Reviewer's report

Title: Model development including interactions with multiple imputed data

Version: 1 Date: 29 September 2014

Reviewer: Libero Fatti

Reviewer's report:

Minor Essential Revisions:

Referee’s report on the revision of the article entitled: “Model development including interactions with multiple imputed data” by GM Hendry et al, submitted for publication to the BMC Medical Research Methodology Journal.

This report will focus on the revisions to the original version of the paper that were required by this referee:

1. Being a methodology journal, it is important to include definitions and explanations of terms that may not be familiar to medical researchers, such as MCAR and MAR. DONE
2. Since Multiple Imputation is central to the article, its methodology and motivation need to be fully described. While the authors don’t necessarily have to go into mathematical detail, the details of how and why it works should be clearly described. DONE
3. The same holds for the EM algorithm. DONE
4. In the last sentence of the subsection on Multiple Imputation in the Methods section, the authors state that 20 sets of data were imputed “on the basis of the percentage of data missing”. The thinking behind this needs to be explained. NOT DONE – WHY 20 AND NOT SOME OTHER NUMBER?
5. In the last sentence of the subsection on Data Review in the Results section, the phrase “…for all variables except ‘food availability’”, needs to be explained and its implications considered. DONE
6. The “numerical problems” referred to in the last paragraph of the subsection on Model Development in the Results section need to be explained and the implications considered. THIS STATEMENT HAS BEEN REMOVED
7. The statement in the first sentence of the Analysis subsection of the Results section, that the standard errors of the parameter estimates (“beta coefficients”) were on average 53.7% larger for the complete case analysis (than for the imputed data analysis) is surprising, as it cannot be explained purely by the decreased size of the complete sample (56.5% of the total available cases). On the basis of the smaller sample size in the complete case analysis one would expect the standard errors to be on average about 33% larger, so it would seem that the multiple imputation is resulting in a smaller average standard error than could be explained on the basis of sample size alone. This needs to be explained, or at least commented on, by the authors. (To this reviewer, the effect
of multiple imputation on results of the analysis may not be as innocent as implied by the article.) DONE

8. The “degrees of freedom associated with the t-value in Rubin’s rules” (last paragraph of the Diagnostics subsection of the Results section) needs to be explained as it would otherwise be meaningless to most readers. DONE

9. The recommendation (paragraph 5 of the Discussion section) that Method 1 should be the method of choice, is based on the findings of this single study. What evidence is there that this is valid in general? DONE – “probably” has been included in this recommendation

Minor comments (Minor Essential Revisions)

a. The singular of “criteria” is “criterion”. NOT DONE. Last sentence of Model development section should read: “Here a p-value of 0.10 was used for the stopping criterion.”

b. Rather use “parameter estimates” or “estimated coefficients” than “beta values”. DONE

Additional Minor Essential Revisions

c. In the second paragraph under Strategy 1 in the section on Model Development with multiple imputation (page 11), the category “female/no” has been left out of the interaction between gender and smoking.

d. Last sentence on page 11. How are the results combined?

e. Under Strategy 2 (page 12), what if there is no selection of variables that is selected in at least 50% of the m sets of imputed data?

f. Page 13 second paragraph in the Results section, there should presumably be only one full stop after “mechanism…”

g. Page 14 in the Model development section. In line 3 it is stated that 17 main effects and 10 interactions were included in the final MVNI model. However, it needs to be stated that far fewer of these were statistically significant. Also, while the FCS1 model includes the same variables as the MVNI model, the set of significant variables in it was very different from those in the MVNI model.

h. Page 20 in the Discussion section. Surely this should be stated the other way round: “…the ratio of cases with completed data to variables should be at least 3:1”? Similarly for the corresponding statement a few lines down.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests.