Author's response to reviews

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

Authors:

Neil Corrigan (n.corrigan@leeds.ac.uk)
Michael JG Bankart (m.j.g.bankart@keele.ac.uk)
Laura J Gray (lg48@leicester.ac.uk)
Karen L Smith (karen.smith@leicester.ac.uk)

Version: 2 Date: 6 March 2014

Author's response to reviews: see over
Dear Editor,

Response to the Reviewer’s comments

We thank both reviewers for their insightful comments, suggestions, and useful references. Our responses to the revisions are given below.

Reviewer: Monica Taljaard

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.

We have included a brief description of the design issues that would need to be considered in making such decisions in the Background section. In the Discussion we have indicated that in the main statistically homogeneous merges can be treated as one cluster without the introduction of bias and conclude that heterogeneous merges should lead to discontinuation of clusters.

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.

Equations 3 to 6 have been amended to allow for attrition within clusters, rather than attrition of entire clusters, to reflect the usual approach for allowance for loss to follow-up in cluster RCTs. We suggest that allowance for loss to follow-up at the cluster level be dealt with using the same methods as used in accounting for variability in cluster size and loss of clusters, and indeed that the method proposed by Taljaard et al (in the above referenced paper) would be one such method.

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.

We agree and emphasise that this scenario is indeed artificial and is included to provide a baseline against which the results from scenario 4 can be assessed. The fourth option suggested by the reviewer is in effect the same as ignoring the cluster merging and consequently would give the same results as with zero merges in Table 2.

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?

A paragraph has been added to the Discussion, which explains that the simulation results showed no difference by cluster size, only by the proportion of clusters merging. The cluster sizes used for the simulation study were chosen to reflect typical cluster sizes in cluster RCTs in primary care as indicated by the paper by Eldridge et al added as reference 18.

Minor Essential Revisions

1. Page 4 first paragraph: Description of number and size of GP practice: specify that this applies in the UK specifically?

This has been amended.

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.

We chose to focus on primary care due in part to the changes that have happened in the UK over recent years, and because two of the authors were asked about this situation in an active, and as yet unreported, cluster RCT in primary care. We have added this rationale to the Background and have indicated in the Discussion that the results would be applicable to other situations in which cluster RCTs are conducted and where there may be a risk of clusters merging. Our literature search was not a systematic review, and therefore the search strategy was simple and probably not exhaustive, but the search terms have now been included.

3. Page 7: (1-beta) tending to infinity: should this be 1?

This has been corrected.

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.

This has been done.

5. Please check the first sentence in Conclusions - does not appear to be complete.

The sentence is complete. A comma has been inserted to aid readability.
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.

I suspect the paper is that by Eldridge, Ashby and Kerry (full reference is in point 9 raised by the other reviewer), in which they examined the values of the coefficient of variation in cluster size which ensures that the maximum underestimate in sample size is <10%. In this paper the authors concluded that when cv <0.23 it would not be necessary to adjust the sample size. However they also indicate that 'Trials randomising UK general practices commonly have cvs ~0.65, which can result in sample size increases of up to 42%. We are comfortable with our choice of the word 'detrimental'.

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.

This has been done.

Reviewer: Clare Rutterford

Major compulsory revisions

1. This is a topic that I have not seen covered before however there were only 1/211 (0.5%) reported incidences of cluster merging and 1/27 (4%) unreported instances, which implies this problem occurs with relatively low frequency. I would have liked to have seen more emphasis about the importance of addressing this at the design stage. The authors recommend that additional clusters be recruited to anticipate cluster merging, however I wonder how realistic this is in practice given the potentially high costs associated with the recruitment of an extra cluster in comparison to the relatively low occurrence of cluster merging, and the fact that the clusters may merge on purely an administrative level only.

We have explicitly acknowledged the low occurrence rate. Two of the authors have previously been asked how to handle a cluster RCT in which two of the clusters were merging, so although this does not happen often it is nonetheless problematic in those trials where it does occur as we discuss in the Background. As we have indicated we don’t think that administrative merges constitute a difficulty so any decisions about allowance for cluster merges would need to be taken by a research team considering their own specific situation.

2. For those who may not be as familiar with cluster randomised trials I think it would be helpful to explicitly describe what is meant by cluster size in this paper. Are the authors assuming that cluster size is the entire cluster, or just the subset from the cluster who are being analysed? I think the authors are assuming the former. The ICC is related to the size of the entire cluster, so as clusters merge the natural cluster size is increasing and this affects the ICC. The underlying value of the ICC does not change as the numbers sampled from a cluster
increase, it is just estimated with greater precision. I think this relationship needed to be more explicitly described.

Indeed we are assuming the former. We have made this clear in the Discussion.

3. This paper started from the assumption of fixed cluster size, so again it is not clear whether the authors are talking about sampling a fixed number or if this refers to the entire cluster. How reasonable is this assumption given the background to the paper is primary care where cluster sizes are large and variable? If variable cluster size is accounted for in the design does cluster merging have much of an effect?

A paragraph has been added to the Discussion. We expect that the impact may not be great if variability in cluster size has already been taken into account but suggest that further work would be needed to fully understand this.

4. In my mind this topic seems to be connected to the sample size issues of variable cluster sizes and cluster drop out, which have been discussed by others. It would be useful if the authors could make the distinction or similarity between these issues and cluster merging more explicit.

This has been raised in the Discussion, in particular with regard to how allowance for loss of clusters might be made using the methods allowing for variability in cluster size.

Minor essential revisions
5. In figure 2, at 0 merges in each treatment group should the blue line not be at 80% power?

This figure has been amended.

Discretionary revisions
6. From a practical perspective I wonder how easy it is for researchers to find or choose appropriate ICC estimates that reflect cluster size as recommended by the authors.

Finding and choosing an appropriate ICC estimate is always difficult. Some of the published papers that include ICC estimates do also contain information on cluster size, but not all. So if researchers follow current guidance on choosing an appropriate ICC, that is to look at multiple sources and consider patterns in ICCs, this should be no more challenging than other issues and we have chosen not to go into the detail in the paper.

7. It would be helpful to state in the introduction that this paper excludes scenarios where a cluster may merge with one which is not already participating in the trial

We have made this point in the Discussion and have chosen not to also discuss this in the Background and Methods since that section is already lengthy.

8. On page 5 the design effect method of sample size calculation is only one approach, albeit the most common, rather than the only approach.

This has been acknowledged.

9. On page 8 equation 7 has also been shown by Eldridge, Kang and Manatunga Eldridge SM, Ashby D, Kerry S. Sample size for cluster randomized trials: effect

These papers have now been cited.

10. On figure 1 it would also be useful to include 5 on the x axis for ease in seeing the relationship as k tends to c/2

This has been added to the figure.

11. Figure 4 should appear prior to the discussion

Reference is now made to this figure in the Results section and the ordering of Figures 3 and 4 changed.

12. Page 19, remind the reader what is meant by the pragmatic analysis

This has been done.

Yours sincerely,

Karen L Smith
Neil Corrigan
Michael JG Bankart
Laura J Gray