Reviewer's report

Title: Accounting for center in the Early External Cephalic Version trials: An empirical comparison of statistical methods to adjust for center in a multicenter trial with binary outcomes

Version: 2
Date: 7 May 2014
Reviewer: Raphael Porcher

Reviewer's report:

The authors have modified the manuscript according to reviewers’ comments, but some comments have been answered or the manuscript was not modified accordingly.

Major Compulsory Revisions

1. There are still many imprecisions in the model formulation and description. Stating that \((\beta_0 + b_{0k})\) follows a normal distribution of mean \(\beta_0\) is not false, but this is not the way the model is fitted (in particular because \(\beta_0\) is unknown). So stating that \((\beta_0 + b_{0k})\) in equation 3 is the random intercept is unusual. Rather \(\beta_0\) is a fixed intercept, and \(b_{0k}\) is a random variable representing how the log odds in centre k deviates from the overall log odds in the control arm, and \(b_{0k}\) follows a normal distribution with mean 0 and variance \(\sigma^2\).

2. I acknowledge that the authors added that they used an exchangeable correlation structure for their GEE models, but the same statement for model 4 (mixed model) is less clear. Actually, the correlation structure refers to the correlation of observations within a same centre. In model 4, what counts is the distribution of the \((b_{0k}, b_{1k})\) random variable. As far as I know, even the syntax of \texttt{lmer()} in the \texttt{lme4} package needs this variance structure to be specified (although there is a value by default). My point here is to make the authors specify if the considered that \((b_{0k}, b_{1k})\) were normally distributed (I assume it was the case), if they were independent and if they had the same marginal variance (I assume this is not the case). The point on independence is important, because forcing \(b_{0k}\) and \(b_{1k}\) to be independent can be problematic. Actually, \(b_{0k}\) measures the centre heterogeneity in log odds in the control group and \(b_{0k} + b_{1k}\) the centre heterogeneity in the experimental group. If both are independent, then the variance of \(b_{0k} + b_{1k}\) is larger than that of \(b_{0k}\), which implies that heterogeneity is more important in the experimental arm. If it is not the case in the data, then this constraint may impair the fit and bias the conclusions on heterogeneity. So this issue should be clarified.

Minor Essential Revisions

1. The authors have added a reference to Liang and Zeger’s paper that I suggested in my first review (reference 11), but only cited it in the introduction, when it should clearly be cited page 13 when referring to GEEs. References 3
and 8 are interesting, but applied works and not seminal papers on the methods, so they cannot serve as a reference for the methodology or theory of GEE.

2. The terminology “random slope” remains, despite being pointed out by two reviewers. Since centre is a categorical variable and treatment indicator a binary variable, the random effect for treatment-by-centre interaction is not really a slope. It would be for treatment by time interaction for instance, if time was continuous. So it would be better to rephrase model v) by “logistic regression with random centre effect and fixed treatment-by-centre interaction” and model vi) by “logistic regression with random centre effect and random treatment-by-centre interaction”

3. By the way, fixed and random “intercepts” are used to describe models ii to vi. “centre effect” would be preferable to “intercept”. The same should apply to the subsection headings.

4. My previous comment on the wording “mixed effects models” and not “random effects models” led to a modified subsection heading, which does not help. Even the model with random treatment-by-centre interaction is a mixed effects model because $\beta_0$ and $\beta_1$ are fixed intercept and treatment effects. So, the authors should modify their manuscript more thoroughly.

5. In equation 4, it is $(\beta_0 + b_{0k})$ and not $(\beta_{0k} + b_{0k})$, the latter being unidentifiable.

6. On the forest plots (figures 3 to 5), the first OR (with confidence interval) should not be labelled as “Fisher’s exact test”, which is a test (as it names says), and not a method to estimate the odds ratio with its confidence interval. By the way, did the authors use an “exact” method to compute the confidence interval or did they use a more usual asymptotic method?

Discretionary Revisions

1. The reference 1 (Biau et al.) is appropriately cited page 5, but perhaps not for the first sentence of the introduction, since it did not specifically evaluate the evolution of the number of RCTs for medical and surgical interventions.

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