



Ultra-processed foods, adiposity and risk of head and neck cancer and oesophageal adenocarcinoma in the European Prospective Investigation into Cancer and Nutrition study: a mediation analysis

Fernanda Morales-Berstein^{1,2} · Carine Biessy³ · Vivian Viallon³ · Ana Goncalves-Soares^{1,2} · Corinne Casagrande³ · Bertrand Hémon³ · Nathalie Kliemann^{3,4} · Manon Cairat^{3,5} · Jessica Blanco Lopez³ · Aline Al Nahas³ · Kiara Chang⁶ · Eszter Vamos⁶ · Fernanda Rauber^{7,8} · Renata Bertazzi Levy^{7,8} · Diana Barbosa Cunha⁹ · Paula Jakszyn^{10,11,12} · Pietro Ferrari³ · Paolo Vineis^{13,14} · Giovanna Masala¹⁵ · Alberto Catalano¹⁶ · Emily Sonestedt¹⁷ · Yan Borné¹⁷ · Verena Katzke¹⁸ · Rashmita Bajracharya¹⁸ · Claudia Agnoli¹⁹ · Marcela Guevara^{20,21,22} · Alicia Heath²³ · Loredana Radoi⁵ · Francesca Mancini⁵ · Elisabete Weiderpass²⁴ · José María Huerta^{21,25} · María-José Sánchez^{21,26,27,28} · Anne Tjønneland^{29,30} · Cecilie Kyrø²⁹ · Matthias B. Schulze^{31,32} · Guri Skeie³³ · Marko Lukic³³ · Tonje Braaten³³ · Marc Gunter^{3,23} · Christopher Millett^{6,34} · Antonio Agudo^{10,11} · Paul Brennan³⁵ · M. Carolina Borges^{1,2} · Rebecca C. Richmond^{1,2} · Tom G. Richardson^{1,2} · George Davey Smith^{1,2} · Caroline L. Relton^{1,2} · Inge Huybrechts³ · on behalf of the EPIC Network

Received: 25 May 2023 / Accepted: 10 October 2023 / Published online: 22 November 2023
© The Author(s) 2023

Abstract

Purpose To investigate the role of adiposity in the associations between ultra-processed food (UPF) consumption and head and neck cancer (HNC) and oesophageal adenocarcinoma (OAC) in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort.

Methods Our study included 450,111 EPIC participants. We used Cox regressions to investigate the associations between the consumption of UPFs and HNC and OAC risk. A mediation analysis was performed to assess the role of body mass index (BMI) and waist-to-hip ratio (WHR) in these associations. In sensitivity analyses, we investigated accidental death as a negative control outcome.

Results During a mean follow-up of 14.13 ± 3.98 years, 910 and 215 participants developed HNC and OAC, respectively. A 10% g/d higher consumption of UPFs was associated with an increased risk of HNC (hazard ratio [HR] = 1.23, 95% confidence interval [CI] 1.14–1.34) and OAC (HR = 1.24, 95% CI 1.05–1.47). WHR mediated 5% (95% CI 3–10%) of the association between the consumption of UPFs and HNC risk, while BMI and WHR, respectively, mediated 13% (95% CI 6–53%) and 15% (95% CI 8–72%) of the association between the consumption of UPFs and OAC risk. UPF consumption was positively associated with accidental death in the negative control analysis.

Conclusions We reaffirmed that higher UPF consumption is associated with greater risk of HNC and OAC in EPIC. The proportion mediated via adiposity was small. Further research is required to investigate other mechanisms that may be at play (if there is indeed any causal effect of UPF consumption on these cancers).

Keywords Epidemiology · Food processing · NOVA classification · Oesophageal cancer · Head and neck cancer · Adiposity · Mediation analysis

Abbreviations

BMI	Body mass index	FFQ	Food frequency questionnaire
CI	Confidence interval	HNC	Head and neck cancer
EPIC	European Prospective Investigation into Cancer and Nutrition	HR	Hazard ratio
		IARC	International Agency for Research on Cancer
		INHANCE	International Head and Neck Cancer Epidemiology
		MICE	Multivariate imputation by chained equations

Extended author information available on the last page of the article

OAC	Oesophageal adenocarcinoma
PM	Proportion mediated
PNDE	Pure natural direct effect
SD	Standard deviation
TE	Total effect
TNIE	Total natural indirect effect
UPF	Ultra-processed food
WHR	Waist-to-hip ratio

Introduction

In recent years, there has been growing interest in the potential role of industrial food processing in disease aetiology. The NOVA (not an abbreviation) classification system developed by Monteiro et al. [1, 2] categorises foods into four groups according to their degree and purpose of processing: (1) unprocessed or minimally processed foods, (2) processed culinary ingredients, (3) processed foods and (4) ultra-processed foods (UPFs). UPFs are industrial formulations manufactured in a complex way using ingredients not usually found in kitchens (e.g. maltodextrin, hydrogenated oils, modified starches) and cosmetic additives (e.g. emulsifiers, flavourings, colourants, artificial sweeteners) [2]. They are typically cheap, highly palatable, and widely available ready-to-eat products which are often consumed in large quantities, replacing more nutritious, unprocessed/minimally processed foods in the diet [3, 4]. Examples of UPFs include soft drinks, sweet or savoury packaged snacks, confectionery, packaged breads and buns, reconstituted meat products and pre-prepared frozen or shelf-stable dishes.

Several studies have shown that the consumption of UPFs may be associated with an increased risk of cancer [5–9]. In the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, Kliemann et al. [9] found positive associations between higher UPF consumption and the risk of head and neck cancer (HNC; hazard ratio [HR] = 1.14 per one standard deviation [SD] higher UPF intake, 95% confidence interval [CI] 1.06–1.24) and oesophageal adenocarcinoma (OAC; HR = 1.21 per 1-SD higher UPF intake, 95% CI 1.05–1.39). They also found an inverse association between UPF consumption and oesophageal squamous cell carcinoma risk (HR = 0.79 per 1-SD higher UPF intake, 95% CI 0.64–0.96), although this did not withstand additional adjustments for alcohol intake, body mass index (BMI) and several dietary factors (HR = 0.90 per 1-SD higher UPF intake, 95% CI 0.72–1.11).

UPF consumption has also been positively associated with higher adiposity (i.e. BMI, fat mass, waist circumference and waist-to-hip ratio [WHR]) [10–14]. Since body fatness

(measured by BMI, waist circumference and WHR) is an established modifiable risk factor for OAC [15–20], and visceral adiposity (i.e. waist circumference and WHR) has been positively associated with HNC risk [21–23], it is plausible that the positive associations between UPF consumption and these upper-aerodigestive tract cancers are mediated via adiposity. Although BMI has been inversely associated with HNC risk, this seems to be a consequence of residual confounding related to smoking (an established risk factor for HNC), as smokers tend to have lower BMIs than non-smokers [23]. In a meta-analysis of 20 prospective cohort studies, BMI was positively associated with HNC risk when the analysis was restricted to never smokers [22]. Although adiposity may be one of the mechanisms underlying the association between UPF consumption and upper-aerodigestive tract cancer, this has not been investigated using mediation analysis.

The aim of this study was to reassess and further investigate the associations between the consumption of UPFs and the risk of HNC and OAC in the EPIC study. As a complement to the study by Kliemann et al. [9] (described above), this study explored the associations between UPF consumption and the risk of HNC and its subtypes (i.e. oral cavity, oropharynx, hypopharynx, larynx and unspecified/overlapping cancers) as defined by the International Head and Neck Cancer Epidemiology (INHANCE) consortium. It also investigated effect modification by smoking status, alcohol intake, sex, physical activity, and education level in the associations between the consumption of UPFs and the risk of upper-aerodigestive tract cancers. Additionally, this study assessed the possibility of residual confounding using accidental death as a negative control outcome. Lastly, it examined the role of BMI and WHR in the associations between UPF consumption and the risk of HNC and OAC by means of a mediation analysis.

Methods

The EPIC cohort

The EPIC study has been fully described elsewhere [24–26]. Briefly, EPIC is one of the largest prospective cohort studies in Europe. It recruited 521,323 participants between 1992 and 2000. Participants were enrolled in 23 centres across 10 European countries, namely Denmark, France, Germany, Greece, Italy, the Netherlands, Norway, Spain, Sweden and the United Kingdom. Most were 35–69 years old at recruitment [24, 25]. They were either volunteers from the general population, blood donors, employees of

local companies, teachers/school employees or individuals enrolled in local ongoing studies. All participants provided written informed consent before completing the dietary and lifestyle questionnaires. Anthropometric and blood pressure data were also obtained at baseline. EPIC was approved by the International Agency for Research on Cancer (IARC) Ethics Committee and the local ethical review boards of all EPIC centres.

Study sample

Participants who withdrew consent from the study were not included in this research. We excluded participants diagnosed with cancer before enrolment ($n = 25,184$) and those with a length of follow-up equal to zero ($n = 4148$). We also excluded participants who did not complete the dietary or lifestyle questionnaires ($n = 6259$). We additionally excluded participants with extreme energy intake versus energy requirement ratios (top and bottom 1%) ($n = 9573$) and participants recruited in Greece due to administrative issues ($n = 26,048$). After exclusions, 450,111 participants were included in the analyses (Supplementary Fig. 1).

Dietary data and food processing variables

Semi-quantitative food frequency questionnaires (FFQs), extensive quantitative dietary questionnaires, and combined methods (i.e. semi-quantitative FFQs combined with 7-day records in the UK, and a non-quantitative FFQ combined with a 14-day record on hot meals in Malmö, Sweden) were used to obtain dietary data at baseline [25]. These were centre specific to account for local dietary habits and were either self-administered or administered in-person by trained interviewers. Furthermore, a standardised 24-h recall was used to obtain supplementary dietary data for a subsample of EPIC participants to calibrate baseline dietary measurements across EPIC centres [25, 27–30]. The dietary questionnaires and their mode of administration were described in detail in previous publications [25, 30].

The NOVA classification was used to categorise foods into four groups according to their extent and purpose of industrial processing [31]. Unprocessed/minimally processed foods (NOVA 1) are natural foods that may have undergone minimal processing for their preservation, storage, safety, or edibility. Processed culinary ingredients (NOVA 2) correspond to substances derived from unprocessed/minimally processed foods

(e.g. oil, butter) or nature (e.g. salt) that are normally consumed in combination with unprocessed/minimally processed foods. Both processed foods (NOVA 3) and UPFs (NOVA 4) are industrial products. The former typically contain two or three common ingredients (i.e. a combination of unprocessed/minimally processed foods and processed culinary ingredients), while the latter contain many ingredients (most of which are rarely used in kitchens) and additives that make the final product tastier and more attractive to consumers.

Food preparations made (at home or elsewhere) using traditional methods were decomposed using standardised recipes. Individual food items were then classified according to their degree of processing. Food items were combined into broader food categories for simplicity. Of a total of 67 food categories in the dietary questionnaires, 19 were classified as unprocessed/minimally processed foods, 5 as culinary ingredients, 13 as processed foods and 30 as UPFs (see Supplementary Table 1 for details).

Here, we used the relative intake of each NOVA group in grams per day (%g/d). We also used the absolute intake in grams per day (g/d) and the absolute and relative intake in kilocalories per day in sensitivity analyses (kcal/d and %kcal/d, respectively).

Ascertainment of cancer cases

Incident cancer cases were identified through population-based cancer registries in Denmark, Italy (except Naples), the Netherlands, Norway, Spain, Sweden and the United Kingdom. Participants in other centres (France, Germany, Greece and Naples) were actively followed up using health insurance records, pathology registries and direct contact with participants or their next of kin.

HNC and OAC were defined using the 2nd and 3rd Revision of the International Classification of Diseases for Oncology (ICDO-2 and ICDO-3). According to the INHANCE consortium [32], HNC cases include malignant neoplasms of the oral cavity (topography codes C00.3–C00.6, C00.8–C00.9, C02.0–C02.3, C03.0–C03.1, C03.9, C04.0–C04.1, C04.8–C04.9, C05.0, C06.0–C06.2, C06.8–C06.9), oropharynx (C01.9, C02.4, C05.1–C05.2, C09.0–C09.1, C09.8–C09.9, C10.0–C10.4, C10.8–C10.9), hypopharynx (C12.9–C13.2, C13.8–C13.9), larynx (C32.0–C32.3, C32.8–C32.9), and oral cavity and pharynx unspecified/overlapping regions (C02.8–C02.9, C05.8–C05.9, C14.0, C14.2, C14.8). We did not exclude any histological subtypes of HNC. Oesophageal cancer cases correspond

to topography codes C15.0–C15.5 and C15.8–C15.9. Among these, OAC cases were identified with codes 8140/3, 8144/3, 8480/3, 8481/3 and 8490/3. Other oesophageal cancer subtypes (e.g. squamous cell carcinoma and small cell carcinoma) were not investigated as outcomes in this study.

Covariates

Data on age at recruitment, sub-centre (22 centres in total, split into 27 sub-centres as follows: Northeast of France, Northwest of France, South of France, South coast of France, Florence, Varese, Ragusa, Turin, Naples, Asturias, Granada, Murcia, Navarra, San Sebastian, Cambridge, Oxford health-conscious population, Oxford general population, Bilthoven, Utrecht, Heidelberg, Potsdam, Malmö, Umeå, Aarhus, Copenhagen, Southeast of Norway, Northwest of Norway), sex (male/female), education level (none, primary, technical/professional, secondary, further education), physical activity based on the Cambridge Physical Activity Index [33] (inactive, moderately inactive, moderately active, active), measured/self-reported height (continuous in cm) and smoking status (never, former, current, unknown) were obtained at baseline through anthropometric measurements and lifestyle questionnaires. Additionally, data on alcohol intake (continuous in g/d) were acquired using dietary questionnaires.

Potential mediators

BMI and WHR were investigated as potential mediators in mediation analyses. BMI (continuous in kg/m²) was calculated from measured height and weight (measured using comparable, standardised methods) [34]. WHR (continuous) was estimated from measured waist and hip circumferences. Waist circumference was measured midway between the iliac crest and the lower ribs or at the narrowest torso circumference. Hip circumference was measured over the buttocks or at the widest point. EPIC-Oxford health-conscious population self-reported data were also used to estimate BMI and WHR, after the application of measurement error corrections [34, 35].

Statistical analysis

Descriptive characteristics

The participants' baseline characteristics were divided into sex-specific quartiles of relative UPF consumption (in

%g/d). Mean and SD estimates were obtained for continuous variables, while frequencies and percentages were obtained for binary/categorical variables. Furthermore, we made a histogram to graphically represent the distribution of UPF consumption (in %g/d) in the EPIC cohort.

Data imputation

We used single-value imputation to deal with missing data in the covariates used to control for potential confounding (i.e. height, physical activity, education level and smoking status). When measured/self-reported height values were not available, missing values were imputed with mean centre-, age- and sex-specific height values [34]. Mode imputation was used for baseline binary and categorical covariates missing less than 5% of their values (i.e. education level: “primary school completed”, physical activity: “moderately inactive”, smoking status: “never”). Multiple imputation was used in sensitivity analyses (details in the “sensitivity analyses” subsection below).

Main association analysis

Cox proportional hazards models with age as the underlying timescale were used to investigate the association between the intake of UPFs and the risk of HNC and OAC. We estimated HRs and 95% CIs per 10% g/d higher consumption of UPFs. Time of entry was defined as age at recruitment, while time of exit was defined as age at first cancer diagnosis (excluding non-melanoma skin cancer) or age at last follow-up (i.e. death, emigration, loss to follow-up or end of follow-up [i.e. between June 2008 and December 2013, depending on the centre]), whichever came first. Model 1 was stratified by age at recruitment in 1-year categories, sex and sub-centre. Model 2 was additionally adjusted for education, physical activity, height and smoking status. Model 3 was additionally adjusted for alcohol intake in g/d to reflect the association between the consumption of UPFs and cancer, regardless of alcohol intake (a well-known cancer risk factor [36–44] that forms part of some processed foods and UPFs).

We graphically assessed the proportional hazards assumption using log–log survival plots. Additionally, we tested proportionality using Schoenfeld residuals. We also used correlation matrices and variance inflation factors to assess the presence of multicollinearity. Non-linearity was assessed using likelihood ratio tests comparing UPF consumption (in %g/d) modelled with and without natural cubic splines.

We undertook additional analyses to investigate the associations between the consumption of UPFs and the risk of HNC subtypes (i.e. oral cavity, oropharynx, hypopharynx, larynx, and oral cavity and pharynx unspecified/overlapping cancers). Heterogeneity tests were used to assess differences between HNC subtype estimates.

Furthermore, we stratified Model 3 (for every exposure–outcome combination) by alcohol intake (as defined by Wozniak et al. [45], i.e. no/light alcohol intake [0.1–6 g/d (men); 0.1–3 g/d (women)], moderate alcohol intake [6.1–24 g/d (men); 3.1–24 g/d (women)], heavy alcohol intake [> 24 g/d]), sex (i.e. male, female), physical activity (i.e. inactive, moderately inactive, moderately active, active), smoking status (i.e. never smoker, former smoker, current smoker) and education level (i.e. primary school or less, secondary or technical/professional school, higher education) and performed likelihood ratio tests to explore interactions. Models were not adjusted for the stratification variable.

Mediation analysis

Under the strong assumption that there is no residual confounding or measurement error in our study, we conducted a mediation analysis using the counterfactual framework [46] to further explore the mediating role of BMI and WHR in the associations between UPF consumption and the risk of HNC and OAC (Fig. 1).

In exploratory analyses, we ran linear regressions to study the associations between UPF consumption (i.e. the exposure) and both WHR and BMI (i.e. the potential mediators). We also ran exposure-adjusted Cox regressions to analyse the associations between the potential mediators and the risk of both HNC and OAC (i.e. the outcomes). Where there was evidence of an association between the potential mediator and both the exposure and the outcome, we used the “cmest” function in the “CMAverse” R package [47] to decompose the Total Effect (TE) of UPF consumption on the corresponding upper-aerodigestive

Fig. 1 Mediation analysis diagram of the counterfactual two-way decomposition of the total effect of UPF consumption on the risk of head and neck cancer and oesophageal adenocarcinoma. All mediation models accounted for potential exposure–mediator interactions and were adjusted for age at recruitment in 1-year categories, sex, sub-centre, education, physical activity, height, smoking status and alcohol intake. The total effect (TE) corresponds to the sum of the pure natural direct effect (PNDE) and the total natural indirect effect (TNIE). Point estimates were obtained by direct counterfactual imputation estimation and confidence intervals were obtained using 1000 bootstrap repetitions. Abbreviations: BMI, body mass index; WHR, waist-to-hip ratio; UPF, ultra-processed food; HNC, head and neck cancer, OAC, oesophageal adenocarcinoma



tract cancer into a Pure Natural Direct Effect (PNDE) and a Total Natural Indirect Effect (TNIE) (on the ratio scale $TE = PNDE \times TNIE$). The proportion mediated was also calculated (i.e. $100 \times (PNDE \times (TNIE - 1)) / (TE - 1)$) for each exposure–mediator–outcome combination [48]. All mediation models accounted for potential exposure–mediator interactions and were adjusted for age at recruitment in 1-year categories, sex, sub-centre, education, physical activity, height, smoking status and alcohol intake. Point estimates were obtained by direct counterfactual imputation estimation and 95% CIs were obtained using 1000 bootstrap repetitions. The results were scaled to reflect a 10% g/d higher consumption of UPFs.

Sensitivity analyses

As a sensitivity analysis, we explored adjusting our Cox models for total water intake (i.e. water content from foods and drinks, in addition to drinking water and water used as an ingredient in preparations). This was to see whether differences in water content across NOVA groups may influence the associations between the relative intake of UPFs and the risk of HNC and OAC. Similarly, we explored adjustments for total energy intake.

We also reran our Cox models after excluding participants who were censored during the first two years of follow-up to avoid reverse causation due to undiagnosed cancer at recruitment.

Additionally, we repeated the analyses using the absolute intake of UPFs in grams per day (g/d) and the absolute and relative intake in kilocalories per day (kcal/d and %kcal/d, respectively) as the exposure.

Moreover, we conducted a complete case analysis excluding participants with missing data for at least one lifestyle covariate (i.e. smoking status, physical activity and education level). In addition, we used the ‘mice’ R package to perform multivariate imputation by chained equations (MICE) [49], whereby smoking status, physical activity and education level were imputed five times by predictive mean matching. We fit our models using the MICE imputed data sets and then pooled the results according to Rubin’s rules [50] to obtain average HR estimates and standard errors for each model. For the complete case analysis and the MICE analysis, we still used centre-, age- and sex-specific imputed height as a covariate, as this is standard practice when dealing with anthropometric variables as confounders in EPIC [34].

Finally, we performed a negative control outcome analysis (i.e. where the outcome is not plausibly linked to the exposure of interest) to help identify any residual confounding that could be biasing our results

[51]. We considered accidental deaths as the outcome (instead of upper-aerodigestive tract cancers) since the consumption of foods by their degree of processing is unlikely associated with the risk of being involved in a deadly accident (e.g. falls, transport accidents, accidental drowning). Any evidence of an association between UPF consumption and accidental deaths would suggest that our main results may be biased by the same factors that biased the negative control outcome results. Accidental deaths were defined as deaths due to events linked to codes V01–X59 in the 10th Revision of the International Classification of Diseases (ICD-10). For the negative control analysis, time of exit was defined as age at the time of death, emigration, loss to follow-up or end of follow-up, whichever came first. Participants were not censored at the time of cancer diagnosis, whereas they were in all other analyses in this study. The accidental death models accounted for the same covariates as the main analysis. BMI and type 2 diabetes mellitus would not normally be adjusted for in this analysis, as they are potential mediators and adjusting for them could induce collider bias (i.e. open backdoor paths from UPF consumption to accidental deaths through unobserved factors) [52]. Here, we did this in an explorative manner, assuming the absence of unobserved confounders of BMI, type 2 diabetes mellitus and accidental deaths.

Statistical software

All statistical analyses and visualisations were performed using R version 4.2.3. We used version 3.2.10 of the ‘survival’ R package for the Cox regressions and version 0.1.0 of the ‘CMAverse’ R package [47] for the mediation analysis. We also used version 0.1.0 of the ‘ggforestplot’ R package to create forest plots. To create tables, we used version 1.3.0 of the ‘tidyverse’ R package and version 0.7.0 of the ‘flextable’ R package. P-values for heterogeneity between HNC subtype estimates were obtained using version 4.18–0 of the ‘meta’ R package. Non-linearity was assessed using version 4.2.3 of the ‘splines’ R package. MICE was performed using version 3.16 of the ‘mice’ R package. Two-sided p -values < 0.05 were considered statistically significant.

Results

Descriptive characteristics

In total, we included 450,111 participants of which 70.8% were female. The mean age at recruitment was 51.1 years

Table 1 Baseline characteristics of study participants by sex-specific quartiles of relative ultra-processed food consumption (%g/d)

	Q1 (M: ≤ 8.95%g/d; F: ≤ 6.94%g/d)	Q2 (M: 8.96– 13.16%g/d; F: 6.95–11.30%g/d)	Q3 (M: 13.17– 18.62%g/d; F: 11.31–17.54%g/d)	Q4 (M: 18.63– 85.16%g/d; F: 17.55–90.53%g/d)	Overall
Participants, <i>n</i>	112,529	112,528	112,527	112,527	450,111
Age at recruitment (years), mean ± SD	52.9 ± 7.74	52.5 ± 8.87	51.1 ± 10.2	48.1 ± 11.1	51.1 ± 9.75
Females, <i>n</i> (%)	79,672 (70.8)	79,672 (70.8)	79,671 (70.8)	79,671 (70.8)	318,686 (70.8)
Country, <i>n</i> (%)					
France	38,677 (34.4)	19,085 (17.0)	7592 (6.7)	2049 (1.8)	67,403 (15.0)
Italy	17,500 (15.6)	13,098 (11.6)	8935 (7.9)	5012 (4.5)	44,545 (9.9)
Spain	23,830 (21.2)	8159 (7.3)	4975 (4.4)	3025 (2.7)	39,989 (8.9)
United Kingdom	2987 (2.7)	13,407 (11.9)	26,794 (23.8)	32,228 (28.6)	75,416 (16.8)
The Netherlands	3217 (2.9)	10,035 (8.9)	13,335 (11.9)	9951 (8.8)	36,538 (8.1)
Germany	6324 (5.6)	11,200 (10.0)	14,136 (12.6)	16,897 (15.0)	48,557 (10.8)
Sweden	8854 (7.9)	16,577 (14.7)	14,299 (12.7)	8944 (7.9)	48,674 (10.8)
Denmark	10,934 (9.7)	19,621 (17.4)	14,384 (12.8)	10,075 (9.0)	55,014 (12.2)
Norway	206 (0.2)	1346 (1.2)	8077 (7.2)	24,346 (21.6)	33,975 (7.5)
Height ^a (cm), mean ± SD	164 ± 8.65	166 ± 8.99	167 ± 8.83	168 ± 8.57	166 ± 8.86
Measured BMI ^b (kg/m ²), mean ± SD	26.2 ± 4.25	25.5 ± 4.10	25.4 ± 4.12	25.5 ± 4.41	25.6 ± 4.22
Missing, <i>n</i> (%)	27,851 (24.8)	15,490 (13.8)	14,348 (12.8)	26,873 (23.9)	84,562 (18.8)
Measured WHR ^c , mean ± SD	0.86 ± 0.1	0.84 ± 0.1	0.83 ± 0.1	0.84 ± 0.1	0.84 ± 0.1
Missing, <i>n</i> (%)	31,561 (28.0)	24,491 (21.8)	23,256 (20.7)	33,945 (30.2)	113,253 (25.2)
Education level, <i>n</i> (%)					
None	8622 (7.7)	3120 (2.8)	2190 (1.9)	1619 (1.4)	15,551 (3.5)
Primary school completed	30,898 (27.5)	28,771 (25.6)	26,533 (23.6)	24,862 (22.1)	111,064 (24.7)
Technical/Professional school	14,072 (12.5)	24,757 (22.0)	30,300 (26.9)	34,653 (30.8)	103,782 (23.1)
Secondary school	28,716 (25.5)	23,685 (21.0)	20,675 (18.4)	20,834 (18.5)	93,910 (20.9)
Higher education	27,810 (24.7)	28,964 (25.7)	27,535 (24.5)	24,622 (21.9)	108,931 (24.2)
Missing	2411 (2.1)	3231 (2.9)	5294 (4.7)	5937 (5.3)	16,873 (3.7)
Physical activity level, <i>n</i> (%)					
Inactive	27,550 (24.5)	21,938 (19.5)	20,355 (18.1)	18,189 (16.2)	88,032 (19.6)
Moderately inactive	39,820 (35.4)	38,982 (34.6)	36,912 (32.8)	34,227 (30.4)	149,941 (33.3)
Moderately active	27,477 (24.4)	28,369 (25.2)	29,751 (26.4)	34,602 (30.8)	120,199 (26.7)
Active	17,185 (15.3)	21,700 (19.3)	22,470 (20.0)	21,760 (19.3)	83,115 (18.5)
Missing	497 (0.4)	1539 (1.4)	3039 (2.7)	3749 (3.3)	8824 (2.0)
Smoking status, <i>n</i> (%)					
Never	57,091 (50.7)	54,677 (48.6)	53,788 (47.8)	53,738 (47.8)	219,294 (48.7)
Former	28,412 (25.2)	31,268 (27.8)	32,246 (28.7)	30,754 (27.3)	122,680 (27.3)
Current	24,549 (21.8)	24,830 (22.1)	24,684 (21.9)	25,651 (22.8)	99,714 (22.2)
Missing	2477 (2.2)	1753 (1.6)	1809 (1.6)	2384 (2.1)	8423 (1.9)
Alcohol intake (g/d), mean ± SD	15.3 ± 21.1	13.0 ± 17.1	10.7 ± 14.4	8.00 ± 12.3	11.7 ± 16.8
Energy intake (kcal/d), mean ± SD	2100 ± 622	2110 ± 598	2140 ± 607	2200 ± 666	2140 ± 625
Carbohydrate intake (g/d), mean ± SD	237 ± 75.7	245 ± 72.3	257 ± 75.6	278 ± 91.6	254 ± 80.7
Fat intake (g/d), mean ± SD	79.4 ± 28.2	81.4 ± 28.6	82.7 ± 29.8	82.6 ± 31.7	81.5 ± 29.6
Fibre intake (g/d), mean ± SD	24.3 ± 8.74	23.3 ± 8.10	23.8 ± 8.54	23.9 ± 9.04	23.8 ± 8.62
Sodium intake (mg/d), mean ± SD	2500 ± 968	2650 ± 1110	2630 ± 1190	2650 ± 1290	2610 ± 1150

Table 1 (continued)

	Q1 (M: $\leq 8.95\%$ g/d; F: $\leq 6.94\%$ g/d)	Q2 (M: 8.96– 13.16%g/d; F: 6.95–11.30%g/d)	Q3 (M: 13.17– 18.62%g/d; F: 11.31–17.54%g/d)	Q4 (M: 18.63– 85.16%g/d; F: 17.55–90.53%g/d)	Overall
Relative intake of ultra-processed foods (%g/d), mean \pm SD	4.95 \pm 1.90	9.65 \pm 1.53	14.6 \pm 1.85	25.7 \pm 7.78	13.7 \pm 8.76
Absolute intake of ultra-processed foods (g/d), mean \pm SD	141 \pm 80.1	273 \pm 104	393 \pm 138	648 \pm 321	364 \pm 264
Relative intake of processed foods (%g/d), mean \pm SD	16.9 \pm 12.0	14.2 \pm 10.0	12.5 \pm 8.85	10.7 \pm 7.43	13.6 \pm 10.0
Absolute intake of processed foods (g/d), mean \pm SD	445 \pm 367	385 \pm 313	328 \pm 270	271 \pm 237	357 \pm 308
Relative intake of unprocessed/minimally processed foods (%g/d), mean \pm SD	76.6 \pm 12.7	74.9 \pm 10.7	71.9 \pm 9.50	62.7 \pm 10.3	71.5 \pm 12.1
Absolute intake of unprocessed/minimally processed foods (g/d), mean \pm SD	2190 \pm 985	2130 \pm 818	1940 \pm 699	1590 \pm 653	1960 \pm 833

Baseline characteristics stratified by sex-specific quartiles of relative intake of ultra-processed foods (%g/d). Abbreviations: BMI, Body mass index; WHR, Waist-to-hip ratio; Q1, quartile 1; Q2, quartile 2; Q3, quartile 3; Q4, quartile 4; M, Males; F, Females; SD, standard deviation

^aWhen measured/self-reported height values were not available, missing values were imputed with mean centre-, age- and sex-specific height values

^bAccording to the World Health Organization [53], a healthy adult BMI lies between 18.5 and 24.9 kg/m²

^cAccording to the World Health Organization [53], a healthy WHR corresponds to <0.9 for males and <0.85 for females

(SD 9.8, range 17.8–98.5 years). The mean consumption of UPFs in the cohort was 13.7% g/d (364 g/d), ranging from a mean intake of 8% g/d (156.9 g/d) in Spain to 18.6% g/d (520.5 g/d) in the United Kingdom. A histogram of the proportion of UPFs in the diet (in %g/d) is available in Supplementary Fig. 2. On average, males consumed a higher proportion of UPFs than females (14.7% vs 13.3%, $p < 0.001$). Participants with technical education were among the highest consumers of UPFs (Table 1). UPFs contributed greatly to the diet of younger, taller, and more physically active participants. Participants who did not provide data on their physical activity and education also tended to consume more UPFs. In terms of diet quality, UPFs were highly consumed by participants who consumed less alcohol and more calories, carbohydrates, fat and sodium.

The UPF group was mainly composed of fizzy drinks (14.1% of absolute UPF consumption in g/d), non-carbonated sweetened beverages (12.1%), ultra-processed dairy products (12.0%), ultra-processed breads (12.0%) and ultra-processed meats (9.9%) (Supplementary Table 1). Beer and wine (46.2%) were the main contributors to the processed

foods group, followed by processed breads (22.7%) and cheese (10.2%). The unprocessed/minimally processed foods group primarily comprised tea and coffee (32.2%), water (17.3%), milk and plain yoghurt (12.4%), fruit (11.2%) and vegetables (8.9%).

Associations between the consumption of ultra-processed foods and upper-aerodigestive tract cancers

During a mean follow-up of 14.13 \pm 3.98 years (6,358,569 person-years, median follow-up = 14.95 years; range 1 day–22.79 years), 1125 incident cases of HNC and OAC were documented. Of these, 910 had HNC (i.e. 234 oral cavity, 235 oropharynx, 66 hypopharynx, 310 larynx and 65 oral cavity and pharynx unspecified/overlapping regions) and 215 had OAC.

The proportional hazards assumption was met in all models (Supplementary Figs. 3, 4, 5, 6), and we did not find evidence of multicollinearity between covariates (Supplementary Fig. 7 and Supplementary Table 2).

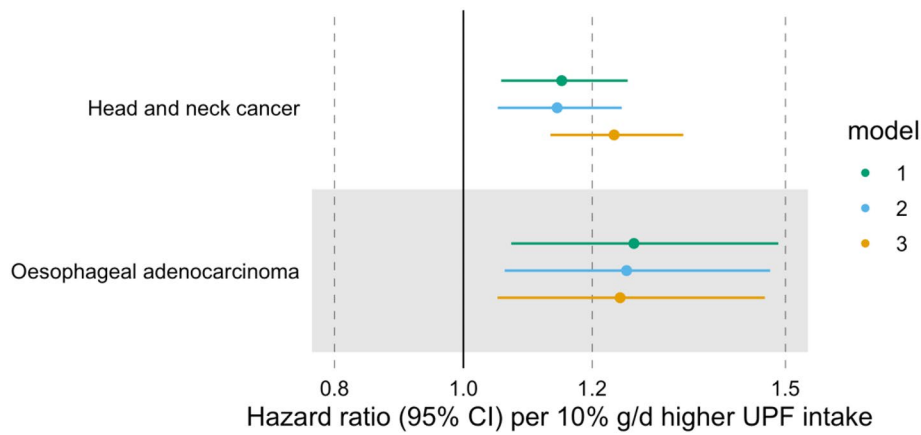
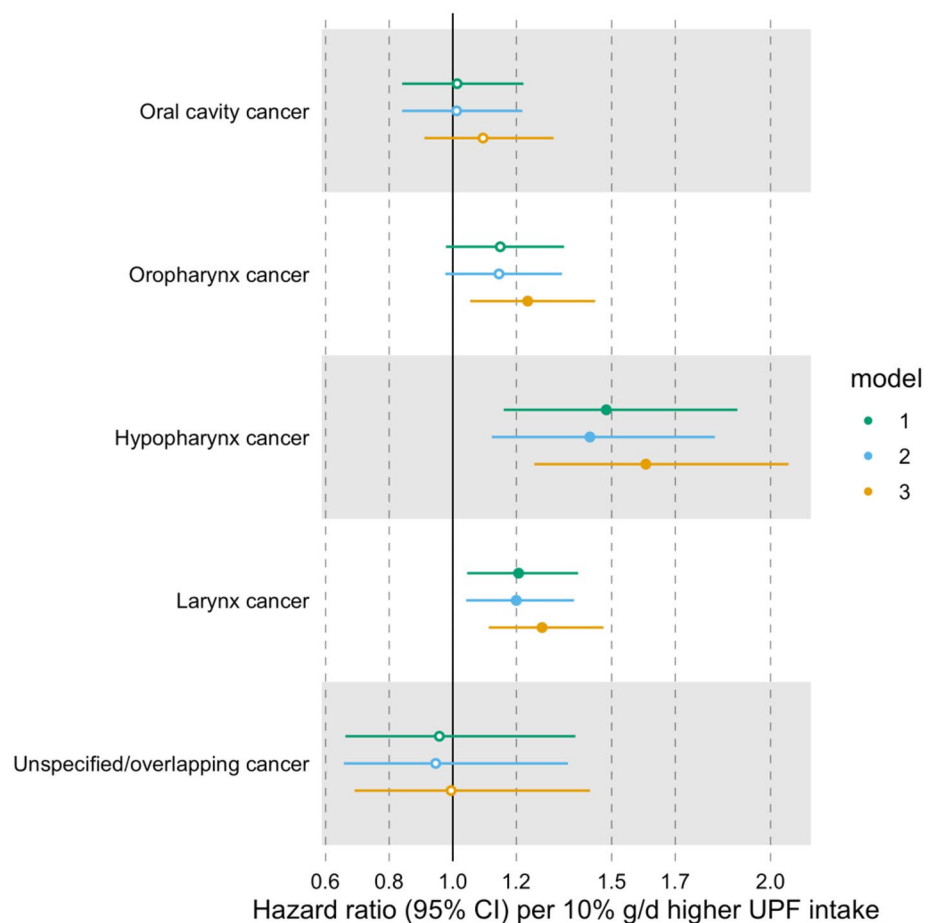


Fig. 2 Associations between the consumption of ultra-processed foods (in %g/d) and the risk of head and neck cancer and oesophageal adenocarcinoma. Hazard ratios per 10% g/d higher ultra-processed food intake. Time of entry was defined as age at recruitment, while time of exit was defined as age at first cancer diagnosis (excluding non-melanoma skin cancer) or age at last follow-up (i.e. death, emigration, loss to follow-up or end of follow-up), whichever came first.

Model 1 was stratified by age at recruitment in 1-year categories, sex, and sub-centre. Model 2 was additionally adjusted for education, physical activity, height, and smoking status. Model 3 was additionally adjusted for alcohol intake. $N=450,111$, of which 910 and 215 had head and neck cancer and oesophageal adenocarcinoma, respectively. Abbreviations: CI, confidence interval; UPF, ultra-processed food

Fig. 3 Associations between the consumption of ultra-processed foods (in %g/d) and head and neck cancer subtypes. Hazard ratios per 10% g/d higher ultra-processed food intake. Time of entry was defined as age at recruitment, while time of exit was defined as age at first cancer diagnosis (excluding non-melanoma skin cancer) or age at last follow-up (i.e. death, emigration, loss to follow-up or end of follow-up), whichever came first. Model 1 was stratified by age at recruitment in 1-year categories, sex, and sub-centre. Model 2 was additionally adjusted for education, physical activity, height and smoking status. Model 3 was additionally adjusted for alcohol intake. $N=450,111$, of which 234, 235, 66, 310 and 65 had cancer of the oral cavity, oropharynx, hypopharynx, larynx, and oral cavity and pharynx unspecified/overlapping regions, respectively. Abbreviations: CI, confidence interval; UPF, ultra-processed food



Furthermore, there was no evidence of non-linearity between UPF consumption and HNC ($p = 0.54$) (Supplementary Fig. 8). The non-linearity test for the association between UPF consumption and the risk of OAC was not informative due to the limited number of OAC cases in the dataset (results not shown).

Head and neck cancer (HNC)

A higher proportion of UPF in the diet (in %g/d) was associated with a higher risk of HNC, even after accounting for alcohol intake (HR = 1.23 per 10% g/d higher UPF intake, 95% CI 1.14–1.34) (Fig. 2 and Supplementary Table 3). We did not find evidence of heterogeneity between HNC subtypes (p -value for heterogeneity = 0.11) (Fig. 3 and Supplementary Table 4).

In stratified analyses for the association between UPF consumption (in %g/d) and HNC risk, we did not find evidence of effect modification by alcohol intake (p -value for interaction = 0.46), physical activity level (p -value for interaction = 0.48), smoking status (p -value for interaction = 0.46) or education level (p -value for interaction = 0.31) (Supplementary Table 5). There was some evidence of an interaction by sex (p -value for interaction = 0.006), with a positive association between UPF consumption and HNC risk among males (HR = 1.34 per 10% g/d higher UPF intake, 95% CI 1.22–1.48,

$N = 131,425$, events = 603) but not among females (HR = 1.03, 95% CI 0.87–1.21, $N = 318,686$, events = 307).

Oesophageal adenocarcinoma (OAC)

After accounting for alcohol intake, UPF consumption (in %g/d) was positively associated with OAC risk (HR = 1.24 per 10% g/d higher UPF intake, 95% CI 1.05–1.47) (Fig. 2 and Supplementary Table 3).

When we stratified the association between UPF consumption (in %g/d) and OAC risk, we did not find evidence of differing estimates across levels of alcohol intake (p -value for interaction = 0.18), physical activity (p -value for interaction = 0.94), smoking status (p -value for interaction = 0.99), sex (p -value for interaction = 0.44) or education (p -value for interaction = 0.83) (Supplementary Table 5).

Mediating role of adiposity in the associations between ultra-processed food consumption and upper-aerodigestive tract cancers

In exploratory analyses (Supplementary Table 6), we found positive associations between UPF consumption and both BMI (mean change = 0.24 kg/m² per 10% g/d higher UPF intake, 95% CI 0.22–0.26) and WHR (mean change = 0.41 per 10% g/d higher UPF intake, 95% CI 0.38–0.43).

Table 2 Mediation analysis for the associations between ultra-processed food consumption (in %g/d) and head and neck cancer and oesophageal adenocarcinoma, where the potential mediators (i.e. waist-to-hip ratio and body mass index) were measured at baseline

Association	Mediator	Effect	Estimate	95% CI	<i>P</i> -value	<i>P</i> -value for interaction
Ultra-processed food consumption and head and neck cancer risk	Waist-to-hip ratio	TE, HR	1.23	1.12–1.33	<0.001	0.03
		PNDE, HR	1.22	1.11–1.32	<0.001	
		TNIE, HR	1.01	1.01–1.01	<0.001	
		PM, %	5.32	2.73–10.27	<0.001	
Ultra-processed food consumption and oesophageal adenocarcinoma risk	Waist-to-hip ratio	TE, HR	1.20	1.02–1.36	0.032	0.38
		PNDE, HR	1.17	0.99–1.33	0.06	
		TNIE, HR	1.03	1.02–1.03	<0.001	
		PM, %	14.85	7.57–72.27	0.03	
	Body mass index	TE, HR	1.21	1.02–1.37	0.038	0.28
		PNDE, HR	1.18	1.00–1.34	0.06	
		TNIE, HR	1.02	1.01–1.03	<0.001	
		PM, %	13.38	5.69–52.86	0.04	

Point estimates were obtained by direct counterfactual imputation estimation and confidence intervals were obtained using 1000 bootstrap repetitions. Results were scaled to reflect a 10% g/d higher consumption of ultra-processed foods. Mediation models were adjusted for age at recruitment in 1-year categories, sex, sub-centre, education, physical activity, height, smoking status and alcohol intake. The total effect corresponds to the multiplication of the pure natural direct effect and the total natural indirect effect. Waist-to-hip ratio models included 336,858 participants, of which 828 and 195 had head and neck cancer and oesophageal adenocarcinoma, respectively. Body mass index models included 365,549 participants, of which 212 had oesophageal adenocarcinoma. *HR* hazard ratio; *CI* confidence interval; *TE* total effect in presence of the mediator; *PNDE* pure natural direct effect; *TNIE* total natural indirect effect; *PM* proportion mediated

Moreover, BMI was positively associated with OAC risk (HR = 1.08 per 1 kg/m² higher BMI, 95% CI 1.04–1.11), but inversely associated with HNC (HR = 0.98 per 1 kg/m² higher BMI, 95% CI 0.96–0.99). WHR was positively associated with the risk of both HNC (HR = 1.02 per 0.01 higher WHR, 95% CI 1.01–1.03) and OAC (HR = 1.06 per 0.01 higher WHR, 95% CI 1.04–1.08).

WHR as a mediator between UPF consumption and HNC risk

Only a small part of the positive association between UPF consumption (in %g/d) and HNC risk was mediated via WHR (5%, 95% CI 3–10%, $p < 0.001$; TNIE HR = 1.01 per 10% g/d higher UPF intake, 95% CI 1.01–1.01) (Table 2). Most of the association was not explained by WHR (PNDE HR = 1.22 per 10% g/d higher UPF intake, 95% CI 1.11–1.32). Furthermore, there was some evidence of an interaction between UPF consumption and WHR (p -value for interaction = 0.03).

WHR as a mediator between UPF consumption and OAC risk

The TE of UPF consumption (in %g/d) on OAC risk was decomposed into a PNDE of 1.17 (95% CI 0.99–1.33) and a TNIE via WHR of 1.03 (95% CI 1.02–1.03) (Table 2). Hence, the proportion mediated by WHR was 15% (95% CI 8–72%, $p = 0.03$) in the association with HNC.

BMI as a mediator between UPF consumption and OAC risk

Most of the association between UPF consumption (in %g/d) and higher OAC risk was not mediated via BMI (PNDE HR = 1.18 per 10% g/d higher UPF intake, 95% CI 1.00–1.34) (Table 2). BMI mediated 13% (95% CI 6–53%, $p = 0.04$) of the association (TNIE HR = 1.02 per 10% g/d higher UPF intake, 95% CI 1.01–1.03).

Sensitivity analyses for the associations between ultra-processed food consumption and upper-aerodigestive tract cancers

Further adjusting for total water intake (including water in foods) or total energy intake produced similar results to those obtained in the main analyses (Supplementary Tables 7 and 8). Excluding participants censored in the first two years of follow-up ($N = 442,536$, since 7575 were excluded) did not substantially affect our results either (Supplementary

Table 9). Likewise, the complete case analysis results ($N = 419,590$, of which 851 had HNC and 191 had OAC) and the results obtained using multiple imputation were similar to those obtained in our main analysis (Supplementary Tables 10 and 11).

Repeating the analyses using either the absolute intake of UPFs in grams per day (g/d) or the relative intake of UPFs in kilocalories per day (%kcal/d) as the exposure produced comparable results to those obtained in the main analysis (Supplementary Table 12 and Supplementary Fig. 9). Nevertheless, using the absolute intake of UPFs in kilocalories per day (kcal/d) as the exposure produced slightly different results to those in the main analysis, namely because UPF consumption was no longer associated with HNC risk (HR = 1.01 per 100 kcal/d higher UPF intake, 95% CI 0.99–1.03) (Supplementary Table 12 and Supplementary Fig. 9).

Lastly, we found a positive association between UPF consumption (%g/d) and accidental deaths (HR = 1.12 per 10% g/d higher UPF intake, 95% CI 1.02–1.23) in the negative control outcome analysis, after accounting for all the covariates included in the upper-aerodigestive tract cancer models (Supplementary Table 13). This association also withstood additional adjustments for BMI and type 2 diabetes mellitus (i.e. factors that may influence recovery after an accident) (results not shown).

Discussion

In this large prospective cohort, UPF consumption (in %g/d) was associated with an increased risk of HNC and OAC. We did not find evidence of heterogeneity between HNC subtype association estimates. Furthermore, the positive association between UPF intake and HNC may be stronger in males than in females. Our negative control analysis suggests that at least part of the observed associations between the consumption of UPFs and the risk of upper-aerodigestive tract cancers is likely due to the influence of residual confounding. However, this does not necessarily mean that the associations are entirely non-causal; only that any causal estimate is likely smaller than we observed. In our mediation analysis, adiposity (i.e. BMI and WHR) only mediated a small proportion of the positive associations between UPF consumption and HNC and OAC.

Apart from the study conducted by Kliemann et al. [9] (which motivated our research), only one other study has investigated the associations between UPF consumption

and HNC and OAC risk. Chang et al. [8] did not find an association between UPF consumption and the risk of HNC and OAC in the UK Biobank, in contrast to the findings in EPIC. A possible explanation for the null results in the UK Biobank could be limited power (197,426 UK Biobank participants of which 342 and 186, respectively, developed HNC and OAC over a median follow-up of 10 years, versus 450,111 EPIC participants of which 910 and 250, respectively, developed HNC and OAC over a median follow-up of 15 years). Residual confounding and the fact that the FFQs used were not designed to capture the extent and purpose of food processing may also partly explain the inconsistencies between studies.

Our mediation results are in line with existing findings. First, UPFs have been associated with excess weight (i.e. obesity and BMI) and central adiposity (i.e. WHR and waist circumference) in several observational studies [10–13]. UPFs are highly palatable energy-dense foods with low nutritional quality. They are convenient, cheap, and often sold in large portions [4, 54]. This, in addition to their reduced satiety potential [55], favours the consumption of large portions and an excessive amount of calories. Some studies even suggest that the consumption of UPFs may disrupt the gut microbiota, induce inflammation, and cause endocrine changes that disturb energy balance and increase the risk of obesity [11, 56, 57]. Second, multiple studies suggest that excess weight and abdominal obesity are positively associated with OAC risk [15–20, 58], and that central adiposity may be a risk factor for HNC [21, 22]. Hence, it is plausible that BMI and WHR mediate the association between UPF consumption and OAC risk, and that WHR mediates the association between UPF consumption and HNC risk. Notwithstanding, our findings indicate that the mediated effects via BMI and WHR are small and that other mechanisms are likely involved.

A review of prospective cohort studies suggested that diet quality did not play an important role in the positive associations between UPF consumption, obesity and obesity-related outcomes (e.g. cancer), since the adjustment for several dietary factors did not substantially attenuate the associations [59]. The authors argue that ultra-processing itself may be associated with disease risk, independent of nutritional quality. This has profound implications for the food industry, as it could mean that UPF reformulation would not be sufficient to tackle the risks associated with UPF consumption. We acknowledge

that this may just be another case of ‘highly consequential but misleading findings’ [60], where the observed associations are not causal but rather an artefact of residual confounding [61, 62] (the association of UPF consumption with accidental death in EPIC provides some support for this interpretation). However, if these associations truly reflect causality, the presence of carcinogenic compounds in UPFs, such as neo-formed contaminants produced during heat treatment, contaminants transferred from packaging materials and additives used to preserve and improve the organoleptic properties of food [5] may partly explain the relation between UPF intake and upper-aerodigestive tract cancer risk.

In sensitivity analyses, we did not find an association between a 100 kcal/d higher UPF consumption and HNC risk. This could be because the consumption of artificially sweetened UPFs (which may contain potentially carcinogenic compounds like aspartame and 4-methylimidazole [63]) is likely disregarded or underestimated when using kcal/d as a measure of UPF intake. We acknowledge that this is an arguable hypothesis since Chazelas et al. [64] did not find an association between artificially sweetened drinks and cancer in the NutriNet-Santé cohort. Nevertheless, a study in the same cohort found a positive association between higher artificial sweetener intake from all food sources and cancer risk, suggesting other artificially sweetened UPFs (e.g. yoghurts, breakfast cereals, gelatine desserts) may play a role in cancer incidence [65].

Some of the strengths of this study include the large sample size of the EPIC cohort and its long follow-up time. The prospective nature of EPIC and the availability of measured rather than self-reported BMI and WHR were also advantages. Moreover, EPIC’s multi-centre design increases the diversity of our study sample. Another advantage is that cancer cases were detected through registries (which provide detailed information on cancer subtypes) and active follow-up methods, both of which are unlikely to be affected by measurement error. Finally, the use of several measures of UPF intake (%g/d, g/d, %kcal/d and kcal/d) as the exposure makes our results more comparable to previous studies investigating the associations between the consumption of UPFs and cancer [5–9].

We acknowledge that this study has several limitations. For instance, in our mediation analysis we assumed that the associations between UPF consumption and upper-aerodigestive tract cancers were not influenced by

residual confounding or measurement error. These are strong assumptions since residual confounding due to unmeasured (e.g. human papilloma-virus infection) or imprecisely measured (e.g. smoking and alcohol intake) confounders inevitably biased our estimates to some extent (since estimates changed when the models were adjusted for potential confounders for which some values were missing or likely measured with error). Indeed, the fact that our negative control outcome analysis suggested that UPF consumption may be associated with accidental deaths points to the possibility of residual confounding.

For our mediation analysis, we used data on BMI and WHR collected at baseline. Admittedly, a limitation with this approach is that the exposure data were not gathered prior to the mediator data, so we cannot be certain that the exposure temporally precedes the mediator. Follow-up data on measured BMI and WHR were only available for 5% and 27% of the participants who answered the lifestyle follow-up questionnaire ($N = 349,283$), respectively. Unfortunately, cancer cases among participants with complete follow-up data were insufficient for us to conduct any sensitivity analyses using follow-up data.

An additional limitation is that we assumed that BMI and WHR mediated the association between UPF consumption and OAC risk through separate pathways. Since BMI and WHR are correlated, it would be incorrect to assume that the proportion mediated via adiposity equals the sum of the proportions mediated through both BMI and WHR. Therefore, in this association, the proportion mediated via adiposity is likely less than 28% (13% via BMI plus 15% via WHR).

Another issue is that the FFQs used at baseline were not designed to distinguish between NOVA groups [66], potentially leading to random misclassification bias and the weakening of our association estimates. Nonetheless, Huybrechts et al. [31] found positive correlations between UPF consumption and food processing biomarkers (i.e. plasma elaidic acid, an unsaturated trans-fatty acid, and urinary 4-methylsyringol sulphate) in EPIC, suggesting that UPFs were likely correctly identified in the dataset. Also, the dietary data used in our analyses were collected

only once, at baseline in the 1990s, when the availability and consumption of UPFs was relatively lower than today. Hence, this study relies on the somewhat unrealistic assumption that UPF intake was rather low and did not increase over time [67]. When Kliemann et al. [9] explored the association between the consumption of UPFs and cancer in EPIC under a hypothetical “upper bound scenario” (where foods were classified into NOVA groups based on their highest degree of processing possible), results were similar to those obtained under the more conservative “middle bound scenario” used in our study (where foods were classified into NOVA groups based on their most likely degree of processing at the time of dietary data collection). Although the use of a hypothetical “upper bound scenario” may account for some of the changes in dietary intake during follow-up, regression dilution bias was still an issue Kliemann et al. [9] could not account for in their analyses. Consequently, dietary questionnaire-related biases may have led to the underestimation of the association between UPF consumption and cancer in the EPIC cohort.

In conclusion, we reaffirmed that UPF intake is associated with an increased risk of HNC and OAC in the EPIC study. Since BMI and WHR explain little of the associations between UPF consumption and upper-aerodigestive tract cancers, further research is required to investigate other mechanisms that may be at play (if there is indeed any causal effect of UPF consumption on these cancers). Our results are likely influenced by residual confounding, as indicated by the negative control analysis. Therefore, our findings should be regarded with caution until they are replicated in other settings (i.e. in populations with different underlying confounding structures) or triangulated with evidence obtained using other methodological approaches.

Appendix

See Table 3

Table 3 STROBE statement—checklist of items that should be included in reports of observational studies

	Item no	Recommendation	Page no	Relevant text from manuscript
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	3	"European Prospective Investigation into Cancer and Nutrition (EPIC) cohort" "We used Cox regressions to investigate the associations..."
<i>Introduction</i>				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4	"In recent years, there has been growing interest ... this has not been investigated using mediation analysis"
Objectives	3	State specific objectives, including any prespecified hypotheses	4–5	"The aim of this study was..."
<i>Methods</i>				
Study design	4	Present key elements of study design early in the paper	6	"The EPIC cohort" subsection
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6	"The EPIC cohort" and "Study sample" subsections
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants (b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	6–7	"Ascertainment of cancer cases" and "Study sample" subsections
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	NA	NA
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6–8	"Dietary data and food processing variables", "Ascertainment of cancer cases", "Covariates" and "Potential mediators" subsections
Bias	9	Describe any efforts to address potential sources of bias	9	"Dietary data and food processing variables", "Ascertainment of cancer cases", "Covariates" and "Potential mediators" subsections "Model 1 was stratified by age at recruitment in 1-year categories, sex, and sub-centre. Model 2 was additionally adjusted for education, physical activity, height, and smoking status. Model 3 was additionally adjusted for alcohol intake in g/d to reflect the association between the consumption of foods by their degree of processing and cancer, regardless of alcohol intake"
Study size	10	Explain how the study size was arrived at	6	"Study sample" subsection
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8–12	"Covariates", "Potential mediators" and "Statistical analysis" subsections

Table 3 (continued)

	Item no	Recommendation	Page no	Relevant text from manuscript
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	8–12	“Statistical analysis” subsection
		(b) Describe any methods used to examine subgroups and interactions	9	“... we stratified Model 3 (for every exposure-outcome combination)”
		(c) Explain how missing data were addressed	9	“We used single-value imputation to deal with missing data among covariates used to control for potential confounding...”
		(d) Cohort study—If applicable, explain how loss to follow-up was addressed	9	“...time of exit was defined as age at first cancer diagnosis (excluding non-melanoma skin cancer) or age at last follow-up (i.e. death, emigration, loss to follow-up or end of follow-up [i.e. between June 2008 and December 2013, depending on the centre]), whichever came first”
		(e) Describe any sensitivity analyses	11	“Sensitivity analyses” within the “Statistical analysis” subsection
Participants	13*	(a) Report numbers of individuals at each stage of study—e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6	“Study sample” subsection in “Methods” section
		(b) Give reasons for non-participation at each stage	NA	NA
		(c) Consider use of a flow diagram	Supplement	Supplementary Fig. 1. Flowchart of European Prospective Investigation into Cancer and Nutrition (EPIC) participants included in the study
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders	13–14	“Descriptive characteristics” subsection
		(b) Indicate number of participants with missing data for each variable of interest	13–14	“Descriptive characteristics” subsection
		(c) Cohort study—Summarise follow-up time (e.g. average and total amount)	14	“During a mean follow-up of 14.13 ± 3.98 years (6,358,569 person-years, median follow-up = 14.95 years; range 1 day to 22.79 years)”
Outcome data	15*	Cohort study—Report numbers of outcome events or summary measures over time	14	“During a mean follow-up of 14.13 ± 3.98 years (6,358,569 person-years, median follow-up = 14.95 years; range 1 day to 22.79 years)”
		Case-control study—Report numbers in each exposure category, or summary measures of exposure		
		Cross-sectional study—Report numbers of outcome events or summary measures		

Table 3 (continued)

	Item no	Recommendation	Page no	Relevant text from manuscript
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorised	14–18	“Associations between the consumption of ultra-processed foods and upper-aerodigestive tract cancers” and “Mediating role of adiposity in the associations between ultra-processed food consumption and upper-aerodigestive tract cancers” subsections “Furthermore, we stratified Model 3 (for every exposure-outcome combination) by alcohol intake (as defined by Wozniak et al. [45], i.e. no/light alcohol intake [0.1–6 g/d (men); 0.1–3 g/d (women)], moderate alcohol intake [6.1–24 g/d (men); 3.1–24 g/d (women)], heavy alcohol intake [> 24 g/d])”
Other analyses	17	(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period Report other analyses done—e.g. analyses of subgroups and interactions, and sensitivity analyses	NA	NA
<i>Discussion</i>				
Key results	18	Summarise key results with reference to study objectives	19	“In this large prospective cohort, UPF consumption (in %g/d) was associated with...”
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20–21	“We acknowledge that this study has several limitations... dietary questionnaire-related biases may have led to the underestimation of the association between UPF consumption and cancer in the EPIC cohort.”
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	19–21	“Discussion” section
Generalisability	21	Discuss the generalisability (external validity) of the study results	19–21	“Discussion” section
<i>Other information</i>				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23	“Funding” subsection in the “Statements & Declarations” section

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00394-023-03270-1>.

Acknowledgements We acknowledge the use of data from the EPIC cohort and thank the participants and investigators for their contributions to the EPIC resource. We also thank the Asturias Government and the National Institute for Public Health and the Environment (RIVM), Bilthoven in the Netherlands, for their contribution and ongoing support to the EPIC study. We additionally thank Richard Wilkinson for proofreading the manuscript.

Author contributions FMB, IH, CB, CC, MG, AA, PB, NK, AGS and VV contributed to the study conception and design. CC, CB and BH prepared the dataset. FMB conducted the analysis and wrote the first draft of the manuscript. IH, CLR, GDS, TGR, MCB, VV and AGS provided comments on the first draft of the manuscript. All authors commented on other versions of the manuscript. All authors read and approved the final version of the manuscript.

Funding FMB was supported by a Wellcome Trust PhD studentship in Molecular, Genetic and Lifecourse Epidemiology (224982/Z/22/Z). RCR was supported by a Cancer Research UK grant (C18281/A29019). AGS is supported by the study of Dynamic longitudinal exposome trajectories in cardiovascular and metabolic non-communicable diseases [H2020-SC1-2019-Single-Stage-RTD, project ID 874739]. MCB is supported by a University of Bristol Vice Chancellor's Fellowship, the British Heart Foundation (AA/18/1/34219) and the UK Medical Research Council (MC_UU_00011/6). GDS works within the MRC Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC_UU_00011/1). CLR was supported by the Medical Research Council (MC_UU_00011/5) and by a Cancer Research UK (C18281/A29019) programme grant (the Integrative Cancer Epidemiology Programme). This work was supported by Cancer Research UK (C33493/A29678), World Cancer Research Fund International (IIG_FULLL_2020_033) and the Institut National du Cancer (INCa number 2021–138). The coordination of EPIC is financially supported by the International Agency for Research on Cancer (IARC) and by the Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London which has additional infrastructure support provided by the NIHR Imperial Biomedical Research Centre (BRC). The national cohorts are supported by: Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Éducation Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), German Institute of Human Nutrition Potsdam-Rehbruecke (DIfE), Federal Ministry of Education and Research (BMBF) (Germany); Associazione Italiana per la Ricerca sul Cancro-AIRC-Italy, Compagnia di SanPaolo and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); Health Research Fund (FIS)—Instituto de Salud Carlos III (ISCIII), Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, and the Catalan Institute of Oncology—ICO (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer Research UK (14136 to EPIC-Norfolk; C8221/A29017 to EPIC-Oxford), Medical Research Council (1000143 to EPIC-Norfolk; MR/M012190/1 to EPIC-Oxford). (United Kingdom).

Data availability and materials The data described in the manuscript will be made available upon request pending application and approval. The analytic code will be made publicly and freely available without restriction at https://github.com/fernandam93/UPF_adiposity_UADT.

Declarations

Conflict of interest Financial interests: TGR is an employee of Glaxo-SmithKline outside of this research. The rest of the authors have no financial interests to disclose. Non-financial interests: The authors have no non-financial interests to declare. The authors declare that no funds, grants or other support were received during the preparation of this manuscript.

Ethical approval The EPIC study was approved by the IARC Ethics Committee and all local ethics committees. This project was conducted under approval number 21–50.

Consent to participate Written informed consent was obtained from all individual participants included in the study.

IARC disclaimer Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Monteiro CA, Levy RB, Claro RM et al (2010) A new classification of foods based on the extent and purpose of their processing. *Cad Saude Publica* 26(11):2039–2049
2. Monteiro CA, Cannon G, Levy RB et al (2019) Ultra-processed foods: what they are and how to identify them. *Public Health Nutr* 22(5):936–941. <https://doi.org/10.1017/S1368980018003762>
3. Monteiro CA, Cannon G, Levy RB et al (2012) The food system ultra-processing: the big issue for nutrition, disease, health, well-being. *World Nutr J World Public Heal Nutr Assoc* 3(12):527–569
4. Monteiro CA, Levy RB, Claro RM et al (2011) Increasing consumption of ultra-processed foods and likely impact on human health: evidence from Brazil. *Public Health Nutr* 14(1):5–13. <https://doi.org/10.1017/S1368980010003241>
5. Fiolet T, Srour B, Sellem L et al (2018) Consumption of ultra-processed foods and cancer risk: results from NutriNet-Sante prospective cohort. *BMJ* 360:k322. <https://doi.org/10.1136/bmj.k322>
6. Romaguera D, Fernandez-Barres S, Gracia-Lavedan E et al (2021) Consumption of ultra-processed foods and drinks and colorectal, breast, and prostate cancer. *Clin Nutr* 40(4):1537–1545. <https://doi.org/10.1016/j.clnu.2021.02.033>
7. El Kinany K, Huybrechts I, Hatime Z et al (2022) Food processing groups and colorectal cancer risk in Morocco: evidence from a nationally representative case-control study. *Eur J Nutr* 61(5):2507–2515. <https://doi.org/10.1007/s00394-022-02820-3>
8. Chang K, Gunter MJ, Rauber F et al (2023) Ultra-processed food consumption, cancer risk and cancer mortality: a large-scale

- prospective analysis within the UK Biobank. *eClinicalMedicine*. <https://doi.org/10.1016/j.eclinm.2023.101840>
9. Kliemann N, Rauber F, Bertazzi Levy R et al (2023) Food processing and cancer risk in Europe: results from the prospective EPIC cohort study. *The Lancet Planetary Health* 7(3):e219–e232. [https://doi.org/10.1016/s2542-5196\(23\)00021-9](https://doi.org/10.1016/s2542-5196(23)00021-9)
 10. Sandoval-Insausti H, Jimenez-Onsurbe M, Donat-Vargas C et al (2020) Ultra-processed food consumption is associated with abdominal obesity: a prospective cohort study in older adults. *Nutrients* 12(8):2368. <https://doi.org/10.3390/nu12082368>
 11. Rauber F, Chang K, Vamos EP et al (2021) Ultra-processed food consumption and risk of obesity: a prospective cohort study of UK Biobank. *Eur J Nutr* 60(4):2169–2180. <https://doi.org/10.1007/s00394-020-02367-1>
 12. Moradi S, Entezari MH, Mohammadi H et al (2021) Ultra-processed food consumption and adult obesity risk: a systematic review and dose-response meta-analysis. *Crit Rev Food Sci Nutr*. <https://doi.org/10.1080/10408398.2021.1946005>
 13. Cordova R, Kliemann N, Huybrechts I et al (2021) Consumption of ultra-processed foods associated with weight gain and obesity in adults: a multi-national cohort study. *Clin Nutr* 40(9):5079–5088. <https://doi.org/10.1016/j.clnu.2021.08.009>
 14. Rudakoff LCS, Magalhaes E, Viola P et al (2022) Ultra-processed food consumption is associated with increase in fat mass and decrease in lean mass in Brazilian women: A cohort study. *Front Nutr* 9:1006018. <https://doi.org/10.3389/fnut.2022.1006018>
 15. World Cancer Research Fund/American Institute for Cancer Research (2018) Diet, nutrition, physical activity and oesophageal cancer. Continuous Update Project Expert Report.
 16. Steffen A, Huerta JM, Weiderpass E et al (2015) General and abdominal obesity and risk of esophageal and gastric adenocarcinoma in the European Prospective Investigation into Cancer and Nutrition. *Int J Cancer* 137(3):646–657. <https://doi.org/10.1002/ijc.29432>
 17. Hoyo C, Cook MB, Kamangar F et al (2012) Body mass index in relation to oesophageal and oesophagogastric junction adenocarcinomas: a pooled analysis from the International BEACON Consortium. *Int J Epidemiol* 41(6):1706–1718. <https://doi.org/10.1093/ije/dys176>
 18. O'Doherty MG, Freedman ND, Hollenbeck AR et al (2012) A prospective cohort study of obesity and risk of oesophageal and gastric adenocarcinoma in the NIH-AARP Diet and Health Study. *Gut* 61(9):1261–1268. <https://doi.org/10.1136/gutjnl-2011-300551>
 19. Lin Y, Ness-Jensen E, Hveem K et al (2015) Metabolic syndrome and esophageal and gastric cancer. *Cancer Causes Control* 26(12):1825–1834. <https://doi.org/10.1007/s10552-015-0675-4>
 20. Du X, Hidayat K, Shi BM (2017) Abdominal obesity and gastroesophageal cancer risk: systematic review and meta-analysis of prospective studies. *Biosci Rep*. <https://doi.org/10.1042/BSR20160474>
 21. Ward HA, Wark PA, Muller DC et al (2017) Measured adiposity in relation to head and neck cancer risk in the European Prospective Investigation into Cancer and Nutrition. *Cancer Epidemiol Biomarkers Prev* 26(6):895–904. <https://doi.org/10.1158/1055-9965.EPI-16-0886>
 22. Gaudet MM, Kitahara CM, Newton CC et al (2015) Anthropometry and head and neck cancer: a pooled analysis of cohort data. *Int J Epidemiol* 44(2):673–681. <https://doi.org/10.1093/ije/dyv059>
 23. World Cancer Research Fund/American Institute for Cancer Research (2018) Diet, nutrition, physical activity and cancers of the mouth, pharynx and larynx. Continuous Update Project Expert Report.
 24. Bingham S, Riboli E (2004) Diet and cancer—the European Prospective Investigation into Cancer and Nutrition. *Nat Rev Cancer* 4(3):206–215. <https://doi.org/10.1038/nrc1298>
 25. Riboli E, Hunt KJ, Slimani N et al (2002) European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* 5(6B):1113–1124. <https://doi.org/10.1079/PHN2002394>
 26. Riboli E, Kaaks R (1997) The EPIC project: rationale and study design. *Int J Epidemiol* 26(1):S6–14
 27. Kroke A, Klipstein-Grobusch K, Voss S et al (1999) Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* 70(4):439–447
 28. Margetts BM, Pietinen P (1997) European Prospective Investigation into Cancer and Nutrition: validity studies on dietary assessment methods. *Int J Epidemiol* 26(1):S1–S5
 29. Slimani N, Bingham S, Runswick S et al (2003) Group level validation of protein intakes estimated by 24-hour diet recall and dietary questionnaires against 24-hour urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. *Cancer Epidemiol Biomark Prev* 12:784–795
 30. Slimani N, Ferrari P, Ocké M et al (2000) Standardization of the 24-hour diet recall calibration method used in the European Prospective Investigation into Cancer and Nutrition (EPIC): general concepts and preliminary results. *Eur J Clin Nutr* 54:900–917
 31. Huybrechts I, Rauber F, Nicolas G et al (2022) Characterization of the degree of food processing in the European Prospective Investigation into Cancer and Nutrition: application of the Nova classification and validation using selected biomarkers of food processing. *Front Nutr* 9:1035580. <https://doi.org/10.3389/fnut.2022.1035580>
 32. Hashibe M, Brennan P, Benhamou S et al (2007) Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *J Natl Cancer Inst* 99(10):777–789. <https://doi.org/10.1093/jnci/djk179>
 33. Wareham NJ, Jakes RW, Rennie KL et al (2003) Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr* 6(4):407–413. <https://doi.org/10.1079/PHN2002439>
 34. Haftenberger M, Lahmann PH, Panico S et al (2002) Overweight, obesity and fat distribution in 50- to 64-year-old participants in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr* 5(6B):1147–1162. <https://doi.org/10.1079/PHN2002396>
 35. Spencer EA, Appleby PN, Davey GK et al (2002) Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutr* 5(4):561–565. <https://doi.org/10.1079/PHN2001322>
 36. Goyal N, Hennessy M, Lehman E et al (2022) Risk factors for head and neck cancer in more and less developed countries: analysis from the INHANCE consortium. *Oral Dis*. <https://doi.org/10.1111/odi.14196>
 37. Hashibe M, Brennan P, Chuang SC et al (2009) Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiol Biomarkers Prev* 18(2):541–550. <https://doi.org/10.1158/1055-9965.EPI-08-0347>
 38. Marziliano A, Teckie S, Diefenbach MA (2020) Alcohol-related head and neck cancer: summary of the literature. *Head Neck* 42(4):732–738. <https://doi.org/10.1002/hed.26023>
 39. Matejic M, Gunter MJ, Ferrari P (2017) Alcohol metabolism and oesophageal cancer: a systematic review of the evidence. *Carcinogenesis* 38(9):859–872. <https://doi.org/10.1093/carcin/bgx067>
 40. Freedman ND, Murray LJ, Kamangar F et al (2011) Alcohol intake and risk of oesophageal adenocarcinoma: a pooled analysis

- from the BEACON Consortium. *Gut* 60(8):1029–1037. <https://doi.org/10.1136/gut.2010.233866>
41. Wilson LF, Baade PD, Green AC et al (2019) The impact of reducing alcohol consumption in Australia: an estimate of the proportion of potentially avoidable cancers 2013–2037. *Int J Cancer* 145(11):2944–2953. <https://doi.org/10.1002/ijc.32204>
 42. Gormley M, Dudding T, Sanderson E et al (2020) A multivariable Mendelian randomization analysis investigating smoking and alcohol consumption in oral and oropharyngeal cancer. *Nat Commun* 11(1):6071. <https://doi.org/10.1038/s41467-020-19822-6>
 43. Boccia S, Hashibe M, Galli P et al (2009) Aldehyde dehydrogenase 2 and head and neck cancer: a meta-analysis implementing a Mendelian randomization approach. *Cancer Epidemiol Biomarkers Prev* 18(1):248–254. <https://doi.org/10.1158/1055-9965.EPI-08-0462>
 44. Weikert C, Dietrich T, Boeing H et al (2009) Lifetime and baseline alcohol intake and risk of cancer of the upper aero-digestive tract in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Int J Cancer* 125(2):406–412. <https://doi.org/10.1002/ijc.24393>
 45. Wozniak MB, Brennan P, Brenner DR et al (2015) Alcohol consumption and the risk of renal cancers in the European prospective investigation into cancer and nutrition (EPIC). *Int J Cancer* 137(8):1953–1966. <https://doi.org/10.1002/ijc.29559>
 46. VanderWeele TJ (2011) Causal mediation analysis with survival data. *Epidemiology* 22(4):582–585. <https://doi.org/10.1097/EDE.0b013e31821db37e>
 47. Shi B, Choirat C, Coull BA et al (2021) CMAverse: a suite of functions for reproducible causal mediation analyses. *Epidemiology* 32(5):e20–e22
 48. VanderWeele TJ (2014) A unification of mediation and interaction: a 4-way decomposition. *Epidemiology* 25(5):749–761. <https://doi.org/10.1097/EDE.0000000000000121>
 49. van Buuren S, Groothuis-Oudshoorn K (2011) Mice: multivariate imputation by chained equations in R. *J Stat Softw* 45(3):1–67
 50. Rubin DB (1987) Multiple imputation for nonresponse in surveys. John Wiley and Sons, New York
 51. Lipsitch M, Tchetgen Tchetgen E, Cohen T (2010) Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* 21(3):383–388. <https://doi.org/10.1097/EDE.0b013e3181d61eeb>
 52. VanderWeele TJ, Robins JM (2007) Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am J Epidemiol* 166(9):1096–1104. <https://doi.org/10.1093/aje/kwm179>
 53. WHO (2011) Waist Circumference and Waist–Hip Ratio: Report of a WHO Expert Consultation, Geneva, 8–11 December 2008.
 54. Monteiro CA (2009) Nutrition and health. the issue is not food, nor nutrients, so much as processing. *Public Health Nutr* 12(5):729–731. <https://doi.org/10.1017/S1368980009005291>
 55. Fardet A (2016) Minimally processed foods are more satiating and less hyperglycemic than ultra-processed foods: a preliminary study with 98 ready-to-eat foods. *Food Funct* 7(5):2338–2346. <https://doi.org/10.1039/c6fo00107f>
 56. Zinocker MK, Lindseth IA (2018) The western diet-microbiome-host interaction and its role in metabolic disease. *Nutrients* 10(3):365. <https://doi.org/10.3390/nu10030365>
 57. Martinez KB, Leone V, Chang EB (2017) Western diets, gut dysbiosis, and metabolic diseases: are they linked? *Gut Microbes* 8(2):130–142. <https://doi.org/10.1080/19490976.2016.1270811>
 58. Elliott JA, Reynolds JV (2021) Visceral obesity, metabolic syndrome, and esophageal adenocarcinoma. *Front Oncol* 11:627270. <https://doi.org/10.3389/fonc.2021.627270>
 59. Dicken SJ, Batterham RL (2021) The role of diet quality in mediating the association between ultra-processed food intake, obesity and health-related outcomes: a review of prospective cohort studies. *Nutrients* 14(1):23. <https://doi.org/10.3390/nu14010023>
 60. Davey Smith G, Holmes MV, Davies NM et al (2020) Mendel’s laws, Mendelian randomization and causal inference in observational data: substantive and nomenclatural issues. *Eur J Epidemiol* 35(2):99–111. <https://doi.org/10.1007/s10654-020-00622-7>
 61. Fewell Z, Davey Smith G, Sterne JA (2007) The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol* 166(6):646–655. <https://doi.org/10.1093/aje/kwm165>
 62. Davey Smith G, Phillips AN (2020) Correlation without a cause: an epidemiological odyssey. *Int J Epidemiol* 49(1):4–14. <https://doi.org/10.1093/ije/dyaa016>
 63. Chazelas E, Druesne-Pecollo N, Esseddik Y et al (2021) Exposure to food additive mixtures in 106,000 French adults from the NutriNet-Sante cohort. *Sci Rep* 11(1):19680. <https://doi.org/10.1038/s41598-021-98496-6>
 64. Chazelas E, Srour B, Desmetz E et al (2019) Sugary drink consumption and risk of cancer: results from NutriNet-Sante prospective cohort. *BMJ* 366:12408. <https://doi.org/10.1136/bmj.12408>
 65. Debras C, Chazelas E, Srour B et al (2022) Artificial sweeteners and cancer risk: results from the NutriNet-Sante population-based cohort study. *PLoS Med* 19(3):e1003950. <https://doi.org/10.1371/journal.pmed.1003950>
 66. Marino M, Puppo F, Del Bo C et al (2021) A systematic review of worldwide consumption of ultra-processed foods: findings and criticisms. *Nutrients* 13(8):2778. <https://doi.org/10.3390/nu13082778>
 67. Baker P, Machado P, Santos T et al (2020) Ultra-processed foods and the nutrition transition: Global, regional and national trends, food systems transformations and political economy drivers. *Obes Rev* 21(12):e13126. <https://doi.org/10.1111/obr.13126>

Authors and Affiliations

Fernanda Morales-Berstein^{1,2}  · Carine Biessy³ · Vivian Viallon³ · Ana Goncalves-Soares^{1,2} · Corinne Casagrande³ · Bertrand Hémon³ · Nathalie Kliemann^{3,4} · Manon Cairat^{3,5} · Jessica Blanco Lopez³ · Aline Al Nahas³ · Kiara Chang⁶ · Eszter Vamos⁶ · Fernanda Rauber^{7,8} · Renata Bertazzi Levy^{7,8} · Diana Barbosa Cunha⁹ · Paula Jakszyn^{10,11,12} · Pietro Ferrari³ · Paolo Vineis^{13,14} · Giovanna Masala¹⁵ · Alberto Catalano¹⁶  · Emily Sonestedt¹⁷ · Yan Borné¹⁷ · Verena Katzke¹⁸ · Rashmita Bajracharya¹⁸ · Claudia Agnoli¹⁹ · Marcela Guevara^{20,21,22} · Alicia Heath²³ · Loredana Radoi⁵ · Francesca Mancini⁵ · Elisabete Weiderpass²⁴ · José María Huerta^{21,25} · María-José Sánchez^{21,26,27,28} · Anne Tjønneland^{29,30} · Cecilie Kyrø²⁹ · Matthias B. Schulze^{31,32} · Guri Skeie³³ · Marko Lukic³³ · Tonje Braaten³³ · Marc Gunter^{3,23} · Christopher Millett^{6,34} · Antonio Agudo^{10,11} · Paul Brennan³⁵ · M. Carolina Borges^{1,2} · Rebecca C. Richmond^{1,2} · Tom G. Richardson^{1,2} · George Davey Smith^{1,2} · Caroline L. Relton^{1,2} · Inge Huybrechts³ · on behalf of the EPIC Network

- ✉ Fernanda Morales-Berstein
dy20206@bristol.ac.uk
- 1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
 - 2 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
 - 3 Nutrition and Metabolism Branch, International Agency for Research on Cancer/World Health Organization, Lyon, France
 - 4 Cancer Research Center of Santa Catarina, CEPON, Florianópolis, Brazil
 - 5 Paris-Saclay University, UVSQ, Inserm “Exposome, Heredity, Cancer and Health” Team, CESP U1018, Gustave Roussy, Villejuif, France
 - 6 Public Health Policy Evaluation Unit, School of Public Health, Imperial College London, London, UK
 - 7 Preventive Medicine Department of the Medical School, University of São Paulo, São Paulo, Brazil
 - 8 Center for Epidemiological Research in Nutrition and Health, University of São Paulo, São Paulo, Brazil
 - 9 Hésio Cordeiro Institute of Social Medicine, Department of Epidemiology, Rio de Janeiro State University, Rio de Janeiro, RJ, Brazil
 - 10 Unit of Nutrition and Cancer, Catalan Institute of Oncology-ICO, L’Hospitalet de Llobregat, Spain
 - 11 Nutrition and Cancer Group; Epidemiology, Public Health, Cancer Prevention and Palliative Care Program, Bellvitge Biomedical Research Institute-IDIBELL, L’Hospitalet de Llobregat, Spain
 - 12 Blanquerna Faculty of Health Sciences, Ramon Llull University, Barcelona, Spain
 - 13 MRC Centre for Environment and Health, School of Public Health, Imperial College London, Norfolk Place, London W2 1PG, UK
 - 14 Italian Institute of Technology, Genoa, Italy
 - 15 Institute for Cancer Research, Prevention and Clinical Network (ISPRO), Florence, Italy
 - 16 Centre for Biostatistics, Epidemiology, and Public Health, Department of Clinical and Biological Sciences, University of Turin, 10043 Orbassano, TO, Italy
 - 17 Nutrition Epidemiology, Department of Clinical Sciences Malmö, Faculty of Medicine, Lund University, Lund, Sweden
 - 18 Department of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
 - 19 Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
 - 20 Instituto de Salud Pública y Laboral de Navarra, 31003 Pamplona, Spain
 - 21 Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP), 28029 Madrid, Spain
 - 22 Navarra Institute for Health Research (IdiSNA), 31008 Pamplona, Spain
 - 23 Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK
 - 24 International Agency for Research on Cancer, World Health Organization, Lyon, France
 - 25 Department of Epidemiology, Murcia Regional Health Council-IMIB, Murcia, Spain
 - 26 Escuela Andaluza de Salud Pública (EASP), 18011 Granada, Spain
 - 27 Instituto de Investigación Biosanitaria ibs.GRANADA, 18012 Granada, Spain
 - 28 Department of Preventive Medicine and Public Health, University of Granada, 18071 Granada, Spain
 - 29 Danish Cancer Society Research Center, Diet, Cancer and Health, Strandboulevarden 49, DK-2100 Copenhagen, Denmark
 - 30 Department of Public Health, University of Copenhagen, DK-2200 Copenhagen, Denmark
 - 31 Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany
 - 32 Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany
 - 33 Department of Community Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway, Tromsø, Norway
 - 34 NOVA National School of Public Health, Public Health Research Centre, Comprehensive Health Research Center, CHRC, NOVA University Lisbon, Lisbon, Portugal
 - 35 Genetic Epidemiology Group, International Agency for Research on Cancer, World Health Organization, Lyon, France