Reviewer's report

Title: Food patterns reflect more than just eating habits - bias in studies on disease etiology is avoided by appropriate exclusions

Version: 2 Date: 21 May 2010

Reviewer: Elisabet Wirfält

Reviewer's report:

MAJOR COMPULSORY REVISIONS

Overall points
1. The research question is well formulated, and addresses an issue of crucial importance to nutrition epidemiology.
2. The Method section needs further clarification (see below).
3. The description of how the original data were edited needs some clarification.
4. Reporting standards are ok.
5. The Discussion and Conclusions needs further clarification (see below).
6. The title could be clearer; maybe along these lines: "Mis-reporting and previous health status may seriously bias the association between food patterns and disease". The Abstract also needs some clarification.
7. Writing is acceptable.

ABSTRACT

Results.
The observed associations could be more clearly described. Avoid the term "normal energy reporters". For instance the text could read: ... "However, when limiting analysis to individuals without previous ill-health and with adequate energy intake reports a protective association was seen with the Fruit and Vegetable pattern."

Conclusion.
This section is ambiguous. What does this sentence mean? ..."singling out the diet as a primary cause for future health problems".... Isn't it one finding of the study that diets high in fruits and vegetables are associated with health protection? The paper also suggests that reported food patterns may not only reflect current and long-term (unchanged) eating habits, but possibly also recent changes eating habits, increased knowledge of "healthy diets" (i.e., not what actually is eaten), active weight reduction (i.e., dietary restriction), or pure mis-reporting of eating habits. Most of these factors likely result in erroneous exposure classification, which cause attenuation of associations between "usual diet" and disease in epidemiological studies.
The conclusion should point out that scrutiny and care is important in cross-sectional studies, but since mis-reporting causes mis-classification of dietary intakes and attenuated associations, efforts to avoid the undue influence of mis-reports are crucial in all nutrition epidemiology studies.

BACKGROUND

1st paragraph 1st sentence

Most studies using food patterns to examine all-cause mortality and CVD have been prospective studies.

2nd and 3rd paragraph.

The authors are aware of the problems with reverse causality, but the wording needs to be clearer; the sentence... “dietary intakes and measurement of health outcomes should be separated in time” ... should indicate that diet needs to be assessed prospectively in epidemiological studies.

METHODS

Study base

This paragraph clearly needs to state that the current study is a cross-sectional study, using data collected at one point in time from each individual (at the first visit to the study center).

Food intake measurements

The food items obtained with the FFQ and the corresponding food groups (i.e., the original 84 items, the subsequent 66 items, and lastly the 35 food groups used in analysis) need to be described in more detail.

Validity coefficients for the foods and nutrients of major importance in the current study need to be presented.

A fuller description of the procedure identifying mis-reporting of energy is required; a crucial section of this paper. As the text reads now it is not clear to the reader how mis-reporting was defined; how low-energy-reporting, versus adequate-energy-reporting, was defined and determined? And how many individuals were classified in either category?

The Method section also needs to give a fuller description of the cluster analysis procedure. It is unclear to the reader how the cluster analysis was performed; if the procedure was sound. The description of each cluster is quite limited (i.e., what else than Tea and ice cream characterized that cluster?). Since a K-means methodology was used, consumption patterns are likely similar across clusters for a number of foods. These foods also need to be commented on.

Subjects

How was “acceptable data” defined? Which individuals (and How many) were not included due to un-acceptable data?
Statistical measures

1st paragraph
Avoid the term "normal" reporters, rather use "adequate".

The data was collected during a 13 year period (from 1992 to 2005). It is plausible that data collection, end-point ascertainment as well as lifestyle habits (general trends) varied during this period. How was the time span accounted for in data analysis? Or in the interpretation of results?

RESULTS

1st paragraph.
Table 1 is difficult to interpret, and it is unclear which comparisons the p-values presented in the text refer to.

When eye-balling the table it reveals that, overall, 62% of the women and 63% of the men were classified as Low Energy Reporters (LER). In women the proportion was higher in the F&V food pattern, where 77% were LER. In contrast, 50% were classified as LER in the T&IC food pattern. In men the proportion of LER in the F&V food pattern was 68% (lower than for women). However, in men the proportion of LER was 74% in the High fat food pattern; in women the proportion of LER in the corresponding pattern was 59%.

Why was the proportion of LERs so high in men in the High fat pattern? Was this difference not obvious in the statistical tests?

DISCUSSION

The 1st paragraph needs to indicate that the current study has a cross-sectional design that may generate hypotheses, but does not allow any conclusions on causality.

2nd paragraph.
How were clusters constructed? Which approach and precautions were taken in the current study to avoid subjectivity when constructing clusters? Fuller description needed in Methods.

3rd paragraph.
How was Low energy reporters identified? A fuller description needed in Methods.

The reported distribution of Low Energy reporters across food pattern groups is different in the Result section compared with the Discussion section (see above).

4th paragraph.
Since the current study is not the first study reporting different food choices or dietary intakes in mis-reporters of energy intakes, or dieters or those trying to lose weight, or have changed their food habits due to ill-health, the authors should compare their results with previous reports.
5th paragraph.
The sentence starting "In the light of ......" appear speculative. Also, the last sentence of this paragraph does not seem to report the data from Table 1 correctly. In men the High fat food pattern clearly had a high proportion of LERs (see above).

8th paragraph.
Given that the authors find different food patterns by gender, and that such food choice differences have been reported in other studies, it is surprising that the current results are not compared with similar findings within Sweden and in other countries.

Each of the papers cited for not stratifying by gender obviously had specific reasons for not doing so. Kesse-Guyot et al combined their data set (rather than stratifying by gender) because patterns were similar. Baily et al used a very small sample and specifically looked at underreporting; the small sample obviously made gender stratification unwise. The methodological study by Wirfält et al noted that "healthy" food patterns were associated with higher prevalence of both food habit change in the past and obesity. A subsequent study examined associations between food pattern clusters and the metabolic syndrome; clear gender differences were seen across patterns, an important finding that may not have been visible with gender stratification when constructing food patterns.

Since the focus of the current paper is the undue influence of under-reporting and other confounders on associations with health outcome, authors need to compare their findings with others that report similar problems. Why not cite the work by Baily et al and Wirfält et al? In order to locate important references, the authors may need to search for references in a wider range of journals.

9th paragraph.
Although, reverse causality is especially important in cross-sectional studies, dietary mis-reporting is also a major problem in all epidemiological studies, because changes in food habits, or incorrect reports, will likely cause misclassification of dietary exposures, which contribute to attenuated associations. Therefore authors should discuss the crucial importance of recognizing that mis-reporting and previous ill health, are major threats to the validity of all nutrition epidemiology studies.

10th paragraph, last sentence.
The suggested subject for future studies may not be realistic. A major reason for conducting food pattern studies is the co-variation of nutrients and food components and the difficulties in separating out individual effects of single food components.

CONCLUSIONS
The Conclusions are unclear, especially the third and fourth sentences. Also see comments regarding the Abstract above.
This paragraph needs to reflect that a cross sectional study is presented, and that such designs may generate hypotheses, but do not allow judgments on causality. Most importantly, the conclusion should emphasize that in order to interpret observed dietary associations correctly, it is crucial for epidemiological studies to identify and examine carefully confounders (like energy mis-reporting and previous health history) that cause biases and attenuated associations in nutrition epidemiology.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.