Author's response to reviews

Title: Assessing the Validity of a Self-administered Food-Frequency Questionnaire (FFQ) in the Adult Population of Newfoundland and Labrador, Canada

Authors:

Lin Liu (liulin0220@gmail.com)
Barbara Roebothan (broeboth@mun.ca)
Ann Ryan (annr@mun.ca)
Yanqing Yi (Yanqing.Yi@med.mun.ca)
Michelle Cotterchio (Michelle.Cotterchio@cancercare.on.ca)
Guang Sun (gsun@mun.ca)
Christina Tucker (cst326@mun.ca)
Jennifer Colbourne (jennifer_colbourne@hotmail.com)
Natasha Baker (natasha--baker@hotmail.com)
Peizhong P Wang (pwang@mun.ca)

Version: 2 Date: 23 March 2013

Author's response to reviews: see over
Dear Editor,

I am pleased to submit the revised manuscript entitled ‘Assessing the Validity of a Self-administered Food-Frequency Questionnaire (FFQ) in the Adult Population of Newfoundland and Labrador, Canada’. For your convenience, the changes are in red-color text.

Thank you very much for sending us the very thoughtful comments and constructive recommendations of the reviewers. The research team carefully reviewed the examiners’ comments and addressed them with our best efforts. In addition to our efforts enhancing the manuscript scientific excellence, we also carefully read the journal’s guideline and formatted this manuscript (e.g. references and tables) according to the journal’s style and requirements. All authors have reviewed the revised manuscript and agreed with its submission.

Thank you in advance for your continued consideration of our paper. We earnestly look forward to hearing from you soon.

Peizhong Peter Wang, M.D, Ph.D
Professor (epidemiology)
Responses to reviewers:

Reviewers' comments:

Reviewer #1: Vanessa Garcia-Larsen

1. Background – Some of the main issues about disseminating the use of FFQs in epidemiological studies are addressed. A major limitation in interpreting data from FFQs is the lack of homogeneity in Food Composition Tables (FCTs). I would suggest this to be mentioned in the Introduction/background section.

Authors: We thank the reviewer for the specific suggestion. The suggested details have now been added to the Background section in the revised manuscript.

“However, investigators have recognized that nutritional values reported from FFQ data are subject to substantial error, both systematic and random, that can profoundly affect the design, analysis, and interpretation of nutritional epidemiologic studies. For example, it is essential to covert food composition values from an FFQ into macronutrient and micronutrient values, but a major limitation in interpreting data from FFQs is the lack of homogeneity in food composition tables.”

2. Methods- My main observation refers to one aspect of their methodology. The authors give very good details of how the FFQ and the chosen gold standard were administered and compared. A major methodological issue is that they use Pearson correlation to compare the correlation between the two instruments (as shown in Table 2). Pearson correlation gives an overall estimate of the association between two instruments, and it is commonly used in studies of validation of dietary questionnaires. However, its use might lead to the wrong interpretation as the correlation is based on the whole sample rather than in the correlation of the intake of one subject against the intake reported by the same subject in the other dietary questionnaire. Using intra-class correlation coefficient (ICC) allows a more accurate assessment and interpretation of the validity of the self-reported intake.

Authors: We thank the reviewer for raising this important point and the professional suggestion. As the reviewer alluded, there are several methods used to assess the validity of dietary assessment instruments and no one is perfect. Masson et.al (2012) summarized current statistical approaches for assessing the relative validity of a food frequency questionnaire, and it was suggested to use Pearson or Spearman correlation coefficients, cross-classification, and the weighted kappa statistic. We also found that intra-class correlation coefficient (ICC) was common used in estimating reliability (reproducibility) of the dietary instruments, as it measured the fraction of total variation that is due to
between-individual variability. We appreciate the reviewer’s comments and tend to use this indicator in further reliability evaluation.

As the reviewer suggested, we have calculated the ICCs of log-transformed nutrient intakes (both unadjusted and adjusted) and the results were shown in the following table. As shown, there were minor numerical differences in findings and the values of ICC were generally lower than our Pearson correlation coefficients. However, the interpretation of the results and conclusion were largely unchanged. Since most publications using Pearson correlation coefficients in the validation studies, it could be easy to compare the results if we have the same indicator. Furthermore, in the present paper, we used both Pearson correlation coefficient and cross-classification into quartiles in order to give credence to the results. We hope the above explanation is acceptable to the reviewer.

*Intra-class correlations between Food Frequency Questionnaire (FFQ) estimates and weighted 24-Hour Recall (24-HDR) estimates*

<table>
<thead>
<tr>
<th>Nutrient</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted</td>
<td>Energy-Adjusted</td>
</tr>
<tr>
<td>Energy (kcal)</td>
<td>0.26        —</td>
<td>0.29</td>
</tr>
<tr>
<td>Protein (g)</td>
<td>0.18        0.24</td>
<td>0.29</td>
</tr>
<tr>
<td>Carbohydrate (g)</td>
<td>0.12    0.21</td>
<td>0.24</td>
</tr>
<tr>
<td>Dietary Fibre (g)</td>
<td>0.23    0.25</td>
<td>0.28</td>
</tr>
<tr>
<td>Total Fat (g)</td>
<td>0.22        0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>Saturated Fat (g)</td>
<td>0.18    0.19</td>
<td>0.22</td>
</tr>
<tr>
<td>Monounsaturated Fat (g)</td>
<td>0.21    0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>Polyunsaturated Fat (g)</td>
<td>0.21    0.17</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>0.21</td>
<td>0.27</td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>Cholesterol (mg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamin A(RAE)</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>Carotene (RE)</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>Vitamin D (IU)</td>
<td>0.19</td>
<td>0.20</td>
</tr>
<tr>
<td>Calcium (mg)</td>
<td>0.18</td>
<td>0.19</td>
</tr>
</tbody>
</table>

3. Given that one of the objectives mentioned by the authors is to have their paper as a methodological referent for further validation studies, I think it is important that this is addressed, and the use of Pearson vs. ICC included in the Discussion section.

Authors: As suggested by the other reviewer, the third objective has been done by other investigators in the past and we have deleted this sentence. Since we tend to use ICC in further reliability evaluation studies, the use of Pearson vs. ICC would not include in the present paper.

4. Minor discretionary comment- In the abstract – indicate that it is an FFQ for adults

Authors: The corresponding change has been made.

Reviewer #2: Adam Bernstein

Minor Essential Revisions:

1. Objective 3 under Background has been done by other investigators in the past.

Authors: As suggested by Dr. Bernstein, this sentence has been taken off.

2. Footnotes are not provided in numerical order in the text.

Authors: We thank the reviewer for pointing this out. We have rechecked the footnotes and kept them in numerical order.

3. Can the authors provide examples of the “1,000 additional sources of data” used by ESHA?

Authors: Sure. The ESHA master database is composed of over 35,000 food items, with data from over 1000 reputable sources. Except the major nutrition databases, such as USDA
database, Canadian Nutrient Files, it also includes nutrient information from manufacturers.
The “1,000 additional sources of data” refers to data from restaurant and literature sources,
including Starbucks, A&W Restaurants, Smoothie King, Produce Marketing Association,
FDA Voluntary Labeling, etc. Detailed information can be found in the following website,

4. Why did the authors only use two 24-hr recalls? The authors should discuss how their
findings may have been different had they used more 24-hr recalls.

Authors: During the 24-hr recall data collection phase, each subject recalled and described
in detail all types and amounts of food and beverages consumed in the previous 24 hours on
two separate occasions, a weekday and a weekend day. Weekend days included Saturday and
Sunday to capture food and beverage consumption patterns which may be different from
those on weekdays (Monday to Friday). We must admit this study may benefit from using
more 24-hr recalls; however, it also brings more burden and discomfort to both participants
and interviewers. Meanwhile, due to the limited funding for this research, we tried to make
the survey cost efficiently and chose using two 24-hr recalls. We hope it is acceptable to the
reviewer.

5. What is the population size of NL? Do the study participants reflect members of the
population in terms of measured characteristics?

Authors: According to the new released 2011 Census Information and Statistic, the
population in NL is about 512,659 with around 60% of the residents living in rural
communities. In the present study, a total of 1834 random telephone numbers were initially
identified. After screening for eligibility, 683 eligible participants were contacted to retrieve
further information. As a result, we have recruited a random population-based sample of NL
adults, aged 35-70 years old, and it should be a representative of the NL population. For
example, 57% were rural residents in the present sample vs. 60% found in the census.
However, as reflected in the results, the majority subjects who intend to participate are
non-smokers (82.6%) and with post-secondary education (60.5%). Our sample may not a
representative for smokers or those with slightly lower educational level, but represents a
good cross-section of the NL adults.

6. In the statistical analysis section, how often were foods not exactly matched? The authors
should provide examples of such items and/or provide this information in a Supplement.

Authors: We thank the reviewer for the specific suggestion. As stated in the paper, the
ESHA Food Processor contains thousands of food and beverage items. Most food items from
manufacturers and restaurants can be exactly matched while some homemade dishes need a
group decision within dietetic professionals. As the reviewer suggested, we have provided examples of such items in the revised manuscript.

7. There were nearly 4 times as many women as men in the study. How does this affect interpretation of the results? Also, the authors should comment on potential reasons for gender differences in their results.

Authors: According to the 2011 Census in NL, the female to male ratio in the age-group we concerned (35-70 years) is around 1.4. We had tended to collect data with no significant proportional difference between the male and female initially; however, due to a high refusal rate in males, it included 297 females and 103 males when we reached the target number (400 subjects). We agree with the reviewer that high number of refusal in males may create potential sources of bias, and corresponding changes have been added to the discussion section.

8. The Bland-Altman plot should be provided if not in the manuscript, then in Supplemental Material, if possible.

Authors: The Bland-Altman plots are provided in the supplemental materials as suggested.

9. How does the ‘moderate relative validity’ of their FFQ impact interpretation of future diet-disease studies that use the FFQ?

Authors: We are grateful to the reviewer’s comments. Generally, correlations above 0.4 are considered relatively strong; correlations between 0.2 and 0.4 are moderate, and those below 0.2 are considered weak. In this study, our correlation values were approximately 0.40 when genders were combined, and thus we made the conclusion that this FFQ has moderate relative validity. Most FFQs are designed for large-scale epidemiological studies of the association between diet and disease in which the expected effects of diet on the disease are often modest. As far as we know, there are several authors have explored the impact of different values for the correlation coefficient between the measured value and true dietary exposure in previous studies. Willett (1998) estimated that if a true relative risk is 2.0, the observed relative risk will be attenuated to 1.62 if the correlation between the estimated and true dietary exposure is 0.7, or to 1.32 if the correlation is 0.4, while Burley et al. (2001) argued that for correlations of below 0.3 or 0.4, attenuation will be so severe that it will be difficult to detect associations. Thus, for nutrients with correlation coefficients between 0.4 and 0.7, these calculations represent ‘worst case’ estimates of the impact of measurement error. Since our correlation values fell in the range of 0.4 to 0.7, the effect of error will be relatively small.