

Estimands in cluster trials

Hemming, Karla; Taljaard, Monica

DOI:

[10.1093/ije/dyac174](https://doi.org/10.1093/ije/dyac174)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Hemming, K & Taljaard, M 2022, 'Estimands in cluster trials: thinking carefully about the target of inference and the consequences for analysis choice', *International Journal of Epidemiology*, pp. 1-3.
<https://doi.org/10.1093/ije/dyac174>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Commentary [Invited only]

Estimands in cluster trials: thinking carefully about the target of inference and the consequences for analysis choice

Karla Hemming ^{1*} and Monica Taljaard^{2,3}

¹Institute of Applied Health Research, The University of Birmingham, Birmingham, UK, ²Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada and ³School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Ottawa, Canada

*Corresponding author. Institute of Applied Health Research, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK. E-mail: k.hemming@bham.ac.uk

Received 8 June 2022; Editorial decision 11 August 2022; Accepted 17 August 2022

Specifying the target of inference in cluster trials

Cluster randomized trials (CRTs) are complex.¹ At the protocol development stage, we have to select an appropriate unit of randomization (which may depend on the unit of intervention delivery) and an appropriate unit of analysis (which may depend on the unit of observation).^{2,3} If the unit of randomization is different from the unit of analysis, we must account for clustering among multiple observations from the same cluster—a requirement that is well appreciated.² The unit of analysis may be either the individual or the cluster, with the choice ideally made on statistical grounds (although in practice it may reflect personal preference, convenience or experience).³ The paper by Kahan and colleagues advises us that we also need to choose an a priori unit of inference and this choice is critical in selecting both the unit and the method of analysis.⁴ We believe that the need to consider the target of inference before specifying the method of analysis has not received adequate attention in the cluster trials literature to date.

Defining the unit of inference, i.e. the estimand of interest, is essentially about carefully specifying the research question.^{5,6} Specifically, we must consider whether interest lies in determining:

- i. the effect of the intervention on a typical individual, or
- ii. the effect of the intervention on a typical cluster.

It is important to realize that the estimated treatment effect for these two questions can differ in the same trial for the same outcome. In particular, the treatment effect will differ when cluster sizes are informative—which essentially means either the outcomes vary across clusters depending on cluster size and/or the treatment effect varies across clusters depending on cluster size (i.e. a cluster size by treatment interaction).^{7,8} Informative cluster sizes are not expected to be uncommon in practice. The precise circumstances under which these treatment effects differ depend on the type of treatment effect (i.e. the measure of association). For mean difference, risk difference or relative risk, differences only arise if there is an interaction between the treatment effect and cluster size; for odds ratios and hazard ratios, the differences arise even if there is no treatment by cluster size interaction, but the baseline prevalence varies across cluster sizes.

Choosing an analytical strategy to match the unit of inference

It is also important to realize that choosing the desired unit of inference is distinct from choosing the unit of analysis: regardless of whether the effect on a typical individual or typical cluster is of interest, it is possible to conduct either

an individual-level analysis or a cluster-level analysis.⁴ However, exactly how to carry out these analyses to ensure they answer the question of interest requires careful consideration. In the case of an individual-level analysis, two commonly used methods are the generalized linear mixed model (GLMM) or generalized estimating equations (GEEs).^{3,9} Whilst these individual-level methods of analysis have advantages, both GLMM and GEE models can produce biased estimates when cluster sizes are informative.^{4,7} Conversely, a cluster-level analysis can yield an unbiased estimate for the effect on the typical cluster even when cluster sizes are informative; however, if we choose a weighting method to preserve statistical efficiency, we can inadvertently introduce a bias.⁴

A cluster-level analysis is probably the least-commonly used approach in practice, even though it generally produces valid inferences.⁹ When clusters vary in size, clusters may be weighted to improve efficiency.¹⁰ However, Kahan *et al.* show that the question of when a cluster-level analysis should be weighted turns out to be subtler. In fact, a cluster-level approach, without any weighting of cluster sizes, even when cluster sizes vary, will allow estimation of the impact of the intervention for the average cluster. Weighting by cluster sizes, i.e. by the number of individuals within that cluster, changes the target of inference to the average individual. Thus, the question of whether a cluster-level analysis should be weighted is not first a question of a gain in statistical efficiency, but rather about whether the objective is to estimate the impact of the intervention for the average cluster or the average individual.

Kahan *et al.* identify what might be considered a more surprising result. Suppose it is of interest to estimate the impact of the intervention on the average individual. Here, a common approach is to use a GLMM or GEEs assuming a working exchangeable correlation structure.³ It transpires that both of these approaches do not target the effect for the average individual when cluster size is informative. This means that whenever a generalized estimating equation (with exchangeable correlation structure) or mixed model is used to evaluate the impact of the intervention for the average individual, this will yield a biased estimate of the effect (again when cluster size is informative). The recommended approach is to use GEEs assuming an independent working correlation structure or ordinary regression with cluster-robust standard errors.⁴

Other issues to consider when specifying the target of inference

Thinking carefully about the target of inference is not a new concept.^{9,11} There are of course other factors to

consider when thinking about the target of inference and choosing the method of analysis in cluster trials—notably whether interest is in the marginal (typically obtained via GEEs, but can be obtained via GLMMs) or cluster-specific effects (obtained via GLMMs).¹² The marginal effect allows us to consider the effect of the intervention for a population of clusters and individuals similar to those included in the trial at hand (useful for making population-level decisions) whereas cluster-specific effects allow consideration of the impact for a typical individual within a specific cluster. Exactly how to estimate a cluster-specific effect in the presence of informative clustering (and without being able to use a GLMM) remains unclear. In longitudinal designs, the target of inference also has a bearing on whether a cohort or cross-sectional design is more appropriate. If the unit of inference is the community, cross-sectional sampling might be best; if the unit of inference is the individual, cohort sampling might be best.¹³

Implications

CRTs are used to evaluate a diverse range of interventions. Sometimes interest will be on the impact on the average individual, perhaps when evaluating a drug. In other settings, they might be used to evaluate an implementation strategy or an education package targeting healthcare providers. In these settings, we might be more interested in the impact of the intervention on the average cluster. The same trial may have several objectives linked to different outcomes and it might be the case that the target of inference differs for different outcomes. Even for the same outcome, we may want to draw inferences at both cluster and individual levels. Kahan's paper underscores the need to clearly define the target of inference, which will involve discussions with trial investigators to determine whether interest is in the impact on the typical cluster or typical individual and then to select an analysis method that is likely to provide an unbiased estimate, irrespective of whether the cluster size is informative (as this will be mostly difficult to rule out).⁴ This is likely to make trials somewhat less efficient. More work is needed to address unanswered questions, such as methods of analysis for multiple period cluster trial designs, power considerations and methods for covariate adjustment particularly when using cluster-level approaches. Future reporting statements in cluster trials could be improved to require authors to clearly define their unit of inference and describe how they handled informative cluster sizes.

Author contributions

Both authors made an equal contribution to the writing of the manuscript.

Conflict of interest

None declared.

References

1. Eldridge S, Kerry S. *A Practical Guide to Cluster Randomised Trials in Health Services Research*. Chichester: Wiley, 2012.
2. Murray DM. Influential methods reports for group-randomized trials and related designs. *Clin Trials* 2022;**19**:353–62.
3. Turner EL, Prague M, Gallis JA, Li F, Murray DM. Review of recent methodological developments in group-randomized trials. Part 2: Analysis. *Am J Public Health* 2017;**107**:1078–86.
4. Kahan BC, Li F, Copas AJ, Harhay MO. Estimands in cluster-randomised trials: choosing analyses that answer the right question. *Int J Epidemiol* 2022;doi:10.1093/ije/dyac131.
5. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. https://www.ema.europa.eu/en/documents/scientific-guide-line/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf (23 May 2022, date last accessed).
6. Jin M, Liu G. Estimand framework: delineating what to be estimated with clinical questions of interest in clinical trials. *Contemp Clin Trials* 2020;**96**:106093.
7. Seaman S, Pavlou M, Copas A. Review of methods for handling confounding by cluster and informative cluster size in clustered data. *Stat Med* 2014;**33**:5371–87.
8. Yelland LN, Sullivan TR, Pavlou M, Seaman SR. Analysis of randomised trials including multiple births when birth size is informative. *Paediatr Perinat Epidemiol* 2015;**29**:567–75.
9. Hayes RJ, Moulton LH, *Cluster Randomised Trials*, 2nd edn. New York: Chapman and Hall/CRC, 2017.
10. Rutterford C, Copas A, Eldridge S. Methods for sample size determination in cluster randomized trials. *Int J Epidemiol* 2015;**44**:1051–67.
11. Campbell MK, Piaggio G, Elbourne DR, Altman DG; CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;**345**:e5661.
12. Youngjo L, Nelder JA. Conditional and marginal models: another view. *Stat Sci* 2004;**9**:219–38.
13. Atienza AA, King AC. Community-based health intervention trials: an overview of ethodological issues. *Epidemiol Rev* 2002;**24**:72–79.