Reviewer’s report

Title: Food purchase patterns: empirical identification and analysis of their association with diet quality, socio-economic factors, and attitudes

Version: 0 Date: 07 Mar 2017

Reviewer: Petra Ryden

Reviewer's report:

Thank you for truly interesting reading. I believe your study provides highly valuable contribution to the knowledge of food purchase patterns and health. It is thoroughly conducted and I am impressed with the quality of the data - well done!

I only have some minor comments and questions about your manuscript, which, if considered, I believe will improve the manuscript even further.

General comments:

* I believe you need to explain more about the food grouping. I cannot quite grasp the construction of the groups, not even by looking in your ref 20. For example - how come juice is grouped with vegetables/salad? What is the connection between brown bread and potatoes? Eggs together with sausages and cream? And the subgroups in Fats and oils - why section vegetable oils in several subgroups, and what is the difference between these (all vegetable) oils and 'vegetable fat' in the Lard, vegetable fat group? Oilseed is - to the best of my knowledge - an umbrella term for several different crops used to produce oils.

* You haven't stated a significance level in the methods section. In the results, there are mentionings of significance level at some points - but this feels like you've selected level for each test. And also, it leaves me wondering what level is set where this isn't stated?

* Have you considered calculating effect size on your comparative tests? You show really low p-values (the majority is stated as 0,000 - please change to p<0,001) when differences are very small - not surprising as you have so many participants. Effect size would be valuable to interpret the impact of significant differences.

* I cannot find drinks other than milk/dairy products and juices in your tables and text. How about other drinks? Sodas etc. Were they included? And if so, which food group were they allocated to? Diet sodas?
Minor

Line 149: This is the first reference to a table - shouldn't they be numbered consequently starting from 1? Thus, this would reference to table 1

Line 168-9: I would like an explanation to how these standardized quantities were calculated.

Line 219: Table 2. To me it makes no sense to present data as mean and SD when data are dichotomous or "number of persons...". It would be so much easier to understand if values were given as number/percentages.

How should we interpret SD when -1SD is below zero? It signals that distribution is really skewed when in fact no measured values can be below zero. If it was up to me, I would consider using median and percentiles instead (given that data is quantitative).

Line 233: I find it a bit confusing to state "the first pattern..." and finishing with the given name of that pattern. Consider 'flipping' the phrasing so that the name is given in the beginning of each description. I found myself leaping to the table to find what "the first pattern" was.

Line 246: Table 3. I would appreciate a footnote or similar explaining why some figures are bold (of course it can be figured out, but better if you state it ☐)

The food group "Meat (higher in fat)" - could you find another wording than higher? Maybe 'moderate fat content'? If I understand correctly, it is an in-between meat group in regards of fat content.

Line 294-5: It is unclear to me what a 'given household size' is.

Line 304: I cannot find a 'Figure 2' in the manuscript? Neither a 'Figure 1'…

Line 312-3: You state that that natural pattern uses supplements - but this result is not significant if your significance level is 0,01. However, it is unclear what your stated level of significance is, as commented above.

Line 319: Table 6. I would like an explanation (footnote maybe) to the "Age 2/100".

Line 344 and 368: This may be due to differences between countries, but in my country margarine is not necessarily lower in fat content than butter.

Line 398: "Households with a higher adherence to the natural and thus healthier..."

The 'thus' implies that natural = healthy which is a generalization that cannot be made. Please rephrase

Line 404-6: I think that one explanation to this controversy lies within the 'natural' thinking. Sugar is 'natural' whilst artificial sweeteners aren't - that's the general view in society, I believe.
So - if you want to have the sweet taste but want to eat natural - you need to go with sugar. This might be extra pronounced in household with kids, not so easy to refuse the sweet taste for the children.

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

An article of importance in its field

**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