Reviewer’s report

Title: Association between diet quality, dietary patterns and cardiometabolic health in Australian adults: a cross-sectional study

Version: 0 Date: 01 Nov 2017

Reviewer: Melanie Deschasaux

Reviewer’s report:

This is an interesting paper reporting on the association between diet and markers of cardiometabolic health. Strengths of this study pertains to the use of several methodological approaches to characterize the "whole diet" of participants including a rather novel RRR approach to derive a posteriori dietary patterns and an a priori approach using the Dietary Guideline index (DGI), both approaches being used in a same representative population.

Please find my comments as follows:

Abstract

-L3: please provide a short explanation of the differences between diet quality and dietary pattern when first mentioned in the abstract (improve understanding for readers)

-L10: "'the' Dietary Guideline Index (DGI)"

-L11-12: "using fiber density, SFA: PUFA and total sugars intake as intermediate markers." => not clear, seems to refer to DGI too

-L13: define DP abbreviation

-L31: "casual" should be "causal"? (same goes L454)

Introduction

-L37-38: "The underlying biological pathways through which these diseases are mediated include markers of poor cardiometabolic health" => this sentence is not clear.

-L41: references [3,4] provided to justify the need for studies on diet and references [3,5] associated with the statement "research to date" are from early 2000s. More recent references would be appreciated

-L45: references [7-9]. Please also provide references from another group
-L46: by "cluster analysis", did you mean dietary patterns derived from factor analysis or principal component analysis? Then cluster is not the appropriate term here

-L62-63: "RRR may better predict risk of disease than purely data-driven DP methodologies" => do you have a reference for that? Is it still true when the dependent variables in the RRR are nutrients and not biomarkers or outcomes as in your study?

Methods

-The method section is very long with a lot of details on the setting of the NHS, NNPAS and NHMS studies or the procedures used to collect data. Please select what is really relevant to the present manuscript. Some details may not be necessary in this manuscript and should be either replaced by a reference to a previous publication or put to some extent in supplementary material

-L80: were these 14,363 private dwellings included in the 15,565 adequate respondents of the NHS? Overall it is not clear how the samples from the NHS, NNPAS and NHMS relate to one another.

-L91: there was an important selection of your sample (from 12,153 to 2121). How representative of the Australian population was your analytical sample?

-L106: was there a correction applied for weight if people did not take out their shoes and heavy clothing?

-L112: please define ABS

-L112-113: how was LDL cholesterol estimated for people with TAG>4.5mmol/L then?

-L114: "without the need for fasting" => only participants that were fasting were included in this study to have data on LDL, TAG and glucose

-L118-119: "Impaired fasting plasma glucose was defined as > 6.0mmol/L and <7.0 mmol/L" => what about participants with fasting plasma glucose >7.0 mmol/L?

-L125-126: "Data on anti-hypertensive and lipid lowering medication were not recorded" => this was mentioned as a limitation in the discussion but how do you think this could have affected your results?

-L129-132: the method to derive the cardiometabolic risk score indicates that this score is actually dependent on your population since you use the mean and sd to normalize your values, then the way people are classified may depend on the overall health status of your population and not on absolute criteria. Why not use absolute thresholds to define "abnormal levels" and derive a risk score. Why not use the metabolic syndrome outcome which gather multiple cardiometabolic risk factors?
- L151-154: the description of the DGI score could be more detailed for some items that may appear as surprising or unclear (e.g. cereals without distinction between refined/whole grains or meat in "recommended" components and unsaturated fat in "discouraged", "extra sugar").

- L191: what do you mean by "standardised food group intakes"?

- L196: no description of DP3?

- L206: "currently on a diet to lose weight" is duplicate

- L228-229: "dieting" and "atyypical dietary intake on day of reporting" => how were these variables distributed among normal weight and overweight/obese people? Among healthy/unhealthy people? There may be a social desirability bias in these variables especially for the "atyypical dietary intake" and these variables may act as a proxy for your outcomes. How were your results without including these variables? Or without including participants that reported an atypical day?

- L234-235: were the weightings applied to correct for the differences between the distribution of sociodemographic characteristics in your sample compared to the Australian population?

- Did you test models in which your outcome variable would be "abnormal value, yes/no" for the studied biomarkers instead of the continuous value?

Results

- L242-244: the comparison should be made between the analytical sample and the excluded participants

- L246: should be high LDL-cholesterol

- L246: how many participants had a plasma glucose above 7?

- L254-256: even if significant, some correlations were very small and don't seem very relevant to characterize your DP (e.g. DP1 and inverse correlation with total sugars, DP2 and positive correlation with fiber density or SFA: PUFA)

- Table 1: the percentage explained in food intakes were very small and could deserve some comments (normal for RR derived DP?). How do the explained variation in responses (total) and in each individual response relate to one another? For example DP3 explains only 7.99% in total but have higher explanatory power for each component compared to DP1 and 2.

- Table 2: is 448 for whole grain bread in T2 (DP1) the correct value?

- L279: "were more physically active"
-I would like to see a comparison between DGI scores and DP1 DP2 scores (for example, scores for DP1 and DP2 in tertiles of DGI and maybe some information about the correlation/% explanation of the DGI to the response variables used in RRR (Fibre density, SFA: PUFA, Sugar))

Discussion

-L352: "healthier" DPs => do you consider that DP2 is a healthy DP? DP2 is positively correlated with sugars and seem to have contradictory effects on health with lower HDL-C (less healthy) and lower DBP (healthier). Please discuss these results as well.

-L358: why "other"?

-L357-358: "Our findings also suggest that there may be differing mechanisms through which dietary components influence other markers of cardiometabolic health." => please elaborate on this. It is not very clear how your findings allow to conclude this (also found in the abstract, l403-404 and l449-450)

-L369-370: "The inconsistency of associations with markers of cardiometabolic health may be partly attributable to differences in the ethnicity of the sample population" => what do you mean by that? What is the ethnic diversity in your sample? Was the ethnicity of your sample different from previous studies? (of note that you observed similar results with a study performed on Hispanic adults). Other hypotheses for the inconsistencies may also be confounding due to the cross-sectional design with differences between countries in metabolic health surveillance and thus awareness of metabolic abnormalities, diet recommendations following detection of abnormalities etc.

-L376-377: I agree that few studies have used RRR methodology but you can still compare the DP you derived from RRR to other DP obtained with unconstrained methods, especially since DP1 can be compared to "healthy" pattern observed in most studies

-L377-380: this is not clear. What was used in other RRR studies? If other studies used cardiometabolic markers as response variables then an association between their RRR-derived DP and the same cardiometabolic markers would just be a confirmation. In addition, what really matters is the resulting DPs and the food characterizing them.

-L391: even if DP2 was positively correlated with all response variable, the correlation was stronger with sugars

-L396: not clear what is counterintuitive here

-L403-407: this paragraph is unclear. As already mentioned, in my opinion your results does not really allow to conclude about "differing mechanisms" as there was not really an evidence for association of the same type of foods with completely different outcomes in your study. I agree that food contains nutrients that may have differential associations with cardiometabolic health
but that does not really justify a whole diet approach. The whole diet approach is rather justified by the fact that foods are not consumed individually and thus some interaction between nutrients consumed together may occur.

-L406-407: this sentence is not clear + ref 51 is about cardiovascular outcomes and not CVD risk markers

-L413: is your sample still representative after the selection you made?

-L417: "data not shown" => these data are found in supplementary table 4

-L418: you used 2 methodologies on the same population which allow comparison but I would have liked to see more direct comparison between the DGI and the RRR DPs

-L425-426: do you know if people were aware of their cardiometabolic health?

-L431: did all participants provide their 24h recall at the same period of the year? Would a "season" effect be relevant for food type intakes in the Australian population?

-L436-438: This sentence gives the impression that no prospective studies already investigated diet quality in relation to cardiometabolic health or diabetes incidence, which is not accurate. Please revise this sentence.

-L440-445: Overall, the use of RRR methodology is original and provide additional insights compared to classic unconstrained approaches. However, I am not sure how generalizable the results are in a context of policy development since the DP derived from RRR are constrained with specific objectives (in your case specific nutrients) which makes the interpretation difficult in a broader public health context. In addition, the interpretation of the DPs is not easy: DP don't really look like "normal" diets people can identify to. Maybe some discussion about the interpretation of the DPs should be added

-L444: "evaluate"

-L448: "healthier DPs" => DP1 may be considered as a healthy DP but DP2 is less interpretable

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal