Confi

Consumption of ultra-processed foods and cancer risk: results from the NutriNet-Santé prospective cohort

Journal:	*BMJ*
Manuscript ID	BMJ.2017.040538.R2
Article Type:	Research
BMJ Journal:	BMJ
Date Submitted by the Author:	08-Dec-2017
Complete List of Authors:	Fiolet, Thibault; Nutritional Epidemiology Research Team (EREN), Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University Srour, Bernard; Inserm - French National Institute of Health and Medical Research, Nutritional Epidemiology Research Team (EREN), Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University Sellem, Laury; Nutritional Epidemiology Research Team (EREN), Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University Kesse-Guyot, Emmanuelle; Université Paris 13, COMUE Sorbonne-Paris-Cité, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre d'Epidémiologie et Biostatistiques Paris Nord, Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne-Paris-Cité, Equipe de Recherche en Epidémiologie Nutritionnelle Alles, Benjamin; Université Paris 13, Equipe de Recherche en Epidémiologie Nutritionnelle, Centre de Recherche en Epidémiologie et Statistiques, Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne Paris Cité Méjean, Caroline; INRA, UMR 1110 MISA, 34000 Montpellier, France Deschasaux, Mélanie; Sorbonne Paris Cité Epidemiology and Statistics Research Center, Inserm U1153, Inra U1125, Cnam, Paris 13 University, Nutritional Epidemiology Research Team (EREN) Fassier, Philippine; Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Nutritional Epidemiology Research Team (EREN):, ; Latino-Martel, Paule; Nutritional Epidemiology Research Team (EREN), Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University Beslay, Marie; Nutritional Epidemiology Research Team (EREN), Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University Hercberg, Serge; Université Paris 13, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre d'Epidémiologie et Statistiques Sorbonne Paris Cité, Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne Paris Cité; Département de Santé Publique, Hôpital Avicenne, F-93017, Bobigny Cedex, France.

	Lavalette, Céline; Nutritional Epidemiology Research Team (EREN), Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University Monteiro, Carlos; University of Sao Paulo, Julia, Chantal; Nutritional Epidemiology Research Team (EREN), Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University; Département de Santé Publique, Hôpital Avicenne, F-93017, Bobigny Cedex, France. Touvier, Mathilde; Université Paris 13, Equipe de Recherche en Epidémiologie Nutritionnelle, Centre de Recherche en Epidémiologie et Statistiques, Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne Paris Cité,
Keywords:	ultra-processed food, food processing, cancer, prospective cohort, NOVA classification

SCHOLARONE™
Manuscripts

For Review Only

Consumption of ultra-processed foods and cancer risk: results from the NutriNet-Santé prospective cohort

Thibault Fiolet[1*], Bernard Srour[1*], Laury Sellem[1], Emmanuelle Kesse-Guyot[1], Benjamin Allès[1], Caroline Méjean[2], Mélanie Deschasaux[1], Philippine Fassier[1], Paule Latino-Martel[1], Marie Beslay[1], Serge Hercberg[1, 4], Céline Lavalette[1], Carlos A. Monteiro[3], Chantal Julia[1, 4], Mathilde Touvier[1]

*Equally contributed

Running Head: Ultra-processed foods and cancer risk

Job titles:

Thibault Fiolet: Epidemiology and biostatistics Master Intern

Bernard Srour: Pharmacist, PhD candidate in epidemiology

Laury Sellem: Nutrition and Public Health Master Intern

Emmanuelle Kesse-Guyot: Senior Researcher in nutritional epidemiology

Benjamin Allès: Junior Researcher in nutritional epidemiology

Caroline Méjean: Senior Researcher in nutritional epidemiology

Mélanie Deschasaux: Post-Doctoral Researcher in nutritional epidemiology

Philippine Fassier: Post-Doctoral Researcher in nutritional epidemiology

Paule Latino-Martel: Senior Researcher in nutritional epidemiology

Marie Beslay: Nutrition and Public Health Master Intern

Serge Hercberg: Professor of Nutrition and Public Health, Head of the EREN team

Céline Lavalette: Epidemiology and Biostatistics Master Intern

Carlos A. Monteiro: Professor of Nutrition and Public Health

Chantal Julia: Senior Researcher in nutritional epidemiology

Mathilde Touvier: Senior Researcher in nutritional epidemiology, Head of the Nutrition and Cancer group at EREN

Confidential: For Review Only

Affiliations:

[1] Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University, Nutritional Epidemiology Research Team (EREN), Bobigny, France

[2] INRA, UMR 1110 MOISA, 34000 Montpellier, France

[3] Department of Nutrition, School of Public Health, University of São Paulo, Av. Dr Arnaldo 715,São Paulo 01246-904, Brazil

[4]Public Health Department, Avicenne Hospital, AP-HP, Bobigny, France

Corresponding author:

Bernard Srour

Equipe de Recherche en Epidémiologie Nutritionnelle (EREN)

Centre de Recherche en Epidémiologie et Statistiques Sorbonne Paris Cité, Inserm (U1153), Inra(U1125), Cnam, Université Paris 13

SMBH Paris 13, 74 rue Marcel Cachin F-93017 Bobigny Cedex. France

Tel : +33 1 48 38 89 68

E-mail : b.srour@eren.smbh.univ-paris13.fr

Word count: 5,364 ; **Tables**: 2 ; **Figures**: 2 ; **Supplemental material** (**Appendixes):** 6

Keywords: ultra-processed food; food processing; cancer; prospective cohort; NOVA classification

Confidential. For Review Only

BOX:

WHAT IS ALREADY KNOWN ON THIS TOPIC

Dietary habits are shifting in many countries through an upsurge in the consumption of ultra-processed foods, which are often characterized by a lower nutritional quality but also the presence of food additives, of substances from materials and packaging in contact with food, and of neoformed compounds during production, processing, and storage.

The few studies performed observed that ultra-processed food intake was associated with a higher incidence of dyslipidaemia in Brazilian children, and higher risks of overweight, obesity and hypertension in a cohort of Spanish University students.

Although epidemiological data regarding their relevance to cancer risk are lacking, mechanistic studies suggest potential carcinogenic effects of several components commonly found in ultra-processed foods.

WHAT THIS STUDY ADDS

For the first time, this study assessed the associations between ultra-processed food consumption and cancer risk in a large prospective cohort (n=104,980).

A 10% increase in the proportion of ultra-processed foods in the diet was associated with a >10% significant increase in overall and breast cancer risks.

If confirmed in other populations and settings in the future, these results suggest that the rapidly increasing consumption of ultra-processed foods may drive an increasing burden of cancer in the next decades. Thus, individual recommendations to improve dietary choices, as well as policy actions targeting product reformulation, taxation and marketing restrictions on ultra-processed products and promotion of fresh or minimally processed foods may contribute to primary cancer prevention.

Confidential: For Review Only

STRUCTURED ABSTRACT:

OBJECTIVE

To assess the prospective associations between ultra-processed food consumption and cancer risk.

DESIGN

Population based cohort study

SETTING AND PARTICIPANTS

In all, 104,980 participants aged ≥18y (median age: 40.4y) from the French NutriNet-Santé cohort (2009-2017) were included. Dietary intakes were collected using repeated 24h-dietary records, designed to register participants' usual consumption for 3300 different food items. These were categorized according to their degree of processing by the NOVA classification.

MAIN OUTCOME MEASURES

Associations between ultra-processed food intake and overall, breast, prostate and colorectal cancer risk were assessed by multivariable Cox Proportional Hazard models adjusted for known risk factors.

RESULTS

Ultra-processed food intake was associated with higher overall cancer risk (n=2,228 cases, $HR_{\text{for a 10\% increment in the proportion of ultra-processed food in the diet}}$=1.12 (1.06-1.18), P-trend<.0001) and breast cancer risk (n=739 cases, HR= 1.11 (1.02-1.22), P-trend=0.02). These results remained statistically significant after adjustment for several markers of the nutritional quality of the diet (lipid, sodium and carbohydrate intakes and/or a Western pattern derived by principal component analysis).

CONCLUSIONS

Confidential: For Review Only

In this large prospective study, a 10% increase in the proportion of ultra-processed foods in the diet was associated with a >10% significant increase in overall and breast cancer risks. Further studies are needed to better understand the relative impact of the various dimensions of processing (nutritional composition, food additives, contact materials, and neoformed contaminants) in these relationships.

Confidential: For Review Only

RESTRUCTURED ABSTRACT:

Study question: The objective was to assess the associations between ultra-processed food consumption and cancer risk in a large population-based cohort study.

Methods: In all, 104,980 participants aged ≥18y (median age: 40.4y) from the French NutriNet-Santé cohort (2009-2017) were included. Dietary intakes were collected using repeated 24h-dietary records, designed to register participants' usual consumption for 3300 different food items. These were categorized according to their degree of processing by the NOVA classification. Associations between ultra-processed food intake and overall, breast, prostate and colorectal cancer risk were assessed by multivariable Cox Proportional Hazard models adjusted for known risk factors.

Study answer and limitations: A 10% increase in the proportion of ultra-processed foods in the diet was associated with a >10% significant increase in overall and breast cancer risks. This study was observational, thus, causality of the observed associations cannot be established. Further studies are needed to investigate these associations in a longer term and to better understand the relative impact of the various dimensions of processing (nutritional composition, food additives, contact materials, and neoformed contaminants) in these relationships.

What this study adds: Our study suggests that the consumption of ultra-processed foods may increase cancer risk.

Funding, competing interests, data sharing: This study was supported by public institutions only and the authors have no competing interest to declare. No additional data is available.

Study registration: The NutriNet-Santé cohort is registered at clinicaltrials.gov (NCT03335644).

Confidential: For Review Only

INTRODUCTION

Cancer represents a major worldwide burden with 14.1 million new cases diagnosed in 2012 [1]. According to the World Cancer Research Fund / American Institute for Cancer Research (WCRF/AICR), about one third of the most common neoplasms could be avoided by changing lifestyle and dietary habits in developed countries [2]. Therefore, reaching a balanced and diversified diet (along with tobacco avoidance and alcohol reduction) should be considered as one of the most important modifiable risk factors in the primary prevention of cancer [3].

At the same time, during the last decades, in many countries, diets have shifted towards a dramatic increase in ultra-processed foods consumption [4-8]. After undergoing multiple physical, biological and/or chemical processes, these food products are conceived to be microbiologically safe, convenient, highly palatable and affordable [9;10]. Several surveys (in Europe, the USA, Canada, New Zealand and Brazil) assessing individual food intake, household food expenses or supermarket sales, suggested that ultra-processed food products contribute to between 25 and 50% of total daily energy intake [10-18].

This dietary trend may be concerning and deserves investigation. Indeed, several characteristics of ultra-processed foods may be involved in disease – in particular cancer – aetiology. First, ultra-processed foods often have a higher content in total fat, saturated fat, added sugar and salt, along with a lower fibre and vitamin density [10-17;19]. Beyond nutritional composition, neoformed contaminants, some of which having carcinogenic properties (such as acrylamide, heterocyclic amines, polycyclic aromatic hydrocarbons, etc.) are present in heat-treated processed food products due to the Maillard reaction [20]. Next, the packaging of ultra-processed foods may contain some contact materials for which carcinogenic and endocrine disruptor properties have been postulated such as Bisphenol A [21]. Finally, ultra-processed foods contain authorized[22] but controversial food additives such as sodium nitrite in processed meat or titanium dioxide (TiO_2, white food pigment), for which carcinogenicity has been suggested in animal or cellular models [23;24].

Studying potential health impacts of ultra-processed foods is a very recent field of research, facilitated by the development of the NOVA classification of products according to their degree of food processing [9]. Nonetheless, epidemiological evidence linking ultra-processed food intake to disease risk is still very

https://mc.manuscriptcentral.com/bmj

Confidential. For Review Only

scarce[25-27] and mostly based on cross-sectional and ecological studies. The few studies performed observed that ultra-processed food intake was associated with a higher incidence of dyslipidaemia in Brazilian children [28], and higher risks of overweight, obesity [29] and hypertension [30] in a prospective cohort of Spanish University students.

To our knowledge, the present prospective study was the first to evaluate the association between the consumption of ultra-processed food products and the incidence of cancer, based on a large cohort study with detailed and up-to-date dietary intake assessment.

MATERIAL AND METHODS

Study population

The NutriNet-Santé study is an ongoing web-based cohort launched in 2009 in France with the objective to study the associations between nutrition and health as well as the determinants of dietary behaviours and nutritional status. This cohort has been previously described in detail [31]. Briefly, participants aged over 18 years with access to the Internet are continuously recruited since May 2009 among the general population by means of vast multimedia campaigns. All questionnaires are completed online using a dedicated website (www.etude-nutrinet-sante.fr). Participants are followed using an online platform connected to their email address. They have the possibility to change their email address, phone number or postal address at any moment on the NutriNet-Santé website. Newsletters and alerts about new questionnaires are sent via email. In case of an "undelivered email to recipient" problem, participants are then contacted by telephone and then by regular mail. The NutriNet-Santé study is conducted according to the Declaration of Helsinki guidelines and was approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm n°0000388FWA00005831) and the "Commission Nationale de l'Informatique et des Libertés" (CNIL n°908450/n°909216). It is registered at clinicaltrials.gov as NCT03335644. Electronic informed consent is obtained from each participant.

Data collection

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

At inclusion, participants completed a set of five questionnaires related to socio-demographic and lifestyle characteristics [32] (e.g. date of birth, sex, occupation, educational level, smoking status, number of children), anthropometry [33;34] (e.g. height, weight), dietary intakes (see below), physical activity (validated 7-day International Physical Activity Questionnaire [IPAQ]) [35], and health status (e.g. personal and family history of diseases, medication use including use of hormonal treatment for menopause, oral contraceptive, and menopausal status).

Participants were invited to complete a series of three non-consecutive validated web-based 24h-dietary records every 6 months (to vary the season of completion), randomly assigned over a 2-week period (2 weekdays and 1 weekend day) [36-38]. To be included in the nutrition component of the NutriNet-Santé cohort, only two dietary records were mandatory. Subjects were not excluded if they did not complete all optional questionnaires. Mean dietary intakes from all the 24h-dietary records available during the first two years of each participant's follow-up were averaged and considered as baseline usual dietary intakes in this prospective analysis. The NutriNet-Santé web-based self-administered 24h dietary records have been tested and validated against an interview by a trained dietitian [36], and against blood and urinary biomarkers [37]. Participants used the dedicated web interface to declare all food and beverages consumed during a 24h-period for each of the three main meals (breakfast, lunch, dinner) and any other eating occasion. Portion sizes were estimated using previously validated photographs or usual containers [39]. 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 [40]. Mean daily alcohol, micro- and macro-nutrient and energy intake were calculated using the NutriNet-Santé food composition database, which contains more than 3,300 different items [41]. Amounts consumed from composite dishes were estimated using French recipes validated by nutrition professionals. Sodium intake was assessed via a specific module included in the 24h records, taking into account native sodium in foods, salt added during the cooking, and salt added in the plate. It has been validated against sodium urinary excretion biomarkers [37].

Confidential: For Review Only

Degree of food processing

All food and beverage items of the NutriNet-Santé composition table were categorized into one of the four food groups in NOVA, a food classification system based on the extent and purpose of industrial food processing[9;42;43]. This study primarily focused on the "ultra-processed foods" NOVA group. This group includes mass-produced packaged breads and buns, sweet or savoury packaged snacks, industrialized confectionery and desserts, sodas and sweetened beverages, meat balls, poultry and fish nuggets and other reconstituted meat products transformed with addition of preservatives other than salt (e.g nitrites), instant noodles and soups, frozen or shelf-stable ready meals, and other food products made mostly or entirely from sugar, oils, and fats and other substances not commonly used in culinary preparations such as hydrogenated oils, modified starches, and protein isolates. 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 or to disguise undesirable qualities of the final product. The ultra-processed food 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[44], home-made and artisanal food preparations were identified and decomposed using standardized recipes, and the NOVA classification was applied to their ingredients. Precisions and examples are presented in Appendix 1.

Case ascertainment

Confidential. For Review Only

Participants self-declared health events through the yearly health status questionnaire, through a specific check-up questionnaire for health events (every three months) or at any time through a specific interface on the study website. For each incident cancer declared, participants were contacted by a physician of the team and asked to provide any relevant medical records. Whenever necessary, the study physicians contacted the physician of the patient and/or hospitals to collect additional information. Afterwards, all medical data were reviewed by a physician expert committee. 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 cohorts to medico-administrative databases of the National health insurance (SNIIRAM databases). Declared health events were therefore completed by the information from these databases, thereby limiting any potential bias due to participants with cancer who may not report their disease to the study investigators. Last, an additional linkage to the French National cause-specific mortality registry (CépiDC) was used to detect death and potentially missed cancer cases for deceased participants. Cancer cases were classified using the International Chronic Diseases Classification, 10th Revision, Clinical Modification (ICD-10). In this study, all first primary cancers diagnosed between the inclusion and January 1st 2017 were considered as cases, except for basal cell skin carcinoma, which was not considered as cancer.

Medical records were obtained for >90% of cancer cases. Because of the high validity of self-reports (95% of self-reported cancers for whom a medical record was obtained were confirmed by our physicians), we included all cases who self-reported incident cancers, unless they were identified as non-case subjects by a pathology report. In the latter situation, they were classified as non-cases.

Statistical analysis

Up to January 1st 2017, 104,980 participants without cancer at baseline and who provided at least 2 valid 24h-dietary records during their 2 first years of follow-up were included. The flow-chart is presented in Appendix 4. For each subject, the proportion (in weight, % g/day) of ultra-processed foods in the total diet was calculated. The proportion of ultra-processed foods in the diet was determined by making a weight ratio rather than an energy ratio in order to take into account processed food that do not provide any energy (in particular artificially sweetened beverages) and non-nutritional issues related to food processing (e.g. neo-

https://mc.manuscriptcentral.com/bmj

Confidential. For Review Only

formed contaminants, food additives and alterations to the structure of raw foods). For all covariates except physical activity, <5% of values were missing and were imputed to the modal value (for categorical variables) or to the median (for continuous variables). Corresponding values are provided in the footnote to Table 1. For physical activity, the proportion of missing values was higher (14%) since the answers of all IPAQ questions were needed to calculate the score. To avoid massive imputation for a non-negligible number of subjects or exclusion of subjects with missing data and risk of selection bias, we included a missing class into the models for this variable. Differences in baseline characteristics of participants between sex-specific quartiles of the proportion of ultra-processed food in the diet were examined using ANOVA or $\chi 2$ tests wherever appropriate. Cox proportional hazards models with age as the primary time-scale were used to evaluate the association between the proportion of ultra-processed foods in the diet (coded as a continuous variable or as sex-specific quartiles) and incidence of overall, breast, prostate and colorectal cancer risk. In these models, cancers of other locations than the one studied were censored at the date of diagnosis (i.e. they were considered as non-cases for the cancer of interest and they contributed person-year until the date of diagnosis of their cancer). Hazard ratios (HR) and 95% confidence intervals (CI) were estimated with the lowest quartile as the reference category. Log-log (survival) vs. log-time plots were generated in order to confirm risk proportionality assumptions. Tests for linear trend were performed using the ordinal score on sex-specific quartiles of ultra-processed food. Participants contributed person-time until the date of cancer diagnosis, the date of last completed questionnaire, the date of death, or January 1st 2017, whichever occurred first. Breast cancer analyses were additionally stratified by menopausal status. For the latter, women contributed person-time to the "pre-menopause model" until their age at menopause and to the "post-menopause model" from their age at menopause. Age at menopause was determined using the yearly health status questionnaires completed during follow-up.

Models were adjusted for age (time-scale), sex, BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low, computed following IPAQ recommendations [35]), smoking status (never or former smokers, current smokers), number of 24h-dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, kcal/d, continuous), family history of cancer (yes/no), and educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school

Confidential: For Review Only

degree). For breast cancer analyses, additional adjustments were performed for the number of biological children (continuous), menopausal status at baseline (menopausal/peri-menopausal/non-menopausal), hormonal treatment for menopause at baseline (for postmenopausal analyses, yes/no) and oral contraception use at baseline (for premenopausal analyses, yes/no) (Model 1=main model). To test for the potential influence of the nutritional quality of the diet in the relationship between ultra-processed food intake and cancer risk, this model was additionally adjusted for lipid, sodium and carbohydrate intakes (Model 2), or for a Western dietary pattern derived from principal component analysis (Model 3) (details in Appendix 2), or for all these nutritional factors together (Model 4). Besides, mediation analyses were carried out according to the method proposed by Lange et al. [45] to evaluate the direct and indirect effect of the relationship between the exposure and the outcome through these following nutritional mediators: intakes of sodium, total lipids, saturated, mono-unsaturated and poly-unsaturated fatty acids, carbohydrates, and a Western-type dietary pattern. The methodology is detailed in Appendix 3.

Sensitivity analyses were performed based on Model 1 by i) excluding cancer cases diagnosed during the first two years of each participant's follow-up to avoid reverse causality bias, ii) testing sex-specific quintiles of the proportion of ultra-processed foods in the diet instead of sex-specific quartiles, and iii) testing further adjustments for prevalent depression at baseline (yes/no), dietary supplement use at baseline (yes/no), healthy dietary pattern (continuous, details in Appendix 2), number of smoked cigarettes in pack-years (continuous), overall fruit and vegetable consumption (continuous) and season of inclusion in the cohort (spring/summer/autumn/winter). The association between ultra-processed food and overall cancer risk was also investigated separately in different strata of the population: men, women, younger adults (<40y), older adults (≥40y), smokers, non-smokers, participants with a high level of physical activity and those with low-to-moderate level of physical activity. Models were also tested after restriction of the population study to the participants with at least six (respectively, at least one) 24h dietary records during the first two years of follow-up. Associations between the quantity (g/d) of each ultra-processed food group and cancer risk were also tested.

Confidential: For Review Only

Secondary analyses were performed by testing the associations between the proportion in the diet of each of the three other NOVA categories of food processing (continuous) with cancer risk, using multivariate Cox models adjusted for model 1 covariates.

All tests were two-sided, and P<0.05 was considered statistically significant. SAS version 9.4 (SAS Institute) was used for the analyses.

Confidential. For Review Only

RESULTS

A total of 104,980 participants with 22,821 (21.7%) men and 82,159 (78.3%) women were included in the present study. Mean age of participants was 42.8y (SD=14.8) years (range: 18.0-72.8y). Mean number of dietary records per subject over their first two years of follow-up was 5.4 (SD=2.9); the minimum was 2, but it only represented 7.2% of the participants (n=7558/104,980). After the launching of the study by the end of May 2009, half of the records were filled between June and November and the other half between December and May. Main baseline characteristics of participants according to quartiles of the proportion of ultra-processed foods in the diet are described in Table 1. Compared to the first quartile, participants among the highest quartile of ultra-processed food intake tended to be younger, current smokers, less educated, with less family history of cancer and a lower physical activity level. Furthermore, they had higher intakes of energy, lipids, carbohydrates and sodium, along with lower alcohol intake. Although there was a higher proportion of women than men in this cohort, the contribution of ultra-processed foods to the overall diet was very similar between men and women (18.74% for men and 18.71% for women, p=0.7). The distribution of the proportion of ultra-processed food in the diet in the study population is presented in Appendix 5. Main food groups contributing to ultra-processed food intake were sugary products (26%) and beverages (20%), followed by starchy foods and breakfast cereals (16%) and ultra-processed fruits and vegetables (15%) (Figure 1).

TABLE 1 Baseline characteristics of the study population according to sex-specific quartiles of ultra-processed food consumption (n=104,980), NutriNet-Santé cohort, France, 2009-2017[a]

		Quartiles of ultra-processed food consumption[b]				
	All participants	**Quartile 1 (n=26,244)**	**Quartile 2 (n=26,245)**	**Quartile 3 (n=26,246)**	**Quartile 4 (n=26,245)**	**P-trend[c]**
Age, years	42.8 ± 14.8	47.9 ± 13.5	45.0 ± 14.0	42.0 ± 14.4	36.5 ± 13.6	<.0001
Sex, n (%)						
Female	82159 (78.3)	20,539 (78.3)	20,540 (78.3)	20,541 (78.3)	205,42 (78.3)	
Male	22821 (21.7)	5,705 (21.7)	5,706 (21.7)	5,707 (21.7)	5,708 (21.7)	
Height, cm	166.8 ± 8.1	166.3 ± 8.0	166.7 ± 8.0	167.0 ± 8.1	167.3 ± 8.2	<.0001
Body mass index, kg/m²	23.8 ± 4.6	23.8 ± 4.3	23.8 ± 4.4	23.8 ± 4.5	23.8 ± 5.0	0.9
Family history of cancer, yes[d]	35668 (34.0)	10,542 (40.2)	9,624 (36.7)	8,625 (32.9)	6,877 (26.2)	<.0001
Higher education, n (%)						0.01
No	19357 (18.4)	5,154 (19.6)	4,961 (18.9)	4,637 (17.7)	4,605 (17.6)	
Yes <2 years	18076 (17.2)	3,938 (15.0)	1,091 (15.6)	4,426 (16.9)	5,621 (21.4)	
Yes ≥2 years	67,547 (64.3)	17,152 (65.4)	17,193 (65.5)	17,183 (65.5)	16,019 (61.0)	
Smoking status, n (%)						<.0001
Current	17,763 (16.9)	4,127 (15.7)	4,065 (15.5)	4,266 (16.3)	5,305 (20.2)	
Never/former	87,217 (83.1)	22,117 (84.3)	22,180 (84.5)	21,980 (83.8)	20,940 (79.8)	
IPAQ Physical activity level, n (%)[e]						<.0001
High	29603 (28.2)	8,753 (33.4)	7,762 (29.6)	6,983 (26.6)	6,105 (23.3)	
Moderate	38874 (37.0)	9,620 (36.7)	9,953 (37.9)	9,814 (37.4)	9,487 (36.2)	
Low	21888 (20.9)	4,407 (13.8)	4,407 (16.8)	5,839 (22.3)	6,490 (24.7)	
Energy intake without alcohol, kcal/d	1879.0±473.7	1,810.6 ± 454.1	1,881.1 ± 457.7	1,908.5 ± 472.3	1,915.8 ± 501.8	<.0001
Alcohol intake, g/d	7.8 ± 11.9	9.3 ± 13.3	8.5 ± 11.9	7.5 ± 11.3	5.9 ± 10.5	<.0001
Total Lipid intake, g/d	80.5 ± 25.5	76.0 ± 24.3	80.3 ± 24.4	82.1 ± 25.3	83.4 ± 27.3	<.0001
Carbohydrate intake, g/d	195.4 ± 57.9	184.6 ± 57.8	193.9 ± 55.3	199.3 ± 56.6	203.6 ± 60.2	<.0001
Sodium intake, mg/d	2,700.1 ± 893.1	2,589.3 ± 881.6	2,731.8 ± 871.0	2,761.9 ± 884.1	2,717.7 ± 925.0	<.0001

Confidential: For Review Only

Number of children	1.3 ± 1.2	1.6±1.2	1.4±1.2	1.3±1.2	1.0±1.2	<.0001
Menopausal status, n (%)[f]						<.0001
Premenopausal	57408 (69.9)	11,797 (57.4)	13,497 (65.7)	14,961 (72.8)	17,153 (83.5)	
Perimenopausal	4282 (5.2)	1,471 (7.16)	1,148 (5.6)	997 (4.9)	666 (3.2)	
Postmenopausal	20469 (24.9)	7,271 (35.4)	5,895 (28.7)	4,582 (22.3)	2,721 (13.3)	
Use of hormonal treatment for menopause, yes n (%)[f]	4324 (5.3)	1602 (7.8)	1242 (6.1)	932 (4.5)	548 (2.7)	<.0001
Oral contraception, yes n (%)[f]	23073 (22.0)	3,779 (14.4)	4,990 (19.0)	6,209 (23.7)	8,095 (30.8)	<.0001
Ultraprocessed food (%)	18.7 ± 10.1	8.5 ± 2.5	14.3 ± 1.4	19.8 ± 1.9	32.3 ± 9.8	-

[a]Values are means ± SDs or n (%). For all covariates except physical activity, a very low proportion of values were missing (0-5%), the latter were replaced by the modal value among the population study: '≥2y of higher education' for educational level, 0 for the number of biological children, 22.9 kg/m2 for BMI, 166 cm for height and non-smoker for smoking status.

[b]Sex specific quartiles of the proportion of ultra-processed food intake in the total quantity of food consumed. Sex-specific cut-offs for quartiles of ultra-processed proportions were 11.8%, 16.8% and 23.3% in men and 11.8%, 16.8% and 23.4% in women.

[c] P_{value} for the comparison between sex-specific quartiles of ultra-processed food consumption, by Fisher test or x^2 test where appropriate.

[d]Among first-degree relatives

[e] Available for 90,365 subjects. Subjects were categorized into the "high", "moderate" and "low" categories according to IPAQ guidelines[35]

[f]Among women

Confidential: For Review Only

During follow-up (426,362 person-years, median follow-up time=5y), 2,228 first incident cancer cases were diagnosed and validated, among which 739 breast cancers (n=264 pre-menopausal and n=475 post-menopausal), 281 prostate cancers and 153 cases of colorectal cancers. Among these 2,228 cases, 108 (4.8%) were identified during mortality follow-up with the national CépiDC database. The abandon rate in the NutriNet-Santé cohort was 6.7%. Associations between the proportion of ultra-processed foods in the diet and overall, breast, prostate and colorectal cancer risks are shown in Table 2. Corresponding cumulative incidence curves are shown in Figure 2. In model 1, ultra-processed food intake was associated with increased risks of overall cancer ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.12 (1.06 to 1.18), P<.0001) and breast cancer (HR=1.11 (1.02 to 1.22), P=0.02). The later association was more specifically observed for post-menopausal breast cancer (P=0.04) but not for pre-menopausal breast cancer (P=0.2). The association with overall cancer risk was statistically significant in all strata of the population investigated, after adjustment for model 1 covariates: in men ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.12 (1.02 to 1.24), P=0.02, 663 cases and 22158 non-cases), in women (HR= 1.13 (1.06 to 1.20), P<0.0001, 1565 cases and 80594 non-cases), in younger adults (<40 years old, HR= 1.21 (1.09 to 1.35), P=0.0006, 287 cases and 48627 non-cases), in older adults (≥40 years old, HR= 1.09 (1.03 to 1.16), P=0.03, 1941 cases and 54485 non-cases), in smokers (including adjustment for pack-years of cigarette smoked, HR =1.18 (1.04 to 1.33), P=0.01, 255 cases and 15355 non-cases), in non-smokers (HR=1.11 (1.05 to 1.17), P=0.0002, 1943 cases and 85219 non-cases), in subjects with low-to-moderate levels of physical activity (HR=1.07 (1.00 to 1.15), P=0.04, 1216 cases and 59546 non-cases), and in those with a high level of physical activity (HR=1.19 (1.09 to 1.30), P<0.0001, 744 cases and 28859 non-cases).

Confidential. For Review Only

TABLE 2 Associations between ultra-processed food intake and overall, prostate, colorectal and breast cancer risk, from multivariable Cox proportional hazard models, NutriNet-Santé cohort, France, 2009 – 2017 (n=104,980)[a]

	Proportion of ultra-processed food intake in the diet						
	Continuous[b]		Sex-specific quartiles[c]				
			Q1	Q2	Q3	Q4	
	HR (95% CI)	*P*-trend	HR	HR (95% CI)	HR (95% CI)	HR (95% CI)	*P*-trend
All cancers							
N for cases/non-cases	2228/102752		712/25532	607/25638	541/25705	368/25877	
Model 1	1.12 (1.06 to 1.18)	<.0001	1	0.99 (0.89 to 1.11)	1.10 (0.99 to 1.24)	1.21 (1.06 to 1.38)	0.002
Model 2	1.12 (1.07 to 1.18)	<.0001	1	1.00 (0.90 to 1.11)	1.11 (0.99 to 1.25)	1.23 (1.08 to 1.40)	0.001
Model 3	1.12 (1.06 to 1.18)	<.0001	1	0.99 (0.89 to 1.11)	1.01 (0.98 to 1.23)	1.21 (1.06 to 1.38)	0.002
Model 4	1.13 (1.07 to 1.18)	<.0001	1	1.00 (0.90 to 1.11)	1.11 (0.99 to 1.24)	1.23 (1.08 to 1.40)	0.001
Prostate cancer							
N for cases/non-cases	281/22540		96/5609	96/5609	59/5647	30/5675	
Model 1	0.98 (0.83 to 1.16)	0.8	1	1.18 (0.89 to 1.57)	0.95 (0.69 to 1.32)	0.93 (0.61 to 1.40)	0.6
Model 2	0.98 (0.83 to 1.16)	0.8	1	1.18 (0.89 to 1.57)	0.95 (0.69 to 1.32)	0.93 (0.61 to 1.40)	0.6
Model 3	0.98 (0.83 to 1.15)	0.8	1	1.18 (0.89 to 1.56)	0.95 (0.68 to 1.31)	0.92 (0.61 to 1.39)	0.6
Model 4	0.98 (0.83 to 1.16)	0.8	1	1.18 (0.89 to 1.57)	0.95 (0.68 to 1.32)	0.93 (0.61 to 1.40)	0.6
Colorectal cancer							
N for cases/non-cases	153/104827		48/26196	43/26202	36/26210	26/26219	
Model 1	1.13 (0.92 to 1.38)	0.2	1	1.10 (0.72 to 1.66)	1.17 (0.76 to 1.81)	1.49 (0.92 to 2.43)	0.1
Model 2	1.16 (0.95 to 1.42)	0.1	1	1.12 (0.74 to 1.70)	1.22 (0.79 to 1.90)	1.59 (0.97 to 2.60)	0.07
Model 3	1.13 (0.92 to 1.38)	0.2	1	1.09 (0.92 to 1.38)	1.16 (0.75 to 1.80)	1.48 (0.91 to 2.41)	0.1
Model 4	1.16 (0.95 to 1.42)	0.1	1	1.12 (0.74 to 1.70)	1.22 (0.79 to 1.89)	1.23 (1.08 to 1.40)	0.07
Breast cancer							
N for cases/non-cases	739/81420		247/20292	202/20338	179/20361	111/20429	
Model 1	1.11 (1.02 to 1.22)	0.02	1	0.97 (0.81 to 1.17)	1.10 (0.90 to 1.34)	1.14 (0.91 to 1.44)	0.2
Model 2	1.11 (1.01 to 1.21)	0.03	1	0.96 (0.80 to 1.16)	1.09 (0.89 to 1.32)	1.12 (0.89 to 1.42)	0.2
Model 3	1.11 (1.02 to 1.22)	0.02	1	0.97 (0.80 to 1.17)	1.09 (0.90 to 1.33)	1.14 (0.91 to 1.44)	0.2
Model 4	1.11 (1.01 to 1.21)	0.03	1	0.96 (0.80 to 1.16)	1.08 (0.89 to 1.32)	1.13 (0.89 to 1.42)	0.2
Pre-menopausal breast cancer							
N for cases/non-cases	264/57151		90/14263	70/14284	55/14299	49/14305	
Model 1	1.09 (0.95 to 1.25)	0.2	1	0.91 (0.67 to 1.25)	0.92 (0.65 to 1.29)	1.30 (0.90 to 1.86)	0.3
Model 2	1.07 (0.93 to 1.23)	0.4	1	0.90 (0.66 to 1.24)	0.90 (0.64 to 1.27)	1.25 (0.87 to 1.80)	0.4
Model 3	1.09 (0.95 to 1.26)	0.2	1	0.91 (0.67 to 1.25)	0.92 (0.66 to 1.30)	1.30 (0.91 to 1.88)	0.3

Confidential: For Review Only

Model 4	1.08 (0.94 to 1.24)	0.3	1	0.91 (0.66 to 1.24)	0.91 (0.64 to 1.28)	1.27 (0.88 to 1.83)	0.4
Post-menopausal							
breast cancer							
N for cases/non-cases	475/29191		107/7309	128/7289	123/7294	117/7299	
Model 1	1.13 (1.01 to 1.27)	0.04	1	1.23 (0.95 to 1.60)	1.28 (0.98 to 1.66)	1.39 (1.07 to 1.82)	0.02
Model 2	1.13 (1.00 to 1.27)	0.05	1	1.23 (0.95 to 1.60)	1.27 (0.98 to 1.65)	1.39 (1.05 to 1.81)	0.02
Model 3	1.13 (1.00 to 1.27)	0.04	1	1.23 (0.95 to 1.59)	1.27 (0.98 to 1.65)	1.38 (1.06 to 1.81)	0.02
Model 4	1.13 (1.00 to 1.27)	0.05	1	1.23 (0.95 to 1.59)	1.27 (0.97 to 1.65)	1.38 (1.05 to 1.81)	0.02

CI, confidence interval, HR, Hazard ratio

[a] Model 1 is a multivariable Cox proportional hazard model adjusted for age (timescale), sex, energy intake without alcohol, number of 24h-dietary records, smoking status, educational level, physical activity, height, BMI, alcohol intake, and family history of cancers. Breast cancer models were additionally adjusted for menopausal status, hormonal treatment for menopause, oral contraception and number of children.

Model 2 = Model 1 + lipid intake, sodium intake, carbohydrate intake

Model 3 = Model 1 + Western dietary pattern (derived by factor analysis)

Model 4 = Model 1 + lipid intake, sodium intake, carbohydrate intake, Western dietary pattern (derived by factor analysis).

Pearson correlation coefficients with the Western dietary pattern were 0.5 for dietary lipids, 0.6 for sodium and 0.40 for carbohydrates.

[b] HR for an increase of 10% of the proportion of ultra-processed food intake in the diet

[c] Sex-specific cut-offs for quartiles of ultra-processed proportions were 11.8% ; 16.8% and 23.3% in men and 11.8% ; 16.8% and 23.4% in women.

In premenopausal women : Cut-offs for quartiles of ultra-processed proportions were 12.8% ; 18.1% and 25.0%. In postmenopausal women : Cut-offs for quartiles of ultra-processed proportions were 10.1% ; 14.3% and 19.5%.

Confidential: For Review Only

More specifically, ultra-processed fats and sauces (P=0.002), sugary products (P=0.03), and beverages (P=0.005) were associated with increased overall cancer risk and ultra-processed sugary products were associated with breast cancer risk (P=0.006) (Appendix 6).

Further adjustment for several indicators of the nutritional quality of the diet (lipid, sodium and salt intakes – model 2; Western pattern – model 3; or both – model 4) did not modify these findings. The Pearson correlation coefficient between the proportion of ultra-processed food in the diet and the Western-type dietary pattern was low (0.06). Consistently, analyses performed according to the method proposed by Lange et al. [45] to assess a potential mediation of the relationship between ultra-processed food and cancer risk by these nutritional factors showed no statistically significant mediation effect of any of the factors tested. The mediated effects ranged between 0 and 2%, with all P>0.05 (Appendix 3).

No association was statistically significant for prostate and colorectal cancers. However, a borderline non-significant trend of increased colorectal cancer risk associated with ultra-processed food intake was observed ($HR_{Q4 \text{ versus } Q1}$=1.23 (1.08 to 1.40), P-trend=0.07 in Model 4).

Sensitivity analyses (adjusted for model 1 covariates, data not tabulated) excluding cancer cases diagnosed during the first two years of follow-up provided similar results ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.10 (1.03 to 1.17), P =0.005 for overall cancer risk, n=1367 cases and 102502 non-cases included; HR=1.15 (1.03 to 1.29), P =0.02 for breast cancer risk, n=441 cases and 80940 non-cases included). Similarly, results were unchanged when non-validated cancer cancers were excluded ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.11 (1.05 to 1.17), P=0.0003 for overall cancer risk, n=1967 cases and 102752 non-cases included; HR=1.12 (1.02 to 1.23), P=0.02 for breast cancer risk, n=677 cases and 81274 non-cases included).

Similar results were observed when i) we included only participants with at least six 24h records on the one hand (overall cancer risk: $HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$= 1.13 (1.06 to 1.21)), P =0.0003, n = 1494 cases and 47 920 non-cases included) and ii) we re-included participants with only one

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

24h record on the other hand (overall cancer risk: $HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.11 (1.06 to 1.16)), P=0.0001, n = 2383 cases and 122 196 non-cases included).

Similar findings were found when the proportion of ultra-processed food in the diet was coded as sex-specific quintiles instead of sex-specific quartiles (overall cancer risk: $HR_{Q5 \text{ versus } Q1}$= 1.25 (1.08 to 1.47), P-trend=0.0003 and breast cancer risk: $HR_{Q5 \text{ versus } Q1}$= 1.25 (0.96 to 1.63), P-trend=0.03).

Further adjustment for the following variables, in addition to model 1 covariates, did not modify the results: dietary supplement use at baseline ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.12 (1.06 to 1.17), P<0.0001 for overall cancer and 1.11 (1.02 to 1.22), P=0.02 for breast cancer), prevalent depression at baseline (HR=1.11 (1.06 to 1.17), P<0.0001 for overall cancer and 1.11 (1.01 to 1.22), P=0.02 for breast cancer), healthy dietary pattern (HR =1.11 (1.05 to 1.17), P<0.0001 for overall cancer and 1.10 (1.00 to 1.21), P=0.04 for breast cancer), overall fruit and vegetable consumption in g/d (HR= 1.10 (1.04 to 1.16), P=0.0009 for overall cancer and 1.11 (1.01 to 1.22), P=0.03 for breast cancer), number of smoked cigarettes in pack-years (HR = 1.13 (1.07 to 1.19), P<0.0001 for overall cancer and 1.13 (1.03 to 1.24), P=0.009 for breast cancer), and season of inclusion in the cohort (HR = 1.12 (1.06 to 1.18), P<0.0001 for overall cancer and 1.12 (1.02 to 1.22), P=0.02 for breast cancer).

Besides, we have tested other methods to deal with missing data, such as multiple imputation[46] and complete case analysis (i.e. exclusion of participants with at least one missing data for a covariate). The results were very similar: for the multiple imputation analysis: $HR_{\text{for a 10-point increment in the proportion of ultraprocessed foods in the diet}}$=1.11 (1.06 to 1.17), P<0.0001, 2228 cases and 102752 non-cases for overall cancer, HR=1.11 (1.01 to 1.21), P=0.02, 739 cases and 81420 non-cases for breast cancer; and for the complete case analysis: HR =1.11 (1.05 to 1.18), P=0.0003, 1813 cases and 82824 non-cases for overall cancer, HR=1.14 (1.03 to 1.26), P=0.01, 579 cases and 64642 non-cases for breast cancer.

As a secondary analysis, associations between the proportions of the three other NOVA degrees of food processing and cancer risk were also tested. No significant associations were found between the proportions of “processed culinary ingredients” nor “processed foods” with cancer risk at any location (all p>0.05). However, and consistently with our findings, the consumption of “minimally/unprocessed foods” was associated with lower risks of overall and breast cancers ($HR_{\text{for a 10-point increment in the proportion of unprocessed foods in the}}$

https://mc.manuscriptcentral.com/bmj

Confidential. For Review Only

$_{diet}$=0.91 (0.87 to 0.95), P<.0001, 2228 cases and 102752 non-cases for overall cancer, HR=0.42 (0.19 to 0.91), P=0.03, 739 cases and 81420 non-cases for breast cancer), in multivariable analyses adjusted for model 1 covariates.

DISCUSSION

Principal findings

In this large prospective cohort, a 10% increase in the proportion of ultra-processed foods in the diet was associated with a 12% and 11% significant increase in overall and breast cancer risks, respectively. While a few studies previously suggested that ultra-processed foods may contribute to increase the risk of cardiometabolic disorders - such as obesity [29], hypertension [30], and dyslipidaemia [28] - no prior prospective epidemiological study evaluated the association between food processing and cancer risk.

Interpretation and comparison with other studies

There is no available estimate of the proportion of ultra-processed food in the diet at the national level in France. However, in the nationally representative INCA3 study conducted by the French Food safety Agency in 2016 [4], "transformed" foods included sweet pastries, biscuits, dairy desserts, ice cream, fruit purée and fruit in syrup, fruit and vegetable juices, soups and broths, sandwiches, pizzas and salted pastries, as well as mixed dishes composed of egg, meat, fish, vegetable and/or starchy foods (cereals, legumes or potatoes). More than half of the "transformed" foods consumed outside catering establishments by adults aged 18-79 were manufactured industrially, about 1/3 were homemade, while the rest was handcrafted (e.g. caterer). These figures illustrate the important share of processed – and especially industrially processed – foods in the diet of French adults.

Several hypotheses could be put forward to explain our findings. The first one relates to the generally poorer nutritional quality of diets rich in ultra-processed foods. Indeed, diets that include a higher proportion of processed food products tended to be richer in energy, sodium, fat and sugar and poorer in fibres and various

Confidential: For Review Only

micronutrients in several studies conducted in various countries [10-17;19]. Ultra-processed foods have also been associated with a higher glycaemic response and a lower satiety effect [47]. Although not being the unique determinant, excessive energy, fat, and sugar intakes contribute to weight gain and obesity risk, the latter being recognized as a major risk factor for the following cancers: post-menopausal breast, stomach, liver, colorectal, oesophagus, pancreas, kidney, gallbladder, endometrium, ovary, liver, prostate (advanced) and hematological malignancies [29]. For instance, body fatness in post-menopausal women is estimated to contribute to 17% of the breast cancer burden [2]. Besides, most of ultra-processed foods, such as dehydrated soups, processed meats, biscuits and sauces, have a high salt content. Salt-preserved foods are associated with increased gastric cancer risk [29]. Conversely dietary fiber intake decreases colorectal cancer risk with a convincing level of evidence [3;29] and may also reduce breast cancer risk [3]. However, the association between ultra-processed food intake and cancer risk observed in this study were statistically significant despite adjustment for BMI, and remained significant after further adjustment for a Western-type dietary pattern and/or energy, fat, sugar and salt content of the diet. Besides, mediation analyses did not support a strong effect of the "nutritional quality" component in this association, thereby suggesting that other bioactive compounds contained in ultra-processed food may contribute to explain the observed relationships.

A second interpretation track concerns the wide range of additives contained in ultra-processed foods. While maximum authorized levels normally protect the consumers against adverse effects of each individual substance in a given food product [48], health impact of the cumulative intake across all ingested foods and potential cocktail/interaction effects remain largely unknown. More than 250 different additives are authorized for an adjunction to food products in Europe and in the US [22;49]. For some of them, experimental studies on animal or cellular models have suggested carcinogenic properties that deserve further investigation in humans [23;24;50-53]. For instance, this is the case for titanium dioxide (TiO_2), a common food additive that contains nanoscale particules and that is used as a whitening agent or in packaging in contact with food or beverages to provide a better texture and anti-microbial properties. Experimental studies, mainly conducted in rodent models, suggested that this additive could initiate or promote the development of colon preneoplastic lesions, as well as chronic intestinal inflammation, thus, TiO_2 was evaluated as "possibly carcinogenic to humans" (Group 2B) by the World Health Organization - International Agency for

Confidential: For Review Only

Research on Cancer (WHO-IARC) [24]. The effects of intense artificial sweeteners such as aspartame on human metabolism and gut microbiota composition/functioning are also controversial [53]. Although previous experimental studies on animals confirmed the safety of aspartame, their relevance to human health outcomes has been questioned, particularly regarding a potential long-term carcinogenicity[51]. Moreover, another concern about sodium nitrite is the formation of carcinogenic nitrosamines in meats containing sodium nitrite when meat is charred or overcooked. These N-nitroso compounds may be involved in the etiology of colorectal cancer [23;52].

Next, food processing and particularly heat treatments produce neoformed contaminants (e.g.acrylamide) in ultraprocessed products such as fried potatoes, biscuits, bread or coffee. A recent meta-analysis underlined a modest association between dietary acrylamide and both kidney and endometrial cancer risks, in non-smokers [54]. In addition, the European Food Safety Agency (EFSA) judged that proofs from animal studies were sufficient to classify acrylamide as genotoxic [20].

Lastly, bisphenol A (BPA) is another contaminant suspected of migrating from plastic packaging of ultra-processed foods. Its endocrine disruptor properties made it judged as "a substance of very high concern" by the European Chemicals Agency (ECHA) [55]. There is increasing evidence for involvement in the development of several non-communicable diseases, including cancer [21] linked to endocrinal disruptors.

Strengths and limitations of the study

Strengths of this study pertained to its prospective design and large sample size, along with a detailed and up-to-date dietary intake assessment. Repeated 24h-dietary records (including 3300 different food items) are more accurate than food frequency questionnaires with aggregated food groups and than household purchasing data. However, some limitations should be acknowledged. First, as it is generally the case in volunteer-based cohorts, participants to the NutriNet-Santé cohort were more often women, with health-conscious behaviours and higher socio-professional and educational levels as compared to the general French population[56]. This might limit the generalizability of the findings and may have resulted in 1) a lower cancer incidence compared to national estimates (age and sex standardized incidence rate per 100,000 persons per year: 786 cases in our cohort vs 972 cases in France [57]) and 2) an overall lower exposure to

Confidential: For Review Only

ultra-processed foods, with less contrast between extreme categories. These points rather tended to underestimate the strength of the associations. However, the possibility that selection bias may have led to an overestimation of some associations cannot be totally excluded. Second, some misclassification in the NOVA 'ultra-processed food' category cannot be ruled out. Third, despite a multi-source strategy for case ascertainment (combining validation of health events declared by participants, medico-administrative databases from the health insurance, and national death registry), exhaustiveness of cancer cases detection cannot be guaranteed. Furthermore, statistical power was limited for some cancer locations (such as colorectal cancer), which may have impaired our ability to detect hypothesized associations. Next, the length of follow-up was relatively limited in time, since the cohort was launched in 2009. Thus, it allowed us to study mostly mid-term associations between ultra-processed food consumption and cancer risk. As it is usually the case in nutritional epidemiology, the assumption is made that the measured exposure at baseline (especially since we averaged a two-year period of exposure) actually reflects more generally the usual eating habits of the individual during adulthood, including several years prior to his/her entry into the cohort. However, since some carcinogenic processes may take several decades, it will be important in the future to re-assess the associations between ultra-processed food and cancer risk in the cohort, in order to investigate longer-term effects. This will be one of the perspectives of the present work for the upcoming 5-10 years. Last, although a large range of confounding factors was included in the analyses, the hypothesis of residual confounding resulting from unmeasured behavioural factors and/or imprecision in the measure of included covariates cannot be entirely excluded due to the observational design of this study. For instance in breast cancer models, oral contraception was a binary variable, since the precise doses, type and duration of contraceptive use across reproductive life were not available. Randomized controlled trials have long been considered the only gold standard to eliminate confounding bias, however, they do not capture consumption as it is in daily life. Moreover, a trial would not be ethically feasible to investigate exposure for which a deleterious effect is suspected. Our large observational cohort was therefore particularly adapted to provide insights in this field.

Confidential. For Review Only

Conclusions and policy implications

To our knowledge, this study was the first to investigate and highlight an increase in overall – and more specifically, breast – cancer risk associated with ultra-processed food intake. These results should be confirmed by other large-scale population-based observational studies in different populations and settings. Further studies are also needed to better understand the relative impact of nutritional composition, food additives, contact materials, and neoformed contaminants in this relationship. Rapidly increasing consumption of ultra-processed foods may drive an increasing burden of cancer and other non-communicable diseases. Thus, policy actions targeting product reformulation, taxation and marketing restrictions on ultra-processed products and promotion of fresh or minimally processed foods may contribute to primary cancer prevention [6;9]. Several countries have already introduced this aspect in their official nutritional recommendations in the name of the precautionary principle [58;59].

Confidential: For Review Only

Acknowledgements

The authors sincerely thank all the volunteers of the NutriNet-Santé cohort. We also thank Véronique Gourlet, Lucien Martinez, Nathalie Arnault, Stephen Besseau, Laurent Bourhis, Yasmina Chelghoum, Than Duong Van, Younes Esseddik, Paul Flanzy, Julien Allègre, Mac Rakotondrazafy, Régis Gatibelza, Fabien Szabo, Roland Andrianasolo, Fatoumata Diallo, Ludivine Ursule, Cédric Agaesse, Claudia Chahine, Anne-Elise Dussoulier and Marion Genest for their technical contribution to the NutriNet-Santé study.

Competing interests statement

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Contributorship statement and guarantor

The authors' contributions were as follows – TF, CJ, EKG, CM, BA and MT: designed the research; SH, MT, CJ, EKG: conducted the research; TF: performed statistical analysis; MT and BS: supervised statistical analysis; TF and MT: wrote the paper; BS performed sensitivity analyses and was in charge of the revision of the paper; TF, BS, LS, MD, PF, PLM, EKG, BA, MB, SH, PG, CL, CM, CJ, and MT: contributed to the data interpretation and revised each draft for important intellectual content. All authors read and approved the final manuscript. MT had primary responsibility for the final content, she is the guarantor. None of the authors reported a conflict of interest related to the study. The funders had no role in the design, implementation, analysis, or interpretation of the data.

Transparency statement

Dr Touvier (the guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Confidential: For Review Only

Copyright / licence for publication statement

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

The default licence, a CC BY NC licence, is needed.

Data sharing statement

No additional data available.

Funding statement and statement of the independence of researchers from funders

The NutriNet-Santé study was supported by the following public institutions: Ministère de la Santé, Institut de Veille Sanitaire (InVS), Institut National de la Prévention et de l'Education pour la Santé (INPES), Région Ile-de-France (CORDDIM), Institut National de la Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche Agronomique (INRA), Conservatoire National des Arts et Métiers (CNAM) and Université Paris 13. Mélanie Deschasaux and Philippine Fassier were funded by a PhD grant from the Cancéropôle Ile de France / Région Ile de France (public funding). Bernard Srour was funded by the French National Cancer Institute (grant number INCa_8085).

Confidential: For Review Only

Researchers were independent from funders. Funders had no role in the study design, the collection, analysis, and interpretation of data, the writing of the report, and the decision to submit the article for publication.

Patient involvement statement

The research question developed in this article corresponds to a strong concern of the participants involved in the NutriNet-Santé cohort, and of the public in general. Participants to the study are thanked in the Acknowledgements section. The results of the present study will be disseminated to the NutriNet-Santé participants through the cohort website, public seminars and a press release.

Confidential: For Review Only

References

(1) Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015; 136(5):E359-E386.

(2) World Cancer Research Fund International/American Institute for Cancer Research. Cancer preventability estimates. Available at: wcrf.org/cancer-preventability-estimates. [Accessed on 03-07-2017], 2017.

(3) Latino-Martel P, Cottet V, Druesne-Pecollo N, Pierre FH, Touillaud M, Touvier M et al. Alcoholic beverages, obesity, physical activity and other nutritional factors, and cancer risk: A review of the evidence. *Crit Rev Oncol Hematol* 2016; 99:308-323.

(4) ANSES (French Agency for Food,Environmental and Occupational Health & Safety). Individual national study for dietary intakes 3 (INCA 3) [Étude individuelle nationale des consommations alimentaires 3 (INCA 3)] (in French), 2017.

(5) Monteiro CA, Moubarac JC, Cannon G, Ng SW, Popkin B. Ultra-processed products are becoming dominant in the global food system. *Obes Rev* 2013; 14 Suppl 2:21-28.

(6) Moodie R, Stuckler D, Monteiro C, Sheron N, Neal B, Thamarangsi T et al. Profits and pandemics: prevention of harmful effects of tobacco, alcohol, and ultra-processed food and drink industries. *Lancet* 2013; 381(9867):670-679.

(7) Moubarac JC, Batal M, Martins AP, Claro R, Levy RB, Cannon G et al. Processed and ultra-processed food products: consumption trends in Canada from 1938 to 2011. *Can J Diet Pract Res* 2014; 75(1):15-21.

(8) Venn D, Banwell C, Dixon J. Australia's evolving food practices: a risky mix of continuity and change. *Public Health Nutr* 2016;1-10.

(9) 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;1-13.

(10) Luiten CM, Steenhuis IH, Eyles H, Ni MC, Waterlander WE. Ultra-processed foods have the worst nutrient profile, yet they are the most available packaged products in a sample of New Zealand supermarkets. *Public Health Nutr* 2016; 19(3):539.

(11) Adams J, White M. Characterisation of UK diets according to degree of food processing and associations with socio-demographics and obesity: cross-sectional analysis of UK National Diet and Nutrition Survey (2008-12). *Int J Behav Nutr Phys Act* 2015; 12:160.

(12) Cediel G, Reyes M, da Costa Louzada ML, Martinez SE, Monteiro CA, Corvalan C et al. Ultra-processed foods and added sugars in the Chilean diet (2010). *Public Health Nutr* 2017;1-9.

(13) Costa Louzada ML, Martins AP, Canella DS, Baraldi LG, Levy RB, Claro RM et al. Ultra-processed foods and the nutritional dietary profile in Brazil. *Rev Saude Publica* 2015; 49:38.

(14) Martinez SE, Baraldi LG, Louzada ML, Moubarac JC, Mozaffarian D, Monteiro CA. Ultra-processed foods and added sugars in the US diet: evidence from a nationally representative cross-sectional study. *BMJ Open* 2016; 6(3):e009892.

Confidential: For Review Only

(15) Moubarac JC, Martins AP, Claro RM, Levy RB, Cannon G, Monteiro CA. Consumption of ultra-processed foods and likely impact on human health. Evidence from Canada. *Public Health Nutr* 2013; 16(12):2240-2248.

(16) Moubarac JC, Batal M, Louzada ML, Martinez SE, Monteiro CA. Consumption of ultra-processed foods predicts diet quality in Canada. *Appetite* 2017; 108:512-520.

(17) Poti JM, Mendez MA, Ng SW, Popkin BM. Is the degree of food processing and convenience linked with the nutritional quality of foods purchased by US households? *Am J Clin Nutr* 2015; 101(6):1251-1262.

(18) Slimani N, Deharveng G, Southgate DA, Biessy C, Chajes V, van Bakel MM et al. Contribution of highly industrially processed foods to the nutrient intakes and patterns of middle-aged populations in the European Prospective Investigation into Cancer and Nutrition study. *Eur J Clin Nutr* 2009; 63 Suppl 4:S206-S225.

(19) Louzada ML, Martins AP, Canella DS, Baraldi LG, Levy RB, Claro RM et al. Impact of ultra-processed foods on micronutrient content in the Brazilian diet. *Rev Saude Publica* 2015; 49:45.

(20) EFSA (European Food Safety Agency). EFSA Panel on Contaminants in the Food Chain (CONTAM). «Scientific Opinion on Acrylamide in Food». *EFSA Journal* 13, n° 6 (1 juin 2015): doi:10.2903/j.efsa.2015.4104, 2015.

(21) Muncke J. Endocrine disrupting chemicals and other substances of concern in food contact materials: an updated review of exposure, effect and risk assessment. *J Steroid Biochem Mol Biol* 2011; 127(1-2):118-127.

(22) European Union (EU) - The database of authorized food additives, Available online on https://ec.europa.eu/food/safety/food_improvement_agents/additives/database_en. [Accessed on 26 October 2017], 2008.

(23) Bouvard V, Loomis D, Guyton KZ, Grosse Y, Ghissassi FE, Benbrahim-Tallaa L et al. Carcinogenicity of consumption of red and processed meat. *Lancet Oncol* 2015; 16(16):1599-1600.

(24) Carbon black, titanium dioxide, and talc. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. IARC Monogr Eval Carcinog Risks Hum. 2010; 93:1-413.

(25) Canella DS, Levy RB, Martins AP, Claro RM, Moubarac JC, Baraldi LG et al. Ultra-processed food products and obesity in Brazilian households (2008-2009). *PLoS One* 2014; 9(3):e92752.

(26) Juul F, Hemmingsson E. Trends in consumption of ultra-processed foods and obesity in Sweden between 1960 and 2010. *Public Health Nutr* 2015; 18(17):3096-3107.

(27) Louzada ML, Baraldi LG, Steele EM, Martins AP, Canella DS, Moubarac JC et al. Consumption of ultra-processed foods and obesity in Brazilian adolescents and adults. *Prev Med* 2015; 81:9-15.

(28) Rauber F, Campagnolo PD, Hoffman DJ, Vitolo MR. Consumption of ultra-processed food products and its effects on children's lipid profiles: a longitudinal study. *Nutr Metab Cardiovasc Dis* 2015; 25(1):116-122.

(29) Mendonca RD, Pimenta AM, Gea A, Fuente-Arrillaga C, Martinez-Gonzalez MA, Lopes AC et al. Ultraprocessed food consumption and risk of overweight and obesity: the University of Navarra Follow-Up (SUN) cohort study. *Am J Clin Nutr* 2016; 104(5):1433-1440.

Confidential: For Review Only

(30) Mendonca RD, Lopes AC, Pimenta AM, Gea A, Martinez-Gonzalez MA, Bes-Rastrollo M. Ultra-Processed Food Consumption and the Incidence of Hypertension in a Mediterranean Cohort: The Seguimiento Universidad de Navarra Project. *Am J Hypertens* 2017; 30(4):358-366.

(31) Hercberg S, Castetbon K, Czernichow S, Malon A, Mejean C, Kesse E et al. The Nutrinet-Sante 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.

(32) Vergnaud AC, Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A et al. Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Sante study. *Int J Public Health* 2011; 56(4):407-417.

(33) Lassale C, Peneau S, Touvier M, Julia C, Galan P, Hercberg S et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Sante study. *J Med Internet Res* 2013; 15(8):e152.

(34) Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur J Epidemiol* 2010; 25(5):287-296.

(35) Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003; 35(8):1381-1395.

(36) Touvier M, Kesse-Guyot E, Mejean C, Pollet C, Malon A, Castetbon K et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* 2011; 105(7):1055-1064.

(37) Lassale C, Castetbon K, Laporte F, Camilleri GM, Deschamps V, Vernay M et al. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br J Nutr* 2015; 113(6):953-962.

(38) Lassale C, Castetbon K, Laporte F, Deschamps V, Vernay M, Camilleri GM et al. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. *J Acad Nutr Diet* 2016; 116(3):427-438.

(39) Le Moulenc N, Deheeger M, Preziozi P, Montero P, Valeix P, Rolland-Cachera M et al. Validation du manuel photo utilisé pour l'enquête alimentaire de l'étude SU.VI.MAX. [Validation of the food portion size booklet used in the SU.VI.MAX study] (in French), *Cah Nutr Diet* 1996;31:158-64.

(40) 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. *Int J Obes Relat Metab Disord* 2000; 24(9):1119-1130.

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

(42) Monteiro CA, Cannon G, Levy RB, Moubarac JC, Jaime PC, Martins AP et al. NOVA. The star shines bright. *World Nutrition* 2016; 7(1-3):28-38.

(43) 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. *Curr Obes Rep* 2014; 3(2):256-272.

https://mc.manuscriptcentral.com/bmj

Confidential for Review Only

(44) Julia C, Martinez L, Alles B, Touvier M, Hercberg S, Mejean C et al. Contribution of ultra-processed foods in the diet of adults from the French NutriNet-Sante study. *Public Health Nutr* 2017;1-11.

(45) Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol* 2012; 176(3):190-195.

(46) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; 338:b2393.

(47) Fardet A. Minimally processed foods are more satiating and less hyperglycemic than ultra-processed foods: a preliminary study with 98 ready-to-eat foods. *Food Funct* 2016; 7(5):2338-2346.

(48) World Health Organization (WHO) - Fact sheet about Food additives, added on July 2017. Available at http://www.who.int/mediacentre/factsheets/food-additives/en/. [Accessed on 26 october 2017], 2017.

(49) US Food & Drug Administration FDA Food Additive Status List. Available at: https://www.fda.gov/food/ingredientspackaginglabeling/foodadditivesingredients/ucm091048.htm [Accessed 03 July 2017], 2017.

(50) Chang X, Zhang Y, Tang M, Wang B. Health effects of exposure to nano-TiO2: a meta-analysis of experimental studies. *Nanoscale Res Lett* 2013; 8(1):51.

(51) EFSA ANS Panel. (EFSA Panel on Food Additives and Nutrient Sources added to Food) - Scientific Opinion on the re-evaluation of aspartame (E 951) as a food additive. EFSA Journal 2013;11(12):3496, 263 pp. doi:10.2903/j.efsa.2013.3496.

(52) Santarelli RL, Vendeuvre JL, Naud N, Tache S, Gueraud F, Viau M et al. Meat processing and colon carcinogenesis: cooked, nitrite-treated, and oxidized high-heme cured meat promotes mucin-depleted foci in rats. *Cancer Prev Res (Phila)* 2010; 3(7):852-864.

(53) Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O et al. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* 2014; 514(7521):181-186.

(54) Virk-Baker MK, Nagy TR, Barnes S, Groopman J. Dietary acrylamide and human cancer: a systematic review of literature. *Nutr Cancer* 2014; 66(5):774-790.

(55) European CHemical Agency (ECHA). Member State Committee support document for identification of 4,4'-isopropylidenediphenol (bisphenol a) as a substance of very high concern because of its toxic for reproduction (Article 57 c) properties. Adopted on 2 December 2016.

(56) Andreeva VA, Salanave B, Castetbon K, Deschamps V, Vernay M, Kesse-Guyot E et al. Comparison of the sociodemographic characteristics of the large NutriNet-Sante e-cohort with French Census data: the issue of volunteer bias revisited. *J Epidemiol Community Health* 2015; 69(9):893-898.

(57) INCa. The French National Institute of Cancer INCA - Cancer in France, 2014 Edition (In French: Les Cancers en France, Editions 2014); Available at: http://www.unicancer.fr/sites/default/files/Les%20cancers%20en%20France%20-%20Edition%202014%20-%20V5.pdf. [Accessed on 26 October 2017], 2017.

(58) Ministry of Health of Brazil. Dietary Guidelines for the Brazilian population, 2014.

Confidential: For Review Only

(59) Haut Conseil de la Santé Publique. Avis relatif à la révision des repères alimentaires pour les adultes du futur Programme National Nutrition Santé 2017-2021, 2017.

Confidential: For Review Only

Figure Legends

Figure 1:

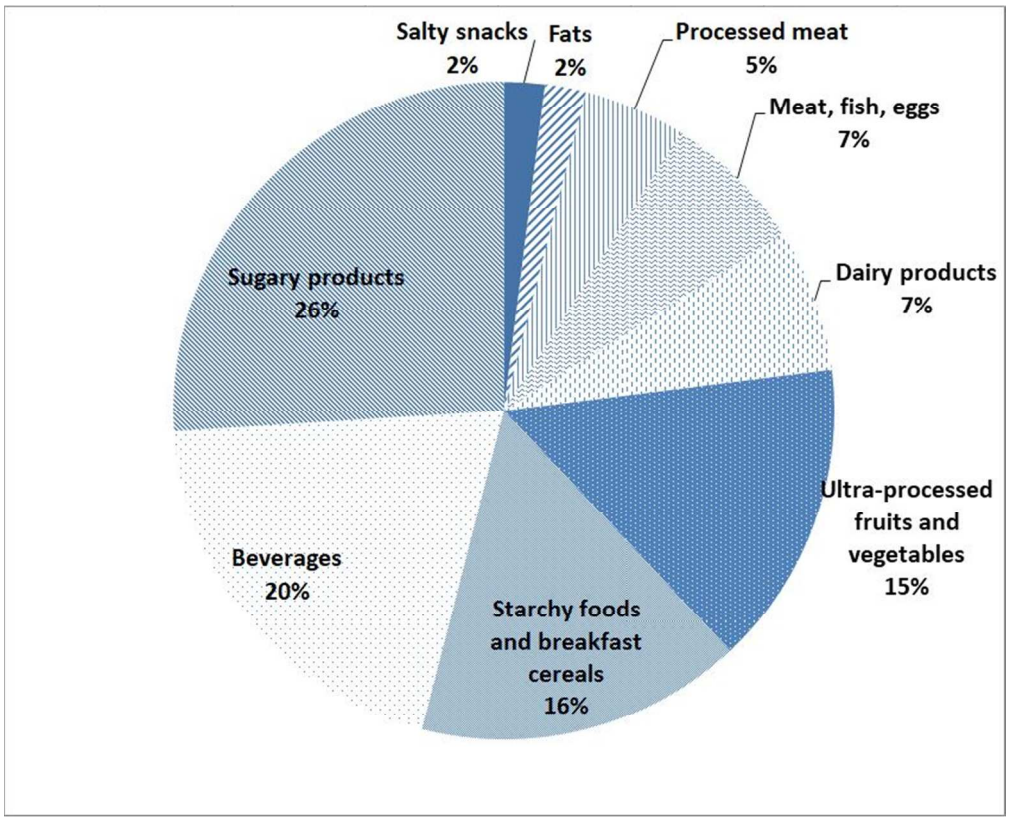
Title: Relative contribution of each food group to ultra-processed consumption in the diet

Figure 2:

Title: Cumulative cancer incidence (overall cancer risk) according to quartiles of ultra-processed food intake

Legend: Q=quartile (1 to 4) of the proportion of ultra-processed food in the diet

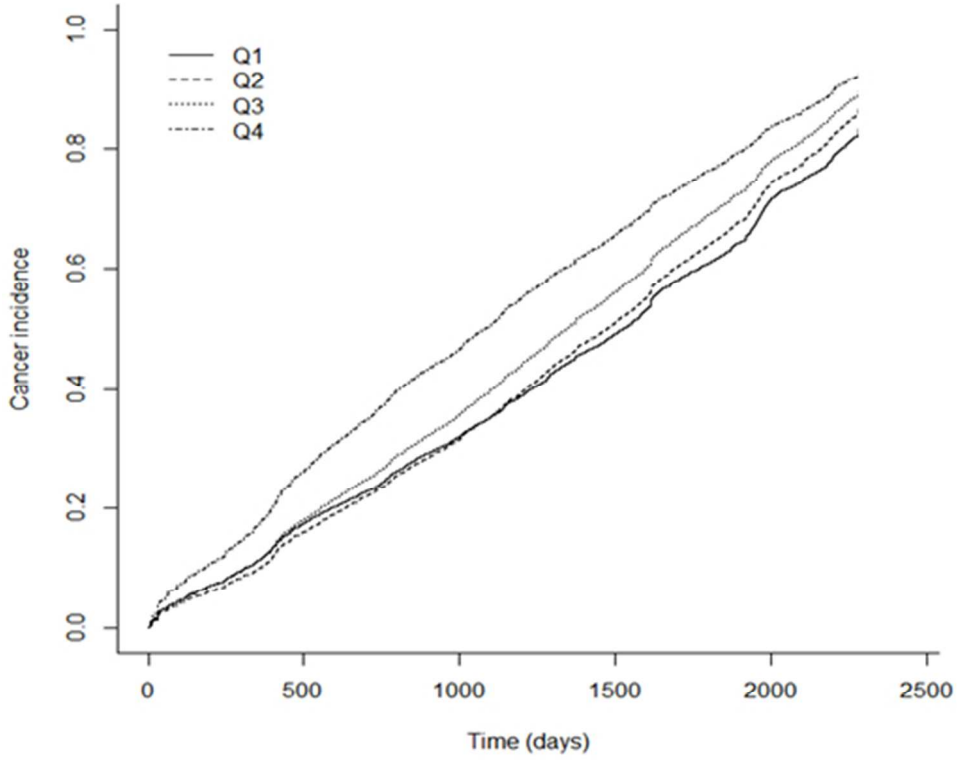
Confidential: For Review Only



Relative contribution of each food group to ultra-processed consumption in the diet

223x181mm (96 x 96 DPI)

For Peer Review Only



Cumulative cancer incidence (overall cancer risk) according to quartiles of ultra-processed food intake.
Legend: Q=quartile (1 to 4) of the proportion of ultra-processed food in the diet

141x115mm (96 x 96 DPI)

For Peer Review Only

Supplemental material

Appendix 1: 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 [1-3]. 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. 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 ultra-processed food according to the NOVA classification:

Carbonated drinks; sweet or savoury packaged snacks; ice-cream, chocolate, candies (confectionery); mass-produced packaged breads and buns; margarines and spreads; industrial cookies (biscuits), pastries, cakes, and cake mixes; breakfast 'cereals', 'cereal' and 'energy' bars; 'energy' drinks; flavoured milk drinks; cocoa drinks; sweet desserts made from fruit with added sugars, artificial flavours and texturizing agents; cooked seasoned vegetables with ready-made sauces; 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; poultry and fish 'nuggets' and 'sticks', sausages, burgers, hot dogs, and other reconstituted meat products, and powdered and packaged 'instant' soups, noodles and desserts.

For instance, fruit compotes with only added sugar are considered as "processed foods", while flavoured fruit desserts with added sugar, texturizing agents and colorants are considered as "ultra-processed foods".

Confidential: For Review Only

Regarding meats, 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".

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)"*.

Confidential: For Review Only

Appendix 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.

Confidential. For Review Only

	Factor loadings	
	Healthy Pattern	Western Pattern
Alcoholic drinks	-.099552	0.284771
Breakfast cereals	0.079447	-.181769
Cakes and biscuits	-.197629	0.003444
Dairy products	0.066066	-.013702
Eggs	0.078582	0.043744
Fats and sauces	0.012600	0.544911
Fish and seafood	0.204373	0.100759
Fruit	0.354075	0.052298
Meat	-.188274	0.318483
Pasta and rice	-.212857	0.341941
Potatoes and tubers	-.029615	0.402694
Poultry	-.030137	0.064064
Processed meat	-.228028	0.207877
Pulses	0.192815	0.026104
Soups and broths	0.264233	0.227787
Sugar andconfectionery	-.088870	0.120660
Sweetened soft drinks	-.288870	-.007506
Unsweetened soft drinks	0.258563	0.152704
Vegetables	0.471255	0.231818
Whole grains	0.380881	-.043132

Confidential: For Review Only

Appendix 3: Methodology and results of the mediation analysis

Mediation analyses were carried out according to the method proposed by Lange et al.[4] in order to evaluate the direct and indirect "effects" in the relationship between the exposure and the outcome, through nutritional mediators. Under the assumption of a causal relationship between quartiles of the proportion of ultra-processed food in the diet (=Exposure, quoted "A") and cancer risk (=Outcome, quoted "Y"), the aim was to estimate how much of this effect was mediated through various factors reflecting the nutritional quality of the diet. The latter factors (dietary intakes of sodium, total lipid, fatty acids, and carbohydrates, and Western-type dietary pattern) were considered as potential Mediators (quoted "M") in each model. The following covariates were considered as potential confounders (quoted "C"): age, sex, BMI, height, physical activity, smoking status, number of 24h-dietary records, alcohol intake, energy intake, family history of cancer, and educational level. To evaluate the direct effect and the indirect effect mediated by each nutritional factor, we applied a mediation analysis in the counterfactual framework. The mediation analyses were implemented according to the following steps for a categorical exposure:

(1) Construction of a new data set by repeating each observation in the original data set. This new variable A* corresponds to the value of the exposure relative to the indirect path. Each observation was repeated four times such that A* got to take all possible values of exposure (quartiles of ultra-processed).

(2) Fitting of a multinomial logistic regression applied to the new data set to estimate the association between ultra-processed food and cancer, conditioned on baseline confounders, and computing predicted values, first using the original variable A and then the new variable A*.

(3) Weighting (W) each observation calculated according to the following formula through applying the fitted models from steps 2 et 3 to the new dataset:

$$W_i = \frac{1}{P(A = A_i \,|C = C_i)} \frac{P(M = M_i|A = A_i^*, C = C_i)}{P(M = M_i|A = A_i, C = C_i)}$$

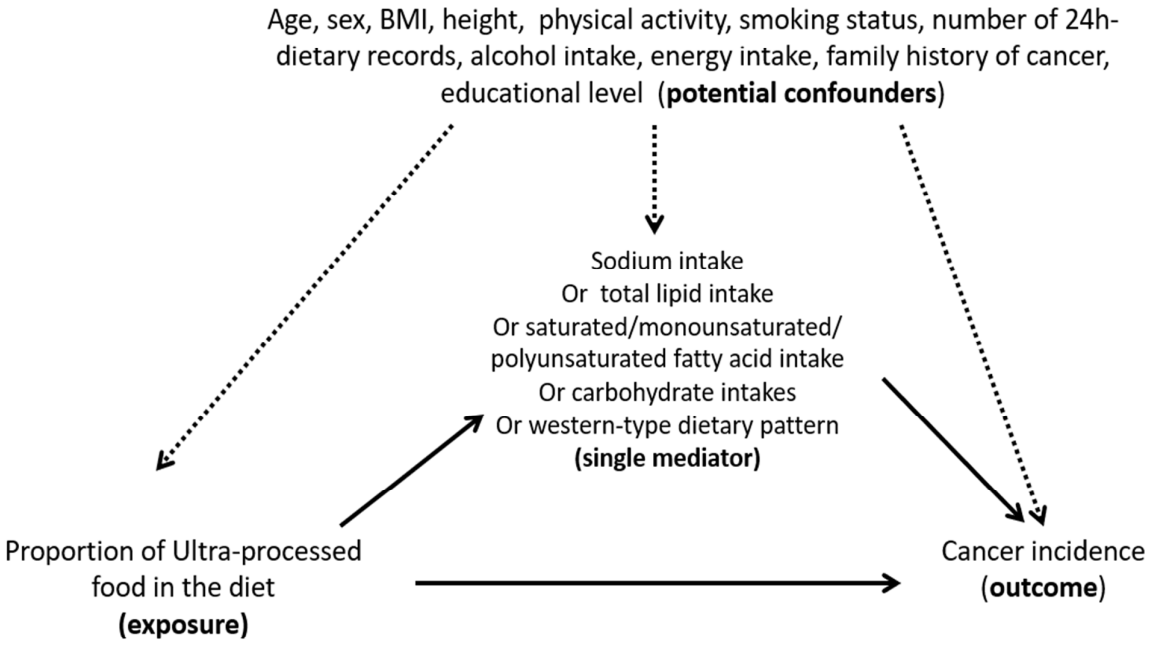
with A, the exposure, M, the mediator, C, the set of baseline confounders

(4) Fitting of a weighted Cox Marginal Structural Model (MSM) for direct and indirect effects controlling for baseline confounders, as the outcome corresponds to a survival time. The "Covsandwich" statement in SAS software allows getting robust standard errors.

(5) To evaluate how much of the total effect was due to the mediator effect, we calculated the 'proportion explained' by each single mediator as ($HR_{\text{total effect}}$ - $HR_{\text{direct effect}}$) / ($HR_{\text{total effect}} - 1$) where $HR_{\text{total effect}}$ and $HR_{\text{direct effect}}$ were respectively, the Hazard Ratios for total effect and for direct effect.

The figure below shows a conceptual model of the association between the proportion of ultra-processed foods in the diet and cancer risk, taking into account nutritional factors as potential single mediators:

Confidential: For Review Only



Conceptual model of the association between ultra-processed food consumption and change in cancer risk taking into account nutritional factors as potential single mediator

For Review Only

The table below shows the results of mediation analyses testing for a potential mediation by total lipid, carbohydrate, sodium, SFA, PUFA and MUFA intakes, and the Western dietary pattern of the association between ultra-processed food intake and cancer risk.

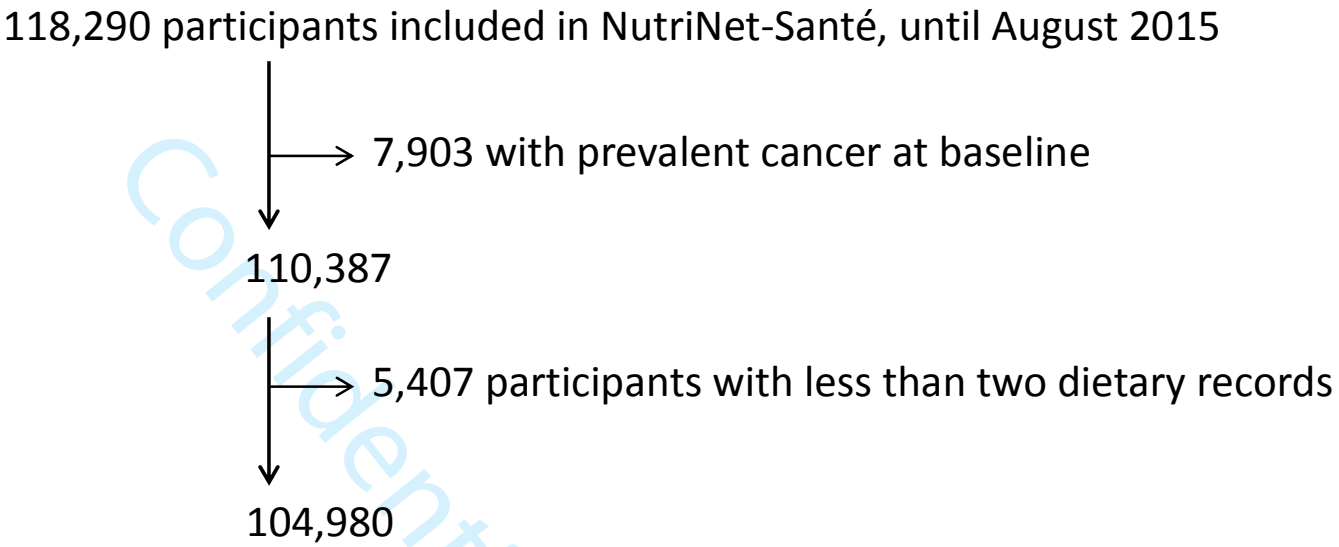
Table 1 – Hazard Ratios of direct, indirect and total effects and proportion of total effects mediated by several nutritional factors in the prospective associations between ultra-processed food and overall cancer risk, N=104980, NutriNet-Santé cohort, France, 2009-2017

	Tested nutritional mediators of the association between ultra-processed foods and overall cancer risk													
	Total lipids		Sodium		Carbohydrates		Western pattern		SFAs		PUFAs		MUFAs	
Effect	HR	p-value	HR	p-value	HR	p-value	HR	p-value	HR	p-value	HR	p-value	HR	p-value
Indirect effect	1.000	0.799	1.003	0.889	1.000	0.900	1.005	0.910	1.000	0.900	1.000	0.900	1.000	0.900
Direct effect	1.302	<0.0001	1.263	<0.0001	1.217	<0.0001	1.317	<0.0001	1.166	0.001	1.319	<0.0001	1.328	<0.0001
Total effect	1.302		1.267		1.217		1.324		1.166		1.319		1.328	
Proportion of the total effect mediated by the nutritional factor	0.00%		1.42%		0.00%		2.04%		0.00%		0.00%		0.00%	

SFAs: saturated fatty acids, PUFAs: poly-unsaturated fatty acids, MUFAs: mono-unsaturated fatty acids, HR: Hazard Ratio

Confidential: For Review Only

Appendix 4: Flow chart

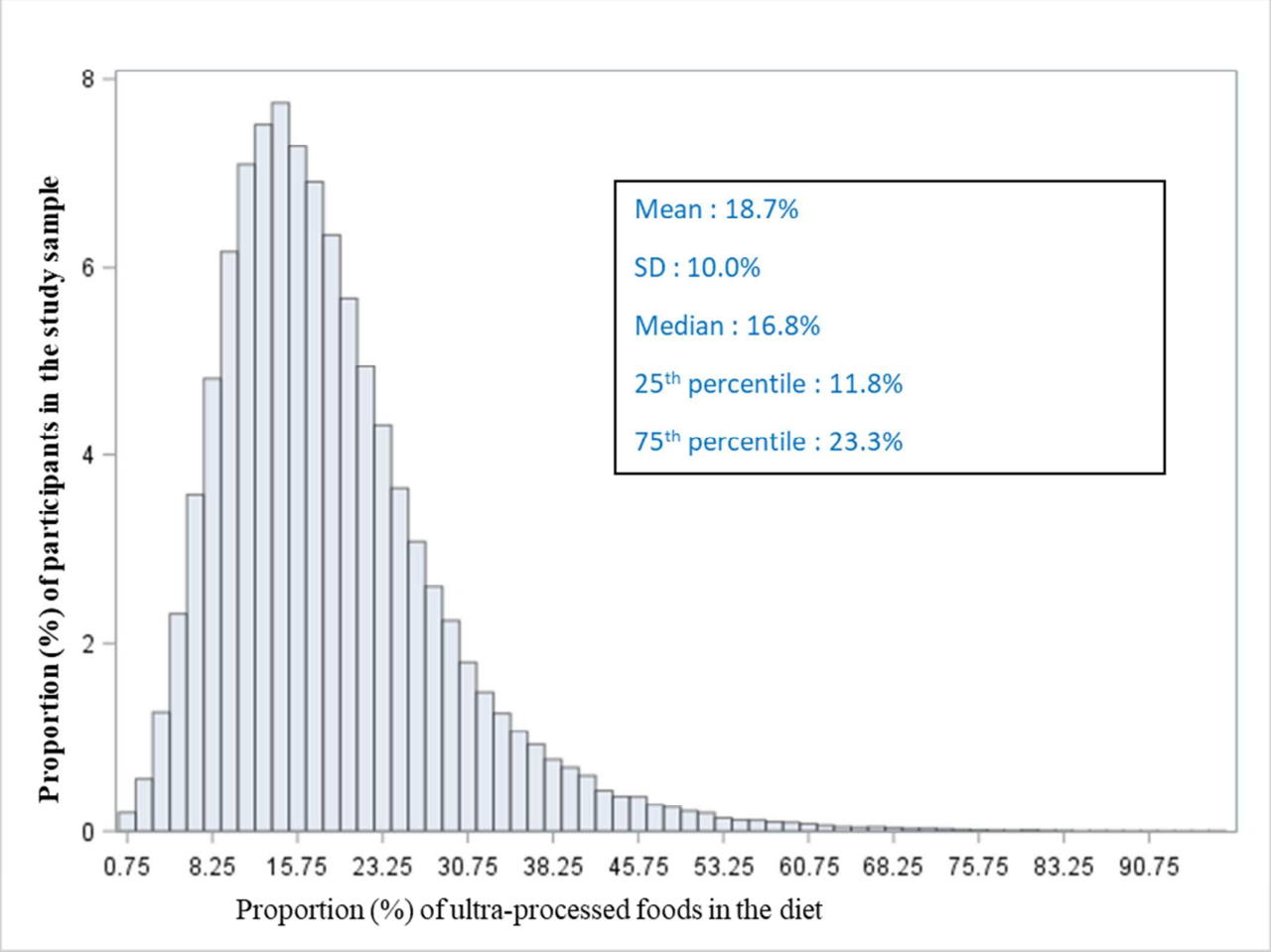


104,980 participants included:

22821 (21.7%) men and 82159 (78.3%) women

Confidential. For Review Only

Appendix 5: Distribution of the main exposure (proportion of ultra-processed food in the diet) in the study sample (N=104 980), NutriNet-Santé, France



For Peer Review Only

Appendix 6: Associations between the quantity (g/d) of each ultra-processed food group and overall and breast cancer risks, from multivariable Cox proportional hazard models, NutriNet-Santé cohort, France, 2009 – 2017 (n=104,980)

	Continuous		
	HR[a,b]	95%CI	P-value
All cancers			
N for cases/non cases	2228/102752		
Starchy foods	1.01	(0.99-1.02)	0.4
Fruits and vegetables	1.00	(0.99-1.01)	0.2
Dairy products	1.01	(1.00-1.02)	0.05
Fats	1.07	(1.03-1.12)	0.002
Salty snacks	0.98	(0.93-1.02)	0.3
Meat, fish, eggs	1.01	(0.99-1.03)	0.4
Processed meat	0.99	(0.97-1.01)	0.5
Sugary products	1.01	(1.00-1.02)	0.03
Beverages	1.00	(1.00-1.01)	0.005
Breast Cancer			
N for cases/non cases	739/81420		
Starchy foods	1.00	(0.98-1.03)	0.7
Fruits and vegetables	1.01	(0.99-1.02)	0.3
Dairy products	1.01	(0.99-1.02)	0.3
Fats	1.06	(0.97-1.14)	0.2
Salty snacks	1.02	(0.95-1.10)	0.6
Meat, fish, eggs	1.01	(0.97-1.04)	0.8
Processed meat	0.98	(0.94-1.02)	0.4
Sugary products	1.02	(1.01-1.03)	0.006
Beverages	1.00	(0.99-1.01)	0.2

CI, confidence interval, HR, Hazard ratio

[a] adjusted for age (timescale), sex, energy intake without alcohol, number of 24h-dietary records, smoking status, educational level, physical activity, height, BMI, alcohol intake, and family history of cancers. Breast cancer models were additionally adjusted for menopausal status, hormonal treatment for menopause, oral contraception and number of children.

[b] HR for an increase of 10g of the quantity (in g/d) of each ultra-processed food group

Confidential: For Review Only

References

(1) Monteiro CA, Cannon G, Levy RB, Moubarac JC, Jaime PC, Martins AP et al. NOVA. The star shines bright. *World Nutrition* 2016; 7(1-3):28-38.

(2) 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;1-13.

(3) 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. *Curr Obes Rep* 2014; 3(2):256-272.

(4) Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol* 2012; 176(3):190-195.

Confidential: For Review Only

Consumption of ultra-processed foods and cancer risk: results from the NutriNet-Santé prospective cohort

Thibault Fiolet[1*], Bernard Srour[1*], Laury Sellem[1], Emmanuelle Kesse-Guyot[1], Benjamin Allès[1], Caroline Méjean[2], Mélanie Deschasaux[1], Philippine Fassier[1], Paule Latino-Martel[1], Marie Beslay[1], Serge Hercberg[1, 4], Céline Lavalette[1], Carlos A. Monteiro[3], Chantal Julia[1, 4], Mathilde Touvier[1]

*Equally contributed

Running Head: Ultra-processed foods and cancer risk

Affiliations:

[1] Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Inserm U1153, Inra U1125, Cnam, Paris 13 University, Nutritional Epidemiology Research Team (EREN), Bobigny, France

[2] INRA, UMR 1110 MOISA, 34000 Montpellier, France

[3] Department of Nutrition, School of Public Health, University of São Paulo, Av. Dr Arnaldo 715,São Paulo 01246-904, Brazil

[4]Public Health Department, Avicenne Hospital, AP-HP, Bobigny, France

Corresponding author:

Bernard Srour

Equipe de Recherche en Epidémiologie Nutritionnelle (EREN)

Centre de Recherche en Epidémiologie et Statistiques Sorbonne Paris Cité, Inserm (U1153), Inra(U1125), Cnam, Université Paris 13

SMBH Paris 13, 74 rue Marcel Cachin F-93017 Bobigny Cedex. France

Tel : +33 1 48 38 89 68

E-mail : b.srour@eren.smbh.univ-paris13.fr

Word count: 5,364 ; **Tables**: 2 ; **Figures**: 2 ; **Supplemental material** (**Appendixes):** 6

Keywords: ultra-processed food; food processing; cancer; prospective cohort; NOVA classification

Confidential: For Review Only

BOX:

WHAT IS ALREADY KNOWN ON THIS TOPIC

Dietary habits are shifting in many countries through an upsurge in the consumption of ultra-processed foods, which are often characterized by a lower nutritional quality but also the presence of food additives, of substances from materials and packaging in contact with food, and of neoformed compounds during production, processing, and storage.

The few studies performed observed that ultra-processed food intake was associated with a higher incidence of dyslipidaemia in Brazilian children, and higher risks of overweight, obesity and hypertension in a cohort of Spanish University students.

Although epidemiological data regarding their relevance to cancer risk are lacking, mechanistic studies suggest potential carcinogenic effects of several components commonly found in ultra-processed foods.

WHAT THIS STUDY ADDS

For the first time, this study assessed the associations between ultra-processed food consumption and cancer risk in a large prospective cohort (n=104,980).

A 10% increase in the proportion of ultra-processed foods in the diet was associated with a >10% significant increase in overall and breast cancer risks.

If confirmed in other populations and settings in the future, these results suggest that the rapidly increasing consumption of ultra-processed foods may drive an increasing burden of cancer in the next decades. Thus, individual recommendations to improve dietary choices, as well as policy actions targeting product reformulation, taxation and marketing restrictions on ultra-processed products and promotion of fresh or minimally processed foods may contribute to primary cancer prevention.

Confidential: For Review Only

STRUCTURED ABSTRACT:

OBJECTIVE

To assess the prospective associations between ultra-processed food consumption and cancer risk.

DESIGN

Population based cohort study

SETTING AND PARTICIPANTS

In all, 104,980 participants aged ≥18y (median age: 40.4y) from the French NutriNet-Santé cohort (2009-2017) were included. Dietary intakes were collected using repeated 24h-dietary records, designed to register participants' usual consumption for 3300 different food items. These were categorized according to their degree of processing by the NOVA classification.

MAIN OUTCOME MEASURES

Associations between ultra-processed food intake and overall, breast, prostate and colorectal cancer risk were assessed by multivariable Cox Proportional Hazard models adjusted for known risk factors.

RESULTS

Ultra-processed food intake was associated with higher overall cancer risk (n=2,228 cases, $HR_{\text{for a 10\% increment in the proportion of ultra-processed food in the diet}}$=1.12 (1.06-1.18), P-trend<.0001) and breast cancer risk (n=739 cases, HR= 1.11 (1.02-1.22), P-trend=0.02). These results remained statistically significant after adjustment for several markers of the nutritional quality of the diet (lipid, sodium and carbohydrate intakes and/or a Western pattern derived by principal component analysis).

CONCLUSIONS

Confidential: For Review Only

In this large prospective study, a 10% increase in the proportion of ultra-processed foods in the diet was associated with a >10% significant increase in overall and breast cancer risks. Further studies are needed to better understand the relative impact of the various dimensions of processing (nutritional composition, food additives, contact materials, and neoformed contaminants) in these relationships.

Confidential: For Review Only

INTRODUCTION

Cancer represents a major worldwide burden with 14.1 million new cases diagnosed in 2012 [1]. According to the World Cancer Research Fund / American Institute for Cancer Research (WCRF/AICR), about one third of the most common neoplasms could be avoided by changing lifestyle and dietary habits in developed countries [2]. Therefore, reaching a balanced and diversified diet (along with tobacco avoidance and alcohol reduction) should be considered as one of the most important modifiable risk factors in the primary prevention of cancer [3].

At the same time, during the last decades, in many countries, diets have shifted towards a dramatic increase in ultra-processed foods consumption [4-8]. After undergoing multiple physical, biological and/or chemical processes, these food products are conceived to be microbiologically safe, convenient, highly palatable and affordable [9;10]. Several surveys (in Europe, the USA, Canada, New Zealand and Brazil) assessing individual food intake, household food expenses or supermarket sales, suggested that ultra-processed food products contribute to between 25 and 50% of total daily energy intake [10-18].

This dietary trend may be concerning and deserves investigation. Indeed, several characteristics of ultra-processed foods may be involved in disease – in particular cancer – aetiology. First, ultra-processed foods often have a higher content in total fat, saturated fat, added sugar and salt, along with a lower fibre and vitamin density [10-17;19]. Beyond nutritional composition, neoformed contaminants, some of which having carcinogenic properties (such as acrylamide, heterocyclic amines, polycyclic aromatic hydrocarbons, etc.) are present in heat-treated processed food products due to the Maillard reaction [20]. Next, the packaging of ultra-processed foods may contain some contact materials for which carcinogenic and endocrine disruptor properties have been postulated such as Bisphenol A [21]. Finally, ultra-processed foods contain authorized[22] but controversial food additives such as sodium nitrite in processed meat or titanium dioxide (TiO_2, white food pigment), for which carcinogenicity has been suggested in animal or cellular models [23;24].

Studying potential health impacts of ultra-processed foods is a very recent field of research, facilitated by the development of the NOVA classification of products according to their degree of food processing [9]. Nonetheless, epidemiological evidence linking ultra-processed food intake to disease risk is still very

Confidential. For Review Only

scarce[25-27] and mostly based on cross-sectional and ecological studies. The few studies performed observed that ultra-processed food intake was associated with a higher incidence of dyslipidaemia in Brazilian children [28], and higher risks of overweight, obesity [29] and hypertension [30] in a prospective cohort of Spanish University students.

To our knowledge, the present prospective study was the first to evaluate the association between the consumption of ultra-processed food products and the incidence of cancer, based on a large cohort study with detailed and up-to-date dietary intake assessment.

MATERIAL AND METHODS

Study population

The NutriNet-Santé study is an ongoing web-based cohort launched in 2009 in France with the objective to study the associations between nutrition and health as well as the determinants of dietary behaviours and nutritional status. This cohort has been previously described in detail [31]. Briefly, participants aged over 18 years with access to the Internet are continuously recruited since May 2009 among the general population by means of vast multimedia campaigns. All questionnaires are completed online using a dedicated website (www.etude-nutrinet-sante.fr). Participants are followed using an online platform connected to their email address. They have the possibility to change their email address, phone number or postal address at any moment on the NutriNet-Santé website. Newsletters and alerts about new questionnaires are sent via email. In case of an "undelivered email to recipient" problem, participants are then contacted by telephone and then by regular mail. The NutriNet-Santé study is conducted according to the Declaration of Helsinki guidelines and was approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm n°0000388FWA00005831) and the "Commission Nationale de l'Informatique et des Libertés" (CNIL n°908450/n°909216). It is registered at clinicaltrials.gov as NCT03335644. Electronic informed consent is obtained from each participant.

Data collection

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

At inclusion, participants completed a set of five questionnaires related to socio-demographic and lifestyle characteristics [32] (e.g. date of birth, sex, occupation, educational level, smoking status, number of children), anthropometry [33;34] (e.g. height, weight), dietary intakes (see below), physical activity (validated 7-day International Physical Activity Questionnaire [IPAQ]) [35], and health status (e.g. personal and family history of diseases, medication use including use of hormonal treatment for menopause, oral contraceptive, and menopausal status).

Participants were invited to complete a series of three non-consecutive validated web-based 24h-dietary records every 6 months (to vary the season of completion), randomly assigned over a 2-week period (2 weekdays and 1 weekend day) [36-38]. To be included in the nutrition component of the NutriNet-Santé cohort, only two dietary records were mandatory. Subjects were not excluded if they did not complete all optional questionnaires. Mean dietary intakes from all the 24h-dietary records available during the first two years of each participant's follow-up were averaged and considered as baseline usual dietary intakes in this prospective analysis. The NutriNet-Santé web-based self-administered 24h dietary records have been tested and validated against an interview by a trained dietitian [36], and against blood and urinary biomarkers [37]. Participants used the dedicated web interface to declare all food and beverages consumed during a 24h-period for each of the three main meals (breakfast, lunch, dinner) and any other eating occasion. Portion sizes were estimated using previously validated photographs or usual containers [39]. 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 [40]. Mean daily alcohol, micro- and macro-nutrient and energy intake were calculated using the NutriNet-Santé food composition database, which contains more than 3,300 different items [41]. Amounts consumed from composite dishes were estimated using French recipes validated by nutrition professionals. Sodium intake was assessed via a specific module included in the 24h records, taking into account native sodium in foods, salt added during the cooking, and salt added in the plate. It has been validated against sodium urinary excretion biomarkers [37].

Confidential: For Review Only

Degree of food processing

All food and beverage items of the NutriNet-Santé composition table were categorized into one of the four food groups in NOVA, a food classification system based on the extent and purpose of industrial food processing[9;42;43]. This study primarily focused on the "ultra-processed foods" NOVA group. This group includes mass-produced packaged breads and buns, sweet or savoury packaged snacks, industrialized confectionery and desserts, sodas and sweetened beverages, meat balls, poultry and fish nuggets and other reconstituted meat products transformed with addition of preservatives other than salt (e.g nitrites), instant noodles and soups, frozen or shelf-stable ready meals, and other food products made mostly or entirely from sugar, oils, and fats and other substances not commonly used in culinary preparations such as hydrogenated oils, modified starches, and protein isolates. 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 or to disguise undesirable qualities of the final product. The ultra-processed food 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[44], home-made and artisanal food preparations were identified and decomposed using standardized recipes, and the NOVA classification was applied to their ingredients. Precisions and examples are presented in Appendix 1.

Case ascertainment

Confidential. For Review Only

Participants self-declared health events through the yearly health status questionnaire, through a specific check-up questionnaire for health events (every three months) or at any time through a specific interface on the study website. For each incident cancer declared, participants were contacted by a physician of the team and asked to provide any relevant medical records. Whenever necessary, the study physicians contacted the physician of the patient and/or hospitals to collect additional information. Afterwards, all medical data were reviewed by a physician expert committee. 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 cohorts to medico-administrative databases of the National health insurance (SNIIRAM databases). Declared health events were therefore completed by the information from these databases, thereby limiting any potential bias due to participants with cancer who may not report their disease to the study investigators. Last, an additional linkage to the French National cause-specific mortality registry (CépiDC) was used to detect death and potentially missed cancer cases for deceased participants. Cancer cases were classified using the International Chronic Diseases Classification, 10th Revision, Clinical Modification (ICD-10). In this study, all first primary cancers diagnosed between the inclusion and January 1st 2017 were considered as cases, except for basal cell skin carcinoma, which was not considered as cancer.

Medical records were obtained for >90% of cancer cases. Because of the high validity of self-reports (95% of self-reported cancers for whom a medical record was obtained were confirmed by our physicians), we included all cases who self-reported incident cancers, unless they were identified as non-case subjects by a pathology report. In the latter situation, they were classified as non-cases.

Statistical analysis

Up to January 1st 2017, 104,980 participants without cancer at baseline and who provided at least 2 valid 24h-dietary records during their 2 first years of follow-up were included. The flow-chart is presented in Appendix 4. For each subject, the proportion (in weight, % g/day) of ultra-processed foods in the total diet was calculated. The proportion of ultra-processed foods in the diet was determined by making a weight ratio rather than an energy ratio in order to take into account processed food that do not provide any energy (in particular artificially sweetened beverages) and non-nutritional issues related to food processing (e.g. neo-

https://mc.manuscriptcentral.com/bmj

Confidential. For Review Only

formed contaminants, food additives and alterations to the structure of raw foods). For all covariates except physical activity, <5% of values were missing and were imputed to the modal value (for categorical variables) or to the median (for continuous variables). Corresponding values are provided in the footnote to Table 1. For physical activity, the proportion of missing values was higher (14%) since the answers of all IPAQ questions were needed to calculate the score. To avoid massive imputation for a non-negligible number of subjects or exclusion of subjects with missing data and risk of selection bias, we included a missing class into the models for this variable. Differences in baseline characteristics of participants between sex-specific quartiles of the proportion of ultra-processed food in the diet were examined using ANOVA or $\chi 2$ tests wherever appropriate. Cox proportional hazards models with age as the primary time-scale were used to evaluate the association between the proportion of ultra-processed foods in the diet (coded as a continuous variable or as sex-specific quartiles) and incidence of overall, breast, prostate and colorectal cancer risk. In these models, cancers of other locations than the one studied were censored at the date of diagnosis (i.e. they were considered as non-cases for the cancer of interest and they contributed person-year until the date of diagnosis of their cancer). Hazard ratios (HR) and 95% confidence intervals (CI) were estimated with the lowest quartile as the reference category. Log-log (survival) vs. log-time plots were generated in order to confirm risk proportionality assumptions. Tests for linear trend were performed using the ordinal score on sex-specific quartiles of ultra-processed food. Participants contributed person-time until the date of cancer diagnosis, the date of last completed questionnaire, the date of death, or January 1st 2017, whichever occurred first. Breast cancer analyses were additionally stratified by menopausal status. For the latter, women contributed person-time to the "pre-menopause model" until their age at menopause and to the "post-menopause model" from their age at menopause. Age at menopause was determined using the yearly health status questionnaires completed during follow-up.

Models were adjusted for age (time-scale), sex, BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low, computed following IPAQ recommendations [35]), smoking status (never or former smokers, current smokers), number of 24h-dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, kcal/d, continuous), family history of cancer (yes/no), and educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school

Confidential: For Review Only

degree). For breast cancer analyses, additional adjustments were performed for the number of biological children (continuous), menopausal status at baseline (menopausal/peri-menopausal/non-menopausal), hormonal treatment for menopause at baseline (for postmenopausal analyses, yes/no) and oral contraception use at baseline (for premenopausal analyses, yes/no) (Model 1=main model). To test for the potential influence of the nutritional quality of the diet in the relationship between ultra-processed food intake and cancer risk, this model was additionally adjusted for lipid, sodium and carbohydrate intakes (Model 2), or for a Western dietary pattern derived from principal component analysis (Model 3) (details in Appendix 2), or for all these nutritional factors together (Model 4). Besides, mediation analyses were carried out according to the method proposed by Lange et al. [45] to evaluate the direct and indirect effect of the relationship between the exposure and the outcome through these following nutritional mediators: intakes of sodium, total lipids, saturated, mono-unsaturated and poly-unsaturated fatty acids, carbohydrates, and a Western-type dietary pattern. The methodology is detailed in Appendix 3.

Sensitivity analyses were performed based on Model 1 by i) excluding cancer cases diagnosed during the first two years of each participant's follow-up to avoid reverse causality bias, ii) testing sex-specific quintiles of the proportion of ultra-processed foods in the diet instead of sex-specific quartiles, and iii) testing further adjustments for prevalent depression at baseline (yes/no), dietary supplement use at baseline (yes/no), healthy dietary pattern (continuous, details in Appendix 2), number of smoked cigarettes in pack-years (continuous), overall fruit and vegetable consumption (continuous) and season of inclusion in the cohort (spring/summer/autumn/winter). The association between ultra-processed food and overall cancer risk was also investigated separately in different strata of the population: men, women, younger adults (<40y), older adults (≥40y), smokers, non-smokers, participants with a high level of physical activity and those with low-to-moderate level of physical activity. Models were also tested after restriction of the population study to the participants with at least six (respectively, at least one) 24h dietary records during the first two years of follow-up. Associations between the quantity (g/d) of each ultra-processed food group and cancer risk were also tested.

Confidential: For Review Only

Secondary analyses were performed by testing the associations between the proportion in the diet of each of the three other NOVA categories of food processing (continuous) with cancer risk, using multivariate Cox models adjusted for model 1 covariates.

All tests were two-sided, and P<0.05 was considered statistically significant. SAS version 9.4 (SAS Institute) was used for the analyses.

Confidential. For Review Only

RESULTS

A total of 104,980 participants with 22,821 (21.7%) men and 82,159 (78.3%) women were included in the present study. Mean age of participants was 42.8y (SD=14.8) years (range: 18.0-72.8y). Mean number of dietary records per subject over their first two years of follow-up was 5.4 (SD=2.9); the minimum was 2, but it only represented 7.2% of the participants. After the launching of the study by the end of May 2009, 49.6% of the records were filled between June and November and 50.4% between December and May. Main baseline characteristics of participants according to quartiles of the proportion of ultra-processed foods in the diet are described in Table 1. Compared to the first quartile, participants among the highest quartile of ultra-processed food intake tended to be younger, current smokers, less educated, with less family history of cancer and a lower physical activity level. Furthermore, they had higher intakes of energy, lipids, carbohydrates and sodium, along with lower alcohol intake. Although there was a higher proportion of women than men in this cohort, the contribution of ultra-processed foods to the overall diet was very similar between men and women (18.74% for men and 18.71% for women, p=0.7). The distribution of the proportion of ultra-processed food in the diet in the study population is presented in Appendix 5. Main food groups contributing to ultra-processed food intake were sugary products (26%) and beverages (20%), followed by starchy foods and breakfast cereals (16%) and ultra-processed fruits and vegetables (15%) (Figure 1).

TABLE 1 Baseline characteristics of the study population according to sex-specific quartiles of ultra-processed food consumption (n=104,980), NutriNet-Santé cohort, France, 2009-2017[a]

		Quartiles of ultra-processed food consumption[b]				
	All participants	**Quartile 1 (n=26,244)**	**Quartile 2 (n=26,245)**	**Quartile 3 (n=26,246)**	**Quartile 4 (n=26,245)**	**P-trend[c]**
Age, years	42.8 ± 14.8	47.9 ± 13.5	45.0 ± 14.0	42.0 ± 14.4	36.5 ± 13.6	<.0001
Sex, n (%)						
Female	82159 (78.3)	20,539 (78.3)	20,540 (78.3)	20,541 (78.3)	205,42 (78.3)	
Male	22821 (21.7)	5,705 (21.7)	5,706 (21.7)	5,707 (21.7)	5,708 (21.7)	
Height, cm	166.8 ± 8.1	166.3 ± 8.0	166.7 ± 8.0	167.0 ± 8.1	167.3 ± 8.2	<.0001
Body mass index, kg/m²	23.8 ± 4.6	23.8 ± 4.3	23.8 ± 4.4	23.8 ± 4.5	23.8 ± 5.0	0.9
Family history of cancer, yes[d]	35668 (34.0)	10,542 (40.2)	9,624 (36.7)	8,625 (32.9)	6,877 (26.2)	<.0001
Higher education, n (%)						0.01
No	19357 (18.4)	5,154 (19.6)	4,961 (18.9)	4,637 (17.7)	4,605 (17.6)	
Yes <2 years	18076 (17.2)	3,938 (15.0)	1,091 (15.6)	4,426 (16.9)	5,621 (21.4)	
Yes ≥2 years	67,547 (64.3)	17,152 (65.4)	17,193 (65.5)	17,183 (65.5)	16,019 (61.0)	
Smoking status, n (%)						<.0001
Current	17,763 (16.9)	4,127 (15.7)	4,065 (15.5)	4,266 (16.3)	5,305 (20.2)	
Never/former	87,217 (83.1)	22,117 (84.3)	22,180 (84.5)	21,980 (83.8)	20,940 (79.8)	
IPAQ Physical activity level, n (%)[e]						<.0001
High	29603 (28.2)	8,753 (33.4)	7,762 (29.6)	6,983 (26.6)	6,105 (23.3)	
Moderate	38874 (37.0)	9,620 (36.7)	9,953 (37.9)	9,814 (37.4)	9,487 (36.2)	
Low	21888 (20.9)	4,407 (13.8)	4,407 (16.8)	5,839 (22.3)	6,490 (24.7)	
Energy intake without alcohol, kcal/d	1879.0±473.7	1,810.6 ± 454.1	1,881.1 ± 457.7	1,908.5 ± 472.3	1,915.8 ± 501.8	<.0001
Alcohol intake, g/d	7.8 ± 11.9	9.3 ± 13.3	8.5 ± 11.9	7.5 ± 11.3	5.9 ± 10.5	<.0001
Total Lipid intake, g/d	80.5 ± 25.5	76.0 ± 24.3	80.3 ± 24.4	82.1 ± 25.3	83.4 ± 27.3	<.0001
Carbohydrate intake, g/d	195.4 ± 57.9	184.6 ± 57.8	193.9 ± 55.3	199.3 ± 56.6	203.6 ± 60.2	<.0001
Sodium intake, mg/d	2,700.1 ± 893.1	2,589.3 ± 881.6	2,731.8 ± 871.0	2,761.9 ± 884.1	2,717.7 ± 925.0	<.0001

Confidential: For Review Only

Number of children	1.3 ± 1.2	1.6±1.2	1.4±1.2	1.3±1.2	1.0±1.2	<.0001
Menopausal status, n (%)[f]						<.0001
Premenopausal	57408 (69.9)	11,797 (57.4)	13,497 (65.7)	14,961 (72.8)	17,153 (83.5)	
Perimenopausal	4282 (5.2)	1,471 (7.16)	1,148 (5.6)	997 (4.9)	666 (3.2)	
Postmenopausal	20469 (24.9)	7,271 (35.4)	5,895 (28.7)	4,582 (22.3)	2,721 (13.3)	
Use of hormonal treatment for menopause, yes n (%)[f]	4324 (5.3)	1602 (7.8)	1242 (6.1)	932 (4.5)	548 (2.7)	<.0001
Oral contraception, yes n (%)[f]	23073 (22.0)	3,779 (14.4)	4,990 (19.0)	6,209 (23.7)	8,095 (30.8)	<.0001
Ultraprocessed food (%)	18.7 ± 10.1	8.5 ± 2.5	14.3 ± 1.4	19.8 ± 1.9	32.3 ± 9.8	-

[a]Values are means ± SDs or n (%). For all covariates except physical activity, a very low proportion of values were missing (0-5%), the latter were replaced by the modal value among the population study: '≥2y of higher education' for educational level, 0 for the number of biological children, 22.9 kg/m2 for BMI, 166 cm for height and non-smoker for smoking status.

[b]Sex specific quartiles of the proportion of ultra-processed food intake in the total quantity of food consumed. Sex-specific cut-offs for quartiles of ultra-processed proportions were 11.8%, 16.8% and 23.3% in men and 11.8%, 16.8% and 23.4% in women.

[c] P_{value} for the comparison between sex-specific quartiles of ultra-processed food consumption, by Fisher test or x^2 test where appropriate.

[d]Among first-degree relatives

[e] Available for 14615 subjects. Subjects were categorized into the "high", "moderate" and "low" categories according to IPAQ guidelines[35]

[f]Among women

Confidential: For Review Only

During follow-up (426,362 person-years, median follow-up time=5y), 2,228 first incident cancer cases were diagnosed and validated, among which 739 breast cancers (n=264 pre-menopausal and n=475 post-menopausal), 281 prostate cancers and 153 cases of colorectal cancers. Among these cases, 108 (4.8%) were identified during mortality follow-up with the national CépiDC database. The abandon rate in the NutriNet-Santé cohort was 6.7%. Associations between the proportion of ultra-processed foods in the diet and overall, breast, prostate and colorectal cancer risks are shown in Table 2. Corresponding cumulative incidence curves are shown in Figure 2. In model 1, ultra-processed food intake was associated with increased risks of overall cancer ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.12 (1.06-1.18), P<.0001) and breast cancer (HR=1.11 (1.02-1.22), P=0.02). The later association was more specifically observed for post-menopausal breast cancer (P=0.04) but not for pre-menopausal breast cancer (P=0.2). The association with overall cancer risk was statistically significant in all strata of the population investigated, after adjustment for model 1 covariates: in men ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.12 (1.02-1.24), P=0.02, 663 cases and 22158 non-cases), in women (HR= 1.13 (1.06-1.20), P<0.0001, 1565 cases and 80594 non-cases), in younger adults (<40 years old, HR= 1.21 (1.09-1.35), P=0.0006, 287 cases and 48627 non-cases), in older adults (≥40 years old, HR= 1.09 (1.03-1.16), P=0.03, 1941 cases and 54485 non-cases), in smokers (including adjustment for pack-years of cigarette smoked, HR =1.18 (1.04-1.33), P=0.01, 255 cases and 15355 non-cases), in non-smokers (HR=1.11 (1.05-1.17), P=0.0002, 1943 cases and 85219 non-cases), in subjects with low-to-moderate levels of physical activity (HR=1.07 (1.00-1.15), P=0.04, 1216 cases and 59546 non-cases), and in those with a high level of physical activity (HR=1.19 (1.09-1.30), P<0.0001, 744 cases and 28859 non-cases).

Confidential. For Review Only

TABLE 2 Associations between ultra-processed food intake and overall, prostate, colorectal and breast cancer risk, from multivariable Cox proportional hazard models, NutriNet-Santé cohort, France, 2009 – 2017 (n=104,980)[a]

	Proportion of ultra-processed food intake in the diet											
	Continuous[b]			Sex-specific quartiles[c]								
				Q1	Q2		Q3			Q4		
	HR	95% CI	*P*-trend	HR	HR	95% CI	HR	95% CI		HR	95% CI	*P*-trend
All cancers												
N for cases/non-cases	2228/102752			712/25532	607/25638		541/25705			368/25877		
Model 1	1.12	1.06 - 1.18	<.0001	1	0.99	0.89 - 1.11	1.10	0.99 - 1.24		1.21	1.06 - 1.38	0.002
Model 2	1.12	1.07 - 1.18	<.0001	1	1.00	0.90 - 1.11	1.11	0.99 - 1.25		1.23	1.08 - 1.40	0.001
Model 3	1.12	1.06 - 1.18	<.0001	1	0.99	0.89 - 1.11	1.01	0.98 - 1.23		1.21	1.06 - 1.38	0.002
Model 4	1.13	1.07 - 1.18	<.0001	1	1	0.90 - 1.11	1.11	0.99 - 1.24		1.23	1.08 - 1.40	0.001
Prostate cancer												
N for cases/non-cases	281/22540			96/5609	96/5609		59/5647			30/5675		
Model 1	0.98	0.83 - 1.16	0.8	1	1.18	0.89 - 1.57	0.95	0.69 - 1.32		0.93	0.61 - 1.40	0.6
Model 2	0.98	0.83 - 1.16	0.8	1	1.18	0.89 - 1.57	0.95	0.69 - 1.32		0.93	0.61 - 1.40	0.6
Model 3	0.98	0.83 - 1.15	0.8	1	1.18	0.89 - 1.56	0.95	0.68 - 1.31		0.92	0.61 - 1.39	0.6
Model 4	0.98	0.83 - 1.16	0.8	1	1.18	0.89 - 1.57	0.95	0.68 - 1.32		0.93	0.61 - 1.40	0.6
Colorectal cancer												
N for cases/non-cases	153/104827			48/26196	43/26202		36/26210			26/26219		
Model 1	1.13	0.92 - 1.38	0.2	1	1.10	0.72, 1.66	1.17	0.76 - 1.81		1.49	0.92 - 2.43	0.1
Model 2	1.16	0.95 - 1.42	0.1	1	1.12	0.74, 1.70	1.22	0.79 - 1.90		1.59	0.97 - 2.60	0.07
Model 3	1.13	0.92 - 1.38	0.2	1	1.09	0.92, 1.38	1.16	0.75 - 1.80		1.48	0.91 - 2.41	0.1
Model 4	1.16	0.95 - 1.42	0.1	1	1.12	0.74, 1.70	1.22	0.79 - 1.89		1.23	1.08 - 1.40	0.07
Breast cancer												
N for cases/non-cases	739/81420			247/20292	202/20338		179/20361			111/20429		
Model 1	1.11	1.02 - 1.22	0.02	1	0.97	0.81 - 1.17	1.10	0.90 - 1.34		1.14	0.91 - 1.44	0.2
Model 2	1.11	1.01 - 1.21	0.03	1	0.96	0.80 - 1.16	1.09	0.89 - 1.32		1.12	0.89 - 1.42	0.2
Model 3	1.11	1.02 - 1.22	0.02	1	0.97	0.80 - 1.17	1.09	0.90 - 1.33		1.14	0.91 - 1.44	0.2
Model 4	1.11	1.01 - 1.21	0.03	1	0.96	0.80 - 1.16	1.08	0.89 - 1.32		1.13	0.89 - 1.42	0.2
Pre-menopausal breast cancer												
N for cases/non-cases	264/57151			90/14263	70/14284		55/14299			49/14305		
Model 1	1.09	0.95 - 1.25	0.2	1	0.91	0.67 - 1.25	0.92	0.65 - 1.29		1.30	0.90 - 1.86	0.3
Model 2	1.07	0.93 - 1.23	0.4	1	0.90	0.66 - 1.24	0.90	0.64 - 1.27		1.25	0.87 - 1.80	0.4
Model 3	1.09	0.95 - 1.26	0.2	1	0.91	0.67 - 1.25	0.92	0.66 - 1.30		1.30	0.91 - 1.88	0.3

Confidential: For Review Only

Model 4	1.08	0.94 - 1.24	0.3	1	0.91	0.66 - 1.24	0.91	0.64 - 1.28	1.27	0.88 - 1.83	0.4
Post-menopausal											
breast cancer											
N for cases/non-cases		475/29191		107/7309		128/7289		123/7294		117/7299	
Model 1	1.13	1.01 - 1.27	0.04	1	1.23	0.95 - 1.60	1.28	0.98 - 1.66	1.39	1.07 - 1.82	0.02
Model 2	1.13	1.00 - 1.27	0.05	1	1.23	0.95 - 1.60	1.27	0.98 - 1.65	1.39	1.05 - 1.81	0.02
Model 3	1.13	1.00 - 1.27	0.04	1	1.23	0.95 - 1.59	1.27	0.98 - 1.65	1.38	1.06 - 1.81	0.02
Model 4	1.13	1.00 - 1.27	0.05	1	1.23	0.95 - 1.59	1.27	0.97 - 1.65	1.38	1.05 - 1.81	0.02

CI, confidence interval, HR, Hazard ratio

[a] Model 1 is a multivariable Cox proportional hazard model adjusted for age (timescale), sex, energy intake without alcohol, number of 24h-dietary records, smoking status, educational level, physical activity, height, BMI, alcohol intake, and family history of cancers. Breast cancer models were additionally adjusted for menopausal status, hormonal treatment for menopause, oral contraception and number of children.

Model 2 = Model 1 + lipid intake, sodium intake, carbohydrate intake

Model 3 = Model 1 + Western dietary pattern (derived by factor analysis)

Model 4 = Model 1 + lipid intake, sodium intake, carbohydrate intake, Western dietary pattern (derived by factor analysis).

Pearson correlation coefficients with the Western dietary pattern were 0.5 for dietary lipids, 0.6 for sodium and 0.40 for carbohydrates.

[b] HR for an increase of 10% of the proportion of ultra-processed food intake in the diet

[c] Sex-specific cut-offs for quartiles of ultra-processed proportions were 11.8% ; 16.8% and 23.3% in men and 11.8% ; 16.8% and 23.4% in women.

In premenopausal women : Cut-offs for quartiles of ultra-processed proportions were 12.8% ; 18.1% and 25.0%. In postmenopausal women : Cut-offs for quartiles of ultra-processed proportions were 10.1% ; 14.3% and 19.5%.

Confidential: For Review Only

More specifically, ultra-processed fats and sauces (P=0.002), sugary products (P=0.03), and beverages (P=0.005) were associated with increased overall cancer risk and ultra-processed sugary products were associated with breast cancer risk (P=0.006) (Appendix 6).

Further adjustment for several indicators of the nutritional quality of the diet (lipid, sodium and salt intakes – model 2; Western pattern – model 3; or both – model 4) did not modify these findings. The Pearson correlation coefficient between the proportion of ultra-processed food in the diet and the Western-type dietary pattern was low (0.06). Consistently, analyses performed according to the method proposed by Lange et al. [45] to assess a potential mediation of the relationship between ultra-processed food and cancer risk by these nutritional factors showed no statistically significant mediation effect of any of the factors tested. The mediated effects ranged between 0 and 2%, with all P>0.05 (Appendix 3).

No association was statistically significant for prostate and colorectal cancers. However, a borderline non-significant trend of increased colorectal cancer risk associated with ultra-processed food intake was observed ($HR_{Q4 \text{ versus } Q1}$=1.23 (1.08, 1.40), P-trend=0.07 in Model 4).

Sensitivity analyses (adjusted for model 1 covariates, data not tabulated) excluding cancer cases diagnosed during the first two years of follow-up provided similar results ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.10 (1.03-1.17), P =0.005 for overall cancer risk, n=1367 cases and 102502 non-cases included; HR=1.15 (1.03-1.29), P =0.02 for breast cancer risk, n=441 cases and 80940 non-cases included). Similarly, results were unchanged when non-validated cancer cancers were excluded ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.11 (1.05-1.17), P=0.0003 for overall cancer risk, n=1967 cases and 102752 non-cases included; HR=1.12 (1.02-1.23), P=0.02 for breast cancer risk, n=677 cases and 81274 non-cases included).

Similar results were observed when i) we included only participants with at least six 24h records on the one hand (overall cancer risk: $HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$= 1.13 (1.06-1.21)), P =0.0003, n = 1494 cases and 47 920 non-cases included) and ii) we re-included participants with only one

Confidential: For Review Only

24h record on the other hand (overall cancer risk: $HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.11 (1.06-1.16)), P=0.0001, n = 2383 cases and 122 196 non-cases included).

Similar findings were found when the proportion of ultra-processed food in the diet was coded as sex-specific quintiles instead of sex-specific quartiles (overall cancer risk: $HR_{Q5 \text{ versus } Q1}$= 1.25 (1.08-1.47), P-trend=0.0003 and breast cancer risk: $HR_{Q5 \text{ versus } Q1}$= 1.25 (0.96-1.63), P-trend=0.03).

Further adjustment for the following variables, in addition to model 1 covariates, did not modify the results: dietary supplement use at baseline ($HR_{\text{for a 10-point increment in the proportion of ultra-processed foods in the diet}}$=1.12 (1.06-1.17), P<0.0001 for overall cancer and 1.11 (1.02-1.22), P=0.02 for breast cancer), prevalent depression at baseline (HR=1.11 (1.06-1.17), P<0.0001 for overall cancer and 1.11 (1.01-1.22), P=0.02 for breast cancer), healthy dietary pattern (HR =1.11 (1.05-1.17), P<0.0001 for overall cancer and 1.10 (1.00-1.21), P=0.04 for breast cancer), overall fruit and vegetable consumption in g/d (HR= 1.10 (1.04-1.16), P=0.0009 for overall cancer and 1.11 (1.01-1.22), P=0.03 for breast cancer), number of smoked cigarettes in pack-years (HR = 1.13 (1.07-1.19), P<0.0001 for overall cancer and 1.13 (1.03-1.24), P=0.009 for breast cancer), and season of inclusion in the cohort (HR = 1.12 (1.06-1.18), P<0.0001 for overall cancer and 1.12 (1.02-1.22), P=0.02 for breast cancer).

Besides, we have tested other methods to deal with missing data, such as multiple imputation[46] and complete case analysis (i.e. exclusion of participants with at least one missing data for a covariate). The results were very similar: for the multiple imputation analysis: $HR_{\text{for a 10-point increment in the proportion of ultraprocessed foods in the diet}}$=1.11 (1.06-1.17), P<0.0001, 2228 cases and 102752 non-cases for overall cancer, HR=1.11 (1.01-1.21), P=0.02, 739 cases and 81420 non-cases for breast cancer; and for the complete case analysis: HR =1.11 (1.05-1.18), P=0.0003, 1813 cases and 82824 non-cases for overall cancer, HR=1.14 (1.03-1.26), P=0.01, 579 cases and 64642 non-cases for breast cancer.

As a secondary analysis, associations between the proportions of the three other NOVA degrees of food processing and cancer risk were also tested. No significant associations were found between the proportions of “processed culinary ingredients” nor “processed foods” with cancer risk at any location (all p>0.05). However, and consistently with our findings, the consumption of “minimally/unprocessed foods” was associated with lower risks of overall and breast cancers ($HR_{\text{for a 10-point increment in the proportion of unprocessed foods in the}}$

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

$_{diet}$=0.91 (0.87-0.95), P<.0001, 2228 cases and 102752 non-cases for overall cancer, HR=0.42 (0.19-0.91), P=0.03, 739 cases and 81420 non-cases for breast cancer), in multivariable analyses adjusted for model 1 covariates.

DISCUSSION

In this large prospective cohort, a 10% increase in the proportion of ultra-processed foods in the diet was associated with a 12% and 11% significant increase in overall and breast cancer risks, respectively. While a few studies previously suggested that ultra-processed foods may contribute to increase the risk of cardiometabolic disorders - such as obesity [29], hypertension [30], and dyslipidaemia [28] - no prior prospective epidemiological study evaluated the association between food processing and cancer risk.

There is no available estimate of the proportion of ultra-processed food in the diet at the national level in France. However, in the nationally representative INCA3 study conducted by the French Food safety Agency in 2016 [4], "transformed" foods included sweet pastries, biscuits, dairy desserts, ice cream, fruit purée and fruit in syrup, fruit and vegetable juices, soups and broths, sandwiches, pizzas and salted pastries, as well as mixed dishes composed of egg, meat, fish, vegetable and/or starchy foods (cereals, legumes or potatoes). More than half of the "transformed" foods consumed outside catering establishments by adults aged 18-79 were manufactured industrially, about 1/3 were homemade, while the rest was handcrafted (e.g. caterer). These figures illustrate the important share of processed – and especially industrially processed – foods in the diet of French adults.

Several hypotheses could be put forward to explain our findings. The first one relates to the generally poorer nutritional quality of diets rich in ultra-processed foods. Indeed, diets that include a higher proportion of processed food products tended to be richer in energy, sodium, fat and sugar and poorer in fibres and various micronutrients in several studies conducted in various countries [10-17;19]. Ultra-processed foods have also been associated with a higher glycaemic response and a lower satiety effect [47]. Although not being the unique

Confidential: For Review Only

determinant, excessive energy, fat, and sugar intakes contribute to weight gain and obesity risk, the latter being recognized as a major risk factor for the following cancers: post-menopausal breast, stomach, liver, colorectal, oesophagus, pancreas, kidney, gallbladder, endometrium, ovary, liver, prostate (advanced) and hematological malignancies [29]. For instance, body fatness in post-menopausal women is estimated to contribute to 17% of the breast cancer burden [2]. Besides, most of ultra-processed foods, such as dehydrated soups, processed meats, biscuits and sauces, have a high salt content. Salt-preserved foods are associated with increased gastric cancer risk [29]. Conversely dietary fiber intake decreases colorectal cancer risk with a convincing level of evidence [3;29] and may also reduce breast cancer risk [3]. However, the association between ultra-processed food intake and cancer risk observed in this study were statistically significant despite adjustment for BMI, and remained significant after further adjustment for a Western-type dietary pattern and/or energy, fat, sugar and salt content of the diet. Besides, mediation analyses did not support a strong effect of the "nutritional quality" component in this association, thereby suggesting that other bioactive compounds contained in ultra-processed food may contribute to explain the observed relationships.

A second interpretation track concerns the wide range of additives contained in ultra-processed foods. While maximum authorized levels normally protect the consumers against adverse effects of each individual substance in a given food product [48], health impact of the cumulative intake across all ingested foods and potential cocktail/interaction effects remain largely unknown. More than 250 different additives are authorized for an adjunction to food products in Europe and in the US [22;49]. For some of them, experimental studies on animal or cellular models have suggested carcinogenic properties that deserve further investigation in humans [23;24;50-53]. For instance, this is the case for titanium dioxide (TiO_2), a common food additive that contains nanoscale particules and that is used as a whitening agent or in packaging in contact with food or beverages to provide a better texture and anti-microbial properties. Experimental studies, mainly conducted in rodent models, suggested that this additive could initiate or promote the development of colon preneoplastic lesions, as well as chronic intestinal inflammation, thus, TiO_2 was evaluated as "possibly carcinogenic to humans" (Group 2B) by the World Health Organization - International Agency for Research on Cancer (WHO-IARC) [24]. The effects of intense artificial sweeteners such as aspartame on human metabolism and gut microbiota composition/functioning are also controversial [53]. Although previous

Confidential: For Review Only

experimental studies on animals confirmed the safety of aspartame, their relevance to human health outcomes has been questioned, particularly regarding a potential long-term carcinogenicity[51]. Moreover, another concern about sodium nitrite is the formation of carcinogenic nitrosamines in meats containing sodium nitrite when meat is charred or overcooked. These N-nitroso compounds may be involved in the etiology of colorectal cancer [23;52].

Next, food processing and particularly heat treatments produce neoformed contaminants (e.g.acrylamide) in ultraprocessed products such as fried potatoes, biscuits, bread or coffee. A recent meta-analysis underlined a modest association between dietary acrylamide and both kidney and endometrial cancer risks, in non-smokers [54]. In addition, the European Food Safety Agency (EFSA) judged that proofs from animal studies were sufficient to classify acrylamide as genotoxic [20].

Lastly, bisphenol A (BPA) is another contaminant suspected of migrating from plastic packaging of ultra-processed foods. Its endocrine disruptor properties made it judged as "a substance of very high concern" by the European Chemicals Agency (ECHA) [55]. There is increasing evidence for involvement in the development of several non-communicable diseases, including cancer [21] linked to endocrinal disruptors.

Strengths of this study pertained to its prospective design and large sample size, along with a detailed and up-to-date dietary intake assessment. Repeated 24h-dietary records (including 3300 different food items) are more accurate than food frequency questionnaires with aggregated food groups and than household purchasing data. However, some limitations should be acknowledged. First, as it is generally the case in volunteer-based cohorts, participants to the NutriNet-Santé cohort were more often women, with health-conscious behaviours and higher socio-professional and educational levels as compared to the general French population[56]. This might limit the generalizability of the findings and may have resulted in 1) a lower cancer incidence compared to national estimates (age and sex standardized incidence rate per 100,000 persons per year: 786 cases in our cohort vs 972 cases in France [57]) and 2) an overall lower exposure to ultra-processed foods, with less contrast between extreme categories. These points rather tended to underestimate the strength of the associations. However, the possibility that selection bias may have led to an overestimation of some associations cannot be totally excluded. Second, some misclassification in the

https://mc.manuscriptcentral.com/bmj

Confidential: For Review Only

NOVA 'ultra-processed food' category cannot be ruled out. Third, despite a multi-source strategy for case ascertainment (combining validation of health events declared by participants, medico-administrative databases from the health insurance, and national death registry), exhaustiveness of cancer cases detection cannot be guaranteed. Furthermore, statistical power was limited for some cancer locations (such as colorectal cancer), which may have impaired our ability to detect hypothesized associations. Next, the length of follow-up was relatively limited in time, since the cohort was launched in 2009. Thus, it allowed us to study mostly mid-term associations between ultra-processed food consumption and cancer risk. As it is usually the case in nutritional epidemiology, the assumption is made that the measured exposure at baseline (especially since we averaged a two-year period of exposure) actually reflects more generally the usual eating habits of the individual during adulthood, including several years prior to his/her entry into the cohort. However, since some carcinogenic processes may take several decades, it will be important in the future to re-assess the associations between ultra-processed food and cancer risk in the cohort, in order to investigate longer-term effects. This will be one of the perspectives of the present work for the upcoming 5-10 years. Last, although a large range of confounding factors was included in the analyses, the hypothesis of residual confounding resulting from unmeasured behavioural factors and/or imprecision in the measure of included covariates cannot be entirely excluded due to the observational design of this study. For instance in breast cancer models, oral contraception was a binary variable, since the precise doses, type and duration of contraceptive use across reproductive life were not available. Randomized controlled trials have long been considered the only gold standard to eliminate confounding bias, however, they do not capture consumption as it is in daily life. Moreover, a trial would not be ethically feasible to investigate exposure for which a deleterious effect is suspected. Our large observational cohort was therefore particularly adapted to provide insights in this field.

To our knowledge, this study was the first to investigate and highlight an increase in overall – and more specifically, breast – cancer risk associated with ultra-processed food intake. These results should be confirmed by other large-scale population-based observational studies in different populations and settings. Further studies are also needed to better understand the relative impact of nutritional composition, food

Confidential: For Review Only

additives, contact materials, and neoformed contaminants in this relationship. Rapidly increasing consumption of ultra-processed foods may drive an increasing burden of cancer and other non-communicable diseases. Thus, policy actions targeting product reformulation, taxation and marketing restrictions on ultra-processed products and promotion of fresh or minimally processed foods may contribute to primary cancer prevention [6;9]. Several countries have already introduced this aspect in their official nutritional recommendations in the name of the precautionary principle [58;59].

Confidential: For Review Only

Acknowledgements

The authors sincerely thank all the volunteers of the NutriNet-Santé cohort. We also thank Véronique Gourlet, Lucien Martinez, Nathalie Arnault, Stephen Besseau, Laurent Bourhis, Yasmina Chelghoum, Than Duong Van, Younes Esseddik, Paul Flanzy, Julien Allègre, Mac Rakotondrazafy, Régis Gatibelza, Fabien Szabo, Roland Andrianasolo, Fatoumata Diallo, Ludivine Ursule, Cédric Agaesse, Claudia Chahine, Anne-Elise Dussoulier and Marion Genest for their technical contribution to the NutriNet-Santé study.

Competing interests statement

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Contributorship statement and guarantor

The authors' contributions were as follows – TF, CJ, EKG, CM, BA and MT: designed the research; SH, MT, CJ, EKG: conducted the research; TF: performed statistical analysis; MT and BS: supervised statistical analysis; TF and MT: wrote the paper; BS performed sensitivity analyses and was in charge of the revision of the paper; TF, BS, LS, MD, PF, PLM, EKG, BA, MB, SH, PG, CL, CM, CJ, and MT: contributed to the data interpretation and revised each draft for important intellectual content. All authors read and approved the final manuscript. MT had primary responsibility for the final content, she is the guarantor. None of the authors reported a conflict of interest related to the study. The funders had no role in the design, implementation, analysis, or interpretation of the data.

Transparency statement

Dr Touvier (the guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Confidential: For Review Only

Copyright / licence for publication statement

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

Data sharing statement

No additional data available.

Funding statement and statement of the independence of researchers from funders

The NutriNet-Santé study was supported by the following public institutions: Ministère de la Santé, Institut de Veille Sanitaire (InVS), Institut National de la Prévention et de l'Education pour la Santé (INPES), Région Ile-de-France (CORDDIM), Institut National de la Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche Agronomique (INRA), Conservatoire National des Arts et Métiers (CNAM) and Université Paris 13. Mélanie Deschasaux and Philippine Fassier were funded by a PhD grant from the Cancéropôle Ile de France / Région Ile de France (public funding). Bernard Srour was funded by the French National Cancer Institute (grant number INCa_8085).

Researchers were independent from funders. Funders had no role in the study design, the collection, analysis, and interpretation of data, the writing of the report, and the decision to submit the article for publication.

Confidential: For Review Only

Patient involvement statement

The research question developed in this article corresponds to a strong concern of the participants involved in the NutriNet-Santé cohort, and of the public in general. Participants to the study are thanked in the Acknowledgements section. The results of the present study will be disseminated to the NutriNet-Santé participants through the cohort website, public seminars and a press release.

Confidential: For Review Only

References

(1) Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015; 136(5):E359-E386.

(2) World Cancer Research Fund International/American Institute for Cancer Research. Cancer preventability estimates. Available at: wcrf.org/cancer-preventability-estimates. [Accessed on 03-07-2017], 2017.

(3) Latino-Martel P, Cottet V, Druesne-Pecollo N, Pierre FH, Touillaud M, Touvier M et al. Alcoholic beverages, obesity, physical activity and other nutritional factors, and cancer risk: A review of the evidence. *Crit Rev Oncol Hematol* 2016; 99:308-323.

(4) ANSES (French Agency for Food,Environmental and Occupational Health & Safety). Individual national study for dietary intakes 3 (INCA 3) [Étude individuelle nationale des consommations alimentaires 3 (INCA 3)] (in French), 2017.

(5) Monteiro CA, Moubarac JC, Cannon G, Ng SW, Popkin B. Ultra-processed products are becoming dominant in the global food system. *Obes Rev* 2013; 14 Suppl 2:21-28.

(6) Moodie R, Stuckler D, Monteiro C, Sheron N, Neal B, Thamarangsi T et al. Profits and pandemics: prevention of harmful effects of tobacco, alcohol, and ultra-processed food and drink industries. *Lancet* 2013; 381(9867):670-679.

(7) Moubarac JC, Batal M, Martins AP, Claro R, Levy RB, Cannon G et al. Processed and ultra-processed food products: consumption trends in Canada from 1938 to 2011. *Can J Diet Pract Res* 2014; 75(1):15-21.

(8) Venn D, Banwell C, Dixon J. Australia's evolving food practices: a risky mix of continuity and change. *Public Health Nutr* 2016;1-10.

(9) 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;1-13.

(10) Luiten CM, Steenhuis IH, Eyles H, Ni MC, Waterlander WE. Ultra-processed foods have the worst nutrient profile, yet they are the most available packaged products in a sample of New Zealand supermarkets. *Public Health Nutr* 2016; 19(3):539.

(11) Adams J, White M. Characterisation of UK diets according to degree of food processing and associations with socio-demographics and obesity: cross-sectional analysis of UK National Diet and Nutrition Survey (2008-12). *Int J Behav Nutr Phys Act* 2015; 12:160.

(12) Cediel G, Reyes M, da Costa Louzada ML, Martinez SE, Monteiro CA, Corvalan C et al. Ultra-processed foods and added sugars in the Chilean diet (2010). *Public Health Nutr* 2017;1-9.

(13) Costa Louzada ML, Martins AP, Canella DS, Baraldi LG, Levy RB, Claro RM et al. Ultra-processed foods and the nutritional dietary profile in Brazil. *Rev Saude Publica* 2015; 49:38.

(14) Martinez SE, Baraldi LG, Louzada ML, Moubarac JC, Mozaffarian D, Monteiro CA. Ultra-processed foods and added sugars in the US diet: evidence from a nationally representative cross-sectional study. *BMJ Open* 2016; 6(3):e009892.

Confidential: For Review Only

(15) Moubarac JC, Martins AP, Claro RM, Levy RB, Cannon G, Monteiro CA. Consumption of ultra-processed foods and likely impact on human health. Evidence from Canada. *Public Health Nutr* 2013; 16(12):2240-2248.

(16) Moubarac JC, Batal M, Louzada ML, Martinez SE, Monteiro CA. Consumption of ultra-processed foods predicts diet quality in Canada. *Appetite* 2017; 108:512-520.

(17) Poti JM, Mendez MA, Ng SW, Popkin BM. Is the degree of food processing and convenience linked with the nutritional quality of foods purchased by US households? *Am J Clin Nutr* 2015; 101(6):1251-1262.

(18) Slimani N, Deharveng G, Southgate DA, Biessy C, Chajes V, van Bakel MM et al. Contribution of highly industrially processed foods to the nutrient intakes and patterns of middle-aged populations in the European Prospective Investigation into Cancer and Nutrition study. *Eur J Clin Nutr* 2009; 63 Suppl 4:S206-S225.

(19) Louzada ML, Martins AP, Canella DS, Baraldi LG, Levy RB, Claro RM et al. Impact of ultra-processed foods on micronutrient content in the Brazilian diet. *Rev Saude Publica* 2015; 49:45.

(20) EFSA (European Food Safety Agency). EFSA Panel on Contaminants in the Food Chain (CONTAM). «Scientific Opinion on Acrylamide in Food». *EFSA Journal* 13, n° 6 (1 juin 2015): doi:10.2903/j.efsa.2015.4104, 2015.

(21) Muncke J. Endocrine disrupting chemicals and other substances of concern in food contact materials: an updated review of exposure, effect and risk assessment. *J Steroid Biochem Mol Biol* 2011; 127(1-2):118-127.

(22) European Union (EU) - The database of authorized food additives, Available online on https://ec.europa.eu/food/safety/food_improvement_agents/additives/database_en. [Accessed on 26 October 2017], 2008.

(23) Bouvard V, Loomis D, Guyton KZ, Grosse Y, Ghissassi FE, Benbrahim-Tallaa L et al. Carcinogenicity of consumption of red and processed meat. *Lancet Oncol* 2015; 16(16):1599-1600.

(24) Carbon black, titanium dioxide, and talc. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. IARC Monogr Eval Carcinog Risks Hum. 2010; 93:1-413.

(25) Canella DS, Levy RB, Martins AP, Claro RM, Moubarac JC, Baraldi LG et al. Ultra-processed food products and obesity in Brazilian households (2008-2009). *PLoS One* 2014; 9(3):e92752.

(26) Juul F, Hemmingsson E. Trends in consumption of ultra-processed foods and obesity in Sweden between 1960 and 2010. *Public Health Nutr* 2015; 18(17):3096-3107.

(27) Louzada ML, Baraldi LG, Steele EM, Martins AP, Canella DS, Moubarac JC et al. Consumption of ultra-processed foods and obesity in Brazilian adolescents and adults. *Prev Med* 2015; 81:9-15.

(28) Rauber F, Campagnolo PD, Hoffman DJ, Vitolo MR. Consumption of ultra-processed food products and its effects on children's lipid profiles: a longitudinal study. *Nutr Metab Cardiovasc Dis* 2015; 25(1):116-122.

(29) Mendonca RD, Pimenta AM, Gea A, Fuente-Arrillaga C, Martinez-Gonzalez MA, Lopes AC et al. Ultraprocessed food consumption and risk of overweight and obesity: the University of Navarra Follow-Up (SUN) cohort study. *Am J Clin Nutr* 2016; 104(5):1433-1440.

Confidential: For Review Only

(30) Mendonca RD, Lopes AC, Pimenta AM, Gea A, Martinez-Gonzalez MA, Bes-Rastrollo M. Ultra-Processed Food Consumption and the Incidence of Hypertension in a Mediterranean Cohort: The Seguimiento Universidad de Navarra Project. *Am J Hypertens* 2017; 30(4):358-366.

(31) Hercberg S, Castetbon K, Czernichow S, Malon A, Mejean C, Kesse E et al. The Nutrinet-Sante 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.

(32) Vergnaud AC, Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A et al. Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Sante study. *Int J Public Health* 2011; 56(4):407-417.

(33) Lassale C, Peneau S, Touvier M, Julia C, Galan P, Hercberg S et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Sante study. *J Med Internet Res* 2013; 15(8):e152.

(34) Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur J Epidemiol* 2010; 25(5):287-296.

(35) Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003; 35(8):1381-1395.

(36) Touvier M, Kesse-Guyot E, Mejean C, Pollet C, Malon A, Castetbon K et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* 2011; 105(7):1055-1064.

(37) Lassale C, Castetbon K, Laporte F, Camilleri GM, Deschamps V, Vernay M et al. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br J Nutr* 2015; 113(6):953-962.

(38) Lassale C, Castetbon K, Laporte F, Deschamps V, Vernay M, Camilleri GM et al. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. *J Acad Nutr Diet* 2016; 116(3):427-438.

(39) Le Moulenc N, Deheeger M, Preziozi P, Montero P, Valeix P, Rolland-Cachera M et al. Validation du manuel photo utilisé pour l'enquête alimentaire de l'étude SU.VI.MAX. [Validation of the food portion size booklet used in the SU.VI.MAX study] (in French), *Cah Nutr Diet* 1996;31:158-64.

(40) 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. *Int J Obes Relat Metab Disord* 2000; 24(9):1119-1130.

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

(42) Monteiro CA, Cannon G, Levy RB, Moubarac JC, Jaime PC, Martins AP et al. NOVA. The star shines bright. *World Nutrition* 2016; 7(1-3):28-38.

(43) 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. *Curr Obes Rep* 2014; 3(2):256-272.

https://mc.manuscriptcentral.com/bmj

Confidential for Review Only

(44) Julia C, Martinez L, Alles B, Touvier M, Hercberg S, Mejean C et al. Contribution of ultra-processed foods in the diet of adults from the French NutriNet-Sante study. *Public Health Nutr* 2017;1-11.

(45) Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol* 2012; 176(3):190-195.

(46) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; 338:b2393.

(47) Fardet A. Minimally processed foods are more satiating and less hyperglycemic than ultra-processed foods: a preliminary study with 98 ready-to-eat foods. *Food Funct* 2016; 7(5):2338-2346.

(48) World Health Organization (WHO) - Fact sheet about Food additives, added on July 2017. Available at http://www.who.int/mediacentre/factsheets/food-additives/en/. [Accessed on 26 october 2017], 2017.

(49) US Food & Drug Administration FDA Food Additive Status List. Available at: https://www.fda.gov/food/ingredientspackaginglabeling/foodadditivesingredients/ucm091048.htm [Accessed 03 July 2017], 2017.

(50) Chang X, Zhang Y, Tang M, Wang B. Health effects of exposure to nano-TiO2: a meta-analysis of experimental studies. *Nanoscale Res Lett* 2013; 8(1):51.

(51) EFSA ANS Panel. (EFSA Panel on Food Additives and Nutrient Sources added to Food) - Scientific Opinion on the re-evaluation of aspartame (E 951) as a food additive. EFSA Journal 2013;11(12):3496, 263 pp. doi:10.2903/j.efsa.2013.3496.

(52) Santarelli RL, Vendeuvre JL, Naud N, Tache S, Gueraud F, Viau M et al. Meat processing and colon carcinogenesis: cooked, nitrite-treated, and oxidized high-heme cured meat promotes mucin-depleted foci in rats. *Cancer Prev Res (Phila)* 2010; 3(7):852-864.

(53) Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O et al. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* 2014; 514(7521):181-186.

(54) Virk-Baker MK, Nagy TR, Barnes S, Groopman J. Dietary acrylamide and human cancer: a systematic review of literature. *Nutr Cancer* 2014; 66(5):774-790.

(55) European CHemical Agency (ECHA). Member State Committee support document for identification of 4,4'-isopropylidenediphenol (bisphenol a) as a substance of very high concern because of its toxic for reproduction (Article 57 c) properties. Adopted on 2 December 2016.

(56) Andreeva VA, Salanave B, Castetbon K, Deschamps V, Vernay M, Kesse-Guyot E et al. Comparison of the sociodemographic characteristics of the large NutriNet-Sante e-cohort with French Census data: the issue of volunteer bias revisited. *J Epidemiol Community Health* 2015; 69(9):893-898.

(57) INCa. The French National Institute of Cancer INCA - Cancer in France, 2014 Edition (In French: Les Cancers en France, Editions 2014); Available at: http://www.unicancer.fr/sites/default/files/Les%20cancers%20en%20France%20-%20Edition%202014%20-%20V5.pdf. [Accessed on 26 October 2017], 2017.

(58) Ministry of Health of Brazil. Dietary Guidelines for the Brazilian population, 2014.

Confidential: For Review Only

(59) Haut Conseil de la Santé Publique. Avis relatif à la révision des repères alimentaires pour les adultes du futur Programme National Nutrition Santé 2017-2021, 2017.

Confidential: For Review Only

Figure Legends

Figure 1:

Title: Relative contribution of each food group to ultra-processed consumption in the diet

Figure 2:

Title: Cumulative cancer incidence (overall cancer risk) according to quartiles of ultra-processed food intake

Legend: Q=quartile (1 to 4) of the proportion of ultra-processed food in the diet

Confidential: For Review Only