Type I Error Control for Cluster Randomized Trials Under Varying Small Sample Structures

CURRENT STATUS: UNDER REVIEW

BMC Medical Research Methodology BMC Series

Joshua Nugent

University of Massachusetts Amherst School of Public Health and Health Sciences

Ken Kleinman

University of Massachusetts Amherst

✉ ken.kleinman@gmail.com *Corresponding Author*

ORCID: <https://orcid.org/0000-0002-5600-5834>

DOI:

10.21203/rs.2.17855/v1

SUBJECT AREAS

Health Economics & Outcomes Research

KEYWORDS

Linear mixed models, Wald test, Likelihood ratio test, Type I error

Abstract

Background: Linear mixed models (LMM) are a common approach to analyzing data from cluster randomized trials (CRTs). Inference on parameters can be performed via Wald tests or likelihood ratio tests (LRT), but both approaches may give incorrect Type I error rates in common finite sample settings. The impact of interactions of cluster size, number of clusters, intraclass correlation coefficient (ICC), and analysis approach on Type I error rates have not been well studied. Reviews of published CRTs find that small sample sizes are not uncommon, so the performance of different inferential approaches in these settings can guide data analysts to the best choices.

Methods: Using a random-intercept LMM structure, we use simulations to study Type I error rates with the LRT and Wald test with different degrees of freedom (DF) choices across different combinations of cluster size, number of clusters, and ICC.

Results: Our simulations show that the LRT can be anti-conservative when the ICC is large and the number of clusters is small, with the effect most pronounced when the cluster size is relatively large. Wald tests with the Between-Within DF method or the Satterthwaite DF approximation maintain Type I error control at the stated level, though they are conservative when the number of clusters, the cluster size, and the ICC are small.

Conclusions: Depending on the structure of the CRT, analysts should choose a hypothesis testing approach that will maintain the appropriate Type I error rate for their data. Wald tests with the Satterthwaite DF approximation work well in many circumstances, but in other cases the LRT may have Type I error rates closer to the nominal level.

Full-text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed.

However, the manuscript can be downloaded and accessed as a PDF.

Figures

Figure 1

Relationship between Type I error rate and design factors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

[tie_rates_LMM.bib](#)
[bmc-mathphys.bst](#)
[vancouver.bst](#)
[tie_rates_Lmm.bbl](#)
[bmcart-biblio.sty](#)
[tie_rates_LMM.tex](#)
[bmcart.cls](#)
[spbasic.bst](#)