Author’s response to reviews

Title: Prospective association between ultra-processed food consumption and incident depressive symptoms in the French NutriNet-Santé cohort

Authors:

Moufidath ADJIBADE (m.adjibade@eren.smbh.univ-paris13.fr)
Chantal Julia (c.julia@uren.smbh.univ-paris13.fr)
Benjamin Allès (b.alles@eren.smbh.univ-paris13.fr)
Mathilde Touvier (m.touvier@eren.smbh.univ-paris13.fr)
Cédric Lemogne (cedric-lemogne@orange.fr)
Bernard Srour (b.srour@eren.smbh.univ-paris13.fr)
Serge Hercberg (s.hercberg@eren.smbh.univ-paris13.fr)
Pilar Galan (p.galan@eren.smbh.univ-paris13.fr)
Karen Assmann (k.assmann@eren.smbh.univ-paris13.fr)
Emmanuelle Kesse-Guyot (e.kesse@eren.smbh.univ-paris13.fr)

Version: 2 Date: 26 Feb 2019

Author’s response to reviews:

RESPONSE TO BMC Medicine

Manuscript No.: BMED-D-18-01380R1

Title of Paper: Prospective association between ultra-processed food consumption and incident depressive symptoms in the French NutriNet-Santé cohort

Point-to-point reply

General note

We appreciate the additional comments made by the Reviewers and believe that their consideration has substantially improved our manuscript. We have extensively revised our manuscript so as to address the points raised.
Reviewer #2:

Can the authors present the model 2 + use of antidepressants during follow-up, the baseline CES-D score and CDS score-adjusted model?

As requested by the Reviewer, we have now presented a new model (model 7) adjusted for the requested variables in model 2: the use of antidepressants during follow-up, baseline CES-D score and CDS score.

Please see below.

Table 3 Association between ultra-processed food intake and incident depressive symptoms, NutriNet-Santé study a

<table>
<thead>
<tr>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>Quartile 3</th>
<th>Quartile 4</th>
<th>P-trendContinuous b</th>
<th>P c</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPF, range</td>
<td>0%-10%</td>
<td>10%-14%</td>
<td>14%-19%</td>
<td>19%-76%</td>
<td></td>
</tr>
<tr>
<td>UPF, median (IQR)</td>
<td>7% (3%)</td>
<td>12% (2%)</td>
<td>16% (2%)</td>
<td>23% (8%)</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>6,682</td>
<td>6,683</td>
<td>6,683</td>
<td>6,682</td>
<td>26,730</td>
</tr>
<tr>
<td>Number of cases</td>
<td>491</td>
<td>459</td>
<td>557</td>
<td>714</td>
<td>2,221</td>
</tr>
<tr>
<td>Person years</td>
<td>21,597</td>
<td>21,097</td>
<td>20,468</td>
<td>19,918</td>
<td>83,080</td>
</tr>
<tr>
<td>Model 1d</td>
<td>1 (ref)</td>
<td>0.90 (0.79; 1.02)</td>
<td>1.07 (0.94; 1.21)</td>
<td>1.31 (1.16; 1.47)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 2e</td>
<td>1 (ref)</td>
<td>0.91 (0.80; 1.04)</td>
<td>1.09 (0.96; 1.23)</td>
<td>1.30 (1.15; 1.47)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 3f</td>
<td>1 (ref)</td>
<td>0.91 (0.80; 1.04)</td>
<td>1.08 (0.95; 1.23)</td>
<td>1.29 (1.13; 1.47)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 4g</td>
<td>1 (ref)</td>
<td>0.92 (0.81; 1.04)</td>
<td>1.09 (0.97; 1.24)</td>
<td>1.31 (1.16; 1.48)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 5h</td>
<td>1 (ref)</td>
<td>0.88 (0.77; 1.00)</td>
<td>1.00 (0.88; 1.13)</td>
<td>1.13 (1.00; 1.28)</td>
<td>0.01</td>
</tr>
<tr>
<td>Model 6i</td>
<td>1 (ref)</td>
<td>0.88 (0.78; 1.00)</td>
<td>1.06 (0.94; 1.20)</td>
<td>1.27 (1.13; 1.44)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Model 7j</td>
<td>1 (ref)</td>
<td>0.86 (0.76; 0.98)</td>
<td>1.00 (0.88; 1.13)</td>
<td>1.13 (1.00; 1.28)</td>
<td>0.01</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>1.15 (1.09; 1.21)</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Values are hazard ratios (95% confidence intervals). CDS Cognitive Difficulties Scale; CES-D Center for Epidemiologic Studies Depression Scale; IQR Interquartile Range; UPF Proportion of ultra-processed food intake

b Hazard ratios for 10% increase in the proportion of ultra-processed food intake

c P for continuous variable

d Adjusted for age, sex and body mass index

e Adjusted for all variables in model 1 + marital status, educational level, occupational categories, household income per consumption unit, residential area, number of 24h-dietary records, inclusion month, energy intake without alcohol, alcohol intake, smoking status and physical activity (main model)

f Adjusted for all variables in model 2 + dietary patterns derived from the factor analysis (“Healthy” and “Western” dietary pattern) and intakes of lipids, sodium, and carbohydrates

g Adjusted for all variables in model 2 + health events during follow-up (cancer, Type 2 diabetes, hypertension and cardiovascular events)

h Adjusted for all variables in model 2 + use of antidepressants during follow-up and baseline CES-D score

i Adjusted for all variables in model 2 + CDS score
j Adjusted for all variables in model 2 + use of antidepressants during follow-up, baseline CES-D score and CDS score

The methods section has been modified accordingly.

“The first model was adjusted for age, sex, and BMI (continuous variable). The second model (main model) was additionally adjusted for marital status, educational level, occupational categories, monthly household income per consumption unit, residential area, energy intake without alcohol, number of 24h records and inclusion month, smoking status, alcohol consumption and physical activity. Five additional models were also performed to account for a) PCA-extracted dietary patterns and intake of carbohydrates, lipids and salt); b) health events occurring during follow-up (cancer, type 2 diabetes and cardiovascular diseases); c) baseline CES-D score (continuous variable) and use of antidepressants during the follow-up; d) CDS score (continuous variable) and e) baseline CES-D score (continuous variable), use of antidepressants during the follow-up and CDS score (continuous variable).” (Methods section - Statistical analysis, page 10, lines 227-232)

Editorial Board Member's advice:

1) The abstract states that participants were "symptom-free". I'm not sure this adequately reflects the inclusion criteria though: < 17 for men and < 23 for women indicate that there is likely no probable diagnosis, but not that there are no symptoms.

We entirely agree with this editorial Board Member’s comment that the CES-D is not used to diagnose depression, even if the validated cut-off values of this scale can be used to classify individuals as having or not depressive symptoms [1, 2]. Note that in this study, we assessed depressive symptomatology and not the diagnosis of depression.

The sentence has been corrected as follows:

“The sample included 20,380 women and 6,350 men (aged 18-86 years) without depressive symptoms at the first Center for Epidemiologic Studies Depression Scale (CES-D) measurement, using the French validated cut-offs (CES-D score ≥17 for men and ≥23 for women).” (Abstract - Methods, page 2, lines 30-32)

2) The abstract states "body max index"; I assume this should be body mass index.

Yes, this has been corrected.
3) On p.6, the authors describe the use of PCA, but in the remainder of the section and other sections, talk about factors and factor loadings. Should these not be components in the case of PCA? This seems an important distinction due to the differences in how PCA (formative latent variable model) and FA (reflective latent variable models) are estimated and deal with measurement error.

We entirely agree with the reviewer. Thus, throughout manuscript, “factors” has been replaced by “principal components”. Also, “factor loading” was replaced by “loading coefficients” or “individual pattern score”.

The different sentences now read as follows:

“The number of dietary patterns retained was determined according to Cattel’s Scree plots and the interpretability of the principal components. Food groups with absolute loading coefficients > 0.3 were considered to be strongly associated with a pattern and an individual pattern score was calculated by summing the intake of the 22 food groups, weighted by their loading coefficients. The first two dietary patterns accounted for about 18% of the initial variance (Supplemental Table 1). The first principal component, corresponding to a” healthy” dietary pattern, was strongly and positively correlated with intake of whole grains, olive oil, vegetables and fruit. The second principal component, corresponding to a “western” dietary pattern, was strongly correlated with refined grains, potatoes, meat and alcoholic beverages.” (Methods section - Dietary data, page 6, lines 140-149)

And

“Similarly, in the stratified analyses according to sex, age (using the median value), BMI (<25 vs. ≥ 25), comorbid conditions and individual pattern score for the « Healthy » pattern (using the sex-specific median value as cut-off: 0.04 for men and -0.11 for women), stronger associations were observed in the subgroups of women, participants aged ≥ 49 years, participants with a chronic disease, participants with a BMI value ≥ 25, and participants with a lower value for the individual score of the « Healthy » pattern, compared to their respective counterparts (Figure 2). However, the estimated HR were similar in the subgroups and no significant interaction was observed (P value for the interaction term was 0.49 for sex, 0.46 for age, 0.59 for BMI, 0.88 for comorbidities and 0.18 for the individual score of the « healthy » dietary pattern). In addition, in the stratified analyses according to the baseline CES-D score (using the sex-specific median value as cut-off: 6 for men and 8 for women) the association between %UPF and the risk of depressive symptoms was significant only among participants with a higher baseline CES-D score (P value for the interaction term was 0.13).” (Results section, page 13, lines 293-306)

4) Do I understand correctly that the first 2 factors/components only accounted for 18% of the variance? This seems extremely low compared to what I'm used to in the literature on variance
decomposition (both for PCA and FA), and would imply that the authors’ models leave over 80% of the variance unmodeled, no?

We agree with the reviewer that the variance explained by the first 2 components is low, but this magnitude is generally observed in epidemiological studies, in particular when they are conducted in very large sample and aim to identify dietary patterns as an indication of the overall quality of the diet [3–6].

5) I cannot find how the authors control for multiple testing (p. 10 bottom where they say they conducted multiple analysis per food group). Would that not be worth considering to balance out type I vs type II error?

As suggested, as these variables are correlated, we have now controlled the False Discovery Rate due to multiple testing using the Benjamini-Hochberg procedure [7].

Please see below.

Table 4 Association between the % of ultra-processed in each food group and incident depressive symptoms, NutriNet-Santé study, N=26,730

<table>
<thead>
<tr>
<th>Quartile 1</th>
<th>Quartile 2</th>
<th>P-trend b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starchy foods</td>
<td>1 (ref) 0.97 (0.86; 1.10)</td>
<td>0.97 (0.86; 1.10)</td>
</tr>
<tr>
<td>Fruit and vegetables</td>
<td>1 (ref) 0.92 (0.81; 1.03)</td>
<td>0.97 (0.86; 1.10)</td>
</tr>
<tr>
<td>Meat, fish, eggs</td>
<td>1 (ref) 1.08 (0.96; 1.22)</td>
<td>0.97 (0.86; 1.10)</td>
</tr>
<tr>
<td>Beverages</td>
<td>1 (ref) 1.19 (0.91; 1.54)</td>
<td>1.00 (0.89; 1.12)</td>
</tr>
<tr>
<td>Dairy products</td>
<td>1 (ref) 1.03 (0.91; 1.16)</td>
<td>1.06 (0.94; 1.20)</td>
</tr>
<tr>
<td>Fatty / sweet products</td>
<td>1 (ref) 1.02 (0.90; 1.16)</td>
<td>1.05 (0.93; 1.18)</td>
</tr>
<tr>
<td>Snacks</td>
<td>1 (ref) 0.97 (0.85; 1.10)</td>
<td>1.10 (0.98; 1.25)</td>
</tr>
<tr>
<td>Sauces / added fats</td>
<td>1 (ref) 1.05 (0.93; 1.19)</td>
<td>0.96 (0.85; 1.09)</td>
</tr>
</tbody>
</table>
Values are hazard ratios (95% confidence intervals) and linear trend tests across the quartiles were assessed by modelling the quartiles of %UPF as ordinal variables. UPF Proportion of ultra-processed food intake.

a Models are adjusted for age, sex, marital status, educational level, occupational categories, household income per consumption unit, residential area, energy intake without alcohol, number of 24h-dietary records, inclusion month, smoking status, physical activity, body mass index, health events during follow-up (cancer, Type 2 diabetes, hypertension and cardiovascular events) and quantity of the equivalent food group

b Corrected using the Benjamini-Hochberg procedure

The different sections have been modified accordingly.

Abstract:
“Considering %UPF in food groups, the association was significant only for beverages and sauces or added fats.” (Abstract - Results, page 2, lines 42-43)

Methods section:
“For this analysis, models were further adjusted for the intake of the considered food group. To account for the multiple testing, False Discovery Rate-corrected P values were estimated using the Benjamini-Hochberg procedure [7].” (Methods section - Statistical analysis, page 10, lines 240-241)

And results section:
“The associations between %UPF in each food group and the risk of depressive symptoms are presented in Table 4. A significantly increased risk was observed for high %UPF in beverages and sauces or added fats.” (Results section, page 13, lines 307-310)
References


