Using Simulation When Only Minimal Information is Available to Estimate the Design Effect for an Ebola Vaccination Evaluation Study

Charles E. Rose, PhD* Paul Gargiullo, PhD† Benjamin Lopman, PhD‡
Manoj Gambhir, PhD§

Abstract
We developed a simulation model to investigate the impact of a proposed cluster stepped-wedge study design on the design effect (D_{EFF}) for the Ebola candidate vaccine evaluation study. The D_{EFF}, i.e., variance inflation factor, inflates the sample size necessary to detect a difference due to randomizing the vaccine to clusters of individuals rather than individuals. We explored several methods using four risk incidence scenarios for estimating the intra-class correlation (ICC) that is used to estimate the D_{EFF}. Our four scenarios assumed an incidence per month of: 1) 1.0%, 2) 2.0 and 1.0%, 3) 2.0, 1.0, and 0.5%, 4) 2.0, 1.0, 0.5, and 0.25%. In addition, we considered sample sizes of 3,600 and 4,500 that are allocated to the step-wedges and clusters within a step-wedge. Our results illustrate that the estimated design effect ranges from 1.02 to 1.14. Hence, we expect, given our currently available information, the cluster step-wedge design to have a minimal impact on the proposed study sample size.

Key Words: intra-class correlation coefficient (*ICC*), parametric, p-value, Fisher

1. Introduction

Randomized controlled trials are usually considered the "gold" standard for determining the efficacy of a vaccine. Randomization often occurs at the individual level but due to practicality it may not be feasible to randomized at the individual level and a cluster, i.e., group, randomized trial may be an attractive and viable alternative (Donner et al., Hayes et al.). When randomization occurs at the cluster level there is generally a loss of statistical efficiency, which is reflected as a loss of degrees of freedom and an inflated standard error, for the estimated vaccine efficacy (VE) (Eldridge et al.). Our purpose is to estimate the design effect with limited available information for a proposed study design to evaluate a candidate Ebola vaccine.

2. Design Effect

The loss of statistical efficiency is usually referred to as the design effect (D_{EFF}) or alternatively the variance inflation factor (VIF). Failure to account for the lack of independence of individuals within a cluster in the design stage may result in an underpowered study (Donner et. al). The D_{EFF} may impact the necessary sample size to detect a difference in clustered randomized trials (Hayes et al.). A feature of cluster randomized trials is that the outcomes of individuals within a cluster are correlated rather than independent. There have

*Quantitative Science and Data Management Branch, Division of HIV and AIDS Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, United States

†Statistical Support Unit, Epidemiology and Prevention Branch, Influenza Division, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, United States

‡Epidemiology Branch, Division of Viral Diseases, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, United States

§IHRC Incorporated, Atlanta, GA, United States

been several D_{EFF} estimators proposed (Eldridge et al.) and here we use the D_{EFF} given by:

$$D_{EFF} = 1 + \rho(m-1), \tag{1}$$

where ρ is the intraclass correlation coefficient (ICC) and m is the mean cluster size. To account for dependence during the design stage an estimate of the ICC and mean cluster size are required. The mean cluster size may be estimated using the best available information, which is usually easily obtained. Whereas, an estimate of the ICC may be more difficult to obtain.

3. Intraclass Correlation Coefficient (ICC)

Here, the ICC can be thought of as the amount of dependency among observations taken on individuals within a cluster, i.e., how related are observations within a cluster for the outcome (Kerry and Bland). The ICC is defined as:

$$ICC = \rho = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2}, \tag{2}$$

where σ_b^2 and σ_w^2 are the between and within cluster variances, respectively. The ICC is well defined for a linear model but for a generalized linear model, e.g., logistic regression, the ICC is more ambiguous. We consider several methods for estimating the ICC on the probability scale using logistic and log-binomial random-effects models.

Our first potential ICC estimator is based on the logistic distribution, where we assume the underlying distribution for the latent variable is logistic. The ICC is estimated using:

$$ICC = \rho = \frac{\sigma_b^2}{\sigma_b^2 + \dfrac{\pi^2}{3}}, \tag{3}$$

where $\pi^2/3$ is the variance of the logistic distribution and σ_b^2 is estimated using the clusters as the random effect in a logistic regression model. Alternatively, it has been suggested that it may be more desirable to estimate the ICC on the scale of interest, probability, rather than on the log-odds scale as for the logistic model, especially when we don't want to assume our outcome is a latent variable with an underlying logistic distribution (Eldridge et al.). Here, because we will have an outcome of Ebola disease (yes / no) with no underlying continuous distribution we will not consider estimator (3) further. Instead, we use the σ_b^2 obtained from the random-effects logistic model and calculate the ICC on the probability scale using the following approximate formulas for estimating σ_b^2 and σ_w^2 when using a Generalized Linear Mixed Model (GLMM) and logit link:

$$\sigma_b^2 \approx \sigma_a^2[\overline{p}(1-\overline{p})]^2, \tag{4}$$

where σ_a^2 is the estimated variance between the clusters on the logit scale and p is the estimated average Ebola probability estimated using the sample data. Our next two estimates for the ICC are obtained by estimating the σ_a^2 on the log scale rather than the $logit$ scale. Here we use a random-effects log-binomial model to estimate σ_a^2 and obtain an approximate estimate of the between cluster variance, σ_b^2, using two estimators:

$$\sigma_b^2 \approx \sigma_a^2\overline{p}^2 \tag{5}$$

$$\sigma_b^2 \approx [e^{\sigma_a^2} - 1]\overline{p}^2 \tag{6}$$

Equation (6) provides an alternative approach to equation (5) based on the log link that uses the properties of the lognormal distribution.

4. Ebola Vaccine Evaluation Design Effect Simulation

Our purpose is to design and conduct a simulation to estimate the ICC and D_{EFF} using our best available limited information. Our simulation steps are:

1. Outline the step-wedge study design
2. Estimate the expected number of clusters (vaccination groups), overall and within a step-wedge
3. Estimate the expected number and range of participants per cluster
4. Calculate the expected incidence for the unvaccinated group
5. Estimate the expected incidence for the vaccinated group based on the assumed VE
6. Generate simulation data based on steps 1-5
7. Fit generalized linear mixed models (GLMM) to the simulated data to estimate the ICC using equation (4), (5), and (6)
8. Calculate the D_{EFF} using the estimated ICC and summarize

Our considered study design is an 18 week step-wedge (Figure 1). Enrolled participants will be vaccinated over an 18 week period and randomized to a cluster within a vaccination week. The usable portion of the step-wedge for the efficacy estimation is within the "black" box. This design results in 14 usable weeks for the vaccinated and unvaccinated participants. Participants will be followed for 14 weeks for the usable portion of the efficacy analysis in both the unvaccinated and vaccinated groups. Hence, the total number of usable weeks per vaccination status group is 105 weeks. All participants in the unvaccinated and vaccinated groups were followed and used in the analysis. Technically those participants in the unvaccinated group who become infected with Ebola would not be followed into the vaccinated group. However, because only a small proportion of unvaccinated participants will become infected with Ebola and would not transition into the vaccinated study arm we keep the number of participants for the vaccinated and unvaccinated groups constant for the purpose of this simulation.

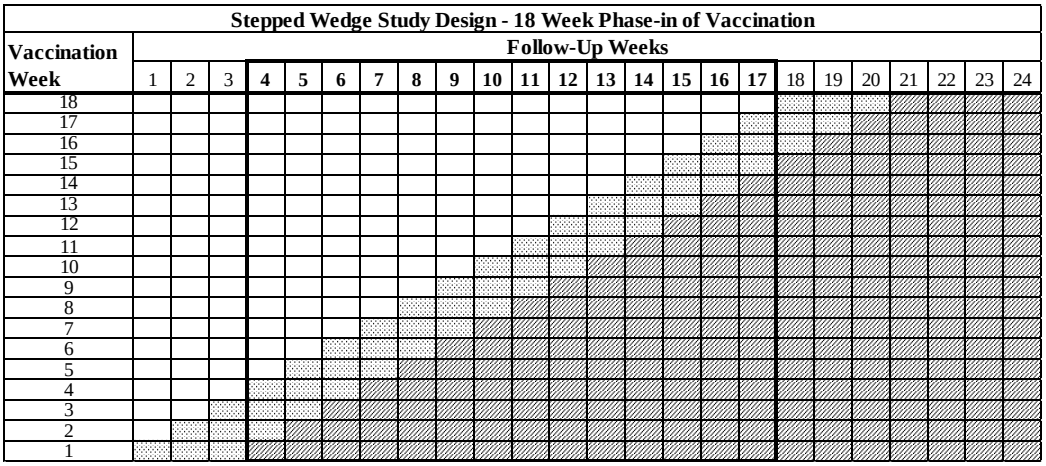

Figure 1: The 18 week step-wedge study design layout considered for an Ebola vaccine evaluation.

The generation of simulation data to estimate the D_{EFF} required several steps. We started our simulation by using currently available information to define the potential cluster types and expected number of clusters. Our currently available information for the type of cluster, number of each type of cluster, the staff per cluster type, total staff, and probability for a given type of cluster are presented in Table 1. There were 4822 staff in 246 clusters with the number of staff per cluster ranging from 2 - 250. We assumed the largest potential cluster for vaccination as a group is 100. The mean cluster size, i.e., staff per unit, assuming we place the tertiary hospitals in with the secondary hospitals is 17.8 staff per cluster.

The provided estimated incidence at the time of the simulation was defined as incidence per month, where a month is defined as 4 weeks. Since participants will be followed for different lengths of time we redefined incidence in terms of one week using a Geometric model. The expected incidence per weeks followed is defined as:

$$I_i = 1 - p^{(\frac{i}{4})}, \tag{7}$$

where i is the number of weeks a participant's step wedge unvaccinated group is followed and ranges from 1 to 14 weeks and p is the estimated per month background incidence. The incidence, p, may be defined by cluster type as we may assume the risk is not constant per group, i.e., there exists heterogeneity, frailty, among the clusters. At the time of the simulation there was minimal available information on incidence, which did not include estimates of heterogeneity.

4.1 Simulation

Our simulation began using the approximate sample size estimates for VE = 50% of 3,600 and 4,500 subjects for the 18 week step-wedge design. We allocated the sample size to the 18 step-wedges by dividing the sample size by 18, e.g., for the sample size of 3,600 we

Table 1: Provided information for the Sierra Leone Ebola health care worker cluster type, number of units, staff per unit, and probability (Prob) of the cluster type.

Cluster	Units	Staff / Unit	Total Staff	Prob
Community Health Clinic	30	30	900	0.122
Clinic	21	15	315	0.085
Community Health Post (CHP)	23	10	230	0.093
Maternal and Child Health Post (MCHP)	30	5	150	0.122
Secondary Hospital	10	100	1000	0.041
Burial Teams	20	12	240	0.081
District Surveillance Officer (DSO)	24	3	72	0.098
Ambulance Teams	22	2	44	0.089
Quarantine	1	25	25	0.004
Nutrition	2	18	36	0.008
Tertiary Hospital	3	250	750	0.012
Contact Tracers	46	10	460	0.187
Ebola Holding Center (EHC)	10	40	400	0.041
Ebola Treatment Unit (ETU)	4	50	200	0.016
Total	246		4822	1.0

divided by 18 giving us 200 subjects per step-wedge. We assume the number of subjects per step-wedge is constant across the study but we could allow this to vary by step-wedge. Hence, assuming 200 participants per wedge and 18 wedges results in a total sample size of 3600 participants. However, because the step-wedge with a three week lag-time results in only 14 wedges of usable time per group we have a total sample size of 2800 per vaccination status group for estimating efficacy. We computed the expected number of Ebola cases in the unvaccinated arm by wedge and summed over all wedges to get the overall expected number Ebola cases. We assumed a 1.0% incidence per month (provided as the estimated incidence at the time of the simulation) and estimated the incidence for each of the 14 wedges using:

$$I_i = 1 - 0.99^{(\frac{i}{4})}, \tag{8}$$

where i ranges from 1 to 14. Next we computed the number of person weeks per wedge, e.g., for vaccination week 18 we have 14 usable weeks of unvaccinated time and 200 participants so we have 2800 weeks of person time (PT). For each wedge we have an estimated incidence per week so we multiply the total person weeks of a wedge times by the incidence per week to estimated the expected number of Ebola cases for a step-wedge. Once we computed the expected Ebola cases for each step-wedge we summed over all step-wedges for the unvaccinated groups to obtain a total expected number of cases.

Alternatively, since our ICC is estimated over all step-wedges we could sum the 14 usable wedges, which results in a total number of weeks of 105, and multiply by the expected participants per week to estimate the total PT. For example, if we have 200 participants per week and 105 total weeks of PT for the 14 usable step-wedges then we estimate 200*105 = 21,000 weeks of follow-up time. Furthermore, since the incidence of 1.0% is based on 4 weeks we divided 21,500 by 4 for a total of 5,250 4 weeks periods of follow-up time. Since our incidence is expressed as a percent we divided the 5,250 by 100 to obtain an expected

number of Ebola cases in the unvaccinated group of 52.5 cases. For the vaccinated group, assuming VE = 50%, we estimate the number of Ebola cases in the vaccinated group as 0.5*52.5 = 26.25 cases.

Our simulation parameters are provided in Table 2. For generating simulation data the total number of clusters was randomly chosen using a Poisson distribution with expected values of 156 and 195. The expected number of clusters was determined by taking the sample size, e.g., 3,600, and dividing by the expected cluster size (17.8, see Table 1). Next, the number of subjects per cluster was randomly chosen using the table of probabilities based on the cluster type (Table 1). Next we randomly drew from a binomial distribution given the cluster size as n and the incidence probability. We assumed several scenarios to add heterogeneity to the estimated incidence for the unvaccinated group and adjusted the vaccinated group by 50%. There are four scenarios conducted under each sample size and ICC estimation method for a total of 12 estimated D_{EFF}. These scenarios are I (only one risk group, 1.0%), II (risk groups of 1.0 and 2.0%, III (risk groups of 2.0, 1.0, and 0.5%), and IV (risk groups of 2.0, 1.0, 0.5, and 0.25%). We assumed that for multiple risk groups that the risk groups are equally likely and we assigned the risk group by drawing a random variable from a Uniform [0, 1] distribution. Our expected number of clusters, subjects per cluster, probability of cluster type, and expected Ebola cases based on our assumptions are given in Tables 1 and 2.

Table 2: The model parameters for generating 1,000 datasets of simulation data.

Simulation Parameters				
Sample Size 3,600				
Parameter	**Mean Estimate**	**Distribution**	**Range**	**Total**
Useable Wedges	14	Constant	NA	NA
Subjects per Wedge	200	Constant	NA	NA
Clusters	156	Poisson	NA	NA
Subjects per Cluster	17.8	Table (see Table 2)	2, 100	NA
Probability of Infection				
Unvaccinated	0.01	Binomial	NA	52.50
Vaccinated	0.005	Binomial	NA	26.25
Sample Size 4,500				
Parameter	**Mean Estimate**	**Distribution**	**Range**	**Total**
Useable Wedges	14	Constant	NA	NA
Subjects per Wedge	250	Constant	NA	NA
Clusters	195	Poisson	NA	NA
Subjects per Cluster	17.8	Table (see Table 2)	2, 100	NA
Probability of Infection				
Unvaccinated	0.01	Binomial	NA	65.625
Vaccinated	0.005	Binomial	NA	32.8125

4.2 Simulation Results

Simulation results were compared against the "truth" for the expected number of clusters and table of cluster type probabilities. The expected number of subjects per cluster and cluster type probabilities are summarized and presented in Figure 2. Our 1,000 generated simulation datasets results in a mean and 95% percentile limits for the sample sizes 3,600 and 4,500 of 155.9 (131.5, 179) and 194.9 (167.5, 221), respectively, which results in the means almost identical to the "truth." Our simulation results using the equations (4), (5), and (6) are presented in table 3. There are no substantial difference in the estimated D_{EFF} using the three different approaches (equations (4), (5), and (6)). As expected as the heterogeneity increases from scenario I to IV the D_{EFF} increases. Scenario IV has the most heterogeneity and the D_{EFF} ranges from 1.08 - 1.14. There was little difference in the estimated D_{EFF} using the two sample sizes of 3,600 and 4,500. For our simulation assumptions there is no evidence of a substantial D_{EFF}.

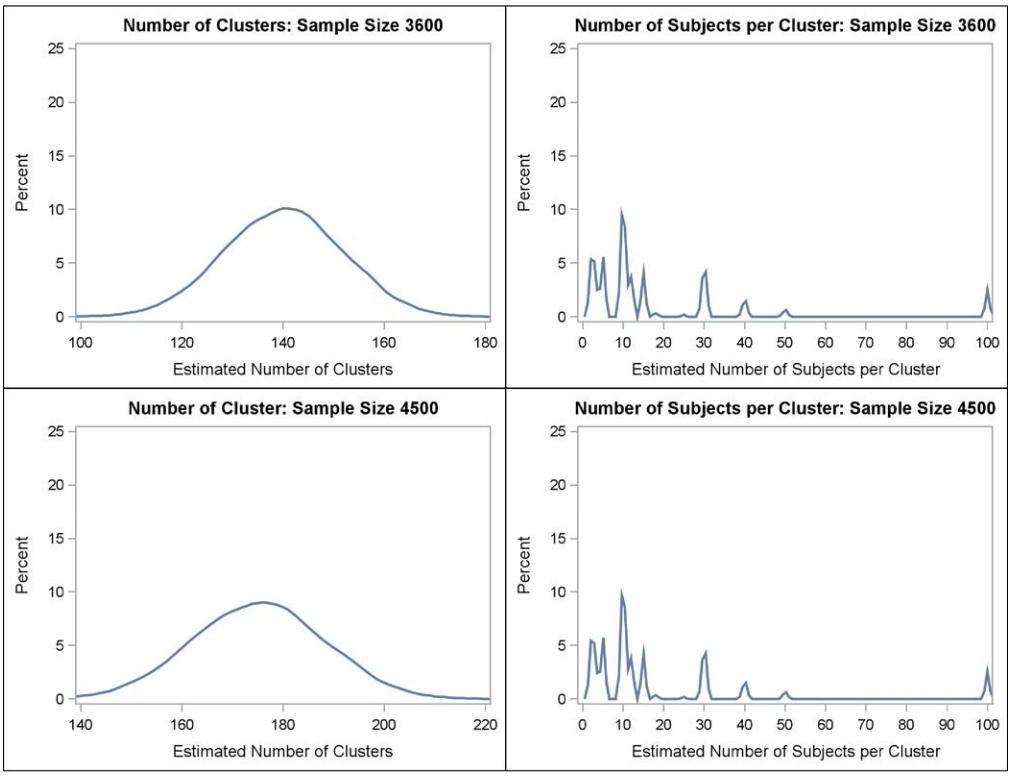

Figure 2: Results from the simulation for the estimated number of health care worker clusters and percent of cluster type.

5. Conclusion

Estimating the D_{EFF} for a cluster randomized trial during the study design phase is crucial for determining an adequate sample size for detecting the effect of interest. The D_{EFF} may be estimated from prior studies, pilot data, or the literature but for rapidly emerging diseases there may be little or no information available for estimating the D_{EFF}. A recently completed Ebola vaccine trial assumed an intra-class correlation coefficient of 0.05,

Table 3: Simulation results based on 1,000 simulations using Methods I, II, and III for estimating the *ICC* and 95% confidence interval (CI), and the D_{EFF} and 95% CI. We considered per-wedge sample sizes of 200 and 250 subjects for the 18 week step-wedge design, which resulted in a total sample size of 3,600 and 4,500 participants, respectively. We conducted four scenarios using each sample size. These scenarios were I (only one risk group, 1.0%), II (risk groups of 1.0 and 2.0%, III (risk groups of 2.0, 1.0, and 0.5%), and IV (risk groups of 2.0, 1.0, 0.5, and 0.25%).

Scenario	*ICC*	95% CI *ICC*	D_{EFF}	95% CI D_{EFF}
Method I: Equation (4)				
Sample Size 3,600				
I	0.0016	0.000061, 0.0050	1.03	1.00, 1.08
II	0.0026	0.000053, 0.0080	1.04	1.00, 1.13
III	0.0040	0.00012, 0.010	1.07	1.00, 1.17
IV	0.0045	0.00014, 0.011	1.07	1.00, 1.18
Sample Size 4,500				
I	0.0014	0.000063, 0.0047	1.02	1.00, 1.08
II	0.0024	0.000076, 0.0070	1.04	1.00, 1.11
III	0.0038	0.00022, 0.0096	1.06	1.00, 1.16
IV	0.0045	0.00044, 0.010	1.08	1.01, 1.17
Method II: Equation (5)				
Sample Size 3,600				
I	0.0016	0.000048, 0.0052	1.03	1.00, 1.08
II	0.0028	0.00013, 0.0082	1.05	1.00, 1.13
III	0.0042	0.00014, 0.011	1.07	1.00, 1.18
IV	0.0049	0.00027, 0.012	1.08	1.00, 1.20
Sample Size 4,500				
I	0.0014	0.000050, 0.0044	1.02	1.00, 1.07
II	0.0024	0.00013, 0.0070	1.04	1.00, 1.12
III	0.0038	0.00023, 0.009	1.06	1.00, 1.15
IV	0.0045	0.00030, 0.011	1.08	1.00, 1.18
Method III: Equation (6)				
Sample Size 3,600				
I	0.0018	0.000048, 0.0062	1.03	1.00, 1.11
II	0.0032	0.00012, 0.0023	1.05	1.00, 1.17
III	0.0059	0.00014, 0.018	1.10	1.00, 1.29
IV	0.0082	0.00029, 0.029	1.14	1.01, 1.47
Sample Size 4,500				
I	0.0016	0.000048, 0.0051	1.03	1.00, 1.09
II	0.0028	0.00012, 0.0084	1.05	1.00, 1.14
III	0.0050	0.00026, 0.015	1.08	1.00, 1.24
IV	0.0074	0.00028, 0.023	1.12	1.00, 1.39

based on unpublished data, whereas our simulation using our available information and considered scenarios estimated the largest *ICC* as approximately 0.008 using our scenarios (Henao-Restrepo et al., Ebola Consortium). Our largest D_{EFF} for our considered scenarios was approximately 1.14, which is substantially less than if we had assumed an *ICC* of 0.05. It may seem counterintuitive to estimate a small D_{EFF}. However, given our expectation of about 150 clusters per vaccination group with an estimated 25-50 Ebola cases per group we would expect most clusters to have zero cases of Ebola. Furthermore, those clusters that have an Ebola case are unlikely to have more than one or two Ebola cases. We demonstrated the feasibility of estimating the D_{EFF} using simulation given a minimal amount of information. In designing a study using simulation the estimate the D_{EFF} will be beneficial.

REFERENCES

Donner, A., Birkett, N., and Buck, C. (1981), "Randomization by cluster. Sample size requirements and analysis," *American Journal of Epidemiology*, 114, 906–914.

Eldridge, S. M., 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, 1292–1300.

Ebola a Suffit Ring Vaccination Trial Consortium. (2015), "The ring vaccination trial: a novel cluster randomized controlled trial design to evaluate vaccine efficacy and effectiveness during outbreaks, with special reference to Ebola," *BMJ*, 351, h3740.

Donner, A., and Klar, N. (1996), "Statistical considerations in the design and analysis of community intervention trials," *Journal of Clinical Epidemiology*, 49, 435–439.

Hayes, R. J., and Bennett, S. (1999), "Simple sample size calculation for cluster randomized trials," *International Journal of Epidemiology*, 28, 319–326.

Henao-Restrepo, A. M., Longini, I. M., Egger, M., Dean, N. E., Edmunds, W. J., Camacho, A., Carroll, M. W., Doumbia, M., Draguez, B., Duraffour, S., Enwere, G., Grais, R., Gunther, S., Hossmann, S., Kond, M. K., Kone, S., Kuisma, E., Levine, M. M., Mandal, S., Norheim, G., Riveros, X., Soumah, A., Trelle, S., Vicari, A. S., Watson, C. H., Kta, S., Kieny, M. P., Rttingen, J. A. (2015), "Efficacy and effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: interim results from the Guinea ring vaccination cluster-randomised trial," *Lancet*, 386(9999), 857-866.

Kerry, S.M., and Bland, J. M. (1998), "The intracluster correlation coefficient in cluster randomization," *BMJ*, 316, 1455.

Lake, S., Kammann, E., Klar, N., and Betensky, R. (2002), "Sample size re-estimation in cluster randomization trials," *Statistics in Medicine*, 21, 1337–1350.