Reviewer's report

Title: How are researchers handling missing data in RCTs? A review of the top medical journals

Version: 2 Date: 22 August 2014

Reviewer: Hua Yun Y Chen

Reviewer's report:

Review report on “How are researchers handling missing data in RCTs? A review of the top medical journals“ by Bell et al.

The article reviews the status of handling missing data in randomized clinical trials in articles published in 4 major medical journals between July-December, 2013. It also compares the results with those in similar studies previously done. Overall, it provide a reasonable summary of the status quo. Here are a few comments and recommendation the authors may consider in the revision.

1. It is important to be accurate in using the statistical terminology and/or making statement about statistical methods. For example, I find the definition of MCAR inaccurate (Line 61 on page 2) in this article. Also, incompletely observed data do not necessarily correspond to dropout and completer (Line 65 on page 2). It can be misleading to make such statements about the validity of a statistical method without qualification. For example, when you say the weighted estimating equation or the mixed model (likelihood approach) or multiple imputation are valid and unbiased method for missing at random data (lines 68-89 on page 2 and other places), it is assumed that the mixed model and the missing data probability model in the case of weighted estimating equation are correctly modeled (and asymptotically unbiasedly estimated). In solving such missing data problem in randomized clinical trials, there is no guarantee that such models used are correct. This might contribute to the lack of use of the more sophisticated methods for handling missing data.

2. The study excludes cluster trails and trials with survival outcomes from consideration. This appears to make results comparable to Wood et al. Since this results in exclusion of about half of the trials (lines 149-151 on page 4), an alternative choice is to study all the trials and perform a subset analysis for the comparison. This reduces the `missing data` in your review of handling missing data.

3. The article provides a nice summary and gives an informative discussion. Points you may consider adding to the discussion are possible reasons for low usage of advanced methods and how to make model-based approaches (weighting, likelihood, and imputations) robust to model assumptions.