Reviewer’s report

Title: The risks and rewards of covariate adjustment in randomised trials: an assessment of 12 outcomes from 8 studies

Version: 1

Date: 20 February 2014

Reviewer: Nick Parsons

Reviewer’s report:

Authors say that “…Adjusted analyses are not often performed in practice…” and also “…unadjusted analyses dominate in practice; reviews have found that between 24-34% of trials use covariate-adjustment for their main analysis…”. This continues to surprise me. I have analysed many clinical trials, and reviewed many statistical analysis plans, and in my experience the majority include covariate adjustment. Clearly I see a biased sample, no doubt due to the assiduousness of my clinical colleagues and the promotion of good practice by the funding bodies, such as NIHR, that require research proposals to be scrutinized carefully by (well-qualified) statisticians!

Personally, I don’t find the concept of retrospective power analysis to be very appealing or informative. The important issue is the modification in the treatment effect estimate and/or the increase in precision of treatment effect estimates that results from the inclusion of covariates. The authors follow this line of reasoning by for instance reporting that in the MIST2 study, “…accounting for baseline effusion size in the analysis resulted in a large reduction in the SE for treatment (unadjusted SE 4.3 vs adjusted SE 2.8, a 35% reduction)…”, and then continue to say that this leads to “…a substantial increase in power….”. I do not see why the increase in power is important or enlightening, after reporting a substantive decrease in the SE’s. Could the authors explain why they think we should be interested particularly in retrospective power? I note their arguments regarding binary and time-to-event outcomes, but to my mind a standardized effect estimates (z-statistic, as they report in text), would be simple enough to understand in the settings described.

The authors state that “…it is recommended that stratification factors be accounted for in the trial analysis….” and that “…pre-specifying which variables will be included in the analysis in the protocol or analysis plan will avoid bias, and give more credibility to the trial results….”. These are drummed into trialists/statisticians here, so should be routine. Again, it surprises me that these issues are not widely known.

The authors are correct that each continuous or binary baseline covariate included in the analysis uses at least one degree of freedom. However, my view is that this is rarely, if ever, an issue of much importance in practice unless the trial is very small. And in those case one would be naïve in the extreme to attempt to fit complex models. The authors state that “…in theory, there should be no missing data for variables collected at baseline…”. In practice there is
often less missing data at baseline, but I don’t agree that in theory there should be none. This is very dependent on trial process and the data being collected; it is usually less than at later follow-up occasions, but there is generally always some missing data at baseline.

Overall, I found this to be a well written manuscript that will be acceptable for publication after some minor revision. As I stated above, I don’t like the concept of retrospective power in this setting, but I do concede that it makes presentation and comparison of the results of the simulation studies much simpler. I would prefer to see data showing effect estimates and SE’s, as to my mind this would be more informative. I leave to the authors to consider either modifying their presentation of results or perhaps providing these data as supplementary material. This is a suggestion (discretionary rather than compulsory revision) only. Overall, although the article was interesting, I found nothing here surprising or that would cause me to change my working practice; other than perhaps to be bolder in suggesting suspected (but not known) prognostic variables in a trial analysis plan. My view is that where there is good practice, then statisticians and trialists will be aware of most or more often all of the material presented here. That is not to diminish the message, but clearly those who are not aware of these issues are probably unlikely to find their way to this article. The fact that, as the authors report, so few trials report adjusted analyses suggest that the issues highlighted here are still not widely appreciated. This may not be the best forum for tackling what are probably issues related to management, training or support for junior staff, but I think the authors should consider saying how they think the poor practice they highlight here might be addressed; for instance, requiring all trial protocols to be published in peer reviewed journals, or encouraging statisticians to disseminate this information as widely as possible. Without some more concerted action, I suspect that the poor practice reported here will continue.

**Level of interest:** An article of importance in its field

**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'