# Impact of subject attrition on sample size determinations for longitudinal cluster randomized clinical trials

**Moonseong Heo**

Division of Biostatistics, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York, USA

## Abstract

Subject attrition is a ubiquitous problem in any type of clinical trials and thus needs to be taken into consideration at the design stage particularly to secure adequate statistical power. Here, we focus on longitudinal cluster randomized clinical trials (cluster-RCT) that aim to test the hypothesis that an intervention has an effect on the rate of change in the outcome over time. In this setting, the cluster-RCT assumes a three level hierarchical data structure in which subjects are nested within a higher level unit such as clinics and are evaluated for outcome repeatedly over the study period. Furthermore, the subject-specific slopes can be modeled in terms of fixed or random coefficients in a mixed-effects linear model. Closed form sample size formulas for testing the hypothesis above have been developed under assumption of no attrition. In this paper, we propose closed form approximate samples size determinations with anticipated attrition rates by modifying those existing sample size formulas. With extensive simulations, we examine performances of the modified formulas under three attrition mechanisms: attrition completely at random, attrition at random and attrition not at random. In conclusion, the proposed modification is very effective under fixed slope models but yields biased, if not substantially, statistical power under random slope models.

### Keywords

longitudinal cluster RCT; three level data; power; sample size; attrition; effect size

## 1. Introduction

Subject attrition during a clinical trial is a norm rather than an exception. For example, a review of attrition problems in geriatric psychiatry clinical trials reveals that average attrition rates over 68 studies is about 27.3% ranging from 3.1% to 54.1% (Heo et al., 2009). Such attrition problems may also apply to longitudinal cluster randomized trials (cluster-RCT), which we are considering in this paper. A cluster-RCT typically assumes a three level data structure in that interventions are randomly assigned to clusters such as clinics (level 3) which follow up subjects (level 2) for repeated assessments (level 1) during the study period.

Address correspondence to: Moonseong Heo, Ph.D., Division of Biostatistics, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Belfer 1312, Bronx, NY 10461, moonseong.heo@einstein.yu.edu.

At the design stage, as important as planning analytic strategies about handling attrition problems, sample size determinations under anticipated attrition rates should also be put into place in part because subjects' attritions compromise statistical power of trials. The most intuitive strategy would be to multiply by some factor a number of subjects determined under assumption of no attrition, i.e., under a hypothetically ideal situation. For example, if an anticipated attrition rate is $\xi$ ($0<\xi<1$), then the multiplication factor would be $1+\xi/(1-\xi)$. Although this strategy could be very effective in a typical parallel group design with only one level data structure, we suspect that it would result in an over-powered study of a cluster-RCT particularly because the number of observations is not taken into account for such sample size determinations.

In this paper, we consider a cluster-RCT in which the primary goal is to compare the longitudinal courses in a continuous outcome between two groups, e.g., control and experimental. For example, this hypothesis has been tested in a cluster randomized trial to evaluate the effect of an intervention for depression in the primary care setting on change in depression symptoms using the Hamilton rating scale for depression (Alexopoulos et al, 2005.; Dietrich et al., 2004). The individual longitudinal courses can be modeled as random or fixed slopes of the outcome over the time at the subject level for the purpose of the comparison. The difference in mean slopes over subjects between groups can then be assessed by including in a linear mixed effects model an interaction term between the treatment and time effects (Laird and Ware, 1982; Longford, 1993). Sample size determination formulas for testing this interaction are available under both a fixed slope model (e.g., Heo and Leon, 2009) and a random slope model (Murray et al., 2007; Roy et al., 2007).

The two parameters in existing formulas that would be affected by subject attrition are the number and the variance of the assessment time points per subject. Therefore, our sample size determination (with respect to the number of subjects) strategy is to replace those two parameters by their corresponding expected number and variance under anticipated attrition rates. While Murray et al. (2007) approaches considered broader and more general models for sample size determinations, they did not examine performances of their approach under anticipated attrition problems. Although Roy et al (2007) also considered general models and examined attrition effects on sample size determinations based on implicit approach using critical regions determined by $\chi^2$ distributions of feasible version generalized least square estimates, they appeared to consider only attrition completely at random mechanism. In contrast, we consider a specific model specified below and our sample size determination strategy mentioned above is simple and straightforward. Furthermore, our approach results in closed form power functions and sample size formulae and we hypothesize that the resulting multiplication factor would be smaller than $1+\xi/(1-\xi)$.

We examine the performance of our approach with extensive simulations considering the following factors among others: fixed and random slope models; different attrition rates; different distributions of attrition time points; and three different attrition mechanisms.

## 2. Statistical Model

A three level mixed-effects linear model for outcome *Y* with subject-specific random slopes can be expressed as follows (Hedeker and Gibbons, 2006):

$$Y_{ijk} = \left(\beta_0 + u_i + u_{j(i)}\right) + \zeta X_{ijk} + \left(\tau + \nu_{j(i)}\right) T_{ijk} + \delta X_{ijk} T_{ijk} + e_{ijk}, \quad (1)$$

where $i = 1, 2, \ldots, 2N_3$ is the index for the level three unit (e.g., clinic); $j = 1, \ldots, N_2$, is the index for the level two unit (e.g., subject) nested within each $i$; and $k = 1, 2, \ldots, N_1$, is the index for the level one unit (e.g., repeated outcomes) within each $j$. The intervention assignment indicator $X_{ijk} = 0$ and 1 if the $i$-th level three unit is assigned to a control intervention and an experimental intervention, respectively. Here we consider a design with $X_{ijk} = X_i$ for all $j$ and $k$ and also a balanced design so that $\Sigma_i X_i = N_3$. In addition, it is assumed that $T_{ijk} = T_k$ for all $i$ and $j$, and that the time from $T_1 = 0$ (the baseline) to $T_{end} = N_1 - 1$ (the last time point) increases by equal unit time intervals.

With respect to the random effects, it is assumed that: 1) $e_{ijk} \sim N\left(0, \sigma_e^2\right), u_{j(i)} \sim N\left(0, \sigma_2^2\right), u_i \sim N\left(0, \sigma_3^2\right)$ and $\nu_{j(i)} \sim N\left(0, \sigma_\tau^2\right)$; 2) these four random components are mutually independent i.e., $u_i \perp u_{j(i)} \perp e_{ijk} \perp \nu_{j(i)}$; and 3) $u_{j(i)}$, $\nu_{j(i)}$ and $e_{ijk}$ are conditionally independent whereas the $u_i$ are *unconditionally* independent—that is, both $u_{j(i)}$ and $\nu_{j(i)}$ are independent conditional on $u_i$, and the $e_{ijk}$ are independent conditional on $u_i$, $\nu_{j(i)}$ and $u_{j(i)}$. When $\sigma_\tau^2 = 0$, model (1) reduces to the fixed slope model.

For the fixed effects, the parameter $\zeta$ represents the intervention effect at baseline, and the parameter $\tau$ represents the slope associated with the time effect, that is, the magnitude of the change in outcome over time, in the control group. Finally, the intervention-by-time effect $\delta$, the parameter of primary interest, represents the difference in *mean* slopes of the outcome *Y* between the intervention groups. The overall intercept (fixed) is denoted by $\beta_0$.

Given that the parameter $\delta$ is of primary interest, the relevant null hypothesis can be expressed as:

$$H_0 : \delta = 0. \quad (2)$$

Under model (1), it can be shown that the elements of the mean vector for the outcome are equal to $E(Y_{ijk}) = \beta_0 + \xi X_i + \tau T_k + \delta X_i T_k$ and the elements of the covariance matrix are:

$$Cov\left(Y_{ijk}, Y_{i'j'k'}\right) = 1(i=i' \& j=j' \& k=k')\sigma_e^2 + 1(i=i' \& j=j')\left(T_k T_{k'}\sigma_\tau^2 + \sigma_2^2\right), + 1(i=i')\sigma_3^2$$

where 1(.) is an indicator function. It follows that:

$$Var\left(Y_{ijk}\right) = Cov\left(Y_{ijk}, Y_{ijk}\right) = \sigma^2 + T_k^2 \sigma_\tau^2$$

where $\sigma^2 \equiv \sigma_e^2 + \sigma_2^2 + \sigma_3^2$, the variance of $Y$ under the fixed slope model with $\sigma_\tau^2 = 0$. Therefore, the correlations among the level two data, i.e., among outcomes from different second level clusters (subjects) but the same third level cluster (clinic), can be expressed for $j \neq j'$ as follows:

$$Corr\left(Y_{ijk}, Y_{ij'k'}\right) = \frac{\sigma_3^2}{\sqrt{\sigma^2 + T_k^2 \sigma_\tau^2} \sqrt{\sigma^2 + T_{k'}^2 \sigma_\tau^2}}.$$

The correlations among the level one data, i.e., among outcomes measured at different time points on the same subject nested within clinics, can be expressed for $k \neq k'$ as:

$$Corr\left(Y_{ijk}, Y_{ijk'}\right) = \frac{\sigma_2^3 + \sigma_3^2 + T_k T_{k'} \sigma_\tau^2}{\sqrt{\sigma^2 + T_k^2 \sigma_\tau^2} \sqrt{\sigma^2 + T_{k'}^2 \sigma_\tau^2}}.$$

Under the fixed slope model, i.e., when $\sigma_\tau^2 = 0$, the correlations reduce to the following, respectively:

$$\rho_2 = \sigma_3^2 / \sigma^2 \quad (3)$$

and

$$\rho_1 = (\sigma_2^2 + \sigma_3^2) / \sigma^2. \quad (4)$$

## 3. Statistical Power and Sample Size with No Subject Attrition

It has been shown under assumption of no subject attrition that the power function to test the null hypothesis (2) based on the ordinary least squares estimate of $\delta$ can be written as follows (e.g., Murray et al., 2007; Heo et al., in press):

$$\varphi = \Phi \left\{ \Delta \sqrt{\frac{N_3 N_2 N_1 \, Var_p(T)}{2\left\{(1 - \rho_1) + r_\tau N_1 \, Var_p(T)\right\}}} - \Phi^{-1}(1 - \alpha/2) \right\}, \quad (5)$$

where $\alpha$ is a two-sided significance level; $\Phi$ is the cumulative distribution function (CDF) of a standard normal distribution and $\Phi^{-1}$ is its inverse;

$$\Delta = \delta / \sigma$$

is a standardized effect size;

$$r_\tau = \sigma_\tau^2 / \left(\sigma_2^2 + \sigma_3^2 + \sigma_e^2\right) = \sigma_\tau^2 / \sigma^2$$

is the ratio of the random slope variance to the sum of the other variances; and

$Var_p(T) = \sum_{k=1}^{N_1} \left( T_k - \overline{T} \right)^2 / N_1$ is the "population variance" of the time variable $T$ where

$\overline{T} = \sum_{k=1}^{N_1} T_k / N_1$ is the "mean" time point. We assume that $\delta = |\delta| > 0$ and also that the probability below a critical value, $\Phi^{-1}(\alpha/2)$, in the other side under the alternative hypothesis is negligible and thus considered as 0. When $\sigma_\tau^2 = 0$ or $r_\tau = 0$, the effect size is identical to the standardized effect size for the slope difference $\delta$ and power function (5) is the same as that derived under a fixed slope model (Heo and Leon, 2009).

It follows that the required sample size per group for the number of subjects $N_2$ per level three unit, for a desired statistical power $\varphi$ with a two-sided significance level $\alpha$, can be calculated from equation (5) as:

$$N_2 = \frac{2 \left\{ (1 - \rho_1) + r_\tau N_1 \, Var_p(T) \right\} \left\{ \Phi^{-1}(1 - \alpha/2) + \Phi(\varphi) \right\}^2}{N_3 N_1 \, Var_p(T) \Delta^2}. \quad (6)$$

More precisely, $N_2$ is the smallest integer greater than the right hand side of equation (6). In sum, the validity of this sample size formula has been supported by extensive simulations (Heo et al., 2012).

## 4. Statistical Power and Sample Size with Anticipated Attritions

A heuristic strategy for derivation of approximate sample sizes would be to replace both $N_1$ and $Var_p(T)$ in equation (6) with the corresponding expected values with an anticipated subject attrition rate. To this end, we assume no attrition at baseline, i.e., when $T = 0$, but consider only monotone attrition pattern in that subject outcome $Y$ is observed at every time point $T$ before attrition but no additional outcomes are observed after attrition. Therefore, the overall attrition rate, $\xi = P(A \quad T_{end})$, where $A$ is attrition time, is equivalent to one minus the proportion of participants who appeared at the last visit. Here, we further assume that the distribution of attrition time $A$ is uniform over or linearly increasing with $T = 0, 1, 2, \ldots, N_1 - 1 = T_{\text{end}}$. For the "uniform" distribution, probability of attrition at time $t$ can be defined as

$$w_t^{(u)} \equiv P\left( A^{(u)} = t | A^{(u)} \leq T_{encl} \right) = 1(t > 0) \frac{\xi}{(N_1 - 1)},$$

which yields $\sum_t w_t^{(u)} = \xi$. For the "linear" distribution in which attrition rates increase over time,

$$w_t^{(l)} \equiv P\left( A^{(l)} = t | A^{(l)} \leq T_{end} \right) = 1(t > 0) \frac{2\xi t}{N_1(N_1 - 1)},$$

which also yields $\sum_t w_t^{(l)} = \xi$. However, we do not make any distributional assumption about $A$ when $A > T_{\text{end}}$ except that $P(A > T_{end}) = 1 - \xi$. From now on, the superscripts "$(u)$" and "$(l)$" will represent "uniform" and "linear" distributions of attrition time, respectively.

The expected number of observations per subject can then be obtained as follow:

$$E^{(u)}(N_1) = \left(1 - \sum_{l=0}^{k-1} W_l^{(u)}\right) = N_1(1 - \xi/2)$$

and

$$E^{(l)}(N_1) = \sum_{k=1}^{N_1} \left(1 - \sum_{l=0}^{k-1} w_l^l\right) = N_1 - \xi(N_1+1)/3.$$

Either of these does not have to be an integer number. It can be seen that

$$E^{(U)}(N_1) \leq E^{(l)}(N_1). \quad (7)$$

This inequality implies that total number of observations will be larger under the linear distribution of attrition time.

On the other hand, the probability distribution of $T$, $P(T = t)$, at which the observations are made is no longer uniform under the assumed monotone attrition pattern regardless of the types of the distribution of the attrition time. Under the uniform and linear attrition time distributions, they can be obtained respectively as follows:

$$P^{(u)}(T=t) = \left(1 - \sum_{l=0}^{t} w_l^{(u)}\right) / \sum_{t=0}^{N_1-1} \left(1 - \sum_{l=0}^{t} w_l^{(u)}\right) = \frac{1 - t\xi/(N_1-1)}{N_1(1-\xi/2)}$$

and

$$P^{(l)}(T=t) = \left(1 - \sum_{l=0}^{t} w_l^{(l)}\right) / \sum_{t=0}^{N_1-1} \left(1 - \sum_{l=0}^{t} w_l^{(l)}\right) = \frac{1 - t(t+1)\xi/\{N_1(N_1-1)\}}{N_1 - \xi(N_1+1)/3}.$$

The first and the second moment of $T$ under these probability distributions can then be obtained as follows:

$$E^{(u)}(T) = \sum_{t=0}^{N_1-1} tP^{(u)}(T=t) = \frac{(N_1-1)(3-2\xi)}{6(1-\xi/2)}$$

$$E^{(u)}(T^2) = \sum_{t=0}^{N_1-1} t^2 P^{(u)}(T=t) = \frac{(N_1-1)\{N_1(4-3\xi)-2\}}{12(1-\xi/2)}$$

$$E^{(l)}(T) = \sum_{t=0}^{N_1-1} tP^{(l)}(T=t) = \frac{N_1(N_1-1)/2-\xi(3N_1-2)(N_1+1)/12}{N_1-\xi(N_1+1)/3}$$

and

$$E^{(l)}(T^2) = \sum_{t=0}^{N_1-1} t^2 P^{(l)}(T=t)$$
$$= \frac{N_1(N_1-1)(2N_1-1)/6-\xi N_1(N_1-1)/4-\xi(2N_1-1)(3N_1^2-3N_1-1)/30}{N_1-\xi(N_1+1)/3}.$$

Consequently, the variances can be obtained as:

$$Var_p^{(u)}(T) = E^{(u)}(T^2) - \left(E^{(u)}(T)\right)^2$$

and

$$Var_p^{(l)}(T) = E^{(l)}(T^2) - \left(E^{(l)}(T)\right)^2.$$

Therefore, the approximate statistical power with the anticipated overall attrition rate $\xi$ can be expressed under the uniform and linear distribution of attrition times respectively as follows:

$$\varphi^{(u)} = \Phi \left\{ \Delta \sqrt{\frac{N_3 N_2^{(u)} E^{(u)}(N_1) Var_p^{(u)}(T)}{2\left\{(1-\rho_1)+r_\tau E^{(u)}(N_1) Var_p^{(u)}(T)\right\}}} - \Phi^{-1}(1-\alpha/2) \right\} \quad (8)$$

and

$$\varphi^{(l)} = \Phi \left\{ \Delta \sqrt{\frac{N_3 N_2^{(l)} E^{(l)}(N_1) Var_p^{(l)}(T)}{2\left\{(1-\rho_1)+r_\tau E^{(l)}(N_1) Var_p^{(l)}(T)\right\}}} - \Phi^{-1}(1-\alpha/2) \right\}. \quad (9)$$

It follows that the approximate sample size determinations with the anticipated overall attrition rate $\xi$ are:

$$N_2^{(u)} = \frac{2\left\{(1-\rho_1)+r_\tau E^{(u)}(N_1) Var_p^{(u)}(T)\right\}\left\{\Phi^{-1}(1-\alpha/2)+\Phi^{-1}(\varphi^{(u)})\right\}^2}{N_3 E^{(u)}(N_1) Var_p^{(u)}(T)\Delta^2} \quad (10)$$

and

$$N_2^{(l)} = \frac{2\left\{(1-\rho_1) + r_\tau E^{(l)}(N_1)\,Var_p^{(l)}(T)\right\}\left\{\Phi^{-1}(1-\alpha/2) + \Phi^{-1}(\varphi^{(l)})\right\}^2}{N_3 E^{(l)}(N_1)\,Var_p^{(l)}(T)\Delta^2}. \quad (11)$$

Here again, both $N_2^{(u)}$ and $N_2^{(l)}$ are the smallest integers greater than their corresponding right hand sides of equations (10) and (11).

Let us denote the ratio, or the multiplication factor, in the number of subjects due to the anticipated attritions by

$$R^{(u)} = N_2^{(u)}/N_2 \quad (12)$$

and

$$R^{(l)} = N_2^{(l)}/N_2 \quad (13)$$

which we expect would be less than $1 + \xi/(1 - \xi)$.

## 5. Attrition Mechanisms

To explore the validity of the approximate sample size formulae $N_2^{(u)}$ (10) and $N_2^{(l)}$ (11) with anticipated attrition rates we consider the following three conventional attrition mechanisms (Little and Rubin, 2002): 1) Attrition completely at random (ACAR), that is, occurrence of a subject's attrition does not depend on any observed or unobserved outcomes; 2) Attrition at random (AAR), that is, occurrence of a subject's attrition depends on observed outcomes; and 3) Attrition not at random (ANAR), that is, occurrence of a subject's attrition depends on unobserved outcomes. To generate missing data based on each mechanism, we first group the subjects who are retained at each time point $t$ into quartiles: $Q_1(t)$, $Q_2(t)$, $Q_3(t)$, and $Q_4(t)$. This quartile grouping is based on values of outcome variable $Y$ among the retainers discounting the dropouts.

Based on the quartile grouping, the conditional distributions of attritions can be formulated in the following way. First, the conditional probability of attrition at time $t$ among retainers up to time $t$ can be written as: with the superscripts indicating "uniform" or "linear" suppressed,

$$z_t \equiv P(A = t \mid t \le A \le T_{end}) = w_t \Big/ \left(1 - \sum_{l=0}^{t} w_l\right).$$

When the quartiles are based on $Y_{t-1}$, the conditional distributions of attrition times for the g-th quartile can be written as

$$\omega_g(t) \equiv P\left(A{=}t|Y_{t-1} \in Q_g(t{-}1), t \le A \le T_{end}\right),$$

which yields $\sum_g \omega_g(t) P\left(Y_{t-1} \in Q_g(t{-}1)\right) = z_t$ and thus $\sum_g \omega_g(t){=}4z_t$. Similarly, when the quartiles are based on $Y_t$, the conditional distributions of attrition times for the g-th quartile can be written as

$$\varpi_g(t) \equiv P\left(A{=}t|Y_t \in Q_g(t), t \le A \le T_{end}\right),$$

which yields $\sum_g \varpi_g(t) P\left(Y_t \in Q_g(t)\right) = z_t$ and thus $\sum_g \varpi_g(t){=}4z_t$.

For the ACAR mechanism, we consider $\omega_g(t) = z_t$ for all $g$ and $t > 0$, i.e., subject attrition at time $t$ does not depend on any previous observations. For the AAR mechanism, we consider that the percentages of dropouts at time $t$ whose observed outcomes at time $t - 1$ belong to $Q_1(t{-}1)$, $Q_2(t{-}1)$, $Q_3(t{-}1)$, and $Q_4(t{-}1)$ are 10%, 20%, 30% and 40%, respectively. Specifically, $\omega_1(t) = .4z_t$, $\omega_2(t) = .8z_t$, $\omega_3(t) = 1.2z_t$ and $\omega_4(t) = 1.6z_t$. Under the AAR mechanism, the subject attrition depends on the observed outcomes at time $t - 1$, that is, the time immediately prior to attrition. For the ANAR mechanism, we similarly consider the percentages of dropouts at time $t$ whose observed outcomes at time $t$ belong to $Q_1(t)$, $Q_2(t)$, $Q_3(t)$, and $Q_4(t)$ are 10%, 20%, 30% and 40%, respectively. Specifically, $\varpi_1(t) = .4z_t$, $\varpi_2(t) = .8z_t$, $\varpi_3(t) = 1.2z_t$ and $\varpi_4(t) = 1.6z_t$. Under the ANAR mechanism, the subject attrition depends on the unobserved outcomes at time $t$, that is, the time of attrition.

## 6. Simulation study

We conduct simulation studies to examine the performance of the sample size $N_2^{(u)}$ and $N_2^{(l)}$ a two-sided significance level $\alpha = 0.05$ and a desired power $\varphi = 0.8$ under the following combinations: $T_{end} = (N_1 - 1) = 0.4, 0.5$; $N_3 = 10, 20$; $N_1 = 5, 9$; $\rho_1 = 0.4, 0.6$; $r_\tau = 0.0$, 0.1, 0.2, while without loss of generality, $\sigma = 1$, $\rho_2 = 0.1$, $\beta_0 = \zeta = 0$, and $\tau = -1$ in model (1) remained fixed. Of note, when $r_\tau = 0.1$ or 0.2 under the random slope models with missing data, $N_1 = 9$ was excluded due to enormous computing times for simulations (Ahn et al., 2000; Overall et al., 1999). Values for $\sigma_3^2$ and $\sigma_2^2$ were determined through $\rho_2$ (3) and $\rho_1$ (4). The effect size of the interaction is specified as a standardized between-group mean difference $T_{end} = (N_1 - 1)$ at the end of trial under a fixed slope model. Effect sizes in the range of 0.4–0.6 have generally been referred to as medium (Cohen, 1988). We further considered two attrition rates, $\xi = 20\%$ and 30%, and two types of distributions, uniform and linear, of the attrition time points as detailed above.

For each combination, we first computed: $E^{(u)}(N_1)$, $Var_\rho^{(u)}(T)$, $E^{(l)}(N_1)$, and $Var_\rho^{(l)}(T)$, and subsequently $N_2^{(u)}$ and $N_2^{(l)}$. We then generate 1000 simulated data sets for each combination with each estimated $N_2^{(u)}$ or $N_2^{(l)}$ in accordance to model (1): $Y_{ijk} = (\beta_0 + u_i + u_{j(i)}) + \zeta X_{ijk} +$

$(\tau + \nu_{j(i)})T_{ijk} + \delta X_{ijk} T_{ijk} + e_{ijk}$. Finally, according to the three ACAR, AAR and ANAR mechanisms specified in the section above, we delete outcomes from each "complete" data set generated with $N_1$ observations per subject, resulting in three data sets with deleted observations.

Although the sample size determinations were derived based on OLS estimates with known variance components, in order to reflect real data analysis with unknown variance components, we fit each deleted data set using SAS PROC MIXED with the maximum likelihood estimation option and retained the resulting $p$-values for testing the null hypothesis (2). We denoted the $p$-value by $p_s(\delta)$ for the $s$-th simulated data set ($s = 1, 2,..,$ 1000) and computed the empirical power $\tilde{\varphi}$, $\tilde{\varphi}^{(u)}$ or $\tilde{\varphi}^{(l)}$, as follows:

$$\tilde{\varphi} = \sum_{s=1}^{1000} 1\left\{p_s(\delta) < \alpha\right\} / 1000. \quad (14)$$

This empirical power is compared with the approximate power $\varphi$ on which the sample sizes $N_2^{(u)}$ and $N_2^{(l)}$ are based. We note that $\varphi$ is never less than the pre-specified power of 0.8 since both $N_2^{(u)}$ and $N_2^{(l)}$ are the smallest integer greater than the right hand side of equation (10) and (11), respectively.

## 7. Simulation study results

### Attrition rates

Over all combinations of the simulation specifications, the empirical attrition rates based on simulated data are virtually identical to the pre-specified attrition rates $\xi = 20\%$ and 30% regardless of the three different attrition mechanisms and distributions of attrition time points, uniform and linear.

### Under fixed slope model

Table 1 summarizes numerical and simulation results when the slopes are considered fixed, i.e., when $r_\tau = 0.0$. The average ratios $R^{(u)}$ (12) and $R^{(l)}$ (13) are both much less than $1 + \xi/(1 - \xi) = 1.25$ and 1.43 for $\xi = 20\%$ and 30%, respectively. Nevertheless, we observe from the evaluation of $\max|\varphi - \tilde{\varphi}|$ that the empirical power estimates are very close to the approximate power regardless of attrition mechanisms and across all simulation parameter combinations. As is foreseen from equation (7), $R^{(u)}$ is no greater than $R^{(l)}$ in every simulation combination. Furthermore, $R^{(l)}$ is 1.0 in many cases for both $\xi = 20\%$ and 30% without loss of statistical power. In general, both $R^{(u)}$ and $R^{(l)}$ are smaller for greater .

### Under random slope model

Table 2 summarizes numerical and simulation results when the slopes are considered random, i.e., when $r_\tau = 0.1$ or 0.2. Again, the average ratios $R^{(u)}$ and $R^{(l)}$ are both much far less than $1 + \xi/(1 - \xi) = 1.25$ and 1.43 for $\xi = 20\%$ and 30%, respectively. In fact, these are even smaller than those under fixed slope model above partly because $N_2$ are much greater due to additional random variations in slopes. However, the simulation-based empirical power estimates underestimate somewhat severely the approximate statistical power.

Although the absolute difference $|\varphi - \tilde{\varphi}|$ is not necessarily associated with attrition mechanism or with distribution of attrition times, it ranges from 0.001 to 0.64 for $\xi = 20\%$; and from 0.019 to 0.078 for $\xi = 30\%$. The underestimation is more severe for the greater attrition rate $\xi = 30\%$. In this case, despite that the biases are about 5%, all of the absolute differences $|\varphi - \tilde{\varphi}|$ are beyond the 95% confidence limit, $\pm 1.96 \sqrt{0.8 \times 0.2/1000} = \pm 0.025$ except only one case. Again, $R^{(u)}$ is no greater than $R^{(l)}$ in every simulation combination and $R^{(l)}$ is 1.0 in many cases for both $\xi = 20\%$ and 30% but seemingly at the cost of statistical power. Again, in general, both $R^{(u)}$ and $R^{(l)}$ are smaller for greater .

## 8. Discussion

The proposed replacement strategies reflected on $N_2^{(u)}$ (10) and $N_2^{(l)}$ (11) are shown very effective resulting nearly unbiased statistical power when the subject-specific slopes are assumed to be fixed despite the fact that the empirical power estimates were obtained based on simulations with maximum likelihood estimates with unknown variances. Furthermore under this fixed slope model, the finding that $R^{(l)}$ (13) in many cases and $R^{(u)}$ (12) in some cases are 1.0 for both $\xi = 20\%$ and 30% implies that no additional recruitment of study subjects may be necessary in those cases. It is because in those cases the statistical power $\varphi$ (5) under assumption of no attrition might be substantially greater than 0.8 with "integer" values of $N_2$ (6).

On the other hand, when the subject-specific slopes are assumed to be random, the replacement strategy yields underestimated statistical power, that is, the magnitudes of empirical power with sample sizes $N_2^{(u)}$ (10) and $N_2^{(l)}$ (11) are smaller than those of the approximate power $\varphi^{(u)}$ (8) and $\varphi^{(l)}$ (9), respectively. The underestimation was more severe for the greater attrition rate. Both $R^{(u)}$ and $R^{(l)}$ are close to 1.0 less than 1.1 in almost all cases regardless of the attrition rates because $N_2$ (6) under no attrition is large (compared to that under fixed model) due to the additional variance of the subject-specific slopes. If the empirical statistical power were more close to unknown "true" statistical power under attrition than the approximate power (8) and (9), then these approximations might overestimate statistical power under random slope models. Although the empirical statistical power is more likely close to the unknown true statistical power, potential sources of the discrepancy between the empirical and the approximate power are unknown. For example, it could be due to inaccuracy of the approximate power, or due to potential loss of power stemming from the unknown variance assumptions in empirical power estimates, or due to both. Regardless, however, some additional adjustments to $N_2^{(u)}$ and $N_2^{(l)}$ deem necessary to yield unbiased statistical power under the random slope model. The adjustment does not appear to be substantial given that underestimation is about 5% point on average. For example, the sample sizes $N_2^{(u)}$ and $N_2^{(l)}$ can be used as a lower reference bound for conducting empirical simulations iteratively by slightly increasing the number of subjects per cluster until empirical power (14) reaches desired statistical power.

In either fixed or random slope model, nevertheless, multiplication of $N_2$(6) by $1 + \xi/(1 - \xi)$ would result in an over-powered study design. However, it is surprising that the performance of the replacement strategy did not depend on attrition mechanisms in view of the empirical

power. In particular, the empirical power under ANAR was virtually identical to those under ACAR and AAR regardless of models, attrition rates, and distribution of attrition times. Although we assume monotone attrition patterns and considered scenarios of the three attrition mechanisms are somewhat limited, this finding implies that the replacement strategy is robust against unknown attrition mechanisms.

Several limitations should be considered when the derived approximate sample size determinations are applied. First, model (1) requires a minimum number of parameters in a class of models for longitudinal cluster trail data and thus may not necessarily reflect a real situation. For instance, when the subject-level random intercepts $u_{j(i)}$ and random slopes $v_{j(i)}$ are correlated, the approximate sample size determination might be even more biased. Therefore, if pilot data were available, testing significance of the correlation in addition to testing significance of variance of random slopes would be important to determine whether to apply the derived sample size formulas. Second, the attrition process may not necessarily be monotone and the distribution of attrition rates may not be uniform or linear. Third, real attrition mechanisms are usually unknown in real practice and may not necessarily be a function of the quintile grouping of the outcome variable. Theoretical derivations, if possible, of sample size determinations addressing all of these concerns that could result in minimized bias should deserve future studies.

In conclusion, even though their application might be limited in real practice, the closed form approximate sample size formulas $N_2^{(u)}$ and $N_2^{(l)}$ should be useful for designing a cluster randomized trial where testing slope differences is a primary goal. If the subject-specific slopes are homogeneous, the approximate determinations should be accurate and unbiased. Otherwise, some adjustments, if not substantial, are needed to secure adequate statistical power.

## Acknowledgments

## References

Ahn C, Tonidandel S, Overall JE. Issues in use of SAS PROC.MIXED to test the significance of treatment effects in controlled clinical trials. Journal of Biopharmaceutical Statistics. 2000; 10:265–286. [PubMed: 10803729]

Alexopoulos GS, Katz IR, Bruce ML, et al. Remission in depressed geriatric primary care patients: A report from the PROSPECT study. American Journal of Psychiatry. 2005; 162:718–724. [PubMed: 15800144]

Cohen, J. Statistical Power Analysis for the Behavioral Science. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

Dietrich AJ, Oxman TE, Williams JW, et al. Re-engineering systems for the treatment of depression in primary care: cluster randomised controlled trial. British Medical Journal. 2004; 329:602–605. [PubMed: 15345600]

Hedeker, D.; Gibbons, RD. Longitudinal Data Analysis. Hoboken, NJ: Wiley; 2006.

Heo M, Leon AC. Sample size requirements to detect an intervention by time interaction in longitudinal cluster randomized clinical trials. Statistics in Medicine. 2009; 28:1017–1027. [PubMed: 19153969]

Heo M, Papademetriou E, Meyers BS. Design characteristics that influence attrition in geriatric antidepressant trials: meta-analysis. International Journal of Geriatric Psychiatry. 2009; 24:990–1001. [PubMed: 19263464]

Heo M, Xue XN, Kim MY. Sample size requirements to detect an intervention by time interaction in longitudinal cluster randomized clinical trials with random slopes. Computational Statistics and Data Analysis. (in press).

Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics. 1982; 38:963–974. [PubMed: 7168798]

Little, RJA.; Rubin, DB. Statistical Analysis with Missing Data. 2. Hoboken, NJ: Wiley; 2002.

Longford, NT. Random Coefficient Models. New York: Oxford University Press; 1993.

Murray DM, Blitstein JL, Hannan PJ, Baker WL, Lytle LA. Sizing a trial to alter the trajectory of health behaviours: Methods, parameter estimates, and their application. Statistics in Medicine. 2007; 26:2297–2316. [PubMed: 17044139]

Overall JE, Ahn C, Shivakumar C, Kalburgi Y. Problematic formulations of SAS PROC.MIXED models for repeated measurements. Journal of Biopharmaceutical Statistics. 1999; 9:189–216. [PubMed: 10091918]

Roy A, Bhaumik DK, Aryal S, Gibbons RD. Sample size determination for hierarchical longitudinal designs with differential attrition rates. Biometrics. 2007; 63:699–707. [PubMed: 17825003]

**Table 1**

Sample sizes $N_2^{(u)}$ and $N_2^{(l)}$, corresponding theoretical power $\varphi$ and empirical power $\tilde{\varphi}$ for testing intervention group by time interaction effect in a three level mixed-effect *fixed slope coefficient* linear regression analysis, i.e., when $r_\tau = 0$, based on 1000 simulations for each combination.

| $\xi$ | $T_{end}$ | $N_3$ | $N_1$ | $\rho_1$ | $N_2$ | $N_2^{(u)}$ | $R^{(u)}$ | $\varphi^{(u)}$ | $\tilde{\varphi}^{(u)}$ ACAR | AAR | ANAR | $N_2^{(l)}$ | $R^{(l)}$ | $\varphi^{(l)}$ | $\tilde{\varphi}^{(l)}$ ACAR | AAR | ANAR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20% | 0.4 | 10 | 5 | 0.4 | 10 | 12 | 1.20 | 0.822 | 0.832 | 0.844 | 0.832 | 11 | 1.10 | 0.819 | 0.795 | 0.795 | 0.798 |
| | | | | 0.6 | 7 | 8 | 1.14 | 0.822 | 0.804 | 0.805 | 0.798 | 7 | 1.00 | 0.802 | 0.788 | 0.785 | 0.790 |
| | | | 9 | 0.4 | 7 | 8 | 1.14 | 0.833 | 0.820 | 0.817 | 0.817 | 7 | 1.00 | 0.804 | 0.767 | 0.777 | 0.769 |
| | | | | 0.6 | 5 | 5 | 1.00 | 0.809 | 0.822 | 0.816 | 0.824 | 5 | 1.00 | 0.831 | 0.823 | 0.815 | 0.818 |
| | | 20 | 5 | 0.4 | 5 | 6 | 1.20 | 0.822 | 0.835 | 0.827 | 0.830 | 6 | 1.20 | 0.851 | 0.836 | 0.835 | 0.842 |
| | | | | 0.6 | 4 | 4 | 1.00 | 0.822 | 0.811 | 0.816 | 0.810 | 4 | 1.00 | 0.851 | 0.821 | 0.823 | 0.818 |
| | | | 9 | 0.4 | 4 | 4 | 1.00 | 0.833 | 0.840 | 0.829 | 0.839 | 4 | 1.00 | 0.854 | 0.830 | 0.824 | 0.827 |
| | | | | 0.6 | 3 | 3 | 1.00 | 0.874 | 0.873 | 0.879 | 0.864 | 3 | 1.00 | 0.892 | 0.890 | 0.889 | 0.876 |
| | 0.5 | 10 | 5 | 0.4 | 7 | 8 | 1.14 | 0.837 | 0.828 | 0.838 | 0.837 | 7 | 1.00 | 0.817 | 0.784 | 0.781 | 0.774 |
| | | | | 0.6 | 5 | 5 | 1.00 | 0.813 | 0.783 | 0.792 | 0.794 | 5 | 1.00 | 0.843 | 0.841 | 0.847 | 0.839 |
| | | | 9 | 0.4 | 5 | 5 | 1.00 | 0.825 | 0.832 | 0.833 | 0.839 | 5 | 1.00 | 0.845 | 0.829 | 0.834 | 0.829 |
| | | | | 0.6 | 3 | 4 | 1.33 | 0.887 | 0.886 | 0.902 | 0.893 | 3 | 1.00 | 0.806 | 0.805 | 0.814 | 0.808 |
| | | 20 | 5 | 0.4 | 4 | 4 | 1.00 | 0.837 | 0.866 | 0.847 | 0.851 | 4 | 1.00 | 0.865 | 0.857 | 0.860 | 0.866 |
| | | | | 0.6 | 3 | 3 | 1.00 | 0.877 | 0.893 | 0.878 | 0.888 | 3 | 1.00 | 0.901 | 0.904 | 0.889 | 0.905 |
| | | | 9 | 0.4 | 3 | 3 | 1.00 | 0.887 | 0.886 | 0.889 | 0.887 | 3 | 1.00 | 0.903 | 0.884 | 0.896 | 0.892 |
| | | | | 0.6 | 2 | 2 | 1.00 | 0.887 | 0.876 | 0.874 | 0.872 | 2 | 1.00 | 0.903 | 0.882 | 0.890 | 0.897 |
| Mean | | | | | | | 1.07 | 0.843 | 0.843 | 0.843 | 0.842 | | 1.02 | 0.849 | 0.834 | 0.835 | 0.834 |
| Min $\lvert \varphi - \tilde{\varphi} \rvert$ | | | | | | | | | 0.001 | 0.001 | 0.000 | | | | 0.001 | 0.003 | 0.001 |
| Max $\lvert \varphi - \tilde{\varphi} \rvert$ | | | | | | | | | 0.030 | 0.022 | 0.024 | | | | 0.037 | 0.036 | 0.043 |
| 30% | 0.4 | 10 | 5 | 0.4 | 10 | 13 | 1.30 | 0.810 | 0.833 | 0.837 | 0.825 | 12 | 1.20 | 0.829 | 0.812 | 0.803 | 0.821 |
| | | | | 0.6 | 7 | 9 | 1.29 | 0.825 | 0.838 | 0.849 | 0.845 | 8 | 1.14 | 0.829 | 0.792 | 0.795 | 0.795 |
| | | | 9 | 0.4 | 7 | 9 | 1.29 | 0.843 | 0.834 | 0.830 | 0.836 | 8 | 1.14 | 0.833 | 0.823 | 0.815 | 0.820 |
| | | | | 0.6 | 5 | 6 | 1.20 | 0.843 | 0.858 | 0.847 | 0.848 | 5 | 1.00 | 0.809 | 0.799 | 0.800 | 0.798 |
| | | 20 | 5 | 0.4 | 5 | 7 | 1.40 | 0.838 | 0.852 | 0.858 | 0.851 | 6 | 1.20 | 0.829 | 0.808 | 0.816 | 0.811 |

| $\xi$ | $T_{end}$ | $N_3$ | $N_1$ | $\rho_1$ | $N_2$ | $N_2^{(u)}$ | $R^{(u)}$ | $\varphi^{(u)}$ | $\tilde{\varphi}^{(u)}$ | | | $N_2^{(l)}$ | $R^{(l)}$ | $\varphi^{(l)}$ | $\tilde{\varphi}^{(l)}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | ACAR | AAR | ANAR | | | | ACAR | AAR | ANAR |
| 0.5 | 10 | 5 | | 0.6 | 4 | 5 | 1.25 | 0.862 | 0.888 | 0.881 | 0.880 | 4 | 1.00 | 0.829 | 0.801 | 0.801 | 0.796 |
| | | | 9 | 0.4 | 4 | 5 | 1.25 | 0.879 | 0.869 | 0.876 | 0.880 | 4 | 1.00 | 0.833 | 0.797 | 0.793 | 0.801 |
| | | | | 0.6 | 3 | 3 | 1.00 | 0.843 | 0.837 | 0.832 | 0.832 | 3 | 1.00 | 0.874 | 0.860 | 0.848 | 0.856 |
| | | 5 | | 0.4 | 7 | 9 | 1.29 | 0.840 | 0.865 | 0.860 | 0.862 | 8 | 1.14 | 0.844 | 0.821 | 0.822 | 0.827 |
| | | | 9 | 0.6 | 5 | 6 | 1.20 | 0.840 | 0.848 | 0.846 | 0.843 | 5 | 1.00 | 0.820 | 0.812 | 0.803 | 0.805 |
| | | | | 0.4 | 5 | 6 | 1.20 | 0.857 | 0.848 | 0.853 | 0.856 | 5 | 1.00 | 0.825 | 0.816 | 0.809 | 0.814 |
| | 20 | 5 | | 0.6 | 3 | 4 | 1.33 | 0.857 | 0.860 | 0.857 | 0.849 | 4 | 1.33 | 0.887 | 0.859 | 0.861 | 0.858 |
| | | | 9 | 0.4 | 4 | 5 | 1.25 | 0.875 | 0.906 | 0.889 | 0.896 | 4 | 1.00 | 0.844 | 0.844 | 0.830 | 0.839 |
| | | | | 0.6 | 3 | 3 | 1.00 | 0.840 | 0.866 | 0.846 | 0.849 | 3 | 1.00 | 0.883 | 0.867 | 0.869 | 0.854 |
| | | | 9 | 0.4 | 3 | 3 | 1.00 | 0.857 | 0.864 | 0.854 | 0.863 | 3 | 1.00 | 0.887 | 0.862 | 0.867 | 0.860 |
| | | | | 0.6 | 2 | 2 | 1.00 | 0.857 | 0.849 | 0.841 | 0.840 | 2 | 1.00 | 0.887 | 0.866 | 0.864 | 0.861 |
| Mean | | | | | | | 1.20 | 0.848 | 0.857 | 0.854 | 0.853 | | 1.07 | 0.846 | 0.827 | 0.825 | 0.826 |
| Min $|\varphi - \tilde{\varphi}|$ | | | | | | | | | 0.003 | 0.000 | 0.001 | | | | 0.000 | 0.009 | 0.005 |
| Max $|\varphi - \tilde{\varphi}|$ | | | | | | | | | 0.031 | 0.027 | 0.022 | | | | 0.037 | 0.040 | 0.034 |

Note: ACAR = Attrition Completely At Random; AAR = Attrition At Random; ANAR = Attrition Not At Random

**Table 2**

Sample sizes $N_2^{(u)}$ and $N_2^{(l)}$, corresponding theoretical power $\varphi$ and empirical power $\tilde{\varphi}$ for testing intervention group by time interaction effect in a three level mixed-effect *random slope coefficient* linear regression analysis, i.e., when $r_\tau > 0$, based on 1000 simulations for each combination.

| $\xi$ | $r_\tau$ | $T_{end}$ | $N_3$ | $\rho_1$ | $N_2$ | $N_2^{(u)}$ | $R^{(u)}$ | $\varphi^{(u)}$ | $\tilde{\varphi}^{(u)}$ ACAR | AAR | ANAR | $N_2^{(l)}$ | $R^{(l)}$ | $\varphi^{(l)}$ | $\tilde{\varphi}^{(l)}$ ACAR | AAR | ANAR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 20% | 0.1 | 0.4 | 10 | 0.4 | 26 | 28 | 1.08 | 0.814 | 0.792 | 0.796 | 0.799 | 27 | 1.04 | 0.812 | 0.780 | 0.764 | 0.774 |
| | | | | 0.6 | 22 | 24 | 1.09 | 0.812 | 0.789 | 0.785 | 0.782 | 23 | 1.05 | 0.806 | 0.785 | 0.779 | 0.784 |
| | | | 20 | 0.4 | 13 | 14 | 1.08 | 0.814 | 0.802 | 0.784 | 0.799 | 14 | 1.08 | 0.826 | 0.800 | 0.808 | 0.803 |
| | | | | 0.6 | 11 | 12 | 1.09 | 0.812 | 0.794 | 0.788 | 0.790 | 12 | 1.09 | 0.822 | 0.786 | 0.787 | 0.783 |
| | | 0.5 | 10 | 0.4 | 17 | 18 | 1.06 | 0.815 | 0.798 | 0.799 | 0.802 | 17 | 1.00 | 0.806 | 0.780 | 0.780 | 0.785 |
| | | | | 0.6 | 15 | 15 | 1.00 | 0.803 | 0.787 | 0.778 | 0.790 | 15 | 1.00 | 0.813 | 0.785 | 0.785 | 0.786 |
| | | | 20 | 0.4 | 9 | 9 | 1.00 | 0.815 | 0.776 | 0.785 | 0.777 | 9 | 1.00 | 0.828 | 0.764 | 0.766 | 0.772 |
| | | | | 0.6 | 8 | 8 | 1.00 | 0.828 | 0.799 | 0.815 | 0.806 | 8 | 1.00 | 0.837 | 0.836 | 0.824 | 0.813 |
| | 0.2 | 0.4 | 10 | 0.4 | 41 | 43 | 1.05 | 0.802 | 0.779 | 0.777 | 0.764 | 42 | 1.02 | 0.801 | 0.781 | 0.759 | 0.773 |
| | | | | 0.6 | 38 | 39 | 1.03 | 0.800 | 0.771 | 0.771 | 0.763 | 39 | 1.03 | 0.806 | 0.789 | 0.797 | 0.776 |
| | | | 20 | 0.4 | 21 | 22 | 1.05 | 0.811 | 0.790 | 0.785 | 0.787 | 21 | 1.00 | 0.801 | 0.772 | 0.775 | 0.765 |
| | | | | 0.6 | 19 | 20 | 1.05 | 0.810 | 0.766 | 0.769 | 0.774 | 20 | 1.05 | 0.816 | 0.788 | 0.794 | 0.793 |
| | | 0.5 | 10 | 0.4 | 27 | 28 | 1.04 | 0.809 | 0.785 | 0.781 | 0.785 | 27 | 1.00 | 0.803 | 0.757 | 0.764 | 0.767 |
| | | | | 0.6 | 25 | 25 | 1.00 | 0.801 | 0.754 | 0.751 | 0.752 | 25 | 1.00 | 0.807 | 0.809 | 0.815 | 0.817 |
| | | | 20 | 0.4 | 14 | 14 | 1.00 | 0.809 | 0.759 | 0.757 | 0.762 | 14 | 1.00 | 0.817 | 0.783 | 0.788 | 0.776 |
| | | | | 0.6 | 13 | 13 | 1.00 | 0.816 | 0.786 | 0.789 | 0.787 | 13 | 1.00 | 0.822 | 0.789 | 0.798 | 0.794 |
| Mean | | | | | | | 1.04 | 0.811 | 0.783 | 0.782 | 0.782 | | 1.02 | 0.814 | 0.787 | 0.786 | 0.785 |
| Min $|\varphi - \tilde{\varphi}|$ | | | | | | | | | 0.012 | 0.013 | 0.013 | | | | 0.001 | 0.008 | 0.010 |
| Max $|\varphi - \tilde{\varphi}|$ | | | | | | | | | 0.050 | 0.052 | 0.049 | | | | 0.064 | 0.062 | 0.056 |
| 30% | 0.1 | 0.4 | 10 | 0.4 | 26 | 29 | 1.12 | 0.809 | 0.777 | 0.775 | 0.778 | 27 | 1.04 | 0.802 | 0.752 | 0.738 | 0.739 |
| | | | | 0.6 | 22 | 25 | 1.14 | 0.814 | 0.745 | 0.745 | 0.750 | 24 | 1.09 | 0.814 | 0.772 | 0.777 | 0.771 |
| | | | 20 | 0.4 | 13 | 15 | 1.15 | 0.822 | 0.783 | 0.786 | 0.788 | 14 | 1.08 | 0.817 | 0.766 | 0.768 | 0.754 |
| | | | | 0.6 | 11 | 13 | 1.18 | 0.828 | 0.771 | 0.788 | 0.769 | 12 | 1.09 | 0.814 | 0.738 | 0.756 | 0.742 |
| | | 0.5 | 10 | 0.4 | 17 | 19 | 1.12 | 0.818 | 0.778 | 0.781 | 0.799 | 18 | 1.06 | 0.818 | 0.742 | 0.756 | 0.774 |

| ξ | $r_\tau$ | $T_{end}$ | $N_3$ | $\rho_1$ | $N_2$ | $N_2^{(u)}$ | $R^{(u)}$ | $\varphi^{(u)}$ | $\tilde{\varphi}^{(u)}$ ACAR | AAR | ANAR | $N_2^{(l)}$ | $R^{(l)}$ | $\varphi^{(l)}$ | $\tilde{\varphi}^{(l)}$ ACAR | AAR | ANAR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 0.6 | 15 | 16 | 1.07 | 0.814 | 0.769 | 0.763 | 0.756 | 15 | 1.00 | 0.805 | 0.741 | 0.744 | 0.730 |
| | | | 20 | 0.4 | 9 | 10 | 1.11 | 0.837 | 0.796 | 0.806 | 0.794 | 9 | 1.00 | 0.818 | 0.773 | 0.765 | 0.763 |
| | | | | 0.6 | 8 | 8 | 1.00 | 0.814 | 0.782 | 0.776 | 0.770 | 8 | 1.00 | 0.830 | 0.770 | 0.770 | 0.776 |
| 0.2 | 0.4 | 10 | | 0.4 | 41 | 45 | 1.10 | 0.808 | 0.763 | 0.746 | 0.766 | 43 | 1.05 | 0.804 | 0.759 | 0.762 | 0.764 |
| | | | | 0.6 | 38 | 40 | 1.05 | 0.802 | 0.740 | 0.742 | 0.739 | 39 | 1.03 | 0.802 | 0.751 | 0.757 | 0.756 |
| | | | 20 | 0.4 | 21 | 23 | 1.10 | 0.817 | 0.774 | 0.766 | 0.759 | 22 | 1.05 | 0.813 | 0.779 | 0.775 | 0.779 |
| | | | | 0.6 | 19 | 20 | 1.05 | 0.802 | 0.749 | 0.769 | 0.748 | 20 | 1.05 | 0.812 | 0.760 | 0.760 | 0.765 |
| | 0.5 | 10 | | 0.4 | 27 | 29 | 1.07 | 0.811 | 0.745 | 0.749 | 0.751 | 28 | 1.04 | 0.811 | 0.757 | 0.757 | 0.742 |
| | | | | 0.6 | 25 | 26 | 1.04 | 0.808 | 0.746 | 0.730 | 0.741 | 25 | 1.00 | 0.802 | 0.744 | 0.742 | 0.736 |
| | | | 20 | 0.4 | 14 | 15 | 1.07 | 0.824 | 0.779 | 0.771 | 0.772 | 14 | 1.00 | 0.811 | 0.762 | 0.767 | 0.759 |
| | | | | 0.6 | 13 | 13 | 1.00 | 0.808 | 0.758 | 0.766 | 0.753 | 13 | 1.00 | 0.818 | 0.763 | 0.765 | 0.774 |
| Mean | | | | | | | 1.09 | 0.815 | 0.766 | 0.766 | 0.765 | | 1.04 | 0.812 | 0.758 | 0.760 | 0.758 |
| Min $|\varphi - \tilde{\varphi}|$ | | | | | | | | | 0.032 | 0.031 | 0.019 | | | | 0.034 | 0.037 | 0.034 |
| Max $|\varphi - \tilde{\varphi}|$ | | | | | | | | | 0.069 | 0.078 | 0.067 | | | | 0.076 | 0.064 | 0.075 |

Note: ACAR = Attrition Completely At Random; AAR = Attrition At Random; ANAR = Attrition Not At Random