Letter to the Editor

Response to: How to design and analyze cluster randomized trials with a small number of clusters? Commentary on Leyrat, Morgan, Leurent, Kahan, by Van Breukelen and Candel.

Clémence Leyrat, Katy E Morgan, Baptiste Leurent and Brennan C Kahan

We would like to thank Van Breukelen and Candel for their comments on our manuscript (1). Although we broadly agree with them, we would like to clarify several points.

First, they argue that our results can be understood in light of the existing literature. We agree that some of the results in our article are well known (e.g. that unweighted cluster-level analyses lose efficiency). However, these approaches are still commonly used (2), and so we included them in order to empirically demonstrate the benefit of other approaches. Furthermore, we are unaware of any empirical comparison between GEEs, mixed-effect models and cluster-level analyses for continuous outcomes. We agree with Van Breukelen and Candel that some theoretical results are available for these approaches, however these are often based on approximations which do not always translate to realistic scenarios (particularly regarding small-sample corrections), and so it is useful to assess the properties of these approaches across a range of realistic scenarios using simulation (3).

Second, Van Breukelen and Candel take issue with the sample size formula used in our simulation study. Because sample size formulae depend on the underlying analysis model, there is no single formula which is appropriate for all the analysis methods being compared. However, our aim was to benchmark the *relative* performance of each analysis method in terms of type-I error rate and power. Given that the specific sample size formula used will have no impact on which analysis approach performs best, we are unsure why Van Breukelen and Candel have taken issue with this. We do however thank them for pointing out the typo in the formula in the appendices. We have now added a corrigendum.

Third, Van Breukelen and Candel describe a straightforward approach for determining the number of additional clusters required to compensate for the loss of power due to a small-sample correction. We agree that increasing the number of clusters is the best approach, however this is not always feasible due to lack of resources. Investigators may therefore wish to increase the number of participants per cluster rather than the number of clusters (4), if feasible. Simulation provides a straightforward approach for this. We also note that estimating the number of clusters and participants via simulation based on the expected characteristics of the trial will provide a more precise estimate of the total number of participants required than the authors' sample size approximation, and will ensure that investigators do not recruit more participants than required (5).

References

1. Leyrat C, Morgan KE, Leurent B, Kahan BC. Cluster randomized trials with a small number of clusters: which analyses should be used? Int J Epidemiol. 2018 Feb 1;47(1):321–31.

2. Kahan BC, Forbes G, Ali Y, Jairath V, Bremner S, Harhay MO, et al. Increased risk of type I errors in cluster randomised trials with small or medium numbers of clusters: a review, reanalysis, and simulation study. Trials. 2016 Sep 6;17(1):438.

3. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. ArXiv171203198 Stat [Internet]. 2017 Dec 8 [cited 2018 Mar 12]; Available from: http://arxiv.org/abs/1712.03198

4. Hemming K, Girling AJ, Sitch AJ, Marsh J, Lilford RJ. Sample size calculations for cluster randomised controlled trials with a fixed number of clusters. BMC Med Res Methodol. 2011 Jun 30;11:102.

5. Hemming K, Taljaard M, Forbes G, Eldridge SM, Weijer C. Ethical implications of excessive cluster sizes in cluster randomised trials. BMJ Qual Saf. 2018 Feb 20;bmjqs-2017-007164.