**Supplementary Material**

[Supplementary Methods 2](#_Toc125096701)

[Supplementary Table 1: Specific food items comprising the ultra-processed food subgroups in the Nurses’ Health Study, the Nurses’ Health Study II and the Health Professionals’ Follow-up Study. 8](#_Toc125096702)

[Supplementary Table 2: Systematic review search strategy.1 10](#_Toc125096703)

[Supplementary Table 3: Age-standardized intakes of UPF of participants in the Nurses’ Health Study (1998), the Nurses’ Health Study II (1999) and the Health Professionals Follow-up Study (1998) (pooled data from the 3 cohorts). 11](#_Toc125096704)

[Supplementary Table 4: Multivariable adjusted hazard ratios (95% confidence intervals) for incident type 2 diabetes according to total ultra-processed food consumption (servings/d) stratified by key variables. 12](#_Toc125096705)

[Supplementary Table 5: Hazard ratio (95% confidence intervals) for incident type 2 diabetes according to alternative metrics of total ultra-processed food consumption. 14](#_Toc125096706)

[Supplementary Table 6: Hazard ratio (95% confidence intervals) for incident type 2 diabetes or symptomatic diabetes according to total ultra-processed food consumption using alternative covariable modeling. 17](#_Toc125096707)

[Supplementary Table 7: Mediation effect of key nutrients in the relationship between ultra-processed food and T2D risk. 19](#_Toc125096708)

[Supplementary Table 8: Characteristics of studies included in the meta-analysis.1 21](#_Toc125096709)

[Supplementary Table 9: List of confounders among studies included in the meta-analysis. 24](#_Toc125096710)

[Supplementary Table 10: Assessment of risk of bias with the Newcastle-Ottawa Scale.1 25](#_Toc125096711)

[Supplementary Table 11: Pre-specified subgroup meta-analyses of the association between total ultra-processed food consumption and risk of type 2 diabetes, using random-effects models.1 26](#_Toc125096712)

[Supplementary Table 12: Assessment of the quality of evidence on the relationship between total ultra-processed food consumption and risk of type 2 diabetes using the NutriGrade scoring system.1 27](#_Toc125096713)

[Supplementary Figure 1: Flowchart of participants. 28](#_Toc125096714)

[Supplementary Figure 2: Meta-analysis search strategy and study selection. 29](#_Toc125096715)

[Supplementary Figure 3: Funnel plot for assessment of publication bias for the association between total ultra-processed food consumption and risk of type 2 diabetes. 30](#_Toc125096716)

[Supplementary Figure 4: Forest plot of influence analysis for the association between total ultra-processed food consumption and risk of type 2 diabetes. 31](#_Toc125096717)

[Supplementary references 32](#_Toc125096718)

# Supplementary Methods

***Nova classification and ultra-processed food (UPF) categorization***

The Nova classification was defined as a system that classifies food based on the extent and purpose of the industrial processing they undergo and accounts for the physical, biological and chemical methods used in their manufacture, including the use of additives. UPF were defined as industrial formulations made mostly or entirely with substances extracted from foods, often chemically modified, and from additives, with little if any whole foods added and specifically identified based on the criteria provided by Monteiro in 2019 (1). Specifically, the approach to the categorization of all food items was driven by whether or not they may have the additives listed in Monteiro et al. 2019 (1), which serve as proxy indicators of ultra-processing. The additives are usually some food substances not used in household kitchens, including hydrolyzed proteins, soya protein isolate, gluten, casein, whey protein, mechanically separated meat, fructose, high-fructose corn syrup, fruit juice concentrate, invert sugar, maltodextrin, dextrose, lactose, soluble or insoluble fiber, hydrogenated or unesterified oil; and also other sources of macronutrients which are neither foods from Nova group 1 or group 3, nor culinary ingredients from Nova group 2. Nutrient composition or the presence of high sodium, saturated fat or added sugars are not considered as additives for UPF with Nova. The categorization of food using the Nova system is informed by the granularity of the information available on the specific food items.

In the process of categorizing food item from the FFQ used in the NHS, NHSII and HPFS per the Nova classification, a multi-stage process was undertaken. In the first phase of categorization, 3 individuals independently assigned a Nova classification to all food items. The classifications were subsequently compared and allowed to identify that over 70% of all items included in the FFQ were categorized within the same group by the 3 reviewers. These foods include, for instance, “raw carrots” (minimally processed food), “hummus” (processed food), “beef, pork hot dogs” (ultra-processed food).

For the remaining 20% food items, the team consulted with research dieticians; used cohort-specific recipe files that contained information on how the food was processed, or specific ingredients used; and scanned online supermarket to better gauge the processing level of the foods. For instance, recipe files supported the categorization of “hamburgers”, “French fried potatoes”, “brownies”, and “pizzas*”* into the appropriate Nova group. No assumptions were made in the categorization of these foods. Their categorization was backed by information available from the sources mentioned above.

Only 9 food items had no unambiguous source of information to support a categorization. The 9 food items are: “popcorn”; “soy milk”; “cream”; “pancakes or waffles”; “pie, home-baked or ready-made”; “chicken sandwich”; “beef, pork, lamb sandwich”; “tomato sauce”; “potato or corn chips”. The team categorized these products into a lower processing category as a conservative approach. The team also recommended to reclassify these foods into the UPF category as sensitivity analyses. This approach was used in the current analysis on type 2 diabetes.

The full description of how the Nova classification was applied in the NHS, NHSII and HPFS FFQs is available from Khandpur et al. 2021 (2).

In the three cohorts, the correlation coefficients between total UPF (servings/day and energy percent from UPF) and the Alternative Healthy Eating Index (AHEI) were -0.27, and -0.45, respectively.

***Systematic review and meta-analysis of UPF consumption and T2D***

We conducted a systematic review and meta-analysis based on the current study and previous prospective cohort studies that evaluated the association between total UPF consumption and T2D risk in the general adult population. The report was conducted using the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines (3). We registered the protocol on the international prospective register of systematic reviews (PROSPERO CRD42022337267). **Supplementary Table 2** presents the strategy used to search *Medline* (via *Pubmed*), *Embase*, and *Web of Science* up to June 6th, 2022. We screened the reference lists of selected studies to identify additional relevant studies. Studies were included if they were of prospective design; if they assessed the association between UPF consumption and incidence of T2D; and if they provided risk estimates for three or more levels of total UPF consumption or a dose-response estimate. Extracted data included first author name, publication year, cohort name, country where the study was conducted, follow-up duration, number of participants, sex, age range at baseline, method used to assess diet, and method used to identify events. Additionally, we obtained data about the number of T2D events, categories of UPF consumption, risk estimates (95% CIs) from the multivariable model, and covariates in the maximally adjusted model.

We used the Newcastle-Ottawa scale to assess the risk of bias (ROB) in included studies (4). The Newcastle-Ottawa scale contains nine items. Each item is scored (0-1 point) according to specific criteria. The maximum possible score is nine. Scores from 7-9 reflect low ROB, scores from 4-6 indicate moderate ROB, and scores 0-3 indicate high ROB. Age, sex, BMI, smoking status, physical activity, alcohol intake, and total energy intake were considered primary confounders of the association between total UPF consumption and T2D risk. Intakes of major foods (e.g., non-ultra-processed fruits, vegetables, whole grains, coffee and tea) or diet quality were considered as secondary confounders. Two authors (ZC and CD) independently screened the literature (title and abstract, then full article) to retrieve relevant studies. Disagreement and discordance were resolved by consensus between the two authors.

Relative risks were used as the common measure of association across studies. HRs and odd ratios (ORs) were considered equivalent to RRs. As most of these previous studies used percent of grams of total UPF intake relative to total diet weight as unit to quantify total UPF intake, we also used percent of grams of UPF intake harmonize exposures between studies and quantify the amount of UPF intake in the meta-analysis. For one study that reported total UPF intake according to absolute grams, we converted grams to percent of grams of UPF (5). To determine the intake levels, we used the median of each UPF intake category, if available, or the midpoint between the lower and upper bound of each category of intake. When the highest category was open, we estimated the intake by multiplying the lower bound of the highest category by 1.75. One study (6) did not provide person-years data. We thus imputed this metric based on available data. For one study that only presented HR and 95% CI of per 10% increment of grams of UPF intake (7), we contacted the authors by email to obtain HRs and 95% CIs of the quintiles.

First, we conducted a meta-analysis of high vs low total UPF intake by pooling the risk estimates from the highest category of intake compared with the lowest category reported in each study using random-effects models. Secondly, we performed a linear dose-response meta-analysis to assess T2D risk associated with a 10% increment of total UPF intake (percent of grams per day) by using the 2-stage generalized least squares trend estimation method, which first estimated study-specific slope lines, and then combined the studies with estimated study-specific slopes with studies in which the slopes were directly reported to obtain an overall average slope (8). Thirdly, a second 2-stage, random effects dose-response meta-analysis was conducted to assess potential nonlinearity of the association between total UPF intake (percent of grams per day) and T2D risk. Restricted cubic splines with 3 knots (set at 10th, 50th, and 90th percentiles for UPF intake) were used to model the association of interest. Using the Orsini method (9), we calculated restricted cubic spline models by a generalized least squares trend estimation method, which takes into account the correlation within each set of reported risk estimates. The study specific estimates were then combined by the restricted maximum likelihood method in a multivariate random effects meta-analysis. A probability value for non-linearity was estimated by null hypothesis testing, in which the coefficient of the second spline was considered equal to zero.

Between-study heterogeneity in the summary estimate was assessed with the *I*2 statistic and interpreted according to the Cochrane Handbook thresholds (0-40%, might not be important; 30-60%, might represent moderate heterogeneity; 50-90%, might represent substantial heterogeneity; 75-100%, considerable heterogeneity) (10). We conducted an influence analysis by systematically removing each study from the meta-analysis and calculating the relative risk to evaluate if any single study caused the heterogeneity. Age, follow-up duration, number of participants, number of events, geographical location, ROB, adjustment for total energy intake, adjustment for hypertension or hypercholesterolemia and dietary assessment method were identified a priori as potential sources of heterogeneity. We tested for publication bias by using Begg’s test and Egger’s test, and visual appreciation of a funnel plot. Statistical analyses for the meta-analysis were preformed using Stata version 17.0 (StataCorp, College Station, TX).

**Assessment of the quality of evidence**

We used the NutriGrade scoring system for cohort studies to assess the quality of evidence supporting the association of total UPF intake and T2D risk (11). The NutriGrade scoring system for meta-analyses of cohort studies includes the eight following items: *1)* ROB, study quality, and study limitations (0–2 points); *2)* precision (0–1 point); *3)* heterogeneity (0–1 point); *4)* directness (0–1 point); *5)* publication bias (0–1 point); *6)* funding bias (0–1 point); *7)* effect size (0–2 points); and *8)* dose-response (0-1 point) (11). A NutriGrade score ≥8/10 points indicates “high-quality evidence,” meaning that there is high confidence in the effect estimate, and further research probably will not change the confidence in the effect estimate (34). Other thresholds are 6–7.99 points for “moderate-quality evidence,” 4–5.99 points for “low-quality evidence,” and 0–3.99 for “very low-quality evidence”. Very low-quality evidence refers to very low confidence in the effect estimate (11). The assessment was conducted independently by two authors (ZC and JPDC). Disagreement and discordance in the scoring were resolved by discussion.

# Supplementary Table 1: Specific food items comprising the ultra-processed food subgroups in the Nurses’ Health Study, the Nurses’ Health Study II and the Health Professionals’ Follow-up Study.

|  |  |
| --- | --- |
| **Ultra-processed food subgroups** | **Specific food items** |
|
| **1.** Ultra-processed breads and cereals |  |
| **1a.** Ultra-processed cereals | Cold breakfast cereal; |
| **1b.** Ultra-processed dark and whole-grain breads | Rye, pumpernickel bread;Whole-grain bread; |
| **1c.** Other ultra-processed refined-grain breads | English muffins, bagels, rolls;White bread; |
| **2.** Sauces, spreads, and condiments | Cream cheese;Ketchup;Margarine;Mayonnaise (regular and low fat);Non-dairy coffee whitener;Red chili sauce;Salad dressings;Salsa;Soy sauce;Spread butter; |
| **3.** Packaged sweet snacks and desserts |  |
| **3a.** Confectioneries | Candy bar with chocolate;Candy bar without chocolate;Chocolate bars;Dark chocolate bars; |
| **3b.** Packaged sweet snacks | Breakfast bar;Brownies;Cookies, brownie, ready-made;Cookies, fat free, reduced fat;Doughnuts;Energy bars;High protein, low carb candy bars;Muffins or biscuits;Ready-made cakes;Ready-made cookies;Ready-made sweet rolls and coffee cakes; |
| **3c.** Fruit-based products | Applesauce;Canned peaches;Canned pears;Jams, jellies, preserves, honey;Ready-made pies; |
| **4.** Packaged savory snacks | Fat-free popcorn;Fat-free, light crackers;Regular crackers; |
| **5.** Artificially and sugar-sweetened beverages |  |
| **5a.** Artificially sweetened beverages  | Caffeinated, caffeine-free, carbonated and noncarbonated low-calorie soda; Other low-calorie Cola with caffeine; |
| **5b.** Sugar-sweetened beverages | 7-up;Coke or Pepsi with caffeine & sugar;Coke or Pepsi without caffeine but with sugar;Hawaiian punch with sugar;Other carbonated beverage;Dairy coffee drinks; |
| **6.** Animal-based products | Bacon;Beef, pork hotdogs;Breaded fish cakes, pieces, sticks;Chicken or turkey hot dogs;Processed meats, sausages;Salami, bologna, processed meat sandwiches; |
| **7.** Ready-to-eat/heat mixed dishes | Chowder or cream soup;French fries potatoes;Pizza;Ready-made soup from cans;Soup made with bouillon; |
| **8.** Yogurt and dairy based desserts | Artificially sweetened yogurt;Flavored yogurt without Nutrasweet;Frozen yogurt;Ice cream;Ice cream;Sherbet; |
| **9.** Other ultra-processed foods | Nutrasweet or equal;Other artificial sweeteners;Splenda; |

# Supplementary Table 2: Systematic review search strategy.1

|  |  |  |  |
| --- | --- | --- | --- |
| **Database** | **Step** | **Search terms** | **Results, n** |
| PubMed | 1 | Ultra-processed food OR NOVA | 50,161 |
| 2 | “Diabetes Mellitus"[MeSH] OR “Diabetes Mellitus, Type 2"[MeSH] OR diabetes OR T2D OR insulin resistance | 943,061 |
| 3 | Prospective OR Cohort OR Longitudinal OR Follow-up OR Case-cohort OR Nested Case-control | 3,548,276 |
| 4 | (#1 AND #2 AND #3) Filters: Humans | 408 |
| Web of Science | 1 | (((TS=(ultra-processed food )) OR TS=(nova)) AND (TS=(diabetes) OR TS=(diabetes mellitus) OR TS=(t2d) OR TS=(insulin resistance)) AND (TS=(prospective) OR TS=(cohort) OR TS=(longitudinal) OR TS=(follow-up) OR TS=(case-cohort) OR TS=(nested case-control))) | 80 |
| Embase | 1 | ('ultra processed' AND ('food'/exp OR food) OR 'nova'/exp OR nova) AND [embase]/lim | 42,619 |
| 2 | ('diabetes mellitus'/exp OR diabetes.ab. OR 'type 2 diabetes.ab.') AND [embase]/lim | 1,006,492 |
| 3 | (prospective OR cohort OR longitudinal OR 'follow up' OR 'case cohort' OR 'nested case control') AND [embase]/lim | 3,775,311 |
| 4 | 'human' | 25,874,978 |
| 5 | (#1 AND #2 AND #3 AND #4) | 529 |

1 Search date: 2022-06-06.

# Supplementary Table 3: Age-standardized intakes of UPF of participants in the Nurses’ Health Study (1998), the Nurses’ Health Study II (1999) and the Health Professionals Follow-up Study (1998) (pooled data from the 3 cohorts).

|  |  |
| --- | --- |
| **Subgroups**  | **Quintiles of ultra-processed foods (UPF) intake (servings/day)** |
| **Q1** | **Q2** | **Q3** | **Q4** | **Q5** |
| Ultra-processed breads and cereals | 0.9 (0.5) | 1.3 (0.6) | 1.6 (0.8) | 2.0 (0.9) | 2.6 (1.3) |
| *Ultra-processed cereals* | 0.3 (0.3) | 0.3 (0.3) | 0.4 (0.3) | 0.4 (0.4) | 0.4 (0.4) |
| *Ultra-processed dark and whole-grain breads* | 0.3 (0.3) | 0.5 (0.5) | 0.6 (0.6) | 0.8 (0.7) | 1.0 (0.9) |
| *Other ultra-processed* refined-grain *breads* | 0.3 (0.3) | 0.5 (0.4) | 0.6 (0.5) | 0.8 (0.7) | 1.2 (1.0) |
| Sauces, spreads, and condiments | 0.7 (0.5) | 1.1 (0.6) | 1.4 (0.8) | 1.7 (1.0) | 2.4 (1.5) |
| Packaged sweet snacks and desserts | 0.5 (0.4) | 0.8 (0.5) | 1.0 (0.6) | 1.3 (0.8) | 1.9 (1.2) |
| *Confectioneries* | 0.1 (0.2) | 0.2 (0.2) | 0.3 (0.3) | 0.3 (0.4) | 0.5 (0.6) |
| *Packaged sweet snacks* | 0.3 (0.2) | 0.4 (0.3) | 0.5 (0.4) | 0.7 (0.6) | 1.0 (0.9) |
| *Fruit-based products* | 0.1 (0.2) | 0.2 (0.2) | 0.2 (0.3) | 0.3 (0.3) | 0.4 (0.5) |
| Packaged savory snacks | 0.1 (0.2) | 0.2 (0.2) | 0.2 (0.3) | 0.3 (0.4) | 0.5 (0.8) |
| Artificially and sugar-sweetened beverages | 0.4 (0.4) | 0.7 (0.6) | 1.0 (0.8) | 1.2 (1.1) | 1.8 (1.6) |
| *Sugar-sweetened beverages* | 0.2 (0.3) | 0.3 (0.4) | 0.3 (0.5) | 0.4 (0.6) | 0.5 (0.8) |
| *Artificially sweetened beverages* | 0.3 (0.4) | 0.5 (0.6) | 0.6 (0.8) | 0.8 (1.0) | 1.3 (1.6) |
| Animal-based products | 0.2 (0.2) | 0.2 (0.2) | 0.3 (0.2) | 0.3 (0.3) | 0.4 (0.4) |
| Ready-to-eat/heat mixed dishes | 0.2 (0.1) | 0.2 (0.1) | 0.2 (0.2) | 0.3 (0.2) | 0.3 (0.2) |
| Yogurt and dairy-based desserts | 0.2 (0.2) | 0.3 (0.2) | 0.3 (0.3) | 0.4 (0.4) | 0.5 (0.6) |
| Other ultra-processed foods | 0.0 (0.1) | 0.1 (0.2) | 0.1 (0.3) | 0.2 (0.4) | 0.3 (0.7) |

Values are means (standard deviation).

# Supplementary Table 4: Multivariable adjusted hazard ratios (95% confidence intervals) for incident type 2 diabetes according to total ultra-processed food consumption (servings/d) stratified by key variables.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Total ultra-processed food consumption (servings/day)** | ***P* value for trend** | ***P* value for interaction** |
| **Quintile 1 (low)** | **Quintile 2** | **Quintile 3** | **Quintile 4** | **Quintile 5 (high)** |
| **Age** |  |  |  |  |  |  |  |
| < 65 years | 1.00(reference) | 1.15(1.07, 1.22) | 1.24(1.16, 1.32) | 1.26(1.18, 1.34) | 1.35(1.26, 1.45) | <0.0001 | <0.0001 |
| ≥ 65 years | 1.00(reference) | 1.01(0.93, 1.10) | 1.12(1.02, 1.22) | 1.12(1.02, 1.23) | 1.18(1.07, 1.30) | 0.0003 |  |
| **Sex** |  |  |  |  |  |  |  |
| Male  | 1.00(reference) | 1.01(0.90, 1.13) | 1.05(0.94, 1.18) | 1.14(1.01, 1.29) | 1.22(1.07, 1.39) | 0.0004 | 0.09 |
| Female | 1.00(reference) | 1.12(1.05, 1.18) | 1.22(1.15, 1.29) | 1.22(1.15, 1.30) | 1.30(1.22, 1.39) | <0.0001 |  |
| **BMI** |  |  |  |  |  |  |  |
| < 25 kg/m2 | 1.00(reference) | 0.99(0.87, 1.12) | 1.02(0.90, 1.15) | 1.08(0.95, 1.23) | 1.10(0.95, 1.26) | 0.10 | 0.12 |
| ≥ 25 kg/m2 | 1.00(reference) | 1.10(1.04, 1.17) | 1.18(1.12, 1.25) | 1.17(1.10, 1.24) | 1.22(1.15, 1.30) | <0.0001 |  |
| **AHEI score** |  |  |  |  |  |  |  |
| < median level | 1.00(reference) | 1.05(0.97, 1.14) | 1.11(1.03, 1.21) | 1.13(1.04, 1.22) | 1.20(1.10, 1.30) | <0.0001 | 0.70 |
| ≥ median level | 1.00(reference) | 1.10(1.02, 1.18) | 1.21(1.12, 1.30) | 1.21(1.12, 1.30) | 1.27(1.17, 1.38) | <0.0001 |  |
| **Physical activity** |  |  |  |  |  |  |  |
| < median level | 1.00(reference) | 1.12(1.05, 1.20) | 1.22(1.13, 1.31) | 1.23(1.15, 1.33) | 1.29(1.20, 1.40) | <0.0001 | 0.70 |
| ≥ median level | 1.00(reference) | 1.05(0.97, 1.13) | 1.13(1.05, 1.22) | 1.14(1.05, 1.23) | 1.23(1.13, 1.34) | <0.0001 |  |
| **Smoking** |  |  |  |  |  |  |  |
| Never | 1.00(reference) | 1.09(1.01, 1.17) | 1.18(1.09, 1.26) | 1.19(1.11, 1.29) | 1.27(1.18, 1.38) | <0.0001 | 0.95 |
| Ever or current | 1.00(reference) | 1.10(1.02, 1.19) | 1.21(1.12, 1.30) | 1.21(1.12, 1.31) | 1.30(1.19, 1.40) | <0.0001 |  |
| **Family history of diabetes** |  |  |  |  |  |  |  |
| No | 1.00(reference) | 1.11(1.03, 1.19) | 1.22(1.13, 1.31) | 1.23(1.15, 1.33) | 1.30(1.20, 1.40) | <0.0001 | 0.36 |
| Yes | 1.00(reference) | 1.07(0.99, 1.16) | 1.14(1.05, 1.23) | 1.15(1.06, 1.25) | 1.25(1.15, 1.36) | <0.0001 |  |

Results were obtained from the pooled multivariable model 2 stratified by calendar time (in 2-year intervals) and cohort (sex), and adjusted for age, race/ethnicity (white/other), family history of diabetes (yes/no), history of hypercholesterolemia at baseline (yes/no), history of hypertension at baseline (yes/no), baseline BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0), smoking status (never, past, current), physical activity (MET-hours/week: <3.0, 3.0-8.9, 9.0-17.9, 18.0-26.9, ≥27.0), oral contraceptive use (never, former, current, in NHSII only), postmenopausal hormone use (premenopausal, never, former, current, in NHS and NHSII only), physical examination in the past 2 years (yes/no), neighborhood income (quintiles), total energy (kcal/d, quintiles), and total alcohol consumption (g/d, quintiles). All covariables (except race/ethnicity, family history of diabetes, baseline hypercholesterolemia, hypertension, and baseline BMI) were updated every 2 years. *P-*values for interaction were calculated using the likelihood ratio test. Given the potential for multiple testing, the statistical level for significance was set at 0.007 (0.05/7 comparisons).

# Supplementary Table 5: Hazard ratio (95% confidence intervals) for incident type 2 diabetes according to alternative metrics of total ultra-processed food consumption.

|  |  |  |  |
| --- | --- | --- | --- |
| **Models** | **Ultra-processed food consumption** | ***P* value for trend** | **Hazard ratio (95% CI) for continuous level of intake†** |
| **Quintile 1 (low)** | **Quintile 2** | **Quintile 3** | **Quintile 4** | **Quintile 5 (high)** |
| **Servings/day with alternative categorization for 9 food items‡**  |  |  |  |  |  |  |  |
| NHS | 1.00 (reference) | 1.04(0.96, 1.12) | 1.12(1.03, 1.21) | 1.14(1.05, 1.24) | 1.19(1.09, 1.30) | <0.0001 | 1.03(1.02, 1.04) |
| NHSII | 1.00 (reference) | 1.12(1.03, 1.23) | 1.24(1.13, 1.35) | 1.25(1.14, 1.37) | 1.40(1.27, 1.53) | <0.0001 | 1.03(1.02, 1.04) |
| HPFS | 1.00 (reference) | 0.99(0.89, 1.11) | 1.12(0.99, 1.26) | 1.14(1.00, 1.29) | 1.27(1.11, 1.45) | <0.0001 | 1.02(1.01, 1.04) |
| Pooled | 1.00 (reference) | 1.06(1.01, 1.12) | 1.16(1.10, 1.22) | 1.17(1.11, 1.24) | 1.27(1.20, 1.35) | <0.0001 | 1.03(1.02, 1.03) |
| **Calories from UPF/day** |  |  |  |  |  |  |  |
| NHS | 1.00 (reference) | 1.01(0.93, 1.09) | 1.08(1.00, 1.18) | 1.15(1.05, 1.25) | 1.20(1.09, 1.32) | <0.0001 | 1.03(1.02, 1.05) |
| NHSII | 1.00 (reference) | 1.05(0.96, 1.14) | 1.03(0.94, 1.13) | 1.04(0.94, 1.14) | 1.23(1.10, 1.36) | <0.0001 | 1.04(1.03, 1.05) |
| HPFS | 1.00 (reference) | 1.02(0.91, 1.14) | 1.13(1.01, 1.28) | 1.19(1.05, 1.35) | 1.30(1.12, 1.50) | 0.0001 | 1.03(1.01, 1.05) |
| Pooled | 1.00 (reference) | 1.03(0.98, 1.08) | 1.08(1.02, 1.14) | 1.12(1.06, 1.18) | 1.24(1.16, 1.32) | <0.0001 | 1.04(1.03, 1.04) |
| **Percent of calories from UPF/day** |  |  |  |  |  |  |  |
| NHS | 1.00 (reference) | 1.06(0.99, 1.14) | 1.10(1.03, 1.19) | 1.16(1.08, 1.24) | 1.14(1.06, 1.23) | <0.0001 | 1.05(1.03, 1.08) |
| NHSII | 1.00 (reference) | 0.98(0.91, 1.07) | 0.98(0.90, 1.06) | 1.03(0.95, 1.12) | 1.06(0.98, 1.15) | 0.03 | 1.04(1.01, 1.07) |
| HPFS | 1.00 (reference) | 1.16(1.04, 1.29) | 1.14(1.02, 1.27) | 1.23(1.10, 1.37) | 1.19(1.06, 1.33) | 0.003 | 1.07(1.03, 1.11) |
| Pooled | 1.00 (reference) | 1.05(1.00, 1.10) | 1.06(1.01, 1.12) | 1.13(1.07, 1.18) | 1.13(1.08, 1.18) | <0.0001 | 1.05(1.04, 1.07) |
| **Percent of grams of UPF/day** |  |  |  |  |  |  |  |
| NHS | 1.00 (reference) | 1.06(0.98, 1.15) | 1.17(1.08, 1.26) | 1.27(1.18, 1.37) | 1.36(1.26, 1.46) | <0.0001 | 1.13(1.11, 1.15) |
| NHSII | 1.00 (reference) | 1.23(1.12, 1.34) | 1.33(1.22, 1.45) | 1.37(1.26, 1.50) | 1.63(1.50, 1.76) | <0.0001 | 1.11(1.09, 1.13) |
| HPFS | 1.00 (reference) | 1.08(0.96, 1.21) | 1.16(1.04, 1.30) | 1.22(1.09, 1.37) | 1.41(1.26, 1.58) | <0.0001 | 1.09(1.06, 1.13) |
| Pooled | 1.00 (reference) | 1.12(1.06, 1.18) | 1.22(1.16, 1.28) | 1.30(1.23, 1.36) | 1.46(1.39, 1.54) | <0.0001 | 1.12(1.10, 1.13) |
| **Energy-adjusted servings of UPF/day** |  |  |  |  |  |  |  |
| NHS | 1.00(reference) | 1.10(1.02, 1.18) | 1.13(1.05, 1.21) | 1.16(1.08, 1.25) | 1.16(1.08, 1.24) | 0.0001 | 1.02(1.01, 1.03) |
| NHSII | 1.00(reference) | 1.11(1.03, 1.21) | 1.11(1.02, 1.21) | 1.17(1.08, 1.27) | 1.20(1.11, 1.30) | <0.0001 | 1.03(1.02, 1.04) |
| HPFS | 1.00(reference) | 1.06(0.94, 1.18) | 1.10(0.99, 1.23) | 1.11(0.99, 1.24) | 1.19(1.07, 1.32) | 0.001 | 1.02(1.01, 1.04) |
| Pooled | 1.00(reference) | 1.10(1.04, 1.15) | 1.12(1.06, 1.17) | 1.16(1.10, 1.22) | 1.18(1.12, 1.24) | <0.0001 | 1.02(1.02, 1.03) |

Results were obtained from multivariable model 2: stratified by calendar time (in 2-year intervals) and cohort (sex; for pooled analyses only), and adjusted for age, race/ethnicity (white/other), family history of diabetes (yes/no), history of hypercholesterolemia at baseline (yes/no), history of hypertension at baseline (yes/no), baseline BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0), smoking status (never, past, current), physical activity (MET-hours/week: <3.0, 3.0-8.9, 9.0-17.9, 18.0-26.9, ≥27.0), oral contraceptive use (never, former, current, in NHSII only), postmenopausal hormone use (premenopausal, never, former, current, in NHS and NHSII only), physical examination in the past 2 years (yes/no), neighborhood income (quintiles), total alcohol consumption (g/d, quintiles), and total energy intake (kcal/d, quintiles). All covariables (except race/ethnicity, family history of diabetes, baseline hypercholesterolemia, hypertension, and BMI) were updated every 2 years. *P*-values for trend were calculated using continuous ultra-processed food variable derived from the median ultra-processed food intake in each category of consumption.

†Hazard ratios (95% CI) for continuous level of intake refer to hazard ratios for increments of 1 ultra-processed food serving/day, 100 kcal from ultra-processed food /day, 10% of calories from ultra-processed food /day and 10% of grams of ultra-processed food /day, respectively.

‡The 9 food items with insufficient information to classify as ultra-processed food or non- ultra-processed food were: popcorn, soy milk, pancakes or waffles, pie home-baked or ready-made, beef, pork, lamb, sandwich, tomato sauce.

# Supplementary Table 6: Hazard ratio (95% confidence intervals) for incident type 2 diabetes or symptomatic diabetes according to total ultra-processed food consumption using alternative covariable modeling.

|  |  |  |  |
| --- | --- | --- | --- |
| **Models** | **Ultra-processed food consumption (servings/day)** | ***P* value for trend** | **Hazard ratio (95% CI) for 1 serving/day increment** |
| **Quintile 1 (low)** | **Quintile 2** | **Quintile 3** | **Quintile 4** | **Quintile 5 (high)** |
| Adjusted for updated BMI†  | 1.00 (reference) | 1.10 (1.04, 1.15) | 1.18 (1.12, 1.25) | 1.20 (1.14, 1.26) | 1.28 (1.21, 1.35) | <0.0001 | 1.03(1.02, 1.03) |
| Additionally adjusted for T2D screening frequency‡ | 1.00 (reference) | 1.08(1.03, 1.14) | 1.18(1.11, 1.24) | 1.19(1.13, 1.26) | 1.27(1.19, 1.34) | <0.0001 | 1.03(1.02, 1.03) |
| Symptomatic diabetes at diagnosis⸸ | 1.00 (reference) | 1.12(1.04, 1.21) | 1.17(1.08, 1.26) | 1.20(1.10, 1.30) | 1.29(1.19, 1.41) | <0.0001 | 1.03(1.02, 1.04) |
| Adjusted for intakes of non-UP foods, instead of total energy intake††  | 1.00 (reference) | 1.09(1.03, 1.14) | 1.18(1.12, 1.24) | 1.20(1.14, 1.26) | 1.30(1.24, 1.36) | <0.0001 | 1.04(1.03, 1.04) |

Results were obtained from the pooled multivariable model 2 stratified by calendar time (in 2-year intervals) and cohort (sex), and adjusted for age, race/ethnicity (white/other), family history of diabetes (yes/no), history of hypercholesterolemia at baseline (yes/no), history of hypertension at baseline (yes/no), smoking status (never, past, current), physical activity (MET-hours/week: <3.0, 3.0-8.9, 9.0-17.9, 18.0-26.9, ≥27.0), oral contraceptive use (never, former, current, in NHSII only), postmenopausal hormone use (premenopausal, never, former, current, in NHS and NHSII only), physical examination in the past 2 years (yes/no), neighborhood income (quintiles), total energy (kcal/d, quintiles), and total alcohol consumption (g/d, quintiles). All covariables (except race/ethnicity, family history of diabetes, baseline hypercholesterolemia, hypertension) were updated every 2 years. *P*-values for trend were calculated using continuous UPF variable derived from the median UPF intake in each category of consumption.

† This model additionally included updated BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0) every 2 years.

‡ This model additionally included baseline BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0) and for T2D screening frequency.

⸸ This model additionally included baseline BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0) and cases were restricted to symptomatic diabetes at diagnosis.

†† This model additionally included baseline BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0), and intakes of non-ultra-processed fruits (servings/day, quintiles), vegetables (servings/day, quintiles), nuts (servings/day, quintiles), legumes (servings/day, quintiles), tea (servings/day, quintiles), coffee (servings/day, quintiles), whole grains (servings/day, quintiles), red meat (servings/day, quintiles), fish (servings/day, quintiles), and poultry (servings/day, quintiles), and excluded total energy intake.

# Supplementary Table 7: Mediation effect of key nutrients in the relationship between ultra-processed food and T2D risk.

|  |  |  |
| --- | --- | --- |
|  | **Hazard ratio (95% CI) for 1 serving/day increment**† | **Mediation effect**‡ |
| **Dietary fibers** | **Refined starch** | **Added sugar** | **Sodium** | **Minerals** | **Partially hydrogenated oils**  | **All 6 nutrients (combined effect)**⸸ |
| **Total UPF** | 1.03(1.02, 1.03) | 5.9 %(3.9%, 8.8%) | --- | --- | 9.9 %(5.9%, 16.2%) | --- | --- | 11.9%(4.6%, 27.7%) |
| **Ultra-processed cereals and breads** | 0.98(0.96, 0.99) | --- | --- | --- | --- | 11.7%(4.2%, 28.9%) | --- | --- |
| *Ultra-processed cereals* | 0.78(0.75, 0.82) | 2.6 %(1.4%, 4.8%) | 4.5%(1.3%, 13.8%) | 5.9%(3.6%, 9.6%) | --- | 8.1%(4.7%, 13.5%) | --- | 18.4%(12.0%, 27.1%) |
| *Ultra-processed dark breads and whole-grain breads* | 0.96(0.94, 0.98) | 5.6%(2.0%, 15.0%) | 15.9%(3.6%, 48.9%) | --- | --- | 7.2%(2.7%, 17.8%) | --- | 27.4%(9.0%, 59.0%) |
| *Other ultra-processed refined breads* | 1.05(1.02, 1.07) | 18.5%(7.9%, 37.3%) | --- | 3.3%(1.4%, 7.8%) | --- | 5.7%(2.5%, 12.7%) | --- | --- |
| **Sauces, spreads, and condiments** | 1.05(1.03, 1.06) | --- | --- | 1.9%(1.0%, 3.8%) | **---** | 4.8%(3.0%, 7.6%) | --- | 13.1%(5.8%, 27.0%) |
| **Packaged sweet snacks and desserts** | 0.89(0.87, 0.91) | --- | --- | 9.4%(6.2%, 13.9%) | --- | --- | 4.4%(1.3%, 13.4%) | --- |
| *Confectioneries* | 0.99(0.95, 1.03) | --- | --- | 29.4%(8.0%, 66.6%) | --- | --- | --- | --- |
| *Packaged sweet snacks* | 0.87(0.84, 0.89) | --- | --- | 6.6%(4.3%, 10.1%) | --- | 1.5%(0.9%, 2.5%) | --- | --- |
| *Fruit-based products* | 0.82(0.77, 0.86) | 1.8% (1.0%, 3.1%) | --- | 8.5% (5.3%, 13.4%) | --- | --- | --- | 5.4%(2.4%, 11.4%) |
| **Packaged savory snacks** | 0.91(0.87, 0.94) | 4.1 %(2.2%, 7.3%) | --- | --- | --- | --- | 4.0%(1.1%, 13.2%) | --- |
| **Artificially and sugar-sweetened beverages** | 1.10(1.09, 1.12) | 1.1%(0.7%, 1.8%) | --- | --- | --- | --- | --- | --- |
| *Artificially sweetened beverages* | 1.09(1.08, 1.11) | --- | --- | 5.2% (3.7%, 7.1%) | --- | --- | --- | 5.3%(3.1%, 9.0%) |
| *Sugar-sweetened beverages* | 1.15(1.12, 1.17) | 3.5%(2.1%, 5.6%) | --- | --- | --- | 4.6%(2.9%, 7.1%) | --- | --- |
| **Animal-based products** | 1.44(1.38, 1.51) | 4.6%(3.0%, 7.0%) | --- | 7.1%(4.7%, 10.5%) | --- | 7.9%(5.2%, 11.6%) | --- | 21.0%(15.5%, 27.8%) |
| **Ready-to-eat/heat mixed dishes** | 1.42(1.31, 1.54) | 3.1%(1.8%, 5.1%) | --- | --- | --- | 3.5%(2.1%,5.8%) | --- | 13.9%(5.8%, 29.9%) |
| **Yogurt and dairy-based desserts** | 0.91(0.88,0.95) | --- | --- | 10.0% (5.1%, 18.6%) | --- | 9.4%(4.9%, 17.4%) | --- | 8.4%(3.4%, 19.3%) |
| **Other UPF** | 1.05(1.02, 1.08) | 2.4%(1.2%, 4.8%) | --- | --- | --- | --- | 1.5%(0.4%, 5.2%) | --- |

† Hazard ratio (95% CI) were obtained from the pooled multivariable model 2 stratified by calendar time (in 2-year intervals) and cohort (sex), and adjusted for age, race/ethnicity (white/other), family history of diabetes (yes/no), history of hypercholesterolemia at baseline (yes/no), history of hypertension at baseline (yes/no), baseline BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0), smoking status (never, past, current), physical activity (MET-hours/week: <3.0, 3.0-8.9, 9.0-17.9, 18.0-26.9, ≥27.0), oral contraceptive use (never, former, current, in NHSII only), postmenopausal hormone use (premenopausal, never, former, current, in NHS and NHSII only), physical examination in the past 2 years (yes/no), neighborhood income (quintiles), total energy (kcal/d, quintiles) and total alcohol consumption (g/d, quintiles). All covariables (except race/ethnicity, family history of diabetes, baseline hypercholesterolemia, hypertension, and BMI) were updated every 2 years. The UPF groups (or subgroups) were included simultaneously in the models as distinct covariables.

‡ The SAS macro %mediate and the formula [1-(βmediator model/βbase model)\*100] were used to evaluate the mediation effect of the 6 key nutrients on the association between total UPF and T2D risk. The model included calendar time (in 2-year intervals), cohort (sex), age, race/ethnicity (white/other), family history of diabetes (yes/no), history of hypercholesterolemia at baseline (yes/no), history of hypertension at baseline (yes/no), smoking status (never, past, current), baseline BMI (kg/m2: <21.0, 21.0-22.9, 23.0-24.9, 25.0-26.9, 27.0-29.9, 30.0-34.9, ≥35.0), physical activity (MET-hours/week: <3.0, 3.0-8.9, 9.0-17.9, 18.0-26.9, ≥27.0), oral contraceptive use (never, former, current, in NHSII only), postmenopausal hormone use (premenopausal, never, former, current, in NHS and NHSII only), physical examination in the past 2 years (yes/no), neighborhood income (quintiles), total energy (kcal/d, quintiles) and total alcohol consumption (g/d, quintiles). For each mediation analysis, each of the 6 nutrients were included individually as potential mediator, while the 5 others were included as covariables in the model. All covariables (except race/ethnicity, family history of diabetes, baseline hypercholesterolemia, hypertension) were updated every 2 years. In analyses on UPF subgroups, UPF subgroups were included simultaneously in the models as distinct covariables.

⸸ To estimate the overall combined effect of all 6 tested nutrients, the 6 nutrients were simultaneously included in the mediation model as potential mediators per the guidelines of the SAS macro %mediate (12).

“---" indicates no significant mediation effect.

# Supplementary Table 8: Characteristics of studies included in the meta-analysis.1

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Author, year** | **Country (cohort)** | **Participants, n** | **Age (years)** | **Follow-up duration (years)** | **Assessment of diet** | **Number of cases of T2D** | **Ascertainment of cases** | **Statistical Model** | **Categories of Exposure** | **Relative Risks (95% CI)** | **Covariates in Multivariable Model** |
| Duan et al. 2022 | The Netherlands (Lifelines) | 70,421 (men=29,178; women=41,243) | 35-70  | 3.4 | Self-administered, validated, 110-item semi-quantitative FFQ. | 1,128 | Self-report and validation with blood analysis (fasting glucose≥7.0 mmol/L, or HbA1c≥48 mmol/mol [6.5%]) | Logistic regression models | Weight (% of grams per day) of UPF1st quartile2nd quartile3rd quartile4th quartileIncrement of 10% consumption of UPF in the total diet | 1.00 (reference)1.04 (0.87, 1.26)1.20 (0.99, 1.45)1.56 (1.27, 1.92)1.17 (1.09, 1.26) | Age, sex, Lifelines diet score, total energy intake, alcohol intake, smoking status, educational level, nonoccupational moderate-to-vigorous physical activity level, TV watching time, BMI. |
| Llavero-Valero, et al. 2021 | Spain (SUN project) | 20,060 (men=7,723; women=12,337) | Mean (SD): 37.4 (12.2)  | 12 | Self-administered, validated, 136-item semi-quantitative FFQ, at baseline and at 10 years of follow-up | 175 | Self-report and validation through an endocrinologist according to American Diabetes Association criteria | Cox proportional hazard models | Grams per day of UPFTertile 1Tertile 2Tertile 3 | 1.00 (reference)0.99 (0.69, 1.43)1.53 (1.06, 2.22) | Sex, age, tertiles of BMI, educational status, family history of diabetes, smoking status, snacking between meals, 8-item active and sedentary lifestyle score, following a special diet at baseline. |
| Levy et al. 2021 | UK (UK Biobank) | 21,730 (men=10,235; women=11,495) | 40-69 | 5.4 | Web-based, self-administered questionnaire that collects quantities of over 200 common food and beverage items consumed in the previous 24 hours | 305 | Self-report and nurse-interview data. Algorithm to identify people with or without incident T2D during follow-up in the UK Biobank. | Cox proportional hazards regression | Weight (% of grams per day) of UPF1st quartile2nd quartile3rd quartile4th quartilePer 10% absolute incrementin UPF | 1.00 (reference)0.98 (0.68, 1.39)1.10 (0.76, 1.55)1.44 (1.04, 2.02)1.12 (1.04, 1.20) | Age, sex, ethnicity, family history of T2D, current smoking status, BMI, quintiles of the Index of Multiple Deprivation (IMD), physical activity level. |
| Srour et al. 2020 | France (NutriNet Santé) | 104,707 (men=21,800; women=82,907) | ≥18 | 6.0 | 3 nonconsecutivevalidated web-based 24-hour dietary records atbaseline and every 6 months,randomly assigned over a 2-week period (2 weekdaysand 1 weekend day) | 821 | Self-report and validation through the medico-administrative databases of the SNIIRAM | Cox proportional hazards regression | Absolute increment of 10% of UPF (in weight) in the Diet | 1.13 (1.01, 1.27) | Age, sex, educational level, BMI, physical activity level, smoking status, alcohol intake, number of 24-hour dietary records, energy intake without alcohol, family history of diabetes, overall nutritional quality of the diet. |
| Chen et al.(current) | United States (NHS) | 71,871 women | 30-55 | 32 | Self-administered 130-items semi-quantitative FFQ, every 4 years | 8,591 | Self-reports of incident T2D by participants identified through follow-up questionnaires and confirmed by a validated supplementary questionnaire | Cox proportional hazards regression | Servings/day of UPF1st quintile2nd quintile3rd quintile4th quintile5th quintile;Weight (% of grams per day) of UPF1st quintile2nd quintile3rd quintile4th quintile5th quintile;Per 10%increment (in % of grams per day) of UPF | 1.00 (Reference)1.07 (0.99, 1.16)1.13 (1.05, 1.22)1.15 (1.06, 1.25)1.19 (1.09, 1.30)1.00 (Reference)1.06 (0.98, 1.15)1.17 (1.08, 1.26)1.27 (1.18, 1.37)1.36 (1.26, 1.46)1.13 (1.11, 1.15) | Age, race/ethnicity, family history of diabetes, history of hypercholesterolemia at baseline, history of hypertension at baseline, baseline BMI, smoking status, physical activity, postmenopausal hormone use, physical examination, neighborhood income, total alcohol consumption, and total energy  |
|  | United States (NHSII) | 87,918 women | 25-44 | 26 | Self-administered 130-items semi-quantitative FFQ, every 4 years | 7,177 | Self-reports of incident T2D by participants identified through follow-up questionnaires and confirmed by a validated supplementary questionnaire | Cox proportional hazards regression | Servings/day of UPF1st quintile2nd quintile3rd quintile4th quintile5th quintile;Weight (% of grams per day) of UPF1st quintile2nd quintile3rd quintile4th quintile5th quintile;Per 10%increment (in % of grams per day) | 1.00 (Reference)1.18 (1.08, 1.29)1.35 (1.23, 1.47)1.32 (1.21, 1.45)1.46 (1.33, 1.60)1.00 (Reference)1.23 (1.12, 1.34)1.33 (1.22, 1.45)1.37 (1.26, 1.50)1.63 (1.50, 1.76)1.11 (1.09, 1.13) | Age, race/ethnicity, family history of diabetes, history of hypercholesterolemia at baseline, history of hypertension at baseline, baseline BMI, smoking status, physical activity, postmenopausal hormone use, oral contraceptive use, physical examination, neighborhood income, total alcohol consumption, and total energy |
|  | United States (HPFS) | 38,847 men | 40-75 | 30 | Self-administered 130-items semi-quantitative FFQ, every 4 years | 3,735 | Self-reports of incident T2D by participants identified through follow-up questionnaires and confirmed by a validated supplementary questionnaire | Cox proportional hazards regression | Servings/day of UPF1st quintile2nd quintile3rd quintile4th quintile5th quintile;Weight (% of grams per day) of UPF1st quintile2nd quintile3rd quintile4th quintile5th quintile;Per 10%increment (in % of grams per day) of UPF | 1.00 (Reference)1.01 (0.90, 1.13)1.05 (0.94, 1.18)1.14 (1.01, 1.29)1.22 (1.07, 1.39)1.00 (Reference)1.08 (0.96, 1.21)1.16 (1.04, 1.30)1.22 (1.09, 1.37)1.41 (1.26, 1.58)1.09 (1.06, 1.13) | Age, race/ethnicity, family history of diabetes, history of hypercholesterolemia at baseline, history of hypertension at baseline, baseline BMI, smoking status, physical activity, physical examination, neighborhood income, total alcohol consumption, and total energy |

1 T2D: type 2 diabetes; UPF: ultra-processed foods; FFQ: food frequency questionnaire; SNIIRAM: *Système national d’information inter-régimes de l’Assurance maladie* [National Health Insurance Inter-Scheme Information System]; BMI: body mass index; UK: United Kingdom; IMD: Index of Multiple Deprivation; ADA: American Diabetes Association;

# Supplementary Table 9: List of confounders among studies included in the meta-analysis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Duan et al. (2022)** | **Llavero-Valero et al. (2021)** | **Levy et al. (2021)** | **Srour et al. (2020)** | **Chen et al. (current)** |
| **Primary confounders** |
| Age | X | X | X | X | X |
| Sex | X | X | X | X | X |
| Body mass index | X | X | X | X | X |
| Smoking status | X | X | X | X | X |
| Physical activity/energy expenditure | X | X | X | X | X |
| Alcohol intake | X |  |  | X | X |
| Total energy intake | X |  |  | X | X |
| **Secondary confounders** |
| Diet quality or major non-ultra processed foods | X |  |  | X |  |
| **Other confounders** |
| Ethnicity |  |  | X |  | X |
| Family history of diabetes |  | X | X | X | X |
| Education | X | X |  | X |  |
| Deprivation level |  |  | X |  |  |
| Television hours | X |  |  |  |  |
| Number of time diet was assessed |  |  |  | X |  |
| Hypertension or hypercholesterolaemia at baseline |  |  |  | X | X |
| Snacking between meals |  | X |  |  |  |
| Following a special diet at baseline |  | X |  |  |  |
| Postmenopausal hormone use |  |  |  |  | X |
| Physical examination |  |  |  |  | X |
| Neighborhood income |  |  |  |  | X |

# Supplementary Table 10: Assessment of risk of bias with the Newcastle-Ottawa Scale.1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author, year (Study ID)** | **Selection** | **Comparability** | **Outcome** | **Total** |
| **Representativeness of the exposed cohort** | **Selection of the non-exposed cohort** | **Ascertainment of exposure** | **Outcome of interest not present at start of the study** | **Control for primary confounders** | **Control for secondary confounders** | **Assessment of outcome** | **Duration of follow-up** | **Adequacy of follow-up** |
| Duan et al. (2022) | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 7 |
| Llavero-Valero et al. (2021) | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 6 |
| Levy et al. (2021) | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 6 |
| Srour et al. (2020) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| Chen et al. (current) | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 7 |

1 Representativeness of the exposed cohort: 1 point awarded if community-based population;

Selection of the non-exposed cohort: 1 point awarded if drawn form the same community as the exposed cohort;

Ascertainment of exposure: 1 point awarded if diet assessed at baseline and at least one time during follow-up;

Outcome of interest not present at start of the study: 1 point awarded if individuals with prevalent diabetes at baseline were excluded;

Control for primary confounders: 1 point awarded if adjustment for age, sex, body mass index (BMI), smoking status, physical activity, alcohol intake, and total energy intake;

Control for secondary confounders: 1 point awarded if adjustment for major food items (e.g., fruits, and vegetables, whole grains) or diet quality;

Assessment of outcome: 1 point awarded if self-reported cases were confirmed using additional measures (e.g. fasting blood glucose, supplemental questionnaires, physician’s diagnosis, etc.);

Duration of follow-up; 1 point awarded if follow-up ≥ 10 years;

Adequacy of follow-up: 1 point awarded if loss to follow-up <20%.

# Supplementary Table 11: Pre-specified subgroup meta-analyses of the association between total ultra-processed food consumption and risk of type 2 diabetes, using random-effects models.1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Stratification** | **Categories** | **Risk estimates, n** | **Pooled Relative Risk (95% CI)** | **I2 (%)** | ***P* interaction2** |
| Mean age of participants | <45 years | 3 | 1.10 (1.07, 1.13) | 0.0% | 0.77 |
|  | ≥45 years | 4 | 1.12 (1.11, 1.14) | 8.6% |  |
| Region | Western European countries | 2 | 1.14 (1.09, 1.20) | 0.0% | 0.41 |
|  | Mediterranean Countries | 2 | 1.14 (1.05, 1.23) | 0.0% |  |
|  | USA | 3 | 1.11 (1.09, 1.13) | 53.9% |  |
| Follow-up duration | <10 years | 3 | 1.14 (1.09, 1.20) | 0.0% | 0.39 |
|  | ≥10 years | 4 | 1.11 (1.10, 1.13) | 33.0% |  |
| Number of participants | <50,000 | 3 | 1.10 (1.07, 1.13) | 0.0% | 0.28 |
|  | ≥50,000 | 4 | 1.12 (1.11, 1.14) | 8.6% |  |
| Number of cases | <1,000 | 3 | 1.13 (1.07, 1.19) | 0.0% | 0.76 |
|  | >1,000 | 4 | 1.12 (1.10, 1.14) | 49.4% |  |
| Main analyses adjusted for total energy intake | No | 2 | 1.13 (1.06, 1.20) | 0.0% | 0.81 |
|  | Yes | 5 | 1.12 (1.10, 1.14) | 32.9% |  |
| Main analyses adjusted for history of hypertension or hypercholesterolemia | No | 3 | 1.14 (1.09, 1.20) | 0.0% | 0.36 |
|  | Yes | 4 | 1.11 (1.10, 1.13) | 31.5% |  |
| Risk of bias | <7 | 3 | 1.14 (1.09, 1.20) | 0.0% | 0.36 |
|  | ≥7 | 4 | 1.11 (1.10, 1.13) | 31.5% |  |
| Dietary assessment method | 24 hour diet recalls | 2 | 1.12 (1.06, 1.19) | 0.0% | 0.88 |
|  | FFQ | 5 | 1.12 (1.10, 1.14) | 33.9% |  |

1 Pooled relative risks are from a random-effects meta-analysis of each 10% increment of ultra-processed food intake (% of gram per day). *I*-squared refers to the proportion of heterogeneity between studies;

2 Calculated using meta-regression.

# Supplementary Table 12: Assessment of the quality of evidence on the relationship between total ultra-processed food consumption and risk of type 2 diabetes using the NutriGrade scoring system.1

|  |  |  |
| --- | --- | --- |
| **CRITERIA** | **MAX score** | **Score for the current study** |
| Risk of bias/study quality/study limitations | 2 | 2 |
| Precision | 1 | 1 |
| Heterogeneity | 1 | 0.5 |
| Directness | 1 | 1 |
| Publication bias | 1 | 0.5 |
| Funding bias | 1 | 1 |
| Effect size | 2 | 1 |
| Dose-response | 1 | 1 |
| **Overall score** | **10** | **8** |

1Evidence grading according to overall score: very-low meta-evidence, 0-3.99; low meta-evidence, 4-5.99; moderate meta-evidence, 6-7.99; high meta-evidence: 8+. The overall score was obtained by summing scores from each component. Maximum possible score for risk of bias/study quality/study limitations, precision, heterogeneity, directness, publication bias, funding bias, effect size, and dose response was 2, 1, 1, 1, 1, 1, 2, and 1, respectively.



# Supplementary Figure 1: Flowchart of participants.

HPFS: Health Professionals’ Follow-Up Study; NHS: Nurses’ Health Study.



# Supplementary Figure 2: Meta-analysis search strategy and study selection.



# Supplementary Figure 3: Funnel plot for assessment of publication bias for the association between total ultra-processed food consumption and risk of type 2 diabetes.

*P*-value for Egger’s test = 0.69; *P*-value for Begg’s test = 0.88. The assessment of publication bias was for the association of each 10% increment of ultra-processed food intake (% of gram per day).



# Supplementary Figure 4: Forest plot of influence analysis for the association between total ultra-processed food consumption and risk of type 2 diabetes.

Each dot represents the pooled relative risk (95% confidence interval) following the exclusion of the study listed on the left using random-effects meta-analysis. The influence analysis was for the association of high vs low ultra-processed food intakes with type 2 diabetes.

# Supplementary references

1. Monteiro CA, Cannon G, Levy RB, Moubarac JC, Louzada ML, Rauber F, Khandpur N, Cediel G, Neri D, Martinez-Steele E, Baraldi LG, Jaime PC: Ultra-processed foods: what they are and how to identify them. Public health nutrition 2019;22:936-941

2. Khandpur N, Rossato S, Drouin-Chartier JP, Du M, Steele EM, Sampson L, Monteiro C, Zhang FF, Willett W, Fung TT, Sun Q: Categorising ultra-processed foods in large-scale cohort studies: evidence from the Nurses' Health Studies, the Health Professionals Follow-up Study, and the Growing Up Today Study. J Nutr Sci 2021;10:e77

3. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D: The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical research ed) 2021;372:n71

4. Wells GASB OCD, Peterson J, Welch V, Losos M, Tugwell P.: The Newcastle–Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. 2009 <http://wwwohrica/PROGRAMS/CLINICAL_EPIDEMIOLOGY/OXFORDASP> 2009;

5. Llavero-Valero M, Escalada-San Martín J, Martínez-González MA, Basterra-Gortari FJ, de la Fuente-Arrillaga C, Bes-Rastrollo M: Ultra-processed foods and type-2 diabetes risk in the SUN project: A prospective cohort study. Clinical nutrition (Edinburgh, Scotland) 2021;40:2817-2824

6. Duan M-J, Vinke PC, Navis G, Corpeleijn E, Dekker LH: Ultra-processed food and incident type 2 diabetes: studying the underlying consumption patterns to unravel the health effects of this heterogeneous food category in the prospective Lifelines cohort. BMC Medicine 2022;20:7

7. Srour B, Fezeu LK, Kesse-Guyot E, Allès B, Debras C, Druesne-Pecollo N, Chazelas E, Deschasaux M, Hercberg S, Galan P, Monteiro CA, Julia C, Touvier M: Ultraprocessed Food Consumption and Risk of Type 2 Diabetes Among Participants of the NutriNet-Santé Prospective Cohort. JAMA Internal Medicine 2020;180:283-291

8. Greenland S, Longnecker MP: Methods for Trend Estimation from Summarized Dose-Response Data, with Applications to Meta-Analysis. American journal of epidemiology 1992;135:1301-1309

9. Orsini N, Bellocco R, Greenland S: Generalized Least Squares for Trend Estimation of Summarized Dose–response Data. The Stata Journal 2006;6:40-57

10. Higgins JP GS: Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons 2011

11. Schwingshackl L, Knuppel S, Schwedhelm C, Hoffmann G, Missbach B, Stelmach-Mardas M, Dietrich S, Eichelmann F, Kontopantelis E, Iqbal K, Aleksandrova K, Lorkowski S, Leitzmann MF, Kroke A, Boeing H: Perspective: NutriGrade: A Scoring System to Assess and Judge the Meta-Evidence of Randomized Controlled Trials and Cohort Studies in Nutrition Research. Adv Nutr 2016;7:994-1004

12. Lin DY, Fleming TR, De Gruttola V: Estimating the proportion of treatment effect explained by a surrogate marker. Stat Med 1997;16:1515-1527