Author's response to reviews

Title: Practical considerations for sensitivity analysis after multiple imputation applied to epidemiological studies with incomplete data.

Authors:

Vanina Héraud-Bousquet (v.bousquet@invs.sante.fr)
Christine Larsen (c.larsen@invs.sante.fr)
James R Carpenter (James.Carpenter@lshtm.ac.uk)
Jean-Claude Desenclos (jc.desenclos@invs.sante.fr)
Yann Le Strat (y.lestrat@invs.sante.fr)

Version: 3 Date: 3 February 2012

Author's response to reviews: see over
Dear Mr. Christopher Morrey,

Thank you for the opportunity to resubmit our paper titled "Practical considerations for sensitivity analysis after multiple imputation applied to epidemiological studies with incomplete data" for publication in BMC Medical Research and Methodology.

We have provided a detailed response to the reviewer's comments below in italic and have incorporated some changes in a revised manuscript. These changes are highlighted in the main text in yellow.

We hope that the manuscript is now suitable for publication in BMC Medical Research and Methodology.

With best regards,

Vanina Héraud-Bousquet
On behalf on all authors
Reviewer’s report

Title: Practical considerations for sensitivity analysis after multiple imputation applied to epidemiological studies with incomplete data.

Version: 2  Date: 17 November 2011

Reviewer: Michel Chavance

Reviewer’s report:
In spite of the fact that the quality of written English in the added sentences needs revision, the authors have substantially improved their manuscript. The section presenting the four-step procedure for choosing delta is now easier to follow. However two problems remain.

1) The authors still perform their sensitivity analyses assuming that the observation mechanism is the same for the cases and the control, a choice which limits the interest of their paper since the article does not present an extension of the already published method nor an in-depth and sound analysis of real data. The outcome they chose to model was SLD at first referral. Thus data were collected when the outcome was already known and considering that the MNAR mechanism was identical for cases and controls is a strong assumption which deserves more consideration than a few lines in the discussion section.

Data were not collected when the outcome (severity of liver disease) was already known: this is not a case-control study where the number of cases and controls is fixed, and then the covariates are observed. Instead, the data are from the first wave (corresponding to first referral) of a prospective longitudinal cohort of patients entering the HCV surveillance system. For our study we then took the 4343 patients from the study who, at first referral (defined in the methods section), reported having injected or snorted drugs at least once in their lifetime.

At first referral, a standardized notification form was used to collect data, including risk factors for HCV transmission for all patients, independently of the outcome.

Thus while data collection may well be more difficult among those with established liver disease diagnosis, this point does not apply here because we restricted our analysis to patients with no prior diagnosis of their liver disease at first referral.

For these two reasons, we believe that the MNAR mechanism assumption for both SLD+ and SLD- patients is reasonable in the first instance.

To clarify this independence between data collection and the outcome, we added two sentences in the Methods section.
2) Reading the article, I do not understand any better the connection between 
“local sensitivity” and overlap maybe because I do not understand what is 
supposed to be local and what is supposed to overlap. I am used to local 
sensitivity performed in the neighborhood of the MAR assumption (delta=0). This 
is apparently not the case here, since this neighborhood is said to be of no 
practical use.

_In the last version of the paper we did not write that the neighbourhood of the 
MAR assumption is of no practical use. We wrote that ‘the central part of the 
hatched zone corresponds to departures from MAR for which the weights are still 
approximately equal, so that MAR and MNAR inferences are essentially the 
same’._

_Thus, we did perform a local sensitivity analysis because the range of delta was 
bounded around zero (which is the central part of this zone corresponding to the 
MAR assumption)._ 

• I do not either know exactly which distributions should overlap and 
cannot choose between several possibilities. Is it really, as stated (page 11, last 
line), the distributions of the parameter of interest. I doubt it, since the approach 
does not seem to be bayesian. Is it the distributions of its estimators ? or is it the 
estimated distributions of its estimators ? Actually the answers to all these 
questions can easily be found in Carpenter et al. 2007, but the reader of BMC-
MMRM should find all the relevant information in the article.

_Thank you for pointing out this remains unclear. We certainly agree all the 
relevant information should be in the article. We have revised the text on p10-11, 
step 4, replacing ‘estimates’ by ‘estimators’ and adjusting the wording to try and 
minimize the scope for misunderstanding._

• The approach assumes that the support (i.e. range of values for which a 
probability density function is non-zero) of the NMAR estimator distribution is 
contained in that of the MAR estimator. This is necessary because the NMAR 
estimate is obtained by re-weighting the average of the MAR estimates and thus 
must be within the range of the MAR estimates. Comment: this is obviously a 
strong assumption; when both estimators are Gaussian, the distributions share 
the same infinite support, but if their expectations are too far apart, the approach 
will not work and this is why such sensitivity analyses are said to be “local”.

_We agree. To convey this point, we write (p 10) ‘if all the weight is accruing to 
estimates at the end of the range of \( \hat{\theta}_m \), this is consistent with the MNAR 
estimator having a distribution lying outside the range of MAR estimates, i.e. a 
‘non-local’ departure from MAR. Under such a non-local MNAR mechanism the 
estimate of \( \theta \) is most likely beyond the smallest (largest) of the MAR imputation 
estimates.’_