Reviewer's report

Title: Planning dietary improvements without additional costs for low-income individuals in Brazil: linear programming optimization as a tool for public policy in nutrition and health

Version: 0  Date: 27 Feb 2019

Reviewer: Sian Robinson

Reviewer's report:

This paper describes the use of linear programming optimization to model the dietary changes needed among low-income individuals in Brazil to enable nutrient requirements to be met, and at minimal cost. It is an important question to address but in its present form the paper is quite challenging for the non-specialist reader to understand - and in some places there is a need for additional detail. My comments are set out below.

General comments:

1. The term 'macro-stratum' is not meaningful without reading the Methods section (eg Abstract). As these strata represent groups defined by geography (state and within-state areas) could the term be changed? This would avoid confusion with 'household' and 'sampling' strata also described in the paper.

2. The overlap between HBS and NDS needs clear explanation. Was NDS a sub-sample of HBS and if so, are the dietary data analysed for this paper restricted only to individuals in NDS? This is difficult to understand as dietary data were available for individuals aged 10 years or older (line 84) but Table 2 shows 24% below 10 years. Line 97 suggests that the linear optimization models are based on dietary data from 18,139 individuals - the number of people studied needs to be explicit throughout, stated in the Abstract, and represented in the Tables. It would be clearer if Table 2 only represented the individuals with modelled dietary data.

3. The dietary data are modelled against recommended cut-offs that were averaged for sex/age groups (p8). Is this justified in a group that ranges from children aged 10 to adults over 60 years of age, and includes men and women? With such a large dataset - could subgroups be considered to avoid the need for weighting (eg children and adults)? My concern is that messages regarding the dietary changes needed to meet nutrient requirements might differ between these subgroups.
4. The separation of cost/no-cost models is straightforward for the reader but the different models considered in terms of 'levels of acceptability' are not. The definition of these levels (p11) needs clearer explanation: 10th, 70th, 80th and 90th percentiles of consumption of individual foods were used but the mapping to these levels is not clear. It would be simpler for the reader if the groups were labelled in a way that would indicate their derivation without reference to the Methods.

5. Detail of the dietary data collection should be included in the paper. It was not possible to evaluate the significance of the information presented without this - but I could not find the cited paper either through PubMed or via the journal site (ref 21). Line 84 states that food consumption data were collected from individuals; detail is needed of who they were and/or how selected from the households studied (line 81). Dietary data always have limitations - these need to be considered in the Discussion.

6. The use of individual vs household data needs to be made clear in the paper. An example is Table 2 where number of household members is reported together with per capita income - but it is not clear what the significance of this information is. Were a number of individuals per household studied (ie all had dietary assessments) and if so, is the per capita income their own or given for the head of household? This also relates to point 2. above.

7. A key point from these analyses is that it was not possible to model a diet that met all the nutrient thresholds. Although this is commented on in the Discussion, if no model achieves recommended intakes, it questions the use of the RDA thresholds to consider the nutritional value of the Brazilian diet. It needs to be explicit throughout the paper that the stage 2 models were used for the data presented and a clearer message that no model met all thresholds (Abstract), such that the focus is on dietary improvement.
Specific points:

* Line 71 - please amend to clarify meaning
* Line 73/Title - this is confusing as half the models allowed for cost increases - please amend to clarify
* Line 99 - explanation is needed of use of sampling weights
* Line 118 - were costs per 100g then multiplied by reported consumption to assess cost per portion consumed?
* Line 127 - 'acceptable range of food contents' needs to be explained
* Line 139 - 'macrostratum m' needs explanation
* Line 186 - why were mean food intakes derived for 550 strata?
* Line 200 - 'five macro-regions' needs explanation
* Table 2 needs to specify number of individuals studied and would be easier to understand if restricted to individuals with dietary data; years of schooling may be unhelpful in a group that includes children
* Tables 3 and 4 should specify number of individuals studied
* Figure 1 - it is difficult to map the data from Table 3 to the figure - for example the cost-constrained flexible model shows mean vitamin C intakes increasing from 44.5mg (observed) to 105.8mg - but the figure shows median increase of about 60%? This suggests that median (IQR) data should be presented rather than means in Table 3?

* To help the reader understand the dietary changes needed to map against the increases in nutrient intakes, it would be helpful to include a figure showing food group intakes that is equivalent to Figure 1

* Discussion - it is a limitation that these food intake data were collected 10 years ago; comment is needed on the relevance of the findings now and implications for costs given in the paper.
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:

Needs some language corrections before being published

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