51**.
- 535 47. Tobias, D.K., et al., *Body-mass index and mortality among adults with incident type 2*  
536 *diabetes*. N Engl J Med, 2014. **370**(3): p. 233-44.
- 537 48. Tseng, C.H., *Diabetes but not insulin increases the risk of lung cancer: a Taiwanese*  
538 *population-based study*. PLoS One, 2014. **9**(7): p. e101553.
- 539 49. Tseng, C.H., *Higher risk of mortality from lung cancer in Taiwanese people with diabetes*.  
540 Diabetes Res Clin Pract, 2013. **102**(3): p. 193-201.
- 541 50. Pearson-Stuttard, J., et al., *Type 2 Diabetes and Cancer: An Umbrella Review of*  
542 *Observational and Mendelian Randomization Studies*. Cancer Epidemiol Biomarkers Prev,  
543 2021. **30**(6): p. 1218-1228.
- 544 51. Yu, D., et al., *Overall and Central Obesity and Risk of Lung Cancer: A Pooled Analysis*. J Natl  
545 Cancer Inst, 2018. **110**(8): p. 831-842.
- 546 52. Hidayat, K., et al., *Abdominal Obesity and Lung Cancer Risk: Systematic Review and Meta-*  
547 *Analysis of Prospective Studies*. Nutrients, 2016. **8**(12).

- 548 53. Barbi, J., et al., *Visceral Obesity Promotes Lung Cancer Progression-Toward Resolution of the*  
549 *Obesity Paradox in Lung Cancer*. J Thorac Oncol, 2021. **16**(8): p. 1333-1348.
- 550 54. Bhupathiraju, S.N. and F.B. Hu, *Epidemiology of Obesity and Diabetes and Their*  
551 *Cardiovascular Complications*. Circ Res, 2016. **118**(11): p. 1723-35.
- 552 55. Khandekar, M.J., P. Cohen, and B.M. Spiegelman, *Molecular mechanisms of cancer*  
553 *development in obesity*. Nat Rev Cancer, 2011. **11**(12): p. 886-95.
- 554 56. Fardet, A. and E. Rock, *Chronic diseases are first associated with the degradation and*  
555 *artificialization of food matrices rather than with food composition: calorie quality matters*  
556 *more than calorie quantity*. Eur J Nutr, 2022. **61**(5): p. 2239-2253.
- 557 57. Badoud, F., et al., *Molecular insights into the role of white adipose tissue in metabolically*  
558 *unhealthy normal weight and metabolically healthy obese individuals*. Faseb j, 2015. **29**(3): p.  
559 748-58.
- 560 58. Karelis, A.D., et al., *Metabolic and body composition factors in subgroups of obesity: what do*  
561 *we know?* J Clin Endocrinol Metab, 2004. **89**(6): p. 2569-75.
- 562 59. Cao, H., et al., *Association between circulating levels of IGF-1 and IGFBP-3 and lung cancer*  
563 *risk: a meta-analysis*. PLoS One, 2012. **7**(11): p. e49884.

564  
565

## 566 **Figure Legends**

567 Figure 1: The relative contribution of food groups to the total diet (a) and to the subtotal of  
568 UPF consumption in men (b) and women (c) in the PLCO trial data.

569 Figure 2: Non-linear relationship of UPF consumption in weight ratio and lung cancer  
570 incidence among older adults participated in the PLCO trial, USA.

571 Figure 3: Subgroup analysis for the association between UPF consumption and risk of lung  
572 cancer incident stratified by demographic, lifestyle and baseline comorbidities in  
573 older adults enrolled in the PLCO trial, the in USA]. The model was adjusted for

574 covariates mentioned in Table 2 footnote. BMI: body mass index; HR: hazard

575 ratio; UPF: ultra-processed foods.

576

577

578

579



Table 1: Distribution of baseline characteristics of older adults enrolled in the PLCO trial across quintiles of UPF consumption in USA (n=96607).

Baseline features	Overall	Proportion of UPF (% weight in the diet)					p-trend
		Q1	Q2	Q3	Q4	Q5	
		0 to 18.8	18.81 to 26.3	26.31 to 33.8	33.81 to 42.7	42.71 to 99	
No.	96607	19622	19440	19482	19552	18511	
<b>Sex = Females, n (%)</b>	50803 (52.6)	13044 (66.5)	10872 (55.9)	9838 (50.5)	8991 (46.0)	8058 (43.5)	<0.001
<b>Trial arm= Controls, n (%)</b>	47428 (49.1)	9563 (48.7)	9545 (49.1)	9551 (49.0)	9507 (48.6)	9262 (50.0)	0.053
<b>Ethnicity, n (%)</b>							<0.001
Non-Hispanic White	88013 (91.1)	17352 (88.4)	17708 (91.1)	18029 (92.5)	18105 (92.6)	16819 (90.9)	
Non-Hispanic Black	3061 (3.2)	485 (2.5)	552 (2.8)	554 (2.8)	597 (3.1)	873 (4.7)	
Hispanic	1374 (1.4)	321 (1.6)	290 (1.5)	250 (1.3)	255 (1.3)	258 (1.4)	
Asian, Pacific Islander and American Indian	4124 (4.3)	1453 (7.4)	886 (4.6)	647 (3.3)	585 (3.0)	553 (3.0)	

**Table 1:** Distribution of baseline characteristics of older adults enrolled in the PLCO trial across quintiles of UPF consumption in USA

(continued)

<b>Age, n (%)</b>								<0.001
	55-59	17011 (17.6)	3081 (15.7)	3130 (16.1)	3223 (16.5)	3396 (17.4)	4181 (22.6)	
	60-64	27232 (28.2)	5257 (26.8)	5246 (27.0)	5372 (27.6)	5585 (28.6)	5772 (31.2)	
	65-69	26188 (27.1)	5517 (28.1)	5414 (27.9)	5418 (27.8)	5306 (27.1)	4533 (24.5)	
	70-83	26173 (27.1)	5767 (29.4)	5649 (29.1)	5468 (28.1)	5264 (26.9)	4025 (21.7)	
	Mean (SD) (years)	65.6 (5.7)	66.0 (5.7)	65.9 (5.7)	65.8 (5.7)	65.6 (5.7)	64.6 (5.7)	<0.001
<b>Education, n (%)</b>								<0.001
	Up to high school or less	27987 (29.0)	4985 (25.5)	5352 (27.6)	5696 (29.3)	5956 (30.5)	5998 (32.4)	
	Post-high school training	33139 (34.4)	6605 (33.7)	6497 (33.5)	6655 (34.2)	6722 (34.5)	6660 (36.0)	
	College graduate	35296 (36.6)	7996 (40.8)	7559 (38.9)	7089 (36.5)	6823 (35.0)	5829 (31.5)	
<b>Occupation, n (%)</b>								<0.001
	Homemaker	11592 (12.1)	2900 (14.9)	2590 (13.4)	2344 (12.1)	2140 (11.0)	1618 (8.8)	
	Employed	38340 (39.9)	7325 (37.5)	7349 (38.0)	7398 (38.2)	7926 (40.7)	8342 (45.2)	
	Retired	41624 (43.3)	8437 (43.2)	8582 (44.3)	8733 (45.0)	8457 (43.5)	7415 (40.2)	
	Others	4617 (4.8)	862 (4.4)	837 (4.3)	916 (4.7)	940 (4.8)	1062 (5.8)	

**Table 1:** Distribution of baseline characteristics of older adults enrolled in the PLCO trial across quintiles of UPF consumption in USA (continued)

<b>Marital status, n (%)</b>							<b>&lt;0.001</b>
Married	75771 (78.6)	14781 (75.5)	15345 (79.1)	15644 (80.5)	15699 (80.5)	14302 (77.4)	
Widowed	7837 (8.1)	1915 (9.8)	1592 (8.2)	1464 (7.5)	1446 (7.4)	1420 (7.7)	
Divorced	9095 (9.4)	2019 (10.3)	1737 (8.9)	1646 (8.5)	1700 (8.7)	1993 (10.8)	
Separated/never married	3732 (3.9)	872 (4.5)	734 (3.8)	690 (3.5)	666 (3.4)	770 (4.2)	
<b>BMI at baseline, n (%)</b>							<b>&lt;0.001</b>
<18.5	648 (0.7)	217 (1.1)	125 (0.7)	114 (0.6)	111 (0.6)	81 (0.4)	
18.5-24.99	32585 (34.2)	8285 (42.8)	7025 (36.6)	6427 (33.4)	5897 (30.6)	4951 (27.1)	
25-29.99	40538 (42.5)	7331 (37.9)	8077 (42.1)	8410 (43.7)	8564 (44.4)	8156 (44.6)	
≥30	21592 (22.6)	3510 (18.1)	3979 (20.7)	4299 (22.3)	4722 (24.5)	5082 (27.8)	
<b>Cigarette smoking, n (%)</b>							<b>&lt;0.001</b>
Never smoker	46560 (48.2)	9877 (50.3)	9599 (49.4)	9357 (48.0)	9437 (48.3)	8290 (44.8)	
Current smoker	8617 (8.9)	1239 (6.3)	1476 (7.6)	1647 (8.5)	1888 (9.7)	2367 (12.8)	
Former smoker	41417 (42.9)	8502 (43.3)	8363 (43.0)	8476 (43.5)	8222 (42.1)	7854 (42.4)	

**Table 1:** Distribution of baseline characteristics of older adults enrolled in the PLCO trial across quintiles of UPF consumption in USA

(continued)

<b>Cigar smoking, n (%)</b>							<b>&lt;0.001</b>
Never smoker	84220 (88.0)	17658 (91.1)	17047 (88.5)	16862 (87.3)	16846 (87.0)	15807 (86.1)	
Current smoker	1518 (1.6)	217 (1.1)	276 (1.4)	319 (1.7)	345 (1.8)	361 (2.0)	
Former smoker	9930 (10.4)	1518 (7.8)	1931 (10.0)	2128 (11.0)	2171 (11.2)	2182 (11.9)	
<b>Pipe smoking, n (%)</b>							<b>&lt;0.001</b>
Never smoker	82588 (86.2)	17410 (89.7)	16769 (87.0)	16420 (85.0)	16429 (84.8)	15560 (84.7)	
Current smoker	837 (0.9)	119 (0.6)	132 (0.7)	176 (0.9)	211 (1.1)	199 (1.1)	
Former smoker	12340 (12.9)	1882 (9.7)	2382 (12.4)	2723 (14.1)	2745 (14.2)	2608 (14.2)	
<b>Alcohol drinking status, n (%)</b>							<b>&lt;0.001</b>
Never drinker	9667 (10.0)	2053 (10.5)	1776 (9.1)	1793 (9.2)	2046 (10.5)	1999 (10.8)	
Current drinker	13814 (14.3)	2611 (13.3)	2429 (12.5)	2512 (12.9)	2805 (14.3)	3457 (18.7)	
Former drinker	70410 (72.9)	14337 (73.1)	14704 (75.6)	14666 (75.3)	14198 (72.6)	12505 (67.6)	
Unknown status	2716 (2.8)	621 (3.2)	531 (2.7)	511 (2.6)	503 (2.6)	550 (3.0)	

**Table 1:** Distribution of baseline characteristics of older adults enrolled in the PLCO trial across quintiles of UPF consumption in USA

(continued)

<b>Family history of lung cancer (%)</b>							<b>&lt;0.001</b>
No	83499 (87.1)	16992 (87.2)	16804 (87.1)	16939 (87.6)	16927 (87.3)	15837 (86.2)	
Yes, family members	10058 (10.5)	2060 (10.6)	2038 (10.6)	1973 (10.2)	1976 (10.2)	2011 (10.9)	
Yes, relatives or unclear	2311 (2.4)	431 (2.2)	444 (2.3)	432 (2.2)	480 (2.5)	524 (2.9)	
cancer types							
Chronic Bronchitis (%)	4070 (4.2)	807 (4.1)	810 (4.2)	793 (4.1)	840 (4.3)	820 (4.4)	0.413
Diabetes (%)	6327 (6.6)	1212 (6.2)	1196 (6.2)	1195 (6.2)	1284 (6.6)	1440 (7.8)	<0.001
Emphysema (%)	1983 (2.1)	325 (1.7)	343 (1.8)	402 (2.1)	475 (2.4)	438 (2.4)	<0.001
Heart attack (%)	7968 (8.3)	1415 (7.3)	1606 (8.3)	1677 (8.6)	1626 (8.4)	1644 (8.9)	<0.001
Hypertension (%)	31174 (32.4)	5869 (30.1)	6125 (31.7)	6359 (32.8)	6443 (33.1)	6378 (34.6)	<0.001
Stroke (%)	1916 (2.0)	365 (1.9)	372 (1.9)	377 (1.9)	387 (2.0)	415 (2.3)	0.073

Other includes unemployed, disabled, and extended sick leave; SD: standard deviation.

Table 2: Cox proportional hazard model with time-varying covariate for the association between proportion of UPF consumption (% gm/day and risk of lung cancer incidence among older population of PLCO trial (Complete cases = 91,453).

Explanatory variables	Proportion of UPF consumption in the diet					UPF as continuous exposure		
	HR (95% CI)					(10% increment)		
	Q1	Q2	Q3	Q4	Q5	p-trend	HR (95% CI)	p-nonlinear
Cases	282	280	311	321	323		1517	
Person-years	161583.5	163458.8	164101.1	164313.0	153729.7		807186.2	
Cumulative Incidence <sup>a</sup>	17.5(15.5, 19.6)	17.1(15.2, 19.3)	18.9(16.9,21.2)	19.5(17.5, 21.8)	21.0(18.8, 23.4)		18.8(17.9,19.8)	
Model 1	Ref	0.96(0.81,1.12)	1.06(0.91,1.24)	1.10(0.95, 1.28)	1.20(1.02,1.39) *	0.07	1.12(0.97, 1.29) *	0.57
Model 2	Ref	0.92(0.78,1.08)	1.01(0.86,1.18)	1.05(0.89, 1.22)	1.19(1.01,1.39) *	0.08	1.09(0.95, 1.26) ***	0.08
Model 3	Ref	0.91(0.77,1.06)	0.95(0.81,1.12)	0.95(0.81, 1.11)	0.98(0.83,1.15)	0.33	0.99(0.87, 1.15)	0.36
Model 4	Ref	0.91(0.77,1.07)	0.96(0.81,1.12)	0.94(0.81,1.11)	0.95(0.81,1.12)	0.30	0.99(0.86,1.14)	0.40

Model 5	Ref	0.92(0.78,1.08)	0.98(0.83,1.15)	0.96(0.82, 1.13)	0.98(0.83,1.15)	0.32	0.98(0.85, 1.13)	0.44
Competing risk Regression								
	Ref	0.92(0.78,1.09)	0.97(0.82,1.14)	0.95(0.81,1.12)	0.95(0.81,1.12)	0.35		

<sup>a</sup> cumulative incidence rate per 10,000 person-years; \* represents significant at p-value <0.05; and \*\*\*significant at p-linear < 0.001.

Model 1: unadjusted model.

Model 2: Adjusted for Age and sex.

Model 3: adjusted for model 2 plus study arm, race, marital status, education, occupation, smoking, BMI, family history of lung cancer and alcohol consumption.

Model 4: Adjusted for model 3 plus chronic comorbidity (diabetes, hypertension, heart attack, stroke, chronic bronchitis, and emphysema).

Model 5: Additionally adjusted for total dietary energy plus model 4 and excluding dietary alcohol consumption.

Table 3: Adjusted HRs (95% CI) and additive interaction between UPF consumption and diabetes status for the risk of LC in US (n = 91,453)<sup>a</sup>

Percentage of UPF intake <sup>b</sup>	Subjects without diabetes		Subjects with diabetes		Diabetic vs non-diabetic in strata of %UPF		Interaction (Additive)		
	Cases/n	HR (95% CI)	Cases/n	HR (95% CI)	Cases/n	HR (95% CI)	RERI (95% CI)	AP (95% CI)	SI (95% CI)
Low	525/85,512	Ref	37/5,941	Ref	562/36,563	1.01(0.71,1.43)	0.13(-0.32,0.58)	0.10(-0.22,0.42)	1.67(0.22,12.46)
High	879/85,512	0.99(0.89,1.11)	76/5,941	1.17(0.79,1.75)	955/54,890	1.22(0.96,1.55)			

AP: Attributable Proportion; CI: Confidence Interval; HR: Hazard Ratio; LC: Lung Cancer; RERI: Relative Excess Risk due to Interaction; SI: Synergy Index; UPF: Ultra-processed Food.

<sup>a</sup> Models were adjusted for age, sex, ethnicity, study arm, marital status, education, occupation, cigarette smoking, BMI, family history of lung cancer, history of hypertension, heart attack, chronic bronchitis, emphysema, alcohol intake and total dietary calory intake.

<sup>b</sup> Percentage of UPF (% gm/day) categorised as high and low based on the mean intake of UPF which was  $\geq 31.2\%$  for high and  $< 31.2\%$  for low).



Table 4: Interaction between UPF consumption and diabetes for the risk of lung cancer among adults using additive hazard models in comparison with Cox proportional hazard models, US (n = 91,453).

	No. additional lung cancer cases per 100, 000 person-years <sup>a</sup> (95% CI)	Cox proportional hazard models, HR (95% CI)
High UPF vs low UPF <sup>b</sup>	6(-13,26)	1.02(0.91, 1.14)
Diabetes vs non-diabetes	14(-50, 77)	1.09(0.78, 1.53)
Stratified Analysis		
People with diabetes		
High UPF vs low UPF <sup>b</sup>	36(-47, 119)	1.12(0.76,1.65)
Q2 vs 1	113 (-12, 238)	1.79(0.93, 3.46)
Q3 vs 1	106 (-17, 229)	1.69(0.87, 3.25)
Q4 vs 1	-19(-120, 82)	0.91(0.43,1.89)
Q5 vs 1	186(59, 313) **	2.18(1.17, 4.04) *
People without diabetes		
High UPF vs low UPF <sup>b</sup>	7(13,26)	0.99(0.89,1.10)

	Q2 vs 1	-26(-54,3)	0.84(0.72, 1.01)
	Q3 vs 1	-9(-39, 21)	0.91(0.77,1.07)
	Q4 vs 1	-1(-32, 29)	0.93(0.79,1.09)
	Q5 vs 1	-9(-41,23)	0.91(0.77,1.07)
Joint effects			
	High UPF * diabetes	35(-50, 120)	1.16(0.78, 1.75)
Quintiles of UPF *diabetes			
	Q2* diabetes	144(16, 272) *	2.18(1.07,4.43) *
	Q3* diabetes	121(-5, 247)	1.97(0.96,4.01)
	Q4* diabetes	-12(-115,92)	0.99(0.44,2.17)
	Q5* diabetes	201(70, 332) **	2.44(1.27,4.67) **

<sup>a</sup> Models were adjusted for age, sex, race, study arm, marital status, education, occupation, family income, smoking, BMI, physical activity, family history of lung cancer, history of hypertension, heart attack, chronic bronchitis, emphysema, alcohol intake and total dietary calory intake.

<sup>b</sup> Percentage of UPF (%gm/day) categorised as high and low based on the mean intake of UPF which was  $\geq 31.2\%$  for high and  $< 31.2\%$  for low).

\* Indicates covariates significant at p-value < 0.05, and \*\* at p-value < 0.01.