Reviewer's report

Title: Changing cluster composition in cluster RCTs: design and analysis considerations

Version: 1 Date: 13 December 2013

Reviewer: Monica Taljaard

Reviewer's report:

This is the first study, to the best of my knowledge, to specifically address the issue of post-randomization cluster merges. Although such merges do not appear to be a common occurrence (at least based on the empirical evidence presented here), it would nevertheless be helpful to have guidance for how to handle such occurrences in the design and analysis of cluster randomized trials. Fortunately the authors' results show that pragmatic common-sense approaches are reasonable without bias or loss of precision. Cluster merging, predictably, can lead to a loss in power, but interestingly, the loss in power may be attenuated by a corresponding change in observed ICC. Also not surprising is that merging of clusters originally allocated to different arms tends to bias towards the null if the merged clusters are retained in the analysis, although I question whether researchers would in reality analyze merged clusters this way.

The authors present some preliminary recommendations, namely that cluster merges be avoided where possible, clusters be discontinued where merges are heterogeneous, and to allow for this possibility in sample size calculation. Although further methodological development is required, this study prepares the groundwork for such development. The manuscript is well-written and the study is well conducted, and I believe it will be a worthwhile contribution to the literature. I would therefore recommend publication after my concerns below have been addressed.

Major Compulsory Revisions

1. In the introduction, I think readers might find it helpful to see a brief mention of or explanation of some of the practical considerations with respect to analyzing homogeneous or heterogeneous merges. For example, under which circumstances is it feasible to analyze the merged clusters separately? Under which circumstances might one consider analyzing them as one cluster? Does it depend on how patient recruitment is done (e.g., sequential recruitment or identification and enrolment prior to randomization? Cohort versus cross-sectional design?) Under which practical circumstances might researchers consider assigning a merged cluster to either treatment group, as opposed to discontinuing the clusters or analyzing them separately. This would go a long way towards setting up the circumstances where guidance is needed.

2. You apply a simple correction for attrition in equation 4, by dividing by 1 minus the attrition rate, while keeping the cluster size constant at m (as opposed to
using $m$ reduced by attrition. Which assumptions underlie this simple correction? In previous work, we have shown that this approach is equivalent to assuming that entire clusters are lost to follow-up. (See Taljaard M, Donner A, Klar N. Accounting for expected attrition in the planning of community intervention trials. Statistics in Medicine 2007; 26(13): 2615-2628.) You conclude that allowance for loss to follow-up at cluster level as well as individual level might be advisable at the planning stage of a cluster RCT, but don’t give further information about how this can be done.

3. Scenario 3: This scenario seems somewhat artificial: why would one analyze clusters merged after completion of the intervention as a single cluster? You then consider three strategies: merged clusters analyzed in the control arm; merged clusters analyzed in the intervention arm; merged clusters dropped from analysis. What about the fourth option: namely analyze merged clusters as separate clusters in the arms to which they were allocated? Some clarification will be helpful here.

4. In your simulation study, you assume three different levels for cluster sizes (20, 40, and 100). In some cluster randomized trials, cluster sizes are often much larger (e.g., up to 500 or up to 1000 patients per cluster). You don’t present detailed results for the different levels of cluster sizes, but is it possible to speculate about how your conclusions might change in the presence of much larger cluster sizes?

Minor Essential Revisions

1. Page 4 first paragraph: Description of number and size of GP practice: specify that this applies in the UK specifically?
2. Page 4: In your review of reported incidences of cluster merging, why did you choose to focus specifically on primary care? Are merges less likely to occur in cluster randomized trials conducted in other settings? As well, it will help to present further details of your Medline search strategy to allow the reader to interpret the results of your literature search.
3. Page 7: (1-beta) tending to infinity: should this be 1?
4. Page 14-15: When you present the results for scenarios 3 and 4, it may be helpful to briefly remind the reader what the scenarios were, as you did for scenarios 1 and 2.
5. Please check the first sentence in Conclusions - does not appear to be complete.

Discretionary Revisions

1. Page 7: You state that "variability in cluster size has a detrimental effect on study power." Although I cannot find the reference at the moment, I thought that previous work has shown that in most cases, an inflation by approximately 10% is adequate to account for variability in cluster sizes, which hardly seems detrimental. I am wondering if this statement needs to be softened.
2. Page 18: You mention the issue of cluster membership which may fluctuate
during the course of a study. This issue was previously discussed by Diehr et al (Optimal survey design for community intervention evaluations: cohort or cross-sectional. J Clin Epidemiol 1995. Vol 48(12): 1461-1472.) It may be worthwhile citing this work.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.