Author’s response to reviews

Title: Importance of Details in Food Descriptions in Estimating Population Nutrient Intake Distributions

Authors:

Liangzi Zhang (liangzi.zhang@wur.nl)
Anouk Geelen (anouk.geelen@wur.nl)
Hendriek Boshuizen (hendriek.boshuizen@wur.nl)
José Ferreira (jose.ferreira@rivm.nl)
Marga Ocké (liangzi.zhang@wur.nl)

Version: 1 Date: 19 Dec 2018

Author’s response to reviews:

NUTJ-D-18-00263

Importance of Details in Food Descriptions in Estimating Population Nutrient Intake Distributions

Dear Prof. Kirkpatrick,

Thank you very much for your willingness to consider a revised version of our manuscript “Importance of Details in Food Descriptions in Estimating Population Nutrient Intake Distributions” for publication in the Nutrition Journal. We believe that our responses to the valuable and helpful comments of the reviewers have resulted in an improved manuscript. Comments from the reviewers and our responds are detailed below. Changes are highlighted directly in the text.

Sincerely,

Liangzi Zhang

Reviewer #1: This is interesting work, and definitely has potential for streamlining data collection methodology. I cannot speak to the validity/reasonableness of the random forest method, but developing a scientific approach to identifying extraneous details in dietary
interviews has definite value for the scientific community. It is also very important to assess the impact of any changes to the interview, which the authors have done. However, the conclusion that elimination of these facets will have no significant effect on nutrient distributions is not really demonstrated by this work. The results show that 727 of the foods reported in the Dutch national study were changed to a new food code following elimination of the facets identified as least predictive. This represents just 2.7% of all the foods reported in the study (26,679). The lack of effect on nutrient distributions is as likely to be due to the small number of changed food codes as it is to changes in the nutrients resulting from the facet elimination. It may be that in a population where these changes were more prevalent, the effect on nutrients may be greater. This may be one of the points - that the elimination of these facets may simplify the data collection and have little effect on nutrients precisely because it affects so few food items, but that is not the point the authors make. The manuscript would benefit from a discussion of the impact of changing so few food codes.

Answer: We appreciate the reviewer for the observation. In order to see a greater effect of facet deletion, we have decided to increase the facet cut-off point from 0.8 to 1.00. Then we did the NEVO code reassignment and the simulation study again based on the new cut-off point. We have added the following sentence After deleting facets with a max. normalized %IncMSE lower than 0.80, trivial effects on population nutrient intake distributions were observed, therefore a cut-off point at 1.00 was chosen for greater effects.” In lines 158-160. In the revised manuscript we now briefly summarized the results for the scenario with 0.8 as cut-off with a reference to Additional file 3 for full results, and more extensively described the results for the cut-off of 1.00.

Moreover, based on the reviewers’ comments, we think it is important to clarify the number of changed NEVO codes. In the original manuscript, we mentioned that 727 out of 26,679 records with a changed NEVO code. 26,679 was the number of unique food combinations, but in the full dataset there was a total of 219,006 food consumed foods available. Therefore the number of changed NEVO codes in the full dataset was also higher than 727. For a cut-off point at 0.8, in the original total food consumption database, 77,529 (22%) of the 350,369 descriptors used were identified as unimportant, which has resulted a NEVO code change of 8% (2251 out of 26,679) of the combinations in the unique food dataset and 1.9% (4232 out of 219,006) of the combinations in the total food consumption dataset.

In the revised manuscript, we explained the meaning of these numbers with an extended sentence: “The total collected consumption data from all participants for the two 24HRs has 219,006 food records, with 350,369 descriptors ranging from 0 to a maximum of 8 for individual foods. This results in a number of 26,679 unique combinations of foods with descriptors.” In lines 126-128.

This has increased the number of changed NEVO codes. However, the number is still quite low compared with all reported foods. We included a sentence to clarify the number of changed NEVO codes: “In the original total food consumption database, 35% (121,015 out of 350,369) of the total number of descriptors used were identified as unimportant, which has resulted a NEVO code change of 11% (2,923 out of 26,679) of the combinations in the unique food dataset
and 3.7% (8,196 out of 219,006) of the combinations in the total food consumption dataset. "In lines 221-224.

In discussion, we added a sentence illustrating that the lack of effect on nutrients was due to the low number of NEVO code changes due to limited food included in the Dutch food composition database, which is a generic food composition database with 2389 foods: “...difference of less than 10% was found for most nutrients. This could be explained by the fact that 96.3% of the combinations were relinked to the same food code in the food composition database. Apparently, the food name and remaining facets provided sufficient information to link to the same food item in the Dutch National Food Composition Database.” In lines 262-265.

It is not clear that saving 3 to 4 minutes of a 45 minute interview time is truly an impact - it is not a significant reduction in respondent burden on an individual level.

Answer: After changing the cut-off point of facet deletion, we recalculated the amount of time that could be saved during the interview: “The time needed to complete the facet collection procedure of a 44 minutes 24HR interview was estimated to be 15 minutes. Without 35% of the unimportant facets, the time saved for one interview would on average be 5 minutes. In a survey with 3819 participants that are interviewed twice, a total of 637 hours would be saved. "In lines 314-317. We agree that 5 minutes are not a large reduction on an individual level, but it is still one third of the time needed for the food description part of the interview and 11% reduction on total interview times. Moreover, the paper aims to reduce the burden and cost of conducting the survey, the total of 668 saved hours can be beneficial for the survey implementer.

The authors claim that the reduction in the number of unique food combinations would result in 200 fewer coding hours. The coding hours are more accurately determined by the number of food items reported by respondents, and as the work evaluated in this manuscript is data processing, it doesn't support a reduction in the number of food items reported. Additionally, when coding 7,638 24-hour recalls, a savings of 200 hours is unlikely to be a meaningful reduction.

Answer: In the data handling procedures of DNFCS, NEVO codes were assigned automatically to repeated food combinations, therefore do not need to be linked manually. We could just focus on the reduced number of unique food–descriptor combinations. After changing the cut-off point of facet deletion, we recalculated the amount of time that could be saved during data handling based on the number of reduced combinations, we now explained this better in the text: “After deleting less important facets from the unique food-descriptor combination list, the number of unique combinations reduced with 3534 (from 26679 to 23145). In the data handling of DNFCS, only new food-descriptor combinations needed to be linked to the food composition database manually. Therefore, around 442 hours would be saved for data handling to link each unique food-descriptor to the food composition code. To sum up, we estimated that a total of 1110 hours would be saved for both data collection and handling if facet deletion would be applied.” In lines 319-324.
Comments on lines from the manuscript:

149: Stratified by food group, the importance of a facet is calculated as the percentage increase in prediction error, denoted by %IncMSE, when data for that facet are permuted in the dataset (implying no predictive power for that facet in predicting the nutrient content), while keeping data for the other facets unchanged.

The sentence is confusing, as it seems to be saying that an increase in the prediction error implies no predictive power for the facet in predicting the nutrient content. Doesn't an INCREASE in prediction error when data for that facet is permuted mean that the facet actually DOES have predictive power for the nutrient content?

Answer: We apologise for this unclear sentence. We have changed the sentence into: “Stratified by food group, the importance of a facet (denoted by %IncMSE), was calculated as the percentage increase in prediction error, when data for that facet were permuted in the dataset, while keeping data for the other facets unchanged.” In lines 139-141.

204 In the food group 'Miscellaneous' most facets were unimportant, followed by 'Meat (products)' and 'Dairy (products)'.

Sentence is unclear. What does it mean that meat and dairy "follow miscellaneous"?

Answer: Since the results has changed due to the increased cut-off point, we rewrote this part of the manuscript. Please see highlighted sentences from line 196 to line 237.

Reviewer #2: This is an interesting paper exploring the use of random forest algorithm to identify the less important characteristics of food descriptions (facets) for the purpose of improving the efficiency of data collection and handling in 24-h dietary recalls. The analyses were carefully done but for clarity (and to avoid confusion), the reviewer has the following comments:

Lines 122-123: There were two 24 HRs, however, it is not clear whether analyses were based on both 24 HRs or only one of them? What was the rationale (using both vs. choosing one 24 HRs over the other)? If both 24HRs were used, how was the two 24 HRs analyzed (i.e., average, or other methods)?

Answer: We appreciate the reviewer raising this point and have added the following sentence to make things clear: “...foods consumed by all participants in both 24HRs were used for predicting facet importance,” in line 137. “The average nutrient intake calculated from two 24 HRs of each
participant was used in the simulation study for estimating population nutrient intake distributions.” In lines 165-167.

Lines 123-125: How many interviews were telephone vs. face-to-face? Were results/estimates obtained from telephone vs. face-to-face interviews comparable (i.e., both formats of interview could be used interchangeably?)? The paper seems to suggest that both 24 HRs (from a same participant) were always conducted in the same format depending on the age of the participant (i.e., adults vs. children), however please confirm this.

Answer: We have added the following sentences in the methods part:“ The 24HRs for 2522 persons aged 16 and older were conducted by trained dieticians through telephone interviews. The 24HRs for 1297 children between 7 to 15 years old were collected by face-to-face interviews with the children and their care takers during home visits. All interviews were conducted following a same data collection and handling protocol.” In lines 106-109.

Lines 128-130: Was the average time (42 minutes) needed to complete one 24HR for telephone or face-to-face interviews? Was there a difference in time regarding these two formats of interview?

Answer: We have incorporated time for both telephone and face-to-face interview:“ The average time needed to complete one face-to-face 24HR interview was 41 minutes and 46 minutes for telephone interviews.” In lines 111-112.

Lines 133-134: What was the rationale behind deciding on the 16 facets? Was it a data-driven or expert knowledge decision?

Answer: We added an explanation on this in the text:“ IARC provided for countries that used Globodiet as their data collection software with the common facets and descriptors. The actual selection of facets and descriptors could be adjusted according to country-specific situations. For the Dutch version of the software, a total of 16 facets with varying numbers of descriptors were selected by experienced dieticians based on knowledge of the food market and insight in the purposes for which the data were collected.” In lines 116-120

Lines 137-139: Please could you explain how you came up with 26,679 unique combinations of foods with descriptors? Was that number obtained from the 2,389 food items and the total number of descriptors from Table 1 (total number of descriptors=244)?

Answer: The 26,679 unique combinations were obtained from foods within a food list incorporated in the Globodiet software and different combinations of descriptors. We have rewrote the sentences into:“ The total collected consumption data from all participants for the two 24HRs has 219,006 food records, with 350,369 descriptors ranging from 0 to a maximum of 8 for individual foods. This results in a number of 26,679 unique combinations of foods with
descriptors. All food records were linked to 1599 most appropriate food codes in the Dutch National Food Composition Database (NEVO table 2011/3.0) by trained dieticians. NEVO 3.0 contains the energy, macro- and micronutrient contents of 2,389 food codes in total.” In lines 126-130

Lines 139-140: If 1,599 codes of 2,389 food items were linked to the consumption data, what happened to the remaining 790 (=2389-1599) food items?

Answer: We apologise for the misleading sentences. It should be “1599 codes of 2389 codes were linked to the consumption data”. The remaining 790 codes were not relevant to the consumption data.

Lines 143-144: The reviewer suggests adding an appendix listing the food items included in each food group.

Answer: We added Additional File 1 with detail information on foods included each food group.

Lines 150-151: The reviewer is confused about the meaning of "when data for that facet are permuted in the dataset (implying no predictive power for that facet in predicting the nutrient content)". Please add details or reformulate the sentence.

Answer: We reformulated the sentence into: “Stratified by food group, the importance of a facet (denoted by %IncMSE), was calculated as the percentage increase in prediction error, when data for that facet were permuted in the dataset, while keeping data for the other facets unchanged.” In lines 139-141.

Lines 182-189: The reviewer is confused regarding the explanation about the NEVO code reassignment protocol. Since those detailed explanations are not included in Figure 1, the reviewer suggests adding them into Figure 1 for clarity.

Answer: We added Figure 2 to illustrate the protocol and reformulated the sentences into: “As illustrated in Figure 1, a NEVO code reassignment protocol was developed to identify NEVO codes of the most similar food-descriptor combinations with the combination of facets that needed to be relinked considering only important facets. For foods-descriptor combinations in the dataset with same food IDs, combinations received a positive score for each identical descriptor (equal to the maximum normalized %IncMSEs) and a penalty for each descriptor that was different (equal to the negative maximum normalized %IncMSEs). The scores were summed and the NEVO code of the food-descriptor combination with the highest score was assigned to the combination that needed to be relinked.” In lines 171-178.
Lines 191-195: It seems both 24 HRs were used at the second step of the simulation study (see "averaged over two days"). Does this mean: only one of the two 24 HRs was used in random forest algorithm, and both 24 HRs were used at the second step of the simulation study? Please add details to avoid confusion. Lastly, did the authors consider using the NCI method (https://epi.grants.cancer.gov/diet/usualintakes/method.html)? If yes, please add details. If not, please discuss in the discussion section.

Answer: We appreciate the reviewer raising this point and have added the following sentence to make things clear: "...foods consumed from all participants in both 24 HRs were used for predicting facet importance," in line 139. "The average nutrient intake calculated from two 24 HRs of each participant was used in the simulation study for estimating population nutrient intake distributions." In lines 165 – 167.

We did consider calculating usual intake, however, since the study is already complicated, and calculating usual intake may not add more insight in the nutrient intake distribution, we did not apply this method in our current study.

Lines 195-197: Were sampling weights used in the analyses in order to get population nutrient intake distributions?

Answer: We added the following to the text: “All results were weighted for small deviances in sociodemographic characteristics (age, sex, region, degree of urbanisation and educational level), day of the week and season of data collection, in order to give results that are representative for the Dutch population and representative for all days of the week and all seasons.” In lines 185-188.

Lines 201-202: The correct values should be "50 out of 111 facets".

Answer: We thank the reviewer for pointing this out. Since we have changed the cut-off point from 0.80 to 1.00, the result has changed. Please see line 192 to 213.

Lines 209-210: 'Packing medium' was unimportant for 6 food groups (because of the value "0.02" (column 1 - Potatoes; Row 5 - Packing medium)).

Answer: We thank the reviewer for pointing this out. We have changed the sentences into: “‘Brand name (yes/no)’ and ‘Packing medium’ were unimportant for the most of the food groups (10 and 7 food groups, respectively).” In lines 211-212.

Lines 217-218: The reviewer is not sure how the value "350,369 facet descriptors" was obtained, knowing there were 16 facets, a total of 244 descriptors and 26,679 unique combinations of foods with descriptors. Please explain.
Answer: This is the total number of descriptors used for the 2007-2010 survey. We added one sentence to clarify this: “The total collected consumption data from all participants for the two 24HRs has 219,006 food records, with 350,369 descriptors ranging from 0 to a maximum of 8 for individual foods. This results in a number of 26,679 unique combinations of foods with descriptors.” In lines 126-128.

Table 2: There should be an asterisk for value "0.02" (column 1 - Potatoes; Row 5 - Packing medium). Consequently, the "# of omitted/# of original" for column 1 (Potatoes) should be 3/8, and the "# of omitted/# of original" for row 5 (Packing medium) should be 6/9. Lastly, the "# of omitted/# of original" for column 2 (Vegetables) should be 0/5.

Answer: We appreciate the reviewer for pointing this out. We highlighted the correction in Additional File 2.

Discussion section: Although information collected from 24 HRs in this study was not self-reported but by telephone or face-to-face interviews, there still exists challenges related to misreporting or other types of measurement error. Please add this topic into the discussion section.

Answer: We appreciate the reviewer for raising this point, we have incorporated it into the discussion: “Finally, the impact of facet reduction on respondents’ answers during the food description part of the interview was not assessed. Although a face-to-face or telephone 24HR interview has generally smaller self-reporting error than other methods, measurement error is likely to be present (i.e. rely on memory, underreporting) [6]. However, we assume that the effect of facet reduction on self-reporting error will be small.” In lines 297-301.