Reviewer's report

Title: A comparison of methods to account for centre effects in multicentre trials with a binary outcome

Version: 1 Date: 30 August 2013

Reviewer: Jinhui Ma

Reviewer's report:

This paper compares four analytical techniques that account for center effects in multicenter trials with a binary outcome. A re-analysis of a previously reported trial was conducted and a simulation study was implemented using different configurations of data. The results are evaluated in terms of Type I error rate and power. The paper concludes that the random-effects and GEE method generally perform better than other methods.

In general, the conclusion of this paper may not be surprising to any or most biostatisticians. The presentation of this paper could be much improved by providing much clear description for the Methods and Simulation Sections.

Major compulsory revisions:

In Section 1.1 Analysis methods, the author briefly described the four statistical analysis methods accounted for center effects. The models were not presented in a way that commonly accepted. For example, the author described that GEEs fit a logistic regression model

\[
\text{Logit}(Y) = \# + \# \text{treat}.
\]

To be precise, it should be described as \( \text{logit}(\text{prob}(Y=1)) = \# + X\# \), where \( x \) is a vector of covariates.

The descriptions for other statistical methods are not clear as well. For example, which working correlation matrix was specified for GEE method, how the weight was obtained for Mantel-Haenszel method? etc.

In section 1.2 Comparison between different analysis methods, the author wrote “A major difference between the analysis methods is the way in which they calculate the treatment effect”. This is so obvious since they are different analysis methods. I think you may want to convey the idea that different statistical methods use different strategies to handle the center effects.

The author mentioned that fixed-effects and Mantel-Haenszel provide a within-center comparison; they both calculated the treatment effect separately within each center, and then combine the results. This description applies to Mantel-Haenszel. I don’t see how it fits the fixed-effects analysis.

The author believe that “a random-effects model uses both within and between-centre estimates of the treatment effect, and generalized estimating equations incorporates neither within- or between-centre estimates, as it
completely ignores centre in its estimate of the treatment effect”. The description of the random-effects model and the GEE method in this section is not quite precise. Actually it is the within- and between-center variability, rather than within- and between-center estimates of the treatment effect, that were captured in random-effects model for estimating the treatment effect. In GEE method, the similarity of patients within the same center is captured by a working correlation matrix and consequently the center effect is accounted for in the model [1].


In Section 1.4 Simulation study, I have difficulties to judge whether the design of the simulation is appropriate or not since it is not clearly presented. First, the simulation model is not presented in a commonly accepted way as I’ve mentioned above. Second, it is not clear how the values of ICC were incorporated into the simulation model. Third, the author mentioned that sigma(i) and epsilon(i) were generated independently, but did not describe how they were generated. Fourth, the author set the event rate in control group to be 20% and 50%. How could readers connect this to the treatment effect, i.e. what is the true treatment effect in the simulation. Fifth, the simulation parameters include number of centers, number of patients, ICC, and block size, etc. The author should provide references to support the choices of the values for these parameters, i.e. are the simulated scenarios relevant in practice? Overall, the design of the simulation should be reported transparently between the author and the reader.

The author used type I error rate and power as the evaluation criteria for methods comparison. Usually bias, mean squared error, etc are also commonly used as evaluation criteria in the simulation study. According to the results from MIST2 data, odds ratios from random-effects model and GEE method are different, which also suggested that bias should be investigated for the comparison of these models.

Minor essential revisions
According to the results from the simulation study, the convergence rate for GEE method is very high (99.4% or higher). However, convergence was not achieved for GEE in re-analysis of MIST2 trial, which indicated that the non-convergence might be an issue in practice. Since the author recommended the use of GEE and random-effects model in practice, it might be helpful to explain in what cases GEE may not converge.

To adjust for other covariates, I think Mantel-Haenszel method can be improved by running standard logistic regression within each center with adjustment for covariates and then pooling the results to get the final estimated treatment effect.

The author used different block sizes to simulate the data, however, the impact of block sizes were not presented/discussed in the results and discussion section.
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 I have no competing interests.