Supplementary Online Content

Srour B, Fezeu LK, Kesse-Guyot E, et al. Ultraprocessed Food Consumption and Risk of Type 2 Diabetes Among Participants of the NutriNet-Santé Prospective Cohort [published online December 16, 2019]. *JAMA Intern Med*. doi:10.1001/jamainternmed.2019.5942

eMethods.

eFigure 1. Distribution of the main exposure (proportion of ultra-processed food in the diet) in the study sample (N=104,707), NutriNet-Santé, France

eFigure 2. Cox model proportional risk assumption testing (Schoenfeld residuals)

eFigure 3. Spline plot for the linearity assumption of the association between the proportion of ultra-processed food in the diet and the risk of Type-2 Diabetes using Restricted cubic spline (RCS) SAS Macro® developed by Desquilbet and Mariotti

eResults.

eReferences.

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods:

Dietary data collection

Participants used the dedicated web interface to declare all foods and beverages consumed during a 24h-period for each of the three main meals (breakfast, lunch, dinner) and any other eating occasion, with an accurate estimation of portion sizes[1]. Dietary underreporting was identified on the basis of the method proposed by Black, using the basal metabolic rate and Goldberg cut-off, and under-energy reporters were excluded[2]. Mean daily alcohol, micro- and macro-nutrient and energy intake were calculated using the NutriNet-Santé food composition database, which contains more than 3,500 different items (plus the possibility to enter new items in an open field)[3]. Overall, 2% of the food and beverage items were added to an open field (278,420 out of the 14,005,464 reported items). All these foods were re-classified among the existing items or lead to the implementation of new food items in the NutriNet-Santé classification. Since the launching of the cohort, 841 new food items among the 3,500 were created following their occurrence in the open fields, and the database is permanently being enriched.

Biological data collection

Participants in the NutriNet-Santé study were invited, on a voluntary basis, for a visit in one of the 83 hospital centers specifically set up in all regions of mainland France for biological sampling and clinical examination. The whole protocol has been published previously[4]. Overall, 19,772 participants attended this visit between 2011 and 2014. Height, weight and other anthropometric parameters were measured. Blood samples were collected after at least a 6h-fasting period and immediately centrifuged, fractionated into sufficient aliquots and stored at -80°C before shipment to the central laboratory for analysis (IRSA, Tours, France). Fasting blood glucose was determined by hexokinase on C8000 automat, Abbott, Suresnes, France.

Definitions, precisions and examples of ultra-processed foods according to the NOVA classification

All food and beverage items of the NutriNet-Santé composition table were categorized by a team of three trained dieticians into one of the four food groups in NOVA, a food classification system based on the extent and purpose of industrial food processing [5–7]. The whole classification was then reviewed by a committee composed of the three dietitians and five researchers, specialists in nutritional epidemiology. In case of uncertainty for a given food/beverage item, a consensus was reached among researchers based on the percentage of home-made and artisanal foods versus industrial brands reported by the participants.

The "ultra-processed foods" group of the NOVA classification is the primarily focus of this study. Products in this group undergo industrial processes notably include hydrogenation, hydrolysis, extruding, moulding, reshaping, and pre-processing by frying. Flavouring agents, colours, emulsifiers, humectants, non-sugar sweeteners and other cosmetic additives are often added to these products to imitate sensorial properties of unprocessed or minimally processed foods and their culinary preparations. The UPF group is defined by opposition to the other NOVA groups: "unprocessed or minimally processed foods" (fresh, dried, grounded, chilled, frozen, pasteurized or fermented staple foods such as fruits, vegetables, pulses, rice, pasta, eggs, meat, fish or milk), "processed culinary ingredients" (salt, vegetable oils, butter, sugar and other substances extracted from foods and used in kitchens to transform unprocessed or minimally processed foods into culinary preparations) and "processed foods" (canned vegetables with added salt, sugar-coated dry fruits, meat products only preserved by salting, cheeses and freshly made unpackaged breads, and other products manufactured with the addition of salt, sugar or other substances of the "processed culinary ingredients" group). As previously described[8], home-made and artisanal food preparations were identified and decomposed using standardized recipes, and the NOVA classification was applied to their ingredients. Examples of such products as well as examples of distinctions between ultra-processed products and products from other NOVA categories are provided below:

Examples of typical ultra-processed food according to the NOVA classification:

Poultry and fish nuggets and sticks and other reconstituted meat products transformed with addition of preservatives other than salt (e.g nitrites); instant noodles and dehydrated soups; carbonated drinks; sweet or savoury packaged snacks; chocolate, candies (confectionery); margarines and spreads; industrial pastries and instant desserts; breakfast 'cereals', 'energy' bars; 'energy' drinks; flavoured milk drinks; sweet desserts made from fruit with added sugars, artificial flavours and texturizing agents; cooked seasoned vegetables with ready-made sauces; vegetable patties (meat substitutes) containing food additives; meat and chicken extracts and 'instant' sauces; 'health' and 'slimming' products such as powdered or 'fortified' meal and dish substitutes; ready to heat products including pre-prepared pies, pasta and pizza dishes.

For instance, salted-only red or white meats are considered as "processed foods" whereas smoked or cured meats with added nitrites and conservatives, such as sausages and ham are classified as "ultra-processed foods".

Similarly, canned salted vegetables are considered as "processed foods" whereas industrial cooked or fried seasoned vegetables, marinated in industrial sauces with added flavourings are considered as "ultra-processed foods".

Flavoured breakfast cereals with added emulsifiers, texturizing agents and/or colorants were included in the ultra-processed food group. Homemade granola, oatmeal, rye and barley flakes without additives were not considered as ultra-processed.

Regarding soups, canned liquid soups with added salts, herbs and spices are considered as "processed foods" while instant dry soup mixes are considered as "ultra-processed foods".

Example of list of ingredients for an industrial Chicken and Leek flavour soup considered as "ultra-processed" according to the NOVA classification: *"Dried Glucose Syrup, Potato Starch, Flavourings, Salt, Leek Powder (3.6%), Dried Leek (3.5%), Onion Powder, Dried Carrot, Palm Oil, Dried Chicken (0.7%), Garlic Powder, Dried Parsley, Colour [Curcumin (contains MILK)], Ground Black Pepper, MILK Protein, Stabilisers (Dipotassium Phosphate, Trisodium Citrate)"*.

Examples of food products considered as ultra-processed according to the NOVA classification

Ultra-processed food group	**Examples of foods**
Beverages	Sugary drinks (e.g. regular sodas, sugary fruit-based beverages, industrial chocolate powder beverages, energy drinks, flavoured waters); artificially sweetened beverages (e.g. diet sodas, artificially sweetened ice teas)
Dairy products	Flavoured or artificially sweetened yoghurts; products such as dairy desserts, cream cheese, milkshakes, dairy beverages, flavoured milk with one or more texturizer, emulsifier, colorant or other cosmetic additives
Fats and sauces	Sauces and dressings (salad dressing, mayonnaise, ketchup, béchamel, and other dressings) containing emulsifiers, texturizers, flavour enhancers or other additives
Fruits and vegetables	Instant powder soups; reconstituted vegetarian/soy steaks with additives; flavoured and artificially sweetened fruit compotes; vegan nuggets
Meat, fish, and eggs	Processed meat with added nitrites; chicken nuggets; fish fingers; industrial ‘cordon bleu’ chicken with wheat dextrose, emulsifiers, preservatives; surimi-crab sticks
Starchy foods and cereals	Flavoured breakfast cereals with added emulsifiers, texturizing agents and/or colorants; industrial pre-baked breads and buns with added dextrose, preservatives or emulsifiers.
Sugary products	Industrially packed cookies, cakes, chocolate/wafer bars, and candies manufactured with glucose syrup, modified starch, hydrogenated oils, colours, flavours.
Salty snacks	Chips, crisps and crackers made with other ingredients than potatoes, oil and salt such as maltodextrin, flavors, dyes, emulsifiers, flavour enhancers

Case ascertainment

Participants were asked to declare major health events though the yearly health questionnaire, through a specific health check-up questionnaire every three months, or at any time through a specific interface on the study website. They were also asked to declare all currently taken medications and treatments via the check-up and yearly questionnaires. A search engine with embedded exhaustive Vidal® drug database is used to facilitate medication data entry for the participants. Besides, our research team was the first in France to obtain the authorization by Decree in the Council of State (n°2013-175) to link data from our general population-based cohorts to medico-administrative databases of the National health insurance (SNIIRAM database). Thus, data from the NutriNet-Santé cohort are linked every year to medico-administrative databases of the SNIIRAM, providing detailed information about the reimbursement of medication and medical consultations.

Regarding T2D specifically: all 821 cases were primarily detected through the declaration by the participants of a T2D diagnosed by a physician and/or diabetes medication use, in follow-up questionnaires. The questions were: "Have you been diagnosed with T2D (if yes, indicate the date of diagnosis)" and "Are you treated for T2D?". ATC codes considered for T2D medication were A10AB01, A10AB03, A10AB04, A10AB05, A10AB06, A10AC01, A10AC03, A10AC04, A10AD01, A10AD03, A10AD04, A10AD05, A10AE01, A10AE02, A10AE03, A10AE04, A10AE05, A10AE30, A10BA02, A10BB01, A10BB03, A10BB04, A10BB06, A10BB07, A10BB09, A10BB12, A10BD02, A10BD03, A10BD05, A10BD07, A10BD08, A10BD10, A10BD15, A10BD16, A10BF01, A10BF02, A10BG02, A10BG03, A10BH01, A10BH02, A10BH03, A10BX02, A10BX04, A10BX07, A10BX09, A10BX10, A10BX11, A10BX12. Following a T2D diagnosis and/or medication declaration, two additional sources of information were considered for confirmation. First, the linkage with the SNIIRAM National health insurance database allowed confirming 85.7% of investigated cases

(ICD-10 codes E11). Of note, about 10-15% of the French population is covered by other social security regimen and would not be correctly captured by the SNIIRAM databases. Besides, the centralization of SNIIRAM data might take up to a year, leading to delays between reported T2D information and health insurance data. Second, among the participants who provided blood sample during the clinical/biological examination, 232 had elevated fasting blood glucose (i.e. >1.26 g/L[9]). Among them, 85.3% had consistently declared a T2D diagnosis and/or medication. Elevated blood glucose only i.e. without any declaration of T2D diagnosis or treatment) was not considered specific enough to classify the participant as a T2D case.

FSAm-NPS DI computation

The Food Standard Agency nutrient profiling system dietary index (FSAm-NPS DI) is based on the British FSA nutrient profiling system. It is the score underlying the official French, Belgian and Spanish front-of-package food labelling (the Nutri-Score). It has been extensively described and validated elsewhere[10–12]. Its computation is detailed below.

1) <u>FSAm-NPS score computation at food/beverage level</u>

Points are allocated according to the nutrient content for 100g of foods or beverages.

Points are allocated for ‘Negative’ nutrients (A points) and can be balanced according to ‘Positive’ nutrients (C points).

A points

Total A points = (points for energy) + (points for saturated fat) + (points for total sugar) + (points for sodium)

Points	**Energy (kJ)**	**Saturated Fat (g)**	**Total Sugars (g)**	**Sodium (mg)**
0	≤ 335	≤ 1	≤ 4.5	≤ 90
1	> 335	> 1	> 4.5	> 90
2	> 670	> 2	> 9	> 180

3	> 1005	> 3	> 13.5	> 270
4	> 1340	> 4	> 18	> 360
5	> 1675	> 5	> 22.5	> 450
6	> 2010	> 6	> 27	> 540
7	> 2345	> 7	> 31	> 630
8	> 2680	> 8	> 36	> 720
9	> 3015	> 9	> 40	> 810
10	> 3350	> 10	> 45	> 900

C points

Total C points = (points for fruits/vegetables/legumes/nuts) + (points for fibres) + (points for proteins)

Points	**Fruits/vegetables/legumes/nuts**	**Fibre (g) ***	**Protein (g)**
0	≤ 40	≤ 0.7	≤ 1.6
1	> 40	> 0.7	> 1.6
2	> 60	> 1.4	> 3.2
3	-	> 2.1	> 4.8
4	-	> 2.8	> 6.4
5	> 80	> 3.5	> 8.0

* FSAm-NPS score allocates different thresholds for fibres, depending on the measurement method used. We used NSP cut-offs to compute fibres score.

For 100g of a given food, the percentage of fruits/vegetables/legumes/nuts is obtained by summing up the amount (in grams) of all fruits, legumes and vegetables (including oleaginous fruits, dried fruits and olives) contained in this food.

Overall score computation

- If Total A points <11, then FSAm-NPS score =Total A points – Total C points
- If Total A points ≥11,
    - If points for fruits/vegetables/legumes/nuts =5, then FSAm-NPS score =Total A points – Total C points
    - Else if points for fruits/vegetables/legumes/nuts <5, then FSAm-NPS score = Total A points – (points for fibre + points for fruits/vegetables/legumes/nuts).

Exceptions were made for cheese, added fat, and drink to better rank them according to their nutrient profile, consistently with nutritional recommendations:

Score computation for cheese

For cheese, the score takes in account the protein content, whether the A score reaches 11 or not, i.e.: FSAm-NPS score =Total A points – Total C points

Score computation for added fat

For added fat, the grid for point attribution is based on the percentage of saturated fat among total lipids (instead of saturated fat (g)) and has a six-point homogenous ascending step, as shown thereafter:

Points	**Saturated Fat/Lipids (%)**
0	< 10
1	< 16
2	< 22
3	< 28
4	< 34
5	< 40
6	< 46
7	< 52
8	< 58
9	< 64
10	≥ 64

Points attribution for the other nutrients follows the grid displayed in “A points” and “C points” above.

Score computation for drinks

For drinks, the grids for point attribution regarding energy, sugars and fruits/vegetables/ legumes/nuts (%) were modified.

Points	Energy (kJ)	Sugars (g)	Fruits/vegetables/legumes/nuts (%)
0	≤ 0	≤ 0	< 40
1	≤ 30	≤ 1.5	
2	≤ 60	≤ 3	> 40
3	≤ 90	≤ 4.5	
4	≤ 120	≤ 6	> 60
5	≤ 150	≤ 7.5	
6	≤ 180	≤ 9	
7	≤ 210	≤ 10.5	
8	≤ 240	≤ 12	
9	≤ 270	≤ 13.5	
10	> 270	> 13.5	> 80

Points attribution for the other nutrients follows the grid displayed in "A points" and "C points" above.

Given the modification of the grid for fruit and vegetables for beverages, the threshold in the final computation to take into account protein content is set at 10 points:

- If Total A points <11, then FSAm-NPS score =Total A points – Total C points
- If Total A points ≥11,
  - If points for fruits/vegetables/legumes/nuts =10, then FSAm-NPS score =Total A points – Total C points
  - Else if points for fruits/vegetables/legumes/nuts <10, then FSAm-NPS score = Total A points – (points for fibre + points for fruits/vegetables/legumes/nuts).

Milk and vegetable milk are not concerned by this exception. Their scores are computed using the overall score computation system.

FSAm-NPS score and Attribution of Nutri-Score colours

Foods (points)	Beverages (points)	Colour	
Min to -1	Water	Dark green	*Highest nutritional quality*
0 to 2	Min to 1	Light green	
3 to 10	2 to 5	Yellow	
11 to 18	6 to 9	Light orange	
19 to max	10 to max	Dark orange	*Lowest nutritional quality*

Santé Publique France 2017, Nutri-Score Logo

2) FSAm-NPS DI score computation at the individual level

An individual consumes many different foods of contrasted nutritional quality, which synergistically influence his/her disease risk. When studying the association between food intakes and chronic diseases, all food items consumed have to be considered (and therefore all associated FSAm-NPS scores) and not just one single food. Therefore, in a second step, the FSAm-NPS DI was computed at the individual level as an energy-weighted mean of the FSAm-NPS scores of all foods and beverages consumed using the following equation (FSi: score of food/beverage i, Ei: energy intake from food/beverage i, n: number of food/beverage consumed):

$$\mathrm{FSAm-NPS\ DI} = \frac{\sum_{i=1}^{n}(\mathrm{FS_i E_i})}{\sum_{i=1}^{n}\mathrm{E_i}}$$

Higher FSAm-NPS DI therefore reflects lower nutritional quality in foods consumed.

Sensitivity analyses

Sensitivity analyses were performed based on Model 1 by excluding T2D cases having occurred during the first two years of each participant's follow-up to avoid reverse causality bias, unadjusting for BMI, and testing further adjustments for "Healthy" and "Western" dietary patterns obtained by Principal Component Analysis (details below) (continuous), number of smoked cigarettes in pack-years (continuous), and the season of inclusion in the cohort (spring/ summer/ autumn/ winter, to account for potential variation across the four seasons for diet and other baseline covariates). Models were also tested after restriction of the population study to the participants with ≥6 24h-dietary records during the first two years of follow-up; after restriction to participants having accomplished at least two, three and four years of follow-up; after starting follow-up two years after enrollment; and after the excluding prevalent cases of hypertension and dyslipidemia. A supplementary analysis was also performed by using the Fine and Gray model[13] as an alternative method to account for competing risks due to death during follow-up. In order to check whether specific food groups of the UPF category were entirely driving the main association between the proportion of UPF in the diet and T2D risk, adjustments for the consumption of different types of ultra-processed beverages and food groups were tested. Adjustments for plain water, coffee and tea consumption were also tested.

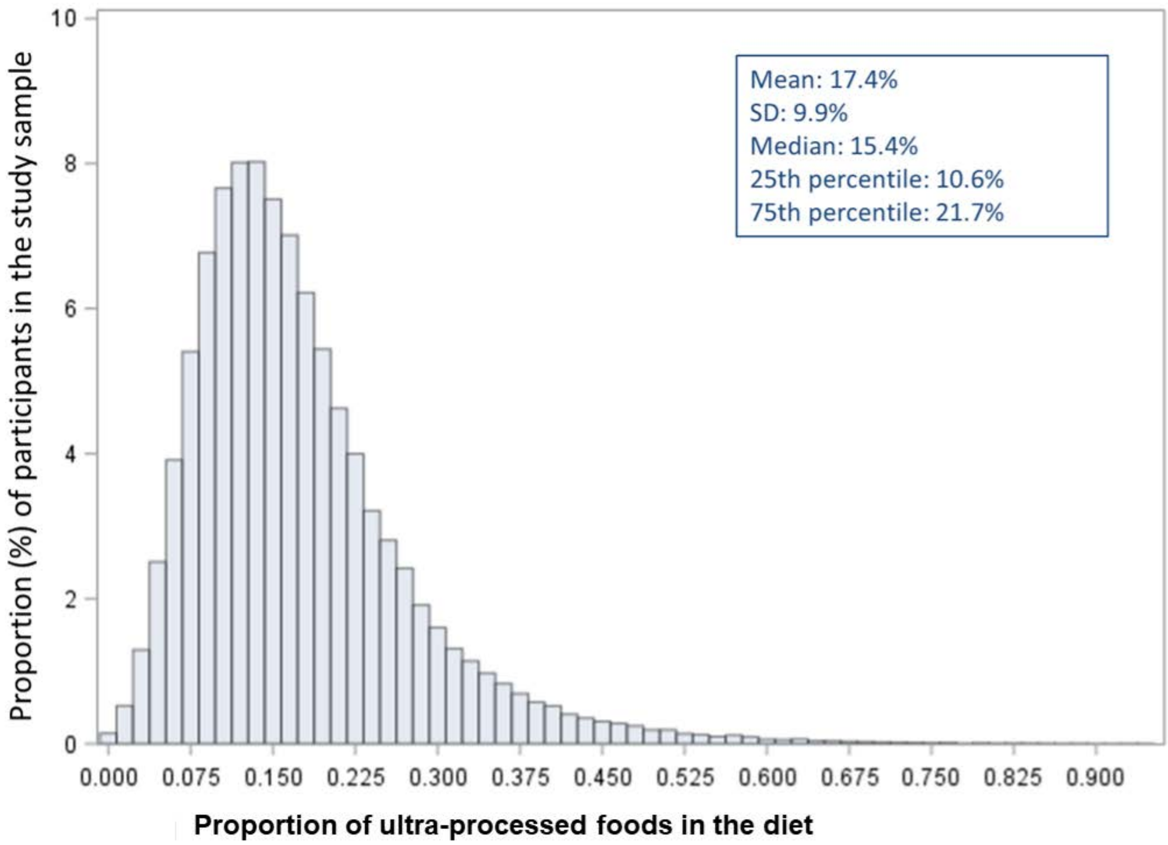
The association between ultra-processed food and overall T2D risk was also investigated separately in different strata of the population: men/women, adults aged <45y / ≥45y, participants with higher sugar intakes (>median)/those with a lower one. All these sensitivity analyses are presented in eTable 1. Finally, we have tested the associations between the proportion of ultra-processed foods in each specific food group and T2D risk (eTable 2).

Method for deriving dietary patterns by principal component analysis and corresponding factor loadings

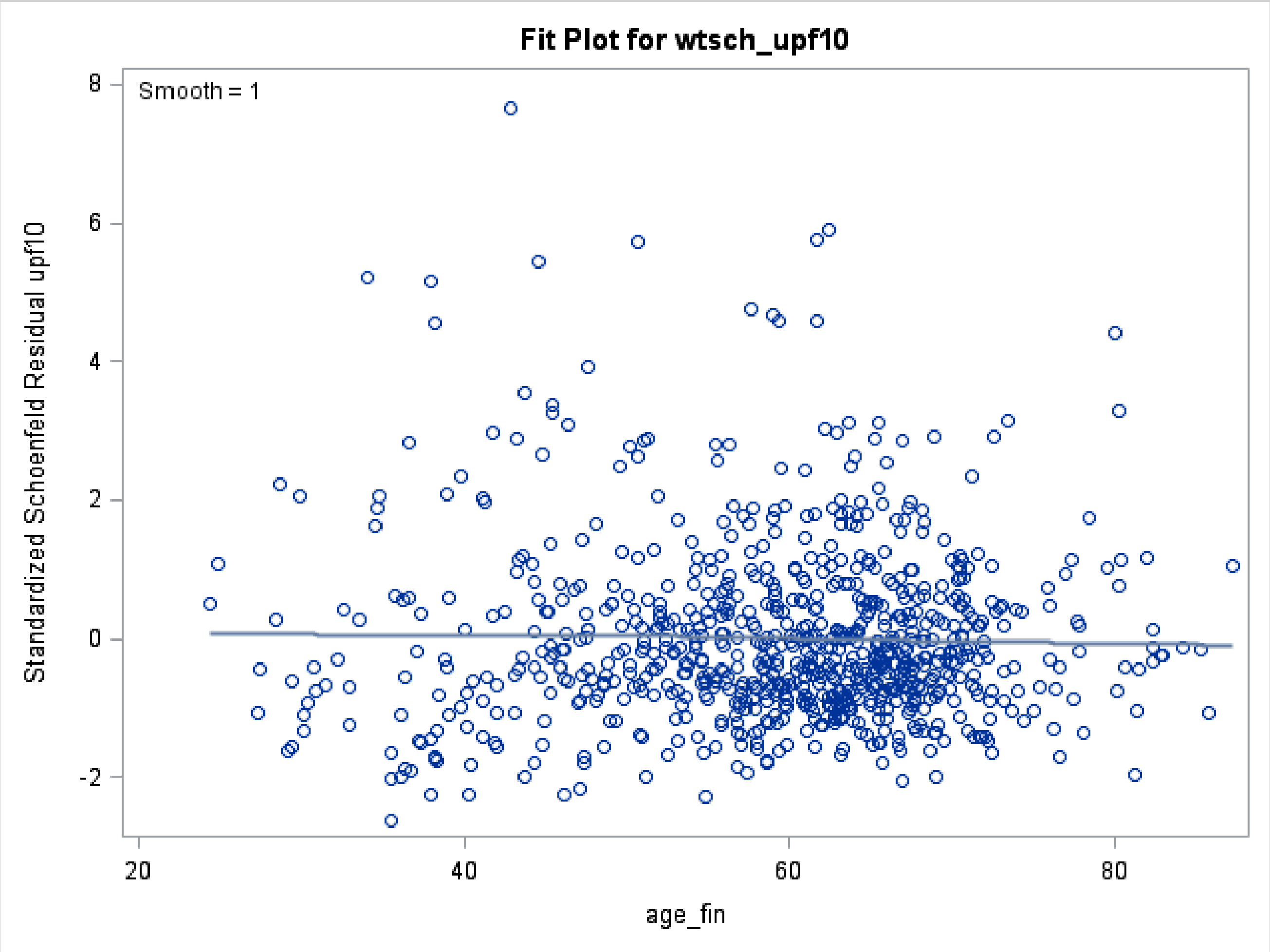
Dietary patterns were produced from principal-components analysis based on 20 predefined food groups, using the SAS ‘‘Proc Factor’’ procedure (SAS Institute Inc., Cary, North Carolina). This factor analysis forms linear combinations of the original food groups, thereby grouping together correlated variables. Coefficients defining these linear combinations are called factor loadings. A positive factor loading means that the food group is positively associated with the factor, whereas a negative loading reflects an inverse association with the factor. For interpreting the data, we considered foods with a loading coefficient under -0.25 or over 0.25. We rotated factors by orthogonal transformation using the SAS ‘‘Varimax’’ option to maximize the independence (orthogonality) of retained factors and obtain a simpler structure for easier interpretation. In determining the number of factors to retain, we considered eigenvalues greater than 1.25, the scree test (with values being retained at the break point between components with large eigenvalues and those with small eigenvalues on the scree plot), and the interpretability of the factors. For each subject, we calculated the factor score for each pattern by summing observed consumption from all food groups, weighted by the food group factor loadings. The factor score measures the conformity of an individual’s diet to the given pattern. Labeling was descriptive, based on foods most strongly associated with the dietary patterns. The healthy pattern (explaining 10.6% of the variance) was characterized by higher intakes of fruit, vegetables, soups and broths, unsweetened soft drinks and whole grains and lower sweetened soft drinks intake. The Western pattern (explaining 7.0% of the variance) was characterized by higher intakes of fat and sauces, alcohol, meat and starchy foods.

	Factor loadings	
	Healthy Pattern	Western Pattern
Alcoholic drinks	-.09	**0.28**
Breakfast cereals	0.07	-.18
Cakes and biscuits	-.19	0.00
Dairy products	0.06	-.01
Eggs	0.07	0.04
Fats and sauces	0.01	**0.54**
Fish and seafood	0.20	0.10
Fruit	**0.35**	0.05
Meat	-.18	**0.31**
Pasta and rice	-.21	**0.34**
Potatoes and tubers	-.02	**0.40**
Poultry	-.03	0.06
Processed meat	-.22	0.20
Pulses	0.19	0.02
Soups and broths	**0.26**	0.22
Sugar and confectionery	-.08	0.12
Sweetened soft drinks	**-.28**	-.00
Unsweetened soft drinks	**0.25**	0.15
Vegetables	**0.47**	0.23
Whole grains	**0.38**	-.04

eFigure 1: Distribution of the main exposure (proportion of ultra-processed food in the diet) in the study sample (N=104,707), NutriNet-Santé, France

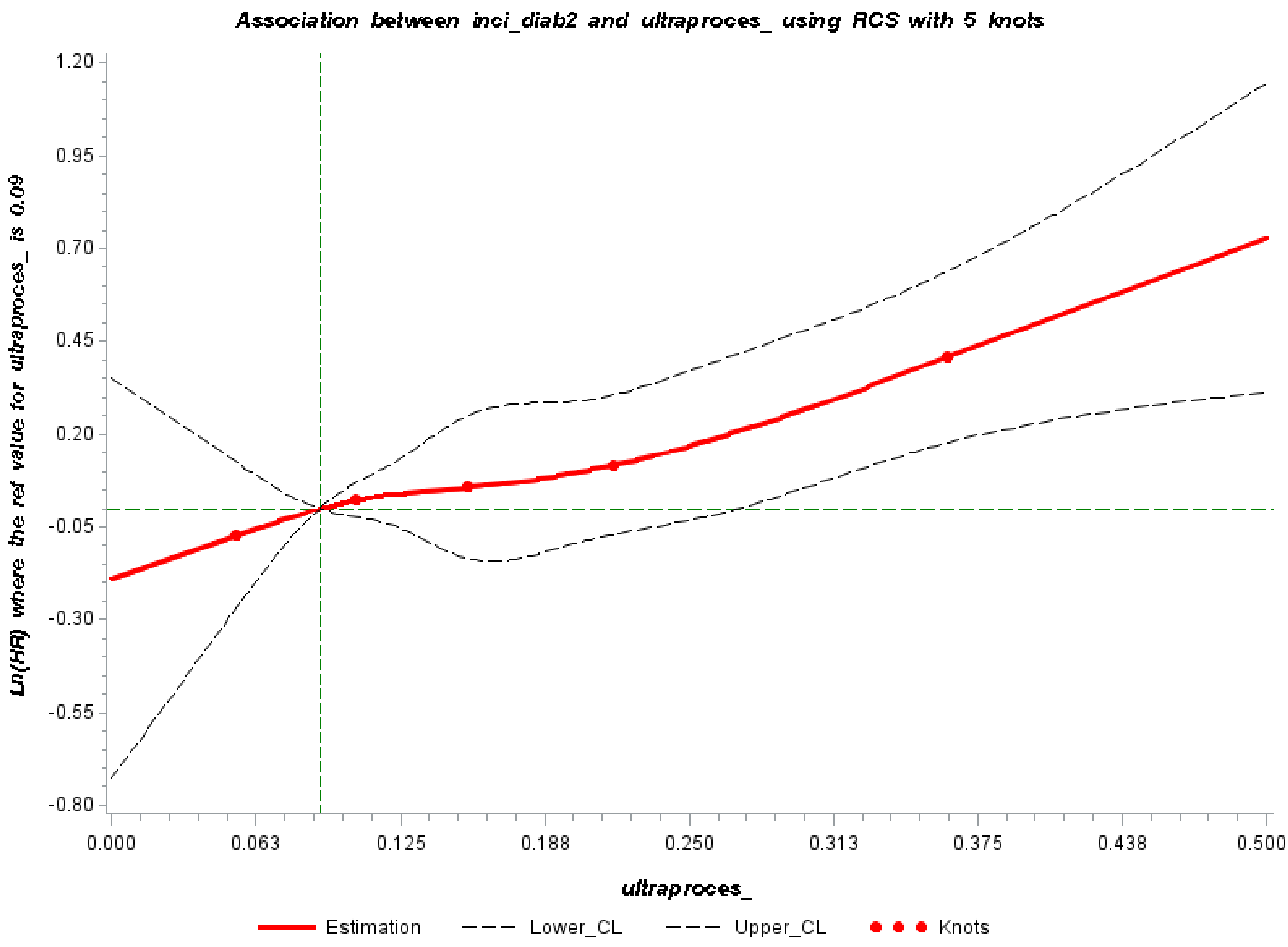


eFigure 2: Cox model proportional risk assumption testing (Schoenfeld residuals)



P-value for correlation between Schoenfeld residuals and time = 0.63

eFigure 3: Spline plot for the linearity assumption of the association between the proportion of ultra-processed food in the diet and the risk of Type-2 Diabetes using Restricted cubic spline (RCS) SAS Macro® developed by Desquilbet and Mariotti[14]



P-value for non-linearity = 0.78

eResults

The mean number of dietary records was 5.7 (SD=3.1); a small proportion of participants (7.6%) had only two dietary records. In a validation analysis we observed that, in participants with ≥8 records, the proportion of UPF was similar when considering all their available records and when considering only the first two ones, with a Pearson correlation coefficient of 0.8.

Main ultra-processed food groups consumed were sugary products (28%) followed by ultra-processed fruits and vegetables (18%), beverages (16%), starchy foods and breakfast cereals (11%), and processed meat and fish (11%).

In higher consumers of UPF (fourth quartile of the % of UPF in the diet), mean proportions of foods were 59.3%, 1.8%, 8.3% and 30.5% for the four NOVA categories, respectively (unprocessed/minimally processed, culinary ingredients, processed, and ultra-processed). Corresponding proportions were 83.1%, 1.5%, 6.3 % and 9.0% in higher consumers of unprocessed/minimally processed foods (fourth quartile).

Results for sensitivity analyses are presented in eTable1 below. The findings remained robust throughout all sensitivity models. In particular, further adjustments tested did not substantially modify the findings, nor did exclusion of T2D cases occurring during the first two years of follow-up, or testing the beginning of follow-up two years after inclusion. Results from the main model (association between the proportion of UPF in the diet and T2D risk, Model 1) also remained statistically significant after adjustment for the consumption of different types of ultra-processed beverages and food groups, showing that no UP food group alone entirely explained the whole association.

eTable 1: Associations between ultra-processed food intake and Type 2-diabetes risk from multivariable Cox proportional hazard models - sensitivity analyses, NutriNet-Santé cohort, France, 2009 – 2019 (n=104,707)

	Cases/total	HR* (95% CI)	P-value
Model a	544/104430	1.16 (1.05 to 1.28)	0.004
Model b1	544/87296	1.15 (1.03 to 1.28)	0.009
Model b2	463/76634	1.18 (1.05 to 1.32)	0.003
Model b3	368/69055	1.20 (1.06 to 1.36)	0.003
Model c	544/87296	1.15 (1.03 to 1.27)	0.01
Model d	821/104707	1.20 (1.11 to 1.31)	<.0001
Model e	821/104707	1.13 (1.04 to 1.24)	0.004
Model f	821/104707	1.15 (1.06 to 1.25)	0.0009
Model g	821/104707	1.15 (1.06 to 1.25)	0.0007
Model h	589/51931	1.16 (1.05 to 1.29)	0.004
Model i	428/90983	1.16 (1.04 to 1.29)	0.008
Model j	821/104707	1.15 (1.05 to 1.25)	0.001
Model k	821/104707	1.12 (1.02 to 1.23)	0.02
Model l	821/104707	1.16 (1.06 to 1.27)	0.01
Model m	821/104707	1.10 (1.00 to 1.21)	0.04
Model n	821/104707	1.14 (1.05 to 1.24)	0.002
Model o	821/104707	1.15 (1.05 to 1.25)	0.003
Model p	821/104707	1.14 (1.04 to 1.24)	0.003
Model q	821/104707	1.20 (1.10 to 1.31)	0.002
Model r	821/104707	1.14 (1.04 to 1.24)	0.003
Model s	821/104707	1.15 (1.06 to 1.26)	0.001
Model t	821/104707	1.23 (1.13 to 1.34)	<0.0001
Model u	821/104707	1.16 (1.07 to 1.26)	0.0005
Model v1	302/21800	1.03 (0.88 to 1.20)	0.7
Model v2	519/82907	1.13 (1.08 to 1.34)	0.0004
Model w1	144/59247	1.19 (1.03 to 1.37)	0.02
Model w2	677/45460	1.13 (1.02 to 1.24)	0.02
Model x1	509/52347	1.13 (1.02 to 1.27)	0.02
Model x2	312/52360	1.22 (1.08 to 1.38)	0.001

CI: confidence interval, HR: cause-specific hazard ratio for T2D risk accounting for competing events (n=340 competing deaths)
*HR for an absolute increment of 10 in the percentage of ultra-processed foods in the diet (except for Fine and Gray Model (model k) for which subdistribution Hazard Ratios sHR are computed[13])
Model 1 is adjusted for age (timescale), sex (except when stratified), educational level, BMI, physical activity, smoking status, alcohol intake, number of 24h-dietary records, energy intake, FSAm-NPS DI and family history of T2D.
Multiple imputation for missing data was performed using the MICE method[15] by fully conditional specification (FCS, 20 imputed datasets) for the following covariates: level of education, physical activity level, smoking status and BMI. Results were combined across imputation based on Rubin's combination rules[16,17] using the SAS PROC MIANALYZE procedure[18].
Model a: Model 1 excluding T2D cases of the first two years of follow-up
Model b1: Model 1 excluding participants having less than two years of follow-up
Model b2: Model 1 excluding participants having less than three years of follow-up
Model b3: Model 1 excluding participants having less than four years of follow-up
Model c: Model 1 considering the beginning of the follow-up 2 years after the inclusion in the cohort
Model d: Model 1 unadjusted for BMI
Model e: Model 1 + Healthy and Western dietary patterns derived from Principal Component Analysis
Model f: Model 1 + number of pack-years
Model g: Model 1 + season of enrolment in the cohort
Model h: Model 1 excluding participants with less than 6 dietary records
Model i: Model 1 excluding prevalent cases of hypertension and dyslipidemia
Model j: Model 1 using Fine and Gray model accounting for competing risks of death (340 competing deaths during follow-up)
Model k: Model 1 + plain water, tea and coffee consumption
Model l: Model 1 + ultra-processed sugar-sweetened beverages consumption
Model m: Model 1 + ultra-processed artificially sweetened beverages consumption
Model n: Model 1 + ultra-processed carbonated drinks consumption
Model o: Model 1 + ultra-processed dairy products consumption
Model p: Model 1 + ultra-processed fats and sauces consumption
Model q: Model 1 + ultra-processed fruits and vegetables consumption
Model r: Model 1 + ultra-processed meat, fish, and eggs consumption
Model s: Model 1 + ultra-processed starchy foods and cereals consumption
Model t: Model 1 + ultra-processed sugary products consumption
Model u: Model 1 + ultra-processed salty snacks consumption
Model v1: Model 1 among men
Model v2: Model 1 among women
Model w1: Model 1 among younger participants (<45 years old)
Model w2: Model 1 among older participants (≥45 years old)
Model x1: Model 1 among participants with sugar intakes below the median (≤89.61 g/d)
Model x2: Model 1 among participants with sugar intakes above the median (>89.61 g/d)

eTable 2: Associations between the proportion of ultra-processed food in each individual food group and Type 2-diabetes risk from multivariable Cox proportional hazard models, NutriNet-Santé cohort, France, 2009 – 2019 (n=104,707)

	Cases/total	HR* (95% CI)	p-value
Ultra-processed beverages	821/104707	1.13 (1.07 to 1.19)	<.0001
Ultra-processed dairy products	821/104707	1.03 (1.00 to 1.07)	0.05
Ultra-processed fats and sauces	821/104707	1.06 (1.03 to 1.09)	0.0002
Ultra-processed fruits and vegetables	821/104707	0.99 (0.95 to 1.04)	0.7
Ultra-processed meat, fish and eggs	821/104707	1.03 (0.98 to 1.09)	0.2
Ultra-processed starchy foods and cereals	821/104707	1.02 (0.97 to 1.08)	0.3
Ultra-processed sugary products	821/104707	1.03 (1.00 to 1.07)	0.02
Ultra-processed salty snacks	821/104707	0.99 (0.97 to 1.01)	0.3

CI: confidence interval, HR: cause-specific hazard ratio for T2D risk accounting for competing events (n=340 competing deaths)
*HR for an absolute increment of 10 in the percentage of the food group consumed in its ultra-processed form
Models are adjusted for age (timescale), sex, educational level, BMI, physical activity, smoking status, alcohol intake, number of 24h-dietary records, energy intake, FSAm-NPS DI, family history of T2D, and the consumption amount of the specific food group (in g/d)
Multiple imputation for missing data was performed using the MICE method[15] by fully conditional specification (FCS, 20 imputed datasets) for the following covariates: level of education, physical activity level, smoking status and BMI. Results were combined across imputation based on Rubin's combination rules[16,17] using the SAS PROC MIANALYZE procedure[18].

References

1. Le Moulec N, Deheerger M, Preziosi P, et al. Validation du manuel-photos utilisé pour l'enquête alimentaire de l'étude SU.VI.MAX. 2016.

2. Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. *IntJObesRelat Metab Disord*. 2000;24:1119-1130.

3. Arnault N, Caillot L, Castetbon K, et al. Table de composition des aliments, Etude NutriNet-Santé. [Food composition table, NutriNet-Santé study] (in French). Paris: Les éditions INSERM/Economica,. 2013.

4. Hercberg S, Castetbon K, Czernichow S, et al. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health*. 2010;10:242. doi:10.1186/1471-2458-10-242

5. Monteiro CA, Cannon G, Levy RB, et al. NOVA. The star shines bright. *World Nutr*. 2016;7(1-3):28-38.

6. Monteiro CA, Cannon G, Moubarac JC, Levy RB, Louzada ML, Jaime PC. The UN Decade of Nutrition, the NOVA food classification and the trouble with ultra-processing. *Public Health Nutr*. 2017;(1475-2727 (Electronic)):1-13. doi:10.1017/S1368980017000234

7. Moubarac JC, Parra DC, Cannon G, Monteiro CA. Food Classification Systems Based on Food Processing: Significance and Implications for Policies and Actions: A Systematic Literature Review and Assessment. *CurrObesRep*. 2014;3(2162-4968 (Electronic)):256-272. doi:10.1007/s13679-014-0092-0

8. Julia C, Martinez L, Alles B, et al. Contribution of ultra-processed foods in the diet of adults from the French NutriNet-Sante study. *Public Health Nutr*. 2017;(1475-2727 (Electronic)):1-11. doi:10.1017/S1368980017001367

9. Forouhi NG, Balkau B, Borch-Johnsen K, et al. The threshold for diagnosing impaired fasting glucose: a position statement by the European Diabetes Epidemiology Group. *Diabetologia*. 2006;49(5):822-827. doi:10.1007/s00125-006-0189-4

10. Julia C, Méjean C, Touvier M, et al. Validation of the FSA nutrient profiling system dietary index in French adults-findings from SUVIMAX study. *Eur J Nutr*. 2016;55(5):1901-1910. doi:10.1007/s00394-015-1006-y

11. Julia C, Etilé F, Hercberg S. Front-of-pack Nutri-Score labelling in France: an evidence-based policy. *Lancet Public Health*. 2018;3(4):e164. doi:10.1016/S2468-2667(18)30009-4

12. Deschasaux M, Huybrechts I, Murphy N, et al. Nutritional quality of food as represented by the FSAm-NPS nutrient profiling system underlying the Nutri-Score label and cancer risk in Europe: Results from the EPIC prospective cohort study. *Plos Med*. 2018;15(9):e1002651. doi:10.1371/journal.pmed.1002651

13. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc*. 1999;94(446):496-509. doi:10.2307/2670170

14. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med*. 2010;29(9):1037-1057. doi:10.1002/sim.3841

15. van BS. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16(0962-2802 (Linking)):219-242. doi:10.1177/0962280206074463

16. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581-592. doi:10.1093/biomet/63.3.581

17. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons; 2004.

18. PROC MIANALYZE: The MIANALYZE Procedure :: SAS/STAT(R) 9.2 User's Guide, Second Edition. https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#mianalyze_toc.htm. Accessed December 19, 2018.