Computational Statistics and Data Analysis 56 (2012) 1174–1187

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Sample size estimation in cluster randomized trials: An evidence-based perspective

Michael Rotondi *, Allan Donner

Department of Epidemiology and Biostatistics, The University of Western Ontario, London, Canada

ARTICLE INFO

Article history:
Available online 25 January 2011

Keywords:
Sample size estimation
Cluster randomized trials
Intracluster correlation coefficient
Meta-analysis
Design of experiments
Sequential methods

ABSTRACT

The evidence-based perspective to sample size estimation determines appropriate trial size by examining its potential impact on the literature. This approach is extended to determine the appropriate size of a planned cluster randomized trial by considering the role of the planned trial on a future meta-analysis (including current literature and the proposed study). A simulation-based algorithm allows consideration of variable cluster sizes and intracluster correlation coefficient values in conjunction with three approaches to sample size estimation, namely the power-based, variance reduction and non-inferiority perspectives. Two examples employing the sample size estimation techniques described are discussed in detail, while appropriate code is provided in the accompanying R package *CRTSize*.

1. Introduction

Cluster randomized trials have become increasingly popular in medical research (Altman, 2000) as the advantages and complexities of this experimental design become increasingly well-understood (Bland, 2004). In contrast to their individually randomized counterparts, a cluster randomized trial randomizes an entire social unit (for example, physician practice, family, or community) to treatment or control group status (Donner and Klar, 2000, Chap. 1).

Reasons for adopting this design include administrative convenience, reduction of between-group contamination, as well as a possible amelioration of ethical concerns (Donner and Klar, 2004). Furthermore, in some cases, such as designing a trial to assess the effect of municipal water fluoridation on tooth decay, or a community-based intervention to reduce smoking rates, randomization by cluster represents the only possible option.

The difficulty of such an allocation strategy is that responses within the same cluster are often positively correlated. That is, individuals within the same social unit tend to respond similarly. The degree of similarity among individual responses within the same cluster is measured by the intracluster correlation coefficient (ICC), denoted by ρ. This parameter, along with the (mean) cluster size (m) forms a natural basis for the required adjustments to classical sample size estimation formulae, ensuring that a proposed cluster randomized trial is adequately powered.

Although a well designed randomized trial is the accepted 'gold standard' for demonstrating the effectiveness of an intervention, a single randomized trial continues to have a smaller role with the advent of meta-analysis and 'evidence-based' medicine (Sutton et al., 2007). Perhaps the greatest indication of this effect is the overwhelming success of the Cochrane Collaboration (2010), which has performed several thousand systematic reviews and meta-analyses for a variety of therapeutic interventions. As such, it may be desirable to design a cluster randomized trial with the intention of ensuring that a subsequent meta-analysis (including both the previously completed and the planned trial) is adequately powered to detect a clinically important and statistically significant treatment effect. Although this approach may require a larger

* Corresponding author.
E-mail address: mrotondi@uwo.ca (M. Rotondi).

0167-9473/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2010.12.010

sample size than may be available to the investigator, these calculations provide insight into the potential impact of the planned trial on the literature.

The method proposed here involves modifications of the method developed by Sutton et al. (2007). We generalize these results to the case of cluster randomized trials with a binary outcome, recognizing that methodological extensions to the case of a continuous outcome are straightforward. Our discussion focuses on the challenges unique to this application, including those involving ρ and m, as well as a review of meta-analysis techniques for cluster randomized trials with a binary outcome. Although the approach exhibits Bayesian aspects, namely through the use of potential prior distributions on ρ and m, we apply the technique to the frequentist concept of hypothesis testing, which underlines the power-based, variance reduction and non-inferiority perspectives. In this manner, the proposed sample size estimation approach may be viewed as a hybrid of the two paradigms.

Section 2 presents traditional formulae for sample size estimation in cluster randomized trials. We also include a brief outline of the suggested method, a discussion of potential prior distributions for various parameters, as well as the approach used to simulate correlated binary data. Following this, we compare the traditional power-based and variance reduction methods in the context of a cluster randomized trial to assess the impact of insecticide treated nets (ITN) on the prevention of severe anaemia in pregnant women. Within the context of a clinic randomized trial, the non-inferiority perspective is applied to determine appropriate sample size requirements to show that a new antenatal care program is not inferior to the standard antenatal care program in an updated meta-analysis. The paper concludes with a brief discussion.

2. Methods

2.1. Traditional sample size estimation techniques in cluster randomized trials

The purpose of sample size estimation is to ensure that the proposed trial is adequately powered to detect a statistically significant and clinically meaningful treatment effect. To ensure consistency with subsequent sections, we suppose that the outcome of interest is binary in nature, and that the investigator would like to design a cluster randomized trial under a completely randomized design.

The primary complication of this design is its dependence on the ICC in both significance testing and sample size estimation. The ICC has the natural interpretation as the Pearson coefficient of correlation between any two responses within the same cluster (Donner and Klar, 2000, p. 8).

An estimate of ρ can be obtained using the standard one-way analysis of variance (ANOVA) among and within clusters. Within this framework, we consider a random sample of k clusters, each of size m. Using MSC and MSW to denote the mean square error among and within clusters, the analysis of variance estimator of ρ is given by

$$\hat{\rho} = \frac{\text{MSC} - \text{MSW}}{\text{MSC} + (m-1)\text{MSW}} = S_A^2/(S_A^2 + S_W^2)$$

where $S_A^2 = (\text{MSC} - \text{MSW})/m$ and $S_W^2 = \text{MSW}$ are sample estimates of σ_A^2 and σ_W^2 respectively (Donner, 1986).
In the case of variable cluster sizes m_j $(j = 1, \ldots, k)$, m may be replaced by

$$m_0 = \left(\frac{1}{k-1}\right)\left(M - \sum_{j=1}^{k} m_j^2/M\right)$$

where M denotes the total number of individuals in the sample (Donner and Klar, 2000, pp. 8–9).

The cluster size m also plays a critical role in the planning and analysis of a cluster randomized trial. In some cases, such as in the randomization of worksites, the values of m_j may be known in advance and can be used in the sample size estimation formula (Eldridge et al., 2006). However, in many cluster randomized trials, an estimate of the mean cluster size is used to replace m even though it leads to a slight underestimation of the required sample size (Donner and Klar, 2000, p. 57). Together, these two parameters give rise to the variance inflation factor (VIF) or design effect:

$$\text{VIF} = 1 + (m-1)\rho. \tag{1}$$

This quantity measures the impact of clustering on the variance of either a sample mean or a proportion as compared to an individually randomized trial.

Although other measures of binary association are possible, we will focus here on the odds ratio (OR) as the effect measure of interest. In this case, the aim of the investigator is to test the hypothesis $H_0 : \text{OR} = 1$ at the two-sided 100α per cent level of significance with power $(1-\beta)$. The parameter, $\text{OR} = \frac{P_1}{(1-P_1)}/\frac{P_2}{(1-P_2)} = \frac{P_1(1-P_2)}{P_2(1-P_1)}$ is defined in terms of P_1 and P_2, the population success rates in the experimental and control groups respectively. Sample estimates of P_1 and P_2 are given by $\hat{P}_1$ and $\hat{P}_2$ as computed over all individuals in all clusters, producing the estimate $\hat{\text{OR}} = \frac{\hat{P}_1(1-\hat{P}_2)}{\hat{P}_2(1-\hat{P}_1)}$.

To further simplify matters, we consider the natural logarithm of the odds ratio as it improves the normality of this effect measure while providing a simple computational expression for its variance (Wang et al., 2002). Specifically, we consider the equivalent hypothesis $H_0 : \log \text{OR} = 0$.