Author’s response to reviews

Title: Reproducibility and validity of a Food Frequency Questionnaire among pregnant women in a Mediterranean area

Authors:

Jesus Vioque (vioque@umh.es)
Eva-Maria Navarrete-Muñoz (enavarrete@umh.es)
Daniel Gimenez-Monzó (dgimenez@umh.es)
Manuela García de la Hera (manoli@umh.es)
Fernando Granado (fgranado.hpth@salud.madrid.org)
Ian S Young (i.Young@qub.ac.uk)
Rosa Ramon (ramon_rosbon@gva.es)
Ferran Ballester (ballester_fer@gva.es)
Mario Murcia (murcia_mar@gva.es)
Marisa Rebagliato (rebagli@umh.es)
Camen Iñiguez (inyiguez_car@gva.es)
for the INMA Valencia Cohort Study (bibliodieta@umh.es)

Version: 2 Date: 19 January 2013

Author’s response to reviews: see over
Dr Nagaraj Nagathihalli  
The Nutrition Journal Editorial Team  

Alicante, 19th January 2013  

Dear Editor,  

Thank you for revising our manuscript entitled “Validity of a Food Frequency Questionnaire to assess diet during pregnancy”.  

We thank the reviewers and the editorial team for the helpful comments.  

Please find enclosed our point by point replay and the revised manuscript. Answers and changes in the manuscript have been highlighted in the text using a red font. We believe we have answered all points appropriately, thereby improving the manuscript.  

Yours faithfully,  

Jesus Vioque, MD, MPH, PhD.  
Departamento de Salud Pública.  
Universidad Miguel Hernández.  
Campus de San Juan.  
03550-San Juan de Alicante, Spain.  
Telephone: +34 6 591 9517. Fax + 34 6 591 9551  
vioque@umh.es
POINT BY POINT REPLY TO THE COMMENTS MADE BY THE MEMBER OF THE EDITORIAL BOARD AND THE REVIEWERS

Paper no.: "Reproducibility and validity of a Food Frequency Questionnaire among pregnant women in a Mediterranean area" by Jesús Vioque, Eva-María Navarrete-Muñoz, Daniel Giménez-Monzó, Manuela García de la Hera, Fernando Granado, Ian S Young, Rosa Ramón, Ferran Ballester, Mario Murcia, Marisa Rebagliato, Carmen Iñiguez, for the INMA-Valencia Cohort Study

Response to the comments from reviewers

Reviewer 1: Kazufumi Nakamura
Reviewer's report:
This is an interesting manuscript describing the reproducibility and the validity of FFQ in pregnant women. This manuscript provides useful information on assessment of food intake during pregnancy.
We thank reviewer 1 for these comments.

Reviewer 2: Maijaliisa Erkkola
Reviewer's report:
In general the manuscript is quite well written and handles a very important topic because of the specific group for which data are still lacking. Novelty of the paper lies in the use of biochemical calibration to assess validity of the food frequency questionnaire. However, it also presents the biggest challenge for the study as the biochemical marker and dietary assessment method do not often measure the same thing. In terms of validation studies, there is a need to be clear about just what the biomarker measures. Many, if not most, biomarkers do not permit an assessment of true absolute intake. I suggest one more editing round before accepting the manuscript to be published in the journal.
We thank this reviewer for the helpful comments and suggestions. We have worked on them, making changes and adding new text in the new version of the article. As suggested, we also made an editing round in the light of all reviewers’ comments.
We agree that vitamin plasma concentration and FFQ nutrient intakes may not measure the same thing. The advantage in using a biochemical indicator is that measurement errors should be essentially uncorrelated with errors in dietary questionnaire. However, they have limitations since they do not provide a quantitative measure of dietary intake but rather a qualitative indicator. In our view, the best option for dietary assessment in most epidemiological studies is the FFQ and the best option to validate a FFQ would be to use as a reference method several food diaries or 24 hr recalls during the time frame of the FFQ.
However, this option was not feasible in our case mainly for operational and resource reasons. As a result, we explored first the reproducibility of the FFQ by comparing nutrient and food intakes at two different moments in pregnancy, thus assuming that the diet of individuals did not change substantially over the elapsed period between. In fact, we performed analysis only with women who reported no major change in their diet during pregnancy but correlations remain basically the same (see response to comment 1, reviewer 4). A sentence was included in Discussion of the article.
We also explored relative validity by comparing nutrient intakes of several vitamins with their respective concentrations in plasma. Although nutrient intake estimates by the FFQ may represent long-term exposure while nutrients in plasma be more related to recent intake, we assume that plasma nutrients were sensitive and acceptable indicators of intakes (with some exceptions, eg, retinol, vitamin E), despite that nutrient levels in plasma may be influenced not only by dietary intake but also by external factors such as the food matrix, the food preparation and by host factors such as gender, smoking status and BMI which did not
substantially change during pregnancy. In addition when we stratified the analysis for some of these variables, correlations did not change. Thus, in a new “editing round” we have included a comment in the Discussion on this issue in which we also point out that there is no reason to assume that vitamin concentrations in plasma should not be representative of the period covered by the FFQ (see response to comment 7 below).

New text added in red
Lines 286-289 new version: “Despite the use of biomarkers as a reference methods may have some limitations since they do not provide a quantitative measure of dietary intake but rather a qualitative indicator and may be more related to recent intake while FFQ to long-term exposure, plasma concentration for some nutrients are sensitive and acceptable indicators of intakes. Therefore, their use for FFQ validation is considered appropriate since the two methods of assessing diet have different sources of error that are unlikely to be correlated with each other [7, 8].”

Line 350-353: “Thus, validity of the FFQ may be even higher than shown in table 3 as correlations were improved when time frame and season between nutrient intakes and plasma concentrations were matched more properly in the analysis.”

Lines 372-377: Therefore, we only explored relative validity by comparing nutrient intakes with their respective concentrations in plasma. Nutrient levels in plasma may be influenced not only by dietary intake but also by external factors such as the food matrix, the food preparation and by host factors such as gender, smoking status and BMI; however, when we stratified the analysis for some of these variables, correlations did not change and showed acceptable degree of validity still.

Lines 342-342 in the original version were deleted: “In addition, food records and 24-hour dietary recalls collect information on nutrient intakes from foods only, leaving out information pertaining to the use of dietary supplements frequently used during pregnancy.”

Lines 378-381: the text in last paragraph of Discussion has been changed a little bit to be more consistent with the Conclusion in the Abstract. The word “acceptable” has been added and the word “foods” deleted.

• Major Compulsory Revisions
1. Abstract: Rewrite the conclusion “…can provide a valid estimate of diet across a wide range of food groups and several important nutrients during pregnancy.” It is too loosely written conclusion as you did not study validity of the FFQ to estimate food intake and the results for nutrients were acceptable only with caution. (lines 53-54)
   We agree. Conclusion has been changed.
   “This study supports that the FFQ has a good reproducibility for nutrient and food intake, and can provide a valid estimate of several important nutrients during pregnancy.”

2. Lines 74-75: I do not see how the frequent use of dietary supplements could limit the use of food records and 24h-recalls to validate an FFQ. The FFQ does not measure the use of supplements if not asked separately as could be done when keeping diaries and answering recalls.
   We agree with this comment. The text has been changed accordingly.
   “The most frequent reference methods to validate FFQ have been food records and 24-hour recalls [6] although they have been used less frequently in pregnant women.”
3. Lines 104-105: the classification of subjects by their socioeconomic status should be clarified. At the present form, it does not open for the reader. The same applies to the table 1.


In any case, we clarify the classification for socioeconomic status. New text has been included in the article: “Social class I included managerial and senior technical staff and freelance professionals; class II included intermediate occupations and managers in commerce; class III included skilled non-manual workers; class IV included skilled and partly skilled manual workers; and class V included unskilled manual workers. Categories were further grouped in three categories (I+II, upper; III, middle; IV+V, lower) in our analysis”.

Dietary assessment

4. Did you use data from previous dietary surveys in pregnant women to be sure that most commonly eaten foods and the foods that most discriminate between pregnant women were included into your FFQ?

We did not carry out previous dietary surveys in pregnant women. As we mention in the article, the 101 food-item FFQ was based on a previous 93 food-item FFQ validated in an adult population of the Valencia Region using 4 one-week food diaries as the reference method. The original 93-food-item FFQ was updated and expanded by including new available foods, commonly eaten by adults in Valencia. Both FFQ are available in Spanish at http://bibliodieta.umh.es/bibliodieta/.

In order to capture any other relevant food, we included a question at the end of the FFQ: “Do you regularly eat any other food at least once a week?” Very few women reported eating new food items so that this information was not included in the analysis. Thus, we do believe that the FFQ was quite comprehensive and included the most relevant and frequently used foods by pregnant women in our region. FFQ estimates for food and nutrient intakes are comparable to other population of childbearing age.

5. Were the FFQs checked with respondents; any checks for coding errors?

Trained interviewers were provided with specific instructions and examples to facilitate FFQ completion. Interviewers were specifically instructed to check the FFQ with women at the end of interviews to ensure that all FFQ food items were completed. Coding errors were checked (double data entry) and filters to avoid implausible values were used (only values from 1 to 9 were allowed for the nine possible frequency responses). Since this type of information is not usually given in Methods we will not add new text in the article unless otherwise required.

6. Lines 123-125: The ranges should be presented separately for the estimates of reproducibility and validity.

As suggested the information has been changed providing more detailed information.

Old text: The validity and reproducibility correlation coefficients (adjusted for energy intake) ranged from 0.38 for reproducibility of carotenoids to 0.44 for validity of vitamin C [16, 17]

New text (lines 129-132): The validity correlation coefficients (adjusted for energy intake) ranged from 0.27 for folate intake to 0.67 for calcium intake (average 0.47) and the reproducibility correlation coefficients ranged from 0.30 for carotene intake to 0.65 for calcium intake (average 0.40).

Biomarkers
7. It is essential that the biomarker information is collected on days that are representative of the total frame of the FFQ – does this apply to all the nutrients assessed?

*We agree with this point. Women were invited to participate the day they attended their first pre-natal visit (between the 10th and 13th weeks of gestation). We think that there is no reason to believe that biomarkers measured that day would not be representative of the total frame of the FFQs. On the other hand, we considered it unfeasible to obtain one or more blood samples from pregnant women by asking them to come a few days/weeks before their first pre-natal visits.*

**Results**

8. Validity should be additionally assessed by dividing subject into categories by the intakes from the FFQ and respective plasma concentrations. This method gives a much clearer and undistorted picture of how well the instrument is doing compared with correlation coefficients.

*We agree in part with this point.*

Alternatives for presenting data from validation studies are mostly based on correlation coefficients although additional methods are also used to explore misclassification (e.g., cross-classification by quintiles, or regression coefficients). Correlation coefficient provides, in a single number, the same information as that contained in a contingency table. In addition, contingency tables may become cumbersome when many nutrients are being studied (Willett, 1998). We provide an example for Vitamin C:

<table>
<thead>
<tr>
<th>Vit C intake (quintiles)</th>
<th>&lt;= 33.60</th>
<th>33.61 - 4</th>
<th>44.76 - 5</th>
<th>55.91 - 6</th>
<th>66.91+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;= 77.54</td>
<td>32</td>
<td>23</td>
<td>17</td>
<td>17</td>
<td>15</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td><strong>30.77</strong></td>
<td><strong>22.12</strong></td>
<td><strong>16.35</strong></td>
<td><strong>16.35</strong></td>
<td><strong>14.42</strong></td>
<td><strong>100.00</strong></td>
</tr>
<tr>
<td>77.55 - 108.99</td>
<td>21</td>
<td>23</td>
<td>21</td>
<td>20</td>
<td>19</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td><strong>20.19</strong></td>
<td><strong>22.12</strong></td>
<td><strong>20.19</strong></td>
<td><strong>19.23</strong></td>
<td><strong>18.27</strong></td>
<td><strong>100.00</strong></td>
</tr>
<tr>
<td>109.00 - 145.17</td>
<td>16</td>
<td>19</td>
<td>28</td>
<td>22</td>
<td>19</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>15.38</td>
<td>18.27</td>
<td>26.92</td>
<td>21.15</td>
<td>18.27</td>
<td>100.00</td>
</tr>
<tr>
<td>145.18 - 200.77</td>
<td>22</td>
<td>19</td>
<td>16</td>
<td>24</td>
<td>23</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>21.15</td>
<td>18.27</td>
<td>15.38</td>
<td>23.08</td>
<td><strong>22.12</strong></td>
<td><strong>100.00</strong></td>
</tr>
<tr>
<td>200.78+</td>
<td>14</td>
<td>19</td>
<td>24</td>
<td>19</td>
<td>28</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>13.46</td>
<td>18.27</td>
<td>23.08</td>
<td>18.27</td>
<td><strong>26.92</strong></td>
<td><strong>100.00</strong></td>
</tr>
<tr>
<td>Total</td>
<td>105</td>
<td>103</td>
<td>106</td>
<td>102</td>
<td>104</td>
<td>520</td>
</tr>
<tr>
<td></td>
<td><strong>20.19</strong></td>
<td><strong>19.81</strong></td>
<td><strong>20.38</strong></td>
<td><strong>19.62</strong></td>
<td><strong>20.00</strong></td>
<td><strong>100.00</strong></td>
</tr>
</tbody>
</table>

Pearson chi2(16) = 21.0110 Pr = 0.178

Likewise, we presented the proportions of correctly categorized subjects in the same or adjacent quintiles for reproducibility using the same FFQ (Last column in table 2). Since dietary assessment methods for validity were different, we did not include the same column. As suggested we have added a new column in table 3 similar to that presented in table 2 to show the percentage of correctly categorized subjects in the same or adjacent quintiles for the validity as well.

**Discussion**

9. Do not present new results in the discussion (lines 311-323). They should be included in the results.

New results are now only presented in the Results although a comment is still kept in Discussion.
10. Line 325-6: “...be valid to estimate of long-term intakes and rank women according to their intakes for relevant nutrients.” The statement of ranking should be based on classification into categories of intake/concentration which I suggest the authors to do. We do think that the FFQ may be valid to estimate long-term nutrient intake as supported by several coefficient correlations of validity. However, as we did not present cross-tabulation, we have changed the word “rank” in the old sentence and used new text instead: “...be valid to estimate long-term intakes and acceptably classify women according to their intakes for relevant nutrients.”

- Minor Essential Revisions

11. Line 94: For a better presentation of results, it would be better to always present both, frequencies and percentages. Percentages have been included together with frequencies.

12. Results: Add subtitles for the better organization of the text. We are not sure if subtitles are allowed by the editorial team. If so, we would propose two subtitles in Results: reproducibility and validity.

13. Table 1. Decimals are not needed when presenting the percentages. Done

Reviewer 3: Joan D. D Fernandez-Ballart

We thank this reviewer for the very detailed comments many of which have helped to detect mistakes and to increase the precision of data presentation and the quality of the article.

Reviewer’s report:

- Minor Essential Revisions

1) The term “calibration” is used twice in the manuscript but possibly in an inappropriate sense. According to Willett (reference 3 of the manuscript, page 101): “calibration refers to a process in which values from one method are quantitatively related to values from a superior, standard method”. Please change line 37 in the Abstract “The aim of this study was to examine the reproducibility and validity (biochemical calibration) of a ... ” for “The aim of this study was to examine the reproducibility and relative validity (against biochemical biomarkers) of a ...”.

   Change line 266 in the Discussion “The use of biochemical measures for calibration of the questionnaire” for “The use of biochemical measures for validation of the questionnaire”.

   Done

2) Please introduce the word “for” between “found” and “major” in line 37 (Abstract): “Significant correlations for reproducibility were found for major food groups and ...”.

   Done

3) Please provide a better explanation as to why “the frequent use of dietary supplements” might be a limitation in the use of food records and 24-hour recall methods during pregnancy. When using food records and 24-hour recall it is also possible to ask women about the use of supplements with the same tools as in the FFQ. This argument appears twice in the manuscript: lines 74-75 and lines 342-344.

   Please see comment 2 from reviewer 2. The sentence has been changed: “The most frequently used reference methods to validate FFQ have been food records and 24-hour recalls [6] although they have been used less frequently in pregnancy.”
4) It is clear that the FFQ was completed by the pregnant women on two occasions but sometimes these are referred to as “... weeks 12 and 32 of pregnancy ...” (as in line 96 or Table 3) and others as “between weeks 10-13th ...” and “between weeks 28-32th ...” (as in Table 2). Please use the same terminology for this topic throughout the text.

The first interview took place between 10-13 weeks (on average the week 12) and the second interview between weeks 28-32 of gestation, although closer to the week 32 of gestation. We mention this the first time in Methods.

5) In line 156 add “between weeks 10-13th of pregnancy” after “… during the first visit” to improve clarity.

Done

6) In line 106: change the position of the dot from “… the British classification system (I+II, III, IV+V). [13]” to “... the British classification system (I+II, III, IV+V) [13].”

Done

7) How “weekly maternal weight gain” is calculated needs to be explained (lines 108-110). Without this explanation the calculation of gestational weight gain can’t be reproduced. This sentence has been changed, providing more details:

Gestational weight gain was categorized according to the American Institute of Medicine guidelines (a weight gain of 12.5–18 kg was low for a pre-pregnancy BMI<19.8, a gain of 11.5–16 kg was normal for a BMI of 19.8–26.0, and a gain of 7–11.5 was high for a BMI> 26.0–29.0).

8) In line 110: please insert “prepregnancy” between “Both” and “BMI”.

The sentence has been changed according to the comment 7.

9) An explanation of how supplement use information was collected is provided in lines 114-116, lines 136-138 and lines 144-147. However, it is not clear whether one method or two different methods were used. In the case of one method, maintain consistency in the description and place it in the “covariates” section of Methods. If two different methods were used (as well as information from both), make it clear(er). In this case, which data are added to dietary intake in the analysis?

The same questions were used to collect information on supplement use in both periods of pregnancy. The method is explained only the first time in the text which corresponds to former lines 114-116. Sentences corresponding to lines 138-138 and lines 144-147 have been deleted since they were repetition.

10) Should a minimum time of use of supplements (such as minimum number of days) not be defined before combining the average daily dose of the micronutrient from supplements with the mean daily intake?

We agree. We forgot to mention the time using the supplement that was also asked and it is necessary to obtain mean daily intake. We include that “time of use” in the text.

11) Please, verify whether “all” should be included in line 207 between “during” and “pregnancy”?

Done

12) Please, verify the percentage of women that smoked throughout pregnancy. In line 207 the authors say that there are 23% and in Table 1 that there are 23.4%.

Decimals in percentage have been deleted in table 1

13) Please verify in line 212, the percentage reported for lycopene (can it be 13.5 %?)
We have included a decimal in the percentage

14) In line 222, in the list of food groups with intake in the second FFQ lower than in the first, nuts are not included. This was forgotten or intentionally omitted?
The same FFQ was used the first and second time (table2). Nuts were mistakenly forgotten. They have been included.

15) It is not clear what is ment by “animal fat” in lines 225 and 229. The same term should be used in the text and the tables.
This is a mistake. As stated in table 2, the lowest correlation was found for nuts (r=0.17) which did not reach statistical significance

16) In line 225, change “and” for “to” in order to say “from r=0.22 … to r=0.61 …”.
Done

17) In line 231, revise percentages because they appear incorrect; 57.7 instead of 52.8% and 80.5 instead of 80.1% for lycopene and iodine respectively.
This mistake has been corrected

18) In line 233, revise percentages because they appear incorrect; 59.3 instead of 58.5% and 70.8 instead of 72.2% for fruits and seafood respectively.
19) In line 233, the highest percentage of agreement is for eggs (96.2%), not for seafood.
These comments are related. The double mistake has been corrected.

20) In line 233, use the term Agreement (used in table 2) in addition to the description provided (“classified in the same or adjacent quintiles by both FFQ”).
Done. The new text is: “…in good agreement, i.e., classified in the same or adjacent quintiles by both FFQ (energy adjusted intakes).”

21) Please define “extreme misclassification” in the Methods section. This term is used for the first time in line 234 (results section), without any prior explanation.
We have deleted that sentence since results and comments in discussion are mainly based on agreement instead.

22) In line 239, please introduce “slightly” between “were” and ““higher”.
Done

23) In line 282, please change “… duration of administration between …” for “… time between the administrations of”.
Done

24) The paragraph that starts with line 317 “In fact, …” and finishes in line 323 with “data not shown” would be more appropriate in the Results section
These results have been moved to the Results. (see comment 9 by reviewer 2)

25) In line 331, please change “Vitamin” for “vitamin” and “(0.12)” for “(from r=0.08 to r=0.12)”.
Done

26) In line 332, please change “(0.20)” for “(from r=0.18 to r=0.20)”.
Done
27) In reference number 10, the volume and pages are missing. Please complete it.
Done

28) Please change the order of references 18 and 19 to match with the text.
Done

29) In Table 1, the heading of column 2 “(Subjects (%))” can’t be applied to “Gestational age (weeks)”. Please specify this.
We agree. We have deleted this variable from the table since it is not relevant for the purpose and the main results of the study.

30) In Table 1, specify the units (years) after “Maternal age”. This saves repeating “years” after each category. Repeat this idea for “Pre-pregnancy body mass index”.
Done

31) In Table 1, the categories are not well defined. In fact there is not a category for 35 year-old women and there are two categories for a woman with exactly 30 kg/m2 of BMI. These mistakes have been corrected in the text.

32) The “Smoking during pregnancy” categories in Table 1 don’t correspond with the text in lines 111-112.
The sentence in lines 111-112 has been changed.
New text: Smoking in pregnancy was categorized as never, first trimester only and entire pregnancy.

33) The “Gestational weight gain” categories should be better defined in the table or in text.
See response in comment 7. The text has been changed in Methods defining categories for gestational weight gain.

34) Is not possible to have 702 woman using folic acid supplements, 38 not using these supplements and 32 missing if the total sample size is 740 as indicated.
Data have been corrected. Only data in the second row are presented:

<table>
<thead>
<tr>
<th>Use of folic acid supplements in pregnancy (Missing=32)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>33 (4.6)</td>
</tr>
<tr>
<td>Yes</td>
<td>675 (95.4)</td>
</tr>
</tbody>
</table>

35) The lines “No” can be deleted for “Medical problems …”, “Folic acid …”, “Vitamin C …” and “Vitamin B12 …”.
Those lines have been deleted.

36) Please verify if the description of variables “Folic acid …”, “Vitamin C …” and “Vitamin B12 …” could be improved by indicating time, perhaps “in the first trimester”?
Supplement use refers to any time during pregnancy. We have added “in pregnancy”

37) In Table 2, use “Nutrients (units/day)” in the heading rather than repeating “/day” for each nutrient.
Done

38) In Table 2, use “Food groups (g/day)” rather than repeating “g/day” for each food group.
Done
39) In Table 2, change “Vegetable Fat” to “Vegetable fat” in the last row.
Done

40) In Table 2, specify the units of cholesterol.
Done (mg)

41) In Table 2, change “Folato” to “Folate”.
Done

42) In Table 2, the heading “Pearson correlations …” can be used for columns “Unadjusted” and “Adjusted” but not for the column “Agreement”.
Done

43) In Table 2, indicate that Agreement is in %.
Done

44) The footnote to Table 2 requires careful revision. It seems that “f” refers to “e” and vice versa. Change “proportion” to “percentage”.
Done

45) In Table 3, there is no obvious reason to change the order of nutrients from that of Table 1. The same order of presentation for nutrients is now in tables 2 and 3

46) In Table 3, for consistency use the same terminology to refer to correlation coefficients as in Table 2.
Done

47) In Table 3, there is no need to repeat means (SD) of nutrient intakes that are exactly the same as in table 2 (except for folate. Is this a mistake?). If “Fruits and vegetables” data are necessary include them in Table 2 and delete this column from table 3.
Done. The differences in folate was due to decimal rounding

48) In the foot of Table 3 change “Acid” for “acid”.
Done

49) What time periods did FFQ1 and FFQ2 refer to? The written instructions for FFQ2 states that women should refer to “the last months of pregnancy”; were they given oral instructions on how many months (perhaps since FFQ1)? And what about FFQ1?
This was mentioned in lines 129-132 of the original version of the article.

• Discretionary Revisions
1) In line 258, the authors emphasize that they are able to declare “statistically significant even small correlation coefficients”. They might consider commenting that validity is more based on the magnitude of the correlation coefficient than on its significance.
We agree that validity should be based on the magnitude of correlation coefficients more than on their statistical significance. It is widely accepted that correlation in the range of 0.3-0.7 are considered acceptable for reproducibility and validity and somehow lower when biomarkers are considered.
We added a few words at the end of the sentence: “which may be considered as acceptably valid (r>0.20)”
2) In Table 2, mean and standard deviation are used to describe log-transformed data on nutrient and food group intake. Mean and standard deviation are not the most appropriate descriptive statistics in the reporting of non-normal data. Median and IQR or others might be suitable alternatives.

We share this comment. However, we decided to present mean and standard deviation because of descriptive purposes and comparability reasons since many studies still use them.

3) In Appendix 1, is the number of foods included in each food group indicated in parentheses? It would be useful to specify the number of items included in each food group (for example: Dairy products (1-11) or White meat (13-15), and so on...).

The answer is yes although they are not ordered according to their appearance in the FFQ as the reviewer seems to suggest. There were quite a few food items which were not included in the food groups of the appendix (eg, salt; beverages, alcoholic drinks; etc). We do not consider them relevant for the main purpose of the study.
Reviewer 4: Henrique Barros
Reviewer’s report:
This is a very interesting, well written paper on an important subject that has not gained the needed attention. The research protocol followed standard steps and seems sound. However, some minor questions and comments follow:
We thank the reviewer for the positive comments and helpful suggestions.

1. FFQs were obtained during the 1st and the 3rd trimester of pregnancy: During the course of pregnancy many different health issues can occur and diet might change a lot because of nausea and vomiting (that disappear), hypertension, edema, unwanted weight gain, etc. Thus, we are not really talking about reproducibility but may be measuring in fact different intakes. This is not addressed in the co-variates section (diseases of symptoms of pregnancy) nor in the discussion as it should
We agree with this comment. More details on when the FFQ were completed are provided in Comments 4 and 49 of Reviewer 3.
We also believe that testing FFQ reproducibility may be difficult to achieve in validation studies with humans as the frame period of the FFQ would not be the same and diet may change because of different reasons (eg, vomiting in pregnancy). However, there are options such as restriction in the analysis by excluding those who manifest changes in diet or stratification in the analysis by variables suspected to influence diet.
When the analysis was restricted only to women who reported no change in their diet (n=606), correlations did not improve

<table>
<thead>
<tr>
<th></th>
<th>Total (n=740)</th>
<th>Women reporting no changes in diet (n=606)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r energy adj</td>
<td>r energy adj</td>
</tr>
<tr>
<td>Retinol (μmol/l)</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>α-carotene (μmol/l)</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>β-carotene (μmol/l)</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>Lutein + Zeaxanthin (μmol/l)</td>
<td>0.29</td>
<td>0.28</td>
</tr>
<tr>
<td>Lycopene (μmol/l)</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>β-Cryptoxanthin (μmol/l)</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>Serum Folate (mmol/dl)</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>Vitamin B12 (pmol/l)</td>
<td>0.08</td>
<td>0.09</td>
</tr>
<tr>
<td>Vitamin C (μmol/l)</td>
<td>0.18</td>
<td>0.16</td>
</tr>
<tr>
<td>Vitamin E (μmol/l)</td>
<td>0.05</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Other analyses were performed according to categories of variables such as age, body mass index, country of origin and self-reported vomiting (no/yes) but main results remained practically unchanged.
We have added a sentence in the text:
“Therefore, we explored whether reproducibility correlations were modified by categories of variables such as self-reported changes in diet (no/yes), age, body mass index, country of origin and self-reported vomiting during pregnancy (no/yes). However, correlations remained practically unchanged (data not shown)” (lines 312-315 of new version)

2. Also, in the Discussion section, the authors stated that “We applied a 20-week interval which made answer memorization more unlikely, thus avoiding an artificially increased reproducibility. However, a long interval could also be a concern since diet in pregnant women
may not be as stable as among non-pregnant women and reproducibility may be compromised by real dietary changes. Is any evidence available from the study protocol or other study’s data to help to go into more detail regarding this possible bias? Could adjustment for weight gain help to understand it?

This is a very good point in our opinion. It may be of interest to explore whether reproducibility may increase by excluding from the analysis those women with altered eating habits because of medical reasons or because they go on a diet. We did not explore the influence of weight gain in pregnancy but we estimated correlations by categories of self-reported changes in overall diet and BMI (see comment before). Correlations were not substantially changed.

As we wrote in the Discussion, we believe that “Despite potential changes in diet during pregnancy, the FFQ showed a satisfactory level of reproducibility for most foods and nutrients, particularly for those more frequently eaten”

3. “We collected detailed information on multivitamins use…” Could you give details? Were the box checked, was it based on prescriptions or self-report. Was this information ever validated? Often, women keep taking vitamins during the all pregnancy period. Was it considered?

Detailed information on supplement use was collected by trained interviewers using specific questions at the end of the FFQ as they were reported by women. It is true that some women took supplements during the whole pregnancy (Navarrete-Muñoz, 2010, ref. 34 in the article). Consequently, we took into account supplement use during pregnancy in the analysis (see table 3 for some nutrients, eg, folic acid). In fact, the results substantially improved for some nutrients according to supplement use (eg. Folic acid), which may be considered an indirect validation of the questions used to collect the information on supplement use even though they have not been formally validated.

4. Was dim light also used for carotenoids handling?

Dim light was only used for vitamin C determination; otherwise, the same protocol for handling blood samples was followed with no delays between blood extraction and preparation of the aliquots which were rapidly frozen at -80ºC.

5. You probably got the information regarding anemia and hemoglobin. Was it considered or do you think it might be useful to take it into consideration to value the nutritional status and “independently” give some additional consistence to the findings?

Non-fasting blood samples were obtained for main research purposes of the main study (antioxidant hypothesis). We did not collect information on anemia or hemoglobin although when we explored the effect of some medical symptoms (eg, self-reported frequent vomiting), the results remained unchanged.

6. The authors stated "All nutrient and food group intakes were log-transformed prior to analysis to improve their normality". As we know, very often this is not enough to guarantee normality - did you always got it or had you to accept robustness of methods and not follow the expected assumptions? If yes, how could this influence the magnitude and significance of relatively weak and imprecise point estimates?

Log transformation was used to estimate energy-adjusted intakes for all nutrients and foods. The same approach was used for plasma vitamins. As mentioned in the text, although energy-adjusted correlations were lower for some nutrients, they were of similar magnitude for all nutrients. Statistical significance was similar for crude and energy adjusted correlations in all cases. We also estimated Spearman correlations but again, as results were very similar, only Pearson correlations were presented.
7. It is true that “most carotenoids are transported in plasma lipoproteins” but it is not obvious that “plasma concentrations of carotenoids and vitamin E” needed adjustment “per plasma cholesterol concentrations using…”. Did it really change anything? Weren’t you just over adjusting?

We agree with the comment. However, when plasma concentrations of carotenoids and vitamin E were adjusted for cholesterol, correlations were slightly higher. Thus, we included both energy/cholesterol adjusted correlations ($r^a$) and unadjusted correlations ($r^b$) in table 3 although the differences were not substantial.

8. “As for nutrients, the energy-adjusted correlations for food groups were slightly lower than unadjusted correlations although all of them were statistically significant (except for animal fat)”. Any possible explanation?

This was the case although differences between adjusted and unadjusted correlations were very small. Although adjusted correlations are supposedly used to reduce variability and increased correlations, we believe that the reason they were lower in our study may be related to the characteristics of the population. Although our study population was pregnant women with a relatively similar age range, they had quite different lifestyles (smoking, alcohol), country of origin (12%), educational level, etc and energy adjustment may not be sufficient to balance the variability in diet due to all these different characteristics as opposed to that observed in other more stable population (eg. Nurses Health Study). However, the season of the year did modify correlations for some nutrients as we mentioned in the article.