NIH PRAGMATIC TRIALS COLLABORATORY
Rethinking Clinical Trials®

Unequal Cluster Sizes in Cluster-Randomized Clinical Trials

Background

Cluster-randomized designs are commonly used in pragmatic clinical trials (PCTs), where individual-level randomization is often infeasible due to practical implementation issues. In addition, cluster-randomized designs allow researchers to avoid contamination across intervention groups. For example, physicians in the same clinic who are randomized to different interventions might change their behaviors by learning from other physicians with different intervention assignments [1,2].

Randomization and the unit of analysis

Furthermore, individual-level randomization may not constitute the desired unit of analysis from which to make inferences. For example, in the case of physician-oriented interventions such as providing information about the likelihood of false-positive test results to help direct better clinical practice, randomization at the physician level rather than the patient level may be more feasible and may allow more accurate estimation of a physician- *vs* patient-level effect. Another common circumstance involves randomization at the level of the clinic, because while an intervention may not be practical to distribute at the physician or patient level, the unit of analysis to generalize the effect estimate may still be at the physician or patient level.

Equal cluster sizes

In the case of equal cluster sizes (i.e., the number of patients per physician or clinic is balanced), most available analytic approaches generalize to all units of analyses, except for some random-effect approaches for binary/survival outcomes yielding within-cluster interpretation. Further, much of the literature addressing cluster-randomized trials has focused on equal clusters of equal size, including the power tradeoffs and simple sample size calculations using different approaches [1,3,4].

Unequal cluster sizes

In the scenario of unequal cluster sizes, which are most common in PCTs, the decisions regarding at what cluster level to randomize and how to account for the clustering in both the sample size calculation and the analysis are harder issues to address. Clustering levels are often chosen on the basis of feasibility, including contamination considerations. However, given the unit of clustering, one has choices in analysis to determine the population to which to weight the estimate. Specifically, one can weight each cluster as equal regardless of the number of observations within the cluster, weight to the lowest individual level available, or somewhere in between.

Prepared by: Elizabeth DeLong and Andrea Cook
Version: 1.0, last updated June 26, 2014

Power tradeoffs

In most cases, these choices will be made with a view to optimally addressing the scientific question of interest, but statistical tradeoffs, including power, must also be considered. Further, given a particular unit of analysis, there are different analytical approaches that will present different power tradeoffs. Some of these have been explored recently for continuous outcomes, but the literature addressing tradeoffs for binary or survival outcomes remains limited [5].

Binary outcomes

For binary outcomes in particular the analysis choice is more complicated, especially when using a standard logistic regression approach to estimate an odds ratio that adds the complexity of within-cluster *vs* marginal-cluster average effects [6].

Next steps

Considering the relative paucity of work addressing unequal cluster-randomized studies and their associated complications, as well as the likelihood that the use of cluster designs in PCTs is likely to increase, a review article summarizing the different approaches and a simulation study to evaluate their comparative performance would constitute valuable contributions to the literature. The first step in such an undertaking would involve explaining the different choices that inform efforts to estimate the three general populations from which to derive inferences: cluster-level, individual-level within cluster, and between-cluster and individual-levels.

Simulation study

Given the choice of estimate and population to generalize, it would be of interest to conduct a simulation study comparing different analytic approaches by varying the number of clusters, the average number of individuals within clusters, and the distribution of individuals within clusters (e.g., a few very large clusters and many smaller clusters; half large clusters and half small; equal-size clusters; etc.) ranging across distributions observed, or likely to be observed, in the NIH Collaboratory, and the strength of correlation within the cluster (the intraclass correlation coefficient, or ICC).

We propose to focus on binary outcomes because the literature with regard to these outcomes is the most limited, despite the fact that binary outcomes are a common outcome type for healthcare intervention studies. We would also vary the outcome rate in the control group and the strength of association of the intervention effect. We would assume that there is approximate balance in the number of individuals and clusters per intervention group because it is standard and good practice to stratify randomization for cluster-randomized studies on cluster size and other important predictors of the outcome [2]. The goals of this review paper and simulation study would be to summarize different approaches to analyzing cluster-randomized studies and the complications of unequal cluster sizes, and to provide guidance through simulation on potential approaches that might be most useful for analyzing such data.

Resources

1. Donner A, Klar N. Design and Analysis of Cluster Randomized Trials in Health Research. Oxford University Press Inc.: New York; 2000.
2. Murray DM, Varnell SP, Blitstein JL. Design and analysis of group-randomized trials: A review of recent methodological developments. Am J Public Health 2004;94(3): 423–32. PMID: 14998806. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1448268/. Accessed May 14, 2014.
3. Ma J, Raina P, Beyene J, Thabane L. Comparison of population-averaged and cluster-specific models for the analysis of cluster randomized trials with missing binary outcomes: a simulation study. BMC Med Res Methodol 2013;13:9. PMID: 23343209. doi: 10.1186/1471-2288-13-9. Available at: http://www.biomedcentral.com/1471-2288/13/9. Accessed May 14, 2014.
4. Ukoumunne OC, Carlin JB, Gulliford MC. A simulation study of odds ratio estimation for binary outcomes from cluster randomized trials. Stat Med 2007;26(18):3415:3428. PMID: 17154246. doi: 10.1002/sim.2769. Available at: http://onlinelibrary.wiley.com/doi/10.1002/sim.2769/abstract. Accessed May 14, 2014.
5. Eldridge SM, Ashby D, Kerry S. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. Int J Epidemiology 2006;35(5) 1292–300. PMID: 16943232. doi: 10.1093/ije/dyl129. Available at: http://ije.oxfordjournals.org/content/35/5/1292.long. Accessed May 14, 2014.
6. Diggle P, Heagerty P, Liang KY, Zeger S. Analysis of Longitudinal Data. Oxford Statistical Science Series, 25. 2nd edition. New York: Oxford University Press; 2002.

Prepared by: Elizabeth DeLong and Andrea Cook

Version: 1.0, last updated June 26, 2014