Published in final edited form as:

Commun Stat Theory Methods. 2020 ; 49(1): 116–124. doi:10.1080/03610926.2018.1532004.

Sample Size Calculation for Count Outcomes in Cluster Randomization Trials with Varying Cluster Sizes

Jijia Wang,

Department of Statistical Science, Southern Methodist University, Dallas, TX

Song Zhang,

Department of Clinical Sciences, UT Southwestern Medical Center, Dallas, TX

Chul Ahn

Department of Clinical Sciences, UT Southwestern Medical Center, Dallas, TX

Abstract

In many cluster randomization studies, cluster sizes are not fixed and may be highly variable. For those studies, sample size estimation assuming a constant cluster size may lead to under-powered studies. Sample size formulas have been developed to incorporate the variability in cluster size for clinical trials with continuous and binary outcomes. Count outcomes frequently occur in cluster randomized studies. In this paper, we derive a closed-form sample size formula for count outcomes accounting for the variability in cluster size. We compare the performance of the proposed method with the average cluster size method through simulation. The simulation study shows that the proposed method has a better performance with empirical powers and type I errors closer to the nominal levels.

Keywords

sample size; count outcome; cluster randomized trial

1 Introduction

Cluster randomization trials are widely used to evaluate the effect of interventions on health outcomes, where clusters of subjects, instead of individual subjects, are randomized to different intervention groups (Eldridge and Kerry, 2012; Campbell and Walters, 2014; Hayes and Moulton, 2017). Donner and Klar (2000) discussed many advantages of cluster randomization studies, which include administrative convenience, ease of obtaining cooperation from investigators, ethical consideration, enhancement of subject compliance, avoidance of treatment group contamination, and intervention naturally applied at the cluster level. A critical feature of cluster randomization trials is the dependency among subjects within the same cluster, quantified by the intracluster correlation coefficient (ICC), denoted as ρ . Dependent observations usually correspond to a smaller effective sample size since

individuals within a cluster are usually positively correlated. Suppose the research question is to investigate if there is a significant difference in mean effect between two intervention groups. Compared to an individual randomization trial, the treatment effect estimated in a cluster randomization trial of equal sample size has a larger variance, by an inflation factor of $1 + (m - 1)\rho$ due to the correlation among subjects within the same cluster (Eldridge et al., 2006). Here m is the fixed cluster size. This variance inflation factor is also called the design effect. The sample size under cluster randomization can be obtained by multiplying the design effect to the sample size calculated assuming individual randomization to maintain the required level of power.

Count outcomes frequently occur in cluster randomized studies. For example, Roig et al. (2010) evaluated the effectiveness of an intervention on reducing the number of cigarettes per day with a cluster randomization trial. Fairall et al. (2005) evaluated an educational outreach program for the case management of respiratory diseases with the outcome being the number of clinic visits using a cluster randomization trial. Sample size formulas have been developed for cluster randomization trials with count outcomes. For example, Amatya et al. (2013) derived the sample size formula for the count outcome in a cluster randomization trial. Hayes and Bennett (1999) provided the sample size estimate based on the coefficient of variation. One limitation of aforementioned methods is that they assume an equal cluster size in the design of cluster randomization trials. This assumption might be unrealistic because in practice clusters (formed by physician, clinic, neighborhood, etc) usually occur naturally with random variation in size due to time, logistics and budget constraint, or differences in recruitment rate, etc. Variation in cluster sizes can affect sample size estimation or the power of the study (Ahn et al., 2014; Guittet et al., 2006). A common practice has been to use the average cluster size to estimate the sample size. The resulting experimental design ignores a source of random variability (in cluster size), which might lead to under-powered trials. To address this problem, Manatunga et al. (2001) proposed a sample size formula accounting for varying cluster sizes for continuous outcomes, which adjusts traditional sample size formula by a variance inflation factor. Design parameters involved in the variance inflation factor include ICC as well as the mean and variance of cluster sizes. Along the same line, Kang et al. (2003) developed a sample size formula for dichotomous outcomes that accounts for varying cluster sizes. They compared their proposed method with the average cluster size method and the maximum cluster size method, where the cluster size is replaced by the average cluster size and the maximum cluster size, respectively. van Breukelen and Candel (2012a) proposed a correction of sample size formula to compensate for the loss of efficiency and power due to varying cluster sizes. Their approach is applicable to continuous and binary outcomes. More details can be found in van Breukelen et al. (2007, 2008); van Breukelen and Candel (2012b). To our knowledge, there has been no investigation on sample size adjustment for varying cluster sizes in the context of count outcomes.

In this study we investigate sample size calculation for cluster randomization trials under randomly varying cluster sizes, in the context of a count outcome. In Section 2, the derivation of sample size formulas under varying cluster sizes is presented. In Section 3, an empirical simulation study is conducted to investigate the performance of the proposed

method compared with the average cluster size method. In Section 4, an application of the proposed method is presented. Finally, we conclude with a discussion.

2 Methods

We use $k = 1/2$ to indicate the treatment/control group in a cluster randomization trial. The number of clusters in each group is denoted by J_k ($k = 1, 2$). The cluster sizes, denoted by n_{kj} ($j = 1, \dots, J_k$), vary across clusters following a certain discrete distribution with mean θ and variance τ^2 . The coefficient of variation (CV) is $\gamma = \tau/\theta$. We use Y_{kji} ($i = 1, \dots, n_{kj}$) to denote the count outcome observed from the i th subject in the j th cluster of the k th group. We assume a Poisson model for Y_{kji} with parameter $\lambda_k = E(Y_{kji}) = \text{Var}(Y_{kji})$. Y_{kji} 's are assumed to be independent across clusters. We also assume a common intracluster correlation coefficient, $\rho = \text{Corr}(Y_{kji}, Y_{kji'})$ for $i \neq i'$ (Rutterford et al., 2015; Zou and Donner, 2004).

Suppose the hypotheses of interest are $H_0: \lambda_1 = \lambda_2$ versus $H_1: \lambda_1 \neq \lambda_2$. For simplicity, we assume balanced randomization, i.e., $J_1 = J_2 = J$. An unbiased estimator of λ_k can be expressed as

$$\hat{\lambda}_k = \frac{\sum_{j=1}^J \sum_{i=1}^{n_{kj}} Y_{kji}}{\sum_{j=1}^J n_{kj}}. \quad (1)$$

Conditional on n_{kj} , $\hat{\lambda}_k$ approximately follows a normal distribution with mean λ_k and variance

$$\text{Var}(\hat{\lambda}_k | n_{kj}) = \frac{\lambda_k \sum_{j=1}^J n_{kj} [1 + (n_{kj} - 1)\rho]}{(\sum_{j=1}^J n_{kj})^2}. \quad (2)$$

The variance of $\hat{\lambda}_k$ can be estimated by

$$s_k^2 = \frac{\hat{\lambda}_k \sum_{j=1}^J n_{kj} [1 + (n_{kj} - 1)\hat{\rho}]}{(\sum_{j=1}^J n_{kj})^2}. \quad (3)$$

At the design stage, ρ can be estimated from the pilot data. A summary of literatures that presents the estimates of ρ can be found in Moerbeek and Teerenstra (2016). However, the exact n_{kj} is usually unknown, which prevents the assessment of $\text{var}(\hat{\lambda}_k)$ and subsequently the calculation of power. In this study, we assume that investigators have good insight about the first two moments (mean and variance) for the distribution of cluster sizes although investigators do not know the exact cluster sizes in advance. This assumption allows us to approximately marginalize with respect to n_{kj} and obtain $\text{var}(\hat{\lambda}_k)$. Specifically, we assume n_{kj} to be independently and identically distributed with mean $\theta = E(n_{kj})$ and variance $\tau^2 = \text{Var}(n_{kj})$. Then, we have

$$\text{Var}(\hat{\lambda}_k) = E\left[\text{Var}(\hat{\lambda}_k | n_{kj})\right] \approx \frac{\lambda_k J \left[(1-\rho)E(n_{kj}) + \rho E(n_{kj}^2) \right]}{(J\theta)^2} = \frac{\lambda_k}{J} \left[(1-\rho)\frac{1}{\theta} + \rho + \rho\gamma^2 \right].$$

The above derivation uses the fact that $E(n_{kj}^2) = \theta^2 + \tau^2$, $\sum_{j=1}^J n_{kj}/J \approx \theta$ for large J , and $\gamma = \tau/\theta$ is the coefficient of variation of cluster sizes.

With a power of $1 - \beta$ at a two-sided significance level of α , the required number of clusters in each intervention arm is

$$J = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 (\lambda_1 + \lambda_2)}{(\lambda_1 - \lambda_2)^2} \left[(1-\rho)\frac{1}{\theta} + \rho + \rho\gamma^2 \right]. \quad (4)$$

This sample size formula can be easily extended to accommodate unbalanced randomization. Suppose the treatment:control randomization ratio is r : 1, the number of clusters required for the control arm is

$$J = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 (\lambda_1/r + \lambda_2)}{(\lambda_1 - \lambda_2)^2} \left[(1-\rho)\frac{1}{\theta} + \rho + \rho\gamma^2 \right]. \quad (5)$$

It is noteworthy that when $\gamma = 0$, i.e., no variability in cluster size so all $n_{kj} = \theta$, the sample size formula becomes

$$J^* = \frac{(z_{1-\alpha/2} + z_{1-\beta})^2 (\lambda_1/r + \lambda_2)}{(\lambda_1 - \lambda_2)^2} \left[(1-\rho)\frac{1}{\theta} + \rho \right].$$

Comparing J and J^* , the relative change in sample size due to random variability in cluster size can be quantified by

$$R = \frac{J}{J^*} - 1 = \frac{\rho\theta}{1-\rho+\rho\theta}\gamma^2.$$

We have several observations about R :

- R does not depend on design parameters such as type I error α , power $1 - \beta$, and λ_1 and λ_2 .
- R is proportional to γ^2 , the squared coefficient of variation in cluster sizes.
- Under a positive ICC, $\rho > 0$, which is usually the case in practice, we have $R > 0$. That is, the variability in cluster size always leads to a larger sample size requirement. Ignoring this source of variability at experimental design will lead to underpowered trials.
- $R \rightarrow 0$ as $\rho \rightarrow 0$. Actually from an equivalent expression,

$$R = \frac{\theta\gamma^2}{\frac{1-\rho}{\rho} + \theta},$$

it is obvious that R is monotonically increasing with ρ for $\rho > 0$. Hence given the same coefficient of variation, the required increase in sample size is greater for trials with stronger dependence among observations from the same clusters.

- From another equivalent expression,

$$R = \frac{\rho\gamma^2}{\frac{1-\rho}{\theta} + \rho},$$

we observe that R is monotonically increasing with θ , the average cluster size. Given the same variability (γ) and ICC, a larger average cluster size is associated with a greater increase in sample size requirement.

In short, the worst scenario to ignore random variability in cluster size is for trials with high ICC and larger cluster sizes.

The first two plots in Figure 1 present the effects of ICC (ρ) and mean cluster size (θ) on relative change in sample size (R) with $\gamma = 1$. The relative change in sample size monotonically increases as ρ and θ increases, respectively. Figure 1 also shows the joint effect of ρ and θ on relative change in sample size. The relative change in sample size increases faster with ρ when θ is large. When $\theta = 1$, the relative change in sample size is equal to ρ .

3 Simulation

We conducted simulation studies to compare the performance of the proposed method and the average cluster size method. In a cluster randomization trial with a count outcome, we assume that the cluster sizes follow discrete uniform distributions DU[5, 15] and DU[25, 85] with mean 10 and 55, variance 10 and 510, and coefficient of variation 0.32 and 0.41, respectively. We explore a wide range of ICC, $\rho = 0.05, 0.15, 0.25, 0.35, 0.45, 0.55$. ICC is usually less than 0.10 in cluster randomization trials in a community setting (Moerbeek and Teerenstra, 2016). However, ICC larger than 0.10 occasionally occurs in a community intervention trials. For example, in a randomization trial that assessed an educational intervention aimed at improving the management of lung disease in adults attending South African primary care clinics, the ICC takes the value of 0.32 (Amatya et al., 2013). ICC is generally larger than 0.10 in dental studies. For example, in three dental studies in which the unit of sampling was subjects and the unit of analysis was tooth (Banting et al., 1985; Donner and Banting, 1988, 1989), the ICC ranged between 0.354 and 0.432. Thus, we chose the value of ICC that ranges between 0.05 and 0.55 in our simulation. We also consider two combinations of (λ_1, λ_2) : (1, 1.5) and (2.5, 2). We specify two-sided type I error $\alpha = 0.05$ and power $1 - \beta = 0.9$. Under each combination of design parameters, the required numbers

of clusters per group are computed for the two methods. The empirical power and empirical type I error are evaluated according to the following algorithm:

1. For each intervention group, generate a set of cluster sizes n_{kj} for $j = 1, \dots, J$, $k = 1, 2$ from the assumed DU distribution.
2. For the (k, j) th cluster, generate a vector of correlated count observations $\mathbf{Y}_{kj} = (Y_{kj1}, \dots, Y_{kjn_{kj}})'$, with mean vector $\mathbf{1}\lambda_k$ and correlation matrix $(1-\rho)\mathbf{I} + \rho\mathbf{1}\mathbf{1}'$. Here $\mathbf{I}$ is the identity matrix and $\mathbf{1}$ is a vector with all elements being 1. One simple approach is to generate $Y_{kji} = Z_{kji} + Z_{kj}^*$, where Z_{kji} are random samples from a Poisson distribution with event rate $\lambda_k(1-\rho)$ and Z_{kj}^* is generated from a Poisson distribution with event rate $\lambda_k\rho$. It is easy to show that under this scheme $E(Y_{kji}) = \text{Var}(Y_{kji}) = \lambda_k$ and $\text{Cov}(Y_{kji}, Y_{kji'}) = \lambda_k\rho$ (Mardia, 1970).
3. For $k = 1, 2$, obtain estimates $\hat{\lambda}_k$ and s_k^2 , respectively. Reject H_0 if $|\hat{\lambda}_1 - \hat{\lambda}_2|/\sqrt{s_1^2 + s_2^2} > z_{1-\alpha/2}$.

Empirical power and type I error are computed as the proportion of times H_0 is rejected under the alternative and null hypothesis, respectively. All the analyses were performed with R version 3.3.3 (<http://www.r-project.org>). In Table 1 we present the estimated number of clusters (empirical power, empirical type I error) obtained over 10,000 iterations. Within each cell, Rows 1 and 2 corresponds to the results under the proposed method and the average cluster size method, respectively. Table 1 shows that by appropriately accounting for random variability in cluster size, the proposed method generally maintains empirical power and type I error at the nominal levels. On the other hand, the average cluster size method consistently leads to underpowered studies. Table 1 also confirms our theoretical observations that the underestimation of sample size is most severe under high ICC and larger cluster sizes. For example, under $\rho = 0.55$ and DU[25, 85] where $\theta = 55$, the average cluster size approach underestimates the number of clusters by 10%. On the other hand, under $\rho = 0.05$ and DU[5, 15] where $\theta = 10$, the estimated numbers of clusters are effectively the same after rounding to integer.

4 Example

To illustrate the application of the proposed sample size estimation method in cluster randomization studies with count data, we consider the cluster randomization trial example provided in Amatya et al. (2013). The investigators were interested in assessing an educational intervention aimed at improving the management of lung disease in adults attending South African primary-care clinics. 40 clinics were included as clusters and randomized into treatment or control group. The goal of enrollment at each clinic was 50 patients and the outcome of interest was the number of clinic visits during the 4-month study period. The analysis reported that $\hat{\lambda}_1 = 4.35$, $\hat{\lambda}_2 = 3.63$, and $\hat{\rho} = 0.32$. Suppose we want to design a new cluster randomization trial based on the estimated design parameters to achieve 90% power at a two-sided significance level of 5%. If there is no variability in cluster size

(every clinic enrolls exactly 50 patients), the required number of cluster is 54. If the variability in cluster sizes is relatively small, say following $DU[40, 60]$, the required number of clusters is 55. Under a larger variability, say $DU[25, 75]$, required number of clusters is 59. Finally, if more patients are expected to be enrolled at each clinic where the cluster sizes follows a distribution of $DU[70, 130]$, the required number of clusters is 55.

5 Discussion

In this paper, we present a realistic sample size estimation formula for count data in cluster randomization trials where clusters sizes are randomly varying. Under a positive ICC, which is usually the case in practice, we demonstrate that ignoring this source of variability would lead to underpowered clinical trials. A higher ICC, a larger averaged cluster size, or a larger coefficient of variation in cluster sizes, is associated with larger percentage increase in sample size requirement. We also show that the traditional sample size formula obtained under constant cluster size assumption is a special case of the proposed method with the coefficient of variation equal to zero.

Compared with traditional sample size methods for cluster randomization trials, the proposed approach requires additional information on the distribution of cluster sizes, specifically, the mean (θ) and variance (τ^2) (equivalently, the coefficient of variation γ^2). In practice, these parameters can be obtained from previous studies and other sources, such as pilot studies, literature, or educated guess by clinicians.

Simulation studies show that the proposed method can maintain empirical power and type I error at their nominal levels over a wide range of design configurations. However, we observe that under small sample sizes (such as when the number of clusters is less than 15), the empirical powers are slightly lower than the nominal level. This is a problem associated with poor approximation by large sample theory, which affects most sample size methods. Since the number of clusters is usually small in community intervention trials, the use of critical values $z_{1-\alpha/2}$ and $z_{1-\beta}$ in Equations (4) and (5) instead of critical values $t_{1-\alpha/2}$ and $t_{1-\beta}$ corresponding to the t-distribution underestimates the required sample size. To adjust for underestimation, Snedecor and Cochran (1989) suggested adding one cluster per group when the sample size is determined with a 5% level of significance, and two clusters per group with a 1% level of significance.

We also consider the sample size estimation method proposed by van Breukelen and Candel (2012a). To compensate for the loss of power due to varying cluster sizes, they adjusted the average cluster size method and derived the sample size formula for the required number of clusters per group (J_{BC}) as

$$J_{BC} = \frac{J^*}{1 - \nu(1 - \nu)\gamma^2}, \text{ where } \nu = \frac{\theta\rho}{\theta\rho + 1 - \rho}.$$

We conducted simulation studies to investigate the performance of van Breukelen and Candel's method using the same parameter settings and algorithm as those in Section 3. Table 2 shows the required number of clusters per group based on van Breukelen and

Candel's method. The estimated numbers of clusters are slightly larger than or equal to those estimated by average cluster size method after rounding to integer, but are still generally too small to maintain the nominal power and type I error levels. It should be pointed out that van Breukelen and Candel's method was originally developed in the context of a continuous outcome. When applying to trial design with a binary outcome, the intraclass correlation ρ has to be specially defined for a latent continuous variable on the logistic scale. For a count outcome, the appropriate definition of ρ remains unclear. In the above application the intraclass correlation ρ was defined on the scale of the count outcome, which is very likely a misuse of van Breukelen and Candel's method to the context of a count outcome. We caution this misuse and further investigation into this important research question is warranted.

Acknowledgements

The work was supported in part by NIH grants 1UL1TR001105, AHRQ grant R24HS22418, and CPRIT grants RP110562-C1 and RP120670-C1.

References

- Ahn C, Heo M and Zhang S (2014), *Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research*, CRC Press.
- Amatya A, Bhaumik D and Gibbons RD (2013), 'Sample size determination for clustered count data', *Statistics in Medicine* 32(24), 4162–4179. [PubMed: 23589228]
- Banting D, Ellen R and Fillery E (1985), 'Clinical science a longitudinal study of root caries: baseline and incidence data', *Journal of Dental Research* 64(9), 1141–1144.
- Campbell MJ and Walters SJ (2014), *How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research*, John Wiley & Sons.
- Donner A and Banting D (1988), 'Analysis of site-specific data in dental studies', *Journal of Dental Research* 67(11), 1392–1395. [PubMed: 3183156]
- Donner A and Banting D (1989), 'Adjustment of frequently used chi-square procedures for the effect of site-to-site dependencies in the analysis of dental data', *Journal of Dental Research* 68(9), 1350–1354. [PubMed: 2778178]
- Donner A and Klar N (2000), *Design and Analysis of Cluster Randomization Trials in Health Research*, London: Arnold.
- Eldridge S and Kerry S (2012), *A Practical Guide to Cluster Randomised Trials in Health Services Research*, Vol. 120, John Wiley & Sons.
- Eldridge SM, Ashby D and Kerry S (2006), 'Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method', *International Journal of Epidemiology* 35(5), 1292–1300. [PubMed: 16943232]
- Fairall LR, Zwarenstein M, Bateman ED, Bachmann M, Lombard C, Majara BP, Joubert G, English RG, Bheekie A, van Rensburg D et al. (2005), 'Effect of educational outreach to nurses on tuberculosis case detection and primary care of respiratory illness: pragmatic cluster randomised controlled trial', *BMJ* 331(7519), 750–754. [PubMed: 16195293]
- Guittet L, Ravaud P and Giraudeau B (2006), 'Planning a cluster randomized trial with unequal cluster sizes: practical issues involving continuous outcomes', *BMC Medical Research Methodology* 6(1), 1. [PubMed: 16412232]
- Hayes R and Bennett S (1999), 'Simple sample size calculation for cluster-randomized trials.', *International Journal of Epidemiology* 28(2), 319–326. [PubMed: 10342698]
- Hayes RJ and Moulton LH (2017), *Cluster Randomised Trials*, CRC press.
- Kang S-H, Ahn CW and Jung S-H (2003), 'Sample size calculation for dichotomous outcomes in cluster randomization trials with varying cluster size', *Drug Information Journal* 37(1), 109–114.

- Manatunga AK, Hudgens MG and Chen S (2001), 'Sample size estimation in cluster randomized studies with varying cluster size', *Biometrical Journal* 43(1), 75–86.
- Mardia KV (1970), *Families of Bivariate Distributions*, Vol. 27, Griffin London.
- Moerbeek M and Teerenstra S (2016), *Power Analysis of Trials with Multilevel Data*, CRC Press.
- Roig L, Perez S, Prieto G, Martin C, Advani M, Armengol A, Roura P, Manresa JM, Briones E, Group IS et al. (2010), 'Cluster randomized trial in smoking cessation with intensive advice in diabetic patients in primary care. itadi study', *BMC Public Health* 10(1), 58. [PubMed: 20132540]
- Rutterford C, Copas A and Eldridge S (2015), 'Methods for sample size determination in cluster randomized trials', *International Journal of Epidemiology* 44(3), 1057–1067.
- Snedecor G and Cochran W (1989), *Statistical Methods*, Eighth Edition, Iowa state University Press Ames, Iowa, USA.
- van Breukelen GJ and Candel MJ (2012a), 'Comment on efficiency loss because of varying cluster size in cluster randomized trials is smaller than literature suggests', *Statistics in Medicine* 31(4), 397–400. [PubMed: 22253143]
- van Breukelen GJ and Candel MJ (2012b), 'Efficiency loss due to varying cluster sizes in cluster randomized trials and how to compensate for it: comment on You et al.(2011)', *Clinical Trials* 9(1), 125–125. [PubMed: 22334469]
- van Breukelen GJ, Candel MJ and Berger MP (2007), 'Relative efficiency of unequal versus equal cluster sizes in cluster randomized and multicentre trials', *Statistics in Medicine* 26(13), 2589–2603. [PubMed: 17094074]
- van Breukelen GJ, Candel MJ and Berger MP (2008), 'Relative efficiency of unequal cluster sizes for variance component estimation in cluster randomized and multicentre trials', *Statistical Methods in Medical Research* 17(4), 439–458. [PubMed: 17698940]
- Zou G and Donner A (2004), 'Confidence interval estimation of the intraclass correlation coefficient for binary outcome data', *Biometrics* 60(3), 807–811. [PubMed: 15339305]

Figure 1:

The effect of ICC (ρ) and mean cluster size (θ) on relative change in sample size (R)

Table 1:

Required number of clusters per group (empirical power, empirical type I error) for the proposed method (p) and the average cluster size method (a)

$\rho/(\lambda_1, \lambda_2)$	Cluster Size Distribution			
	DU[5, 15]		DU[25, 85]	
	(1, 1.5)	(2.5, 2)	(1, 1.5)	(2.5, 2)
0.05	16 (90.3%, 0.055) ^p	29 (90.4%, 0.055)	8 (88.2%, 0.049)	14 (89.2%, 0.054)
	16 (90.5%, 0.052) ^a	28 (89.6%, 0.052)	8 (88.0%, 0.051)	13 (87.5%, 0.053)
0.15	27 (90.9%, 0.053)	48 (90.3%, 0.052)	20 (91.5%, 0.052)	35 (90.5%, 0.052)
	25 (88.7%, 0.053)	45 (88.7%, 0.053)	18 (88.2%, 0.054)	32 (87.9%, 0.053)
0.25	37 (90.7%, 0.508)	67 (90.4%, 0.050)	31 (90.8%, 0.055)	55 (90.1%, 0.051)
	35 (88.9%, 0.055)	62 (87.9%, 0.048)	28 (88.1%, 0.054)	50 (87.4%, 0.049)
0.35	48 (91.1%, 0.048)	86 (90.1%, 0.051)	42 (90.8%, 0.047)	76 (90.4%, 0.053)
	44 (88.3%, 0.051)	79 (88.0%, 0.046)	39 (88.4%, 0.052)	69 (87.3%, 0.050)
0.45	58 (90.8%, 0.054)	105 (90.9%, 0.049)	54 (91.2%, 0.054)	96 (89.7%, 0.054)
	54 (89.1%, 0.051)	96 (88.1%, 0.051)	49 (87.9%, 0.051)	88 (88.0%, 0.053)
0.55	69 (90.6%, 0.052)	123 (90.1%, 0.051)	65 (90.3%, 0.050)	117 (90.6%, 0.050)
	63 (88.1%, 0.053)	113 (87.3%, 0.052)	59 (87.9%, 0.053)	106 (87.1%, 0.048)

Table 2:

Required number of clusters per group (empirical power, empirical type I error) for the van Breukelen and Candel's method

$\rho/(\lambda_1, \lambda_2)$	Cluster Size Distribution			
	DU[5, 15]		DU[25, 85]	
	(1, 1.5)	(2.5, 2)	(1, 1.5)	(2.5, 2)
0.05	16 (90.1%, 0.056)	29 (91.2%, 0.052)	8 (88.7%, 0.047)	13 (87.9%, 0.056)
0.15	26 (89.8%, 0.053)	46 (89.3%, 0.055)	18 (88.4%, 0.056)	32 (87.9%, 0.057)
0.25	35 (89.2%, 0.055)	63 (88.7%, 0.050)	28 (88.0%, 0.055)	51 (88.2%, 0.056)
0.35	45 (89.0%, 0.054)	80 (87.9%, 0.056)	39 (88.4%, 0.054)	69 (87.6%, 0.054)
0.45	54 (88.5%, 0.053)	97 (88.9%, 0.049)	49 (87.8%, 0.056)	88 (87.8%, 0.054)
0.55	63 (87.4%, 0.053)	114 (88.0%, 0.051)	59 (87.9%, 0.054)	106 (87.6%, 0.051)