Ultra-processed food consumption and type 2 diabetes incidence: a prospective cohort study

Renata B. Levy, [1, 2, 4] PhD; Fernanda Rauber,[1,2,3] PhD; Kiara Chang,[4] PhD; Maria Laura C. Louzada,[2,3] PhD; Carlos A. Monteiro,[2,3] PhD; Christopher Millett,[1,2,4] PhD; and Eszter P. Vamos,[4] PhD

[1] Department of Preventive Medicine, School of Medicine, University of São Paulo, São Paulo 01246-903, Brazil.
[2] Center for Epidemiological Research in Nutrition and Health, University of São Paulo, São Paulo 01246-904, Brazil.
[3] Department of Nutrition, School of Public Health, University of São Paulo, São Paulo 01246-904, Brazil.
[4] Public Health Policy Evaluation Unit, School of Public Health, Imperial College London, London W6 8RP, United Kingdom

Renata Bertazzi Levy – ORCID - 0000-0001-5388-7002

Corresponding author:
Renata Bertazzi Levy
Email: rlevy@usp.br

Funding statement: RBL is funded by the São Paulo Research Foundation (FAPESP) 2015/14900-9 and 2018/26660-0. FR is funded by the São Paulo Research Foundation (FAPESP) 2016/14302-7 and 2018/19820-1. This study is funded by the National Institute for Health Research (NIHR) School for Public Health Research (Grant Reference Number PD-SPH-2015). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The funders had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Competing interests: None declared.

Word count: 3111

Abstract word count: 271

Tables: 2

Figures: 1

References: 39

ABSTRACT

Background

Ultra-processed foods account for more than 50% of daily calories consumed in several high-income countries, with sales of ultra-processed foods soaring globally, especially in middle-income countries. The objective of this study is to investigate the association between ultra-processed food (UPF) consumption and risk of type 2 diabetes (T2D) in a UK-based prospective cohort study.

Methods

Participants of the UK Biobank (2007-2019) aged 40-69 years without diabetes at recruitment who provided 24-hour dietary recall and follow-up data were included. UPFs were defined using the NOVA food classification. Multivariable Cox proportional hazards regression models were used to evaluate the association between UPF consumption and the risk of T2D adjusting for socio-demographic, anthropometric and lifestyle characteristics.

Results

A total of 21,730 participants with a mean age of 55.8 years and mean UPF intake of 22.1% at baseline were included. During a mean follow-up of 5.4 years (116,956 person-years), 305 incident T2D cases were identified. In the fully adjusted model, compared with the group in the lowest quartile of UPF intake, the hazard ratio for T2D was 1.44, 1.04-2.02 in the group with the highest quartile of UPF consumption. A gradient of elevated risk of T2D associated with increasing quartiles of UPF intake was consistently observed (p value for trend <0.028). A significantly increased risk of T2D was observed per 10 percentage points increment in UPF consumption ([adjusted HR]: 1.12, 95% confidence interval [CI]: 1.04-1.20).

Conclusions

Our findings demonstrate that a diet high in UPFs is associated with a clinically important increased risk of T2D. Identifying and implementing effective public health actions to reduce UPF consumption in the UK and globally are urgently required.

Keywords: Ultra-processed foods, diabetes, Biobank cohort, United Kingdom, prospective study

INTRODUCTION

In recent decades, the emergence of a global, industrialised food system has displaced traditional dietary patterns based on fresh or minimally processed foods in favour of packaged, ready-to-consume foods.[1] As defined by NOVA, a food classification based on the extent and purpose of food processing, ultra-processed foods are highly palatable, durable and profitable food and drink formulations in which intact foods and their associated health benefits are largely absent.[1] At present, ultra-processed foods account for 54%, 58% and 48% of daily calories consumed in the United Kingdom (UK), the United States and Canada respectively, with sales of ultra-processed foods soaring globally, especially in middle-income countries.[2-9]

Ultra-processed foods are high in saturated fat, sugar and salt, and the growth in consumption may be an important driver of the global obesity and diabetes epidemics. Prospective cohort studies have demonstrated a link between higher consumption and greater risk of overweight and obesity.[10-12] These findings are supported by a recent clinical trial which found that ultra-processed food consumption causes excessive energy intake and weight gain.[13]

To date, only one prospective study has investigated the association between ultra-processed food consumption and type 2 diabetes (T2D) incidence. Based on a French cohort of 104,707 adult participants, the study reported a 13% increased risk for T2D incidence (1.01-1.27) for every 10 percentage points increase in the contribution of ultra-processed foods in the diet.[14] However, this association has not be examined in settings with substantially higher consumption of ultra-processed foods such as in North America and the UK (where annual per capita sales of ultra-processed foods is 140.7 kg vs. 79.0 kg in France), nor in populations with a higher burden of obesity.[15,16] This study investigated the association between ultra-processed food consumption and risk of T2D in a UK-based prospective cohort study.

METHODS

Data source and study population

This study is based on a subsample of 21,730 UK Biobank participants without diabetes at baseline and with valid 24-hour dietary recall and follow-up data available.

UK Biobank is a prospective cohort study of participants aged 40–69 years conducted at 22 assessment centres across the UK, with initial enrolment between 2007-2010 and follow-up assessments between 2012-2019.[17,18] During baseline and follow-up assessments, participants completed questionnaires about their socio-demographic, lifestyle and psychosocial characteristics, and their anthropometric measurements were taken and a detailed medical history and current medication use were recorded by trained research staff.[17]

Participants' dietary intake was assessed using a web-based, self-administered questionnaire that collects quantities of over 200 common food and beverage items consumed in the previous 24 hours.[19] This web-based questionnaire has been shown to capture similar food and beverage items as well as estimated energy and nutrient intakes as an interviewer-administered 24-hour recall.[20] It was introduced towards the end of the recruitment period (2009-2010) and only a subset of participants with known email address were invited to participate online between 2011-2012. Due to the inconsistencies in the timing of assessment centre visits and administrations of 24-hour recalls, we considered participants' first 24-hour recall as this best represents their dietary intake at baseline, and only participants who completed a 24-hour recall within 36 months of their baseline assessment were considered relevant. Pregnant women (n=176) and those with implausible dietary values (n=641) were excluded from our sample. As a result, a total of 23,009 participants with valid 24-hour recall and follow-up data were considered for inclusion.

Categorisation according to food processing

Participants' ultra-processed food consumption was derived from their 24-hour recall. We classified each food and beverage item captured by the 24-hour dietary recall into one of the four food groups according to their extent and purpose of food processing as described by the NOVA food classification system,[1] these are: 1) unprocessed or minimally processed foods, e.g. fruits, vegetables, legumes, roots and tubers, milk and meat; 2) processed culinary ingredients, e.g. table salt, sugar, vegetable oils and butter; 3) processed foods, e.g. canned vegetables in brine, salted or sugar-coated nuts, canned fish, freshly made breads and cheeses; 4) ultra-processed foods, e.g. soft drinks, sweet or savoury packaged snacks, confectionery, breakfast 'cereals', flavoured yoghurts, industrial-processed breads and buns, reconstituted meat products and pre-prepared frozen or shelf-stable ready-to-eat/heat meals.[1]

Since the UK Biobank provided respondents' quantity of each food (and beverage) item consumed but not the nutritional information for each item, we derived these estimates by assigning each food item a typical portion size and a nutrient profile using UK's standard food composition database.[21,22] The estimated weight and energy intake for each food item was calculated as a product of the quantity consumed, its assigned portion size and corresponding nutrient profile. Individuals' ultra-processed food consumption was computed by dividing the amount of ultra-processed foods consumed (i.e. the 4th group of NOVA) by the total amount of foods consumed in grams per day, and expressed as a proportion. This

was preferred over an energy ratio as it accounts for the ultra-processed foods with no energy content, such as artificially sweetened beverages, but additional analyses were performed using ultra-processed consumption derived from the proportion of daily energy intake and are presented as supplementary. In addition, a categorical variable was constructed which represents the sex-specific quartiles of the proportion of ultra-processed food consumption out of total food intake.

Identification of incident T2D cases

The identification of diabetes cases was based on self-reported and nurse-interview data, which enquired whether participants have ever been told by a doctor that they have diabetes. Participants with diabetes were asked whether the diagnosis was only during pregnancy; the type of diabetes; the age when diabetes was first diagnosed; and whether they started insulin within one year of diagnosis. Those reported using glucose-lowering medication were asked to provide details of the medication.

We derived an algorithm based on previously published work using UK Biobank to identify people with or without incident T2D during follow-up (Supplementary Figure 1).[23] Exclusion criteria were applied to 274 participants with no diabetes status provided at baseline (n=68) and follow up (n=206), 688 participants who were prevalent diabetes cases (n=52 type 1 and n=636 type 2) and 104 participants who were diagnosed with T2D at an earlier age than their age at dietary assessment (these were considered prevalent rather than incidence cases). We also excluded participants who were followed for less than 12 months (n=213, including 62 people with T2D diagnosis) to minimise the potential for reverse causality bias, however, sensitivity analyses including these participants were performed and presented in Supplementary Table S1.

Covariates

Socio-demographic and lifestyle factors at baseline included age, sex, ethnicity (white/black/other), family history of T2D (yes/no), current smoking status (smoker/non-smoker), body mass index (BMI, kg/m^2), quintiles of the Index of Multiple Deprivation (IMD), and physical activity level (low/moderate/high).[24,25] We used the previously-derived physical activity variable that was based on the validated and self-administered International Physical Activity Questionnaire (IPAQ).[25] IMD is a composite measure of deprivation for each small area of the UK based on participants' postcode, and we derived IMD quintiles based on deprivation scores.[26] A separate category was introduced for those who had no available data on physical activity (13.5% missing) or IMD (2.1% missing). Less than 0.3% of the study cohort had missing data on ethnicity (0.26%), smoking status (0.01%) and BMI (0.1%). They were included in the unadjusted analyses but no 'missing' category was created for these variables due to their small sample size.

Statistical analysis

We compared participants' baseline characteristics by their quartile of ultra-processed food consumption, using the analysis of variance for continuous variables and χ^2 tests for categorical variables. We examined the shape of survival functions between quartiles of ultra-processed food consumption and between subgroups of other covariates using Kaplan-Meier plots, and assessed the equality of survival functions between subgroups using log rank tests. We used Cox proportional hazards regression models with age as the underlying time metric to estimate the hazard ratios (HRs) and their corresponding 95% confidence intervals (CIs) of the incidence of newly diagnosed T2D, separately for both exposures described above.

Participants contributed person time until the date of diagnosis for T2D or the date of their last assessment, whichever occurred first. We derived the date of diabetes diagnosis from participants' age at follow-up assessment and their self-reported age when diabetes was first diagnosed. When the age of diagnosis was not available, the mid-point between the date of follow-up assessment when diabetes was identified and the date of previous assessment was used instead. Covariates were added in a stepwise manner: Model 1 included age (timescale) and the exposure variable; Model 2 additionally included sex, ethnicity and family history of T2D; Model 3 additionally included quintiles of IMD; Model 4 additionally included current smoking status, physical activity level and total energy intake; and Model 5 additionally included baseline BMI. Potential interaction between the exposure variable and baseline BMI, physical activity level, current smoking status and Index of Multiple Deprivation were assessed but no statistically significant interactions were identified. Proportionality assumption was verified by testing Schoenfeld residuals against survival time, which revealed violations for sex and ethnicity and these variables were therefore incorporated by stratification that allows baseline hazards to differ between sex and ethnicity strata. The linearity assumption between proportion of ultra-processed food consumption and T2D incidence was confirmed by restricted cubic spline (P value for the test of linearity was 0.339).

All analyses were performed using Stata version 16.

RESULTS

Mean age of the study cohort at baseline was 55.8 years (SD=7.4), 52.9% were women and most participants were white (97.9%). The mean contribution of ultra-processed foods to the overall diet (in grams) was 22.1% (Table 1). The most commonly consumed ultra-processed foods were beverages that contributed to 8.6% of total grams per day, followed by 6.6% from bakery products and breakfast cereals and 4.4% from industrial-processed frozen/self-stable prepared meals and salty snacks (Figure 1). Compared with participants who consumed the lowest amount of ultra-processed foods (1st quartile), those with the highest consumption (4th quartile) were more likely to be younger, living in more deprived areas, physically inactive and overweight at baseline.

A total of 305 incident T2D cases were identified amongst participants during a mean follow-up of 5.4 years (116,956 person years, 260 per 100,000 person-years). Across quartiles of ultra-processed food intake (from lowest to highest) the incidence of T2D increased from 209, 216, 256, 362 per 100,000 person years, respectively.

When compared with study participants with the lowest ultra-processed food intake, a gradient of elevated risk of T2D with increasing levels of ultra-processed food intake was consistently observed in all unadjusted and adjusted regression models (p value for trend <0.001). In the fully adjusted model (Model 5), compared with the group with the lowest ultra-processed food intake (1st quartile), the hazard ratio for T2D was 1.44, 1.04-2.02 for the highest (4th) quartile of ultra-processed food consumers. The hazard ratio estimated per 10% absolute increment in ultra-processed food consumption was 1.22 (95% Confidence Interval [CI]: 1.14-1.31) in the unadjusted model (Model 1) and remained significant after all covariates were accounted for (Model 5: 1.12, 1.04-1.20) (Table 2).

The results of sensitivity analyses that included participants who were followed for less than 12 months were consistent with the main findings (Supplementary Table S1). Findings from the analyses that considered ultra-processed food consumption as a proportion of daily energy intake are presented in Supplementary Table 2. All models, except for the one when baseline BMI was adjusted for, have consistently shown a significant dose-response relationship between increased ultra-processed food consumption and risk of T2D, albeit to a lesser extent. The model with BMI adjustment, however, showed an association which was marginally significant statistically (p = 0.053).

DISCUSSION

In this UK-based cohort, participants with the highest levels of ultra-processed food consumption had a 44% increased risk of developing T2D compared with those with the lowest intake over a 5.4-year follow-up time, and the risk of incident T2D increased by 12% for every 10 percentage points increment in ultra-processed food consumption. Given the current high levels and growing consumption of ultra-processed foods in the UK and many countries worldwide,[5-9] these findings highlight an alarming threat to the health of populations.

Our findings are consistent with the only prospective study conducted in France which identified a HR of 1.13 (95% CI: 1.01-1.27) for T2D incidence per 10 percentage points increase in ultra-processed foods consumption.[14] While both the French and British cohorts were followed-up for similar times and consumed substantial amounts of ultra-processed foods (17.3% and 22.1%, respectively), categories of ultra-processed food intake largely differed between cohort participants. The most prominent ultra-processed foods consumed by the French cohort were sugary products (28%), ultra-processed fruits and vegetables (18%) and beverages (16%), whereas in the British cohort, ultra-processed foods were mainly sourced from beverages (39.1%), industrial-processed bakery products and breakfast cereals (29.9%), and industrial-processed frozen/self-stable prepared meals (19.3%). This suggests that ultra-processed foods may be associated with an increased risk of T2D incidence regardless of the specific categories of foods and beverages consumed.

There are several potential underlying mechanisms that may plausibly explain the associations observed. First, diets rich in ultra-processed foods are of poor nutritional quality. In the past decade, analyses of nationally representative food intake surveys conducted in several countries, including the UK, have consistently reported that increasing dietary share of ultra-processed foods results in higher intakes of free sugars and saturated fat, but lower content of fibre and protein in the overall diet.[3-5,8] Furthermore, research has demonstrated that ultra-processed foods have a higher glycaemic response but lower satiety potential, when compared with their less processed alternatives.[27]

The vast variety of chemical compounds that are commonly added to ultra-processed foods, formed during their manufacturing processes or released from their packaging materials are another potential mechanism. For instance, emulsifiers and carrageenan (a thickening and stabilizing agent) are common ingredients of ultra-processed foods and the former has been found to be associated with the development of metabolic syndrome and other chronic inflammatory diseases whereas the latter has been shown to be associated with increased glucose intolerance and insulin resistance.[28,29]

The concentration of bisphenol-A, present in food packaging materials and drink containers, has been found to interfere with the physiological effects of hormones such as oestrogen, androgen and thyroids and with cell pathways related to weight and glucose homeostasis, and a positive association between bisphenol-A and T2D has been previously documented.[30] Generation of neo-formed contaminants, such as acrylamide and acrolein metabolites, during the production of ultra-processed foods using high-temperature heating and extruding methods have been found associated with insulin resistance.[31,32] Polycyclic aromatic hydrocarbons that are released during the manufacturing of ultra-processed foods have also been shown associated with diabetes.[33]

The final potential mechanism is through obesity. Prospective cohort studies have shown a positive association between ultra-processed food consumption and increased risk of

overweight and obesity, including in a Spanish cohort study with 8.9 years of follow-up (adjusted hazard ratio: 1.26, 95% CI: 1.10-1.45), and in a Brazilian cohort study with 3.8 years of follow-up (relative risk: 1.20, 95% CI: 1.03-1.40), and recently in a British cohort study with 5.6 years of follow-up (adjusted hazard ratio: 1.82, 95% CI: 1.06-3.06).[1,10-12] These findings are further supported by a recent cross-over clinical trial that allocated 20 weight-stable adults into an ultra-processed diet or an unprocessed diet with matched calories, sugar, fat, fibre and macronutrients for 14 days each.[13] The authors concluded that ultra-processed food consumption contributed to a significantly higher energy intake (508±106 kcal/day) and weight gain (0.9±0.3 kg) among study participants.[13]

This research builds on an increasingly robust and consistent evidence base which suggests that continued growth in ultra-processed food consumption, especially in middle-income countries, represents an important driver of disease burden and a potential threat to health systems globally. This urges the identification and implementation of effective public health actions to counteract the growth in ultra-processed food consumption. Health authorities in Brazil, Uruguay, Ecuador and Peru have taken the lead in explicitly recommending the avoidance of ultra-processed foods in their food-based dietary guidelines.[34-37] France recently set an ambitious goal of a 20% reduction in ultra-processed food consumption by 2022.[38] Evaluation of the actions to be undertaken to meet this goal is likely to provide valuable lessons for other countries. Further research is needed to elucidate the mechanistic pathways between ultra-processed food consumption and the development of T2D.

Limitations

This study has several important limitations. First, identification of incident T2D cases was based on self-reported and nurse-interview data. Self-reported data are often limited by under-reporting and misclassification resulting in the under-detection of diabetes cases although the nurse interview included detailed questions about age at diabetes diagnosis and history of glucose-lowering medications. Second, the classification of some food items was limited by a generic description used in 24-hour recall. However, this was addressed by assigning them into the most probable food group based on published findings of common foods and drinks consumed in the UK.[5] Additionally, misclassifications in the NOVA categories would lead to a non-differential measurement error, probably bringing the results towards the null hypothesis. Third, the estimation of ultra-processed food consumption was based on one 24-hour recall instead of multiple days. Due to the inconsistencies in the timing of diabetes measurements and the administrations of 24-hour recalls, we considered the first completed 24-hour recall as the best representation of participants' dietary intake at baseline though this does not fully represent their usual consumption. Also, participants' diet may have changed during follow-up that the current analyses do not account for. Fourth, because our main analyses considered UPF consumption as the percentage of total dietary intake in grams, the effect of liquid as soft drinks could have more importance. Nevertheless, we performed additional analyses to further adjust the final statistical model for percentage of calories from sugar-sweetened beverages, and the association remained similar (adjusted hazard ratio 1.11, 95% CI: 1.02-1.21 for each 10 percentage points increment in UPF consumption). Fifth, missing data is a limitation although this was very low for smoking status, weight measurement and ethnicity. We incorporated a missing category for IMD and physical activity to capture any common effects among those with missing data. Sixth, our study cohort is healthier than the general population, with fewer smokers, lower BMI levels, and healthier dietary pattern[39] Therefore, the results may be an under-estimation the association due to a lower contrast between extreme quartiles of ultra-processed consumption, the burden of

ultra-processed food consumption in the UK should consider this. Finally, the observational nature of the study means that residual confounding cannot be completely ruled out.

Conclusion

Our findings demonstrate that ultra-processed food consumption is associated with a clinically important increased risk of T2D. Recommendations to highlight the harmful effects of ultra-processed foods in dietary guidelines, and identifying and implementing effective public health actions to reduce the growth in ultra-processed food consumption are urgently required.

Acknowledgements
This research study was conducted using the UK Biobank database (April 2019 release). The authors would like to thank the participants of the UK Biobank study. Access to the UK Biobank database was granted by the UK Biobank's Access Sub-Committee under Application Number 29239.

Ethics declarations
The UK Biobank received ethical approval from the North West Multi-Centre Research Ethics Committee (16/NW/0274). Participants provided written informed consent before enrolment in the study, which was conducted in accord with the principles of the Declaration of Helsinki.

Contributors
All authors contributed to the concept and design of this study. FR and RBL cleaned the data and RBL performed all statistical analyses and drafted the manuscript. All authors provided inputs in the interpretation of findings and edited and reviewed the manuscript. RBL is the guarantor of this work and, as such, had full access to all of the data used in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding
RBL is funded by the São Paulo Research Foundation (FAPESP) 2015/14900-9 and 2018/26660-0. FR is funded by the São Paulo Research Foundation (FAPESP) 2016/14302-7 and 2018/19820-1. This study is funded by the National Institute for Health Research (NIHR) School for Public Health Research (Grant Reference Number PD-SPH-2015). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The funders had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Competing interests

All authors report no conflict of interest.

Availability of data and materials

The data that support the findings of this study are available from UK Biobank but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission from UK Biobank.

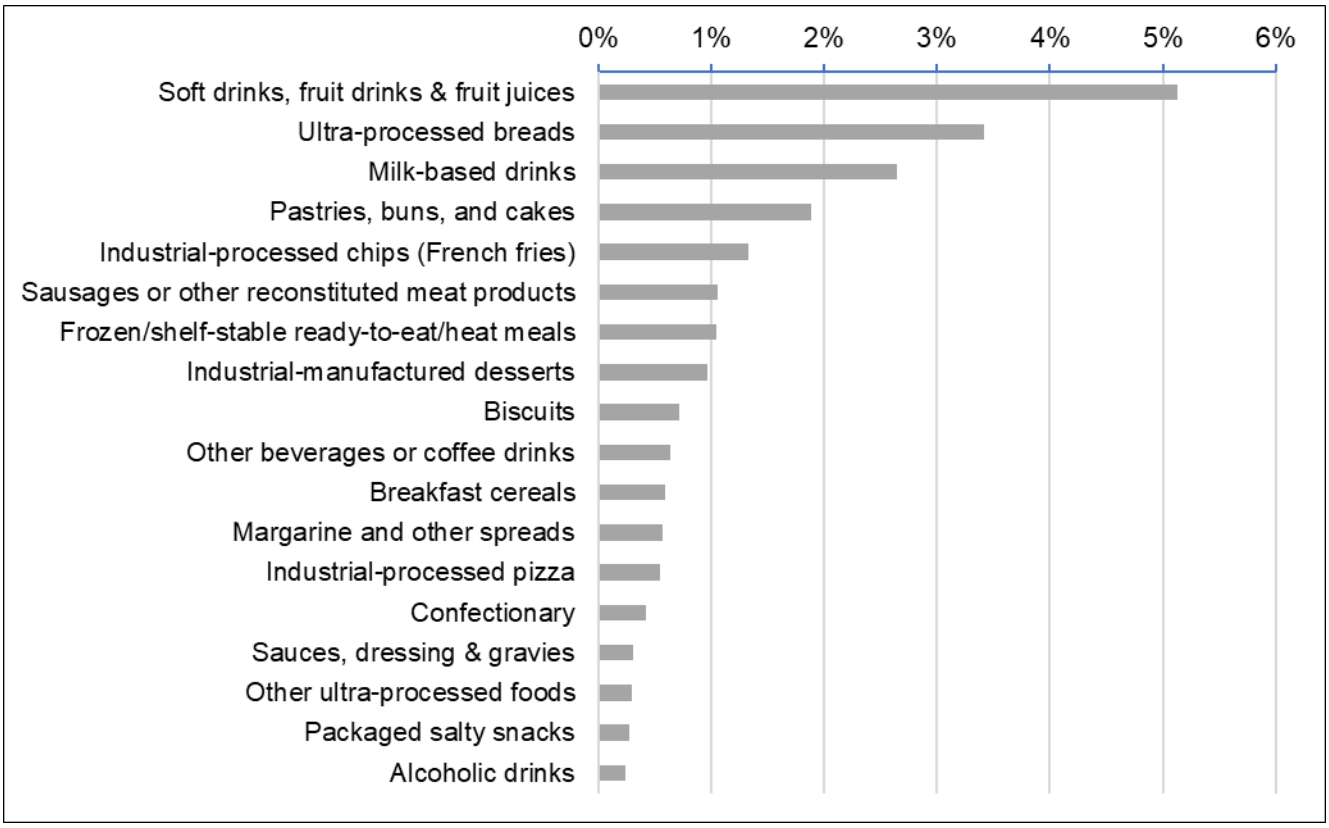
REFERENCES

1. Monteiro CA, Cannon G, Levy RB, et al. Ultra-processed foods: what they are and how to identify them. *Public health nutrition* 2019; **22**(5): 936-41.
2. Rauber F, Steele EM, Louzada MLdC, Millett C, Monteiro CA, Levy RB. Ultra-processed food consumption and indicators of obesity in the United Kingdom population (2008-2016). *PLOS ONE* 2020; **15**(5): e0232676.
3. Martinez Steele E, Popkin BM, Swinburn B, Monteiro CA. The share of ultra-processed foods and the overall nutritional quality of diets in the US: evidence from a nationally representative cross-sectional study. *Population health metrics* 2017; **15**(1): 6.
4. Moubarac JC, Batal M, Louzada ML, Martinez Steele E, Monteiro CA. Consumption of ultra-processed foods predicts diet quality in Canada. *Appetite* 2017; **108**: 512-20.
5. Rauber F, da Costa Louzada ML, Steele EM, Millett C, Monteiro CA, Levy RB. Ultra-Processed Food Consumption and Chronic Non-Communicable Diseases-Related Dietary Nutrient Profile in the UK (2008-2014). *Nutrients* 2018; **10**(5).
6. 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? *The American journal of clinical nutrition* 2015; **101**(6): 1251-62.
7. Monteiro C, Moubarac J, Cannon G, Ng SW, Popkin B. Ultra-processed products are becoming dominant in the global food system. *Obesity Reviews* 2013; **14**: 21-8.
8. Louzada M, Ricardo CZ, Steele EM, Levy RB, Cannon G, Monteiro CA. The share of ultra-processed foods determines the overall nutritional quality of diets in Brazil. *Public Health Nutr* 2018; **21**(1): 94-102.
9. Marrón-Ponce JA, Sánchez-Pimienta TG, da Costa Louzada ML, Batis C. Energy contribution of NOVA food groups and sociodemographic determinants of ultra-processed food consumption in the Mexican population. *Public Health Nutr* 2018; **21**(1): 87-93.
10. Mendonça RD, Pimenta AM, Gea A, et al. Ultraprocessed food consumption and risk of overweight and obesity: the University of Navarra Follow-Up (SUN) cohort study. *The American journal of clinical nutrition* 2016; **104**(5): 1433-40.
11. Canhada SL, Luft VC, Giatti L, et al. Ultra-processed foods, incident overweight and obesity, and longitudinal changes in weight and waist circumference: the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *Public Health Nutr* 2019: 1-11.
12. Rauber F, Chang K, Vamos EP, et al. Ultra-processed food consumption and risk of obesity: a prospective cohort study of UK Biobank. *European Journal of Nutrition (in press)* 2020.
13. Hall KD, Ayuketah A, Brychta R, et al. Ultra-Processed Diets Cause Excess Calorie Intake and Weight Gain: An Inpatient Randomized Controlled Trial of Ad Libitum Food Intake. *Cell Metabolism* 2019.
14. Srour B, Fezeu LK, Kesse-Guyot E, et al. Ultraprocessed Food Consumption and Risk of Type 2 Diabetes Among Participants of the NutriNet-Santé Prospective Cohort. *JAMA Internal Medicine* 2019.
15. Vandevijvere S, Jaacks LM, Monteiro CA, et al. Global trends in ultraprocessed food and drink product sales and their association with adult body mass index trajectories. *Obesity Reviews* 2019.
16. World Health Organization. Noncommunicable diseases country profiles 2018.
17. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine* 2015; **12**(3).
18. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018; **562**(7726): 203-9.

19. Galante J, Adamska L, Young A, et al. The acceptability of repeat Internet-based hybrid diet assessment of previous 24-h dietary intake: administration of the Oxford WebQ in UK Biobank. *British Journal of Nutrition* 2016; **115**(4): 681-6.
20. Liu B, Young H, Crowe FL, et al. Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. *Public Health Nutr* 2011; **14**(11): 1998-2005.
21. Ministry of Agriculture Fisheries and Food. Food portion sizes. HMSO London, United Kingdom; 1993.
22. Public Health England. Composition of foods integrated dataset (CoFID). 2019. https://www.gov.uk/government/publications/composition-of-foods-integrated-dataset-cofid [Last accessed: 04 Feb 2020].
23. Eastwood SV, Mathur R, Atkinson M, et al. Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. *PloS one* 2016; **11**(9).
24. UK Biobank. UK Biobank - Anthropometry. Version 1.0. 2014. http://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/Anthropometry.pdf [Last accessed: 07 Feb 2020].
25. Cassidy S, Chau JY, Catt M, Bauman A, Trenell MI. Cross-sectional study of diet, physical activity, television viewing and sleep duration in 233 110 adults from the UK Biobank; the behavioural phenotype of cardiovascular disease and type 2 diabetes. *BMJ open* 2016; **6**(3): e010038.
26. UK Biobank. UK Biobank - Index of Multiple Deprivation Scores. Version 1.0. 2018. http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/imd_baseline.pdf [Last accessed: 04 Feb 2020] (accessed 04 Feb 2020 2020).
27. 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 & function* 2016; **7**(5): 2338-46.
28. Chassaing B, Koren O, Goodrich JK, et al. Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature* 2015; **519**(7541): 92-6.
29. Bhattacharyya S, Feferman L, Tobacman JK. Carrageenan inhibits insulin signaling through GRB10-mediated decrease in Tyr (P)-IRS1 and through inflammation-induced increase in Ser (P) 307-IRS1. *Journal of Biological Chemistry* 2015; **290**(17): 10764-74.
30. Hwang S, Lim J-E, Choi Y, Jee SH. Bisphenol A exposure and type 2 diabetes mellitus risk: a meta-analysis. *BMC endocrine disorders* 2018; **18**(1): 81.
31. Lin C-Y, Lin Y-C, Kuo H-K, et al. Association among acrylamide, blood insulin, and insulin resistance in adults. *Diabetes Care* 2009; **32**(12): 2206-11.
32. Feroe AG, Attanasio R, Scinicariello F. Acrolein metabolites, diabetes and insulin resistance. *Environmental research* 2016; **148**: 1-6.
33. Stallings-Smith S, Mease A, Johnson TM, Arikawa AY. Exploring the association between polycyclic aromatic hydrocarbons and diabetes among adults in the United States. *Environmental research* 2018; **166**: 588-94.
34. Monteiro CA, Cannon G, Moubarac J-C, et al. Dietary guidelines to nourish humanity and the planet in the twenty-first century. A blueprint from Brazil. *Public Health Nutr* 2015; **18**(13): 2311-22.
35. Food and Agriculture Organization of the United Nations. Food-based dietary guidelines – Uruguay. 2016. http://www.fao.org/nutrition/education/food-dietary-guidelines/regions/uruguay/en/ [Last accessed: 09 Feb 2020].
36. Food and Agriculture Organization of the United Nations. Food-based dietary guidelines - Ecuador. 2018. http://www.fao.org/nutrition/education/food-dietary-guidelines/regions/countries/ecuador/en/ [Last accessed: 09 Feb 2020].

37. Food and Agriculture Organization of the United Nations. Food-based dietary guidelines - Peru. 2019. http://www.fao.org/nutrition/education/food-dietary-guidelines/regions/countries/peru/en/ [Last accessed: 09 Feb 2020].
38. 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. https://www.hcsp.fr/Explore.cgi/Telecharger?NomFichier=hcspa20170216_reperesalimentairesactua2017.pdf [Last accessed: 17 Mar 2020].
39. The NHS Information Centre. Health Survey for England - 2009, Trend tables. https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2009-trend-tables [Last accessed: 10 Feb 2020].

Figure 1: Sub-categories of ultra-processed foods contributed to mean weight of foods consumed



Note: Other ultra-processed foods' included chocolate/nut spread, spreadable cheese, sweeteners, meat alternatives.

Table 1: Baseline characteristics of the study cohort by sex-specific quartiles of ultra-processed food consumption derived from weight of foods consumed

		Quartile of ultra-processed food consumption				
	All participants *n=21,730*	**1st quartile (lowest intake)** *n=5,446*	**2nd quartile** *n=5,419*	**3rd quartile** *n= 5,444*	**4th quartile (highest intake)** *n=5,421*	**p-value**
Mean ultra-processed foods consumption (%)						
Relative to total grams consumed	22.1	7.7	15.4	23.6	41.9	<0.01
Total Energy, mean (SD)						
Kcal	2065.3 (679.2)	1799.3 (565.1)	2011.1 (597.0)	2170.0 (670.5)	2281.6 (767.6)	<0.01
Age, mean (SD)						
Years	55.8 (7.4)	56.6 (7.0)	56.5 (7.2)	55.9 (7.5)	54.2 (7.7)	<0.01
Sex (%)						
Women	52.9	52.8	52.9	52.9	53.0	0.996
Men	47.1	47.2	47.1	47.1	47.0	
Ethnicity (%)						
White	97.9	98.2	98.0	97.9	97.4	0.121
Black	0.8	0.7	0.8	0.8	1.1	
Others	1.3	1.1	1.2	1.3	1.5	
Family history of T2D (%)						
No	80.8	82.5	82.2	80.4	78.3	<0.01
Yes	19.2	17.5	17.8	19.6	21.7	
Index of Multiple Deprivation (%)						
Least derived quintile	19.6	20.6	20.8	19.9	17.2	<0.01
2nd quintile	19.5	20.2	19.5	19.7	18.6	
3rd quintile	19.6	20.5	19.3	19.6	19.1	
4th quintile	19.6	19.1	20.4	19.4	19.3	
Most deprived quintile	19.6	17.4	18.1	19.3	23.6	
Missing	2.1	2.3	1.9	2.0	2.1	
Current smoker (%)						
No	94.5	94.4	94.2	95.0	94.4	0.231
Yes	5.5	5.6	5.8	5.0	5.6	
Physical activity level (%)						
Low	15.7	14.2	15.2	15.3	18.0	<0.01
Moderate	36.4	36.2	36.4	37.7	35.4	
High	34.4	37.1	34.1	33.8	32.5	
Missing	13.5	12.5	14.3	13.2	14.2	
BMI mean (SD)						
Kg/m^2	26.5 (4.2)	26.1(3.8)	26.2(4.0)	26.4(4.1)	27.2(4.6)	<0.01

Table 2: Hazard ratios (95% confidence interval) for the associations between ultra-processed food consumption (derived from weight of foods consumed) and incidence of T2D, estimated by multivariable Cox proportional hazards regression.

			Quartile of ultra-processed food consumption				
	Per 10% absolute increment in ultra-processed foods intake	**p-value**	**1st quartile (lowest intake)**	**2nd quartile**	**3rd quartile**	**4th quartile (highest intake)**	**p-value for trend**
Number of cases/non-cases	*305/21,425*		*62/5,384*	*63/5,356*	*74/5,370*	*106/5,315*	
Model 1	1.22 (1.14-1.31)	<0.001	1	1.05 (0.74-1.49)	1.27 (0.91-1.79)	1.97 (1.44-2.70)	<0.001
Model 2†	1.21 (1.13-1.30)	<0.001	1	1.06 (0.74-1.50)	1.23 (0.88-1.73)	1.88 (1.37-2.59)	<0.001
Model 3	1.20 (1.12-1.29)	<0.001	1	1.04 (0.74-1.48)	1.22 (0.87-1.71)	1.85 (1.34-2.53)	<0.001
Model 4‡	1.20 (1.12-1.29)	<0.001	1	1.03 (0.73-1.48)	1.22 (0.86-1.72)	1.84 (1.32-2.55)	<0.001
Model 5§	1.12 (1.04-1.20)	<0.002	1	0.98 (0.68-1.39)	1.10 (0.76-1.55)	1.44 (1.04-2.02)	<0.028

Model 1 included age (timescale) and the exposure variable.
Model 2† = Model 1 + family history of T2D (yes/no), and stratification by sex and ethnicity.
Model 3 = Model 2 + Index of Multiple Deprivation (least derived quintile/2nd quintile/3rd quintile/4th quintile/most deprived quintile/missing).
Model 4‡ = Model 3 + physical activity level (low/moderate/high/missing), current smoking status (yes/no), and total energy intake.
Model 5§ = Model 4 + BMI continuous at baseline
* <0.3% of the participants with missing data on ethnicity, smoking status or weight measurement were omitted from subsequent regression models due to the small sample size of the 'missing' category for these variables